7th Sem

1st internal preparation

(5) subject

⇒ DST - 1 unit full
⇒ DL - 1 and half unit till hebbian smthg
ME - 2 units
IS - 1 and half unit till modular operator diffe fnctn
DSE - may be 2 unit

Sub ① DST (Distributed storage Technologies) (IS72)

⇒ * information storage
* evolution of storage architecture
* Data centre infrastructure
* virtualization & cloud computing
⇒ * Data centre env application
* DBMS
* Host * Connectivity * Storage
* Disk Drive components
* — performance
* Host acces to Data
* Direct Attached storage
* Storage Design Based on application
* Disk Native Command Queuing
* Introduction to flash Drives

Sub ② DL (Deep Learning) (ISE741)

⇒ Unit ① * Human brain * neuron models
* neural nets as directed graphs * feedback
* neural architectures
* Knowledge representation
* Connection to artificial intelligence

Unit - (2) Learning process
* Error-Correction learning
* memory based learning
* Hebbian learning.

Sub (3) IS (Information Security)

=> Unit (1) * Symmetric cipher model
* Cryptography * Cryptanalysis
* substitution techniques * Transposition Techniques

Block ciphers & the Data encryption standard
* simplified DES * Block cipher principles
* DES * Strength of DES
* Differential & linear cryptanalysis
* Block cipher Design principles
* Block cipher modes of operation.

=> Unit (2)
* public key algorithms : introduction
* modular arithmetic * RSA
* Diffie Hellman.

Sub (4) MAE (Management and Entrepreneurship)

=> Unit (1) * management - Nature & fun^{n} of management
* importance * Definition * management fun^{n}
* levels of management * Roles of a senior manager
* managerial skills
* Development of management Thought - Early classical approaches - Scientific management.
* administrative management * Bureaucracy.

-> unit ② planning : * nature * importance
* Types of plans (Definitions & meaning only)
* Steps in planning * strategic planning proces

Decision making : * meaning * Types
* steps in rational decision making
* difficulties in decision making

Coordination : * need for coordination
* Requisites of Excellent coordination & types of coordination.

Sub ⑤ DS (Data Science)

DST

Subj ①

=> information storage:

* information is the knowledge derived from data.

Data

It is a collection of raw facts from which conclusions may be drawn

2 types

* structured (10%)
* unstructured (90%) (majority of data)

factor of digital data growth

* increase in data processing capabilities
* Lower cost of digital storage
* Affordable & faster com^n technology
* proliferation of app^s & smart devices

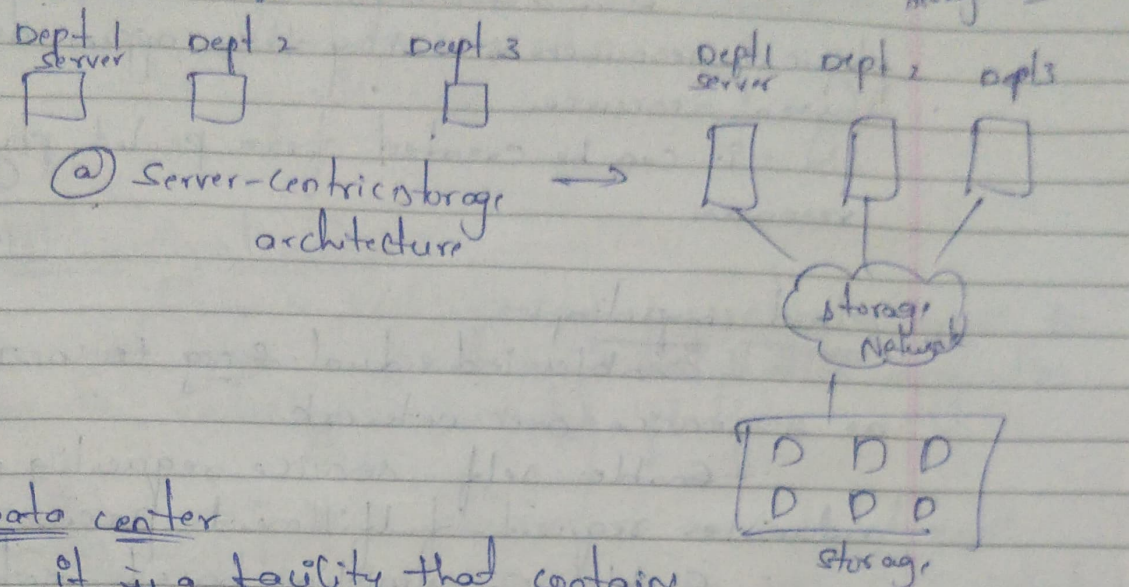
Bigdata

it refers to data sets whose sizes are beyond the ability of commonly used software tools to capture, store, manage and process within acceptable time limits.

storage

it's created by individual & organizations (provides access to data for further processing)

evolution of storage architecture



(a) Server-centric storage architecture → (b) information-centric storage archi

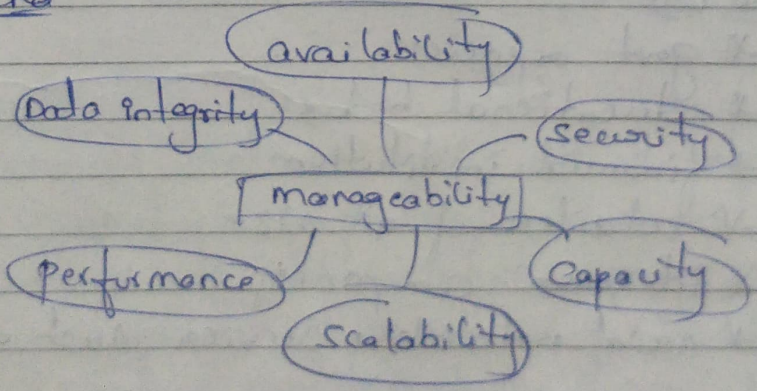
Data center

it is a facility that contain storage, Compute, network, & other IT resource to provide centralized data processing capabilities

Core elements

* Application * DBMS * Host / compute
* Network * Storage

Key char



Key management activities

* monitoring
* Reporting
* provisioning

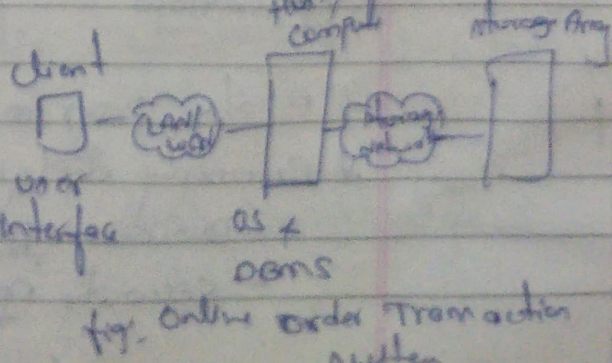


fig. Online order Transaction system

Managing Data Center

* monitoring * Reporting
* Provisioning.

virtualization :-
It is a technique of abstracting physical resources & making them appear as logical resources.
- It can be created from pooled physical resources

Cloud computing :-
- Enables individual & org to use IT resource as a service over network
- Enables self-service requesting & automates request-fulfillment process
- Enables consumption based metering

=> Data Center
It is a facility that centralizes an organization's shared IT operations & equipment for the purpose of storing, processing & disseminating data & applications.

who use
* govt agencies
* educational bodies
* financial institutions
* retailers of all size
* telecomm companies
* social networking service such as google & fb

Applications
* A s/w program that provide logic for computing oper
* Commonly deployed app^n in data center
  - Business appl^n * management app^n
  - security Appl^n - Data protection appl^n

* Key I/O char of an appln
  - Read intensive vs write intensive
  - sequential v.s random
  - I/O request size

Appln virtualization:-
it is a technique of presenting an appln to an end user without any installation, integration on the underlying computing platform

* Allows appln to be delivered in an isolated envt
  - aggregate OS resources & the appln in to virtualized container
  - ensure integrity of OS & appln
  - Avoids conflicts blw diff appln

DBMS:- it is a structured way to store data in logically organized tables that are interrelated
  - MySQL, oracle RDBMS, SQL Server
  - Controls creation, maintenance

Host (compute):-
* Resource that runs appln with the help of underlying computing components
* Consists of hlw & slw components
* Hlw components - CPU, memory, i/o devices
* Slw - OS, device driver, file system

Connectivity :- interconnection blw hosts & blw a host & peripheral devices, such as storage

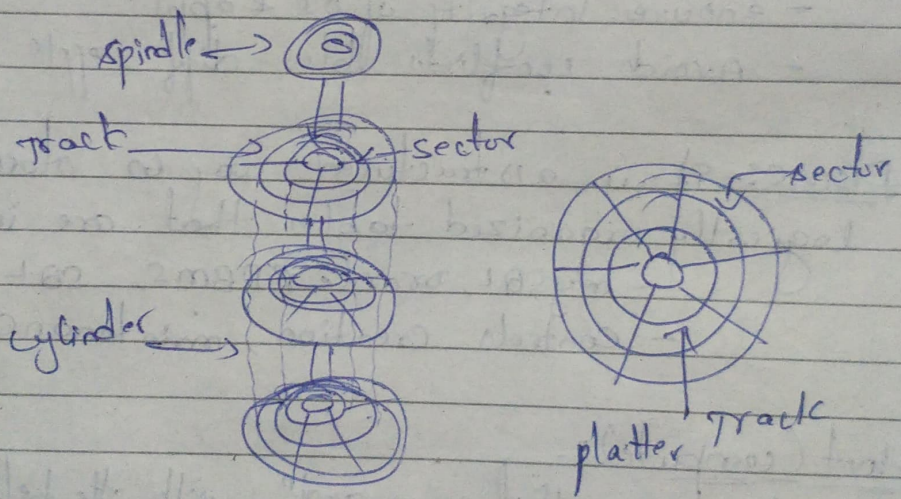
Protocol = a defined format for comm blw sending & receiving devices

Interface protocols
* IDE (ATA (Serial ATA)
* SATA
* SCSI & SAS
* Fibre Channel & IP

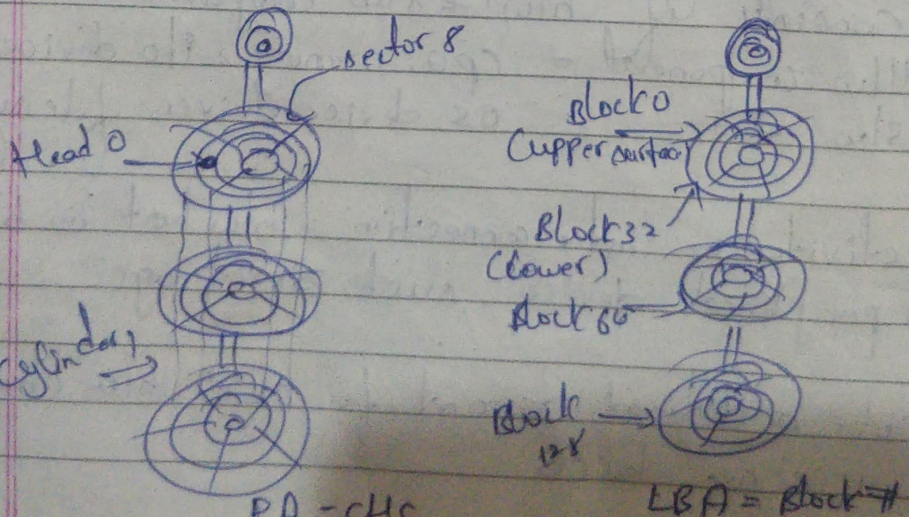
Storage :-
* magnetic Tape
* optical disk
* Disk drive
* flash drives

Disk drive components :- → * platter * spindle * Read/write head
* Actuator Arm assembly
* Drive controller board

* physical disk structure



* Logical Block addressing



Disk drive performance

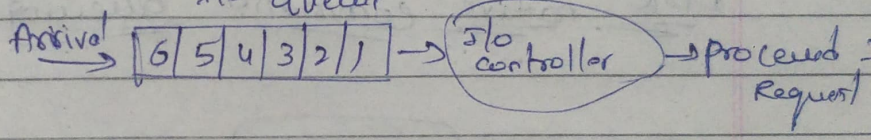
* electromechanical device
* disk service time
  - Seek time → time taken to position the read/write head (= full stroke, = average, = track to track)
  - Rotational latency → The time taken by the platter to rotate & position the data under the R/W head.
  - Data transfer rate → avg amount of data per unit time that the drive can deliver to HBA.

Disk Service time = Seek time + rotational latency + data transfer time

Disk I/O Controller utilization

Q elem^{ts} * Queue * Disk I/O controller

I/O processing



Host access to data

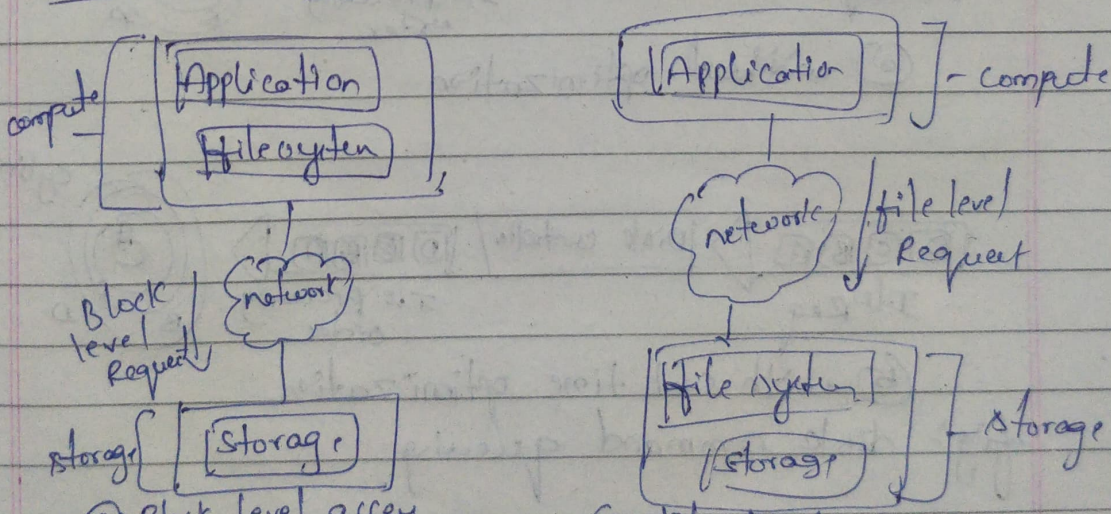


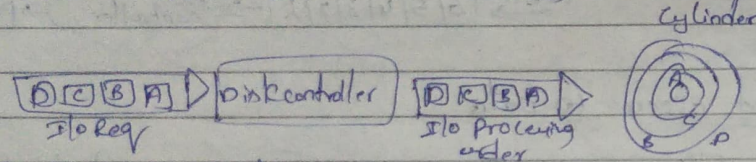
fig: Host access to storage

DAS (Direct Attached storage)

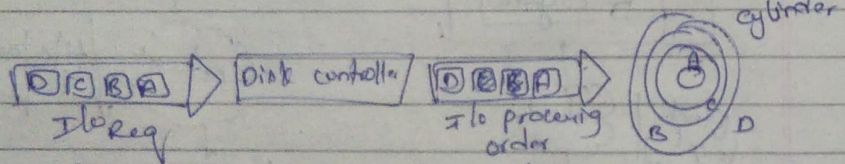
Benefits: * its configuration is simple & can be deployed easily & rapidly
* it require a relatively lower initial investment than storage networking architecture
* it require fewer management tasks & less hlw & slw elements to setup & operate

Limitations :- * DAS does not scale well
* storage array has a limited no of ports
* DAS does not make optimal use of resource due to its limited capability to share front end ports.

Disk native command Queuing



(a) without optimization



(b) with seek time optimization

fig:- disk command queuing.

flash drives (solid state drives (SSD))

Key features

* NAND flash memory technology
* single level cell (SLC) based flash
* write leveling technique

DL Subj (2)

Benefits of neural networks:
* non linearity
* input - output mapping
* adaptivity
* evidential Response
* Neurobiological Analogy
* contextual infor
* fault tolerance
* VLSI implementability
* uniformity of analysis & design

Human Brain

Human nervous system may be viewed as 3 stage system.

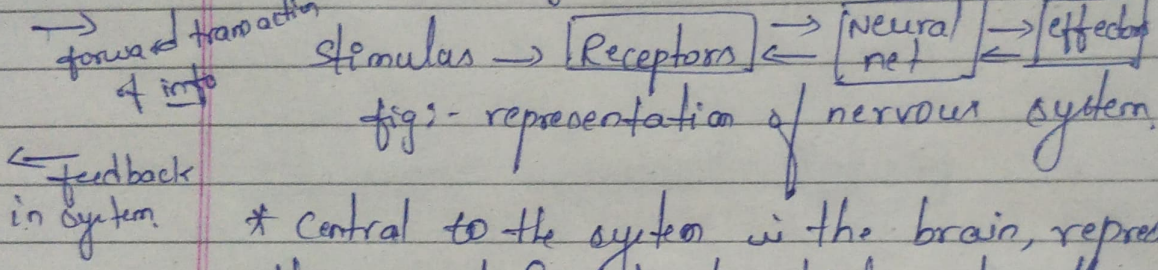
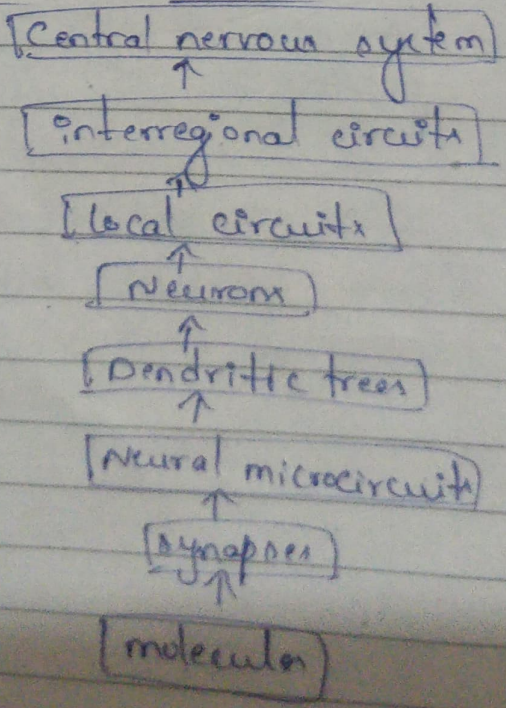


fig:- representation of nervous system

* Central to the system is the brain, represented by the neural (nerve) net, which continually receives information, perceives it & make appropriate decision.

Structural org of levels in brain :-



models of neuron:

A neuron is an information-processing unit that is fundamental to the operation of a neural network.

* A set of synapses/connecting links, each of which is characterized by a weight (or) strength of its own
* An adder for summing the input signals, weighted by the respective synaptic strengths of the neuron
* An activation fun^n for limiting the amplitude of the output of a neuron

Types of Activation fun^n

* threshold fun^n
* sigmoid fun^n

neural network viewed as directed graphs:-

A signal flow graphs is a network of directed links (branches) that are interconnected at certain points called nodes.

(3) rules

* A signal flows along a link only in the direction defined by the arrow on the link

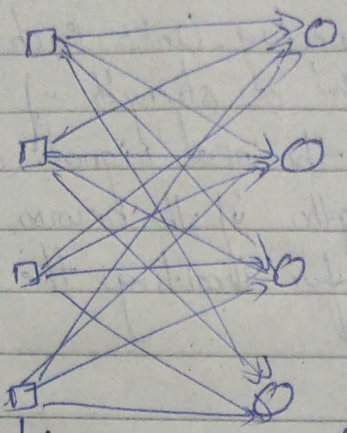
(2) Links (1) synaptic links
(2) activation links

* A node signal equals the algebraic sum of all signals entering the pertinent node via the incoming links
* The signal at a node is transmitted to each outgoing link originating from that node, with the transmission being entirely independent of the transfer fun^n of the outgoing links

feedback :- it is said to exist in a dynamic system whenever the output of an element in the system influence in part the input applied to that particular element, thereby giving rise to one/more closed paths for the transmission of signals around the system.

Network architecture:-

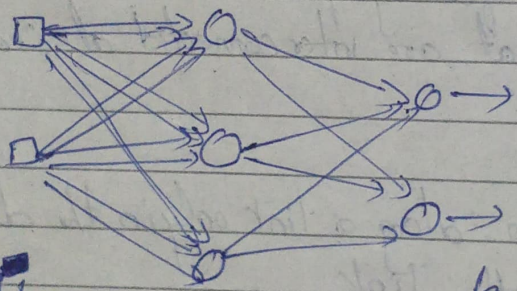
(1) Single-layer feedforward networks



input layer of source node

output layer of neurons

(2) multilayer feedforward networks



input layer of source node

layer of hidden neuron

layer of output neurons

fully connected - every node in each layer of the network is connected to every other node in the single forward layer, if some synaptic connection (connection which are missing from the network then the network is partially connected

(3) Recurrent networks - A recurrent neural network distinguishes itself from a feed forward neural network in that it has at least one feedback loop.

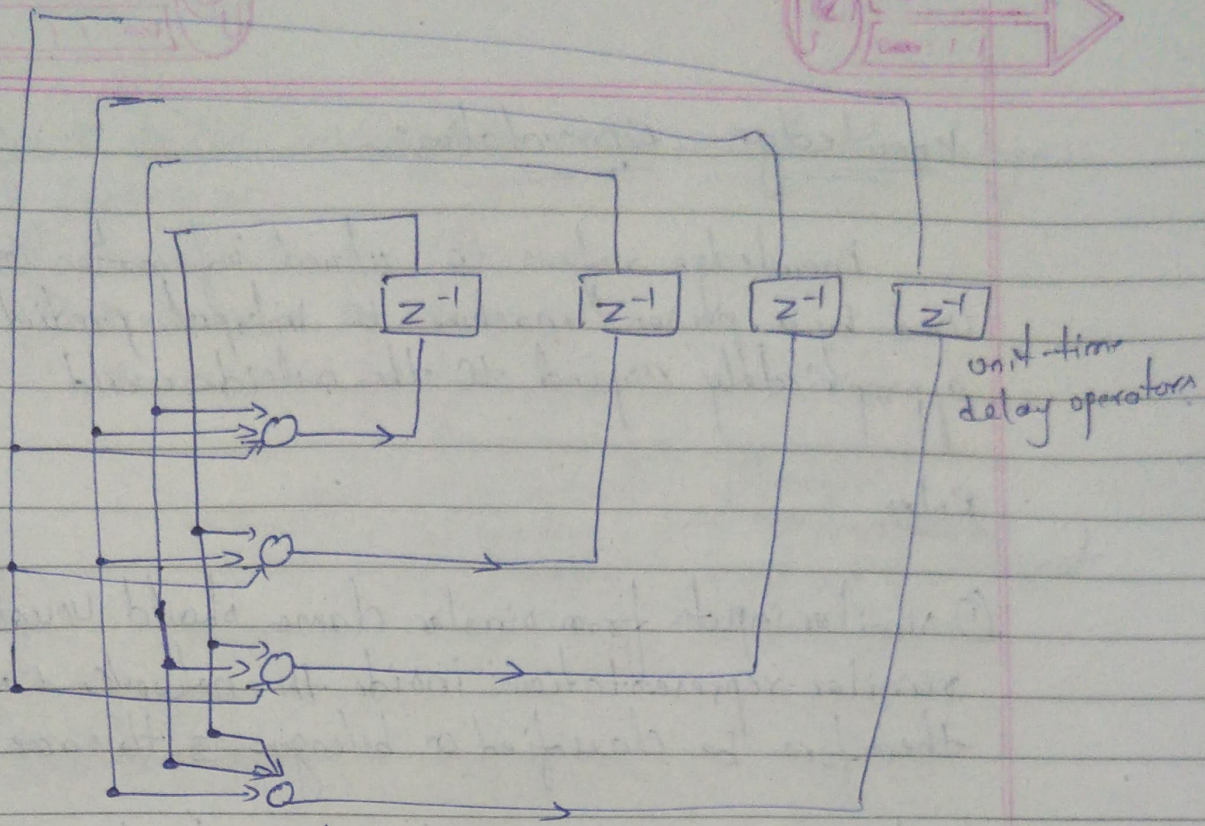


fig:- with no self feedback loops & no hidden neurons

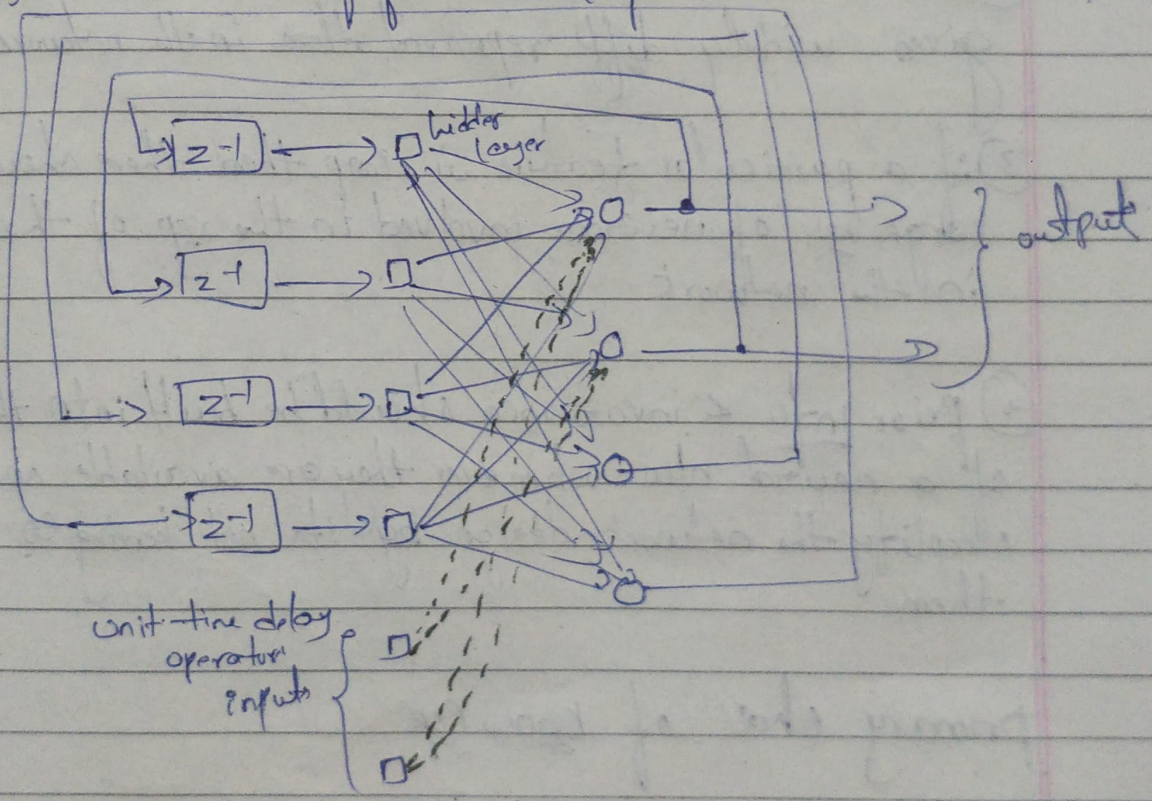


fig: Recurrent network with hidden neurons

Knowledge Representation:-

Knowledge refers to stored information or models used by a person / machine to interpret, predict & appropriately respond to the outside world.

Roles

① similar inputs from similar classes should usually produce similar representations inside the network, & should therefore be classified or belonging to the same class.

② items to be categorized as seperate classes should be given widely diff representation in the network.

③ if a particular feature is imp, then there should be a large no of neurons involved in the rep of that item in the network

④ Prior info & invariance should be built into the design of a neural n/w whenever they are available, so as to simplify the network design by its not having to learn them.

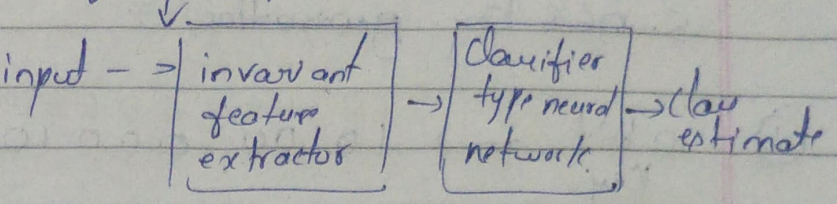
primary char of Know Rep

* what info is actually made explicit
* how the info is physically encoded for subsequent code.

types:- ① learning
② testing
③ generalization

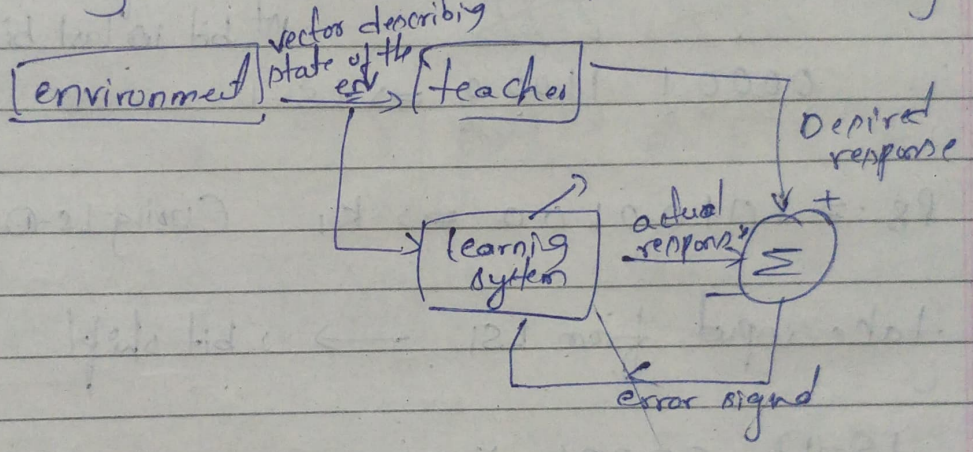
how to build invariance into neural nlw design:

③ techniques
① invariance by structure
② invariance by training
③ invariance feature space



unit ② Learning proces

* Learning with a Teacher (supervised learning)



* Learning without a teacher
- Reinforcement learning
- unsupervised learning

~~Artificial Intelligence~~

⇒ * advantage & disadvantage of neural networks
* diff btw DL & AI
* problem & solution solved
* unit ② need more info

IS

Sub ⑧

Problems

S-DES key generation

① 10 bit key → 1 0 1 0 0 0 0 0 1 0
(1 2 3 4 5 6 7 8 9 10)

$P_{10} = 3,5,2,7,4,10,1,9,8,6$

$P_8 = 6,3,7,4,8,5,10,9$

$P_{10} =$ 1 0 0 0 0 0 1 1 0 0 → 1 bit shift.
1st bit in last bit.

0 0 0 0 1 1 1 0 0 0 ⇒ LS-1
(1 2 3 4 5 6 7 8 9 10)

$P_8 =$ 1 0 1 0 0 1 0 0 ⇒ K_1 (using LS-0)

take input from LS1 → 2 bit shift.

LS-1 0 0 0 0 1 1 1 0 0 0

0 0 1 0 0 0 0 0 1 1 → LS-2
(1 2 3 4 5 6 7 8 9 10)

$P_8 =$ 0 1 0 0 0 0 1 1 ⇒ K_2 //

② 1 2 3 4 5 6 7 8 9 10
0 0 1 0 0 1 0 1 1 1

P_{10} → 1 0 0 0 0 1 0 1 1 1

LS_1 → 0 0 0 0 1 . 0 1 1 1 1

P_8 → 0 0 1 0 1 1 1 1 ⇒ K_1

LS_2 → 0 0 1 0 0 1 1 1 0 1 → 2 bit shift (i/p LS_1)

P_8 → 1 1 1 0 1 0 1 0 ⇒ K_2 //

(3) Using keyword :-

Keyword – GRADE 5X5 matrix

Results will be announced tomorrow.

G	R	A	D	E
B	C	F	H	I/J
K	L	M	N	O
P	Q	S	T	U
V	W	X	Y	Z

RESULTSWILLBEANNOUNCED TOMARROW

AG TP NQ QX CO KC GD UZ LH GE UN SF LZ

(4) Hill cipher

$K = \begin{bmatrix} 7 & 8 \\ 11 & 11 \end{bmatrix}$ $P = ?$

A	B	C	D	E	F	G	H
0	1	2	3	4	5	6	7

I	J	K	L	M	N	O	P
8	9	10	11	12	13	14	15

Q	R	S	T	U	V	W	X
16	17	18	19	20	21	22	23

Y	Z
24	25

Short example

$P = (18,7)\ (14,17)\ (19,4)\ (23,0)\ (12,15)\ (11,4)$

$C = KP \bmod 26$

$= \begin{bmatrix} 7 & 8 \\ 11 & 11 \end{bmatrix} \begin{bmatrix} 18 \\ 7 \end{bmatrix} \bmod 26$

$= \begin{bmatrix} (7\times18)+(8\times7) \\ (11\times18)+(11\times7) \end{bmatrix} \bmod 26 = \begin{bmatrix} 182 \\ 275 \end{bmatrix} \bmod 26$

$= \begin{bmatrix} 0 \\ 15 \end{bmatrix} = \begin{bmatrix} A \\ P \end{bmatrix}$ //

26) 182 (7
182
0 ← Rem

26) 275 (10
260
15 ← Rem

Same continue for all the values

$\therefore P = \begin{bmatrix} 14 \\ 17 \end{bmatrix}$ ans $\begin{bmatrix} A \\ D \end{bmatrix}$

$P = \begin{bmatrix} 19 \\ 4 \end{bmatrix}$ Ans $\begin{bmatrix} J \\ T \end{bmatrix}$

$P = \begin{bmatrix} 23 \\ 0 \end{bmatrix}$ Ans $\begin{bmatrix} F \\ T \end{bmatrix}$

$P = \begin{bmatrix} 12 \\ 15 \end{bmatrix}$ Ans $\begin{bmatrix} W \\ L \end{bmatrix}$

$P = \begin{bmatrix} 11 \\ 4 \end{bmatrix}$ Ans $\begin{bmatrix} F \\ J \end{bmatrix}$

DES :-

S-DES	DES
* 8 bit block	* 64 bit
* 10 bit key : 2*	* 56-bit key : 16*
* 8 bit round key	* 48 bit round keys
* IP : 8 bits	* IP : 64 bits
* 2 rounds	* 16 rounds
* 2 S-Boxes	* 8 S-boxes
* f-operate on 4 bits	* f-operation on 32 bits

Unit (1)

Block ciphers & Data encryption standard

* S-DES. (Simplified DES) Algorithm.

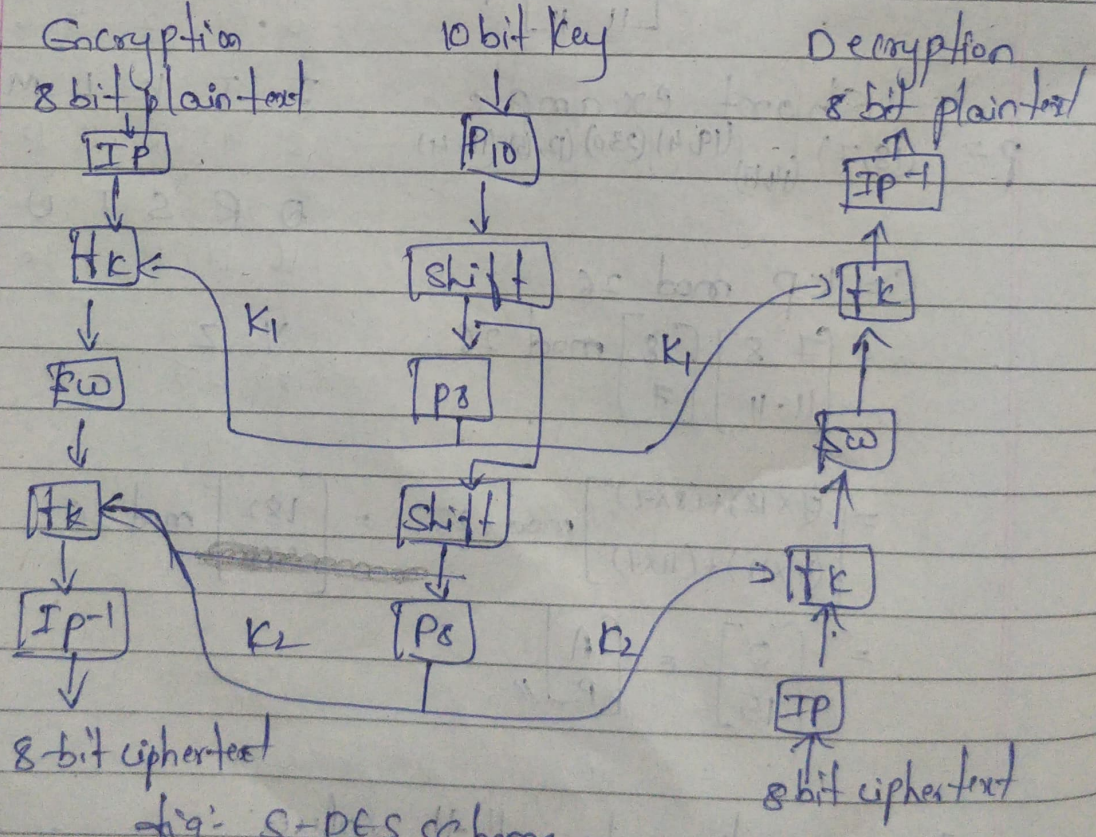
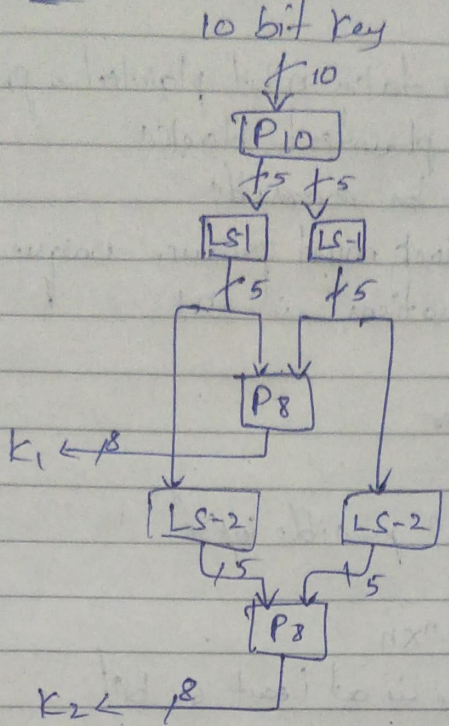


fig:- S-DES scheme

Key generation

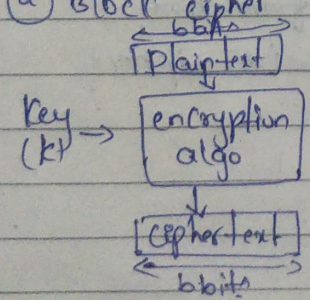


Block cipher

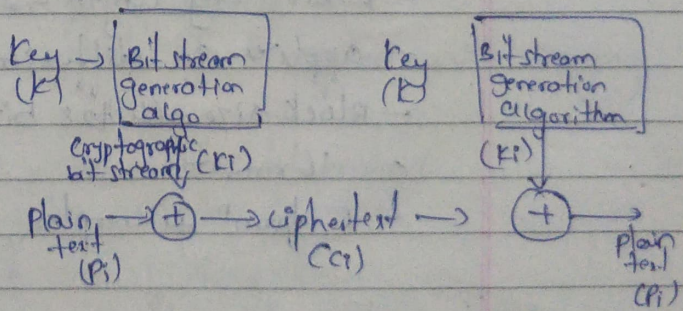
=> Encrypt data one block at a time.
=> block size 64-128 bits 128 bits
=> most algorithm based on a struc referred to as feistel block cipher.
=> Used in broader range of application

Block v/s stream cipher

(a) Block cipher



(b) Stream cipher



-> A block cipher is one in which a block of plaintext is treated as a whole & used to produce a ciphertext block of equal length.
=> A stream cipher is one that encrypts a digital data stream one bit (or) one byte at a time.

Block cipher principle

* n bit block cipher takes n bit plaintext & produces n bit ciphertext
* 2^n possible diff plaintext blocks
* encryption must be reversible
* each plaintext block must produce unique ciphertext block
* total transformations is $2^n!$

Ideal block cipher

* n bit input — 2^n possible input
* 2^n output states
* key length is $2^n \times n$
* Actual block size is at least 64 bit.

Confusion :- make relationship b/w ciphertext & key as complex as possible. even if attacker can find some statistical char of ciphertext, still hard to find key

Diffusion :- statistical nature of plaintext is reduced in ciphertext.

* feistel structure for block ciphers
  - n bit block length
  - k bit key length - 2^k transformations
  - applies on diffusion, confusion, ciphers
  - Block size - 64, 128 bit
  - 16 rounds
  - F fun^n is complex

DES (Data encryption standard)

* Symmetric block cipher (56 bit key, 64 input block / output)
* one of most used encryption system in world - 1977
* simplified DES (S-DES)

Developed NBS / NIST
design - IBM.

=> diff btw S-DES & DES.
=> problems:

Attacks on DES:

* Timing Attack → attacks actual implementation of cipher.
  → use fact that CPU can take varying time depending on value of inputs to it.
* Differential cryptanalysis
* Linear cryptanalysis

Strength of DES

* 56 bit keys have $2^{56} = 7.2 \times 10^{16}$ values
* brute force search looked hard
* advances have shown is possible
* still must be able to recognize plaintext
* forced to consider alternative to DES

=> Differential / Linear cryptanalysis

* one of the most significant recent advances in cryptanalysis
* powerful method to analyze block ciphers
* used to analyze most current block ciphers with varying degrees of success
* A statistical attack against feistel ciphers
* uses cipher structure not previously used
* it compares 2 related pairs of encryptions
* have some i/p diff giving some o/p diff with probability p.
* can infer subkey that was used in round

Linear
- another fairly recent development
- also a statistical method
- based on finding linear approximations
- must be iterated over rounds, with decreasing probabilities
- given linear eqn for key bits
- effectiveness given by : $(p - 1/2)$
- using large no of trial encryptions

DES Design principles

- as reported by Coppersmith in [Copp94]
- 7 Criteria for s-boxes provide for
  - non linearity - good confusion
  - resistance to diff cryptanalysis
- 3 criteria for permutation P provide for
  - increased diffusion

modes of operation:

- Block cipher encrypt fixed size blocks
- have block & stream modes
- to cover a wide variety of applications
- can be used with any block cipher
- NIST SP 800-38A define 5 modes
- need some way to en/decrypt arbitrary amounts of data in practice

Time table for 1st internal

monday	6/12/21	* Data Science * DST	— 9:30/10:30 Am 2/3 pm
Tuesday	7/12/21	*MAE * DL	— 9.30/10.30 Am 2/3 pm
wednesday	8/12/21	* IS	— 9.30/10.30 Am