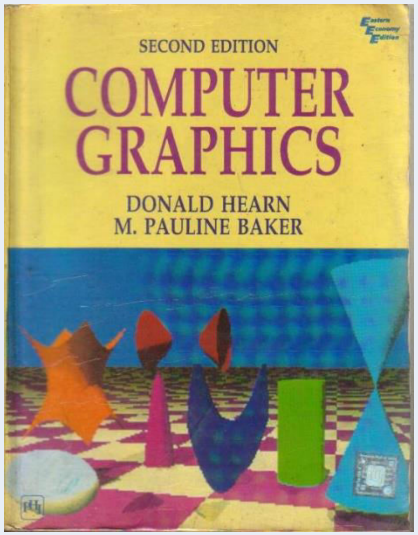
CLASS-1(1)

Simulation and Modeling

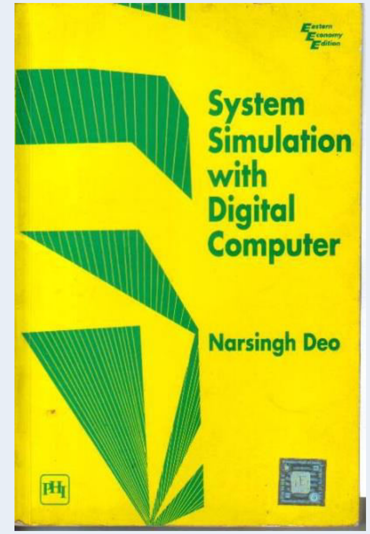
CSE4131_CSE4132

Recommended Books



- COMPUTER GRAPHICS

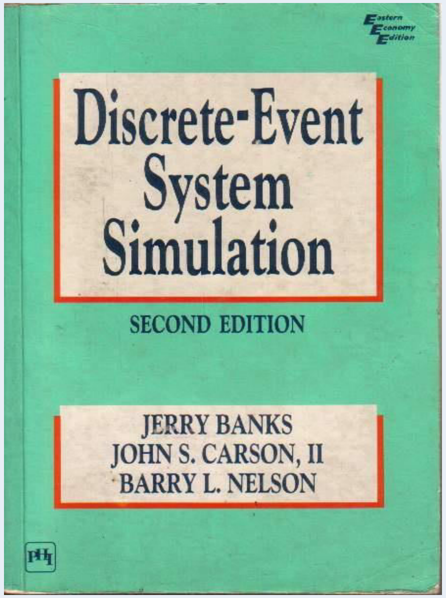
  - By Donald **Hearn &** M. Pauline **Baker**



- SYSTEM SIMULATION WITH DIGITAL COMPUTER

  – By **Narsingh Deo**

Recommended Books



- DISCRETE-EVENT SYSTEM SIMULATION
  - Jerry Banks, John S. Carson, Barry L. Nelson
  - Publisher: Prentice- Hall

Simulation

- What is simulation? (Formal Definition)
  - Simulation → অনুকরন করা
  - *A simulation is the imitation (নকল করা) of the operation of a **real world process or system over time**.*
  - Can be done:
    - by hand **or** on a computer

Definition of System

- *A system is defined as a group of objects that are joined together in some regular interaction or interdependence toward the accomplishment of some purpose.*

- A system is often affected by changes occurring outside the system.
  - Such changes are said to occur in the **system environment**
  - it is necessary to decide on the **boundary** between the system and its environment

Simulation

– Simulation involves

- **Generation** of an artificial history of a system **+**
  Observation of that artificial history

=➢Conclusion -> operating characteristics of the real system.

When Simulation Is the Appropriate Tool

- To study of **internal interaction** of a subsystem and **complex system**
- Changes can be Simulated
  - Informational (say false fire alarm)
  - Organizational (say removing supervisor)
  - Environmental changes (say introducing lunch)

    → **find their effects**
- To **gain knowledge** about **improvement** of system
- Finding important **input parameters** with changing simulation inputs

When Simulation Is the Appropriate Tool (cont'd)

– with **new design and policies** before implementation
– **different capabilities** for a machine can help determine the requirement.
– Simulation models designed for **training, allow learning** without the **cost** and disruption of on-the-job learning.
– A **plan can be visualized** with animated simulation
– The **modern system** (*factory, wafer fabrication plant, service organization*) is **too complex** that its **internal interaction** can be treated only by simulation

When Simulation Is **Not Appropriate**

- When the problem can be solved by **common sense**.
- When the problem can be solved **analytically (logically)**.
- If it is **easier** to perform **direct experiments**.
- If **cost** exceed savings.
- If **resource or time** are not available.
- If system behavior is too **complex**.
    - Like human behavior

Advantages

1. New policies, operating procedures, decision rules, information flows, organizational procedures, and so on **can be explored without disrupting** ongoing operations of the real system.
2. New hardware designs, physical layouts, transportation systems, and so on, **can be tested without committing** resources for their acquisition.
3. **Hypotheses** about how or why certain phenomena occur **can be tested for feasibility.**
4. **Time** can be **compressed or expanded** allowing for a speedup or slowdown of the phenomena under investigation.

Advantages (Cont'd)

5. **Insight** can be obtained about the **interaction of variables**.

6. **Insight** can be obtained about the **importance of variables** to the **performance of** the system.

7. **Bottleneck analysis (?)** can be performed indicating where work-in-process, information, materials, and so on are being excessively delayed.

8. A simulation study can help in **understanding** how the system operates rather than how **individuals think** the system operates.

9. **What-if** questions can be answered. This is particularly useful in the design of new systems.

Disadvantages

1. Model building requires **special training**.
2. Simulation results can be **difficult to interpret**.
3. Simulation modeling and analysis **can be time consuming** and **expensive**.
4. Simulation is used in sometime when an **analytical solution** is possible or even preferable.

Overcome the Limitations

1. Vendors of simulation software have been actively developing **packages** that contain models that need only **input data for their operation**.
2. Many simulation software vendors have **developed output analysis capabilities** within their packages for performing very thorough analysis.
3. Simulation can be performed **faster today than yesterday,** and even faster tomorrow.
4. Closed-form models are not able to analyze most of the complex systems that are encountered in practice.

Area of application

- **Manufacturing Systems**
  - Aircraft, Automobile
  - Semiconductor Manufacturing
- **Public systems**
  - Health care –prediction of pharmaceutical costs and outcome
  - Military – army officer professional development
  - Natural resource – solid waste management system
- **Transportation System**
  - i.e. cargo transport system

Area of application (Cont'd)

- **Construction systems**
  - i.e. cable-stayed bridge
- **Restaurant and entertainment system**
  - Quick service restaurant traffic analysis
- **Business process reengineering**
  - Business process modeling and analysis tools
- **Food processing**
  - Fishing boat operations in fish processing industries
- **Computer system performance**
  - Client-server system architecture

Definition of System

- *A system is defined as a group of objects that are joined together in some regular interaction or interdependence toward the accomplishment of some purpose.*

- A system is often affected by changes occurring outside the system.
  - Such changes are said to occur in the **system environment**
  - … it is necessary to decide on the **boundary** between the system and its environment

Components of a System

- **Entity-** …
  An object of interest in the system : *Machines in factory*
- **Attribute-** …
  The property of an entity : *speed, capacity*
- **Activity-** …
  A time period of specified length *:welding, stamping*
- **State-** …
  A collection of variables that describe the system in any time : *status of machine (busy, idle, down,…)*

Components of a System

- **Event-**…
  A instantaneous occurrence that might change the state of the system: *breakdown*
- **Endogenous-** …
  Activities and events occurring with the system.
- **Exogenous-**… Activities and events occurring with the environment.

Example of System and its components…

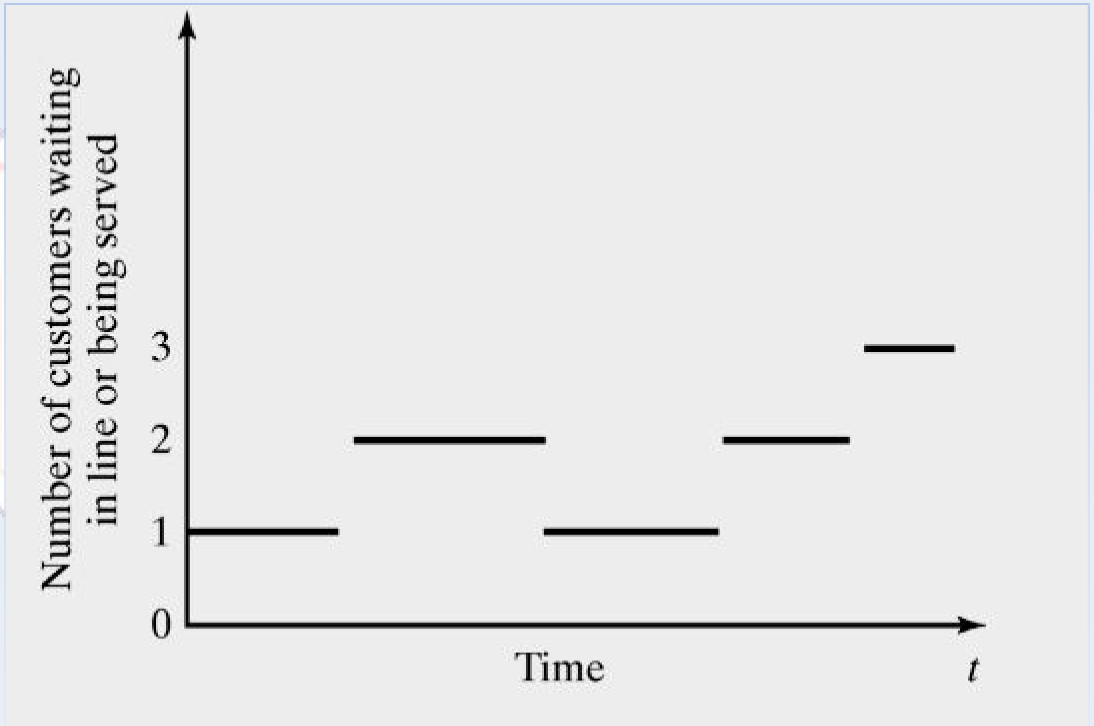
System	*Entities*	*Attributes*	*Activities*	*Events*	*State Variables*
Banking	Customers	Checking account balance	Making deposits	Arrival; departure	Number of busy tellers; number of customers waiting
Rapid rail	Riders	Origination; destination	Traveling	Arrival at station; arrival at destination	Number of riders waiting at each station; number of riders in transit
Production	Machines	Speed; capacity; breakdown rate	Welding; stamping	Breakdown	Status of machines (busy, idle, or down)
Communications	Messages	Length; destination	Transmitting	Arrival at destination	Number waiting to be transmitted
Inventory	Warehouse	Capacity	Withdrawing	Demand	Levels of inventory; backlogged demands

System

- System can be categorized into
  1. Discrete System
  2. Continuous System

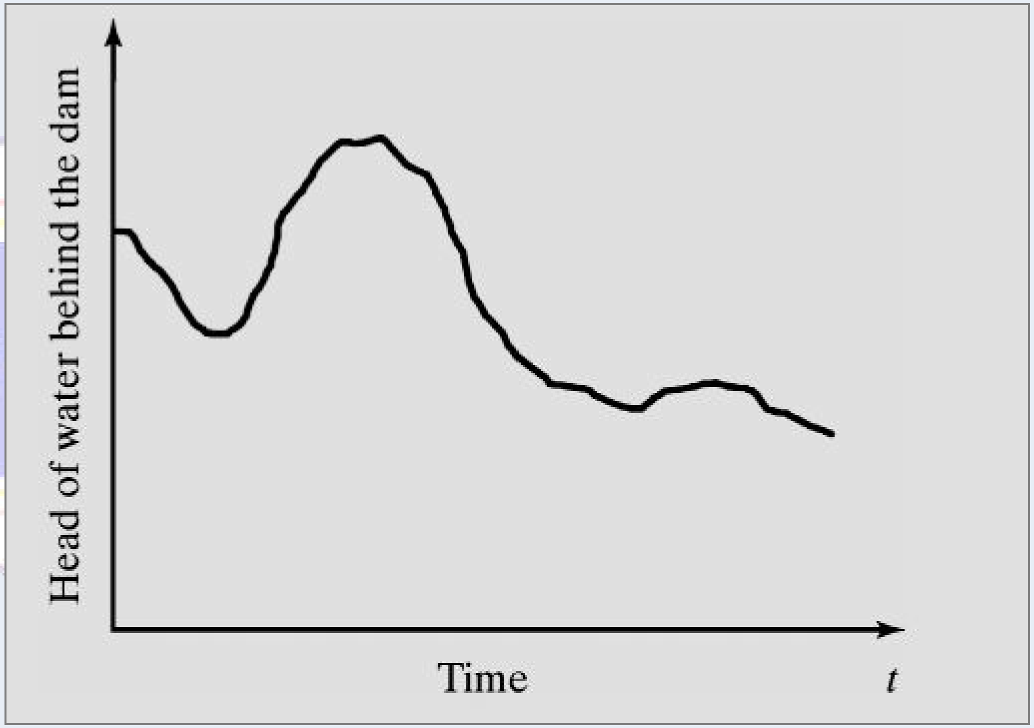
Discrete and Continues Systems

- **A discrete system** is one in which the *state variables change only at a discrete set of points in time* : Bank example



Discrete and Continues Systems (cont.)

- A **continues system** is one in which the state variables change continuously over time: Head of water behind the dam



Model of a System

- To study the system
  - it is sometimes possible to experiment
  - This is not always possible (bank, factory,…)
  - A new system may not yet exist

- Model:
  - **construct a conceptual framework that describes a system**
  - It is necessary to consider those accepts of systems that affect the problem under investigation (unnecessary details must remove)

Types of Models & System Models

- Models can be classified →
  mathematical, **physical** & **process.**
  - A mathematical model uses **symbolic notation and mathematical equations** to represent a system.
    - A simulation model is a particular type of mathematical model of a system.
- System Models may be further classified as:
  - STATIC OR DYNAMIC,
  - DETERMINISTIC OR STOCHASTIC,
  - DISCRETE OR CONTINUOUS

STATIC & DYNAMIC Models

- A STATIC SIMULATION MODEL :
  represents a system at a particular point in time.
  – sometimes called a Monte Carlo simulation

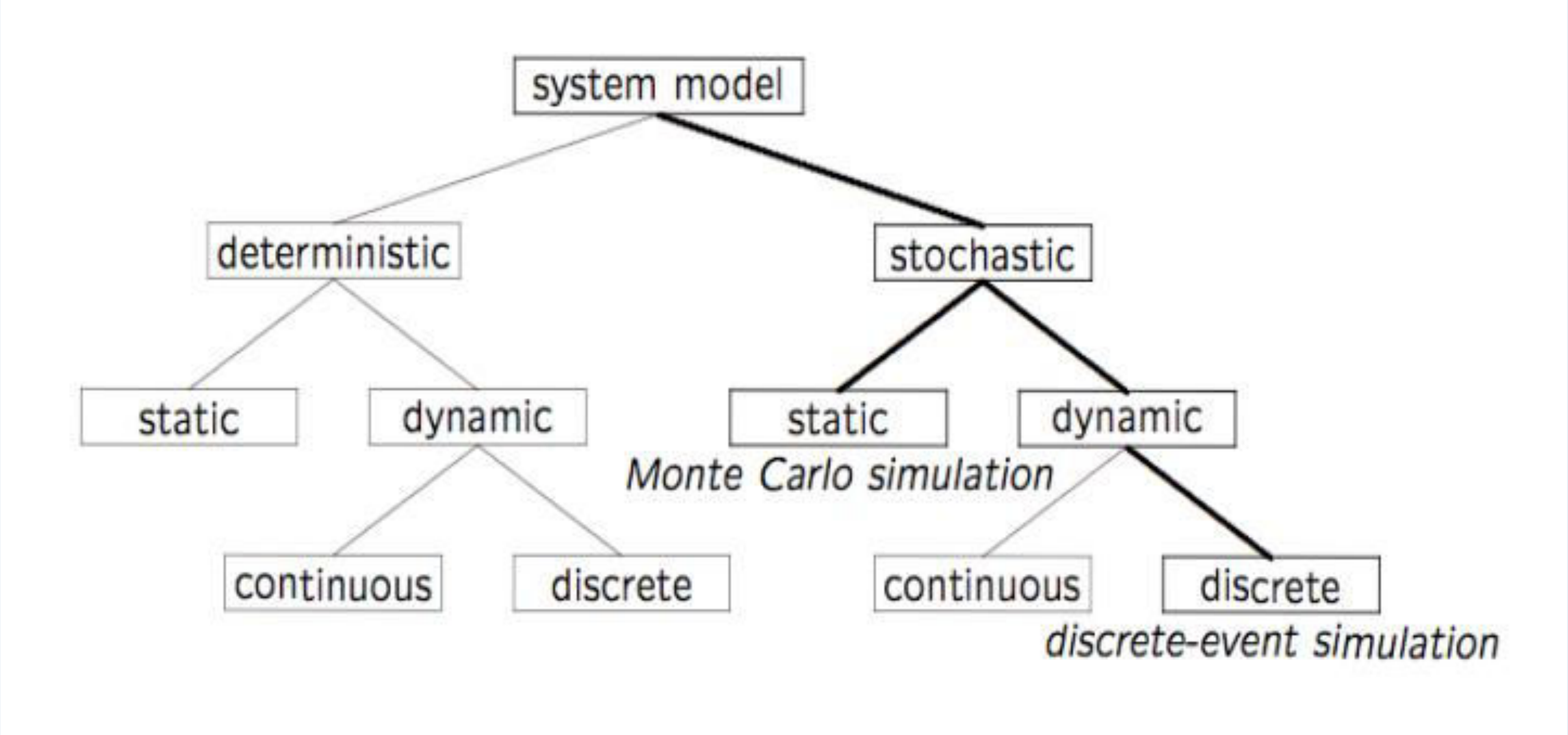
- DYNAMIC SIMULATION MODELS:
  represent systems as they change over time.
  – The simulation of a bank from 9:00 A.M. to 4:00 P.M. is an example of a dynamic simulation

DETERMINISTIC & STOCHASTIC

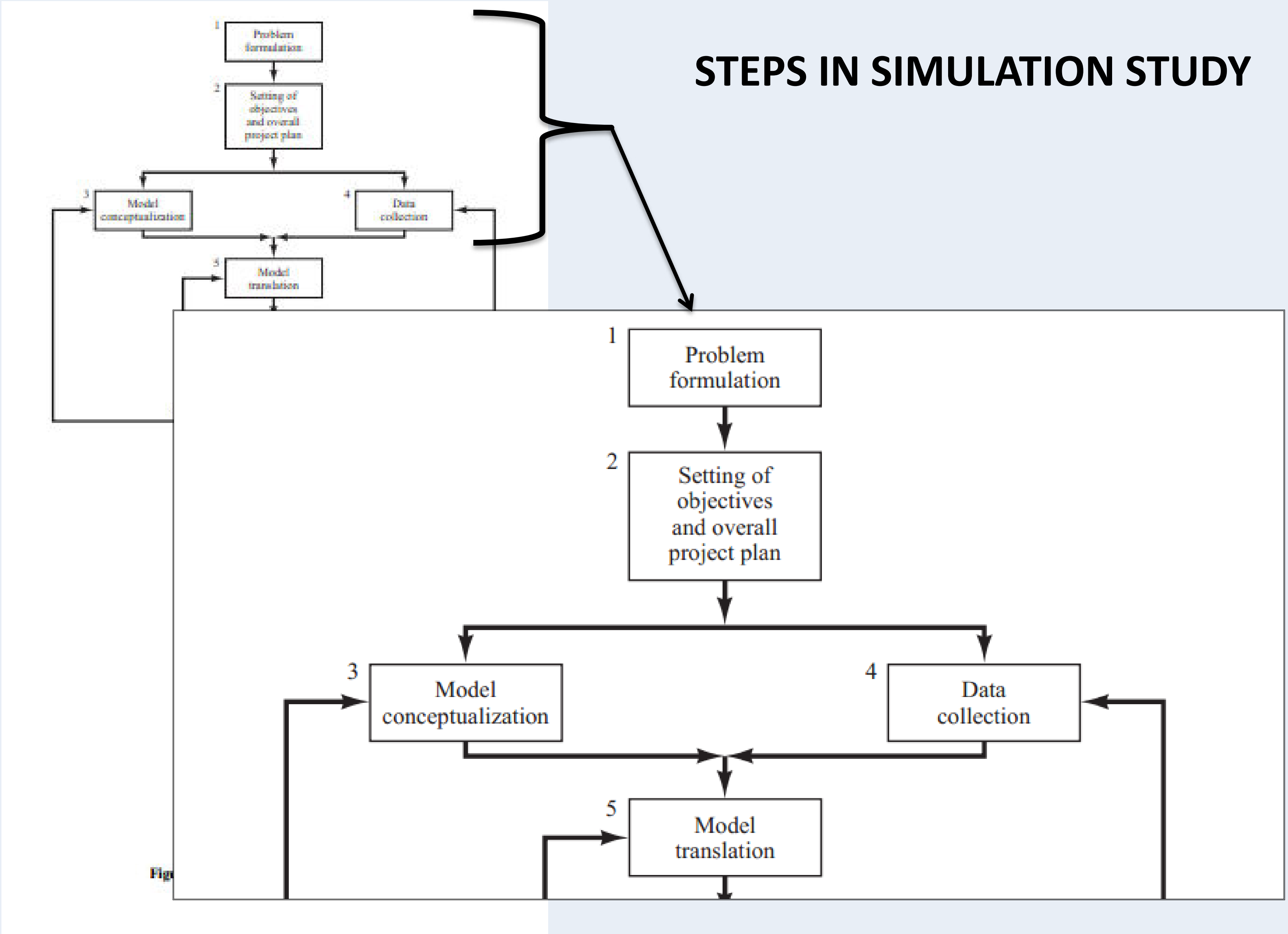
- Simulation models that **contain no random** variables are classified as **deterministic** .
  - Deterministic models have a known set of inputs which will result in a unique set of outputs.

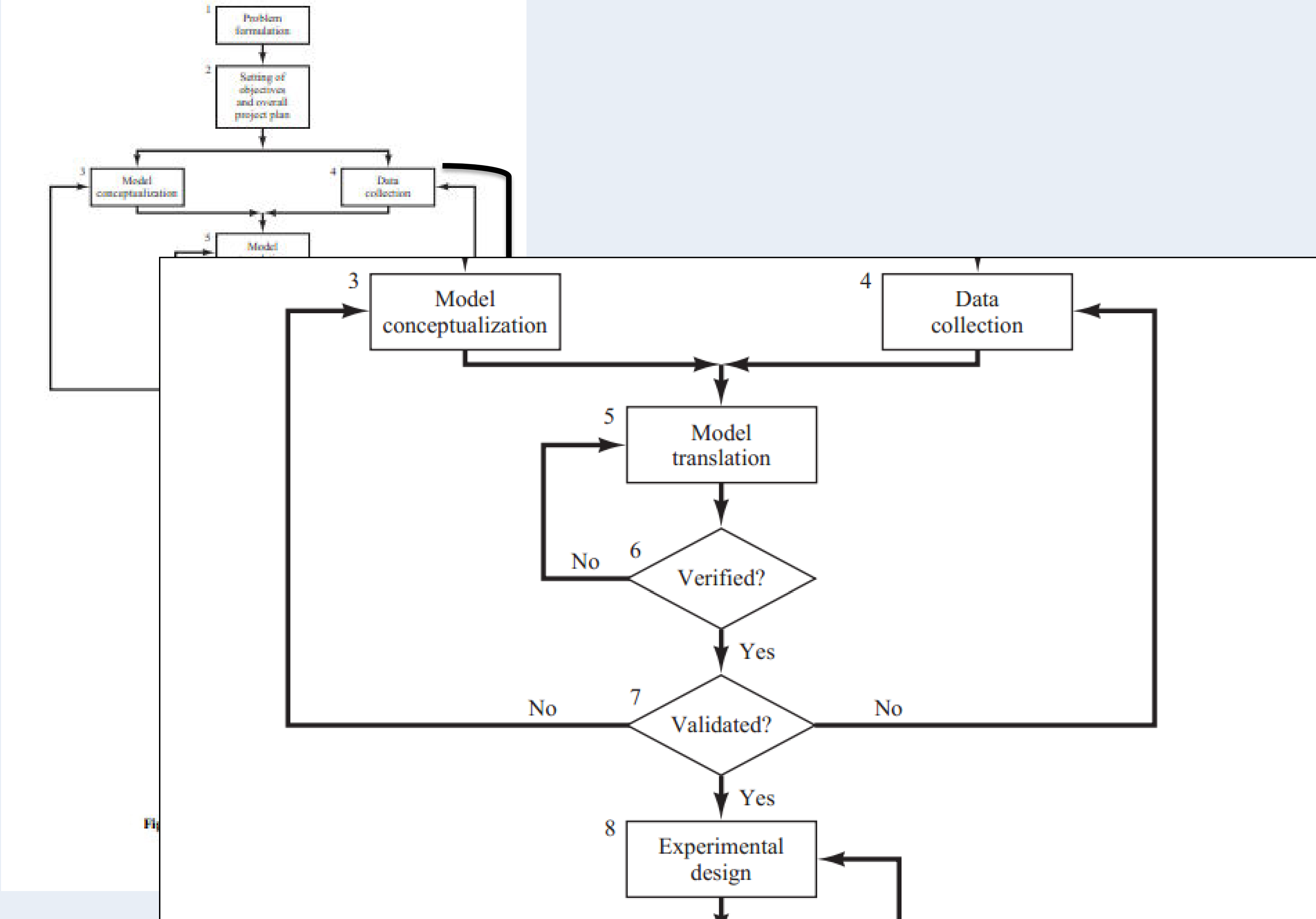
- A **stochastic** *(sto·chas·tic)* simulation model has **one or more random variables** as inputs.
  - Random inputs lead to random outputs.

Types of System Models



STEPS IN SIMULATION STUDY




1
Problem formulation
2
Setting of objectives and overall project plan
3
Model conceptualization
4
Data collection
3
Model conceptualization
4
Data collection
5
Model translation
6
No
Verified?
Yes
7
No
Validated?
No
Yes
8
Experimental design

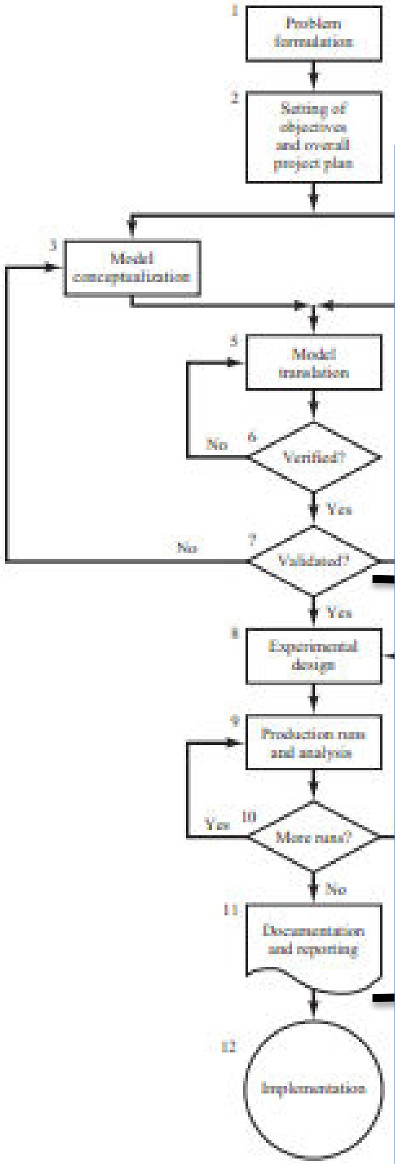
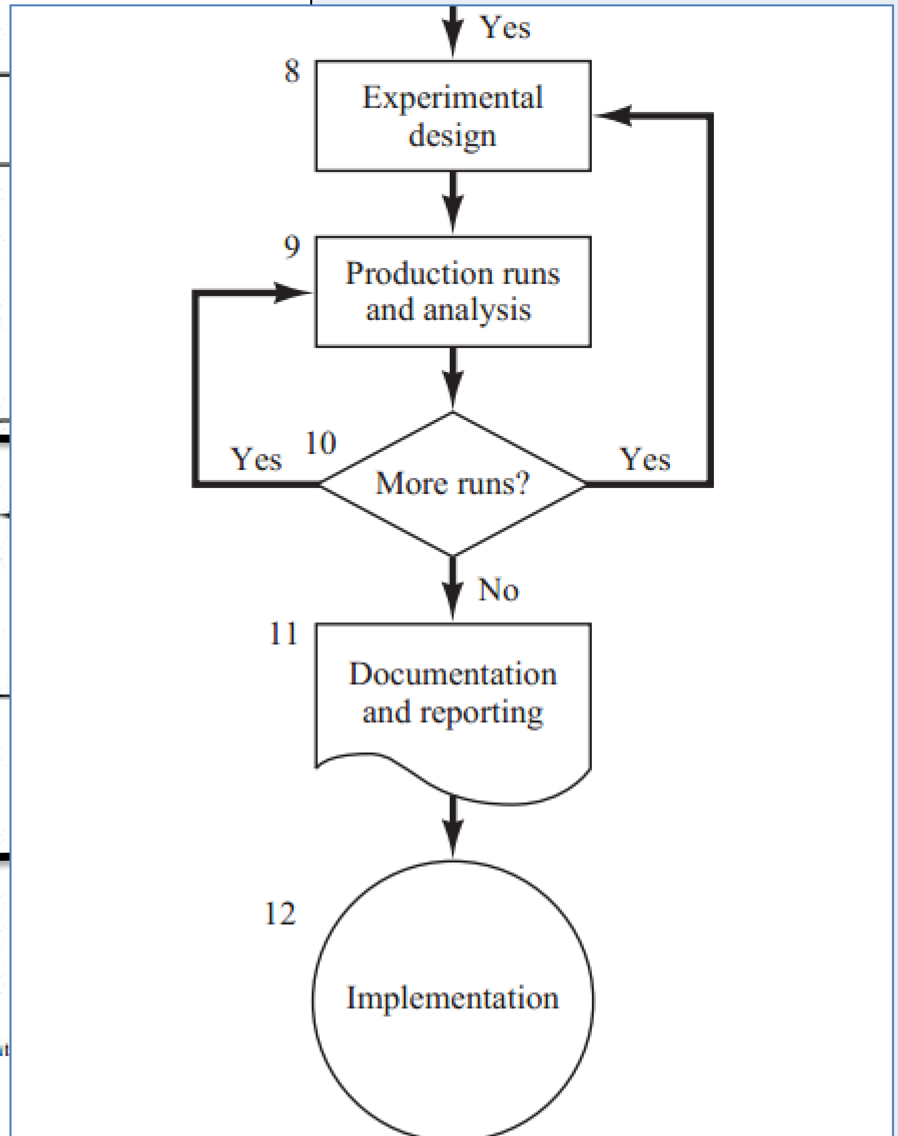


Figure 1.3 Steps in a simulat



Introduction To Modeling and Simulation

Lecture 1

What Is A Model ?

A Representation of an object, a system, or an idea in some form other than that of the entity itself.

(Shannon)

Types of Models:

Physical

(Scale models, prototype plants,…)

Mathematical

(Analytical queueing models, linear programs, simulation)

What is Simulation?

- A Simulation of a system is the operation of a model, which is a representation of that system.

open and responsive to suggestion

- The model is amenable to manipulation which would be impossible, too expensive, or too impractical to perform on the system which it portrays.
- The operation of the model can be studied, and, from this, properties concerning the behavior of the actual system can be inferred.

Applications:

- Designing and analyzing manufacturing systems
- Evaluating H/W and S/W requirements for a computer system
- Evaluating a new military weapons system
- Determining ordering policies for an inventory system
- Designing communications systems and message protocols for them

Applications:(continued)

- Designing and operating transportation facilities such as freeways, airports, subways, or ports
- Evaluating designs for service organizations such as hospitals, post offices, or fast-food restaurants
- Analyzing financial or economic systems

Steps In Simulation and Model Building

1. Define an achievable goal
2. Put together a complete mix of skills on the team
3. Involve the end-user
4. Choose the appropriate simulation tools
5. Model the appropriate level(s) of detail
6. Start early to collect the necessary input data

Steps In Simulation and Model Building(cont'd)

7. Provide adequate and on-going documentation
8. Develop a plan for adequate model verification

   (Did we get the "right answers ?")

9. Develop a plan for model validation

   (Did we ask the "right questions ?")

10. Develop a plan for statistical output analysis

Define An Achievable Goal

"To model the…" is NOT a goal!

"To model the…in order to select/determine feasibility/…is a goal.

Goal selection is not cast in concrete

Goals change with increasing insight

Put together a complete mix of skills on the team

We Need:

-Knowledge of the system under investigation

-System analyst skills (model formulation)

-Model building skills (model Programming)

-Data collection skills

-Statistical skills (input data representation)

Put together a complete mix of skills on the team(continued)

We Need:

-More statistical skills (output data analysis)

-Even more statistical skills (design of experiments)

-Management skills (to get everyone pulling in the same direction)

INVOLVE THE END USER

-Modeling is a selling job!

-Does anyone believe the results?

-Will anyone put the results into action?

-The End-user (your customer) can (and must) do all of the above BUT, first he must be convinced!

-He must believe it is HIS Model!

Choose The Appropriate Simulation Tools

Assuming Simulation is the appropriate means, three alternatives exist:

1. Build Model in a General Purpose Language

2. Build Model in a General Simulation Language

3. Use a Special Purpose Simulation Package

MODELLING W/ GENERAL PURPOSE LANGUAGES

- Advantages:
  – Little or no additional software cost
  – Universally available (portable)
  – No additional training (Everybody knows…(language X) !)
- Disadvantages:
  – Every model starts from scratch
  – Very little reusable code
  – Long development cycle for each model
  – Difficult verification phase

GEN. PURPOSE LANGUAGES USED FOR SIMULATION

FORTRAN
- Probably more models than any other language.

PASCAL
- Not as universal as FORTRAN

MODULA
- Many improvements over PASCAL

ADA
- Department of Defense attempt at standardization

C, C++
- Object-oriented programming language

MODELING W/ GENERAL SIMULATION LANGUAGES

- Advantages:
  - Standardized features often needed in modeling
  - Shorter development cycle for each model
  - Much assistance in model verification
  - Very readable code
- Disadvantages:
  - Higher software cost (up-front)
  - Additional training required
  - Limited portability

GENERAL PURPOSE SIMULATION LANGUAGES

- GPSS
  - Block-structured Language
  - Interpretive Execution
  - FORTRAN-based (Help blocks)
  - *World-view:* Transactions/Facilities
- SIMSCRIPT II.5
  - English-like Problem Description Language
  - Compiled Programs
  - Complete language (no other underlying language)
  - *World-view:* Processes/ Resources/ Continuous

GEN. PURPOSE SIMULATION LANGUAGES (continued)

- MODSIM III
  - Modern Object-Oriented Language
  - Modularity Compiled Programs
  - Based on Modula2 (but compiles into C)
  - *World-view:* Processes
- SIMULA
  - ALGOL-based Problem Description Language
  - Compiled Programs
  - *World-view:* Processes

GEN. PURPOSE SIMULATION LANGUAGES (continued)

- SLAM
  - Block-structured Language
  - Interpretive Execution
  - FORTRAN-based (and extended)
  - *World-view:* Network / event / continuous
- CSIM
  - *process-oriented language*
  - *C-based* (C++ based)
  - *World-view:* Processes

MODELING W/ SPECIAL-PURPOSE SIMUL. PACKAGES

- Advantages
  - Very quick development of complex models
  - Short learning cycle
  - No programming--minimal errors in usage
- Disadvantages
  - High cost of software
  - Limited scope of applicability
  - Limited flexibility (may not fit your specific application)

THE REAL COST OF SIMULATION

Many people think of the cost of a simulation only in terms of the software package price.

There are actually at least three components to the cost of simulation:

1. Purchase price of the software
2. Programmer / Analyst time
3. "Timeliness of Results"

TERMINOLOGY

- System
  - A group of objects that are joined together in some regular interaction or interdependence toward the accomplishment of some purpose.
  - Entity
  - An object of interest in the system.
  - E.g., customers at a bank

TERMINOLOGY (continued)

- Attribute
  - a property of an entity
  - E.g., checking account balance
- Activity
  - Represents a time period of specified length.
  - Collection of operations that transform the state of an entity
  - E.g., making bank deposits

TERMINOLOGY (continued)

- Event:
  - change in the system state.
  - E.g., arrival; beginning of a new execution; departure
- State Variables
  - Define the state of the system
  - Can restart simulation from state variables
  - E.g., length of the job queue.

TERMINOLOGY (continued)

- Process
  - Sequence of events ordered on time
- Note:
  - the three concepts(event, process,and activity) give rise to three alternative ways of building discrete simulation models

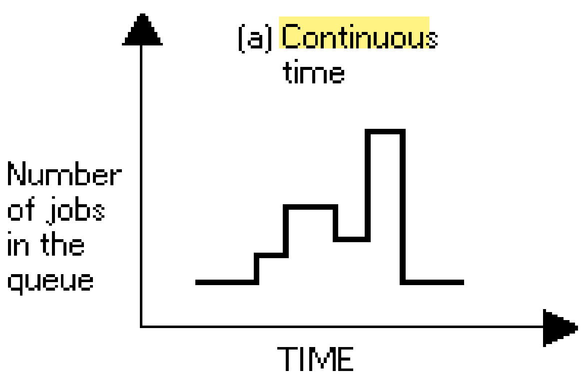
SIMULATION "WORLD-VIEWS"

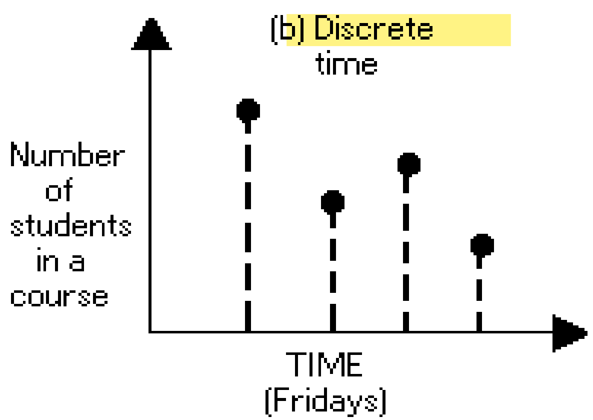
- Pure Continuous Simulation
- Pure Discrete Simulation
  - Event-oriented
  - Activity-oriented
  - Process-oriented
- Combined Discrete / Continuous Simulation

Examples Of Both Type Models

- Continuous Time and Discrete Time Models:

  CPU scheduling model vs. number of students attending the class.





MODEL THE APPROPRIATE LEVEL(S) OF DETAIL

- Define the boundaries of the system to be modeled.
- Some characteristics of “the environment” (outside the boundaries) may need to be included in the model.
- Not all subsystems will require the same level of detail.
- Control the tendency to model in great detail those elements of the system which are well understood, while skimming over other, less well - understood sections.

START EARLY TO COLLECT THE NECESSARY INPUT DATA

Data comes in two quantities:

TOO MUCH!!

TOO LITTLE!!

With too much data, we need techniques for reducing it to a form usable in our model.

With too little data, we need information which can be represented by statistical distributions.

PROVIDE ADEQUATE AND ON-GOING DOCUMENTATION

In general, programmers hate to document.

Documentation is always their lowest priority item.

They believe that "only wimps read manuals."

What can we do?

- Use self-documenting languages
- Insist on built-in user instructions(help screens)
- Set (or insist on) standards for coding style

DEVELOP PLAN FOR ADEQUATE MODEL VERIFICATION

Did we get the "right answers?"

(No such thing!!)

Simulation provides something that no other technique does:

Step by step tracing of the model execution.

This provides a very natural way of checking the internal consistency of the model.

DEVELOP A PLAN FOR MODEL VALIDATION

VALIDATION: "Doing the right thing"

Or "Asking the right questions"

How do we know our model represents the system under investigation?

- Compare to existing system?
- Deterministic Case?

DEVELOP A PLAN FOR STATISTICAL OUTPUT ANALYSIS

- How much is enough?

  Long runs versus Replications

- Techniques for Analysis

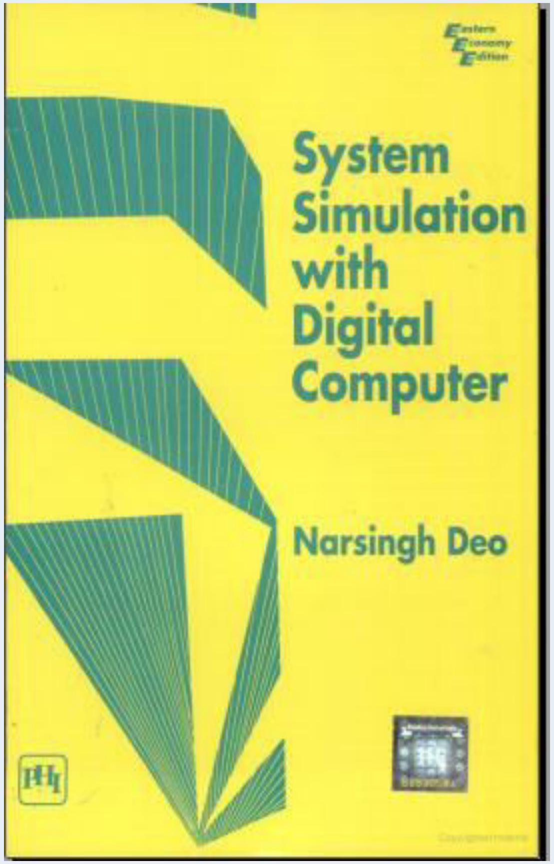
Class-2

Simulation and Modeling

CSE4131_CSE4132

Few examples…

- Reference Book:
  - System Simulation with Digital Computer
  - By: Narsingh Deo
- Chapter No.: 1 (Introduction)



Simulation of a pure pursuit problem

- Pure Pursuit
  - Pure: Clean or unpolluted
  - Pursuit: Chase or Hunt

- Here we are going to study some hunting or chasing problem.

Simulation of a pure pursuit problem

- Problem Statement:
  - A fighter aircraft sights an enemy bomber and flies directly toward it, in order to catch up with the bomber and destroy it. The bomber (the target) continues flying (along a specified curve) so the fighter (the pursuer) has to change its direction to keep pointed toward the target. <u>*We are interested in determining the attack course of the fighter and in knowing how long it would take for it to catch up with the bomber*</u>
- Target – Bomber
- Pursuer – Fighter

Simulation of a pure pursuit problem

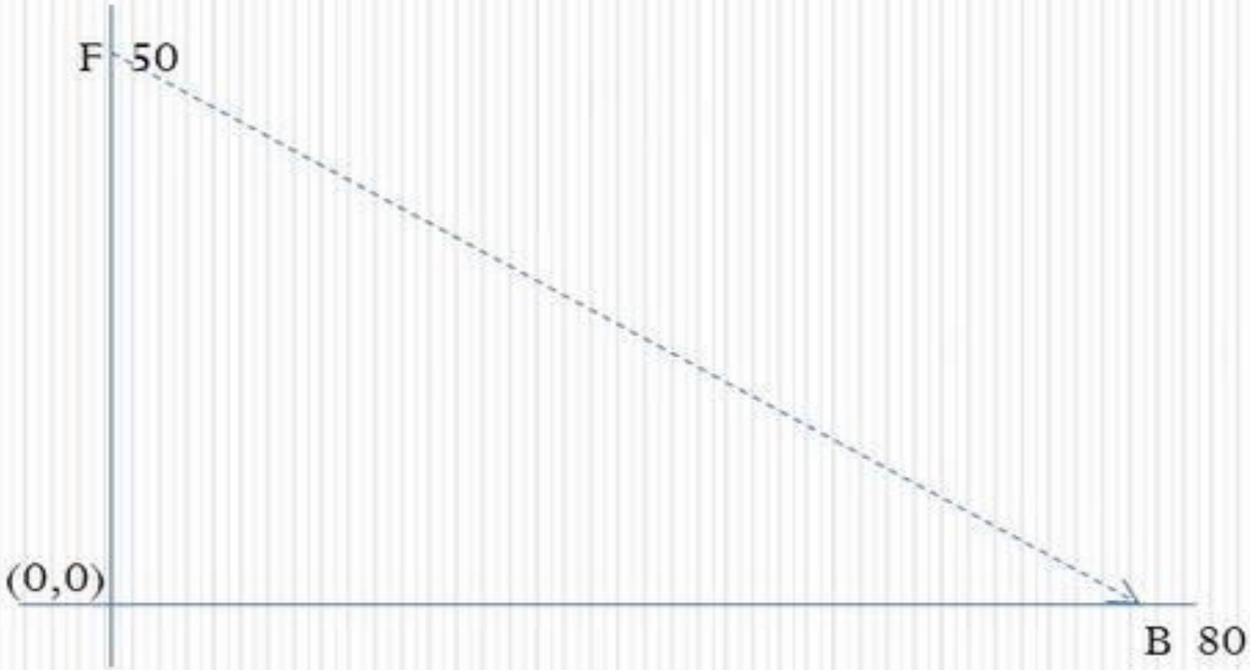
- If the target flies along a straight line, the problem can be solved directly with analytic techniques.
- However, if the path of the target is curved, the problem is much more difficult and normally cannot be solved directly.

Simulation of a pure pursuit problem

- Simplifying conditions:
  - The target and the pursuer are flying in the **same horizontal plane** when the fighter first sights the bomber, and both stay in that plane. This makes the pursuit model 2-D.
  - The **fighter's speed VF is constant** (20 kms/minute).
  - The **target's** path is **specified**.
  - After a fixed time span at (every minute, in this case) the fighter changes its direction in order to point itself toward the bomber.
  - The minimum distance required by the fighter to fire missile at bomber is 10 units.

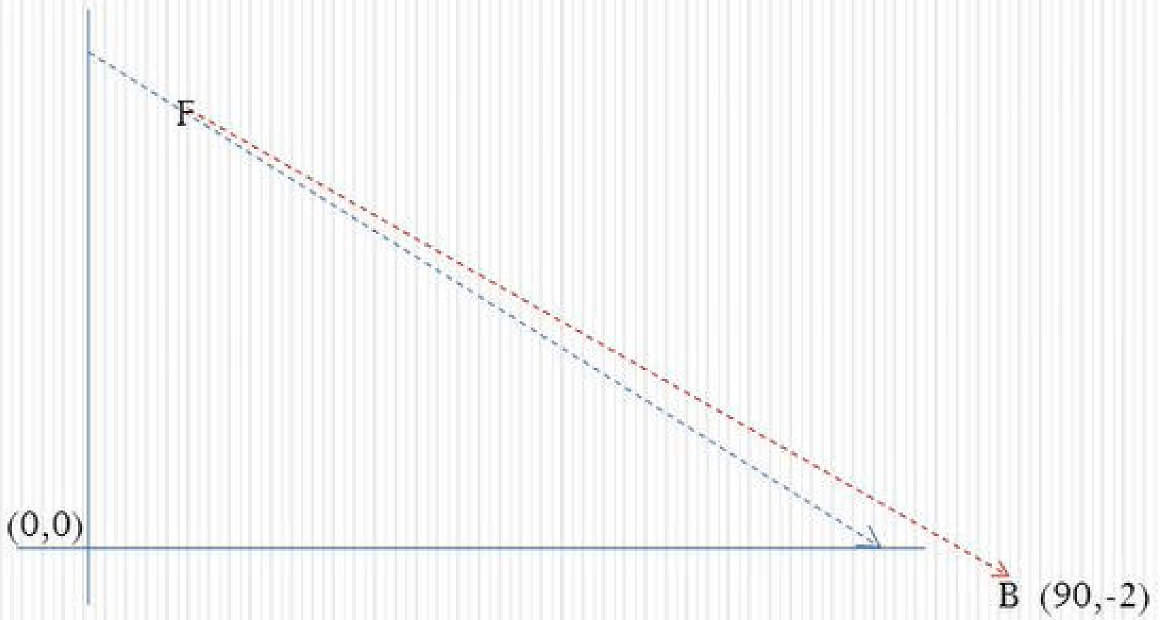
Simulation of a pure pursuit problem

- Figure showing F and B at time 0 and the direction of motion of F



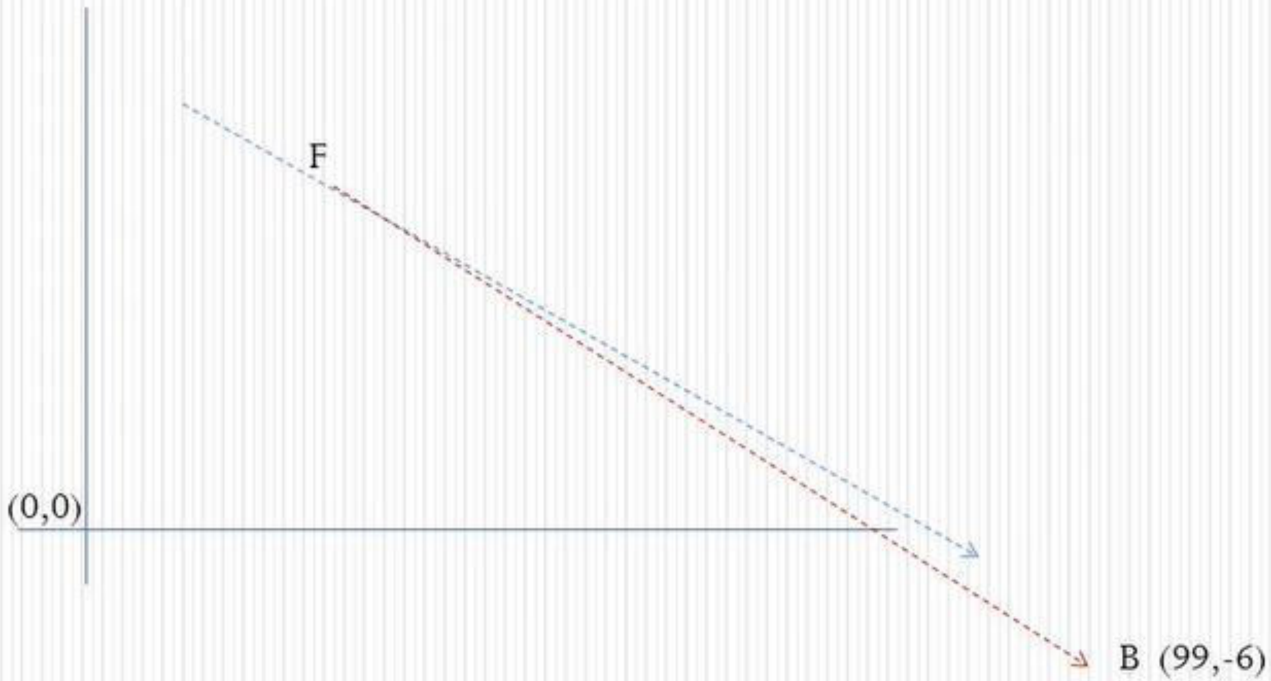
Simulation of a pure pursuit problem

- Figure showing F and B at time 1 and time 2 and the direction of motion of F



Simulation of a pure pursuit problem

- Figure showing F and B at time 2 and the direction of motion of F



Simulation of a pure pursuit problem

- The Coordinate of Bomber is known to us

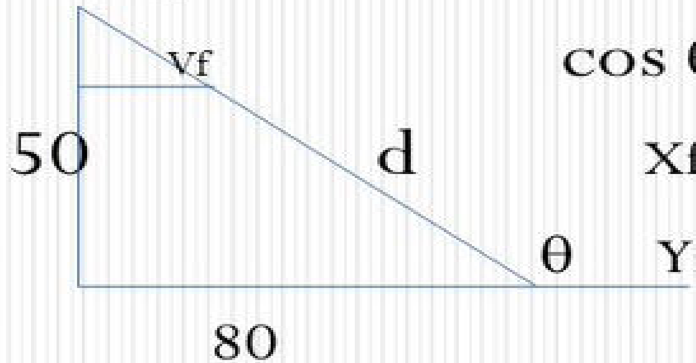
Time, t	0	1	2	3	4	5	6	7	8	9	10	11	12
$XB(t)$	80	90	99	108	116	125	133	141	151	160	169	179	180
$YB(t)$	0	−2	−5	−9	−15	−18	−23	−29	−28	−25	−21	−20	−17

Table 1-1.

Simulation of a pure pursuit problem

Fighter's position

- At time 0 fighter is at (0,50)
- $Xf(0)=0, Yf(0)=50$
- At time 0 bomber is at (80,0)
- Distance between F and B is sqrt($(80-0)^2+(0-50)^2$)) = d

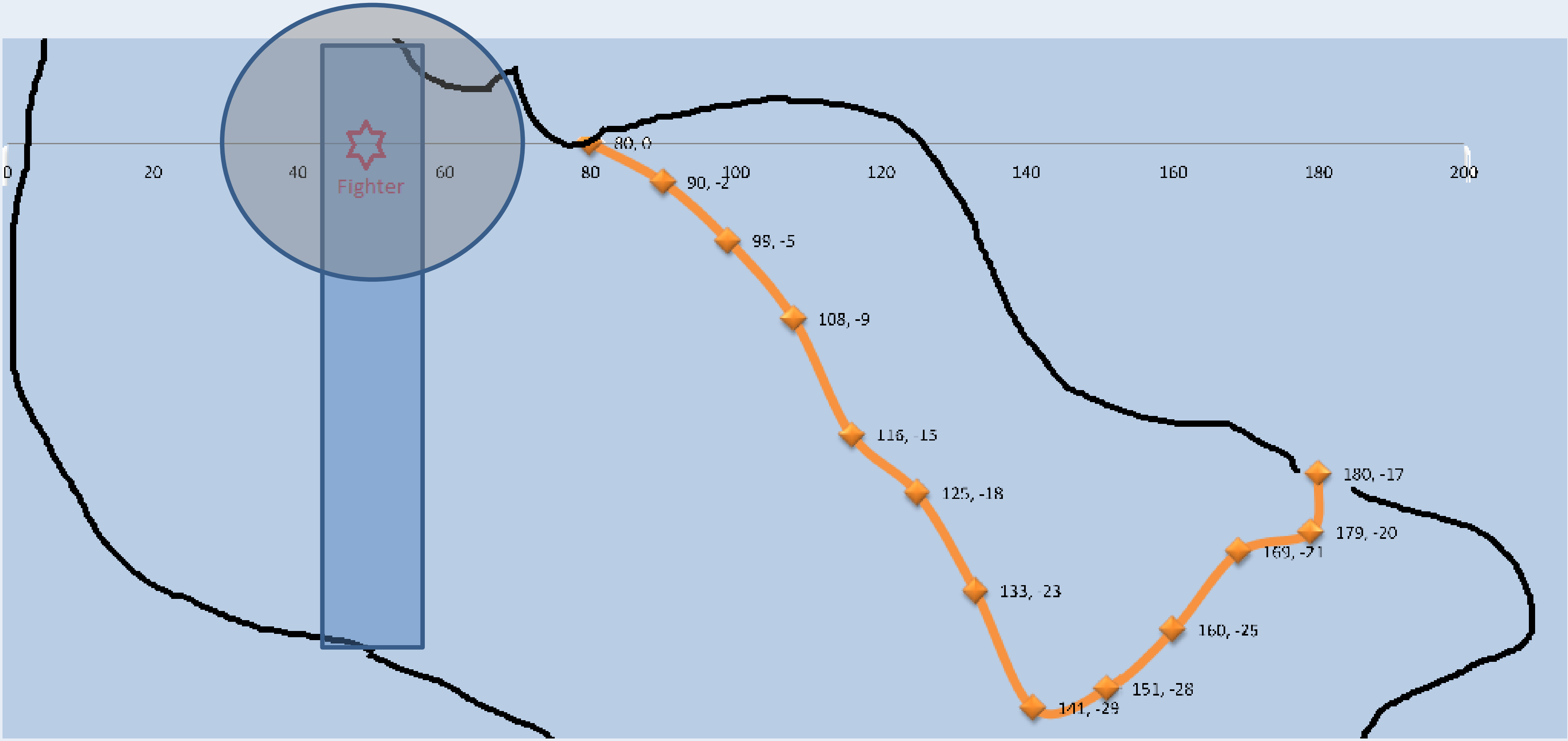


$\cos\theta = 80/d \quad \sin\theta = -50/d$

$Xf(1)=Xf(0)+Vf^{*} \cos\theta$

$Yf(1)=Yf(0)+Vf^{*} \sin\theta$

Simulation of a pure pursuit problem



Simulation of a pure pursuit problem

General formula

- $Xf(t+1)=Xf(t)+Vf*\cos\theta$
- $Yf(t+1)=Yf(t)+Vf*\sin\theta$
- $\sin\theta = [Yb(t)-Yf(t)]/Dist(t)$
- $\cos\theta = [Xb(t)-Xf(t)]/Dist(t)$

Simulation of a pure pursuit problem

The angle θ of the line from the fighter to the target at a given time t is

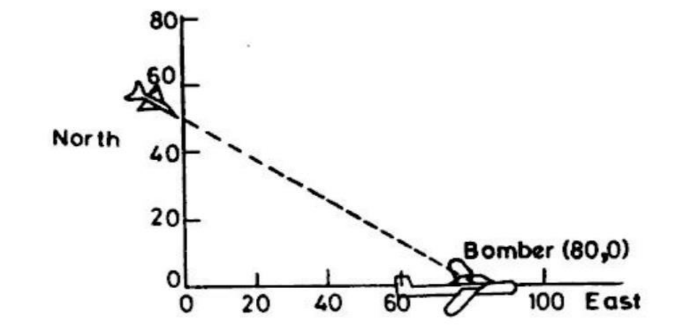


Fig. 1-1: Positions of Pursuer and Target at Time Zero.

$$\text{Sin}\,\theta = \frac{YB(t) - YF(t)}{\text{DIST}(t)} \quad \ldots(1\text{-}2)$$

$$\text{Cos}\,\theta = \frac{XB(t) - XF(t)}{\text{DIST}(t)} \quad \ldots(1\text{-}3)$$

Using this value of the position of the fighter at time $(t + 1)$ is determined by

$$XF(t+1) = XF(t) + VF\,\text{Cos}\,\theta \quad \ldots(1\text{-}4)$$

$$YF(t+1) = YF(t) + VF\,\text{Sin}\,\theta \quad \ldots(1\text{-}5)$$

Simulation of a pure pursuit problem

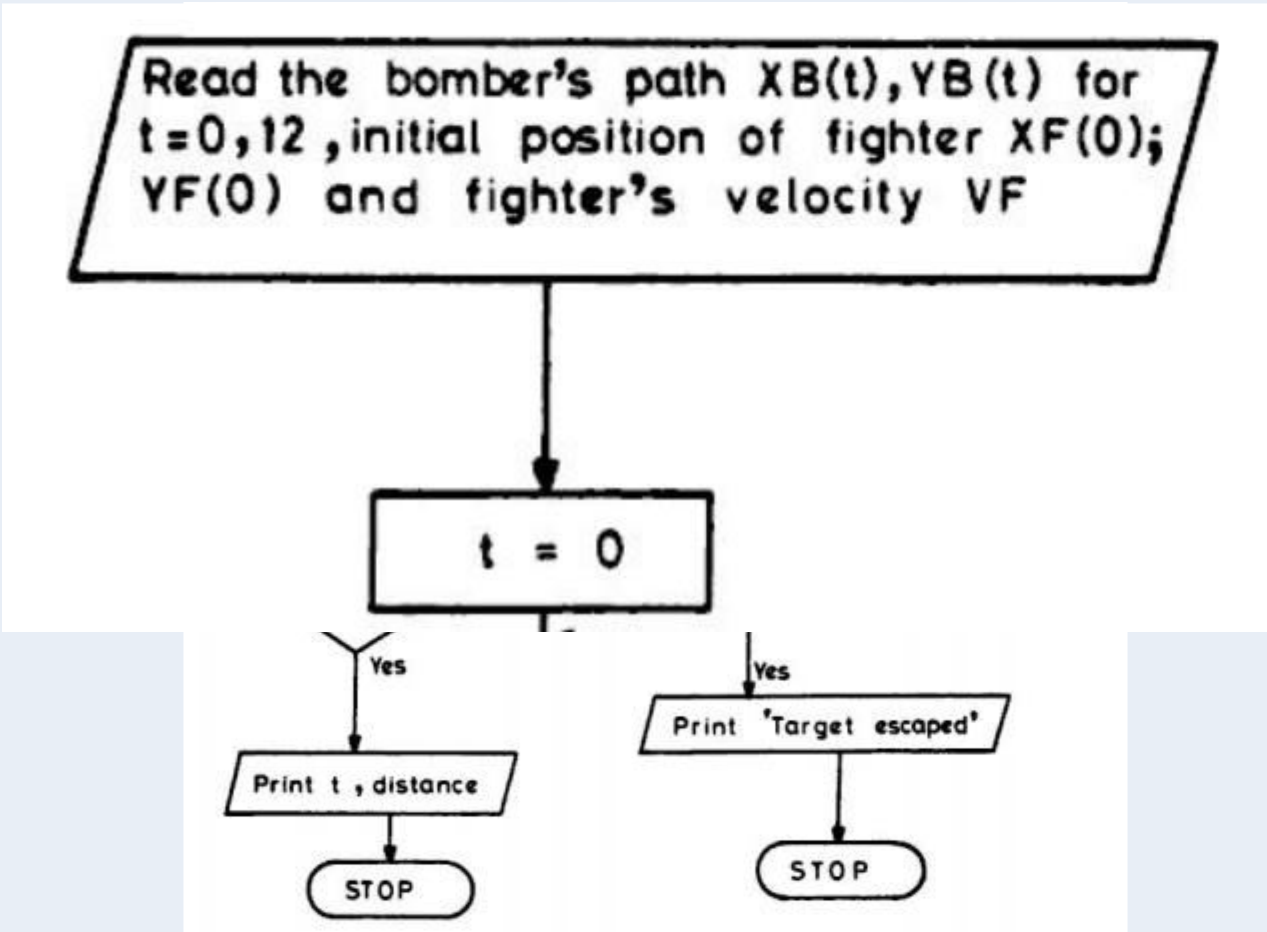


Fig. 1-2: Flowchart of pursuit problem.

Simulation of a pure pursuit problem

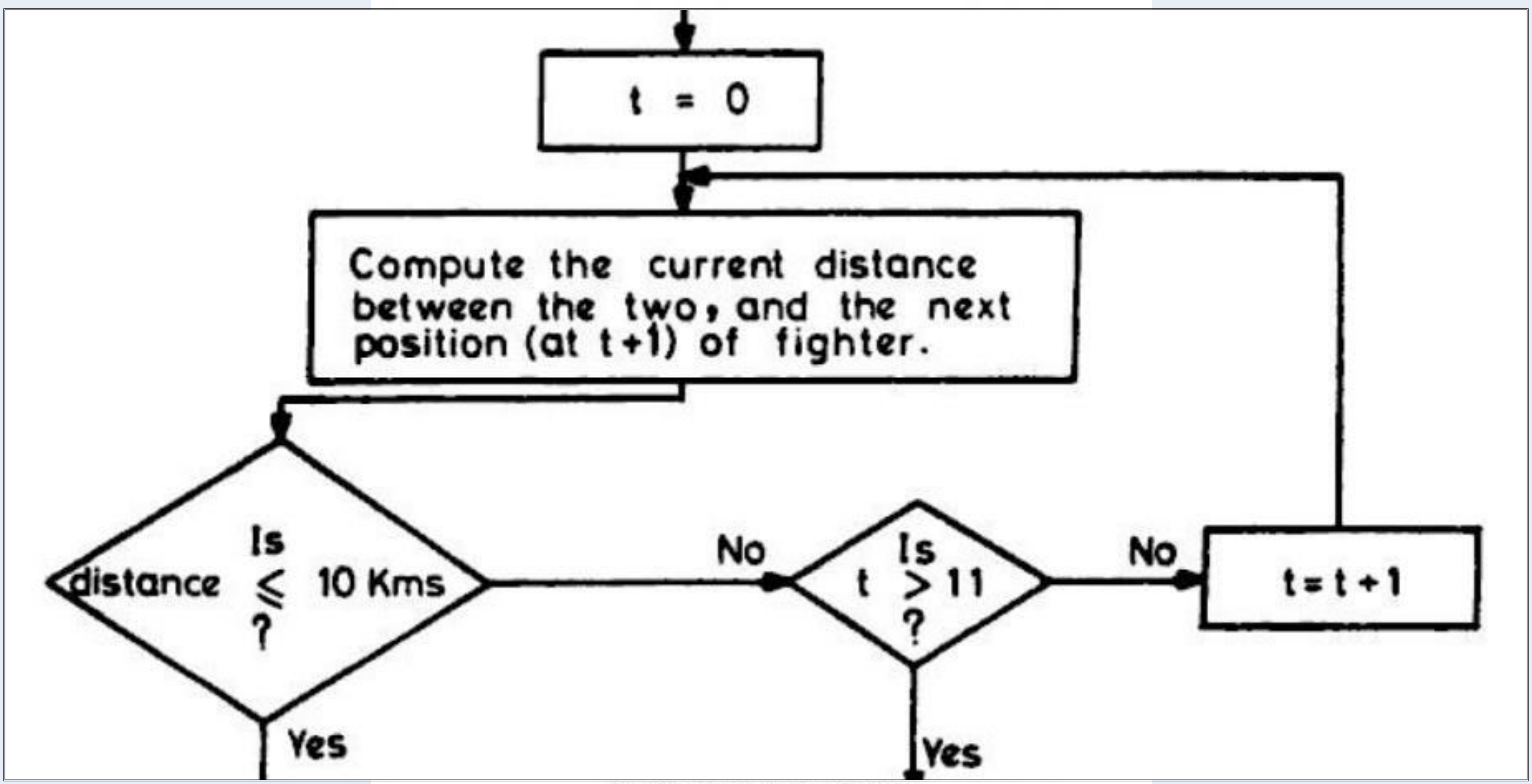


Fig. 1-2: Flowchart of pursuit problem.

Simulation of a pure pursuit problem

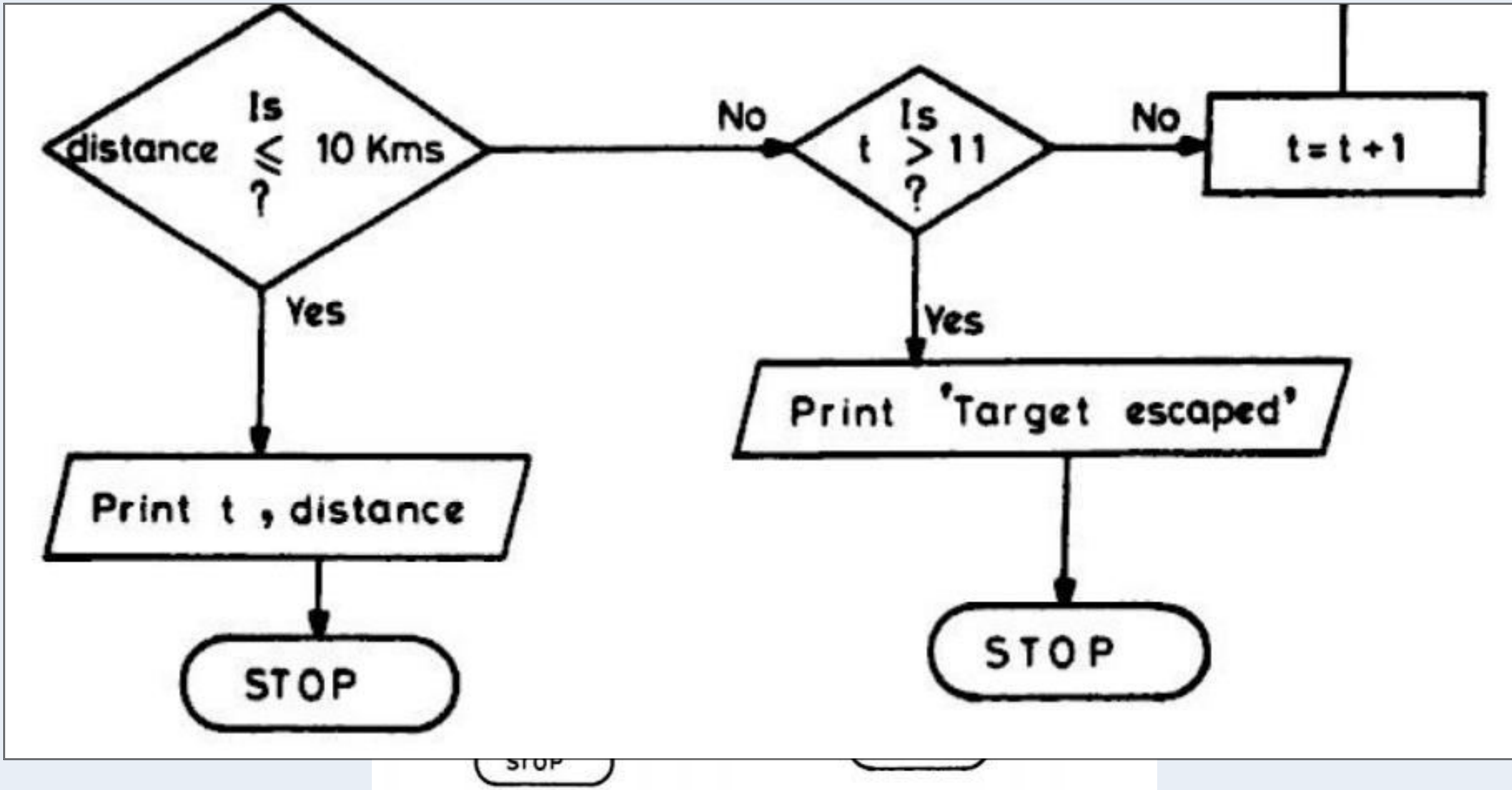


Fig. 1-2: Flowchart of pursuit problem.

Simulation of a pure pursuit problem

1. Read endt, thresh, vf
2. Read xb(t), yb(t) for t=1 to endt
3. Read xf(0), yf(0)
4. For t=1 to endt
5. $dist=sqrt((xb(t)-xf(t))^2+(yb(t)-yf(t)^2))$
6. If dist < thresh then exit
   Else
7. $\sin\theta = [yb(t)-yf(t)]/dist$
8. $\cos\theta = [xb(t)-xf(t)]/dist$
9. t=t+1
10. $xf(t+1)=xf(t)+vf*\cos\theta$
11. $yf(t+1) = yf(t)+vf*\sin\theta$
    End for
12. If t > endt then Print "target escaped", t
13. Stop
    Else
14. Print "target caught"
15. End

Class-3

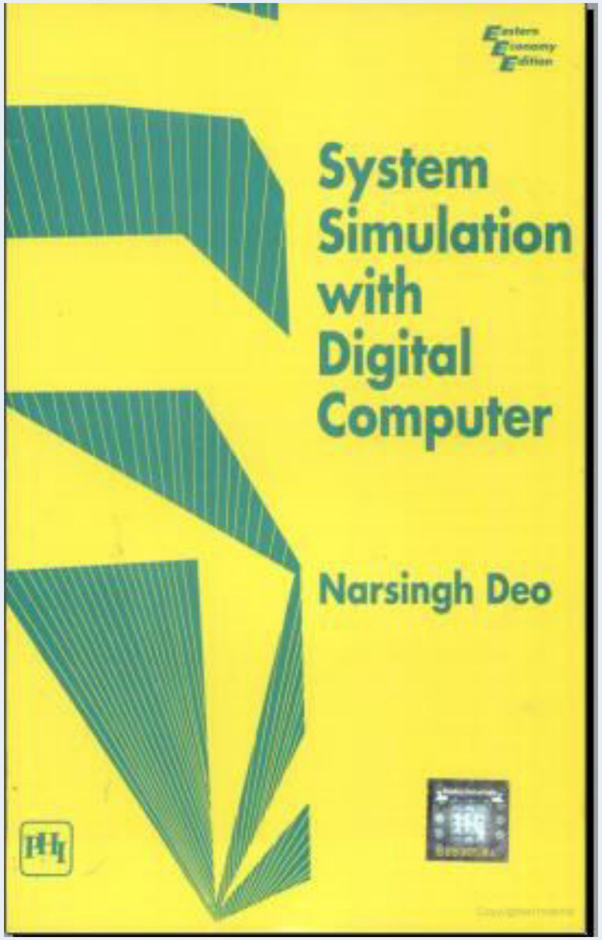
CSE4131_CSE4132

Simulation and Modeling

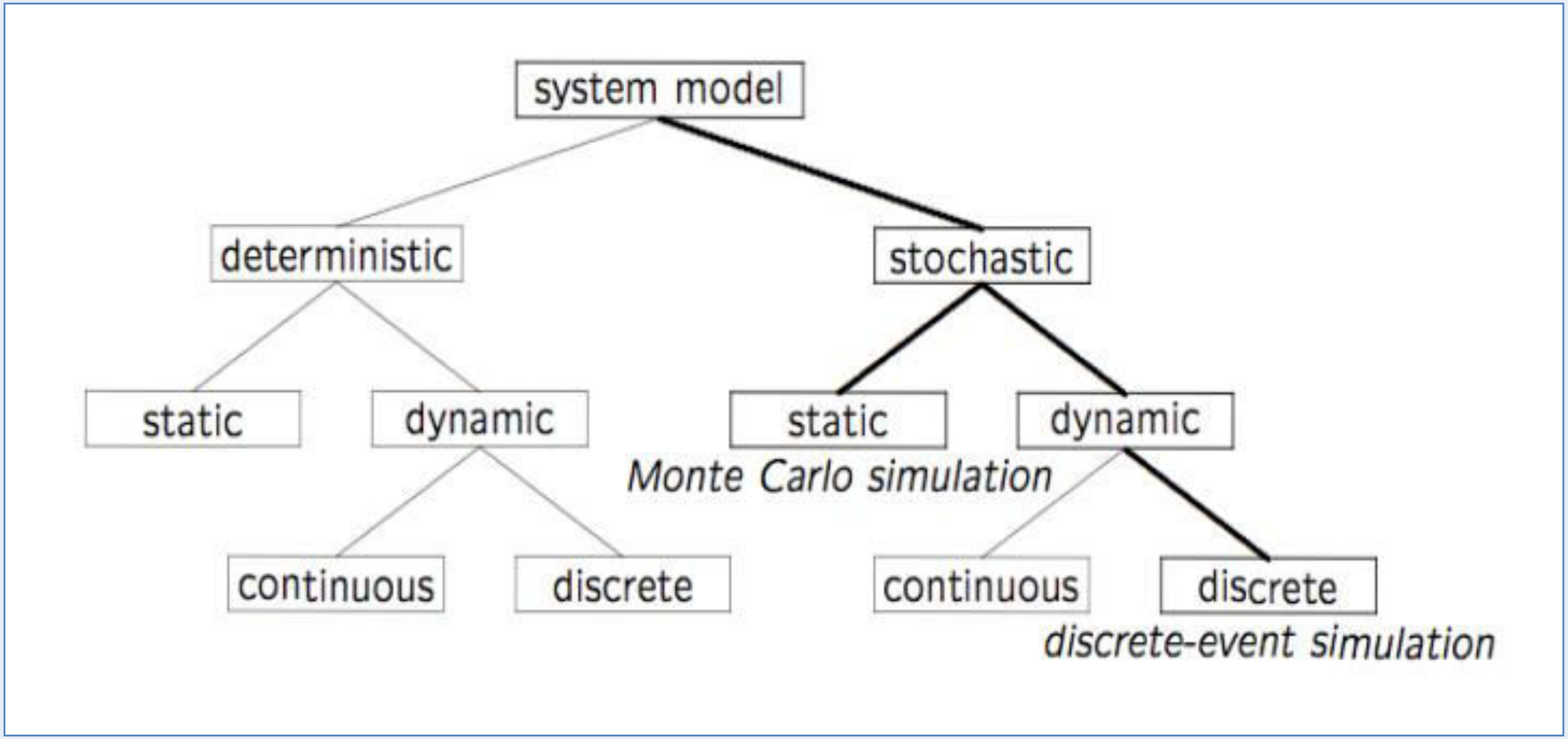
CSE4131_CSE4132

Few examples…

- Reference Book:
  - System Simulation with Digital Computer
  - By- N. Deo
- Chapter No.: 1 (Introduction)



Classifications



An other Example

- Simulation of a Chemical Reactor
  - Its an example of a **continuous system**
  - Its an example of **dynamic system**
  - Question may come:

As an elementary example of a continuous dynamic system, describe the chemical reactor plant.

Problem Statement

- In a chemical reactor:
  - There are two substances: A & B
  - A + B = third chemical substance C
  - Its known:
    - **1 gram of A + 1 gram of B ↔ produces 2 grams of C**
  - The rate of formation of C is:
    - proportional to presence of A & B
  - Backward reaction - decomposition of C back in to A & B:
    - The rate of decomposition proportional to the amount of C present in the reactor

Simulation of a Chemical Reactor

- Let a, b & c … are quantities of A, B & C at any time t
  - Then the rate of increase…

$$\frac{da}{dt} = k_2c - k_1ab, \qquad \ldots(2\text{-}1)$$

$$\frac{db}{dt} = k_2c - k_1ab, \qquad \ldots(2\text{-}2)$$

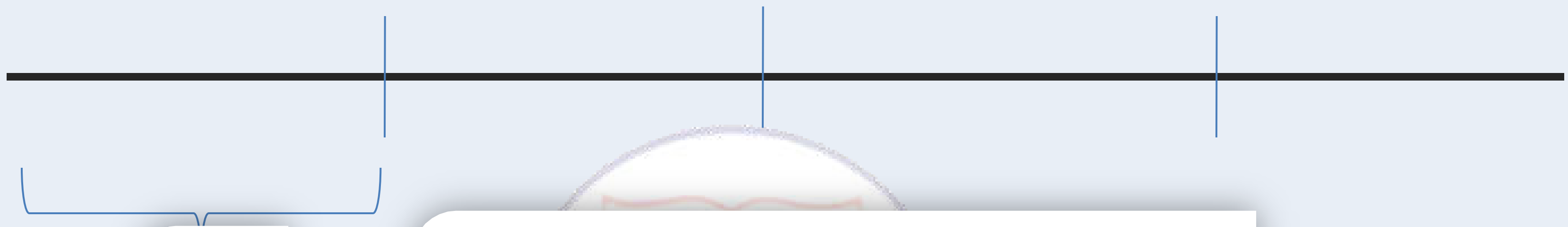
$$\frac{dc}{dt} = 2k_1ab - 2k_2c, \qquad \ldots(2\text{-}3)$$

where k_1 and k_2 are the *rate constants.*

Simulation of a Chemical Reactor

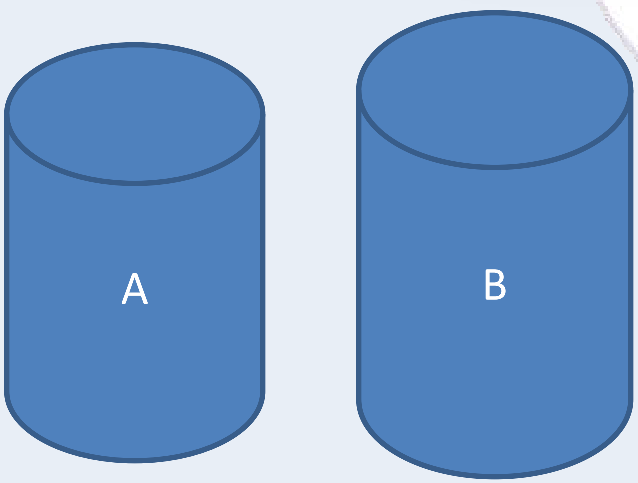
- Given the values of the
  - constants k_1 and k_2 and
  - the initial quantities of the chemicals A and B (and C = 0),
- we wish to determine **how much of C has been produced as a function of time**.
  - Determination of the rate of such chemical reactions is important in many industrial processes
- A straightforward method of simulating this system is to start at time zero and
  - increment time in small steps of Δt.

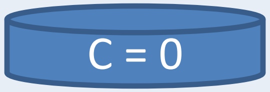
Simulation of a Chemical Reactor



Δt

$$a(t+\Delta t) = a(t) + \frac{da(t)}{dt} \cdot \Delta t$$





Simulation of a Chemical Reactor

- We assume that the quantities of chemicals remain unaltered during each step and only change 'instantaneously' at the end of the step.
- Thus the quantity of A (or B or C) at the end of one such step is given in terms of the quantity at the beginning of the step as

$$a(t + \Delta t) = a(t) + \frac{da(t)}{dt} \cdot \Delta t$$

- Identical equitation can be written for **b(t+Δt) and c(t+Δt)**

Simulation of a Chemical Reactor

- Suppose we want to do the simulation for a period **T**
  - We will divide T in to a large number N of small periods Δt
  - That is,

$$T = N . \Delta t$$

- At the time zero, we know a(0), b(0) and c(0)
- We also know k_1 & k_2

Simulation of a Chemical Reactor

$$\frac{da}{dt} = k_2 c - k_1 ab, \qquad \ldots(2\text{-}1)$$

$$a(t + \Delta t) = a(t) + \frac{da(t)}{dt} . \Delta t$$

- From these values we compute the amount of chemical at the time Δt :

$$a(\Delta t) = a(0) + [k_2 . c(0) - k_1 . a(0) . b(0)]\, \Delta t$$
$$b(\Delta t) = b(0) + [k_2 . c(0) - k_1 . a(0) . b(0)]\, \Delta t$$
$$c(\Delta t) = c(0) + [2k_1 . a(0) . b(0) - 2k_2 . c(0)]\, \Delta t$$

Simulation of a Chemical Reactor

$$a(\Delta t) = a(0) + [k_2 . c(0) - k_1 . a(0) . b(0)]\, \Delta t$$

$$b(\Delta t) = b(0) + [k_2 . c(0) - k_1 . a(0) . b(0)]\, \Delta t$$

$$c(\Delta t) = c(0) + [2k_1 . a(0) . b(0) - 2k_2 . c(0)]\, \Delta t$$

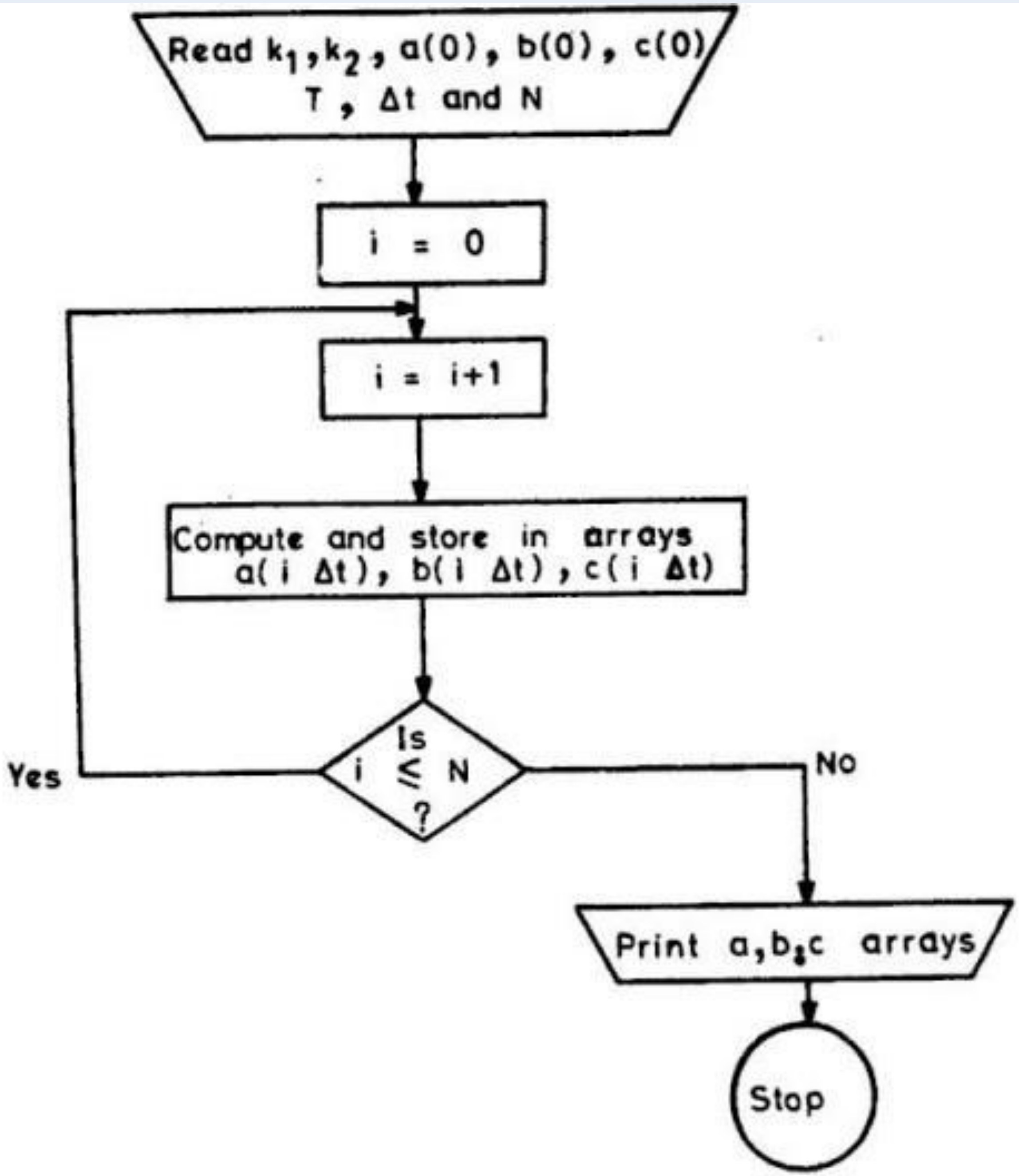
Using these values we calculate the next state of the system, i.e., at time 2*t* as

$$a(2\,\Delta t) = a(\Delta t) + [k_2 . c(\Delta t) - k_1 a(\Delta t) . b(\Delta t)] . \Delta t$$

$$b(2\,\Delta t) = b(\Delta t) + [k_2 . c(\Delta t) - k_1 a(\Delta t) . b(\Delta t)] . \Delta t$$

$$c(2\,\Delta t) = c(\Delta t) + [2k_1 . a(\Delta t) . b(\Delta t) - 2k_2 . c(\Delta t)] . \Delta t$$

- Using the steps we calculate 3Δt, 4Δt … NΔt → T


Read k_1, k_2, a(0), b(0), c(0)
T, Δt and N
i = 0
i = i+1
Compute and store in arrays
a(i Δt), b(i Δt), c(i Δt)
Is
i ≤ N
?
Yes
No
Print a, b, c arrays
Stop

Sample output

TIME	A(I)	B(I)	C(I)
0.00	100.00	50.00	0.00
0.10	96.00	46.00	8.00
0.20	92.47	42.47	15.06
0.30	89.33	39.33	21.34
0.40	86.52	36.52	26.95
0.50	84.00	34.00	32.00
0.60	81.72	31.72	36.55
0.70	79.66	29.66	40.69
0.80	77.77	27.77	44.45
0.90	76.05	26.05	47.89
1.00	74.48	24.48	51.04
1.10	73.03	23.03	53.94
1.20	71.70	21.70	56.61
1.30	70.46	20.46	59.07
1.40	69.32	19.32	61.36
1.50	68.26	18.26	63.48
1.60	67.28	17.28	65.44
1.70	66.36	16.36	67.28
1.80	65.51	15.51	68.99
1.90	64.71	14.71	70.59
2.00	63.96	13.96	72.08

(. . . Continued)

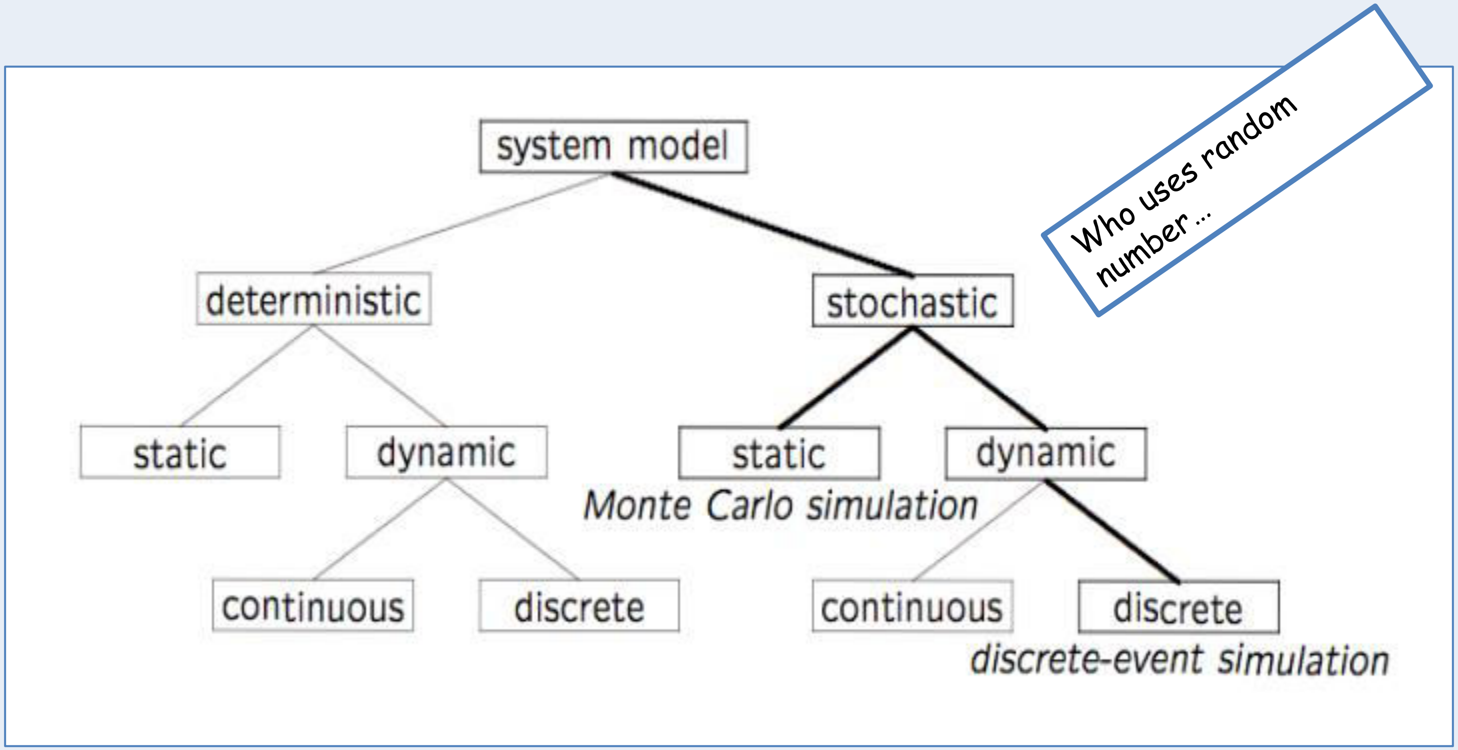
TIME	A(I)	B(I)	C(I)
2.10	63.26	13.26	73.48
2.20	62.60	12.60	74.79
2.30	61.99	11.99	76.03
2.40	61.41	11.41	77.18
2.50	60.86	10.86	78.27
2.60	60.35	10.35	79.30
2.70	59.87	9.87	80.27
2.80	59.41	9.41	81.18
2.90	58.98	8.98	82.04
3.00	58.57	8.57	82.86
3.10	58.19	8.19	83.63
3.20	57.82	7.82	84.36
3.30	57.48	7.48	85.05
3.40	57.15	7.15	85.70
3.50	56.84	6.84	86.32
3.60	56.55	6.55	86.91
3.70	56.27	6.27	87.46
3.80	56.00	6.00	87.99
3.90	55.75	5.75	88.50
4.00	55.51	5.51	88.97
4.10	55.29	5.29	89.43
4.20	55.07	5.07	89.86
4.30	54.86	4.86	90.27
4.40	54.67	4.67	90.66
4.50	54.48	4.48	91.03
4.60	54.31	4.31	91.39
4.70	54.14	4.14	91.73
4.80	53.98	3.98	92.05
4.90	53.82	3.82	92.35
5.00	53.68	3.68	92.65

Class-4

Simulation and Modeling

CSE4131_CSE4132

System Model Classification



Random Numbers

- What is Random Number?
  - A random number is a number generated by a process,
    - whose outcome is **unpredictable**,
    - and which **cannot be** sub sequentially reliably reproduced.
  - This definition … one has some kind of a **black box** – such a black box is usually called a **random number generator** – that fulfills this task.
  - Random number plays important role in simulation tasks.

Properties of Random Numbers

- A sequence of random numbers, R_1, R_2 … must have two important statistical properties:

1. Uniformity → ?
   - They are equally probable every where
2. Independence → ?
   - The current value of a random variable has no relation with the previous value.

Generation of Pseudo Random Numbers

- Pseudo → False / Fake
- Generation of False or Fake Random Numbers?
  - **Pseudo →** because generating numbers using a known method removes the potential for **true** randomness
  - **Goal →** To produce a sequence of numbers in [0,1] that simulates, or imitates, the ideal properties of random numbers (RN)

Generation of Pseudo Random Numbers

- Important considerations in RN routines:
  - The routine should be fast
  - Portable to different computers… (different programming languages)
  - Have sufficiently *long cycle*
  - *Replicable*
  - Closely approximate the ideal statistical properties of uniformity and independence

Generation of Pseudo Random Numbers

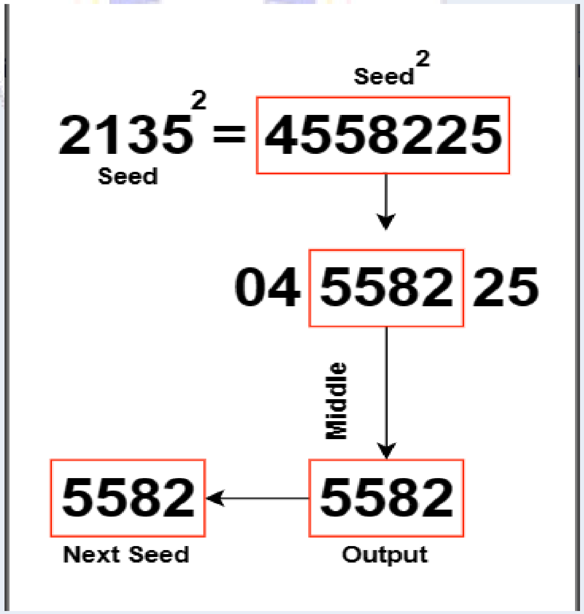
- When generating pseudo random numbers certain problem may occur:
  - Generated number may ***not be uniformly distributed***
  - Generated number may be ***discrete valued, instead of continuous valued***
  - The **mean** of generated numbers can be **too high or too low**
  - The **variance** can be may **too high or too low**.
  - There may be **cyclic variation**: i.e.
    - Autocorrelation between numbers
    - Numbers successively higher or lower than successive numbers
    - Several numbers above the mean followed by several numbers below the mean

Techniques of Generating Random Numbers

- Different methods have been proposed:
  - MidSquare method
  - Linear Congruential Method (LCM).
  - Combined Linear Congruential Generators (CLCG).
  - Random-Number Streams.

MidSquare method

- First arithmetic generator: Midsquare method
  - von Neumann and Metropolis in 1940s
- The Midsquare method:
  - Start with a four-digit positive integer Z_0
  - Compute: $Z_0^2 = Z_0 \times Z_0$ to obtain an integer with up to eight digits
  - Take the middle four digits for the next four-digit number



Midsquare method

Problem: Generated numbers tend to 0

i	Z_i	U_i	$Z_i \times Z_i$
0	7182	-	51581124
1	5811	0,5811	33767721
2	7677	0,7677	58936329
3	9363	0,9363	87665769
4	6657	0,6657	44315649
5	3156	0,3156	09960336
6	9603	0,9603	92217609
7	2176	0,2176	04734976
8	7349	0,7349	54007801
9	78	0,0078	00006084
10	60	0,006	00003600
11	36	0,0036	00001296
12	12	0,0012	00000144
13	1	0,0001	00000001
14	0	0	00000000
15	0	0	00000000

Midsquare method

- A good algorithm is basically the one which does not depend on the seed and the period should also be maximally long .
- In practice it is a highly faulty method for many practical purposes, since its period is usually very short.
- It has some severe weaknesses; repeated enough times, the middle-square method will either begin repeatedly generating the same number or cycle to a previous number in the sequence and loop indefinitely.

Linear Congruential Method (LCM)

- Congruential → Congruent → Similar/ fitting /matching
- Proposed by Lehmer (1951)
- To produce a sequence of integers, X_1, X_2, …between 0 and m-1 by following a recursive relationship

$$X_{i+1} = (\underbrace{a}_{\text{The multiplier}} X_i + \underbrace{c}_{\text{The increment}}) \bmod \underbrace{m}_{\text{The modulus}}, \quad i = 0,1,2,\ldots$$

The initial value of X_0 is called **seed.**
a → **constant multiplier.**
c → **increment**
m → **modulus**

Linear Congruential Method (LCM)

$$X_{i+1} = (\underbrace{a}_{\text{The multiplier}} X_i + \underbrace{c}_{\text{The increment}}) \bmod \underbrace{m}_{\text{The modulus}}, \quad i = 0,1,2,\ldots$$

- If c ≠ 0 then → **Mixed Congruential method**
- If c = 0 then → **Multiplicative Congruential method**

- The random integers are being generated [0,m-1], and to convert the integers to random numbers:

$$R_i = \frac{X_i}{m}, \quad i = 1,2,\ldots$$

$$X_{i+1} = (aX_i + c) \bmod m, \quad i = 0,1,2,\ldots$$

The multiplier

The increment

The modulus

- Given,
  - Seed (X_0)= 27 ,
  - Constant multiplier (a) = 17 ,
  - Increment (c) = 43 , and
  - modulus (m) = 100.

The sequence of X_i and subsequent R_i values is computed as follows:

$$X_0 = 27$$

$$X_1 = (17 \cdot 27 + 43) \bmod 100 = 502 \bmod 100 = 2$$

$$R_1 = \frac{2}{100} = 0.02$$

$$X_2 = (17 \cdot 2 + 43) \bmod 100 = 77 \bmod 100 = 77$$

$$R_2 = \frac{77}{100} = 0.77$$

$$X_3 = (17 \cdot 77 + 43) \bmod 100 = 1352 \bmod 100 = 52$$

$$R_3 = \frac{52}{100} = 0.52$$

$$\vdots$$

Linear Congruential Method:

Characteristics of a good Generator

- Maximum Density
  - The values assumed by R_i, $i=1,2,\ldots$ leave no large gaps on [0,1]
  - Problem: Instead of continuous, each R_i is discrete
  - Solution: a very large integer for modulus m
    - Approximation appears to be of little consequence

- Maximum Period
  - To achieve maximum density and avoid cycling
  - Achieved by proper choice of a, c, m, and X_0

- Most digital computers use a binary representation of numbers
  - Speed and efficiency are aided by a modulus, m, to be (or close to) a power of 2.

Linear Congruential Method:

Characteristics of a good Generator

- The LCG has full period if and only if the following three conditions hold (Hull and Dobell, 1962):
  1. The only positive integer that (exactly) divides both m and c is 1
  2. If q is a prime number that divides m, then q divides a-1
  3. If 4 divides m, then 4 divides a-1

Linear Congruential Method:
Proper choice of parameters

- For m a power 2, $m=2^b$, and $c \neq 0$
  - Longest possible period $P=m=2^b$ is achieved
    if c is relative prime to m and $a=1+4k$, where k is an integer

- For m a power 2, $m=2^b$, and $c=0$
  - Longest possible period $P=m/4=2^{b-2}$ is achieved
    if the seed X_0 is odd and $a=3+8k$ or $a=5+8k$, for $k=0,1,...$

- For m a prime and $c=0$
  - Longest possible period $P=m-1$ is achieved
    if the multiplier a has property that smallest integer k such that a^k-1 is divisible by m is $k = \mathrm{m}-1$

Linear Congruential Method: Example

- Use $a = 13, c = 0$, and $m = 64$
- The period of the generator is very low
- Seed X_0 influences the sequence

i	X_i X_0=1	X_i X_0=2	X_i X_0=3	X_i X_0=4
0	1	2	3	4
1	13	26	39	52
2	41	18	59	36
3	21	42	63	20
4	17	34	51	4
5	29	58	23	
6	57	50	43	
7	37	10	47	
8	33	2	35	
9	45		7	
10	9		27	
11	53		31	
12	49		19	
13	61		55	
14	25		11	
15	5		15	
16	1		3	

Class-5

Simulation and Modeling

CSE4131_CSE4132

Tests For Random Numbers

- The desired properties of Random Numbers:
  - Uniformity & Independence
- To ensure … desired properties… **number of tests** can be performed

Tests For Random Numbers

- Five types of tests:
  1. Frequency test — Uniformity test
  2. Runs Test
  3. Autocorrelation test
  4. Gap test
  5. Poker test

Independence test

Frequency Test

- Used to test uniformity
- Uses:
  - Kolmogorov-Smirnov test or
  - Chi-square test

Ks-test

We will study from N.Deo's Book 68

Frequency Test

- There should be **no favored number.**
  - That is, no number should occur more frequently than what is expected from chance variation.
  - For example,
    - If we have 5000 3-digit numbers (000-999)
    - We should expect-
      - 000-099 → 500 numbers
      - 100-199 → 500 numbers
      - 200-299 → 500 numbers
      - … … …
      - 900-999 → 500 numbers

Frequency Test

- We do not expect … exact 500 … in each of the 10 ranges
- In fact, if we found… we should suspect **nonrandomness.**
- … deviation from 500 should not be too much
  - Suspect → **nonuniformity**
- **How much deviation should be accepted?**
  - 500±x ; x = ?
  - Chi-square test

Chi-square test

- Steps
  1. Dividing the observed data into k non overlapping classes.
     - Observed data → generated random number
     - $K \geq 3$ (must be)

If we have 5000 3-digit numbers (000-999) → 10 non overlapping classes:

i	Range
1	000 - 099
2	100 - 199
3	200 - 299
4	300 - 399
5	400 - 499
6	500 - 599
7	600 - 699
8	700 - 799
9	800 - 899
10	900 - 999

Chi-square test

2. Then we count O_i the number of times observed data fall into each class i
   - For i=1,2,….k

i	Range	O_i (No. of observed occourrence)
1	000 - 099	468
2	100 - 199	519
3	200 - 299	480
4	300 - 399	495
5	400 - 499	508
6	500 - 599	426
7	600 - 699	497
8	700 - 799	515
9	800 - 899	463
10	900 - 999	529

Chi-square test

3. We determine the expected number occurrences E_i in each class i
4. Then we calculate chi-square value:

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

i	Range	O_i (No. of observed occourrence)	E_i (No of Expected occurrences)
1	000 - 099	468	500
2	100 - 199	519	500
3	200 - 299	480	500
4	300 - 399	495	500
5	400 - 499	508	500
6	500 - 599	426	500
7	600 - 699	497	500
8	700 - 799	515	500
9	800 - 899	463	500
10	900 - 999	529	500

Example

$(468-500)^2=(-32)^2=1024$

$1024/500= 2.048$

i	Range	O_i (No. of observed occourrence)	E_i (No of Expected occurrences)	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
1	000 - 099	468	500	1024	2.048
2	100 - 199	519	500	361	0.722
3	200 - 299	480	500	400	0.8
4	300 - 399	495	500	25	0.05
5	400 - 499	508	500	64	0.128
6	500 - 599	426	500	5476	10.952
7	600 - 699	497	500	9	0.018
8	700 - 799	515	500	225	0.45
9	800 - 899	463	500	1369	2.738
10	900 - 999	529	500	841	1.682
				Total = chi² =	**19.588**

Chi-square test

- **The question now arises**:
  - how **large or small** a computed value of chi-square can we accept.
- New term:
  - chi-square table…. Predefined
- New Parameter:
  - ***degree of freedom***
  - Defined by: $v=k-1$, $k \rightarrow$ no of sets into which data is divided
  - In our example $k=10$, that is $v=10-1=9$

Chi-square table

	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588
11	2.603	3.816	14.631	17.275	19.675	21.920	22.618	24.725	26.757	29.354	31.264
12	3.074	4.404	15.812	18.549	21.026	23.337	24.054	26.217	28.300	30.957	32.909
17	5.697	7.564	21.615	24.769	27.587	30.191	30.995	33.409	35.718	38.648	40.790
18	6.265	8.231	22.760	25.989	28.869	31.526	32.346	34.805	37.156	40.136	42.312
19	6.844	8.907	23.900	27.204	30.144	32.852	33.687	36.191	38.582	41.610	43.820
20	7.434	9.591	25.038	28.412	31.410	34.170	35.020	37.566	39.997	43.072	45.315
21	8.034	10.283	26.171	29.615	32.671	35.479	36.343	38.932	41.401	44.522	46.797
22	8.643	10.982	27.301	30.813	33.924	36.781	37.659	40.289	42.796	45.962	48.268
23	9.260	11.689	28.429	32.007	35.172	38.076	38.968	41.638	44.181	47.391	49.728
24	9.886	12.401	29.553	33.196	36.415	39.364	40.270	42.980	45.559	48.812	51.179

8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588

Chi-square table

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588

99.5% probability chi^2 will exceeding → 1.735
97.5% probability chi^2 will exceeding → 2.700
0.5% probability chi^2 will exceeding → 23.589

Thus the probability of the chi^2 statistic (for v=9) falling below 1.73 and above 23.6 is only 1%

That is 990 out of 1000 sequences from a perfectly uniform random number generator would have given:

$$1.73 \leq chi^2 \leq 23.6$$

Tests For Random Numbers

- Five types of tests:
    1. Frequency test — Uniformity test — Just Discussed
    2. **Runs Test**
    3. Autocorrelation test
    4. Gap test
    5. Poker test

Independence test

Runs Test

- Usually 3 types of tests…
    1. Runs up and down
    2. Runs above and below the mean.
    3. Runs test: length of runs.

Class-6

Simulation and Modeling

CSE4131_CSE4132

Tests For Random Numbers

- Five types of tests:
  1. Frequency test
  2. Runs Test
  3. Autocorrelation test
  4. Gap test
  5. Poker test

Uniformity test

Independence test

Frequency Test

0.08	0.09	0.23	0.29	0.42	0.55	0.58	0.72	0.89	0.91
0.11	0.16	0.18	0.31	0.41	0.53	0.71	0.73	0.74	0.84
0.02	0.09	0.3	0.32	0.45	0.47	0.69	0.74	0.91	0.95
0.12	0.13	0.29	0.36	0.38	0.54	0.68	0.86	0.88	0.91

0.08	0.09	0.23	0.29	0.42	0.55	0.58	0.72	0.89	0.91
8	9	23	29	42	55	58	72	89	91
0.11	0.16	0.18	0.31	0.41	0.53	0.71	0.73	0.74	0.84
11	16	18	31	41	53	71	73	74	84
0.02	0.09	0.3	0.32	0.45	0.47	0.69	0.74	0.91	0.95
2	9	30	32	45	47	69	74	91	95
0.12	0.13	0.29	0.36	0.38	0.54	0.68	0.86	0.88	0.91
12	13	29	36	38	54	68	86	88	91

8	9	23	29	42	55	58	72	89	91
11	16	18	31	41	53	71	73	74	84
2	9	30	32	45	47	69	74	91	95
12	13	29	36	38	54	68	86	88	91

2	9	30	32	45	47	69	74	91	95
8	9	23	29	42	55	58	72	89	91
11	16	18	31	41	53	71	73	74	84
12	13	29	36	38	54	68	86	88	91

		Oi	Ei	
1	0-19	9	8	0.125
2	20-39	8	8	0
3	40-59	8	8	0
4	60-79	7	8	0.125
5	80-99	8	8	0
		Chi-square		0.25

$0.207 \le chi^2 \le 14.86$

Runs Test

- Usually 3 types of tests…
  1. *Runs up and down*.
  2. *Runs above and below the mean*.
  3. Runs test: length of runs.

Runs Test

- Suppose a random number generator generates the following numbers:

0.08	0.09	0.23	0.29	0.42	0.55	0.58	0.72	0.89	0.91
0.11	0.16	0.18	0.31	0.41	0.53	0.71	0.73	0.74	0.84
0.02	0.09	0.30	0.32	0.45	0.47	0.69	0.74	0.91	0.95
0.12	0.13	0.29	0.36	0.38	0.54	0.68	0.86	0.88	0.91

Frequency test will say answer → generated numbers are uniformly distributed

But…

Runs Tests

- If we re-arrange the numbers…then its ok

0.41	0.68	0.89	0.84	0.74	0.91	0.55	0.71	0.36	0.30
0.09	0.72	0.86	0.08	0.54	0.02	0.11	0.29	0.16	0.18
0.88	0.91	0.95	0.69	0.09	0.38	0.23	0.32	0.91	0.53
0.31	0.42	0.73	0.12	0.74	0.45	0.13	0.47	0.58	0.29

The runs test examines the arrangement of numbers in a sequence to test the hypothesis of independence

Runs Test

The test of significance devised to test the randomness of a sample is the "RUN TEST", which is based on the theory of "RUNS"

What is RUN ?

A run is a sequence of symbols or occurrences of one kind surrounded (followed or proceeded) by a sequence of symbols or occurrences of other kind or by none at all.

Runs Test

The number of elements in a run is usually referred as length (l) of the run.

Examples of Run:

1) A data scientist carrying out research interviewed 10 persons during a survey. We denote the genders of the people by M for men and F for women. Assuming the respondents were chosen as follows:

Scenario 1

M M M M M F F F F F

Scenario 1 has only 2 runs

Scenario 2

F M F M F M F M F M

Scenario 2 has 10 runs

Scenario 3

F F F M M F M M F F

Scenario 3 has 5 runs

Runs Test

Examples of Run:

2) Suppose the applicants for the post of Professor in a particular University arrive as follows:

M F F MMM F F M F F MMM F F F M M

Here: M – Male; F – Female

There are in all nine runs in this sequence.

Runs Test

When Considered RUN and when NOT RUN?

The sequence of sample observations cannot be considered to be random:

1) If the similar item cluster together, resulting in too few runs.

   Example: Consider an experiment of tossing a coin 10 times and the results are

   T T T T T H H H H H

Here, the proportion (number) of heads and tails is same, the sequence CANNOT be considered as random BECAUSE there are only two runs, i.e., too few runs (2).

Runs Test

The sequence of sample observations cannot be considered to be random:

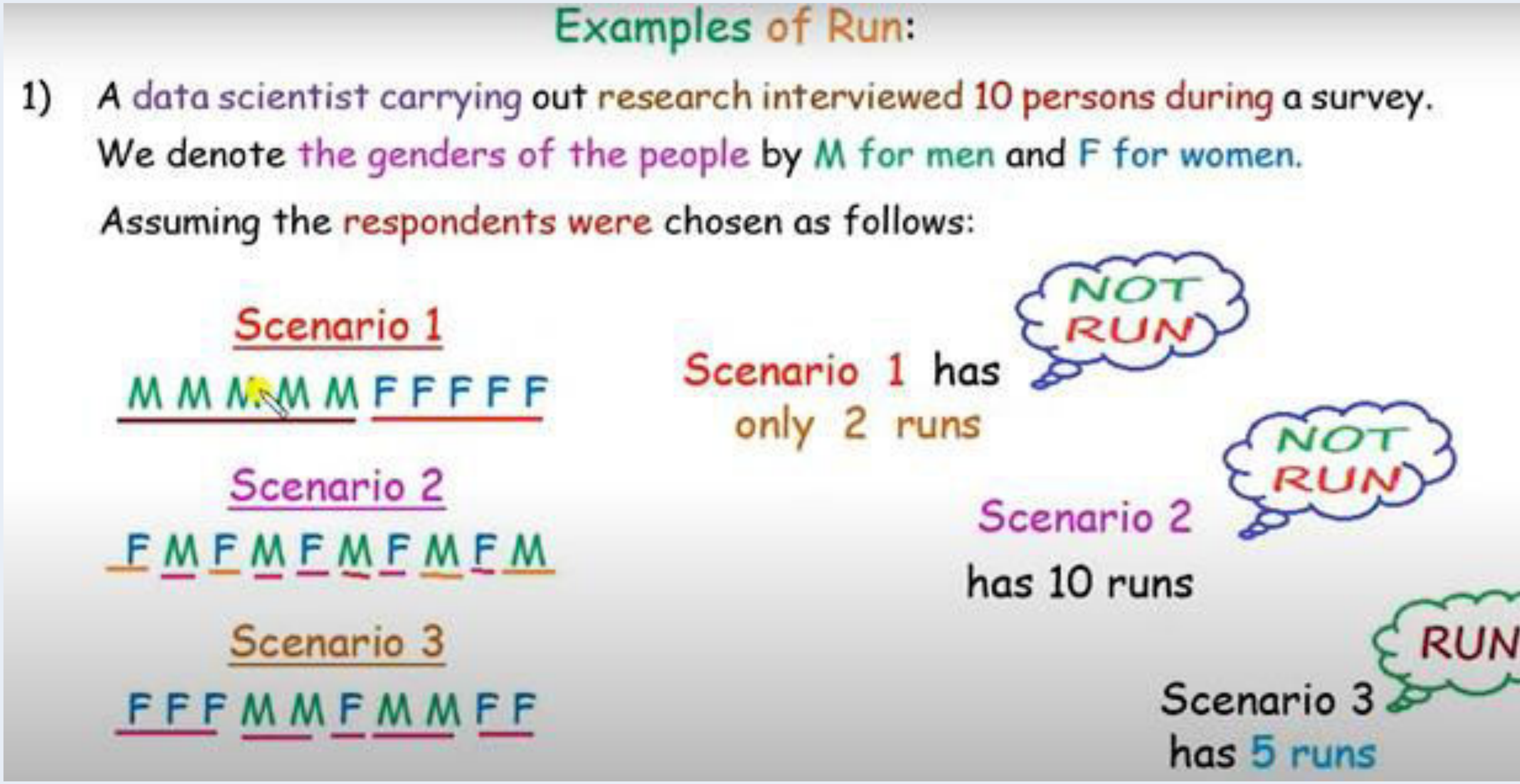
2) If similar items would alternatively mix, resulting in too many runs.

Example: Consider the following outcome of tossing of a coin 10 times.

T H T H T H T H T H

Here, also proportion of head and tail is same, the sequence CANNOT be considered as random

Runs Test


Examples of Run:
1) A data scientist carrying out research interviewed 10 persons during a survey.
We denote the genders of the people by M for men and F for women.
Assuming the respondents were chosen as follows:
Scenario 1
M M M M M F F F F F
Scenario 2
F M F M F M F M F M
Scenario 3
F F F M M F M M F F
NOT RUN
Scenario 1 has only 2 runs
NOT RUN
Scenario 2 has 10 runs
RUN
Scenario 3 has 5 runs

Runs Test

Notation:

n_1: number of symbols of first kind

n_2: number of symbols of second kind

$n = n_1 + n_2$ Total sample size

G: number of "runs" in the sequence

Example: Suppose the applicants for the post of Professor in a particular University arrive as follows:

M F F MMM F F M F F MMM F F F M M

Here: M – Male; F – Female

n_1: number of males = 10

n_2: number of females = 9

G = number of runs = 9

Runs Test

Remark:

It is Not necessary that the sample data are given in the form of symbols.

However, it may be given in the form of the number.

For example,

Marks obtained by the students in a class are

45, 65, 35, 76, 11, 09, 55, 87, 91, 25

Then How to find

$n_1, n_2,$ and G?

Runs Test

For such cases, we firstly convert numbers into symbolic form by using the median value and hence into runs.

Suppose $x_1, x_2, \dots, x_n$ is the set of n observations (in numerical form)

We obtain the median value (say Md) of the given set of samples and Replace the number

Ignore when value = Md

with symbol, say "A" if the observation is > Md

&

With symbol, say "B" if the observation is <Md

Runs Test

For example, Marks obtained by the students in a class are

45, 65, 35, 76, 11, 09, 55, 87, 91, 25

The median of this series is 50

Ascending order:
09, 11, 25,35, 45,
55, 65, 76, 87, 91

Replace the number with "A" if marks > 50

Replace the number with "B" if marks < 50

The sequence is

B A B A B B A A A B

n_1 = number of A runs = 5

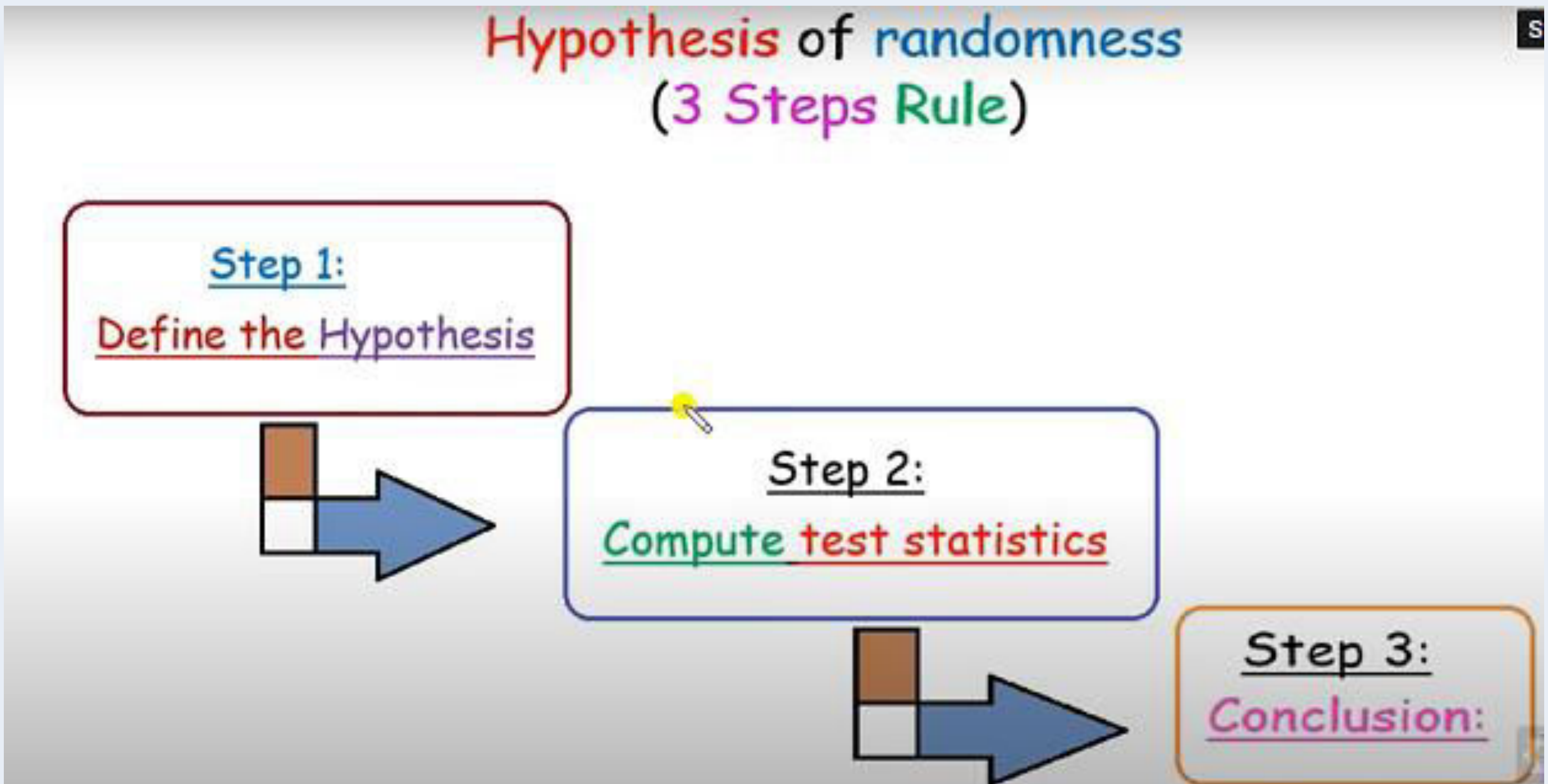
n_2 = number of B runs = 5

G = number of runs = 7

Runs Test

Single Sample RUN test

(Small Samples, $n_1 \leq 20, n_2 \leq 20$)



Runs Test

In a single sample run-test

Step 1: Define the Hypothesis

Null hypothesis — H_0: the process that generates sample data is random

Alternative hypothesis — H_1: the sample data is NOT random (two-tailed)

Note: Since "too few" or "too many" runs indicate lack of randomness of the sample, a two-tailed alternative hypothesis is appropriate.

find

n_1, n_2, G

Step 2:

Compute test statistics

Runs Test

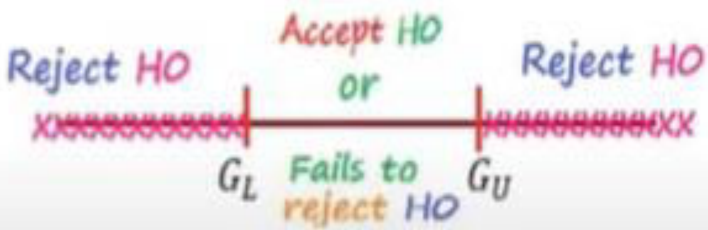
Step 3: Conclusion:

At the values of n_1, n_2 and α level of significance.

From RUN table

Find G_L = Lower critical value of G

G_U = upper critical value of G



Runs Test

TABLE G

Critical values of r in the runs test*

Given in the tables are various critical values of r for values of m and n less than or equal to 20. For the one-sample runs test, any observed value of r which is less than or equal to the smaller value, or is greater than or equal to the larger value in a pair is significant at the $\alpha = .05$ level.

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –
3					2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	2 –	3 –	3 –	3 –	3 –	3 –	3 –
4				2 9	2 9	2 –	3 –	3 –	3 –	3 –	3 –	3 –	3 –	3 –	4 –	4 –	4 –	4 –	4 –
5			2 9	2 10	3 10	3 11	3 11	3 –	3 –	4 –	4 –	4 –	4 –	4 –	4 –	4 –	5 –	5 –	5 –
6		2 –	2 9	3 10	3 11	3 12	3 12	4 13	4 13	4 13	4 13	5 –	5 –	5 –	5 –	5 –	5 –	6 –	6 –
7		2 –	2 –	3 11	3 12	3 13	4 13	4 14	5 14	5 14	5 14	5 15	5 15	6 15	6 –	6 –	6 –	6 –	6 –
8		2 –	3 –	3 11	3 12	4 13	4 14	5 14	5 15	5 15	6 16	6 16	6 16	6 16	6 17	7 17	7 17	7 17	7 17

Runs Test

Step 3: Conclusion:

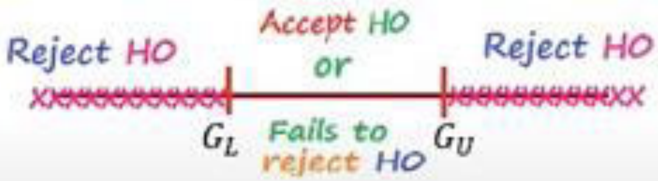
From RUN table

At the values of n_1, n_2 and α level of significance.

Find $G_L =$ Lower critical value of G

$G_U =$ upper critical value of G

reject the H_0,
if either $G \leq G_L$
or $G \geq G_U$



Fails to reject H_0,
if $G_L \leq G \leq G_U$

Runs Test

Example: A coin is tossed 20 times and the following sequence of heads (H) and tails (T) is obtained.

H T T H H T T H T H H H

T T H H T H H H

Use Run test to determine at 5% level of significance if the coin is unbiased.

Solution:

Step 1: Define the Hypothesis

Null Hypothesis H_0: The sequence of heads and tails is random, i.e., coin is unbiased

Alternative Hypothesis H_1: coin is biased

Runs Test

Step 2: Compute test statistics

For sequence

H TT HH TT H T HHH
TT HH T HHH

n_1 = number of heads = 12

n_2 = number of tails = 8

G = number of runs = 11

Thus, the test statistics is

$G = 11$

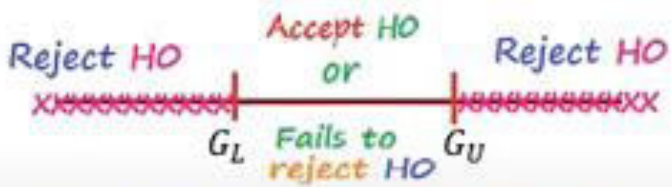
Step 3: Conclusion

Since $n_1 < 20, n_2 < 20$,
so it is small sample.

The critical value for $n_1 = 12, n_2 = 8, \alpha = 0.05$ is

G_L = Lower critical value of G = 6

G_U = upper critical value of G = 16



Since $6 < 11 < 16$, i.e., $G_L < G < G_U$,

Hence, we fail to reject H_0 and
conclude that the coin
may be regarded as unbiased.

Runs Test

Example: On a commuter train, the conductor wants to see whether the passengers entering a train enter in a random manner. He observes the first 25 people, with the following sequence of males (M) and females (F).

F F F M M F F F F M F M M M F F F F M M F F F M M

Test for randomness at 5% level of significance.

Solution:

Step 1: Define the Hypothesis

Null Hypothesis H_0: The pattern of occurrence of M and F enter the train is random.

Alternative Hypothesis H_1: The pattern of occurrence of M and F is not random

Runs Test

Step 2: Compute test statistic

The sequence is

FFF MM FFFF M F MMM

FFFF MM FFF MM

n_1: number of females = 15

n_2: number of males = 10

$G =$ number of runs = 10

Thus, the test statistics is

$G = 10$

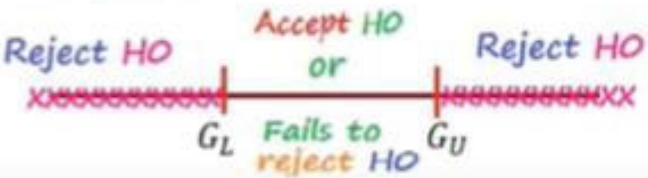
Step 3: Conclusion

Since $n_1 < 20, n_2 < 20$, so, it is small sample.

The critical value for $n_1 = 15, n_2 = 10$, $\alpha = 0.05$ is

$G_L =$ Lower critical value of $G = 7$

$G_U =$ upper critical value of $G = 18$



Since $7 < 10 < 18$, we may accept the H_0, i.e., we fail to reject H_0.

Hence, there are not enough evidence to reject the claim

Runs Test

Example: We have 20 people that enrolled in a drug abuse program. Test the claim that the ages of the people, according to the order in which they enrol occur at random, at $\alpha = 0.05$. The data are as follows:

18, 36, 19, 22, 25, 44, 23, 27, 27, 35, 19, 43,

37, 32, 28, 43, 46, 19, 20, 22

Solution:

Step 1: Define the Hypothesis

Null Hypothesis (H_0): The pattern of occurrence of ages of the people enrolled in a drug abuse program is determined by a random process

Alternative Hypothesis (H_1): The pattern is not random

Runs Test

Step 2: Compute Test statistic

To find the number of runs we first find the median.

18, 36, 19, 22, 25, 44, 23, 27, 27, 35, 19, 43,

37, 32, 28, 43, 46, 19, 20, 22

To it, arrange the data in ascending order and get Median of the data is 27

Replace the number with "A" if observation is > 27	Replace the number with "B" if observation is < 27

The sequence is

B, A, B, B, B, A, B, , , A, B, A,

A, A, A, A, A, B, B, B

Runs Test

Step 2: Compute Test statistic

The sequence is

B, A, B, B, B, A, B, A, , , B, A,

A, A, A, A, A, B, B, B

n_1 = number of A runs = 9

n_2 = number of B runs = 9

G = number of runs = 9

Thus, the test statistics is

G = 9

Step 3: Conclusion

Since $n_1 < 20, n_2 < 20$ so, it is small sample.

The critical value for $n_1 = 9, n_2 = 9, \alpha = 0.05$ is

G_L = Lower critical value of G = 5

G_U = upper critical value of G = 15

Since $5 < 9 < 15$, we may accept the H_0,
i.e., we fail to reject H_0.

Runs Test

Large Sample:

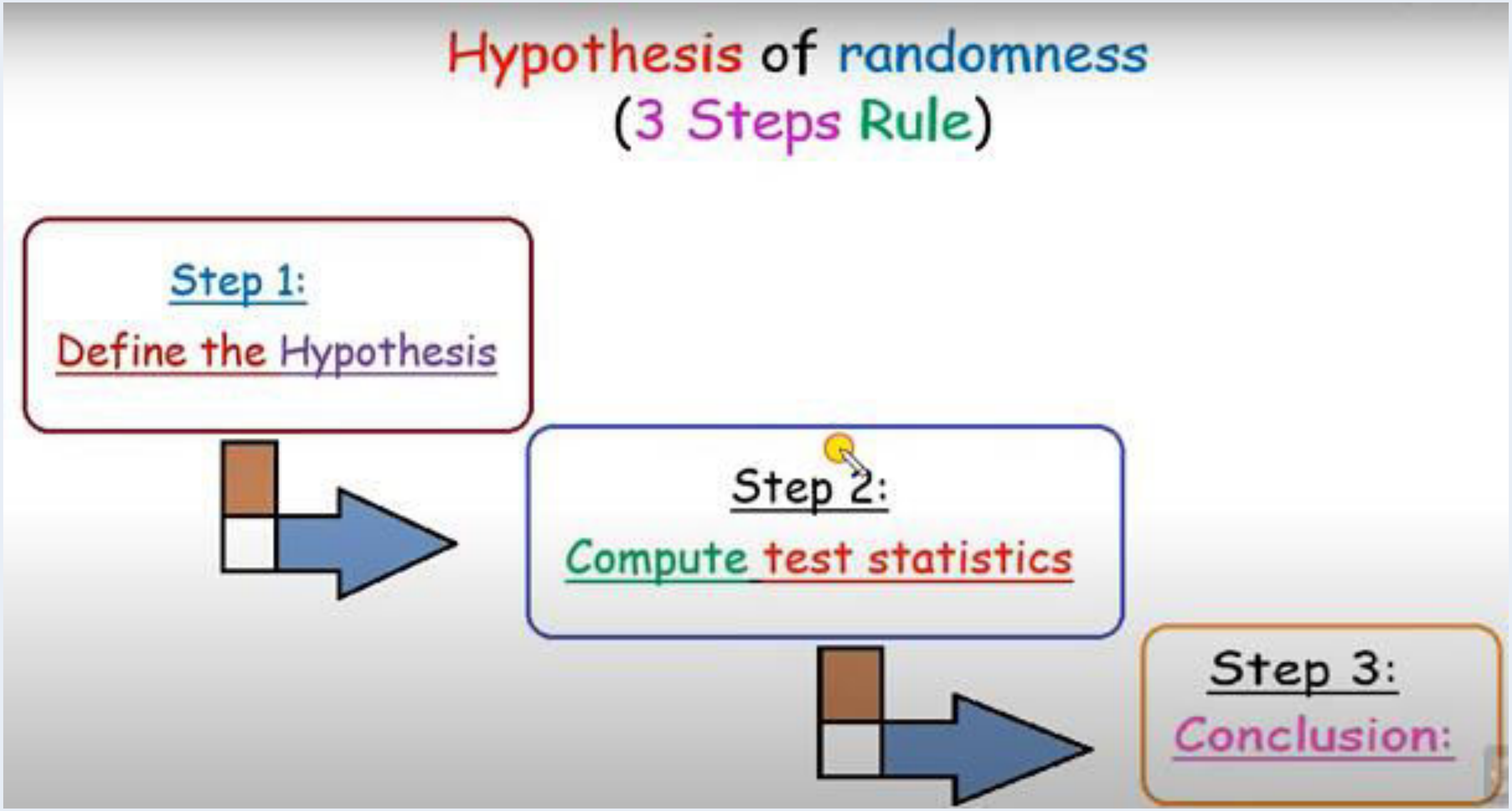
When $n_1 > 20$ or $n_2 > 20$, the sampling distribution of the number of runs (G) is approximately normal distributed with mean μ_G and standard error σ_G given by

$$\boxed{\mu_G = \frac{2n_1 n_2}{n} + 1}$$

$$\boxed{\sigma_G = \sqrt{\frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{n^2(n-1)}}}$$

where $n = n_1 + n_2$

Runs Test



Hypothesis of randomness
(3 Steps Rule)
Step 1:
Define the Hypothesis
Step 2:
Compute test statistics
Step 3:
Conclusion:

Runs Test

Step 2: Compute Test statistic

For large sample and under H_0,

the test statistics is

$$Z = \frac{G - \mu_G}{\sigma_G} \sim N(0,1)$$

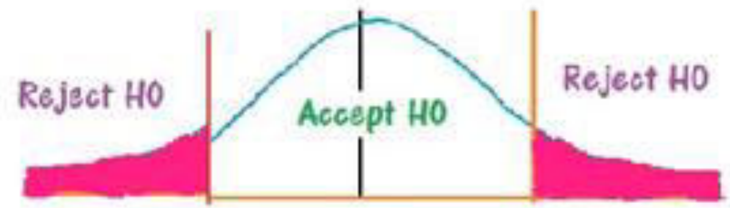
$$\mu_G = \frac{2n_1 n_2}{n} + 1$$

$$\sigma_G = \sqrt{\frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{n^2(n-1)}}$$

where $n = n_1 + n_2$

Step 3: Conclusion

The critical value of Z for two-tailed test at α (level of significance) can be taken from Z-table.



If computed $Z >$ critical value of Z at given level of significance, we REJECT H_0, otherwise we fail to reject H_0.

Settings

Runs Test

Example: A sample of 48 tools produced by a machine shows the following sequence of good (G) and defective (D) tools:

G G G G G D D G G G G G G G G G D D D D G G G G G G

D G G G G G G G G G D D G G G G G D G G

Test the randomness of the sequence at 5% significance level.

Solution:

Step 1: Define the Hypothesis

Null Hypothesis (H_0): The pattern is determined by a random process

Alternative Hypothesis (H_1): The pattern is not random

Runs Test

Step 2: Compute Test statistic

The sequence

G G G G G D D G G G G G G G G G D D D D

G G G G G G D G G G G G G G G G D D

G G G G G D G G

n_1 = number of G runs = 38

n_2 = number of D runs = 10

G = number of runs = 11

$n = n_1 + n_2$

$= 48$

Since $n_1 > 20$ so for large sample,

$$\mu_G = \frac{2n_1 n_2}{n} + 1$$

$$= \frac{2(38)(10)}{48} + 1$$

$$= 16.83$$

$$\sigma_G = \sqrt{\frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{n^2(n-1)}}$$

$$= \sqrt{\frac{2(38)(10)[2(38)(10) - 38 - 10]}{48^2(48-1)}}$$

$$= 2.235$$

Runs Test

Thus, the test statistics is

$$Z = \frac{G - \mu_G}{\sigma_G}$$

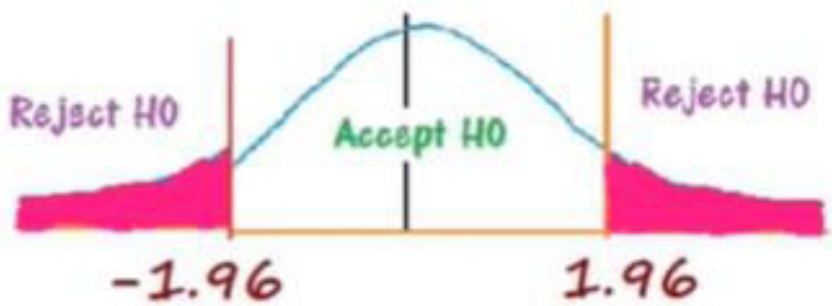
$$= \frac{11 - 16.83}{2.235}$$

$$= -2.61$$

Step 3: Conclusion

At 5% level of significance,

the critical value is 1.96 for two-tail.



Since -2.61 < -1.96, so we Reject H0.

Hence, we conclude that random process is not random.

Class-7

Simulation and Modeling

CSE4131_CSE4132

Tests For Random Numbers

- Five types of tests:
  1. Frequency test — Uniformity test
  2. **Runs Test**
  3. Autocorrelation test
  4. Gap test
  5. Poker test

  Tests 2–5: Independence test

1. Runs up and down
2. Runs above and below the mean.
3. Runs test: length of runs.

Runs above and below mean

- Lets consider the following generated number set:

0.63	0.72	0.79	0.81	0.52	0.94	0.83	0.93	0.87	0.67
0.54	0.83	0.89	0.55	0.88	0.77	0.74	0.95	0.82	0.86
0.43	0.32	0.36	0.18	0.08	0.19	0.18	0.27	0.36	0.34
0.31	0.45	0.49	0.43	0.46	0.35	0.25	0.39	0.47	0.41

The sequence of runs up and down is as follows:

+ + + − + − + − − − + + − + − − + − + − − + − −
+ − + + − − + + − + − − + + −

Runs above and below mean

0.63	0.72	0.79	0.81	0.52	0.94	0.83	0.93	0.87	0.67
0.54	0.83	0.89	0.55	0.88	0.77	0.74	0.95	0.82	0.86
0.43	0.32	0.36	0.18	0.08	0.19	0.18	0.27	0.36	0.34
0.31	0.45	0.49	0.43	0.46	0.35	0.25	0.39	0.47	0.41

- N.B. First 20 numbers are above the mean and last 20 numbers are below the mean -→ **Not Acceptable**

Runs above and below mean

0.40	0.84	0.75	0.18	0.13	0.92	0.57	0.77	0.30	0.71
0.42	0.05	0.78	0.74	0.68	0.03	0.18	0.51	0.10	0.37

- ‘**+**’ denotes an observation is **above** the mean
- ‘**-**’ denotes an observation is **below** the mean

The pluses and minuses are as follows:

− + + − − + + + − + − − + + + − − + − −

\- + + - - + + + - + - - + + + - - + - -

Runs above and below mean

- + + - - + + + - + - - + + + - - + - -

Run with length 2 and it’s a above the mean run

Run with length 1 and it’s a blow the mean run

- There are 11 runs
  - 5 of them above the mean
  - 6 of them below the mean

Runs above and below mean

\-   +   +   -   -   +   +   +   -   +   -   -   +   +   +   -   -   +   -   -

- Let, *b* = total number of runs (11)
- Again
  – Let n_1 be the no of **observations** … **above** the mean
  – And n_2 be the no of **observations** … **below** the mean

- Then maximum number of possible runs $N=n_1+n_2$
  – And minimum number of runs = one
- Given n_1 and n_2 then the mean and variance suggested by **Swed and Eisenhart [1943]….**

Runs above and below mean

- Given n_1 and n_2 then the mean and varience suggested by Swed and Eisenhart [1943]….

$$\mu_b = \frac{2n_1n_2}{N} + \frac{1}{2}$$

$$\sigma_b^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}$$

Either n_1 or n_2 > 20 … b will be approximated … normal distribution

Runs above and below mean

- For normal distribution Z-score calculation

Let Y be a random variable with expected value μ and standard deviation σ, the <u>Z score random variable</u> is defined as follows:

$$Z = \frac{Y - \mu}{\sigma}$$

$$\mu_b = \frac{2n_1n_2}{N} + \frac{1}{2}$$

$$\sigma_b^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}$$

$$Z_0 = \frac{b - \mu_b}{\sigma_b} = \frac{b - (2n_1n_2/N) - 1/2}{\left[\frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}\right]^{1/2}}$$

Runs above and below mean

Failure to reject the hypothesis of independence occurs when :

$$-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$$

α → level of significance

$\alpha = 0.05$ → $\alpha/2 = 0.025$

$z_{\alpha/2}$ → $z_{0.025} = 1.96$

Critical Values of the Standard Normal Distribution

α	z_α
0.1000	1.28
0.0500	1.64
0.0250	1.96
0.0100	2.33
0.0050	2.58
0.0025	2.81
0.0010	3.09

Example

− + + + + + + + − − − + + − + − − − − −

− − + + − − − − + + − − + − + − − + + −

The values of n_1, n_2, and b are as follows:

$$n_1 = 18$$

$$n_2 = 22$$

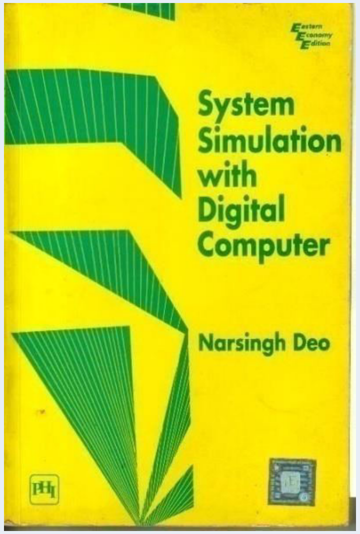
$$N = n_1 + n_2 = 40$$

$$b = 17$$

Class-8

Recommended Book

- System Simulation With Digital Computer
  - By Narshing Deo
  - Chapter 5:
    Simulation of a PERT network



Project Management

- **Introduction**
- **Network Planning**
- **Estimating Time**
- **CPM** (Critical Path Method)
- **PERT** (Program Evaluation and Review Technique)

Project Management

Project

- A project is an interrelated set of activities that has a definite starting and ending point and that results in a unique product or service.

Project management

- Project management is a scientific way of planning, implementing, monitoring & controlling the various aspects of a project such as time, money, materials, manpower & other resources.

Network Planning Methods

- Methods used for network planning are:
  - CPM
  - PERT

- Managing a project with network planning methods involves four steps:

1. Describing the Project.
2. Diagramming the Network.
3. Estimating time of completion.
4. Monitoring Project Progress.

Network Diagram

Concepts

- Activity
- Precedence relationship
- Successor
- Event

Network model of a project

- Projects → **number of separate activities**
  - For example:
    - Building construction → Selection of site + decision of architecture + laying of foundation + etc.
- Due to technical reasons:
  - Some activities can not be started before some others are complete

  →Thus there will be some **PRECEDENCE RELATIONSHIP**
- In addition, each activity require certain time → **DURATION**

Network model of a project

- **ACTIVITIES + PRECEDENCE + DURATION →** by means of an **arrow diagram**
  - Called: ***Activity Network*** or ***Activity Graph***
- **Activities** are shown as: **directed lines** or **edges**
  - Also called branches or arrows or arcs.
- And the **nodes** represents the ***beginning*** and ***completion*** of Activities → called: **Event** or **milestones** in the project

Guidelines for network diagram

1. Before an activity can begin, its preceding activities must be completed.
2. Arrows indicate logical precedence.
3. Flow of the diagram is from left to right.
4. Arrows should not intersect.
5. **Dangling** should be avoided.

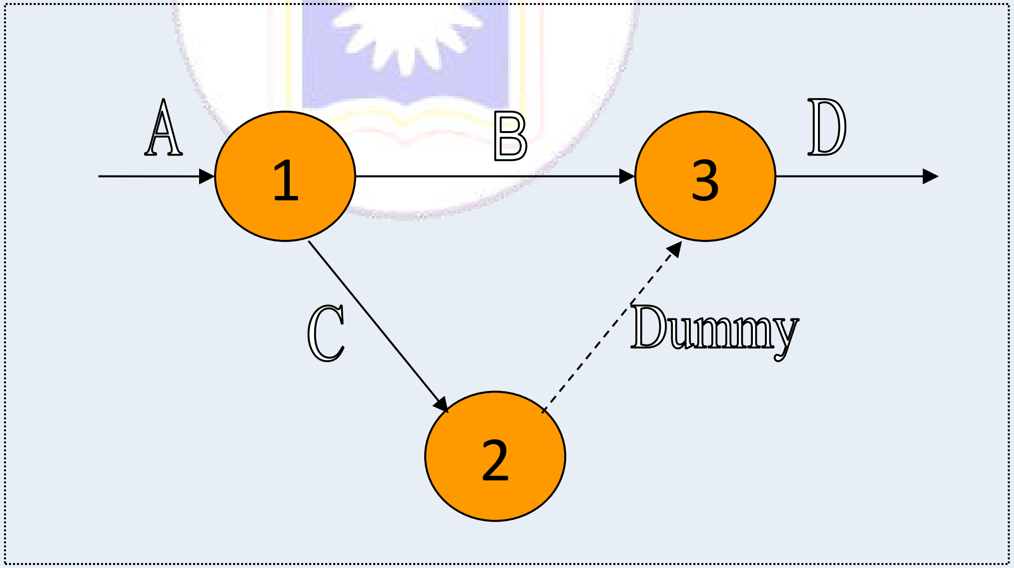
Dangling - the act of suspending something

Approaches for Network Diagram

ACTIVITY ON ARC (AOA) approach requires the addition of a ***Dummy Activity*** to clarify the precedence relationships between the two activities. It is a zero time activity and consumes no resources.

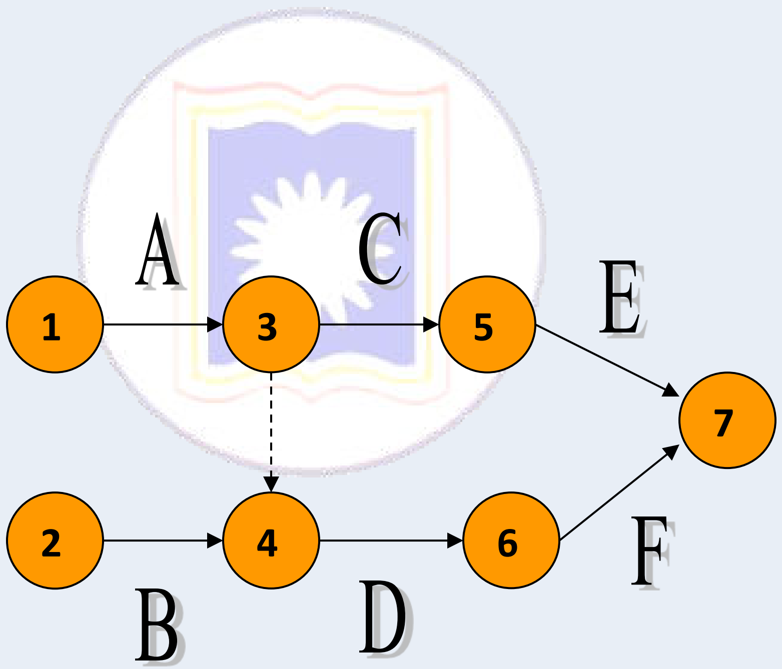
Dummy Activity is used in two situations:

1) When two or more activities start and end at the same nodes



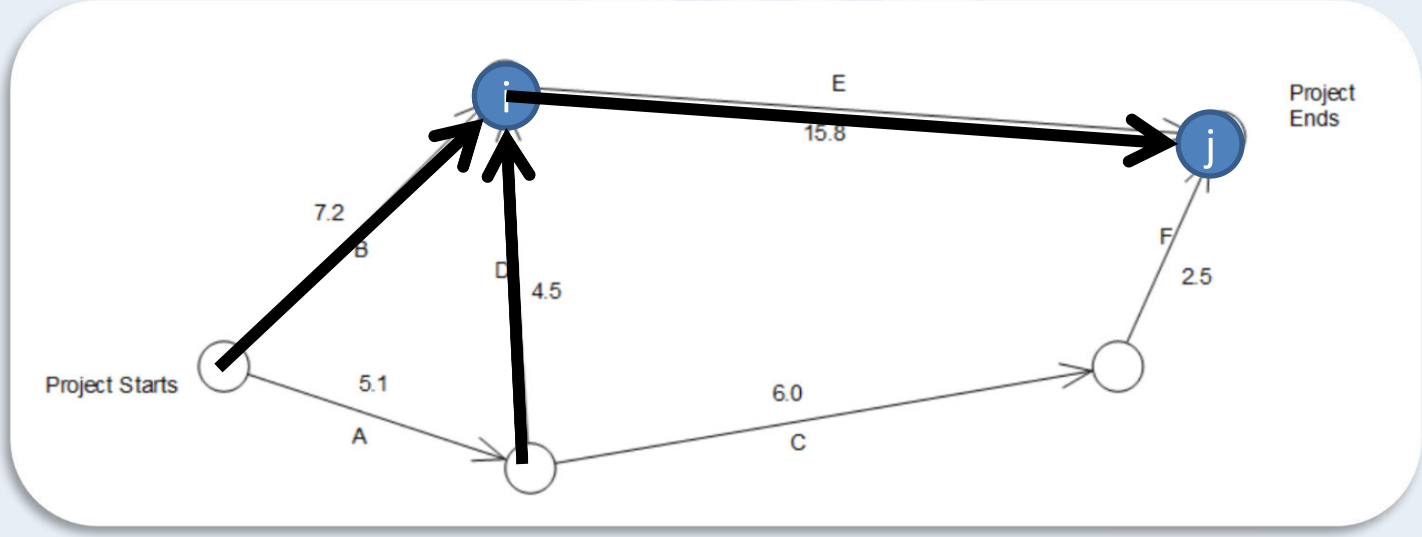
Approaches for Network Diagram

2) When two or more activities share the same precedence activity but not all the precedence are shared.

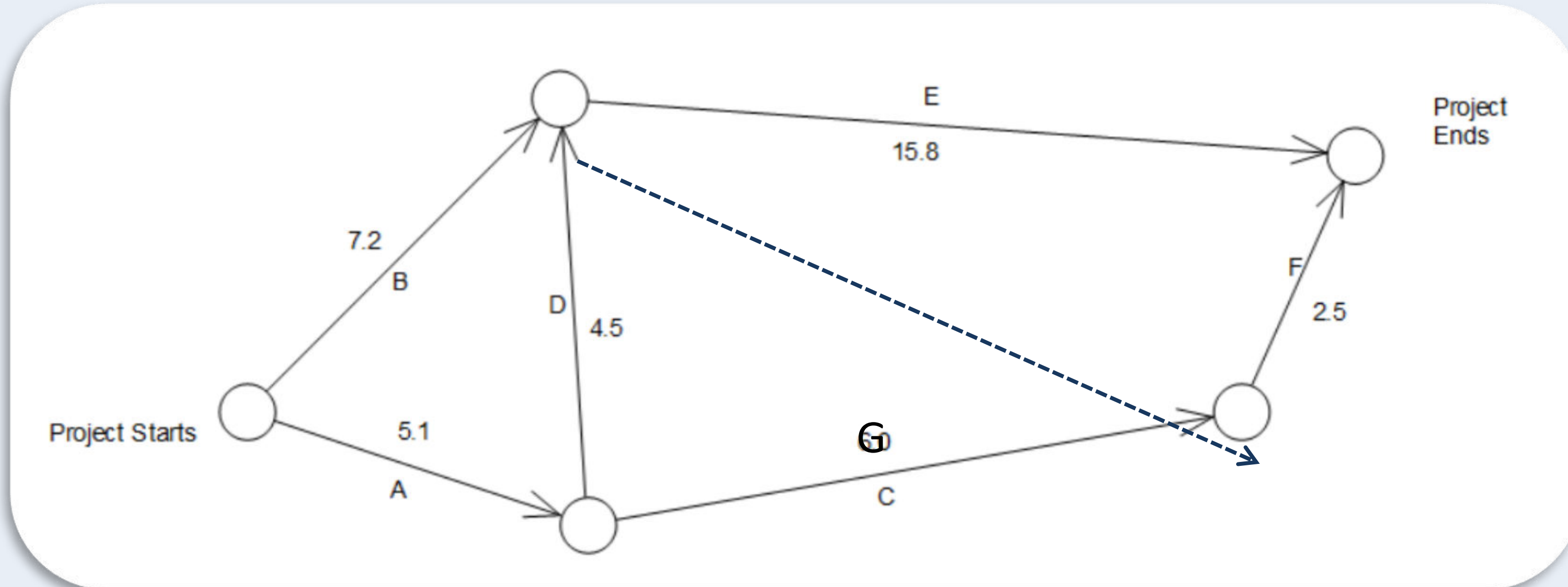


Network model of a project

- An activity represented by edge ***(i , j)***
  - can not be started before all activities leading to node ***i*** have been completed.
- Let,… we have a project … 6 well defined, non overlapping individual jobs. → A, B, C, D, E, F
- **Restriction:**
  - A must proceed C and D
  - B and D must proceed E
  - C must proceed F



Network model of a project



- **Dummy Activity Example:**
  - Suppose we had an additional restriction→ F can't be initiated before completion of B and D
  - Dummy activities are necessary → when existing activities are not sufficient enough to describe relationship
  - Duration = 0

Estimating Time of Completion

Planning the schedule of the project

Time estimates include:

1) Total time for completion.
2) **Earliest start time (ES):** the earliest time at which the activity can start given that its precedent activities must be completed first.
3) **Earliest finish time (EF):** equals to the earliest start time for the activity plus the time required to complete the activity.
4) **Latest finish time (LF):** the latest time in which the activity can be completed without delaying the project.
5) **Latest start time (LS):** equal to the latest finish time minus the time required to complete the activity.

Estimating Time Of Completion

6) **Forward Pass:**

The **early start** and **early finish** times are calculated by moving **forward** through the network and considering the predecessor activities. Considers maximum

7) **Backward Pass:**

The **latest start** and **finish times** are calculated by moving **backward** through the network. Considers minimum

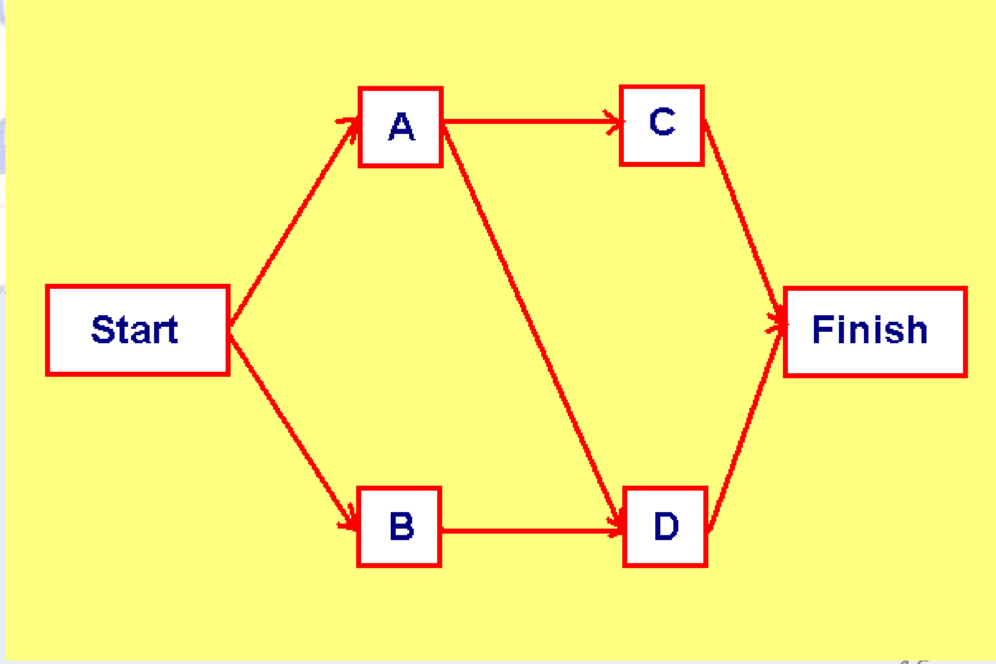
8) **Slack Time:**

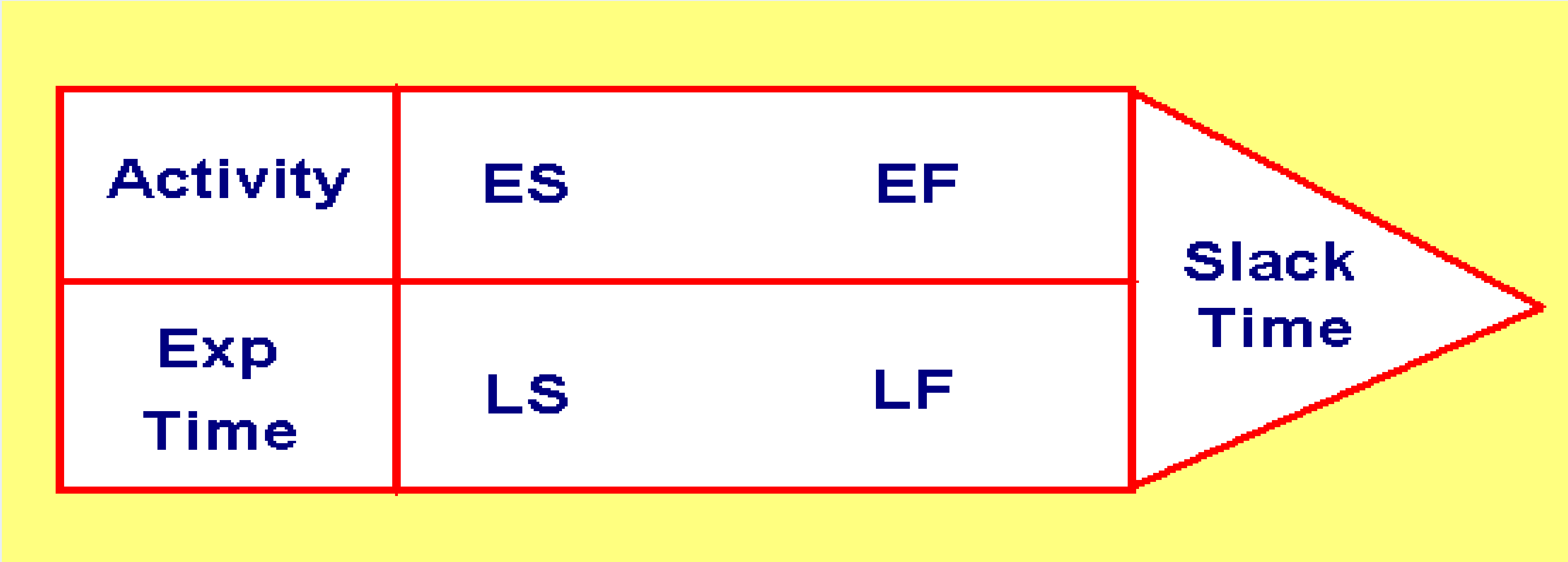
Slack time for an activity is the difference between its **earliest** and **latest start time** or between the **earliest** and **latest finish time**.

Critical path is the path of activities having zero Slack time.

A Simple Project

Activity	Immediate Predecessor	Expected Time
A	-	5
B	-	6
C	A	4
D	A, B	2

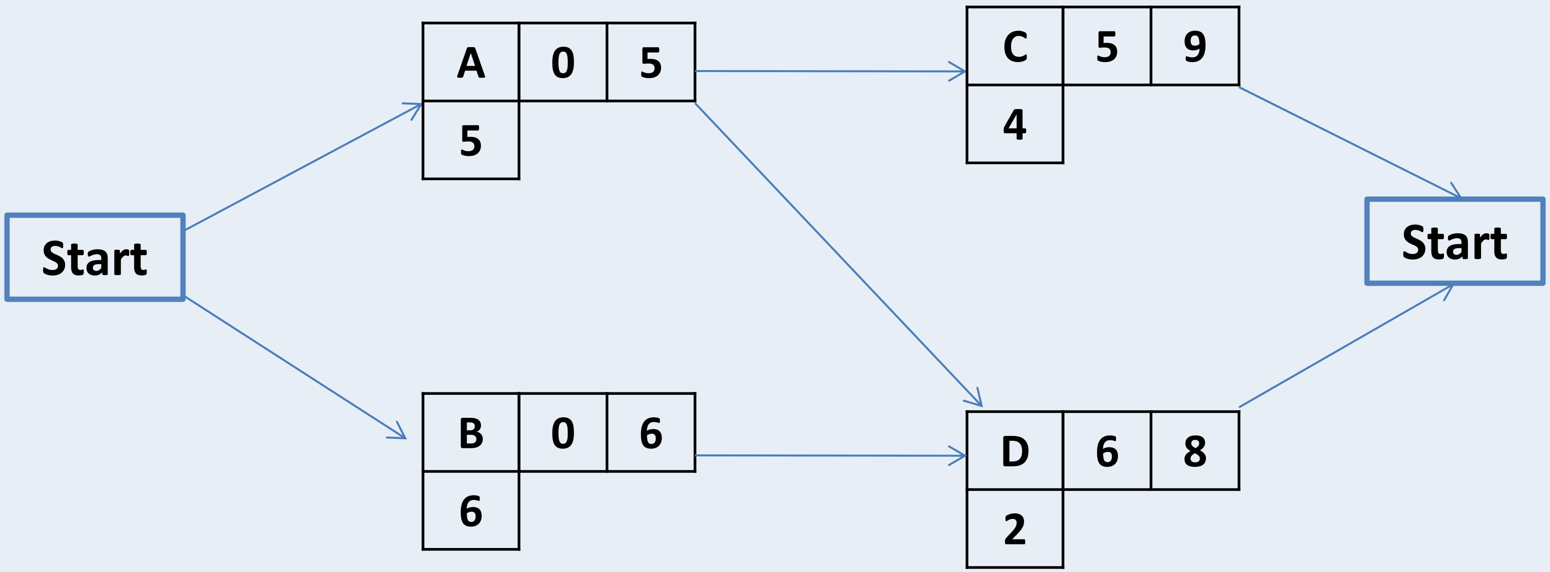




ES	Earliest Starting (time)
EF	Earliest Finishing
LS	Latest Starting
LF	Latest Finishing
Slack	Difference Time

Activity	Immediate Predecessor	Expected Time
A	-	5
B	-	6
C	A	4
D	A, B	2

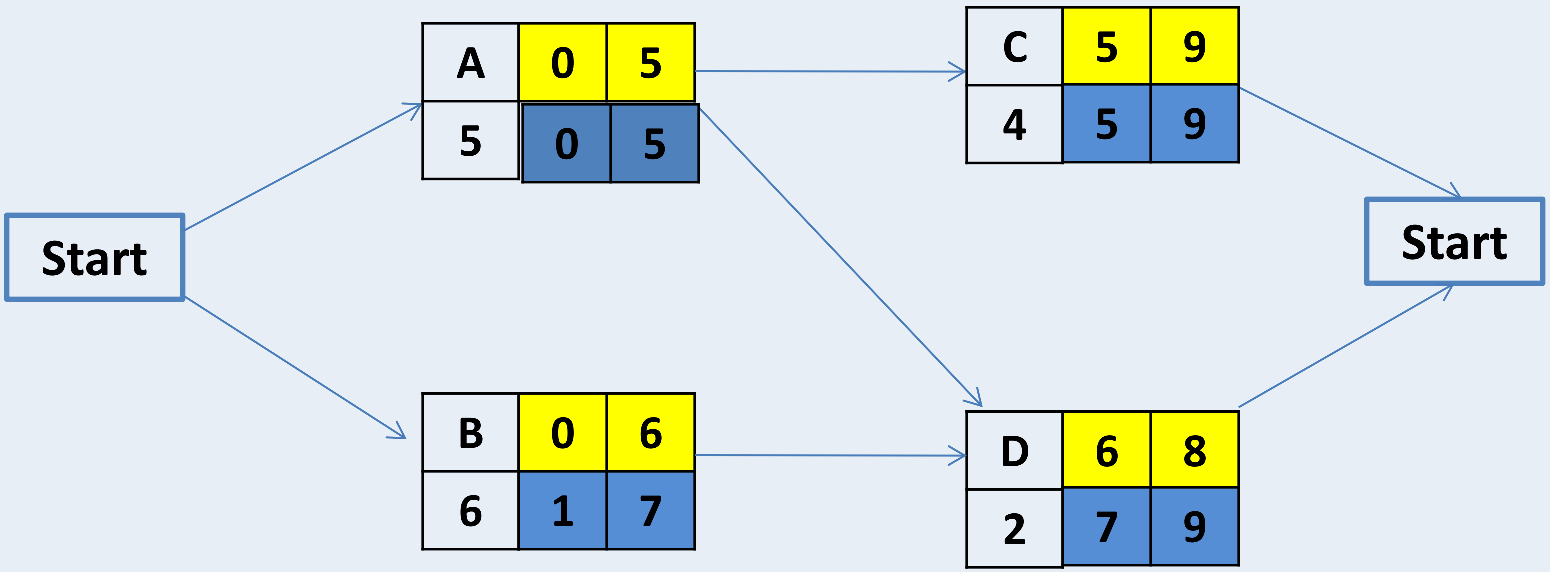
Forward Pass:



ES[D] = max{ EF[A], EF[B] }

Activity	Immediate Predecessor	Expected Time
A	-	5
B	-	6
C	A	4
D	A, B	2

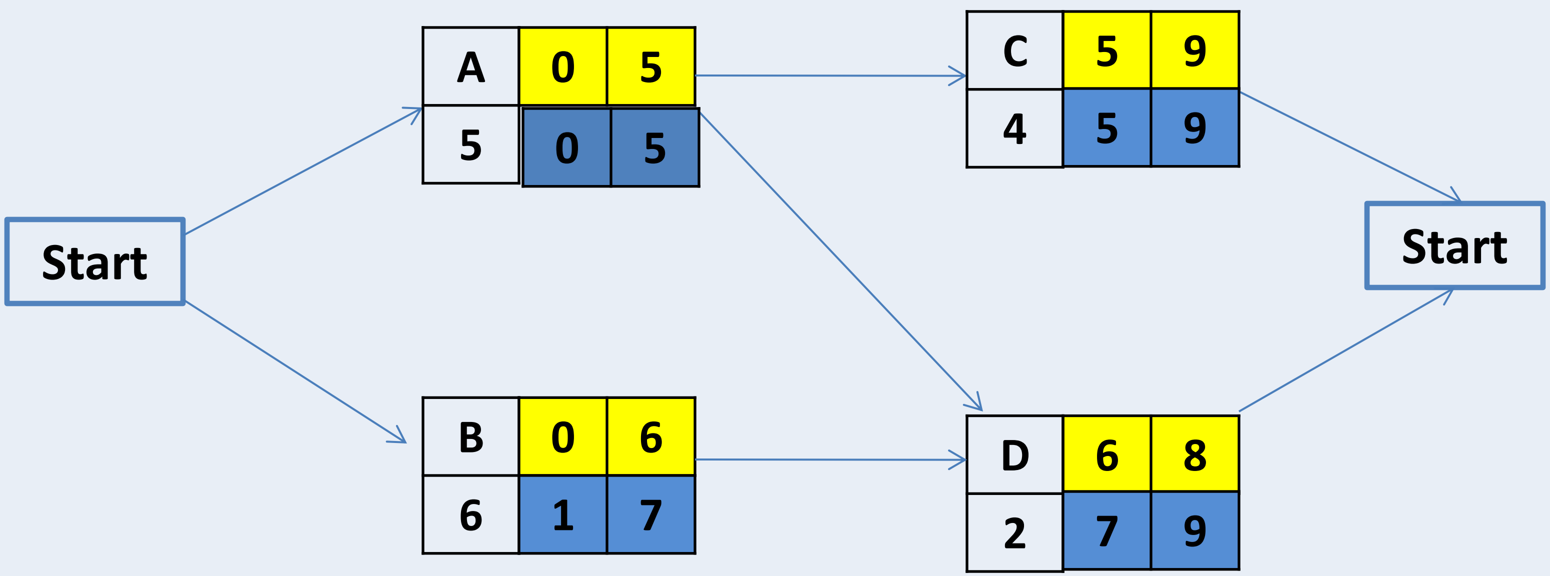
Backward Pass:



LF[A] = min{ LS[C], LS[D] }

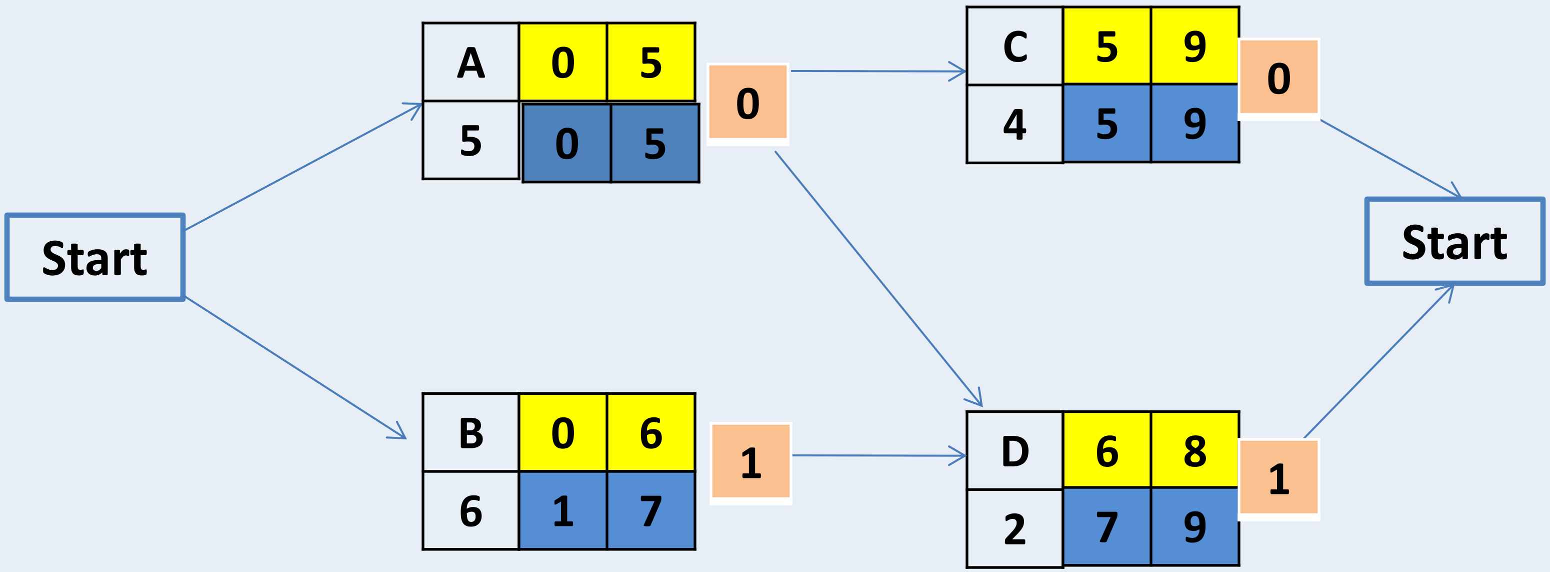
Activity	Immediate Predecessor	Expected Time
A	-	5
B	-	6
C	A	4
D	A, B	2

Backward Pass:



LF[A] = min{ LS[C], LS[D] }

Activity	Immediate Predecessor	Expected Time
A	-	5
B	-	6
C	A	4
D	A, B	2



Critical Part = {A, C}

Home Work

Activity	Predecessor	Duration (days)
A	-	3
B	A	4
C	A	2
D	B	5
E	C	1
F	C	2
G	D,E	4
H	F,G	3

Class-9

Simulation and Modeling

CSE4131_CSE4132

PERT (*Program Evaluation and Review Technique*)

The **PERT** is a network model, a statistical tool that takes into account uncertainty in completion times owing to the various project activities. It was created in the late 1950's for the Polaris Project corresponding to the U.S. Navy fleet with thousands of contractors.

PERT steps –

- Identify the particular activities and their respective milestones.
- Determine the best possible succession of the activities.
- Designing the network diagram
- Estimate the time needed for every activity
- Determine the Critical Path.
- Update the PERT diagram as the project advances.

PERT (*Program Evaluation and Review Technique*)

For every project activity, the PERT network model generally incorporates three types of time-estimates.

Optimistic time (t_o): It is the shortest time in which the activity can be completed.

Most likely time (t_m): It is the probable time required to perform the activity.

Pessimistic time (t_p): It is the longest estimated time required to perform an activity.

Expected time

$$t_e = \frac{t_o + 4t_m + t_p}{6}$$

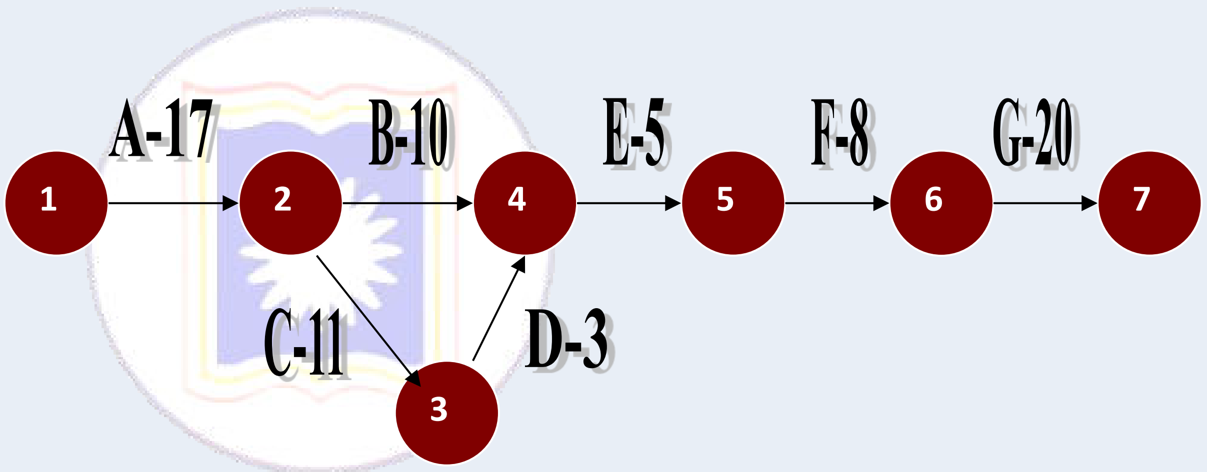
Activity	Description	Precedence	Optimistic time	Most Likely time	Pessimistic time	Expected time
A	Initial design	-	12	16	26	17
B	Survey market	A	6	9	18	10
C	Build prototype	A	8	10	18	11
D	Test prototype	C	2	3	4	3
E	Redesigning	B,D	3	4	11	5
F	Market testing	E	6	8	10	8
G	Set up production	F	15	20	25	20

Expected time (t_e) for A: (12+4*16+26)/6 = 17

Expected time (t_e) for B: (6+4*9+18)/6 = 10

Etc.

Activity	Precedence	Expected time
A	-	17
B	A	10
C	A	11
D	C	3
E	B, D	5
F	E	8
G	F	20



A-B-E-F-G = 60

A-C-D-E-F-G = 64 (CRITICAL PATH)

Example

An R & D project has a list of tasks to be performed whose time estimates are given in the Table , as follows.

Table 8.11: Time Estimates for R & D Project

Activity i j	Activity Name	T_0	t_m (in days)	t_p
1-2	A	4	6	8
1-3	B	2	3	10
1-4	C	6	8	16
2-4	D	1	2	3
3-4	E	6	7	8
3-5	F	6	7	14
4-6	G	3	5	7
4-7	H	4	11	12
5-7	I	2	4	6
6-7	J	2	9	10

a. Draw the project network.

b. Find the critical path.

c. Find the probability that the project is completed in 19 days. If the probability is less than 20%, find the probability of completing it in 24 days.

The variance of the activity time is calculated using the formula,

$$\sigma_i^2 = \left(\frac{t_p - t_0}{6} \right)^2$$

Activity	T_o	T_m	T_p	T_e	σ^2
1-2	4	6	8	6	0.444
1-3	2	3	10	4	1.777
1-4	6	8	16	9	2.777
2-4	1	2	3	2	0.111
3-4	6	7	8	7	0.111
3-5	6	7	14	8	1.777
4-6	3	5	7	5	0.444
4-7	4	11	12	10	1.777
5-7	2	4	6	4	0.444
6-7	2	9	10	8	1.777

Construct a network diagram:

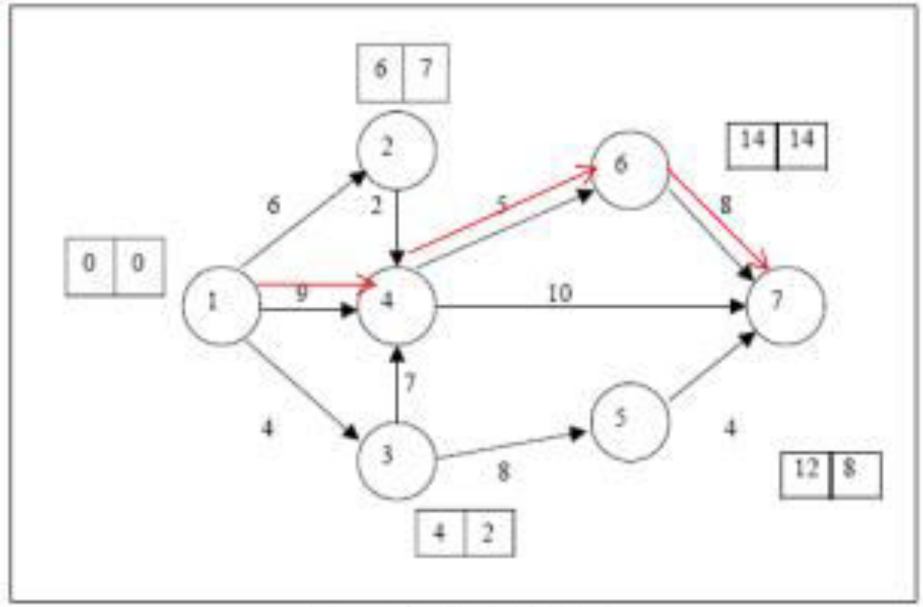


Figure 8.23: Network Diagram

calculate the time earliest (TE) and time Latest (TL) for all the activities.

From the network diagram Figure , the critical path is identified as 1-4, 4-6, 6-7, with a project duration of 22 days.

The probability of completing the project within 19 days is given by, P (Z< Zo)

To find Zo ,

$$Z_0 = \left(\frac{T_s - T_e}{\sqrt{\Sigma\sigma \text{ in critical path}}} \right)$$

$$= \left(\frac{19 - 22}{\sqrt{2.777 + 0.444 + 1.777}} \right) \qquad = \left(\frac{-3}{\sqrt{5}} \right) = -1.3416$$

we know, P (Z <Z Network Model o) = 0.5 – z (1.3416) (from normal tables, z (1.3416) = 0.4099)
= 0.5 – 0.4099
= 0.0901
= 9.01% Thus, the probability of completing the R & D project in 19 days is 9.01%.

Since the probability of completing the project in 19 days is less than 20% As in question, we find the probability of completing it in 24 days.

$$Z_0 = \frac{T_s - T_e}{\sqrt{\Sigma\sigma \text{ in critical path}}}$$

$$= \left(\frac{24 - 22}{\sqrt{5}} \right) = \left(\frac{2}{\sqrt{5}} \right) = 0.8944 \text{ days}$$

$$P(Z \le Z_0) = 0.5 - Y\ (0.8944) \qquad \text{(from normal tables, } Y\ (0.8944) = 0.3133)$$

$$= 0.5 + 0.3133$$

$$= 0.8133$$

$$= 81.33\%$$

Advantages of PERT

- Expected project completion time.
- Probability of completion before a specified date.
- The critical path activities that directly impact the completion time.
- The activities that have slack time and that can lend resources to critical path activities.
- Activity start and end dates.

LIMITATIONS

- The PERT Formula Requires Too Much Work.
- The network charts tend to be large and unwieldy.
- Calculating the time estimates is very complex for all the activities.
- Updating of the project is time consuming and requires high costs.
- Emphasis is laid only on time factors and cost factors are neglected.

Difference between CPM & PERT

CPM	PERT
• CPM works with fixed deterministic time	• PERT works with probabilistic time
• CPM is useful for repetitive and non complex projects with a certain degree of time estimates.	• PERT is useful for non repetitive and complex projects with uncertain time estimates.
• CPM includes time-cost trade off.	• PERT is restricted to time variable.
• CPM- for construction projects.	• PERT- used for R&D programs.

Class-10

Simulation and Modeling

CSE4131_CSE4132

Modeling

Model:

We define a model as **the body of information** about a system(?) → gathered for the purpose of studying the System.

System:

A group of interacting, interrelated, or interdependent elements forming a complex whole.

Why Modeling

- To study a system …
  - sometime … possible to experiment with system
  - Most time **impossible** to experiment with system itself.

- You can’t modify supply and demand … to study economy system
  - So we have to build a model

Points to remember

- Since the purpose of the study will determine the nature of information that is gathered.
  - there is **no unique model of a system.**

- Different models of same system … produced by different analysts different aspect of the system and
- Or by same analyst → as understanding the system changes.

Types of CNN Models.

LeNet

AlexNet

ResNet

GoogleNet/InceptionNet

MobileNetV1

ZfNet

Depth based CNNs

Highway Networks

Wide ResNet

VGG

PolyNet

Inception v2

Inception v3, V4, V5 and Inception-ResNet.

DenseNet

Pyramidal Net

Xception

Channel Boosted CNN using TL

Residual Attention NN

Attention Based CNNS

Feautre-Map based CNNS

Squeeze and Excitation Networks

Competitive Squeeze and Excitation Networks

Deriving a Model

- Task of deriving a model ➔ two subtasks
  1) Establishing the model structure
  2) Supplying the data

– Establishing the **model structure**:
  - Determines the **system boundary** +
  - Identifies the **entities**, **attributes** and **activities** of the system

– The data provides the **values the attributes** can have and define the relationships involved in the activities.
  – N.B.:… they are defined as parts of one task rather than two different task

System Boundaries

System boundaries are established to define what is inside and what is outside the system.

- They show other systems that are used or depend on the system being developed.

The position of the system boundary has a profound (insight) effect on the system requirements.

Defining a system boundary is a political judgment

- There may be pressures to develop system boundaries that increase/decrease the influence or workload of different parts of an organization.

Entities, Attributes and Activities

System Boundary:

The line that separate the system with its environment.

Entity:

Objects of interest within a system.

Attribute:

Denote a property of an entity.

Activity:

Any process that cause change within a system is called activity.

Some example

System	Entities	Attributes	Activities
Traffic	Cars	Speed Distance	Driving
Bank	Customers	Balance Credit Status	Depositing
Communications	Message	Length Priority	Transmitting
Supermarket	Customer	Shopping List	Checking-out

Types of Model

- Basically … model … two category
  - Physical
  - Mathematical

- Physical → Static or dynamic
- Mathematical → Static or dynamic

Physical Model

- Based on some analogy (**similarity**) between such system as mechanical and electrical or electrical and hydraulic (driven by fluid)
- System attributes are represent… voltage or position of a shaft
- Reflects the physical laws that drive the model.

Example

Bridges, cell models, human organs, crash test dummies, landform models, models of buildings, and models of chemical compounds are examples of physical models. These only qualify as physical models if they are actually built, and can be touched, as opposed to being build through computer software and only seen on a screen.

Mathematical model

- Mathematical model:
  - … some **symbolic notation and mathematical equations** to represent the system
  - System **attributes** are represented by **variables**
  - and **activities** ..by mathematical **functions**

Predicting the Future

Mathematical models can also be used to forecast future behavior.

Example: An ice cream company keeps track of how many ice creams get sold on different days.

By comparing this to the weather on each day they can make a mathematical model of **sales versus weather**.

They can then predict future sales based on the weather forecast, and decide how many ice creams they need to make ... ahead of time!

Another way to classify

- Second distinguish ….
  - **Static** and **dynamic** models

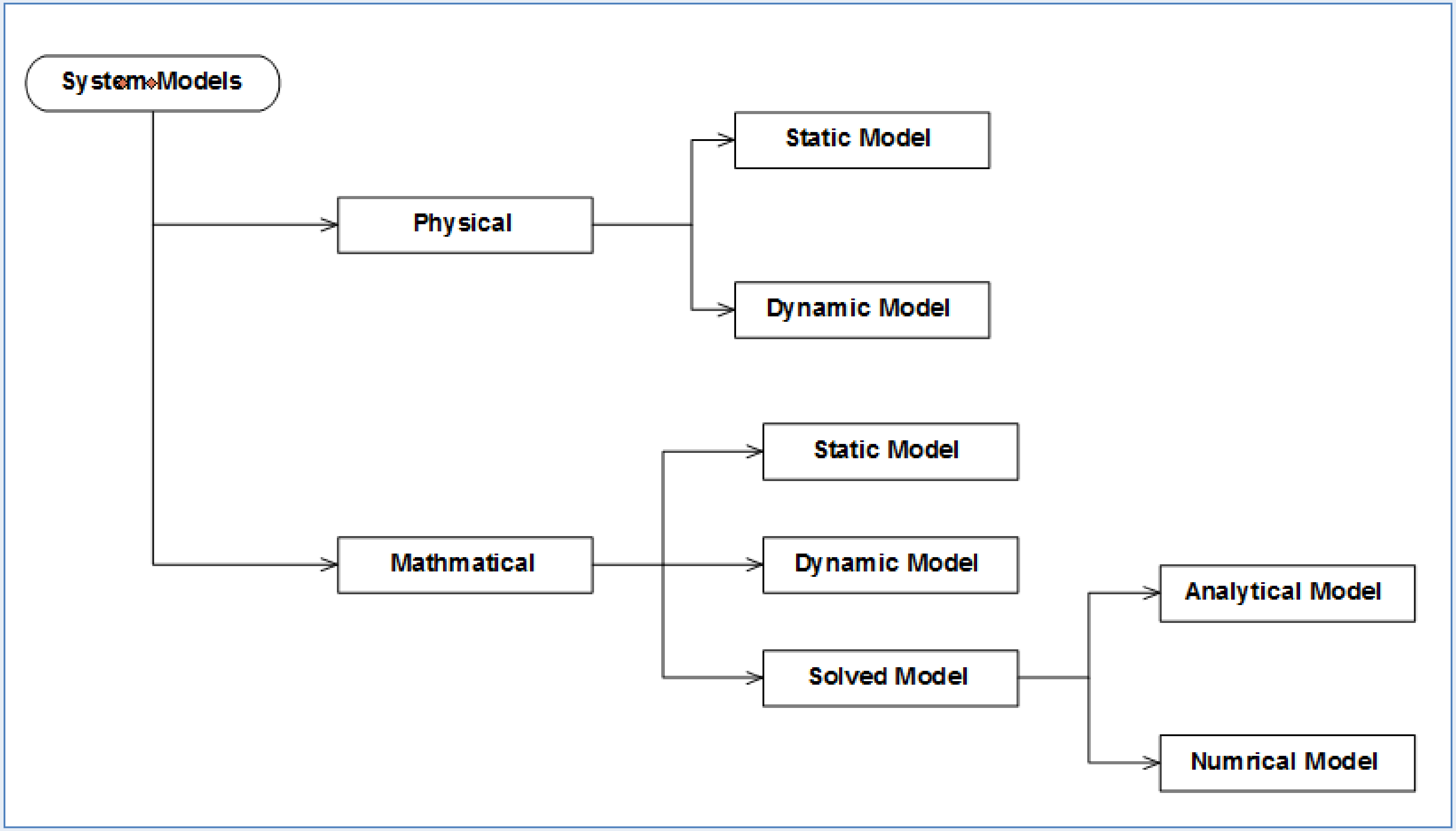
- **Static** models can only show the values that system attributes take when the system is in balance.
  - Values at some particular time
- **Dynamic** models follow the changes over time that result from the system activities

Mathematical Model

- In case of mathematical models… Solved Model
  - a third distinction is the **technique by which the model is solved**
  - That is actual values are assigned to the system attributes

- A distinction can be made:
  - Analytical and
  - Numerical methods
- **Analytical** … using the deductive reasoning of mathematical theory to solve a model.
- **Numerical methods**… involve applying computational procedure to solve equations

Summary of Classification



Principles used in modeling

- Its not possible to provide rules by which mathematical models are built.
  - But a number of principles can be stated

- Principles:
  - **Block** building… should be organized in series of blocks.
  - **Relevance** … relevant to the study objective.
  - **Accuracy** … of information gathered form the system should be considered
  - **Aggregation** … in the extent to which the number of individual entities can be grouped together into a large entity

Subsystems

- Simple definition of a system:
  - A set of interacting objects
  - But description of a system can be made many levels of details.
  - A system → consisting of interacting subsystem
  - A subsystem → consisting of subsystems… lower level of details
  - A system study must begin by deciding on the level of subsystem details to be used.

  - Remember the principle…. Block-building principle
  - Sub system → sub models or blocks

Block building

- The description of System → Series of blocks
- Few/ Preferably one input → block
- Results in few output variables
- System → interconnections between blocks.

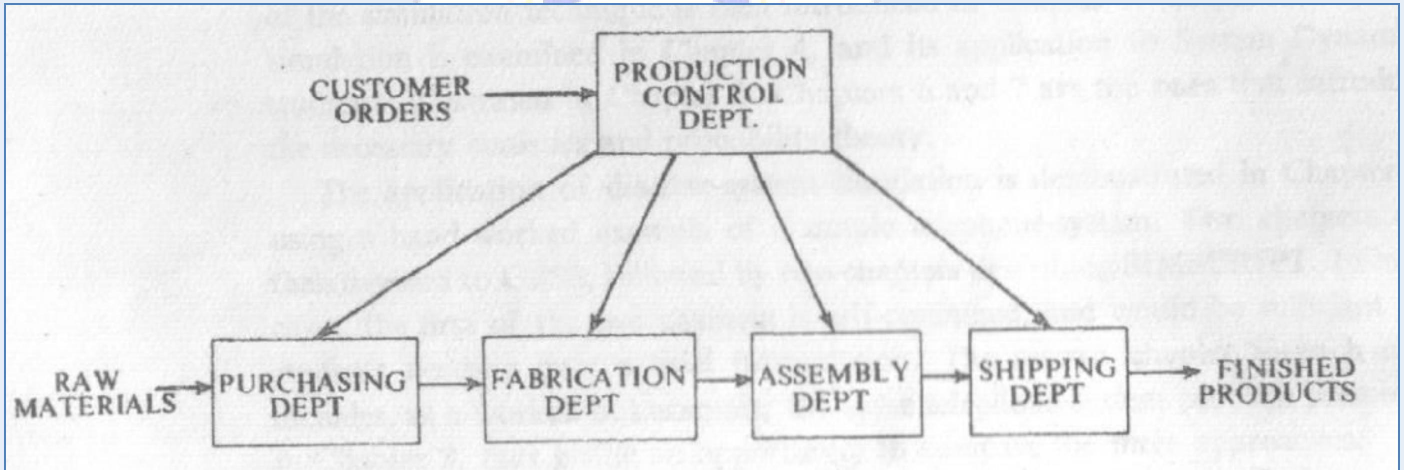


Figure 1-2. A factory system.

Relevance

- The model should only include → relevant to study objective
- For example:
  - If factory system study aim – to compare the effects of different operating rules on efficiency
  - Then its not relevance to consider … hiring of employees as activity

- If consider then → no harm to the system but make the system more complex.

Accuracy & Aggregation

- Accuracy →
  - The accuracy of information … for model … should be considered.
  - Wrong information or inaccurate information → may mislead …

- Aggregation →
  - … in the extent to which the number of individual entities can be grouped together into a large entity

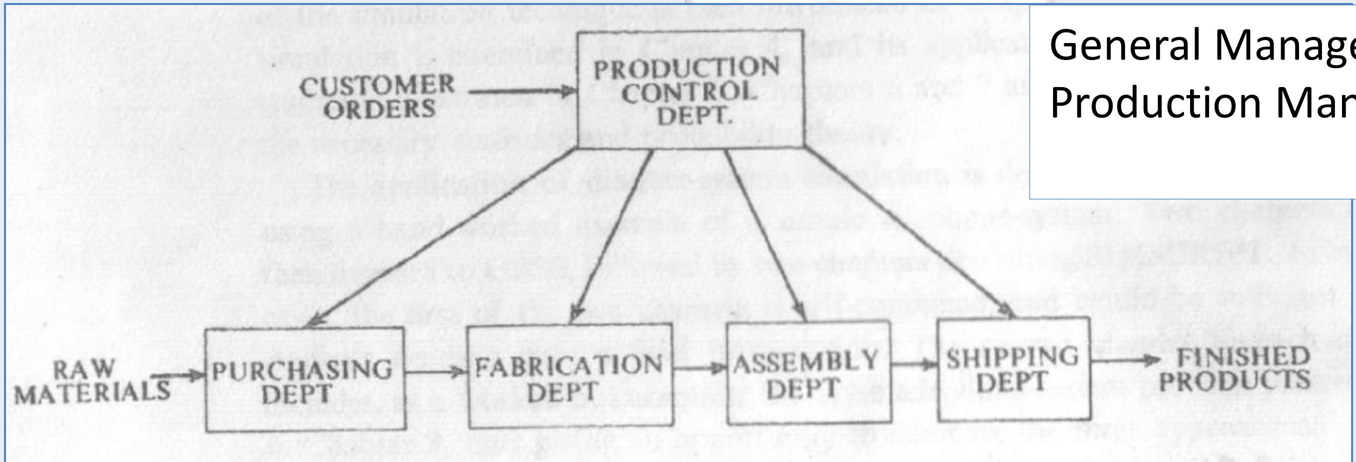


Figure 1-2. A factory system.

General Manager – Satisfied
Production Manager – more details..

The Corporate Model

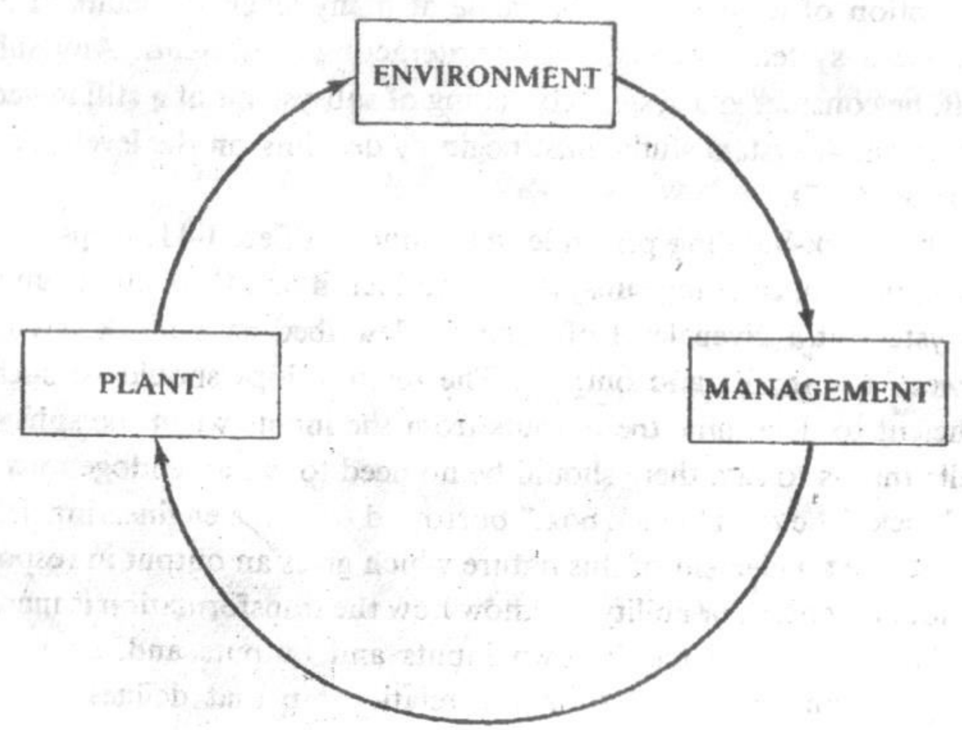


Figure 2-1. Major segments of a corporate model.

- Models of this nature called: **Corporate Model**
- Different corporation use this model in various aspects for **planning their operations**

The Corporate Model

- Assume the corporate is a **manufacturing industry** and
  - Its planning to **produce and market some new product**

- A first level of details might consider the complete model as consisting of three parts … each of a different general nature.
  - The corporation operates within an **environment**
  - A **physical plant** …. Means of production
  - The **management segment**… policy making aspects

The Environment Segment

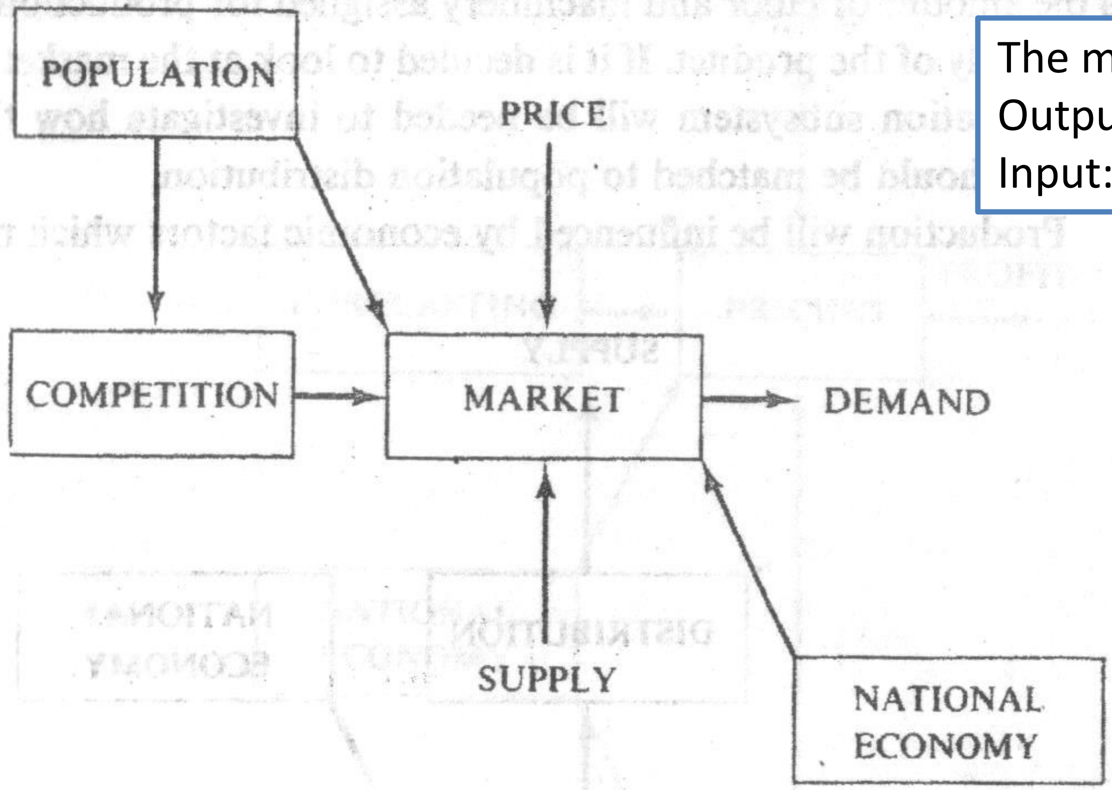


Figure 2-2. The environment segment.

The major element: **MARKET**
Output: **Demand**
Input: **Price & Supply**

The Production Segment

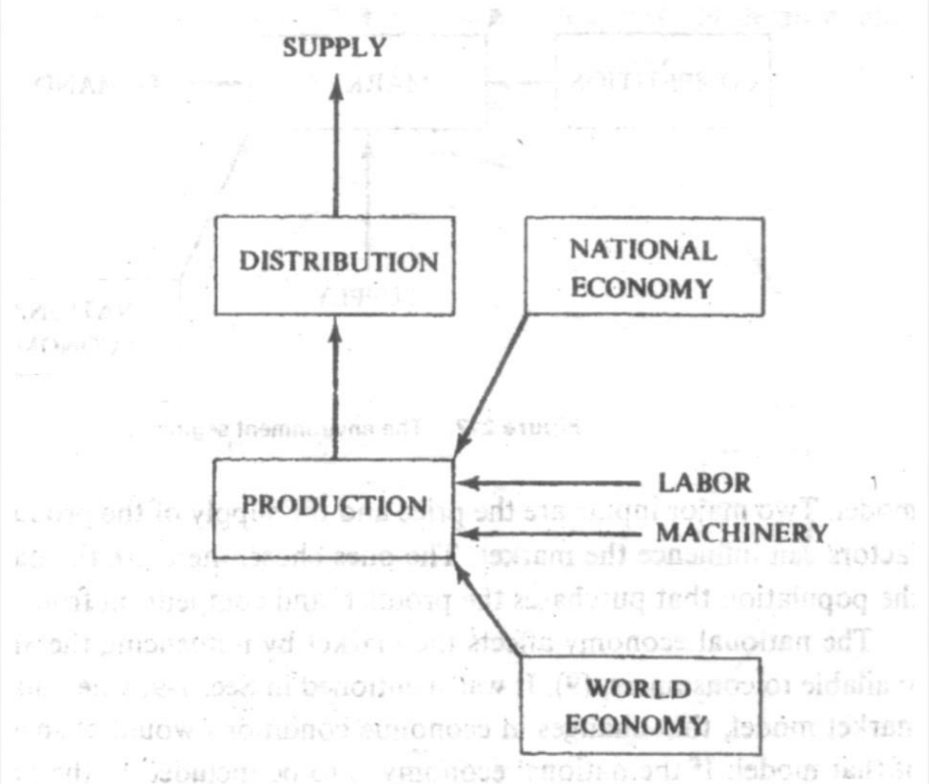


Figure 2-3. The production segment.

Main Input: **LABOR & MACHINERY**
Main Output: **SUPPLY**

The management Segment

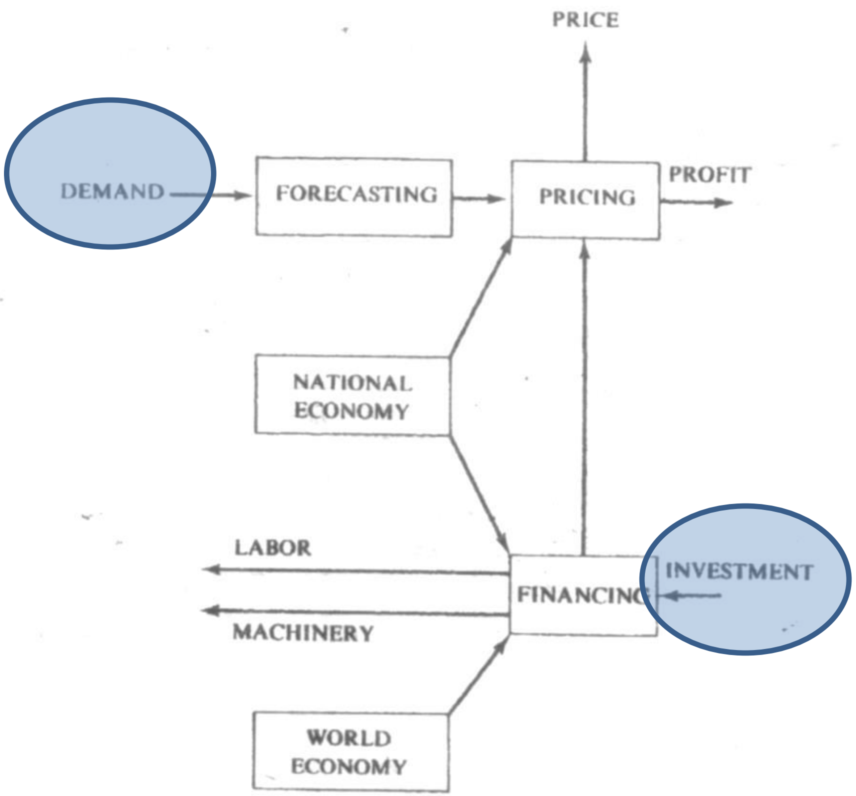


Figure 2-4. The management segment.

Main Input: **DEMAND & INVESTMENT**
Main Output: **Price to be set & profit to be achieved**

Full Corporate Model

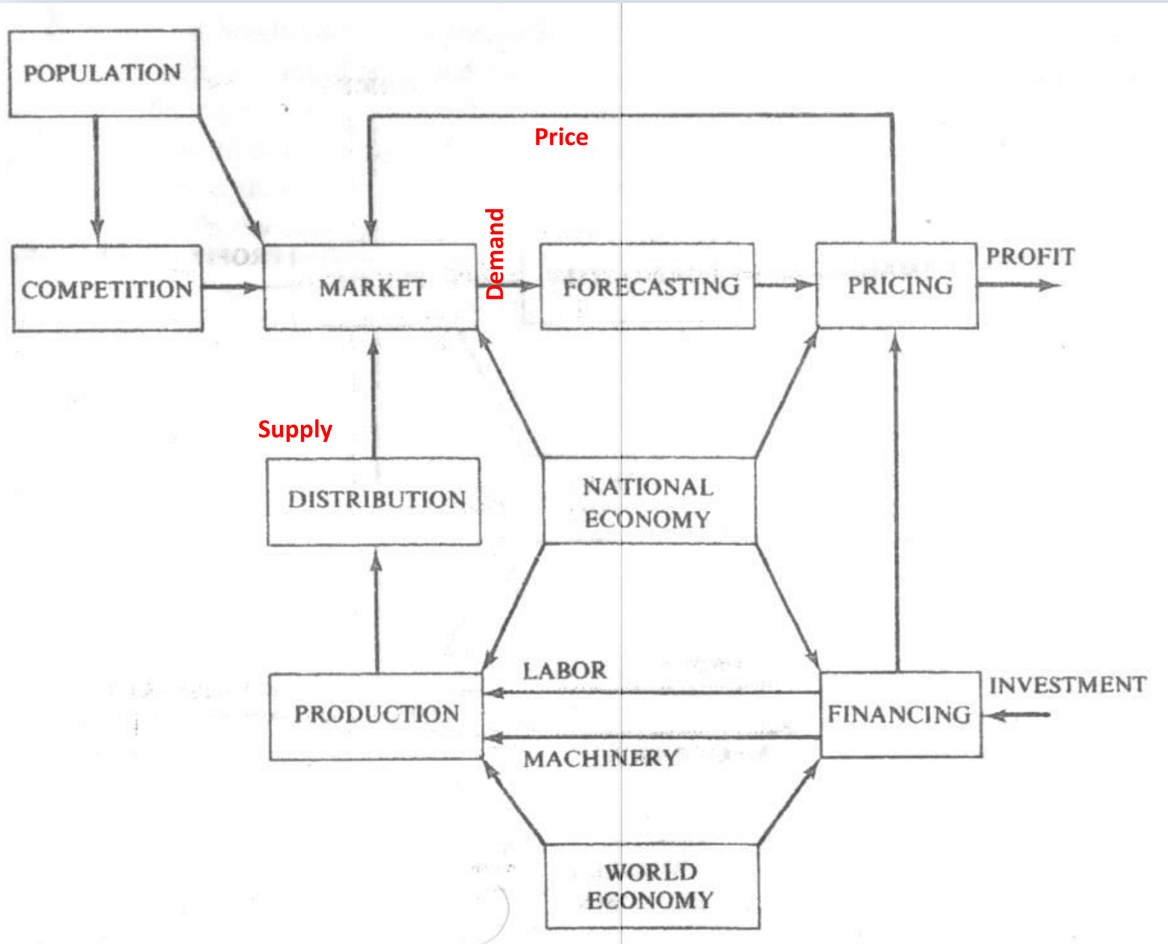


Figure 2-5. The full corporate model.

Types of System Study

- Use of the corporate model?
- Having developed a model….
  - It could be used to study a system.
- System study → 3 ways
  - System analysis
  - System Design
  - System Postulation

System Analysis

- System analysis Aims to
  - Understand how an existing system or a proposed system operates
  - In ideal situation …. Investigator … … … able to do some experiment with the system
  - What actually done … to construct a model of the system and investigate the model.

System Design

- The objective
  - is to produce a system that meets some specification.
  - Certain system parameters or components are chosen by designer
    - And conceptually he chooses a combination ….. to construct the system.

System Postulation

- Postulation → hypothesis
- …is the characteristic of the way models are employed in social, economic, political and medical studies.
  - Where the behavior of the system Unknown
  - But the Process that produce the behavior is known.

Class-11

Simulation and Modeling

CSE4131_CSE4132

The Corporate Model

- Models of this nature called:

  Corporate Model

- Different corporation use this model in various aspects for **planning their operations**

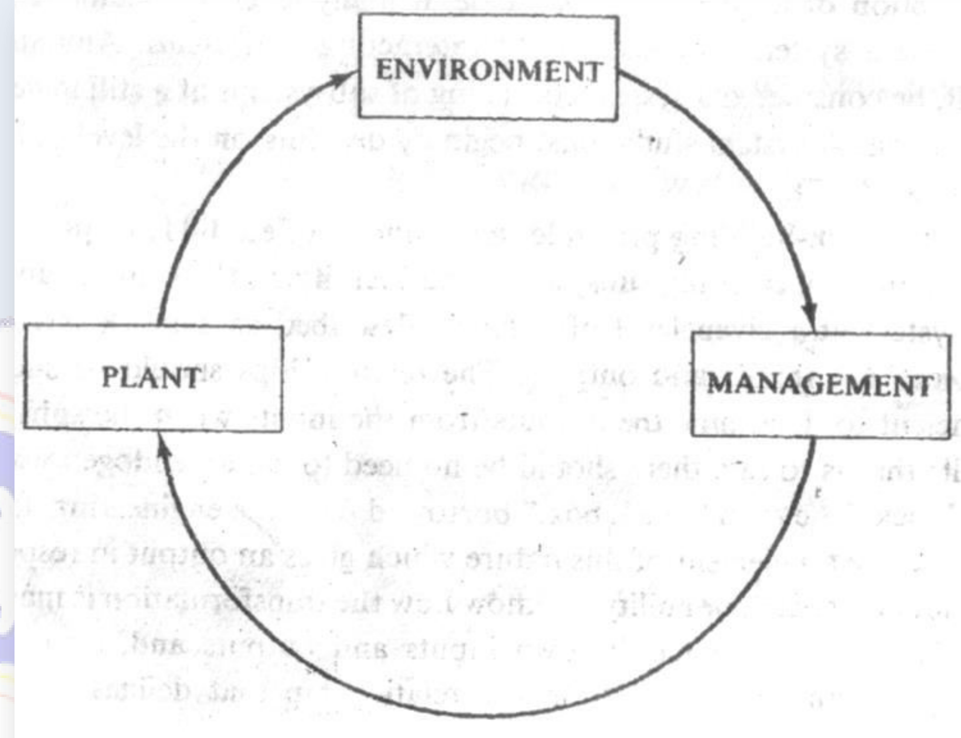


Figure 2-1. Major segments of a corporate model.

The Corporate Model

- Assume the corporate is a **manufacturing industry** and
  - Its planning to **produce and market some new product**

- A first level of details might consider the complete model as consisting of three parts … each of a different general nature.
  - The corporation operates within an **environment**
  - A **physical plant** …. Means of production
  - The **management segment**… policy making aspects

The Environment Segment

The major element: **MARKET**

Output: **Demand**

Input: **Price & Supply**

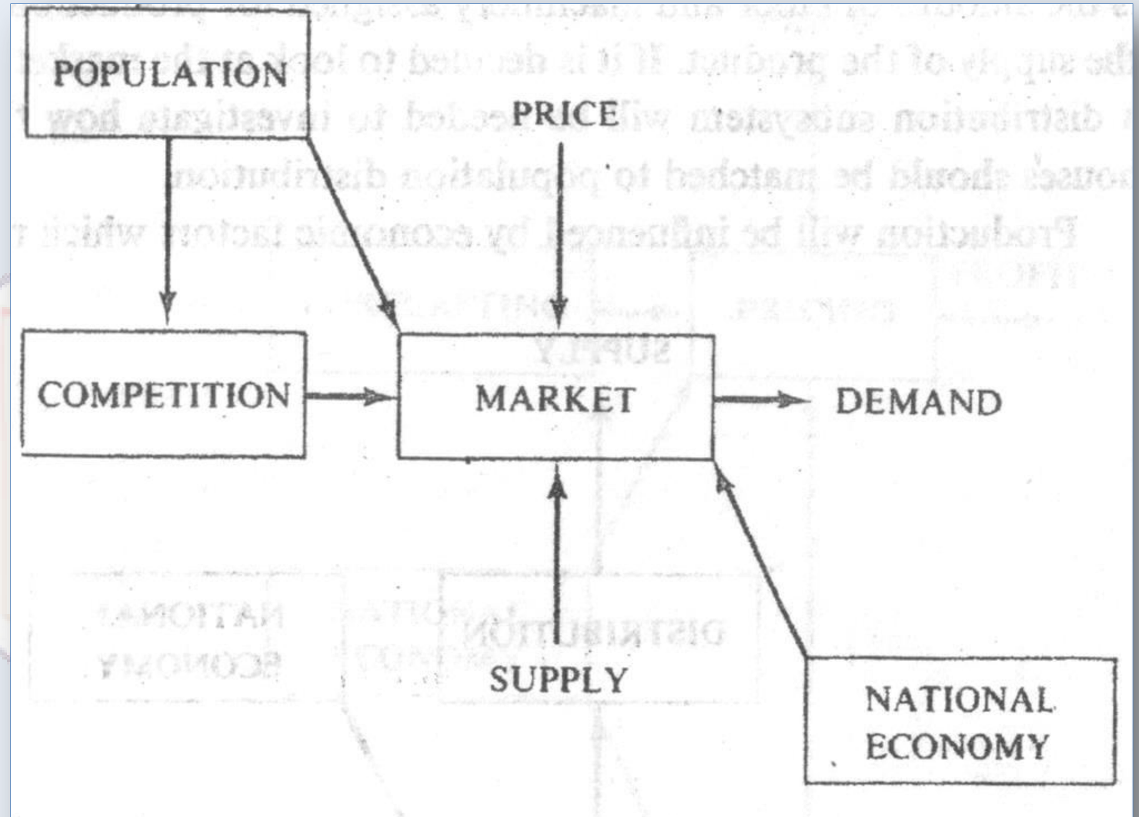


Figure 2-2. The environment segment.

The Production Segment

Main Input: **LABOR & MACHINERY**

Main Output: **SUPPLY**

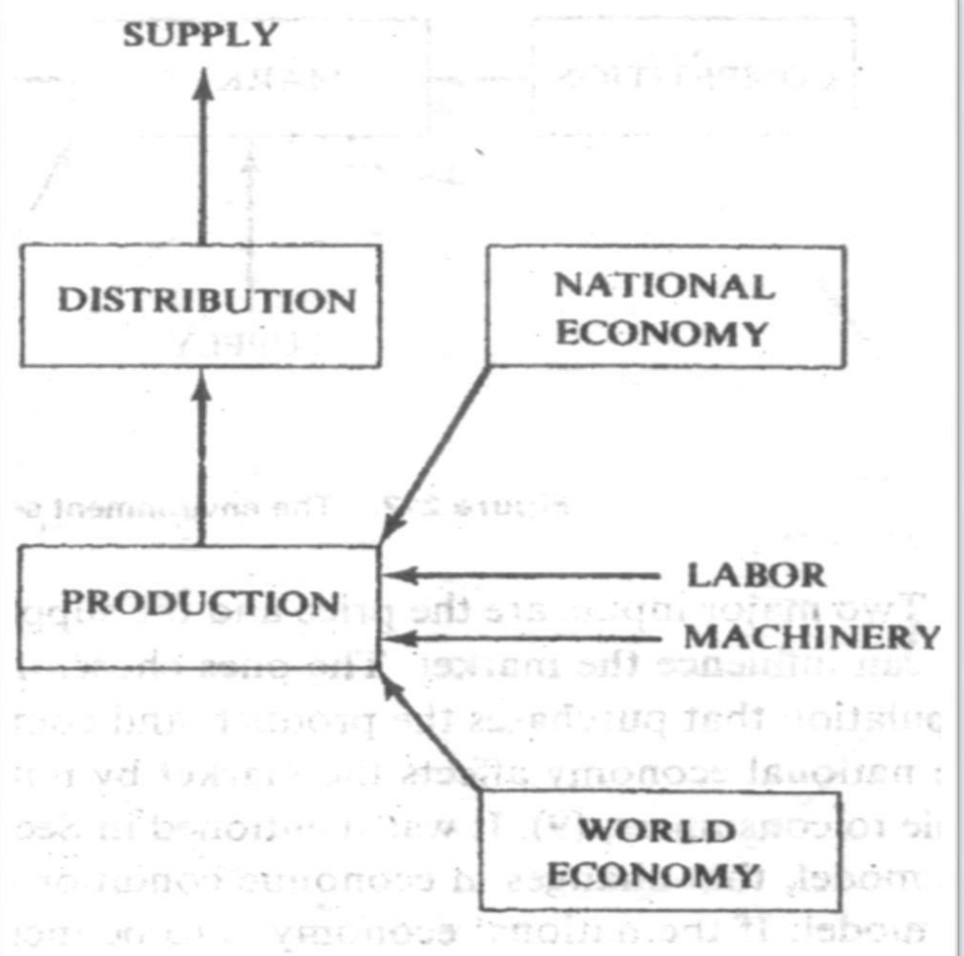


Figure 2-3. The production segment.

The management Segment

Main Input: **DEMAND & INVESTMENT**

Main Output: **Price to be set & profit to be achieved**

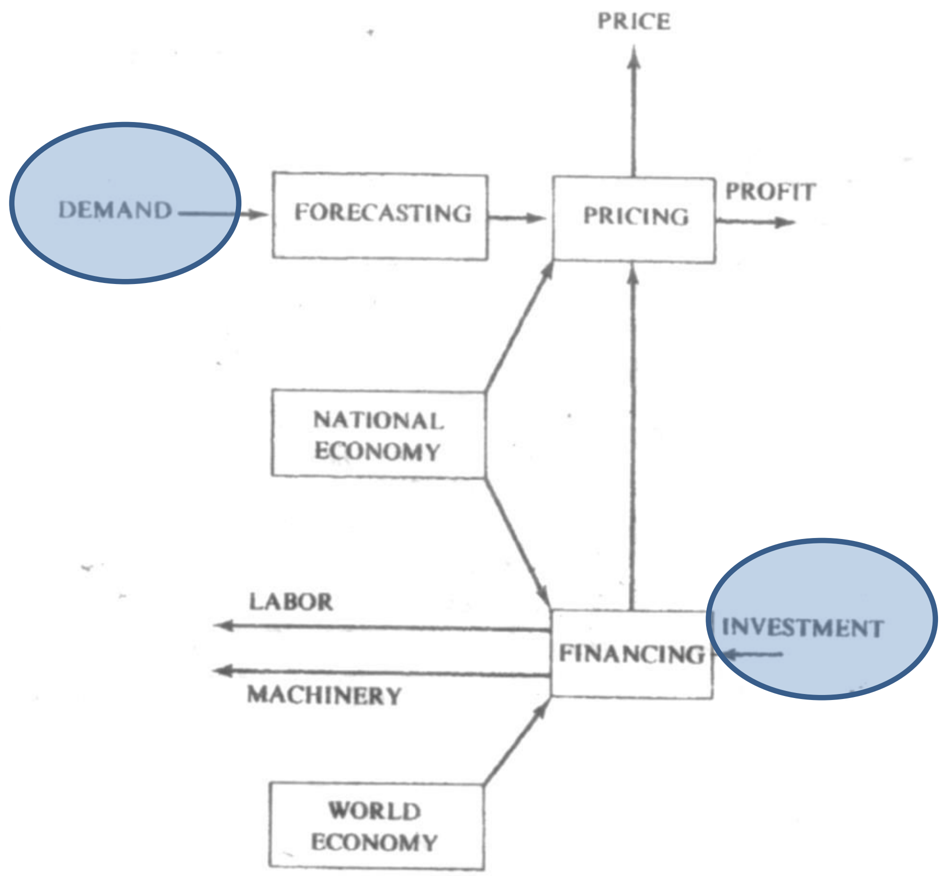


Figure 2-4. The management segment.

Full Corporate Model

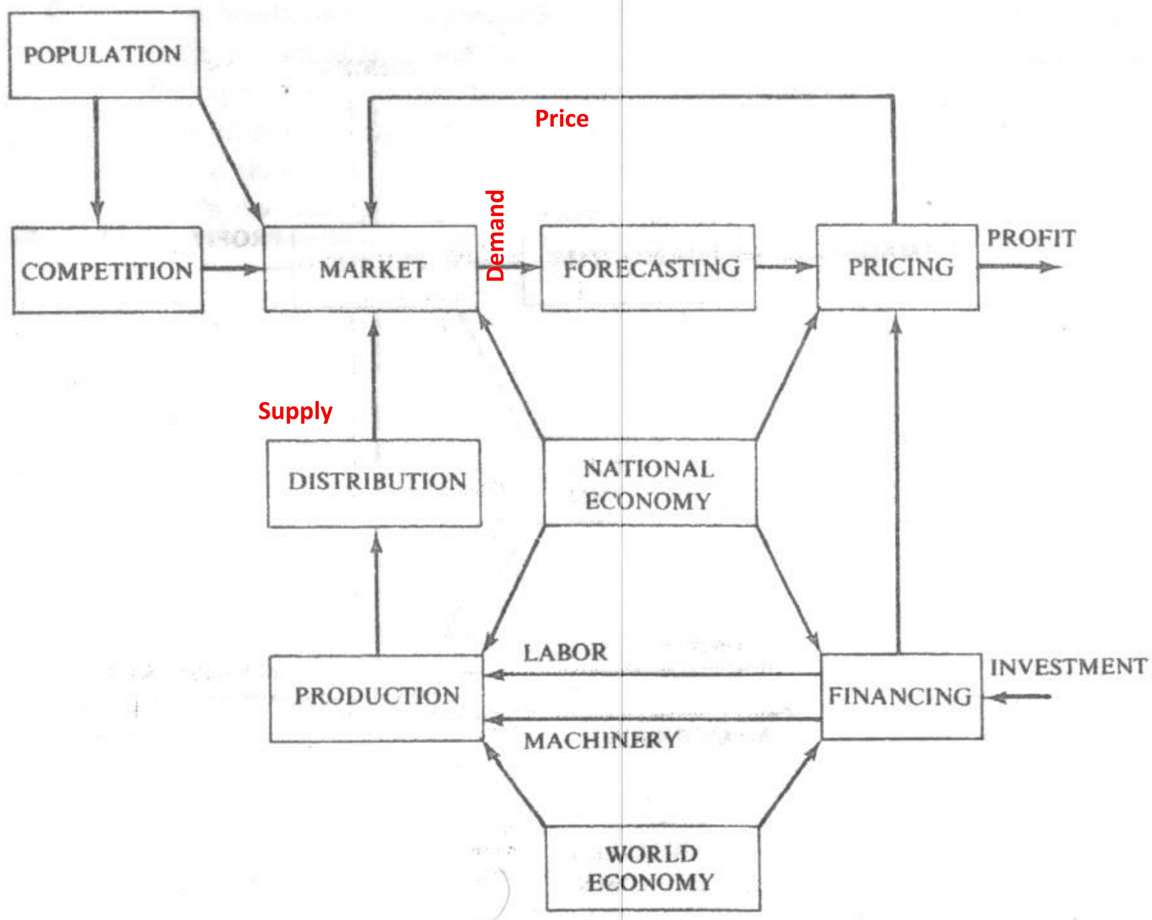


Figure 2-5. The full corporate model.

Full Corporate Model

Types of System Study

- Use of the corporate model?
- Having developed a model….
  - It could be used to study a system.
- System study → 3 ways
  - System analysis
  - System Design
  - System Postulation → hypothesis

System Analysis

- System analysis Aims to
  - Understand how an existing system or a proposed system operates
  - In ideal situation …. Investigator … … … able to do some experiment with the system
  - What actually done … to construct a model of the system and investigate the model.

System Design

- The objective
  - is to produce a system that meets some specification.
  - Certain system parameters or components are chosen by designer
    - And conceptually he chooses a combination ….. to construct the system.

System Postulation

- Postulation → hypothesis
- …is the characteristic of the way models are employed in social, economic, political and medical studies.
  - Where the behavior of the system Unknown
  - But the Process that produce the behavior is known.

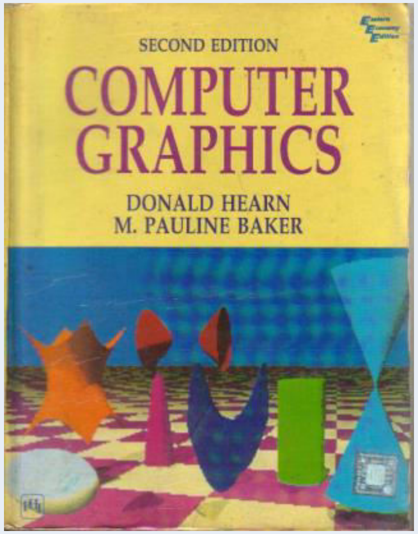
Class-12

Simulation and Modeling

CSE4131_CSE4132

Reference Book

- Computer Graphics (2nd Edition)
  - D. HEARN
  - M. P. BAKER



A Silly Question

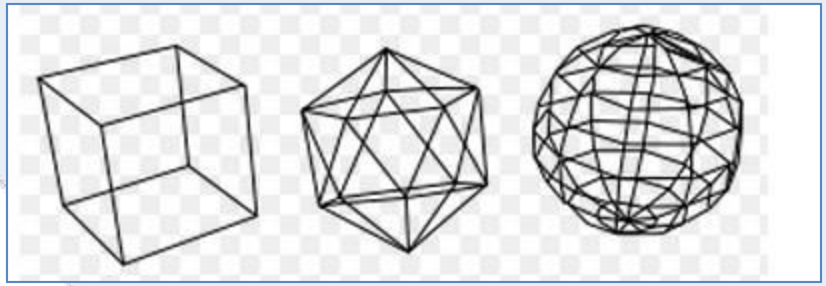
- Which one is to build first?
  - Simulation or Model?
  - Definitely we have to build Model first
  - Then use the model for simulation

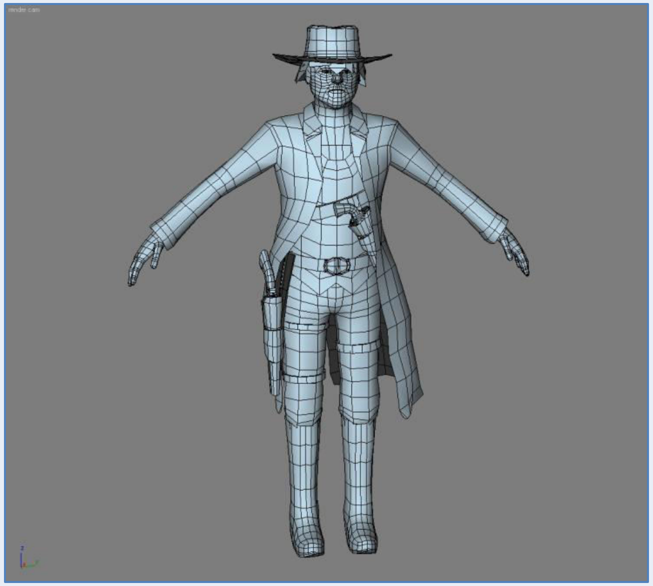
Computer Representation of a Model

- Recall –
  - We are interested in mathematical model
  - Mathematical model ⇨ number +symbols + equations
  - Computer representation ⇨ numbers and equations
  - But more perceivable if graphically represented…
  - Graphically representation ⇨ 2D or 3D
  - Usually we prefer 3D graphics representation.

Computer Representation of a Model

- 3D Model
  - Wireframe model
  - Solid Model
- A wire frame model is a visual presentation of a 3 dimensional or physical object used in 3D computer graphics.
  - created by specifying each edge of the physical object where two mathematically continuous smooth surfaces meet,
  - or by connecting an object's constituent vertices using straight lines or curves.





3D Object Representation

- Graphics scenes can contain many different kinds of objects:
  - flowers**,** clouds, rocks, water, bricks, wood paneling, rubber, paper, marble, steel, glass, plastic, cloth….

- There is <u>no one method that</u> we can use to describe objects that will include all characteristics of these different materials.

- …. we need to use representations that accurately model object characteristics.

3D Object Representation

- Representation of Object
  - 2 broad categories
- **Boundary representations (B-reps)**
  - describe a three-dimensional object as a set of surfaces that separate the object interior from the environment.
  - **E.g. polygon facets** and **spline patches**.
- **Space-partitioning**
  - representations are used to describe interior properties, by partitioning the spatial region containing an object into a set of small, non-overlapping, contiguous solids (usually **cubes).**
  - **E.g. Octrees**

Boundary representations (B-reps)

- The most commonly used boundary representation … for a 3-dimensional graphics object is a set of **surface polygons** that enclose the object interior**.**
- This **simplifies and speeds** up the surface rendering (depiction) and display of objects, since all surfaces are described with linear equations.
- Often referred as ” **Standard Graphics Object**”

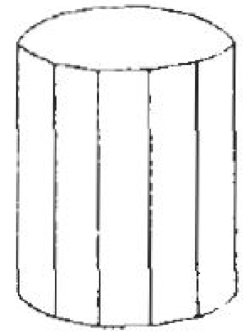


Figure 10-1
Wireframe representation of a cylinder with back (hidden) lines removed.

(***Polygon:*** *a two-dimensional geometric figure formed of three or more straight lines)*

Geometric Tables

- A convenient organization for storing geometric data is to create three lists:

1. A Vertex Table,
   (Vertex → highest point **)**
2. An Edge Table, And
   (Edge → boundary/ border)
3. A Polygon Table Or Polygon Surface Table

Polygon Tables

- We specify a polygon surface with
  - a set of **vertex coordinates** and
  - **associated attribute** parameters.
    - Vertex – highest point

- **Polygon data tables** can be organized into two groups:
  - Geometric tables and
  - Attribute tables.

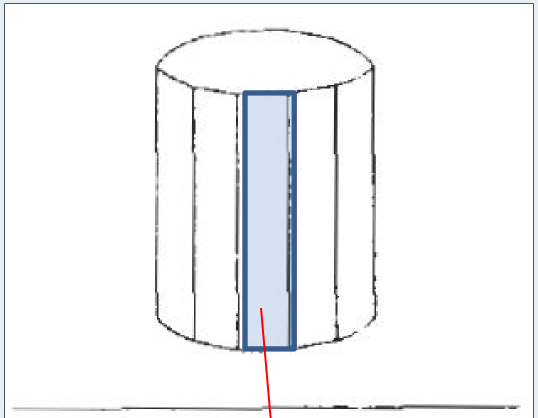
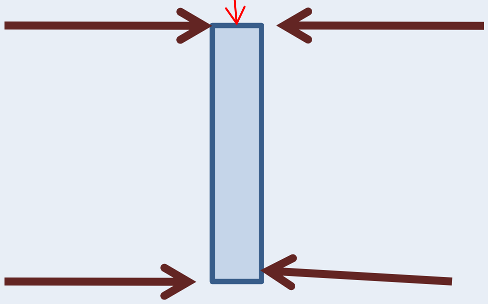
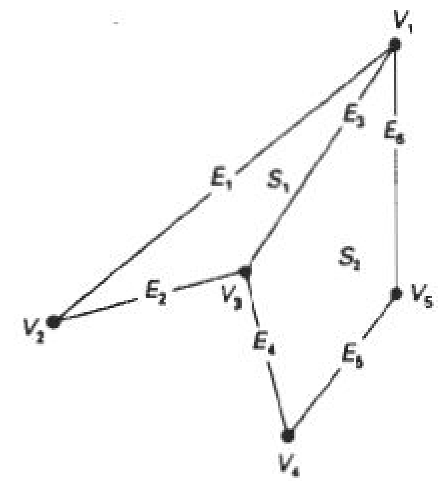


Figure 10-1
Wireframe representation of a cylinder with back (hidden) lines removed.



Geometric Tables

- **Coordinate values** for each vertex in the object are stored in the ***vertex table.***
- The ***edge table*** contains pointers back into the vertex table to identify the vertices for each **polygon** edge.
- And the **polygon table** contains pointers back into the edge table to identify the edges for each polygon.



VERTEX TABLE	
V_1:	x_1, y_1, z_1
V_2:	x_2, y_2, z_2
V_3:	x_3, y_3, z_3
V_4:	x_4, y_4, z_4
V_5:	x_5, y_5, z_5

EDGE TABLE	
E_1:	V_1, V_2
E_2:	V_2, V_3
E_3:	V_3, V_1
E_4:	V_3, V_4
E_5:	V_4, V_5
E_6:	V_5, V_1

POLYGON-SURFACE TABLE	
S_1:	E_1, E_2, E_3
S_2:	E_3, E_4, E_5, E_6

Figure 10-2
Geometric data table representation for two adjacent polygon surfaces, formed with six edges and five vertices.

Blobby Objects

- Some objects do not maintain a fixed shape, but change their surface characteristics in certain motions or when in proximity to other objects

- Examples in this class of objects include:
  - molecular structures, water droplets and other liquid effects, melting objects, and muscle shapes in the human body

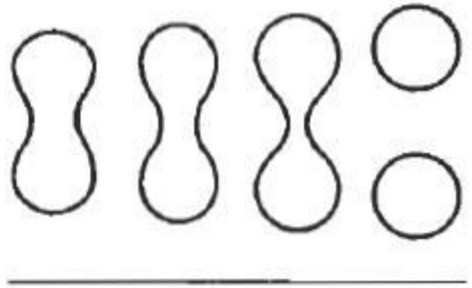


Figure 10-14
Molecular bonding. As two molecules move away from each other, the surface shapes stretch, snap, and finally contract into spheres.

Spline Representations

- A **spline** is a flexible strip **used to produce a smooth** curve through a designated **set of points.**
- **Several small weights are distributed** along the length of the strip to hold it in position on the drafting table as the curve is drawn.
- The term ***spline curve originally referred to a curve drawn in this*** manner.
- We can mathematically describe such a curve with a piecewise **cubic polynomial function**

whose first and second derivatives are continuous across the various curve sections

Spline Representations

- The term **spline curve** now refers to any composite Curve formed with polynomial section satisfying specified continuity conditions at the boundary of the pieces.

- **A spline surface can be** described with two sets of orthogonal spline curves.
  - **Orthogonal** means mutually independent, non-redundant, non-overlapping, or irrelevant

Interpolation and Approximation Spline

- We specify a spline curve by giving a set of coordinate positions, called control points, which indicates the general shape of the curve.

- When polynomial sections are fitted so that the curve passes through each control point, **the resulting curve is said to interpolate the** set of control points.

Figure 10-19
A set of six control points interpolated with piecewise continuous polynomial sections.

Interpolation and Approximation Spline

- When the polynomials are fitted to the general control-point path **without necessarily passing through any control point,** the resulting curve is said to **approximate** the set of control points

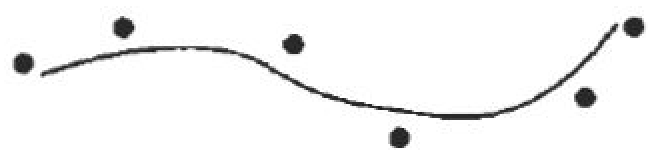


Figure 10-20
A set of six control points approximated with piecewise continuous polynomial sections

Convex Hull

- Convex → Curved
- Hull → Outer Covering

- The convex polygon boundary that encloses a set of control points is called the **convex hull**.

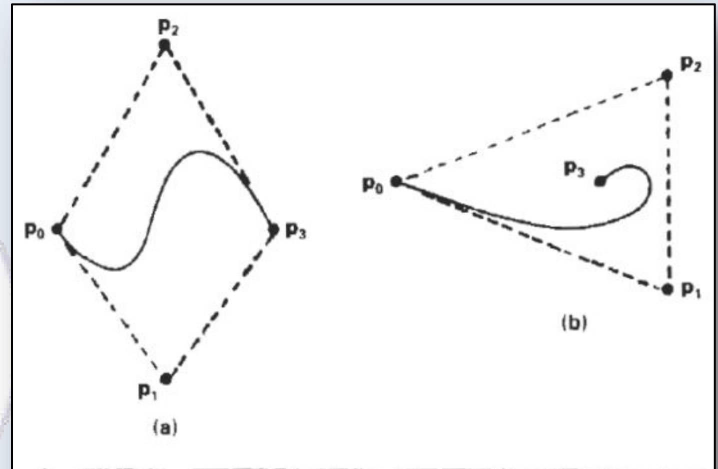


Figure 10-22
Convex-hull shapes (dashed lines) for two sets of control points.

Control Graph

- A **polyline** connecting the sequence of control points for an **approximation spline** is usually <u>displayed to remind a designer</u> of the control-point ordering.
- This set of connected line segments is often referred to as the ***control graph*** of the curve.
- Also Called: **control polygon and characteristic polygon**.

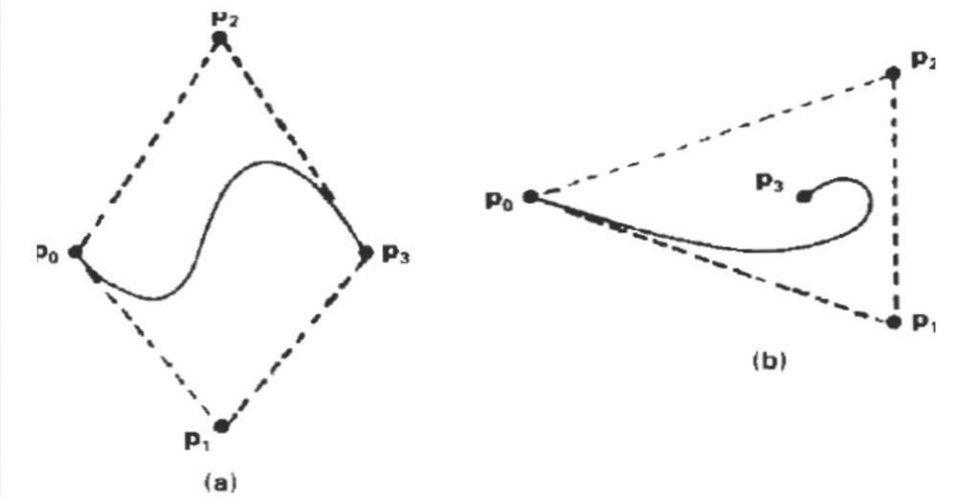


Figure 10-22
Convex-hull shapes (dashed lines) for two sets of control points.

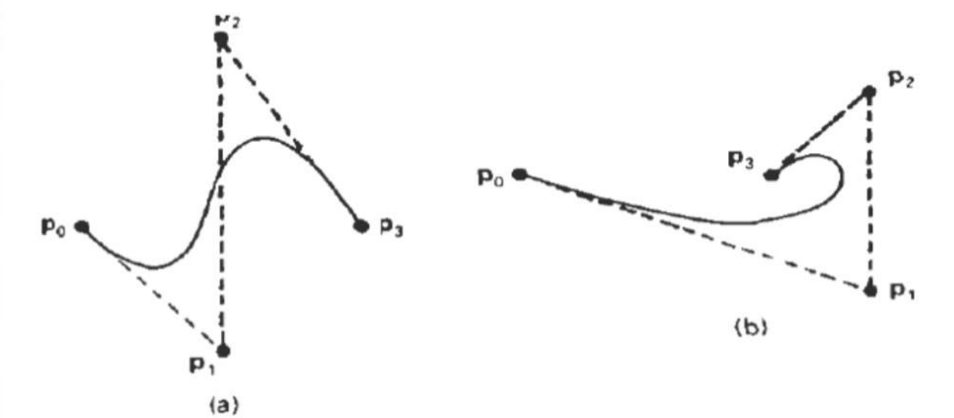


Figure 10-23
Control-graph shapes (dashed lines) for two different sets of

Parametric Continuity Conditions

- To ensure a smooth transition from one section **of a piecewise parametric curve** to the next, **we can impose various continuity conditions at the connection** points.
- **If each section of a spline is described with a set of parametric coordinate** functions of the form:

$$x = x(u), \qquad y = y(u), \qquad z = z(u), \qquad u_1 \le u \le u_2 \qquad (10\text{-}20)$$

We set parametric continuity by matching the parametric derivatives of adjoining curve sections at their common boundary.

Parametric Continuity Conditions

- Zero-order parametric continuity,
  - described as C^0 continuity, means simply that the curves meet. That is, the values of x, y, and *z evaluated at* u_2 *for the first* curve section are equal, respectively, to the values of x, y, and *z evaluated at* u_1 for the next curve section
- First-order parametric continuity,
  - referred to as C^1 continuity, means that the first parametric derivatives (tangent lines) of the coordinate functions in Eq. 10-20 for two successive curve sections are equal at their joining point.
- Second-order parametric continuity,
  - Or C^2 continuity, means that both the first and second parametric derivatives of the two curve sections are the same at the intersection, Higher-order parametric continuity conditions are defined similarly.
- Figure 10-24 shows examples of C^0, C^1, *and* C^2 *continuity.*

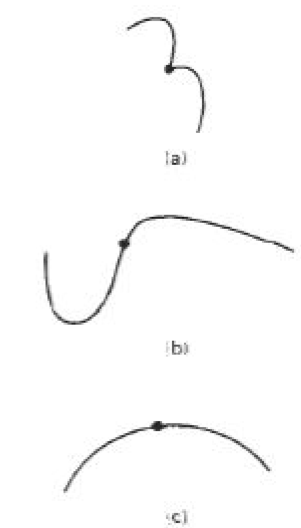


Figure 10-24
Piecewise construction of a curve by joining two curve segments using different orders of continuity: (a) zero-order continuity only, (b) first-order continuity, and (c) second-order continuity.

Spline Specifications

- There are **three equivalent methods** for specifying a particular spline representation:
  1) We can state the **set of boundary conditions** that are imposed on the spline;
  2) We can **state the matrix** that characterizes the spline;
  3) We can state the **set of blending functions** (or **basis functions**) that determine how specified geometric constraints on the curve are combined to calculate positions along the curve path.

Some Commonly Used Splines

- Some commonly used splines and their matrix and blending-function specifications
  - Cubic Spline Interpolation Method
    - Natural Cubic Splines
    - Hermit Spline
    - Cardinal Splines
    - Kochanek-Bartels Splines

Cubic Equation

A cubic equation is an equation involving a cubic polynomial, i.e., one of the form

$$a_3 x^3 + a_2 x^2 + a_1 x + a_0 = 0.$$

Since $a_3 \neq 0$ (or else the polynomial would be quadratic and not cubic), this can without loss of generality be divided through by a_3, giving

$$x^3 + a_2' x^2 + a_1' x + a_0' = 0.$$

– to **provide a representation for an existing object** or **drawing**, +
– also used sometimes **to design object shapes.**

- Cubic polynomials offer a reasonable compromise between **flexibility and speed of computation.**
  – Compared to higher-order polynomials, cubic splines require **less calculations and memory** and they are **more stable**.
  – Compared to lower-order polynomials, cubic splines are **more flexible for modeling** arbitrary curve shapes

Cubic Spline Interpolation Method

- Given a set of control points,
  - cubic interpolation splines are obtained by fit ting the input points with a piecewise cubic polynomial curve that **passes through every control point**.
- Suppose we have n + 1 control points specified with coordinates

$$\mathbf{p}_k = (x_k, y_k, z_k), \qquad k = 0, 1, 2, \ldots, n$$

- A cubic interpolation fit of these points is illustrated in Fig. 10-26

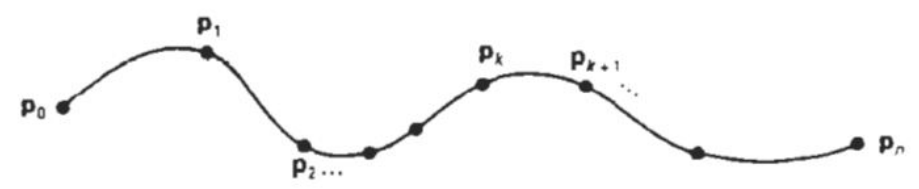


Figure 10-26
A piecewise continuous cubic-spline interpolation of $n + 1$ control .points.

Cubic Spline Interpolation Method

- We can describe the parametric cubic polynomial that is to be fitted between each pair of control points with the following set of equations:

$$x(u) = a_x u^3 + b_x u^2 + c_x u + d_x$$

$$y(u) = a_y u^3 + b_y u^2 + c_y u + d_y, \qquad (0 \le u \le 1) \qquad (10\text{-}26)$$

$$z(u) = a_z u^3 + b_z u^2 + c_z u + d_z$$

- For each of these three equations, we need to determine
  - We do this by the values of the four coefficients **a, b, c,** and **d** in the polynomial representation for each of the n curve sections between the n + 1 control points.