Uncertain Domain

- Agent may need to handle uncertainty
- Environment is partially observable
  - Fully observable: An agent knows where it is.
  - The drone never knows its exact position, only an estimated one.
  - The true user intent is not directly observable, only noisy speech data.
  - It cannot see the whole environment or detect humans perfectly.
  - The true health state (disease present or not) is hidden.
- Environment is non-deterministic
  - AI Trading Agent
    - even if the agent takes the same action, the outcome can vary dramatically.
    - from market prices changes unpredictably due to external events (news, other traders, economic data).
    - network latency, competing traders' behavior are stochasic and independent, random fluctuations can affect the outcome.
    - buy action could lead to profit, loss, or no change depending on uncontrollable external factors.
  - AI Helpdesk Chatbot
    - even when the chatbot executes the same action, it sends queries to multiple backend systems, the result may differ each time.
    - network latency or failures, backend load, concurrent requests, and external dependencies can lead to different response times or outcomes.
    - same command doesn't guarantee the same outcome.
- Autonomous Driving
  - Partial Observability
    - Sensors cannot detect everything, blind spots, weather effects, occluded vehicles.
  - Non-Determinism
    - Even if the car signals a turn and starts moving, other drivers might brake suddenly, pedestrians might cross unexpectedly, or traffic lights might malfunction.
    - depending on random external events.
- Medical Diagnosis & Treatment Assistant
  - Partial Observability
    - The patient's internal health state is not directly visible.
    - tests can be inaccurate or incomplete: false positives/negatives, missing data
  - Non-Determinism
    - The same treatment can have different effects on different patients due to genetics, lifestyle, or random biological responses.

        - A drug may succeed, fail, or cause side effects.

Bayes' Theorem

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

- $P(A|B)$: Posterior probability, the probability of event A given taht B has occurred.
- $P(B|A)$: Likelihood, the probability of event B occurring given that A is true.
- $P(A)$: Prior probability, the initial probability of A being true before considering B.
- $P(B)$: Evidence, the total probability of B occurring.
- examples
    - 1% of people have a certain disease (prior probability). $P(Disease) = 0.01$
    - the test for the disease is 99% accurate (likelihood).
        - $P(PositiveTest|Disease) = 0.99$
        - $P(NegativeTest|NoDisease) = 0.99$
        - so $P(PositiveTest|NoDisease) = 0.01$
    - $P(PositiveTest) = P(PositiveTest|Disease) \cdot P(Disease) + P(PositiveTest|NoDisease) \cdot P(NoDisease)$
        - $= 0.99 \cdot 0.01 + 0.01 \cdot 0.99 = 0.0198$
    - $P(Disease|PositiveTest) = \frac{P(PositiveTest|Disease) \cdot P(Disease)}{P(PositiveTest)}$
        - $= \frac{0.99 \cdot 0.01}{0.0198} \approx 0.5$
- the conditional probability $P(effect|cause)$ quantifies the relationship in **the causal direction** from cause to effect.
- but $P(cause|effect)$ is often what we really want to know, describing the relationship in **the diagnostic direction** from effect to cause.
    - In medical dignosis, the doctor knows $P(symptoms|disease)$ from medical studies, and want to derive a $P(disease|symptoms)$ for a particular patient.
    - $$P(disease|symptoms) = \frac{P(symptoms|disease) \cdot P(disease)}{P(symptoms)}$$

General Form of Bayes' Rule

$$P(Y|X) = \alpha \cdot P(X|Y) \cdot P(Y)$$

- α is the normalization constant needed to make the entries in $P(Y|X)$ sum to 1 for each value of X.
- conditional independency

- two variables X and Y are conditionally independent given a third variable Z if the value of X provides no information about Y once Z is known.
- $P(X, Y|Z) = P(X|Z) \cdot P(Y|Z)$

Constructing Bayesian Network

- constructing a Bayesian network in which resulting joint probability distributions are a good representation of the agent's knowledge.
- i. $P(x_1, ..., x_n) = \prod_{i=1}^{n} P(x_i|\text{parents}(x_i))$
- ii. rewrite the entries in the joint probability distribution in terms of conditional probabilities using the product rule.
  - $P(x_1 \wedge ... \wedge x_n) = P(x_n|x_{n-1} \wedge ... \wedge x_1) \cdot P(x_{n-1} \wedge ... \wedge x_1)$
- iii. repeat step 2 until all variables are included.
  - $P(x_1, ..., x_n) = P(x_n|x_{n-1} \wedge ... \wedge x_1) \cdot P(x_{n-1}|x_{n-2} \wedge ... \wedge x_1) \cdots P(x_2|x_1) \cdot P(x_1)$

Knowledge Representation

- The agent's knowledge can at best provide only a degree of belief in the relevant logical sentences.
- A probabilistic agent may have a numerical degree of belief between 0 and 1.

Bayesian Network

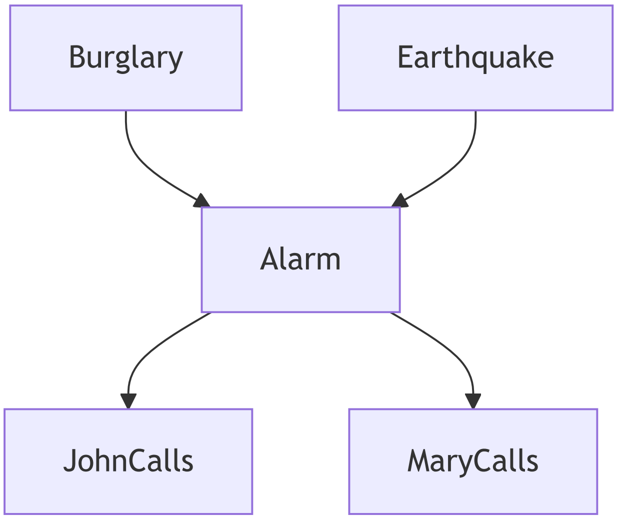
- a directed graph
- each node repressents a random variable and edge with an arrow from X to Y represents that X is a parent of Y.
- each node is associated with a probability distribution
- if a node Y has parents X, the Y node has a conditional probability distribution with respect to its parents X, i.e. $P(Y|\text{parents}(Y))$
- if a node has no any parents, it has an unconditional probability distribution.

Burglary Example

- a new burglar **alarm** installed at home.
- it is reliable at detecting a **burglary**, but also responds on occasion to minor **earthquakes**
- you have two neighbors, **John** and **Mary** who promised to call you at work when they hear the alarm.
- given the evidence of who has or has not called, estimate the probability of a burglary.
- Knowledge representation as the probability of each event that can occur in this case.

$$P(\text{event}_j) = \frac{\text{number of ccurrences of event }_j}{\text{total number of experiments}}$$

- calulate the full joint probability distribution for all relevant variables.
  - specifying probabilities for all possible events one by one is unnatural and tedious.



$$P(B, E, A, J, M) = \prod_{i}^{5} P(X_i|\text{parents}(X_i)) = P(B) \cdot P(E) \cdot P(A|B, E) \cdot P(J|A) \cdot P(M|A)$$

$$\begin{aligned}
&P(M, J, A, E, B) = P(M|J, A, E, B) \cdot P(J, A, E, B) \\
&P(J|A, E, B) \cdot P(A, E, B) \\
&\quad P(A|E, B) \cdot P(E, B) \\
&\quad\quad P(E|B) \cdot P(B) \\
&= P(M|J, A, E, B) \cdot P(J|A, E, B) \cdot P(A|E, B) \cdot P(E|B) \cdot P(B) \\
&= P(M|J, A, E, B) \cdot P(J|A, E, B) \cdot P(A|E, B) \cdot P(E) \cdot P(B) \\
&\quad i.e.\ P(E|B) = P(E) \text{ (Earthquake is independent of Burglary)} \\
&= P(M|A) \cdot P(J|A) \cdot P(A|E, B) \cdot P(E) \cdot P(B) \\
&\quad i.e.\ P(J|A, E, B) = P(J|A) \text{ (JohnCalls depends only on Alarm)} \\
&\quad i.e.\ P(M|J, A, E, B) = P(M|A) \text{ (MaryCalls depends only on Alarm)}
\end{aligned}$$

$$P(M, J, A, E, B) = P(B) \cdot P(E) \cdot P(A|B, E) \cdot P(J|A) \cdot P(M|A)$$

Probabilistic reasnoing

- to explain how to use network models to reason under uncertainty.
- **Exact inference**: calculate the exact posterior probabilities.

- **Approximate inference**: Stochastic approximation techniques such as likelihood weighting and Markov Chain Monte Carlo
  - can give reasonable estimates of the true posterior probabilities in a network
  - can cope with much larger networks than exact inference algorithms.
- **Static worlds**: Each random variable has a single fixed value.
- **Dynamic worlds**: view the world as a series of snapshots, or time slices.
  - Each of snapshots contains a set of random variables, some observed and some hidden.
  - The values of these variables can change over time.
  - With the assumption of state sequences are Markov, the future is independent of the past given the present.

Inference in Bayesian Networks

- Computing the posterior probability distribution of a set of query variables, given some observed events, called some assignment of values to a set of evidence variables.
- X: query variable(s)
  - is it raining?
- E: the set of evidence variables with observed values $E_1, ..., E_m$
  - weather report says it is cloudy.
  - weather report says there is high humidity.
- e: a particular observed event.
  - $E_1 = \text{cloudy}$
  - $E_2 = \text{high humidity}$
- Y: the non-evidence, non-query variables.
  - the other variables in the network.
- $X = \{X\} \cup E \cup Y$
  - the complete set of variables in the network.
- $P(X|e)$: the posterior probability distribution of the query variable(s) given the evidence.
- $P(X|e) = \alpha \sum_y P(X, e, y)$

Exact inference in Bayesian Networks

- One exact inference method is called by inference by **enumeration**.
- Observed some event $A = true$ and $B = true$, want to compute the posterior probability distribution $P(C|A, B)$.
  - Use the full conditional probability table
  - Use the Bayesian Network
- It calculates all the required probability components in each branch in the evaluation tree, and this results in repetitive calculations.
- To avoid this, we can use **variable elimination algorithm** and **clustering algorithms**.

- Calculation of probability distribution for a query variable is complexity exponential.
- Enumeration method can become intractable in large, multiple connected networks. –> use approximate inference methods.

For any Bayesian network with given nodes, $X = \{X_1, X_2, ..., X_n\}$, the joint probability distribution is given by:

$$P(X) = P(X_1 \wedge X_2 \wedge \ldots \wedge X_n) = \prod_{i=1}^{n} P(X_i|\text{parents}(X_i))$$

Using the Bayesian network, we can compute the conditional probability.

$$P(B|event) = \alpha \sum_e \sum_a P(B, E, A, j, m)$$

- $i = 1, P(X_1|\text{parents}(X_1)) = P(B)$
- $i = 2, P(X_2|\text{parents}(X_2)) = P(E)$
- $i = 3, P(X_3|\text{parents}(X_3)) = P(A|B, E)$
- $i = 4, P(X_4|\text{parents}(X_4)) = P(j|A)$
- $i = 5, P(X_5|\text{parents}(X_5)) = P(m|A)$
- $P(B, E, A, j, m) = P(B) \cdot P(E) \cdot P(A|B, E) \cdot P(j|A) \cdot P(m|A)$

$$\begin{aligned} P(B|event) &= \alpha \sum_e \sum_a P(B, E, A, j, m) \\ &= \alpha \sum_e \sum_a P(B) \cdot P(E) \cdot P(A|B, E) \cdot P(j|A) \cdot P(m|A) \\ &= \alpha \cdot P(B) \sum_e P(E) \sum_a P(A|B, E) \cdot P(j|A) \cdot P(m|A) \end{aligned}$$

Berglary example calculation:

$$P(b|event) = \alpha P(b) \sum_e P(E) \sum_a P(A|b,E) \cdot P(j|A) \cdot P(m|A)$$

$$= \alpha P(b) \left[P(e) \sum_a P(A|b,e) \cdot P(j|A) \cdot P(m|A) + P(\neg e) \sum_a P(A|b,\neg e) \cdot P(j|A) \cdot P(m|A) \right]$$

$$\sum_a P(A,b,e) \cdot P(j|A) \cdot P(m|A)$$

$$= P(a,b,e) \cdot P(j|a) \cdot P(m|a) + P(\neg a,b,e) \cdot P(j|\neg a) \cdot P(m|\neg a)$$

$$\sum_a P(A|b,\neg e) \cdot P(j|A) \cdot P(m|A)$$

$$= P(a,b,\neg e) \cdot P(j|a) \cdot P(m|a) + P(\neg a,b,\neg e) \cdot P(j|\neg a) \cdot P(m|\neg a)$$

$$= \alpha P(b) \Big[P(e) \big(P(a|b,e) \cdot P(j|a) \cdot P(m|a) + P(\neg a|b,e) \cdot P(j|\neg a) \cdot P(m|\neg a) \big)$$

$$+ P(\neg e) \big(P(a|b,\neg e) \cdot P(j|a) \cdot P(m|a) + P(\neg a|b,\neg e) \cdot P(j|\neg a) \cdot P(m|\neg a) \big) \Big]$$