BAYESIAN LEARNING

— Bayesian reasoning provides a probabilistic approch to inference.

— Bayesian learning methods are relant to our study of machine learning for two different reasons.

— First, Bayesian learning algorithms that calculate explicit probabilities for hypothesis, such as naive Bayes classifier.

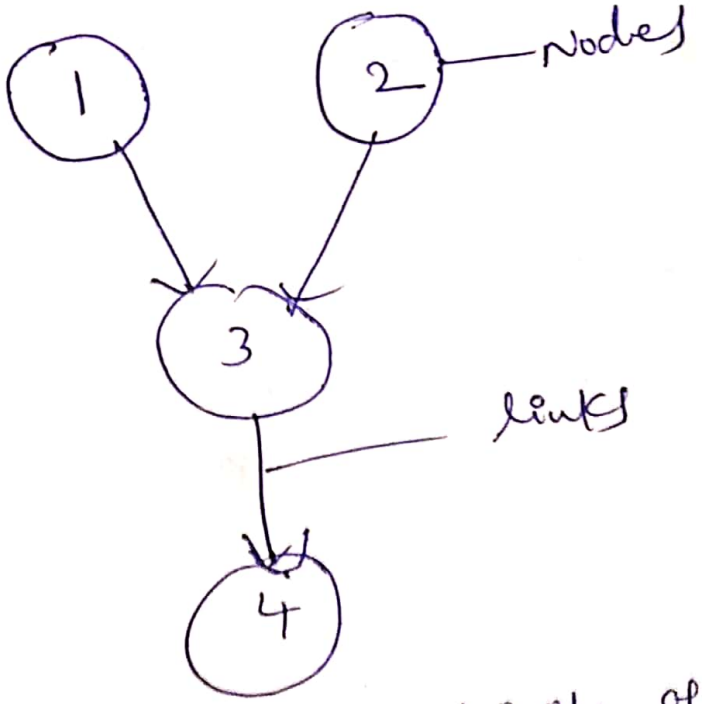
— Second Bayesian methods are important to our study of machine learning is that they provide useful perspective for understanding many learing algorithm that do not explicitly manipulate probabilities.

Ex: FIND-S, Candidate elimination algo.

Bayesian Network ?

A Bayesian Network falls under the cate
of probablistic Graphical modeling (PGM)
technique that is used to compute
uncertainties by using the concept of probability

What is Directed Acyclic graphs?



A DAG models the uncertainty of an event
occuring based on the Conditional probability
Distribution (CDP) of each random variable.

Joint Probability is a measure of two events happening at the same time i.e., P(A and B). The probability of the intersection of A and B may be written P(A ∩ B).

Conditional Probability :- of an event B is the probability that the event will occur given that an event A has already occured.

P(B/A) : probability of event B occuring, given that event A occur.

If A and B are dependent events

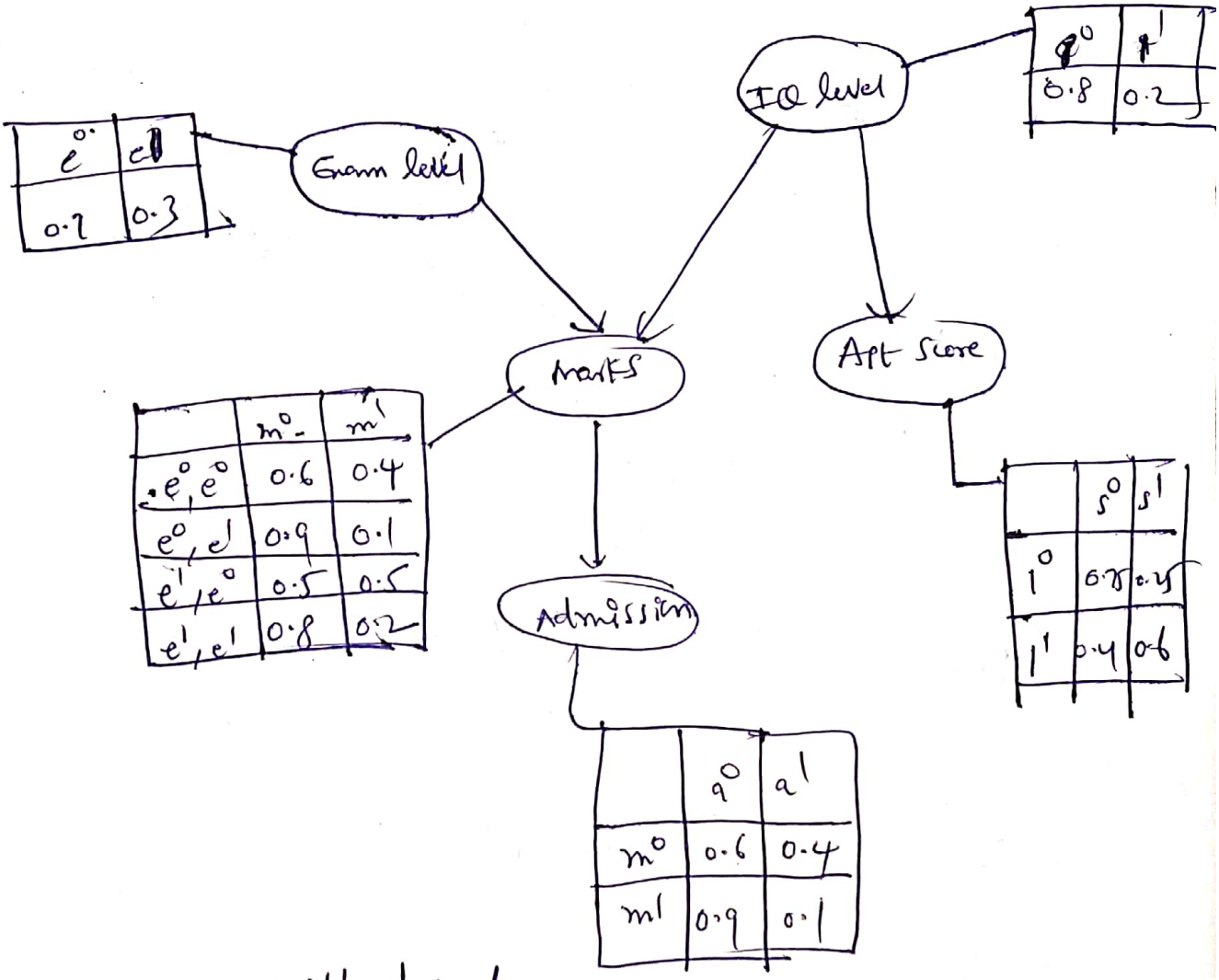
$$P(B/A) = \frac{P(A \text{ and } B)}{P(A)}$$

If A and B are independent events :

$$P(B/A) = P(B)$$

BAYESIAN NETWORKS EXAMPLE

Create a Bayesian Network that will model the marks (m) of a student on his examination.



The marks will depend on:

- Exam level (e) : (difficult, easy)
- IQ of the student (i) (high, low)
- Marks → admitted (a) to a university.
- The IQ → aptitude score (s) of the student.

Factorising Joint Probability Distibution:

$$P(a, m, i, e, s) = P(a/m)\, P(m/i, e)\, P(i)\, P(e)\, P(s/i)$$

- $P(a/m)$: cp of Studet admit → marks.
- $P(m/i, e)$: cp of students marks → (IQ & enam level)
- $P(i)$: probability → IQ level
- $P(e)$: probability → enam level
- $P(s/i)$: cp of aptitude score → IQ level.

The probability of a random variable depends on his parents. Therefore we can formulate Bayesian networks of:

$$P(X_1 \ldots X_n) = \prod_{i=1}^{n} P(x_i \mid \text{parents}(x_i))$$

↓

random variable depends on point probability.

— Applied in predictive modeling and descriptive analysis.