Q.1 i) Solution :-
Without doing normalization, still we can see from $\phi_1(A,B)$, where A is Alice and B is Bob.

ii) The value of $\phi_1(a^0, b^0) = 30$ and $\phi_1(a_1, b_1) = 10$, so we can conclude that $\phi_1(a^0, b^0) > \phi_1(a_1, b_1)$, so
$\therefore$ there is high chance that Alice and Bob would like to disagree than to agree.

Q.2 i) A Markov network is a probabilistic graphical model that represents the relationships between variables using an undirected graph.

ii) In computer vision Markov networks are often used to model spatial dependencies and contextual relationships within images.

iii) For example, in image segmentation, each pixel or region can be treated as a node, and edges between nodes represent the dependency between neighbouring pixels.

iv) These connections allow Markov networks to enforce smoothness in label assignments, helping to produce coherent segments where nearby pixels tend to have similar labels.

Q.3 To check independencies, we prove that

$P(X_1 = x_1, X_3 = x_3) = P(X_1 = x_1) \cdot P(X_3 = x_3)$

But $P(X_1 = 0) = P(X_1 = 1) = 4/8 = 1/2$

and $P(X_3 = 0) = P(X_3 = 1) = 4/8 = 1/2$

$P(X_1 = 0, X_3 = 0) = 2/8 = 1/4$

~~$P(X_1 = 1, P$~~

$P(X_1 = 0) \cdot P(X_3 = 0) = 1/2 \cdot 1/2 = 1/4$

$P(X_1 = 0, X_3 = 1) = 2/8 = 1/4$

$P(X_1 = 0) \cdot P(X_3 = 1) = 1/2 \cdot 1/2 = 1/4$

$P(X_1 = 1, X_3 = 0) = 2/8 = 1/4$

$P(X_1 = 1) \cdot P(X_3 = 0) = 1/2 \cdot 1/2 = 1/4$

$P(X_1 = 1, X_3 = 1) = 2/8 = 1/4$

$P(X_1 = 1) \cdot P(X_3 = 1) = 1/2 \cdot 1/2 = 1/4$

$\because P(X_1, X_3) = P(X_1) \cdot P(X_3)$ for all combination of X_1 and X_3, we can say that X_1 is independent of X_3.

Q.4 Given that probability ~~or~~ is only given for where all is 0 or where all is 1.

like $P(A=0, B=0, C=0, D=0) = 0.5$

or $P(A=1, B=1, C=1, D=1) = 0.5$

Markov blanket is a set of node that define condition independence of that node with remaining conditioning markov blanket

$\because$ only knowing one variable we can define all other variable this is perfectly correlated if we know $A=1$ then all B, C, D is 1.

1] ~~pef~~ perfect correlation

if $A=1$ all other are 1

if $A=0$ all other are 0

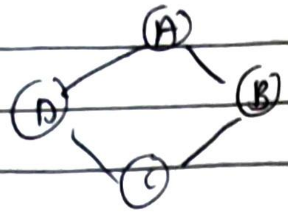
2] Markov blanket

We take one node, A, its markov blanket will be B, C, D, i.e. all other nodes. Markov blanket condition is it give conditional independency given markov blanket in this case, when making markow blanket for any of node it is all remaining node so it is trivially independant to its subset.

Overall, distribution is highly degenerate as all variable are perfectly correlated.

Thus, the markov blanket condition is ~~stat~~ satisfied for all nodes in this network

Q.5



We have four students A, B, C, D
we have relationship as (A, B), (B, C), (C, D), (D, A)

clique potential :-
The factor for each pair is given as $\phi_1(A, B)$
$\phi_2(B, C)$, $\phi_3(C, D)$, $\phi_4(D, A)$

Logarithmic ~~Re~~ Representation

$\phi_1(A, B)$

a_0	b_0	log (30)
a_0	b_1	log (5)
a_1	b_0	log 1 = 0
a_1	b_1	~~tg~~ log (100)

$\phi_2(B, C)$

b_0	c_0	log (100)
b_0	c_1	log (1) = 0
b_1	c_0	0
b_1	c_1	log (100)

$\phi_3(C, D)$

c_0	d_0	0
c_0	d_1	log (100)
c_1	d_0	log (100)
c_1	d_1	log (1) = 0

$\phi_4(D, A)$

d_0	a_0	log (100)
d_0	a_1	0
d_1	a_0	0
d_1	a_1	log (100)

log representation of any clique potential is
log (ϕ_1(A, C)) of ϕ_1(A, C)

Q.6. For calculating canonical energy function for a clique D is given as:-

$$\epsilon^*_D(d) = \sum_{Z \subseteq D} (-1)^{|D-Z|} \, l(d_Z, \epsilon^*_{-Z})$$

where sum is all subsets of D.

For $\epsilon_{AB}(d_{AB})$
$d_{AB} = a^0, b^0$ taking all subset $\{A,B\}, \{A\}, \{B\}, \{\phi\}$

$\epsilon_{AB}(a^0, b^0) = l(a^0, b^0) - l(a^0, \epsilon_B) - l(\epsilon_A, b^0) + l(\epsilon_\phi)$
where ϵ_B, ϵ_A and ϵ_ϕ are default values

$\therefore \epsilon_{AB}(a^0, b^0) = \ln(80) - \ln(30) - \ln(30) + \ln(30)$
$= 0$

$\epsilon_{AB}(a^0, b^1) = 0$
$\epsilon_{AB}(a^1, b^0) = 0$
$\epsilon_{AB}(a^1, b^1) = 4.09$

For $\epsilon_{BC}(d_{BC})$ subsets are $\{B, C\}, \{B\}, \{C\}, \{\phi\}$

$\epsilon_{BC}(b^0, c^0) = 0$

$\epsilon_{BC}(b^0, c^1) = 0$
$\epsilon_{BC}(b^1, c^0) = 0$
$\epsilon_{BC}(b^1, c^1) = 9.24$

For ϵ_{CD}

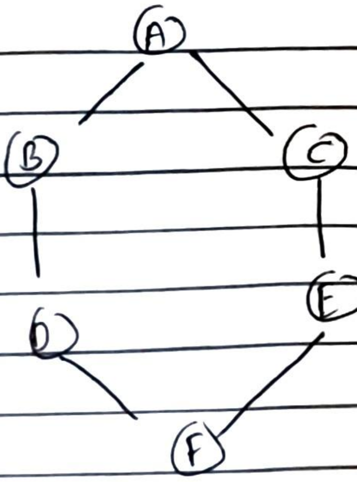
$\epsilon_{CD}(c^0, d^0) = 0$
$\epsilon_{CD}(c^0, d^1) = 0$
$\epsilon_{CD}(c^1, d^0) = 0$
$\epsilon_{CD}(c^1, d^1) = -9.21$

Q.6 For ϵ_{DA}

$$\epsilon_{DA}(d^0, a^0) = 0$$
$$\epsilon_{DA}(d^0, a^1) = 0$$
$$\epsilon_{DA}(d^1, a^0) = 0$$
$$\epsilon_{DA}(d^1, a^1) = 9.21$$

For $\epsilon_A(A)$	$\epsilon_B(B)$	$\epsilon_C(C)$	$\epsilon_D(D)$	$\epsilon(\emptyset)$
$a_0 = 0$	$b_0 = 0$	$c_0 = 0$	$d^0 = 0$	-3.18
$a_1 = -8.01$	$b_1 = -6.4$	$c_1 = 0$	$d^1 = 0$	

Q.7

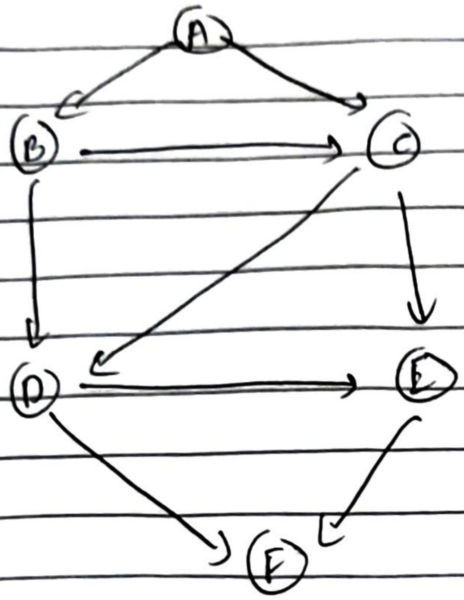


By enumerating the nodes in x in some ordering and define the parent set for each one in turn according to the independencies in distribution

Toking enumeration as A, B, C, D, E, F

- We can see that A is parent of B
- A is parent of C, C is not independant of B
- B is parent of D and C is not independant of D given B so C is also parent.
- For F, C is direct parent of F, B is independant of F given C and D is not independant of F.

overall



Bayesian network for the

Q.8 To represent vehicle localization task ~~and~~ using a dynamic Bayesian network (DBN) key variables are location, velocity, weather, sensor failure and observation.

From given question we can see that,

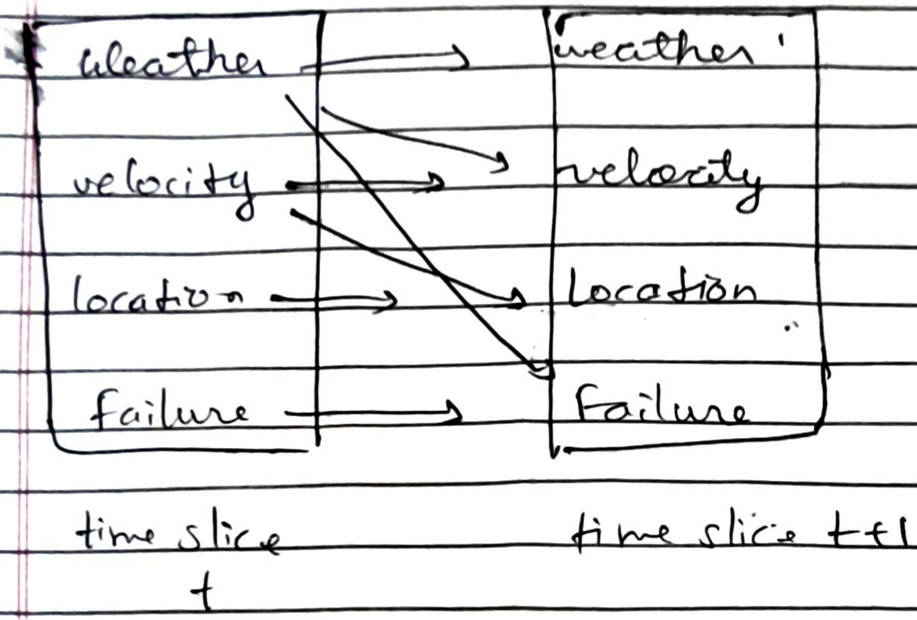
- weather (t+1) depends on ~~and~~ weather and velocity of (t).

- velocity at (t+1) depends on weather and velocity of (t)

- location at (t+1) depends on previous location and velocity.

- ~~Fa~~ Failure at (t+1) is as per previous failure and weather.

Q.8 and at time t location depends on velocity and observation t on location and failure.



Q.9 we can see that at time t+1 e is depended on value of failure and location so if failure malfunction then z i.e sensor failure.

- it indicates that sensor has malfunctioned because of inaccurate reading or missing value.
- obs will reflect this error.

From diagram we can see that value of failure affects obs and because of that failure it can be detected.

R.E → DFA, NFA

Q.10 HMM's are probabilistic graphical models that use hidden states to represent different genomic regions in DNA sequence. The model assigns probabilities to transitions between these states, allowing it to predict likely gene regions based on observed nucleotide sequence.

Phylo-HMM's are an extension of HMMs that incorporate evolutionary relationships using multiple aligned sequences from different species. They combine phylogenetic models with HMM states, improving sensitivity in detecting conserved regions by considering evolutionary conservation.

Q.11 In PGM, ~~the~~ the state observation model defines the the link between hidden states and observed data.

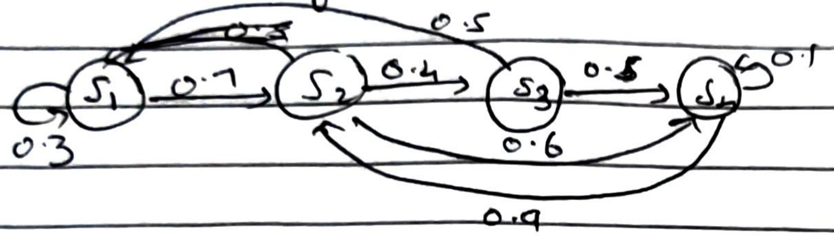
Key elements:-

i) Hidden states - Represent underlying variables or conditions that we want to infer but can't directly observe

ii) Observation: The actual data or signal we observe

iii) Emission probabilities: These probabilities define how likely each observation is given & each hidden state, helping us infer hidden states from the ~~obs~~ observed data.

Q.11 The observation thus play a crucial role in connecting hidden state to observable outcomes, allowing the model to make inference about unobserved factor based on observed data

Q.12 

$\therefore$ Transition matrix $T = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0 & 0.4 & 0.4 & 0.6 \\ 0 & 0.6 & 0.5 & 0 \\ 0.9 & 0 & 0 & 0.1 \end{bmatrix}$

Transition matrix $T = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.4 & 0.6 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0.9 & 0 & 0.1 \end{bmatrix}$

Q.13 A Hidden markov model is a statistical model used in speech recognition to represent sequence of speech sounds (like phonemes).

It consists of :-

i) States : Represent speech units

ii) Observations: Acoustic features extracted from speech

iii) Transitions: Probabilities of moving between states.

iv) Emission: Likelihood of an observation given a state

v) Start / End states: Marks the beginning and end of speech

In training HMM's learn transition and emission probabilities from speech data. During recognition, they decode the most probable sequence of states from a sequence of acoustic features, often using the 'Viterbi' algorithm.

Q.14 A linear dynamical system is a mathematical model used to describe systems that evolve over time according to linear equations. It is widely used in fields like control theory, signal processing and speech recognition.

i) State vector: Represents system's internal state at any given time

ii) State transition: Describes how the state evolves over time, typically expressed as a linear equation.

$$x_{t+1} = Ax_t + Bu_t$$

iv) Output equation: describes how the system state is related to its observed output, often written as

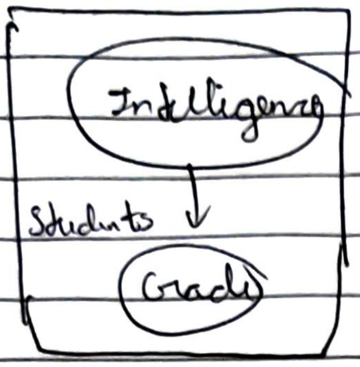
$$y_t = Cx_t + Du_t$$

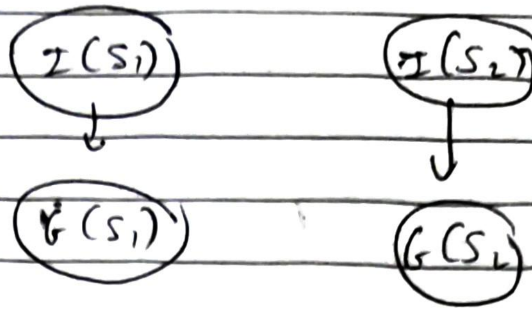
Q.14 4. Assumptions: The system is assumed to be linear, meaning that the relationship between state, input, and output is better.

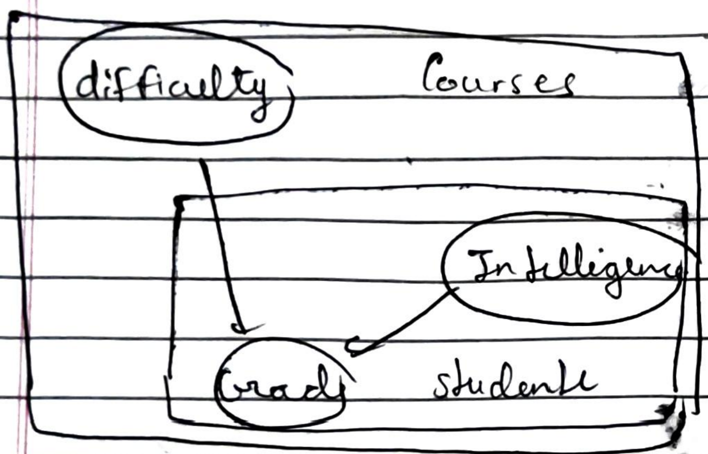
Applications:

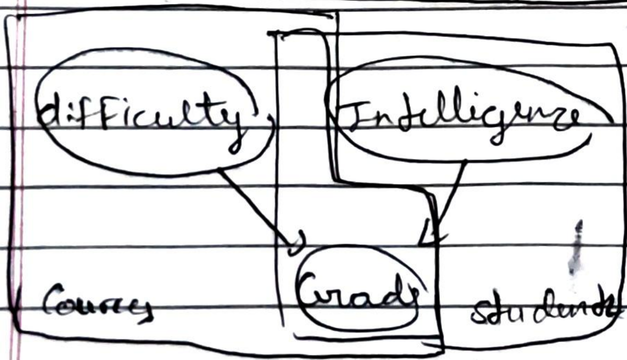
LDS models are used in areas like Kalman filtering, speech recognition, and robotics for tasks such as state estimation, prediction and control.

Q.15


Intelligence
Students
Grade


I(S1)
G(S1)
I(S2)
G(S2)


difficulty
Courses
Intelligence
Grade
students


difficulty
Intelligence
Grade
Courses
students