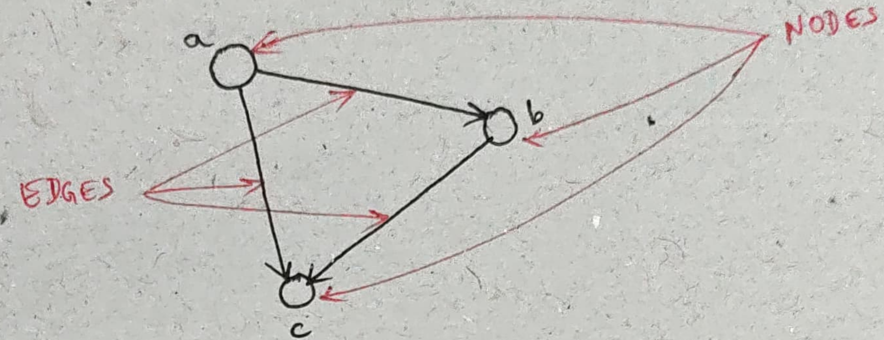
Week 8 : Session 1.

## 12. Probabilistic Graphical Models (PGMs)

Probabilistic Graphical Models are VISUAL REPRESENTATION of probability distributions. PGM is a fusion of Probability and Graph theories.

- Graph consists of nodes and edges.




- Each node represents a Random variable.
- Edge represents a Probabilistic Relationship.

Bayesian Networks (BNs)

- BNs are subset of PGMs
- These are Directed Graphical Models
- The arrows show Directionality.
- Arrows capture Causal Relationships between Random Variables.

---

Consider Random variables $x_1, x_2, \ldots, x_k$

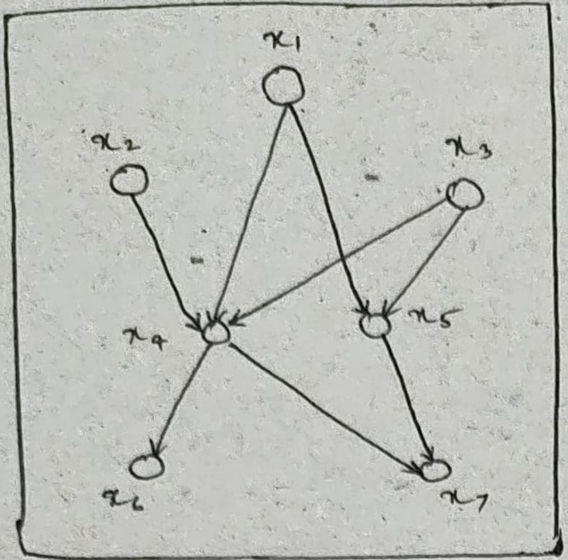
What is the Joint PDF of these Random variables ?

The Joint PDF can be simplified as

$$p(x_1, x_2, \ldots, x_k)$$
$$= \underbrace{p(x_1) \times p(x_2 | x_1) \times p(x_3 | x_1, x_2) \times \ldots \times p(x_k | x_1, x_2, \ldots, x_{k-1})}_{\text{Chain Rule}}.$$

---

Example BN.




The Joint PDF for this example BN can be simplified as follows.

$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) =$$
$$p(x_1) \times p(x_2|x_1) \times p(x_3|x_1, x_2)$$
$$\times p(x_4|x_1, x_2, x_3)$$
$$\times p(x_5|x_1, x_2, x_3, x_4)$$
$$\times p(x_6|x_1, x_2, x_3, x_4, x_5)$$
$$\times p(x_7|x_1, x_2, x_3, x_4, x_5, x_6)$$

The following property holds for BN.

(ii) $p(x_k|x_1, x_2, \ldots, x_{k-1}) = p(x_k | P_k)$

where, $P_k \rightarrow$ Set of Parents of $x_k$

Example: $P_5 = \{x_1, x_3\} \rightarrow$ Parents of $x_5$

$P_7 = \{x_4, x_5\} \rightarrow$ Parents of $x_7$

Given Parents $P_k$, $x_k$ is CONDITIONALLY INDEPENDENT of others!

Therefore, for this graph

$$p(x_2|x_1) = p(x_2|P_2) = p(x_2).$$
$$p(x_5|x_1, x_2, x_3, x_4) = p(x_5|P_5) = p(x_5|x_1, x_3)$$

Now, the Joint PDF can be simplified as

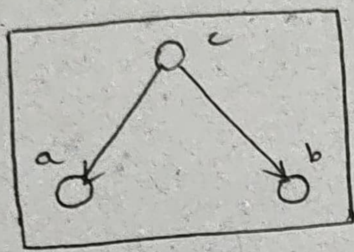
$$p(x_1, x_2, x_3, x_4, x_5, x_6, x_7) =$$
$$p(x_1) \times p(x_2) \times p(x_3)$$
$$\times p(x_4|x_1, x_2, x_3) \times p(x_5|x_1, x_3)$$
$$\times p(x_6|x_4) \times p(x_7|x_4, x_5).$$

In general,

$$p(x_1, x_2, \ldots, x_k) = \prod_{k=1}^{K} \underbrace{p(x_k \mid P_k)}_{\uparrow \text{ Product of Conditional Factors.}}$$

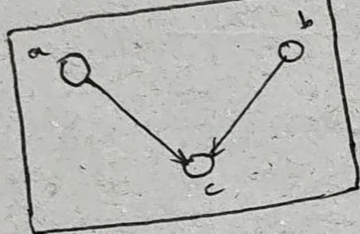
This is the principle of Bayesian Network.

<u>Simple Examples</u>.

① 

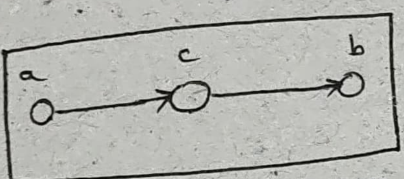

Joint PDF is given as

$$p(a, b, c) = p(c) \times p(a \mid c) \times p(b \mid c).$$

② 


Joint PDF is given as

$$p(a, b, c) = p(a) \times p(b) \times p(c \mid a, b)$$

③ 


Joint PDF is given as

$$p(a, b, c) = p(a) \times p(c \mid a) \times p(b \mid c)$$

This is a Markov chain.

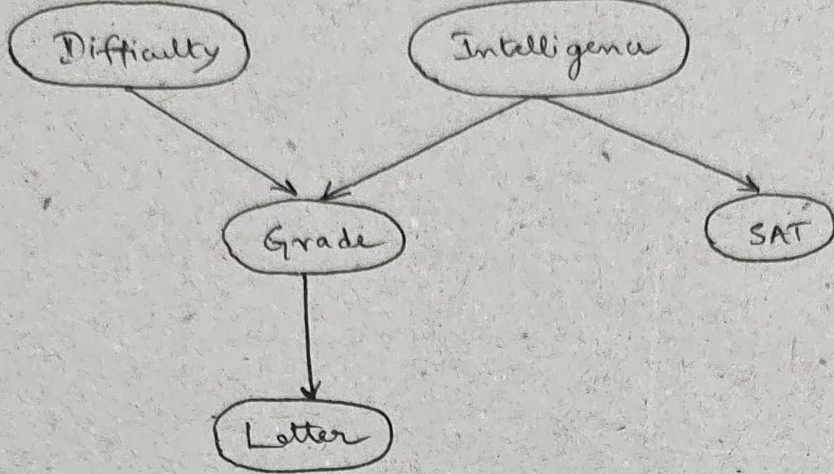
Markov chain is characterized as

$$p(x_k \mid \underbrace{x_1, x_2, \ldots, x_{k-1}}_{\text{History}}) = p(x_k \mid \underbrace{x_{k-1}}_{\uparrow \text{ Immediate past}})$$

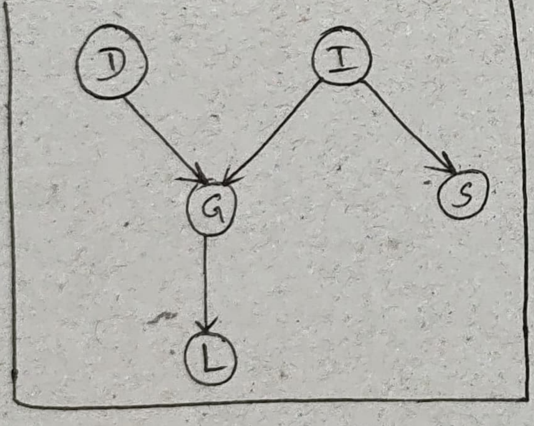
Week 8 : Session 2

## BN Example

Consider the BN example shown below.




- Student has certain Intelligence (I) level and takes a course of certain Difficulty (D) level.
- Grade (G) in this course is determined by D and I.
- Quality of Letter (L) determined by G.
- SAT (S) score exclusively determined by I.

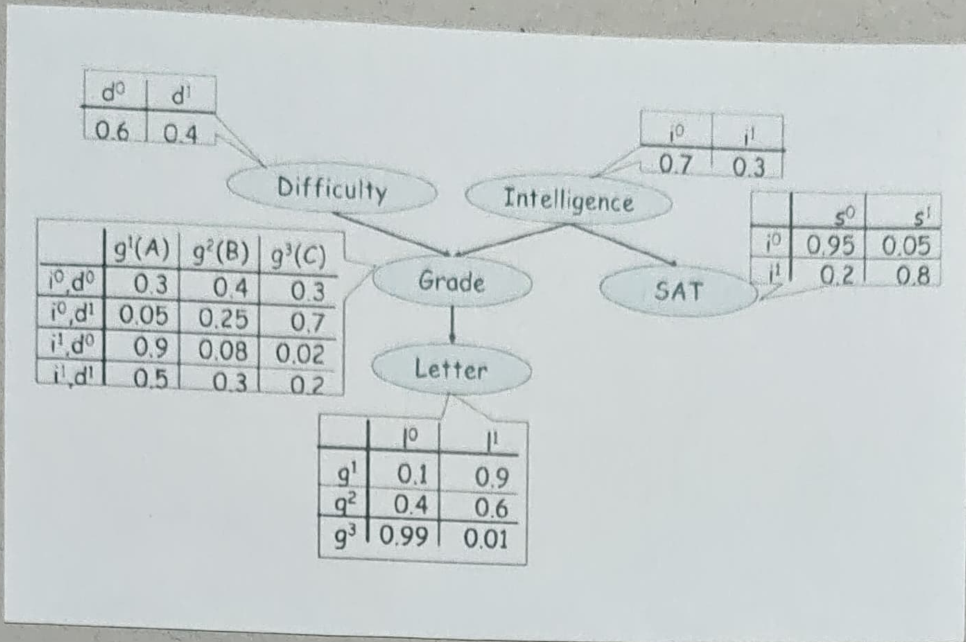



The Joint PDF can be evaluated as

$$p(D, I, G, L, S) = p(D) \times p(I) \times p(G \mid D, I) \times p(L \mid G) \times p(S \mid I)$$

So, one can write very efficiently the Joint PDF as product of factors.

Joint PDF can be represented as CPDs. (conditional probability distributions) which is very compact.




In this BN example, let us characterize each Random Variable as follows.

<u>Difficulty</u>

- $d^0$ = Easy
- $d^1$ = Difficult

<u>Intelligence</u>

- $i^0$ = Low intelligence
- $i^1$ = High intelligence

<u>Grade</u>

- $g^1$ = A Grade
- $g^2$ = B Grade
- $g^3$ = C Grade

<u>Letter</u>

- $l^0$ = Poor letter
- $l^1$ = Good letter

<u>SAT</u>

- $s^0$ = Poor SAT score
- $s^1$ = Good SAT score

Each entry is Conditional Probability of Column Element given row Element.

Say, $p(g^3 | i^0, d^1) = 0.7$

$p(s^1 | i^0) = 0.05$

Thus, the CPDs enable very compact representation!

---

The Joint probability can be evaluated as product of factors based on the Bayesian Network.

$$p(d^0, i^1, g^3, s^1, l^1) = p(d^0) \times p(i^1) \times p(g^3 | d^0, i^1) \times p(l^1 | g^3) \times p(s^1 | i^1)$$

$$= 0.6 \times 0.3 \times 0.02 \times 0.01 \times 0.8$$

$$= 0.0000288.$$

---

How to evaluate $p(i^1 | s^1, l^0)$ ?

$$p(i^1 | s^1, l^0) = \frac{p(i^1, s^1, l^0)}{p(s^1, l^0)}$$

$p(A|B) = \frac{P(A \cap B)}{P(B)}$

where,

$A = \{i^1\}$

$B = \{s^1, l^0\}$

Now, $p(i^1, s^1, l^0) = \sum_{D,G} p(D, i^1, G, l^0, s^1)$

$$p(s^1, l^0) = \sum_{I,D,G} p(D, I, G, l^0, s^1)$$

---

$$\sum_{D,G} p(D, i^1, G, l^0, s^1) = p(d^0, i^1, g^1, l^0, s^1) + p(d^0, i^1, g^2, l^0, s^1) + p(d^0, i^1, g^3, l^0, s^1) + p(d^1, i^1, g^1, l^0, s^1) + p(d^1, i^1, g^2, l^0, s^1) + p(d^1, i^1, g^3, l^0, s^1)$$

Week 8 : Session 3.

## Causal Reasoning.

- Bayesian Network represents the Causal probabilistic relationships.
- The Difficulty, Intelligence together explain the Grade, Letter.
- We can use that to analyze the Cause / Evidence relationship.
- Basically, it shows the effect of the cause on the Evidence.
- "Given the cause, what is the effect?"
- Cause (Difficulty / Intelligence) explain Evidence (Letter).
- Hence the name "Causal Reasoning".

Say, the Probability of getting good letter,

$$p(l^1) \approx 0.5.$$

But, if the Student's intelligence is Low, then the probability of getting good letter decreases.

$$p(l^1 | i^0) \approx 0.39$$

And now, in addition, if the course is easy, then the probability of getting good letter increases.

$$p(l^1 | i^0, d^0) \approx 0.51$$

---

## Evidential Reasoning

- This is other way around. (ie)

"Given the Evidence, what is the cause?"

- Used in Medical Diagnosis.

(ie) The doctor observes the symptoms and try to decide / estimate the disease.

- The Evidence (Grade) explains the causes (Difficulty / Intelligence)

Say, the Probability of Difficult course

$$p(d^1) = 0.4$$

And, the probability of High level intelligence

$$p(i^1) = 0.3$$

Now, the above probabilities, given the grade is poor increases.

$$p(d^1 | g^3) = 0.63$$

$$p(i^1 | g^3) = 0.08$$

---

## Intercausal Reasoning

- Here, One cause explains the other cause.

say, Difficulty explains Intelligence.

- Example: Student gets $g^3$ = C Grade.
- The Bayesian Interface activates V shaped structure.
- Originally, the Difficulty and Intelligence is independent.
- Here, the difficulty of the course conveys something about the intelligence of student, to the Grade.

- That is why, Grade is needed to activate this V structure only when the grade is known.
(i) Grade has to be known / Some descendant of Grade has to be known (i.e. Letter)

Example.

(i) $p(i^1 | g^3) = 0.08$

$p(i^1 | g^3, d^1) \approx 0.11 \uparrow$

(ii) $p(i^1 | g^2) = 0.175$

$p(i^1 | g^2, d^1) \approx 0.34 \uparrow$

---

Conclusion:

Bayesian Networks are powerful tools for

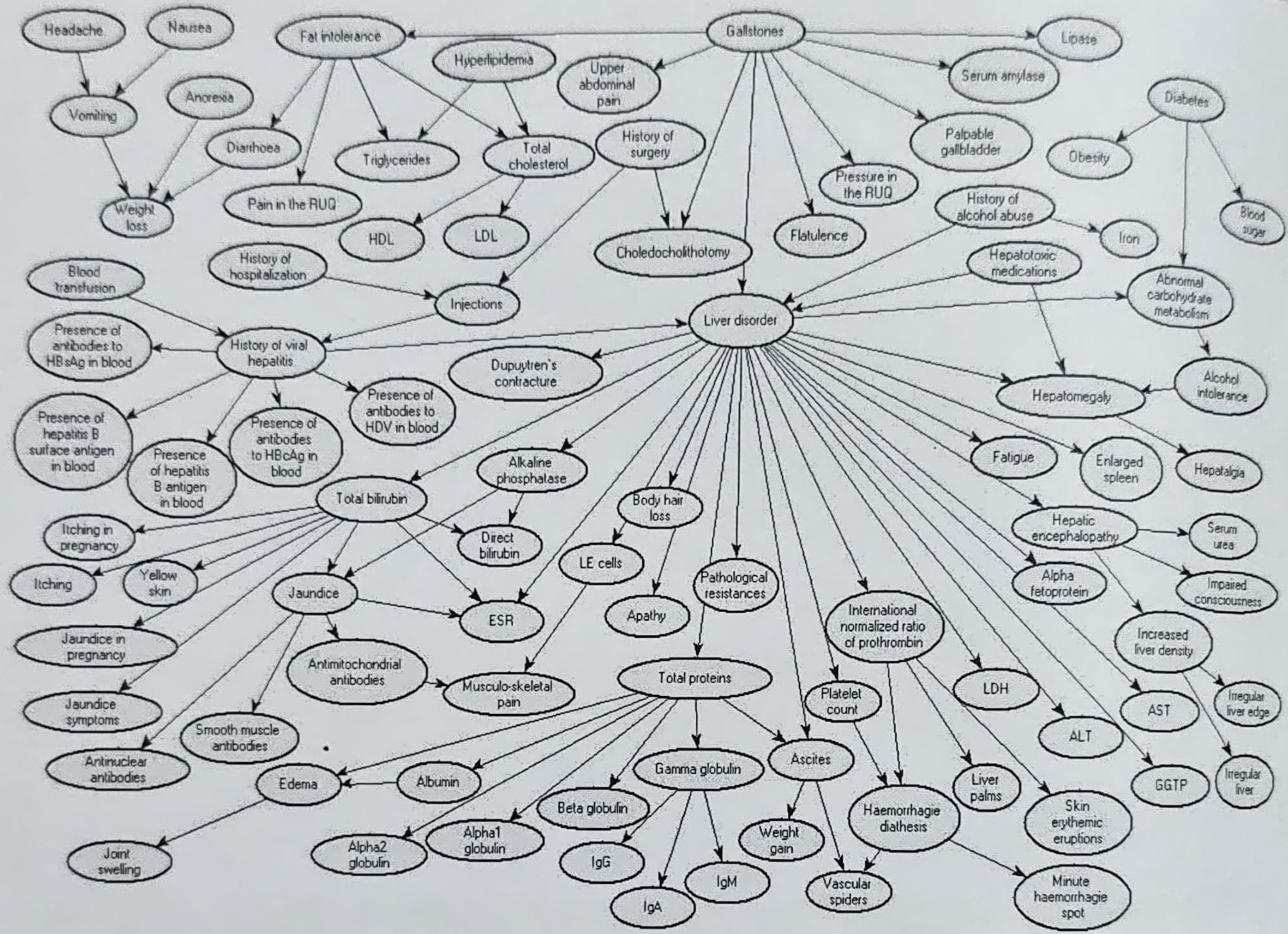
(i) Reasoning
- Causal
- Evidential
- Inter causal

and

(ii) Inference.

Bayesian Networks are used in several applications such as Medical Diagnosis.

Fig: BN for diagnosis of Liver disorders.


Headache
Nausea
Fat intolerance
Gallstones
Lipase
Hyperlipidemia
Upper abdominal pain
Serum amylase
Vomiting
Anorexia
Diabetes
Diarrhoea
Triglycerides
Total cholesterol
History of surgery
Palpable gallbladder
Obesity
Pain in the RUQ
Pressure in the RUQ
Weight loss
HDL
LDL
Choledocholithotomy
Flatulence
History of alcohol abuse
Iron
Blood sugar
History of hospitalization
Hepatotoxic medications
Blood transfusion
Injections
Liver disorder
Abnormal carbohydrate metabolism
Presence of antibodies to HBsAg in blood
History of viral hepatitis
Dupuytren`s contracture
Alcohol intolerance
Hepatomegaly
Presence of hepatitis B surface antigen in blood
Presence of antibodies to HDV in blood
Presence of antibodies to HBcAg in blood
Presence of hepatitis B antigen in blood
Alkaline phosphatase
Fatigue
Enlarged spleen
Hepatalgia
Total bilirubin
Body hair loss
Hepatic encephalopathy
Serum urea
Itching in pregnancy
Direct bilirubin
LE cells
Pathological resistances
Alpha fetoprotein
Impaired consciousness
Itching
Yellow skin
Jaundice
ESR
Apathy
International normalized ratio of prothrombin
Increased liver density
Jaundice in pregnancy
Antimitochondrial antibodies
Musculo-skeletal pain
Total proteins
Platelet count
LDH
AST
Irregular liver edge
Jaundice symptoms
Smooth muscle antibodies
ALT
Antinuclear antibodies
Edema
Albumin
Gamma globulin
Ascites
Liver palms
GGTP
Irregular liver
Beta globulin
Haemorrhagie diathesis
Skin erythemic eruptions
Alpha2 globulin
Alpha1 globulin
IgG
Weight gain
Joint swelling
IgM
Vascular spiders
Minute haemorrhagie spot
IgA