PROBABILITY THEORY 1

ARTIFICIAL INTELLIGENCE | COMP 131

TODAY ON AI

- Uncertainty
- Probability Theory
- Bayes' rule
- Questions?

SECTION 01

Uncertainty

- There are many **situations** where uncertainty arises:
  - When you travel, you reason about the possibility of delays
  - When an insurance company offers a policy, it has calculated the risk that one will claim
  - When you play a game, one cannot be certain what the other player will do
  - A medical expert system that diagnoses disease has to deal with the results of tests that are sometimes incorrect
- Agents which can reason about the effects of **uncertainty** should do better than those that don't.
- How should uncertainty be **represented**?

Emergence and Uncertainty in Gestaltism.

- I have toothache. What is the cause?

  There are many possible causes of an observed event.

- If I go to the dentist and he examines me, when the probe catches this indicates there may be a cavity, rather than another cause.

  The likelihood of a hypothesised cause will change as additional pieces of evidence arrive.

- Bob lives in San Francisco. He has a burglar alarm on his house, which can be triggered by burglars and earthquakes. He has two neighbours, John and Mary, who will call him if the alarm goes off while he is at work, but each is unreliable in their own way. All these sources of uncertainty can be quantified. Mary calls, how likely is it that there has been a burglary?

  Using probabilistic reasoning we can calculate how likely a hypothesised cause is.

Let's say we are planning a trip. We need to go to the airport. Will **leaving for the airport t minutes before the flight** (A_t) get me there on time?

- Partially observable state (state of the road, other drivers' plans, etc.)
- Noisy sensors (traffic reports, etc.)
- Uncertainty in action outcomes (flat tire, etc.)
- Complexity of the model (predicting traffic is hard, etc.)

- Given the following beliefs, which action do I chose?

$$P(A_{25m}|all\_known\_variables) = 0.04$$

$$P(A_{90m}|all\_known\_variables) = 0.70$$

$$P(A_{2h}|all\_known\_variables) = 0.95$$

$$P(A_{24h}|all\_known\_variables) = 0.9999$$

- **Decision theory** combines the agent's beliefs (Probability Theory) and desires (**Utility Theory**), defining the best action as the one that maximizes expected utility

SECTION 02

Probability Theory

- Probability assertions summarize the effect of:
  - **Laziness**: failure to enumerate exceptions, and qualifications of actions, etc.
  - **Theoretical ignorance**: complex models, etc.
  - **Practical ignorance**: lack of relevant facts, initial conditions, etc.

- **Bayesian** or **Subjective** probability relates propositions to one's own state of knowledge
- Probabilities do assert a **belief** and **not facts**
- Probabilities of propositions change with new evidence

$$P(A_{25} | \mathbf{no\_reported\_accident}) = 0.80$$

$$P(A_{25} | \mathrm{no\_reported\_accident}, \mathbf{5AM}) = 0.90$$

- A **random variable** is some aspect of the world that is uncertain:
  - R = Is it raining?
  - T = Is it hot or cold?
  - D = How long will it take to drive to the airport?

- Like **variables** in a **CSP**, random variables have **domains**:
  - **Propositional** or **Boolean**: R in {**T**, **F**} (often write as {r, ¬r})
  - **Multi-valued**: T in $\{\text{hot, cold}\}$
  - **Discrete** or **continuous** (**finite** or **infinite**): D in $[0, \infty)$

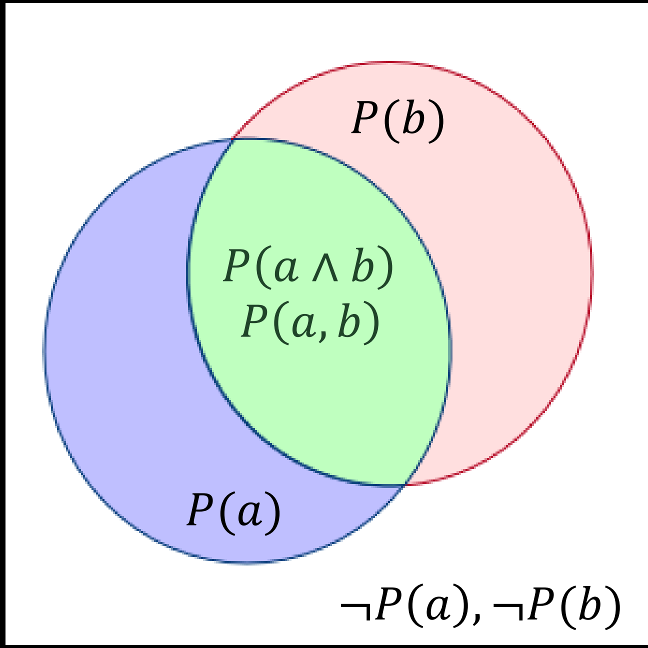
- The domain of coin tossing (let's call it **Throw**) is $\{\text{head}, \text{tail}\}$

- A Boolean random variable has **two** outcomes

- Cavity has the domain {**T**, **F**}

- Toothache has the domain {**T**, **F**}

- A **probability** x is a number defined by the following properties:
  1. $\forall x \; 0 \leq P(X = x) \leq 1 \; and \; \sum_x P(X = x) = 1$
  2. $P(\textbf{True}) = 1, P(\textbf{False}) = 0$
  3. $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$



- Each variable's value has an associated probability called **prior** that corresponds to the **belief** prior to the arrival of any evidence:

$$P(W = \text{rain}) = 0.1$$

- A **distribution** is a table of probability values:

T	P
hot	0.5
cold	0.5

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

- We can create events out of **combinations** of the outcomes of **random variables**

- An **atomic event** is a complete specification of the values of the random variables of interest:
  - Example: if our world consists of only **two** Boolean random variables, then the world has a **four possible atomic events**:

$$
\begin{aligned}
Toothache &= \textbf{true} \land Cavity = \textbf{true} \\
Toothache &= \textbf{true} \land Cavity = \textbf{false} \\
Toothache &= \textbf{false} \land Cavity = \textbf{true} \\
Toothache &= \textbf{false} \land Cavity = \textbf{false}
\end{aligned}
$$

- The set of **all possible atomic events** has two properties:
  - It is **mutually exhaustive** (nothing else can happen)
  - It is **mutually exclusive** (only one of the four can happen at one time)

- An **atomic event** (or **event**) is a set E of outcomes:

$$X_1, X_2, \dots, X_n \quad P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \quad P(x_1, x_2, \dots, x_n)$$

- We can use atomic events to calculate the probability of any new combination:
  - Probability that it's $\text{hot} \land \text{sunny}$?
  - Probability that it's hot?
  - Probability that it's $\text{hot} \lor \text{sunny}$?

- Typically, the events we care about are **partial assignments**, like $P(T = \text{hot})$.

A **joint distribution** ($\wedge$ distribution) over a set of random variables is a table that specifies a probability for each assignment (or **outcome**) when the **random variables happen all at the same time**:

- Must respect the following rules:

$$P(x_1, x_2, \dots, x_n) \geq 0$$

$$\sum_{(x_1, x_2, \dots, x_n)} P(x_1, x_2, \dots, x_n) = 1$$

T	W	$P(T,W)$
hot	sun	0.4
hot	$rain$	0.1
$cold$	sun	0.2
$cold$	$rain$	0.3

- A **probabilistic model** is a joint distribution over a set of random variables

Marginal distributions are sub-tables of joint distributions in which some variables are eliminated:

$$P(X_1 = x_1) = \sum_{x_2} P(X_1 = x_1, X_2 = x_2)$$

$P(Temperature, Weather)$

T	W	$P(T,W)$
hot	sun	0.4
hot	$rain$	0.1
$cold$	sun	0.2
$cold$	$rain$	0.3

$$\forall t \in T: \ P(t) = \sum_{w=\{sun,rain\}} P(t,w)$$

$P(Temperature)$

T	P
hot	0.5
$cold$	0.5

$$\forall w \in W: \ P(w) = \sum_{t=\{hot,cold\}} P(t,w)$$

$P(Weather)$

W	P
sun	0.6
$rain$	0.4

- A **conditional probability** expresses the likelihood that one event a will occur if b occurs:

$$P(a \mid b)$$

$$P(Toothace = \mathbf{T}) = 0.2$$

$$P(Toothace = \mathbf{T} \mid Cavity = \mathbf{T}) = 0.6$$

- So conditional probabilities reflect the fact that **some events** make other **events more** (or **less**) **likely**

- If one event doesn't affect the likelihood of another event, they are said to be independent and therefore:

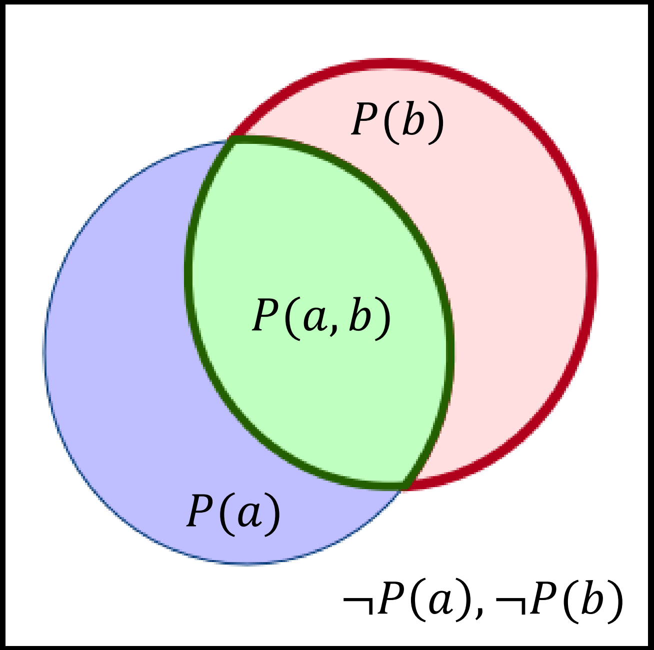
$$P(a \mid b) = P(a)$$

if you roll a 6 on a die, it doesn't make it more or less likely that you will roll a 6 on the next throw. The rolls are independent.

A conditional probability is defined as:

T	W	$P(T,W)$
hot	*sun*	0.4
hot	*rain*	0.1
cold	*sun*	0.2
cold	*rain*	0.3

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



$$P(W = sun|T = cold) \;=\; \frac{P(T = cold, W = sun)}{P(T = cold)}$$

$$P(T = cold, W = sun) \;=\; \mathbf{0.2}$$

$$P(T = cold) \;=\; P(T = cold, W = sun) + (T = cold, W = rain)$$

$$P(T = cold) = 0.2 + 0.3 \;=\; \mathbf{0.5}$$

$$P(W = sun|T = cold) = \frac{0.2}{0.5} = \mathbf{0.4}$$

- How we can work out the likelihood of two events occurring together given their **prior** and **conditional probabilities**? We use the **product rule**:

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

- So in our example:

$$\begin{aligned} P(Tootache, Cavity) &= P(Tootache \mid Cavity)\, P(Cavity) \\ &= \; P(Cavity \mid Toothache) P(Toothache) \end{aligned}$$

- Unfortunately, this doesn’t answer the question: “I have toothache. Do I have a cavity?”

SECTION 03

Bayes' rule

- We can rearrange the two parts of the product rule:

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

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

- **Bayes' rule** states that: $P(a \mid b) = \dfrac{P(b \mid a)}{P(b)} P(a)$

- Then Bayes' rule can be seen as a **relationship cause-effect**
  - allows us to use that model to infer the likelihood of the hidden cause (and thus answer our question)

POSTERIOR BELIEF CONDITIONAL BELIEF PRIOR BELIEF

$$P(cause|effect) = \frac{P(effect|cause)}{P(effect)} P(cause)$$

HYPOTHESIS EVIDENCE

- We can think about some events as being **hidden** causes: not necessarily directly observed (i.e. a cavity).

- Sometimes is **easier** to model how **likely observable effects** are given **hidden causes** (how likely toothache is given a cavity)

- In fact good models of $P(effect|cause)$ are **often available** to us in real domains (i.e. medical diagnosis)

- It is normally **harder** to find out $P(effect|cause)$ for **all** causes independently than it is simply to find out $P(effect)$

Suppose a **doctor** knows that:

- **Meningitis** causes a **stiff neck** in 50% of cases: $P(s \mid m) = 0.5$
- She also knows that the **probability** in the general population of someone having a **stiff neck** at **any time** is $1/20$: $P(s) = 0.05$
- She also has to know the incidence of meningitis in the population is $1/50{,}000$: $P(m) = 0.00002$
- Using Bayes' rule, she can calculate the **probability** the patient has meningitis:

$$P(m \mid s) = \frac{P(s \mid m)\, P(m)}{P(s)} = \frac{0.5 \times 0.00002}{0.05} = 0.0002 = 1/5000$$

- Why wouldn't the doctor be better off if she just knew the **likelihood of meningitis** given a **stiff neck**? I.e. information in the diagnostic direction from symptoms to causes?

- Because **diagnostic knowledge** is often more fragile than **causal knowledge**

- Suppose there was a meningitis epidemic? The rate of meningitis goes up 20 times within a group:

$$P(m \mid s) = \frac{P(s \mid m)\,P(m)}{P(s)} = \frac{0.5 \times \boxed{0.0004}}{0.05} = 0.004 = 1/250$$

- The **conditional belief** $P(s|m)$ is unaffected by the change in $P(m)$, whereas the diagnostic model $P(m|s) = 1/5000$ is now **completely wrong**.

- **Tests and events are separate things:** there is a cancer test result, separate from the event of actually having cancer.
- **Tests are flawed:** Tests detect things that don't exist (false positive), and miss things that do exist (false negative).
- **Tests give us test probabilities, not the real probabilities:** Often the test error is not factored in.
- **False positives skew results:** even with a good test, finding something rare (1 over a million) will likely result in a false positive.
- **Even science is a test:** At a philosophical level, scientific experiments can be considered "potentially flawed tests" and need to be treated accordingly.

PRACTICE

Exercises from the textbook:

any exercise of chapter 13

READINGS

Chapters 13

QUESTIONS ?

ARTIFICIAL INTELLIGENCE

COMP 131

FABRIZIO SANTINI

VERSION 3.2