Artificial Intelligence Foundation – JC3001

Lecture 19: Quantifying Uncertainty and Reasoning with Probabilities I

Prof. Aladdin Ayesh (aladdin.ayesh@abdn.ac.uk)

Dr. Binod Bhattarai (binod.bhattarai@abdn.ac.uk)

Dr. Gideon Ogunniye, (g.ogunniye@abdn.ac.uk)

September 2025

Material adapted from:

Russell and Norvig (AIMA Book): Chapters 12 and 13

Sebastian Thrun — Stanford University / Udacity

Course Progression

- Part 1: Introduction
  1. Introduction to AI ✓
  2. Agents ✓
- Part 2: Problem-solving
  1. Search 1: Uninformed Search ✓
  2. Search 2: Heuristic Search ✓
  3. Search 3: Local Search ✓
  4. Search 4: Adversarial Search ✓
- Part 3: Reasoning and Uncertainty
  1. Reasoning 1: Constraint Satisfaction ✓
  2. Reasoning 2: Logic and Inference ✓
  3. **Probabilistic Reasoning 1: BNs**
  4. Probabilistic Reasoning 2: HMMs
- Part 4: Planning
  1. Planning 1: Intro and Formalism
  2. Planning 2: Algos and Heuristics
  3. Planning 3: Hierarchical Planning
  4. Planning 4: Stochastic Planning
- Part 5: Learning
  1. Learning 1: Intro to ML
  2. Learning 2: Regression
  3. Learning 3: Neural Networks
  4. Learning 4: Reinforcement Learning
- Part 6: Conclusion
  1. Ethical Issues in AI
  2. Conclusions and Discussion

Objectives

- Probability Theory
- Probabilistic Inference
- Bayes Rule
- Bayesian Networks
  - Graphical Semantics
  - Inference

- Most environments contain some element of uncertainty.
  - A robot might not be sure where it is due to drift in the motors.
  - Sensors may be unreliable
  - Only partial information is available
- How can we reason about such domains?

- Reasoning Under Uncertainty
- Probability Theory

- Real world problems contain uncertainties due to:
  - partial observability,
  - nondeterminism, or
  - adversaries.
- Example of dental diagnosis using propositional logic

$$Toothache \Rightarrow Cavity$$

- However inaccurate, not all patients with toothaches have cavities

$$Toothache \Rightarrow Cavity \vee GumProblem \vee Abscess \ldots$$

- In order to make the rule true, we have to add an almost unlimited list of possible problems.
- The only way to fix the rule is to make it logically exhaustive

- Rational agents much choose the **right thing**, which depends on:
  - relative importance of the various goals
  - the likelihood that, and degree to which, goals will be achieved.
- Large domains such as medical diagnosis fail for three main reasons:
  - **Laziness:** it is too much work to list complete set of logic rules
  - **Theoretical ignorance:** medical science has no complete theory for the domain
  - **Practical ignorance:** even if we know all rules, uncertainty from partial observability
- An agent only has a degree of belief in the relevant sentences.

- **Probability Theory**
  - tool to deal with degrees of belief of relevant sentences.
  - summarises the uncertainty that comes from our laziness and ignorance.
- Uncertainty and rational decisions
  - An agent requires **preference** among **different possible outcomes** of various plans
  - **Utility Theory:** defines the quality of the outcome being useful
    - Every state has a degree of usefulness/utility
    - Agent prefers higher utility
  - **Decision Theory:** Preferences (Utility Theory) combined with probabilities
    - Decision theory = probability theory + utility theory
    - agent is rational if and only if it chooses the action that yields the highest expected utility, averaged over all the possible outcomes of the action.
    - principle of maximum expected utility (MEU).

Function of a decision-theoretic agent that selects rational actions

function DT-Agent($percept$)
returns an action
  persistent: $beliefState$ ▷ *probabilistic beliefs about the current state of the world*
  $action$ ▷ *the agent's action*
  update $beliefState$ based on $action$ and $percept$
  calculate outcome probabilities for actions,
    given action descriptions and current belief state
  select action with the highest expected utility
    given probabilities of outcomes and utility information
  return $action$

- Reasoning Under Uncertainty

- Probability Theory

Language	Ontological Commitment	Epistemological Commitment
Propositional logic	facts	true/false/unknown
First-order logic	facts, objects, relations	true/false/unknown
Temporal logic	facts, objects, relations, times	true/false/unknown
Probability theory	facts	degree of belief
Fuzzy logic	facts + degree of truth	known interval value

- How can we deal with complex rules which are not always true?
- We extend propositional logic to deal with probabilities.
- We associate a degree of belief with each proposition.
  - $P(h) = 0.5$
- Here, h is a random variable.
- It has a domain of values it can take on (e.g., $\{true, false\}$)
- Random variables can be
  - Boolean: as above, with domain $\{true, false\}$
  - Discrete: taking on values from some countable domain. E.g. Weather can be sunny, rainy, windy, or cloudy.
  - Continuous: taking on values from the set of real numbers.

- In logic, we had a number of possible worlds
  - one had to be true;
  - all others had to be false
- Probability theory talks about how probable each possible world is:
  - Ω (uppercase omega) refers to the **sample space** (the set of all possible worlds)
  - ω (lowercase omega) refers to one such world
- A fully specified **probability model** associates a probability $P(\omega) \in [0, 1]$ to each possible world
  - $\sum_{\omega \in \Omega} P(\omega) = 1$
  - We can use logical formula to specify possible worlds and obtain the probability of all the worlds in which the formula holds, so for any proposition φ, $P(\varphi) = \sum_{\omega \in \phi} P(\omega)$

- Prior, or unconditional probabilities measure the degree of belief associated with some proposition in the absence of any other information.
- For example: $P(coin = heads) = 0.5$ (abbreviated $P(heads) = 0.5$)
- A probability distribution captures the probability of each possible value of the proposition
  E.g. Fair Coin $P(coin)$
  - $P(heads) = 0.5$
  - $P(tails) =$

- Prior, or unconditional probabilities measure the degree of belief associated with some proposition in the absence of any other information.
- For example: $P(coin = heads) = 0.5$ (abbreviated $P(heads) = 0.5$)
- A probability distribution captures the probability of each possible value of the proposition
  E.g. Fair Coin $P(coin)$
  - $P(heads) = 0.5$
  - $P(tails) = 0.5$
  - We write this $P(coin) : P(coin = h) = 0.5, P(coin = t) = 0.5$

- **Conditional** or **posterior probability**: given evidence that has happened, degree of belief of new event
- Conditional probabilities:
  - Probability of a given b:

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

  - Can also be written as (Product Rule):

$$P(a \wedge b) = P(a \mid b)P(b)$$

  - Example of rolling fair dice, rolling doubles when the first dice is 5

$$P(doubles \mid Die_1 = 5) = \frac{P(doubles \wedge Die_1 = 5)}{P(Die_1 = 5)}$$

- Loaded coin — always comes heads up
  $P(heads) =$

Probabilities: Flipping Coins

- Loaded coin — always comes heads up
  $P(heads) = 1 \quad = 100\%$
  $P(tails) = 0$
- $P(heads) + P(tails) = 1$

- Loaded coin
  $P(heads) = 0.75$
  $P(tails) = ?$
- $P(heads) + P(tails) = 1$

- Loaded coin
  $P(heads) = 0.75$
  $P(tails) = 0.25$
- $P(heads) + P(tails) = 1$

Probabilities: Flipping Coins 2

- Loaded coin
  $P(heads) = 0.75$
  $P(tails) = 0.25$
- $P(heads) + P(tails) = 1$
- $P(A) = 1 - P(\neg A)$

$$\{heads, heads\} \qquad P(H) = 0.5$$
$$P(H, H) = ?$$

$$\{heads, heads\} \qquad P(H) = 0.5$$
$$P(H, H) = .25$$

$$\{heads, heads\} \qquad P(H) = 0.5$$
$$P(H, H) = .25$$

$$P(H,H) = \quad P(H) \quad *P(H)$$
$$0.25 \qquad 0.5 \qquad 0.5$$

Truth Table

Flip-1	Flip-2	
H	H	0.25
H	T	0.25
T	H	0.25
T	T	0.25

$$\{heads, heads\} \qquad P(H) = 0.6$$

$$P(T) = ?$$

$$P(H, H) = ?$$

$$P(H, H) = \quad P(H) \quad * P(H)$$

Truth Table

Flip-1	Flip-2	
H	H	?
H	T	?
T	H	?
T	T	?

$$\{heads, heads\} \qquad P(H) = 0.6$$
$$P(T) = 0.4$$
$$P(H,H) = ?$$

$$P(H,H) = \quad P(H) \quad *P(H)$$

Truth Table

Flip-1	Flip-2	
H	H	?
H	T	?
T	H	?
T	T	?

$$\{heads, heads\} \qquad P(H) = 0.6$$
$$P(T) = 0.4$$
$$P(H, H) = ?$$

$$P(H, H) = \quad P(H) \quad * P(H)$$

Truth Table

Flip-1	Flip-2	
H	H	0.36
H	T	0.24
T	H	0.24
T	T	0.16

Two Coin Flips 2

$$\{heads, heads\} \qquad P(H) = 0.6$$
$$P(T) = 0.4$$
$$P(H,H) = \quad 0.36$$

$$P(H,H) = \quad P(H) \quad *P(H)$$

Truth Table

Flip-1	Flip-2	
H	H	0.36
H	T	0.24
T	H	0.24
T	T	0.16

$$\{heads, heads\} \qquad P(H) = 0.6$$
$$P(T) = 0.4$$
$$P(H,H) = \ 0.36$$

$$P(H,H) = \quad P(H) \quad *P(H)$$
$$0.36 \qquad 0.6 \qquad 0.6$$

Truth Table

Flip-1	Flip-2	
H	H	0.36
H	T	0.24
T	H	0.24
T	T	0.16

$P(exactly\ one\ H) = ?$ $\qquad P(H) = 0.5$

Flip-1	Flip-2	
H	H	0.25
H	T	0.25
T	H	0.25
T	T	0.25

$$P(\text{exactly one } H) = 0.5 \qquad P(H) = 0.5$$

Flip-1	Flip-2	
H	H	0.25
H	T	0.25
T	H	0.25
T	T	0.25

$$P(exactly\ one\ H) = 0.5 \qquad P(H) = 0.5$$

Flip-1	Flip-2	
H	H	0.25
H	**T**	$\mathbf{0.25}$ ✓
T	**H**	$\mathbf{0.25}$ ✓
T	T	0.25

Joint Probability Distribution

- A joint probability distribution captures the probability distribution of a set of variables. E.g. $P(coin, die)$
- In the above examples, we have seen that $P(a \wedge b) = P(a) * P(b)$
  Warning: not always true!

Joint Probability Distribution

- A joint probability distribution captures the probability distribution of a set of variables. E.g. $P(coin, die)$
- In the above examples, we have seen that $P(a \wedge b) = P(a) * P(b)$
  Warning: not always true!
- This only works when a and b are independent, formally, when
  $P(b \mid a) = P(b)$

Joint Probability Distribution

- A joint probability distribution captures the probability distribution of a set of variables. E.g. $P(coin, die)$
- In the above examples, we have seen that $P(a \wedge b) = P(a) * P(b)$
  Warning: not always true!
- This only works when a and b are independent, formally, when
  $P(b \mid a) = P(b)$
- E.g. $a \equiv$ it rains today, $b \equiv$ it rains tomorrow:
  $P(b \mid a) > P(b)$

- A joint probability distribution captures the probability distribution of a set of variables. E.g. $P(coin, die)$
- In the above examples, we have seen that $P(a \wedge b) = P(a) * P(b)$
  Warning: not always true!
- This only works when a and b are independent, formally, when
  $P(b \mid a) = P(b)$
- E.g. $a \equiv$ it rains today, $b \equiv$ it rains tomorrow:
  $P(b \mid a) > P(b)$
- Thus, in general: $P(a \wedge b) = P(b \mid a)P(a)$ (conditional probability)
  Notice this also means that $P(a \wedge b) = P(a \mid b)P(b)$

To continue in the next session.