Bayesian belief network (BBN).

- BBN is a probabilistic graphical model which represents a set a variables and their conditional dependencies using a directed acyclic graph.
- They are probabilistic, becoz these networks are built from a probability distribution, & also use probability theory for prediction & anamoly detection.

- It describes the prob distribution governing a set of variables by specifying a set of conditional independence assumptions along with a set of conditional probabilities.

- It represents the joint probability distribution by specifying a set of conditional independence assumptions with set of local conditional probabilities.

- Each variable in the joint space is represented by a node in Bayesian ntwrk. For each variable two types of info are specified:

  i) The network arcs represent the assertion that the variable is conditionally independent of its non-descendants in the ntwrk given its immediate predecessors in ntwrk.

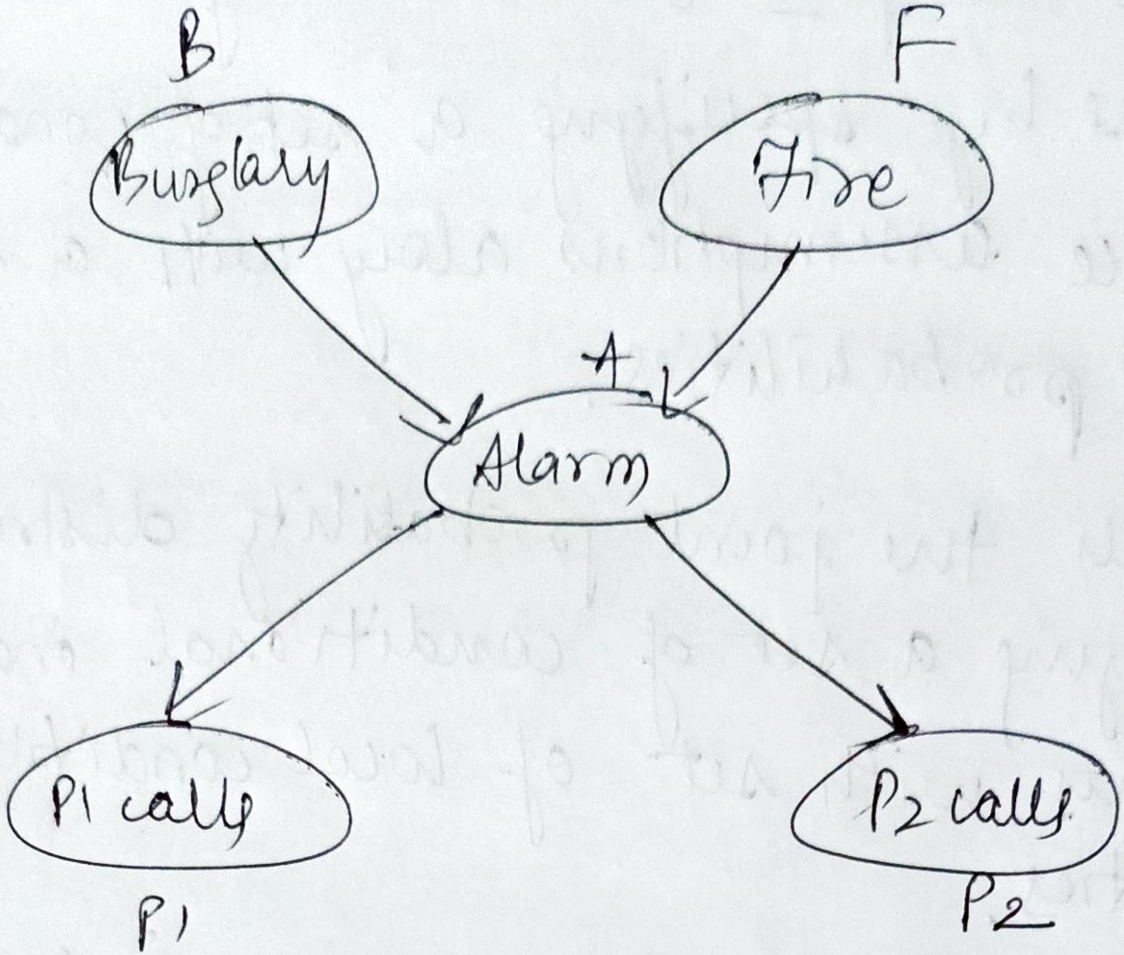
  ii) A conditional prob table is given for each variable, describing the prob distribution for that variable given value of its immediate predecessors.

- The joint prob for any desired assignment of values $(y_1, \ldots y_n)$ to the tuple of ntwrk variables $(Y_1, \ldots Y_1)$ can be computed by formula:-

$$P(y_1 \ldots y_n) = \prod_{i=1}^{n} P(y_i / Parents(Y_i))$$

where Parents (Y_i) denotes set of immediate predecessors of Y_i in the network.

EX :-



- ________
- ________ } geeks.

Inference :- . . . .

6.11.3 Inference

We might wish to use a Bayesian network to infer the value of some target variable (e.g., **ForestFire)** given the observed values of the other variables. We really wish to infer is the probability distribution for the target variable, which specifies the probability that it will take on each of its possible values given the observed values of the other variables. This inference step can be straightforward if values for all of the other variables in the network are known exactly. In the more general case we may wish to infer the probability distribution for some variable (e.g., **ForestFire**) given observed values for only a subset of the other variables (e.g., **Thunder** and **BusTourGroup** may be the only observed values available). In general, a Bayesian network can be used to compute the probability distribution for any subset of network variables given the values or distributions for any subset of the remaining variables.

Exact inference of probabilities in general for an arbitrary Bayesian network is known to be NP-hard. Numerous methods have been proposed for probabilistic inference in Bayesian networks.

Gradient ascent procedure that learns the entries in the conditional probability tables is proposed. This gradient ascent procedure searches through a space of hypotheses that corresponds to the set of all possible entries for the conditional probability tables. The objective function that is maximized during gradient ascent is the probability P(D/h) of the observed training data D given the hypothesis h. By definition, this corresponds to searching for the maximum likelihood hypothesis for the table entries.