- Rational Goals / Decisions : Depend on relative importance of various goals & the likelihood that, & degree to which, they will be achieved.
- Probability provides a way of summarizing the uncertainty that comes from our laziness, ignorance.
- Degree of belief is not equal to degree of truth.
- Degree of truth is subject to fuzzy logic.
- Beliefs are based on percepts that are received to date
- These percepts constitute the evidence on which probability assertions are based.
- Prior / Unconditional probability : Probability before evidence
- Posterior / ~~Un~~conditional probability : " After evidence ;
+ Refers to degree of belief, in absence of any other info.

Uncertainty & rational decisions

- An agent must have preferences betw diff possible outcomes from various outcomes
- Outcomes is completely specified state
- Utility : (In AI) is quality of being useful.
- Utility Theory : Says that every state has usefulness, or utility, to an agent & that an agent will prefer state with higher utility.
- Utility is relative to the agent.
- Decision theory = probability theory + utility theory.

* - Maximum ~~Utility~~ Expected Utility : An agent is rational iff it chooses an action that yields highest expected utility, averaged over all the possible outcomes of the action. (Actions are repeated multiple times; can have diff utility each time)
- Belief state : Reflect the history of percepts to date.

Because the utilities may differ each time, avg. is taken out to find utility of an action

- Agent's belief state not just represent the possibilities of world but these probabilities.
- In probability theory, set of all possible worlds is called sample space. The possible words are mutually exhaustive - two posssible worlds simultaneously not possible. & Ω refers to sample space & ω refers to elements of the space (possible worlds)
- A fully specified probability model associates a numerical probability $P(\omega)$ with every possible world & sum is 1.

  $0 \leq P(\omega) \leq 1$ for each ω & $\sum_{\omega = \Omega} P(\omega) = 1$.
- In probability theory, set of possible worlds are refer to events. Sets are always described by proposition in a formal language.
- ~~For~~ A proposition contains possible worlds in which proposition holds (true).
  + Its probability is the probability sum of probabilities of individuals possible worlds.
- Variables in probability theory are called random variables specified with names with first char capital & values in small case.
- A propositon of random variable can be in small case eg A = true cab be written as 'a' & A = false as ¬a. Also just value can be used instead of proposition.
- Probabilities $P(\text{weather} = \text{sunny}) = 0.6$, $P(\text{weather} = \text{rain}) = 0.1$, & $P(\text{weather} = \text{snow}) = 0.01$ & can be written as

  $\mathbf{P}(\text{Weather}) = \langle 0.6, 0.1, 0.01 \rangle$ where $\mathbf{P}$ is vector.

  Here $\mathbf{P}$ defines probability distribution of Weather.
- For continuous variables, we can define the probability that a random variable takes on some value x as parameterized f^n of x.

  eg $P(\text{MoonTemp} = x) = \text{Uniform}[18C, 26C](x)$ expresses the belief that moon-temp is betw 18 to 26C. We call this a probabity distribution function (PDF)

+ PDF should be interpreted as:
Eg the probability of temp betw $18^c - 26^c$ is 100%,
that in $18^c - 22^c$ value is 50%, $18^c - 20^c$ - 25% & so on.

+ PDF of X at 'x' is written as $P(x)$; is the probability that X̂ falls within arbitrarily small region beginning at x, divided by width of a region.

- Joint Probability ~~proba~~ distribution: Probability dist^t on 2 or more variables. Eg P(weather, cavity)

- Full Joint Probability ~~Function~~ Distribution: JPD for all the random variables of the worlds.

- $P(\neg a) = 1 - P(a)$.

- Inclusion - Exclusion principle:

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$

- Marginal probability = Conditional probability
+ The process is called marginalization / summing out - because we sum of the probabilities of each possible value of the other variables. thereby taking them out of the eq^n; written as

$$P(Y) = \sum_{z \in Z} P(Y, z) \quad \text{Eg} \quad P(\text{Cavity}) = \sum_{z \in \{\text{Catch, Tootache}\}} P(\text{Cavity}, z)$$

- Conditional probability

$$P(Y) = \sum_{x} P(Y|z) P(z).$$

Baye's rule & its use

- Bayes' rule $P(b|a) = \frac{P(a|b)P(b)}{P(a)}$

- General form of Bayes' rule

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

- Conditional independence

$$P(X, Y|Z) = P(X|Z) P(Y|Z)$$

$$P(X|Y, Z) = P(X|Z) \quad \& \quad P(Y|X, Z) = P(Y|Z)$$

- Full Joint distribution

$$P(\text{Cause}, \text{Effect}_1, \ldots, \text{Effect}_n) = P(\text{Cause}) \prod_i P(\text{Effect}_i | \text{Cause}).$$

* Also called „ Naive Bayes Model.

Efficient Representation of Conditional Distribution

- Canonical distribution: Consists of complete CPT with some standard pattern but with fewer parameters.
- Deterministic nodes:- Node whose value is exactly specified by its parents, & with no uncertainty.
- Uncertain relⁿ can be characterized by so-called noisy relⁿ.
- Leak node: Covers miscellaneous causes.

* - The inhibition probability can be used to generate ~~for~~ CPT / full CPT. Only "K" ~~parameters~~ rows are required instead of 2^k

- Ashish R. Gavande