Chapter 12: Quantifying Uncertainty

This is a short chapter (26 pages but not dense) all of which is important.

12.1 – Acting Under Uncertainty

This section connects the topic of uncertainty to previous sections on deliberative agents and maximizing expected utility. Good, focused introduction before starting on probabilities

1. **Uncertainty**:
   - Uncertainty refers to not knowing for sure what state the world is in or what will happen next.
   - For instance, when you're driving to the airport, you might encounter various uncertainties like traffic, car breakdowns, road closures, etc.
   - Agents, like automated taxis, need to handle uncertainty by considering all possible scenarios, even unlikely ones.
2. **Rational Decision-Making**:
   - Rational decision-making involves choosing the best action based on the available information and preferences.
   - Consider the example of choosing a plan to get to the airport on time. Each plan has different probabilities of success and consequences.
   - The rational choice is the one that maximizes the expected utility, which is a combination of the likelihood of success and the desirability of outcomes.
3. **Example**:
   - Let's say you're planning a picnic, but the weather forecast predicts a 70% chance of rain.
   - You have two options: go ahead with the picnic or cancel it and stay indoors.
   - Going ahead with the picnic might give you a 30% chance of enjoying a lovely day outdoors, but a 70% chance of getting rained on.
   - Cancelling the picnic means you'll miss out on the outdoor fun but avoid the risk of getting soaked.
   - To make a rational decision, you consider your preferences (how much you value outdoor activities versus staying dr[illegible]nd the probabilities of each outcome (rain or no rain).

Sure, here are simple definitions for the terms:

1. Uncertainty: The condition of not knowing for sure what will happen or what state the world is currently in. It arises due to factors like incomplete information, randomness, or unpredictability.

2. Rational Decision-Making: Making choices based on a systematic process that takes into account available information, preferences, and goals to select the option that maximizes expected utility. It involves weighing the probabilities of different outcomes and choosing the action with the best overall outcome.

3. Expected Utility: The average or statistical mean of the outcomes' utilities, weighted by their probabilities. It represents the overall desirability or usefulness of an action or outcome based on an agent's preferences and the likelihood of different outcomes occurring.

12.2 Basic Probability Notation

1. What probabilities are, what a probability distribution is. Short, relevant
   Sample space:
   - Sample space refers to the set of all possible outcomes or scenarios in a given situation.
   - For example, when rolling two dice, the sample space consists of all possible combinations of outcomes, such as (1,1), (1,2), (2,3), etc.
2. **Probabilistic assertions and events**:
   - Probabilistic assertions are statements about the likelihood of events or outcomes occurring.
   - Events in probability theory are sets of possible outcomes within the sample space.
   - For instance, when rolling two dice, the event of rolling doubles consists of all outcomes where both dice show the same number.
3. **Unconditional and conditional probabilities**:
   - Unconditional (prior) probabilities represent the likelihood of an event occurring without any additional information.
   - Conditional (posterior) probabilities consider specific evidence or conditions and represent the likelihood of an event given that evidence.
   - For example, the unconditional probability of rolling doubles when rolling two fair dice is 1661, while the conditional probability of rolling doubles given that the first die shows a 5 might be different.
4. **Random variables and probability distributions**:
   - Random variables represent numerical outcomes of random processes, often denoted by uppercase letters.

- Probability distributions describe the likelihood of each possible value of a random variable.
- For example, in weather forecasting, "Temperature at noon" could be a random variable with a probability distribution indicating the likelihood of different temperature values.

5. **Joint probability distributions**:
   - Joint probability distributions describe the likelihood of combinations of values for multiple random variables.
   - They are represented as tables or matrices, with entries indicating the probability of each combination of values.
   - For instance, the joint probability distribution for "Weather" and "Cavity" could show the probability of each possible weather condition and whether a patient has a cavity.

2.3 Inference Using the Full Joint Distribution

Marginalization, conditioning, normalization. Short, relevant

1. **Overview:**
   - Probabilistic inference involves computing posterior probabilities given observed evidence.
   - Full joint distribution serves as the knowledge base for deriving answers.
2. **Example Scenario:**
   - Domain with three Boolean variables: Toothache, Cavity, and Catch.
   - Full joint distribution table provides probabilities for all possible combinations of variables.
3. **Marginalization:**
   - Technique to extract distribution over a subset of variables or a single variable.
   - Achieved by summing probabilities for each possible value of other variables.
   - Ensures normalization of probabilities to sum up to 1.
4. **Conditioning:**
   - Derives conditional probabilities by using the product rule and marginalization.
   - Useful for computing probabilities of variables given evidence about others.
5. **Inference Procedure:**
   - For single-variable queries, given evidence and observed values, compute the probability using a summation over all possible combinations of unobserved variables.
6. **Challenges:**
   - Full joint distribution approach is not scalable for large domains.
   - Requires a large input table and extensive computational resources.
   - Estimating probabilities separately for each entry can be impractical due to the vast number of examples needed.
7. **Practical Considerations:**
   - Full joint distribution serves as a theoretical foundation rather than a practical tool.
   - Realistic reasoning systems employ more efficient approaches for inference.
   - Chapter introduces basic concepts essential for developing practical systems in subsequent chapters.

Normalization is a process used in probability theory to ensure that the probabilities of all possible outcomes of an event add up to 1. This is important because probabilities represent the likelihood of events occurring, and the total probability of all possible outcomes in an event space should always equal 1.

In simpler terms, normalization helps us scale probabilities so they accurately represent the likelihood of each outcome relative to the entire event space.

Example:

Let's say we have a simple event where the outcome can be either heads or tails when flipping a fair coin. The probability of getting heads is 0.5 and the probability of getting tails is also 0.5. To ensure that the probabilities are normalized, we add them together:

Probability of heads + Probability of tails = 0.5 + 0.5 = 1

So, after normalization, the total probability adds up to 1, indicating that one of the two outcomes (heads or tails) is certain to occur when flipping the coin.

Normalization is particularly useful when dealing with conditional probabilities, where we divide the probability of a specific outcome by the total probability of the given condition. This ensures that the resulting probabilities are properly scaled and can be interpreted correctly.

12.4 Independence

Short definition and examples of independence. Crucial to understanding the rest of this module.

12.5 Bayes Rule and its Use

Also central to the module

12.6 Naïve Bayes Models -- Skip

12.7 The Wumpus World Revisited

This is the probabilistic equivalent of what we did in propositional logic, figuring out what the agent knows on the basis of evidence it has collected.

- Uncertainty arises because of both laziness and ignorance. It is inescapable in complex, nondeterministic, or partially observable environments.
- **Probabilities** express the agent's inability to reach a definite decision regarding the truth of a sentence. Probabilities summarize the agent's beliefs relative to the evidence.
- **Decision theory** combines the agent's beliefs and desires, defining the best action as the one that maximizes expected **utility**.
- Basic probability statements include **prior** or **unconditional probabilities** and **posterior** or **conditional probabilities** over simple and complex propositions.
- The axioms of probability constrain the probabilities of logically related propositions. An agent that violates the axioms must behave irrationally in some cases.
- The **full joint probability distribution** specifies the probability of each complete assignment of values to random variables. It is usually too large to create or use in its explicit form, but when it is available it can be used to answer queries simply by adding up entries for the possible worlds corresponding to the query propositions.
- **Absolute independence** between subsets of random variables allows the full joint distribution to be factored into smaller joint distributions, greatly reducing its complexity.
- **Bayes' rule** allows unknown probabilities to be computed from known conditional probabilities, usually in the causal direction. Applying Bayes' rule with many pieces of evidence runs into the same scaling problems as does the full joint distribution.
- **Conditional independence** brought about by direct causal relationships in the domain allows the full joint distribution to be factored into smaller, conditional distributions. The **naive Bayes** model assumes the conditional independence of all effect variables, given a single cause variable; its size grows linearly with the number of effects.
- A wumpus-world agent can calculate probabilities for unobserved aspects of the world, thereby improving on the decisions of a purely logical agent. Conditional independence makes these calculations tractable.