unit -02 : probabilistic Reasoning.

01) Acting under uncertainty:

* uncertainty representation using first order logic and propositional logic with certainty, which means we were sure about the predicates.

* With this knowledge representation, we might write A → B, which means if A is true then B is true.

* But consider a situation where we not sure about wheather a A is true or not then we cannot express this statement.

* This situation is called uncertainty.

causes of uncertainty:

uncertainty occur in the real world.

01) Information occured from unreliable sources.

02) Experimental Errors.

03) Equipment fault

04) Temperature variation

05) climate change

probabilistic Reasoning:

probabilistic reasoning is a way of knowledge representation where we apply the concept of probability to indicate the uncertainty in knowledge.

* We use probability in probabilistic reasoning because it provides a way to handle the uncertainty that is the result of someone's laziness and ignorance.

* In the real worlds, there are lots of scenarios, where the certainty of something is not confirmed, such as "It will rain today", "behaviour of someone for some situation", "A match between two teams or two players".

* These are probable sentence for which we can assume that it will happens but not sure about it, so here we use probabilistic reasoning.

Need of probabilistic reasoning in AI:

* When there are unpredictable outcomes
* When specifications or possibilities of predicates becomes too large to handle.
* When an unknown error occur during an experiment.

There are two ways to solve problem

i) Baye's rule
ii) Bayesian statistics.

Lets understand some common terms :

01) Probability
02) Sample space
03) Random variable
04) Prior probability
05) posterior probability
06) conditional probability
07) Event

probability :

probability can be defined as a chance that an uncertain event will occur

It is the numerical measure of the likelihood than an event occur.

The value of probability always remain between 0 and 1.

$0 \leq P(A) \leq 1$, where $P(A)$ is the probability of event A.

$P(A) = 0$, indicates total uncertainty in an event A.

$P(A) = 1$, indicates total certainty in an event A.

$$\boxed{\text{probability of occurance} = \frac{\text{Number of desired outcomes}}{\text{Total number of outcomes}}}$$

$P(\neg A)$ = probability of a not happening event.

$P(\neg A) + P(A) = 1$

Event :

Each possible outcomes of a variable is called an event.

Sample space:

The collection of all possible events is called Sample space.

Random variable:

Random variable are used to represent the events and object in the real world.

prior probability:

The prior probability of an event is probability computed before observing new information.

posterior probability:

* The probability that is calculated after all evidence or information has taken into account.
* It is a combination of prior probability and new information

conditional probability:

conditional probability is a probability of occuring an event when another event has already happened.

Let's suppose we want to calculate the event A when event B has already occured, "the probability of A under the condition of B" it can be written as

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

where, $P(A \cap B)$ = joint probability of A and B

$P(B)$ = Marginal probability of B.

If the probability of A is given and we need to find the probability of B, then it will be given us

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

Example :

In a class there are 70 % of the student who like English and 40 % students who likes English and mathematics and then what is the percentarge of student those who like English also like mathematics

Soln

Let A is an event that a student likes mathematics

B is an event that a student like English

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.4}{0.7} = 57\%$$

Hence, 57% are the students who like English also like mathematics.

Basic probability Notation

⇒ Logic Language with additional expressivness.

Basic probability Notation are :

1. Preposition.
2. Atomic Events.
3. Unconditional probability.
4. conditional probability
5. Inference using full jaint Distribution.
6. Independence.
7. Bay's Rule.

Preposition

* Preposition are attached with the degree of belief.
* complex preposition can be formed using standard logical connectivity.

ex:

$$[(cavity = true) \wedge (Toothache = False) \;\&\; (cavity \wedge \neg Toothache)]$$

Random variable (3)

⇒ Boolean Variable (Boolean Value)

ex: True

⇒ Discrete Variable

ex: countable.

⇒ continuous Random Variable

ex: Real Number.

$x \leq 4.1$

Atomic Events

↳ complete specification of state.

→ Agent is uncertain

ex:

Cavity = False ∧ Toothache = True

cavity = False ∧ Toothache = False

cavity = True ∧ Toothache = False

Cavity = True ∧ Toothache = True.

Properties:

① Atomic Events are mutually Exclusive.

② Events is exhaustive (Atleast one equivalent to True)

③ Atomic Events entails True or False.

④ Disjunction of atomic Events } = Logically equivalent to true.

③ Unconditional (or) Prior (or) Joint probability

⇒ Preposition in absence of any other information.

⇒ All Random Variables combination

ex:

$P(a_1, a_2)$

P (Weather, Cavity)

↓

Represented as

P (4 × 2)

↙ ↘

Weather Cavity.

④ conditional probability

⇒ concering previously unknown Random variable.

⇒ prior probability are not used.

conditional probability can be defined interms of unconditional probability.

$$P(a/b) = \frac{P(a \wedge b)}{P(b)}$$

Product Rule

$$P(a \wedge b) = P(a/b)\, P(b)$$

ex:

$a \wedge b = \frac{20}{100}$

$\boxed{a \wedge b = 0.2}$

$P(b) = 30/100$

$\boxed{P(b) = 0.3}$

$$P(a/b) = \frac{P(a \wedge b)}{P(b)} = \frac{0.2}{0.3}$$

$\boxed{P(a/b) = 0.61}$

Inference Using Full Joint Distribution

⇒ Computation from observed evidence of posterior probabilities.

⇒ Distribution over some variables (or) single variables called marginal probability.

ex:

	Toothache		¬Toothache	
	catch	¬catch	catch	¬catch
Cavity	0.108	0.012	0.072	0.008
¬Cavity	0.016	0.064	0.144	0.576

$$P(\text{cavity} \mid \text{Toothache}) = \frac{P(\text{cavity} \times \text{Toothache})}{P(\text{Toothache})}$$

$$= \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

$$\boxed{P(\text{cavity} \mid \text{Toothache}) = 0.6}$$

$$P(\text{cavity} \mid \text{Toothache}) = \frac{P[\neg \text{Cavity} \wedge \text{Toothache})}{P(\text{Toothache})}$$

$$= \frac{0.016 + 0.064}{0.012 + 0.108 + 0.016 + 0.064}$$

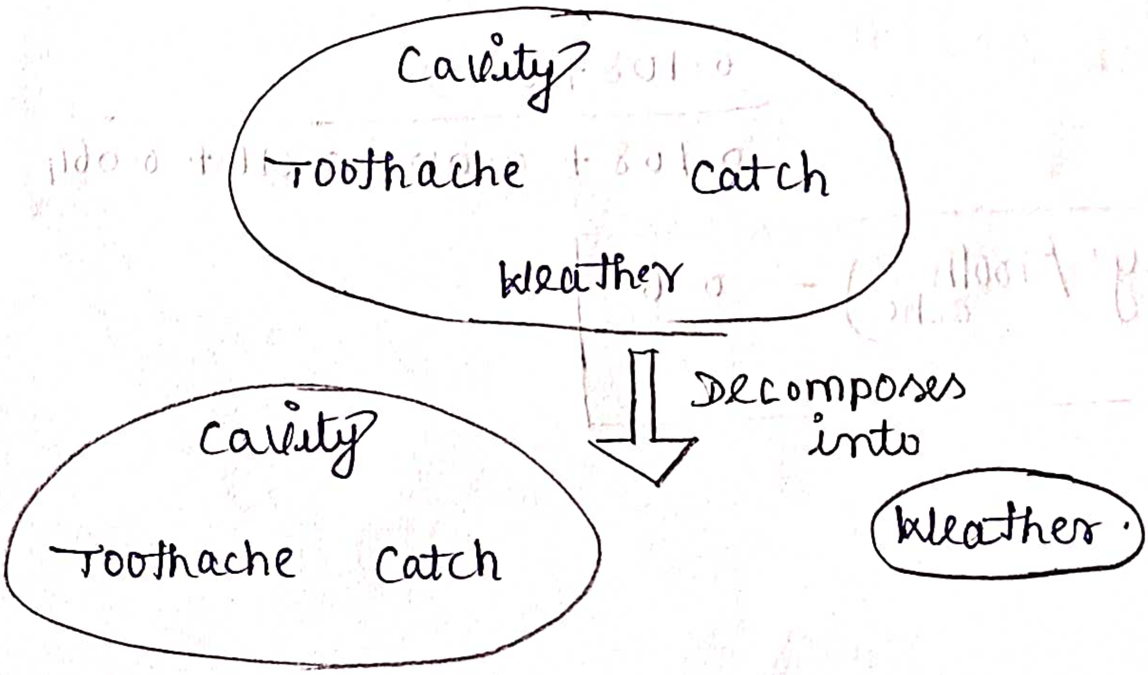
$$\boxed{P(\text{cavity} \mid \text{Toothache}) = 0.4}$$

Independence

⇒ Relationship between two different sets of full joint distribution.

⇒ It is also called Marginal (or) absolute Independence.

ex:



Baye's Theorem

=> Baye's Theorem known as Baye rule or Baye Law.

=> Baye's Theorem is derive from the Conditional probability.

=> The General statement of Baye's Theorem is the conditional probability of an event.

=> A given the occurence of another event B is equal to the product of the event B given A & the probability of A divided by the probability of event B.

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

Where,

=> $P(A)$ & $P(B)$ are the probabilities of events A & B

=> $P(A/B)$ is the probability of event A when event B happens.

=> $P(B/A)$ is the probability of event B when event A happens.

Generalized Baye's rule is

$$P(Y/X) = \frac{P(X/Y)\,P(Y)}{P(X)}$$

Normaligzed Bayle's Rule

$$P(y/x) = \alpha \; P(x/y) \; P(y)$$

Applying Bayle's Rule :

① It requires total three terms one condition & 2 unconditional probabilities for computing one conditional probability.

ex :

Probability of patient having too sugar has high blood pressure is 50%.

Let $M \rightarrow$ be preposition patient has Low sugar.

$S \rightarrow$ patient has high blood pressure.

Suppose we assume that, doctor known following unconditional fact,

(i) Prior probability of $(M) = \frac{1}{50000}$.

(ii) Prior probability of $(S) = 1/20$.

$$P(S/M) = 0.5$$

$P(M) = 1/50000$.

$P(S) = 1/20$.

$P(S/M) = 0.5$

$$P(M/S) = \frac{P(S/M)\;(P(M))}{P(S)}$$

$$= \frac{0.5 \times 20}{50000}$$

$\boxed{P(M/S) = 0.0002}$ → i.e We can expect that 1 in 50000 with high BP will has Low sugar.

② Combining evidence in Baye's Rule

BR is helpful for answering question Conditional on evidences.

ex:

Toothache & catch both evidences are avaiable then cavity is sure to exist.

$P(\text{cavity} \mid \text{Toothache} \wedge \text{catch}) = \alpha \langle 0.108, 0.016 \rangle$
$\approx \langle 0.871, 0.129 \rangle$

Baye's Rule

$P(\text{cavity} \mid \text{Toothache} \wedge \text{catch}) = \alpha P(\text{Toothache} \wedge \text{catch} \mid \text{cavity}) \, P(\text{cavity})$

For this reformulation to work we need to know the conditional probability of the Confection

$\boxed{\text{Toothache} \wedge \text{Catch}}$ → for each value of cavity

That might be feasible for just evidences but again it will not scale up.

⇒ If there are n possible evidence variable x-rays, diet, oral hygiene etc then 2^n possible combinations of observed values.

⇒ Need to know conditional probabilities.

The Notion of Independence can be used here

$$P(\text{Toothache} \wedge \text{Catch} / \text{cavity}) = P(\text{Toothache}/\text{cavity}) \times P(\text{catch}/\text{cavity})$$

Conditional Independence of toothache & catch given cavity.

$$P(\text{cavity} / \text{Toothache} \wedge \text{Catch}) = \alpha P(\text{Toothache}/\text{cavity}) \times P(\text{catch}/\text{cavity}) \times P(\text{cavity})$$

Bayesian Network

⇒ Data structure which is a graph, each node is annotated with quantitative probability information.

⇒ Set of nodes & link ⇒ Topology of the Network
↓
Conditional Independence relationships.

⇒ Joint distribution represented by product of the appropriate elements of the conditional probability Tables [CPTs]

⇒ Full Joint distribution can answer any query in the domain.

⇒ Bayesian Network is correct representation of the domain in conditionally Independent node.

⇒ Nodes in Bayesian Network

Correct order: Add nodes (Root causes)
Reach (Leaves)

Wrong order: More complicated Network.

Bayesian Network both Topology and conditional probability?

① (i) Each Distribution in conditional probability table.

(ii) Row must sum upto 1 because entries represent an exhaustive set of case variables.

② Boolean Variable Probability

True → P

False → 1 − P

③

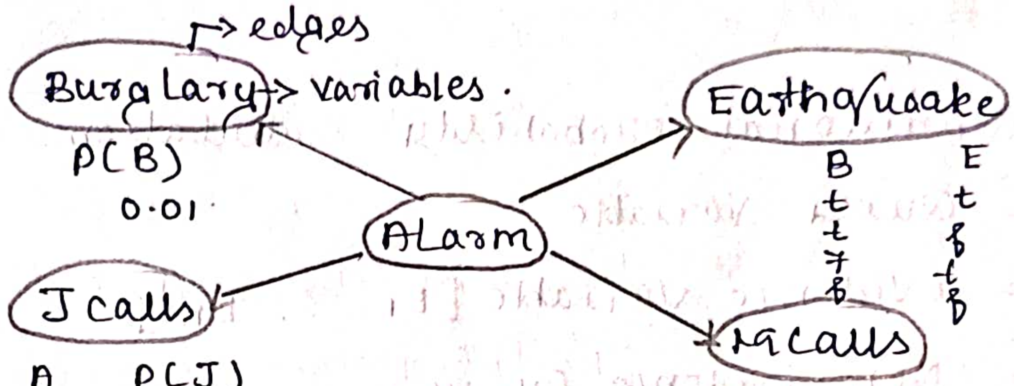
(i) Boolean variables with K boolean parents contain 2^K Independently probabilities.

(ii) Node with No parents has only one ~~No~~ row

↓

PRIOR probability

Bayesian Networks:



$P(B)$ = 0.01

$P(E)$ = 0.02

B	E	P(A)
t	t	.95
t	f	0.91
f	t	0.29
f	f	0.01

A	P(J)
t	0.90
f	0.05

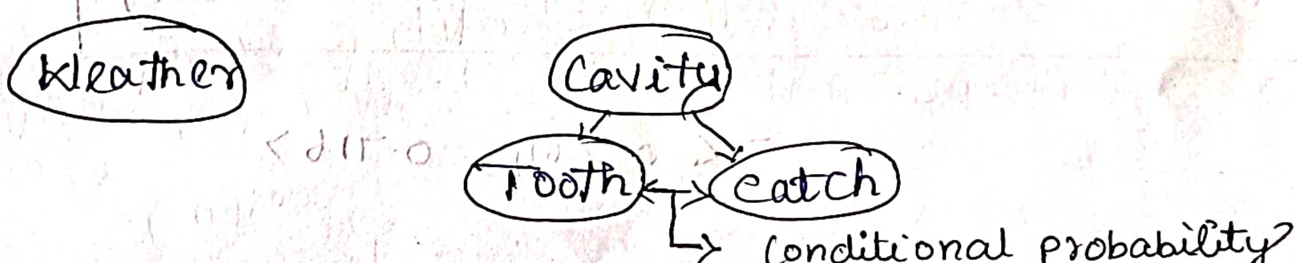
A	P(M)
t	0.70
f	0.01

* Set of nodes & links called topology of Network.
* Set of Random Variable makes nodes of the Network.
* Set of Directly links connects of parent nodes.

Condition

$$x_i = P(x_i \mid Parent(x_i))$$

Example:



* Weather & cavity are Independent.
* Tooth & catch are conditionally Independent.

Inferencing in Bayesian Network

Exact

(1) conditional probability calculation.

X = Query Variable.

E = Evidence Variable $[E_1 \ldots \ldots E_m]$

Y = Non-Evidence $[Y_1, Y_2, \ldots \ldots Y_k]$.
(or) Hidden Variable

$$\boxed{X = \{X\} \cup E \cup Y}$$

ex :

Burglary case

Jcalls = true and Mcalls = true.

Probability Distribution

$$\boxed{P(\text{Burglary} \mid \text{Jcalls} = \text{true}, \text{Mcalls} = \text{true})}$$

$$= \langle 0.284, 0.716 \rangle$$

Bayesian Network Algorithm:

Function : ENUMERATION - ASK (x, e, bn) Returns a distribution over x.

Inputs : x, the query variables.

e, observed values for variables E

bn, a Bayes net with variables.

{x} ∪ E ∪ Y /* Y = Hidden variables */

Q(x) ← A distribution over x, inital empty for each value xi of x do extend e with value xi for x.

Q(xi) ← ENUMERATE - ALL (VARS [bn] e) Return NORMALIZE (Q(x))

Function ENUMERATE - ALL (Vars, e) return a real Number.

if Empty? (Vars) then return 1.0

Y ← First (Vars)

If Y has value y in e

Then return P(y / return(y)) × ENUMERATE - ALL (REST (Vars), e)

else

return Σy P(y / parents(y)) × ENUMERATE - ALL (REST (Vars), ey)

Where,

ey is e extended with Y = y

ex :

$$P(\text{Buralary} \mid j, m) = \alpha P(\text{Buralary}, J, M)$$
$$= \alpha \Sigma_e \Sigma_a P(\text{Buralary}, e, a, j, m)$$

$$P(P \mid j, m) = \alpha \Sigma_e \Sigma_a P(b)\, P(e)\, P(a \mid b, e)\, P(j/a)\, P(m/a)$$

$$P(P \mid j, m) = \alpha P(b) \Sigma_e P(e) \Sigma_a \{P(a/b, e)\, P(j/a)\, P(m/a)$$

$$P(P \mid j, m) = \alpha \langle 0.00059224, 0.0014919 \rangle$$
$$= \langle 0.284, 0.716 \rangle$$

Structure of expression