CS 161 Intro. To Artificial Intelligence

Week 7, Discussion 1C

Li Cheng Lan

Today's Topics

- **Midterm**
- First Order Logic Inference
  - Universal and Existential Instantiation
  - Reduction to Propositional Logic
  - Definite Clauses, Forward Chaining and Backward Chaining
  - Resolution in FOL
- Reasoning Under Uncertainty
  - Terms and Properties
  - Conditional Probability and Independence
  - Probability Inference – Bayes Rule
- Hints for HW7

Midterm – Question 5

- Q: Resolution is a complete inference rule when applied to CNF?
- A: No, resolution is only **refutation complete** if applied to CNF
  - **Refutation complete:** will derive a contradiction if one exists
  - It's not complete in general

Today’s Topics

- Midterm
- **First Order Logic Inference**
  - Universal and Existential Instantiation
  - Reduction to Propositional Logic
  - Definite Clauses, Forward Chaining and Backward Chaining
  - Resolution in FOL
- Reasoning Under Uncertainty
  - Terms and Properties
  - Conditional Probability and Independence
  - Probability Inference – Bayes Rule
- Hints for HW7

FOL Inference – Instantiation

- A Ground term (literal)
  - A ground term is a term without variables
  - E.g., Apple, Color(Apple)
- Variable
  - Can be any specific object <u>in a certain domain</u>
  - E.g. x, y → we often write variable in lower-case letters

If we have only ground terms, we can infer any sentence obtained by substituting a ground term (a term without variables) for the variable.

Universal Instantiation

Universal Instantiation (UI):

For every object in the KB, just write out the rule with the variables substituted.

- UI can be applied **several times** to add new sentences;
  - The new KB is logically **equivalent** to the old
  - Can produce infinite instances if we have a function

Father(John)

$\forall x,\ King(x) \rightarrow Person(x)$

$King(Father(John)) \rightarrow Person(\ldots)$

$King(Father(Father\ldots)) \rightarrow Person(\ldots)$

Example: The objects in our KB include Apple, Orange, MyCar, TheSky

$$\forall x, Fruit(x) \implies Tasty(x)$$

$$\left\{ \begin{array}{l} Fruit(Apple) \implies Tasty(Apple) \\ Fruit(Orange) \implies Tasty(Orange) \\ Fruit(MyCar) \implies Tasty(MyCar) \\ Fruit(TheSky) \implies Tasty(TheSky) \end{array} \right.$$

Existential Instantiation

Existential Instantiation (EI):

Assign a new constant (**Skolem constant**) to the variable.

- Constant name cannot be one we've already used.
- Quantifier can then be discarded.
- EI can be **applied only once** to replace the existential sentence;
  - The new KB is **not equivalent** to the old,
  - But it's satisfiable iff the old KB was satisfiable.

Example:

$$\exists x, Car(x) \wedge ParkedIn(x, E23)$$

$$Car(C) \wedge ParkedIn(C, E23)$$

← C is the Skolem constant in this example

Reduce FOL to Propositional Logic

How to determine if $\forall x\ P$ is True?

- We can reduce it to propositional inference, which is also called **grounding**
- Instantiating all quantified sentences allows us to ground the KB, that is, to make KB propositional
- Then we can apply the inference rules on the propositional KB to show $\Delta \vDash \alpha$

Entailment in FOL is semidecidable:

Herbrand (1930): If sentence α is entailed by an FOL KB, it's entailed by a finite subset of the propositional KB

Joe ← 0
King(F(Joe)) ← 1
King(F(F(Joe))) ← 2
...

- If $\Delta \vDash \alpha$, then $\Delta' \vDash \alpha$ where Δ' is a subset of Δ
- For nesting situation (when we have a function), such as Father(Father(John))), we increase the nesting index by one each time to see if α is entailed by this KB.

Definite Clauses

Definite clauses: has exactly one positive literal

- E.g. $A \vee \neg B$, A, $\neg A \vee \neg B \vee C$ → can also be written as $(A \wedge B \Rightarrow C)$
  if then
- Similar to if-then rule, with positive literals on each side

If we have definite clause KB, we can do **forward chaining** and **backward chaining!**

Definite Clause KB

Example:

The law says that it is a crime for an American to sell weapons to hostile nations.

The country Nono, an enemy of America, has some missiles, and all of its missiles were sold to it by Colonel West, who is American.

Prove that Colonel West is a criminal.

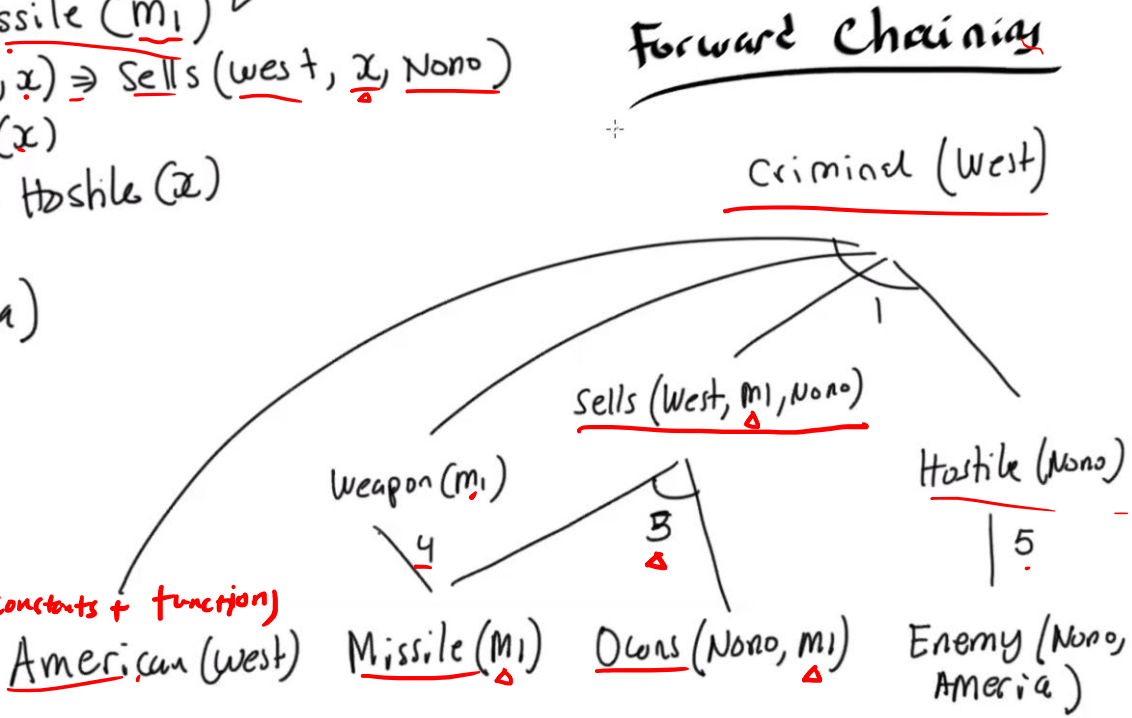
Definite Clause KB – Forward Chaining

KB (Definite clause)

1. American $(x) \wedge$ Weapon $(y) \wedge$ Sells $(x, y, z) \wedge$ Hostile $(z) \Rightarrow$ Criminal (x)
2. ~~∃x~~ Owns (Nono, x) $\wedge$ Missile (x)
   $\rightarrow$ Owns (Nono, M_1) $\wedge$ Missile (M_1) ✓
3. Missile $(x) \wedge$ Owns (Nono, x) $\Rightarrow$ Sells (West, x, Nono)
4. Missile $(x) \Rightarrow$ Weapon (x)
5. Enemy (x, America) $\Rightarrow$ Hostile (x)
6. American (West) ✓
7. Enemy (Nono, America)

Criminal (West)

"Datalog"



constants or (constants + function)

Ground Terms:

Definite Clause KB – Backward Chaining

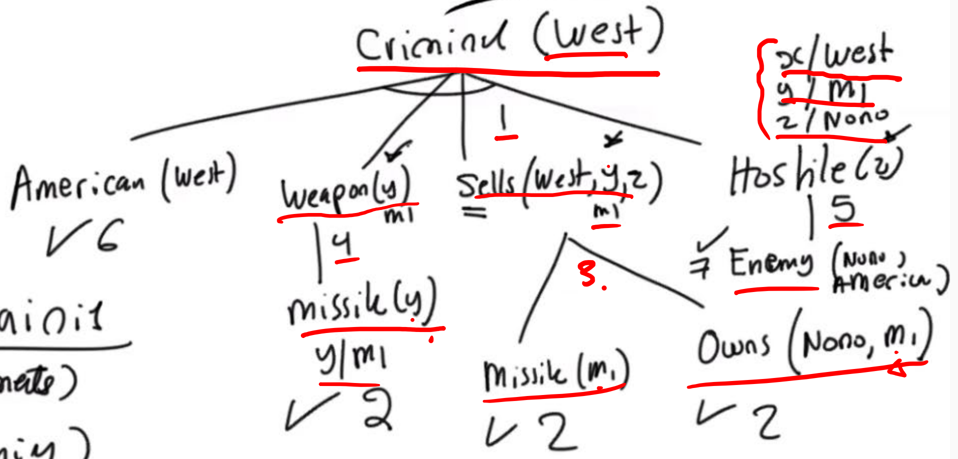
1. $American(x) \wedge Weapon(y) \wedge Sells(x,y,z) \wedge Hostile(z) \Rightarrow Criminal(x)$
2. $\exists x\ Owns(Nono, x) \wedge Missile(x)$ ✓

   $\rightarrow Owns(Nono, M_1) \wedge Missile(M_1)$ ✓
3. $Missile(x) \wedge Owns(Nono, x) \Rightarrow Sells(West, x, Nono)$
4. $Missile(x) \Rightarrow Weapon(x)$
5. $Enemy(x, America) \Rightarrow Hostile(x)$
6. $American(West)$
7. $Enemy(Nono, America)$

$Criminal(West)$

"Datalog"

Backward Chaining



Backward chaining

(with improvements)

(logic programming)

(Prolog)

Resolution in FOL - Unification

Idea: $\Delta \models \alpha$ **iff** $\Delta \wedge \neg\alpha$ **is unsatisfiable**

Unification: finding substitutions that make different logical expressions look identical.

- i.e. Unify(α, β) = θ if SUBST(θ, α) = SUBST(θ, β)
- takes two **atomic** (i.e. single predicates) sentences α, β
- returns a substitution θ that makes α, β identical

$$\text{UNIFY}(Knows(John, x),\ Knows(John, Jane)) = \{x/Jane\}$$
$$\text{UNIFY}(Knows(John, x),\ Knows(y, Bill)) = \{x/Bill, y/John\}$$
$$\text{UNIFY}(Knows(John, x),\ Knows(y, Mother(y))) = \{y/John, x/Mother(John)\}$$
$$\text{UNIFY}(Knows(John, x),\ Knows(x, Elizabeth)) = fail\ .$$

Resolution in FOL - Unification

Standardizing apart: eliminates overlap of variables

$$\text{UNIFY}(Knows(John, x),\ Knows(John, Jane)) = \{x/Jane\}$$
$$\text{UNIFY}(Knows(John, x),\ Knows(y, Bill)) = \{x/Bill, y/John\}$$
$$\text{UNIFY}(Knows(John, x),\ Knows(y, Mother(y))) = \{y/John, x/Mother(John)\}$$
$$\boxed{\text{UNIFY}(Knows(John, x),\ Knows(x, Elizabeth)) = fail\ .}$$

This fails because the two atomic sentences use the same variable name.
If we use Knows(z, Elizabeth) instead of Knows(x, Elizabeth), it should be fine

Most General Unifier

Most General Unifier (MGU):

For every unifiable pair of expressions, there is a single most general unifier (MGU) that is unique up to renaming and substitution of variables

- Unify(Knows(John, x), Knows(y, z)) could return the following two:
  - {y/John, x/z}
    - Then both sentences become Knows{John,z}
    - This is **more general!**
    - This is MGU for this pair of sentences
  - {y/John, x/John, z/John}
    - Then both sentences become Knows(John, John)

Conversion to CNF

Every FOL sentence can be converted into an inferentially equivalent CNF sentence.

- Inferentially equivalent: it is satisfiable exactly when the original sentence is satisfiable

We usually do the following things:

- Eliminate $\Rightarrow$, $\Leftrightarrow$
- Move ¬ down to the atomic formulas
- Eliminate $\forall$, $\exists$
  - Eliminate $\exists$ is **skolemize**
- Rename the variables, if necessary
  - This is called **standardize**
- Distribute $\wedge$ over $\vee$
  - Move $\vee$ down to the literals

Conversion to CNF

Example: Everyone who loves all animals is loved by someone:

$$\forall x \; [\forall y \;\; Animal(y) \Rightarrow Loves(x,y)] \Rightarrow [\exists y \;\; Loves(y,x)]$$

1. Eliminate biconditionals and implications

$A \Leftrightarrow B$: $(A \Rightarrow B) \wedge (B \Rightarrow A)$

$$\forall x \; [\neg\forall y \;\; \neg Animal(y) \vee Loves(x,y)] \vee [\exists y \;\; Loves(y,x)]$$

2. Move $\neg$ inwards: $\neg\forall x, p \;\; \equiv \exists x \;\; \neg p$, $\quad \neg\exists x, p \;\; \equiv \forall x \;\; \neg p$:

$$\forall x \; [\exists y \;\; \neg(\neg Animal(y) \vee Loves(x,y))] \vee [\exists y \;\; Loves(y,x)]$$
$$\forall x \; [\exists y \;\; \neg\neg Animal(y) \wedge \neg Loves(x,y)] \vee [\exists y \;\; Loves(y,x)]$$
$$\forall x \; [\exists y \;\; Animal(y) \wedge \neg Loves(x,y)] \vee [\exists y \;\; Loves(y,x)]$$

Conversion to CNF

3. Standardize apart variables: each quantifier should use a different one

governs

$$\forall x\ [\exists y\ Animal(y) \wedge \neg Loves(x, y)] \vee [\exists z\ Loves(z, x)]$$

4. Skolemize: a more general form of existential instantiation. Each existential variable is replaced by a **Skolem function** of the enclosing universally quantified variables:

$$\forall x\ [Animal(F(x)) \wedge \neg Loves(x, F(x))] \vee Loves(G(x), x)$$

5. Drop universal quantifiers:

$$[Animal(F(x)) \wedge \neg Loves(x, F(x))] \vee Loves(G(x), x)$$

6. Distribute $\wedge$ over $\vee$:

$$[Animal(F(x)) \vee Loves(G(x), x)] \wedge [\neg Loves(x, F(x)) \vee Loves(G(x), x)]$$

Resolution in FOL

- Resolution in Propositional Logic:

$$\frac{\ell_1 \vee \cdots \vee \ell_k, \qquad m_1 \vee \cdots \vee m_n}{\ell_1 \vee \cdots \vee \ell_{i-1} \vee \ell_{i+1} \vee \cdots \vee \ell_k \vee m_1 \vee \cdots \vee m_{j-1} \vee m_{j+1} \vee \cdots \vee m_n}$$

  - each l and each m is a literal and exist one li and mj are complementary literals (i.e. one is the negation of the other)

- **Resolution in FOL:**

Full first-order version: $l_i = \neg m_j$

$$\frac{\ell_1 \vee \cdots \vee \ell_k, \qquad m_1 \vee \cdots \vee m_n}{\text{SUBST}(\theta, (\cdots \vee \ell_{i-1} \vee \ell_{i+1} \vee \cdots \vee m_{j-1} \vee m_{j+1} \vee \cdots))}$$

where $\text{UNIFY}(\ell_i, \neg m_j) = \theta$.$(Unify(\alpha, \beta) = \theta$ if $SUBST(\theta, \alpha) = SUBST(\theta, \beta))$

Resolution in FOL – Example 1

For example,

- $[Animal(F(x)) \vee Loves(G(x), x)]$ and $[\neg Loves(u, v) \vee \neg Kills(u, v)]$
- We could eliminate $Loves(G(x), x)$ and $\neg Loves(u, v)$ with unifier $\theta = \{u/G(x), v/x\}$ to produce the resolvent clause $Animal(F(x)) \vee \neg Kills(G(x), x)$

Resolution in FOL – Example 2

Same example as before:
The law says that it is a crime for an American to sell weapons to hostile nations.
The country Nono, an enemy of America, has some missiles, and all of its missiles were sold to it by Colonel West, who is American.

Prove that Colonel West is a criminal.

Resolution in FOL – Example 2

Convert sentences to FOL:

... it is a crime for an American to sell weapons to hostile nations:

$American(x) \land Weapon(y) \land Sells(x, y, z) \land Hostile(z) \Rightarrow Criminal(x)$

Nono ... has some missiles, i.e.,

$\exists x\ Owns(Nono, x) \land Missile(x)$:

$Owns(Nono, M_1)$ and $Missile(M_1)$

... all of its missiles were sold to it by Colonel West

$\forall x\ \ Missile(x) \land Owns(Nono, x) \Rightarrow Sells(West, x, Nono)$

Missiles are weapons:

$Missile(x) \Rightarrow Weapon(x)$

An enemy of America counts as "hostile":

$Enemy(x, America) \Rightarrow Hostile(x)$

West, who is American ...

$American(West)$

The country Nono, an enemy of America ...

$Enemy(Nono, America)$

Resolution in FOL – Example 2

Transfer FOL into CNF:

- $American(x) \land Weapon(y) \land Sells(x, y, z) \land Hostile(z) \Rightarrow Criminal(x)$
  - $\neg American(x) \lor \neg Weapon(y) \lor \neg Sells(x, y, z) \lor \neg Hostile(z) \lor Criminal(x)$
- $\forall x \ \ Missile(x) \land Owns(Nono, x) \Rightarrow Sells(West, x, Nono)$
  - $\neg Missile(x) \lor \neg Owns(Nono, x) \lor Sells(West, x, Nono)$
- $Missile(x) \Rightarrow Weapon(x)$
  - $\neg Missile(x) \lor Weapon(x)$
- ...

Resolution in FOL – Example 2

Now we have KB:

CNF

- $\neg American(x) \lor \neg Weapon(y) \lor \neg Sells(x, y, z) \lor \neg Hostile(z) \lor Criminal(x)$
- $\neg Missile(x) \lor \neg Owns(Nono, x) \lor Sells(West, x, Nono)$
- $\neg Enemy(x, America) \lor Hostile(x)$
- $\neg Missile(x) \lor Weapon(x)$
- $Owns(Nono, M_1)$
- $\underline{American}(West)$
- $Missile(M_1)$
- $Enemy(Nono, America)$

We want to prove $Criminal(West)$

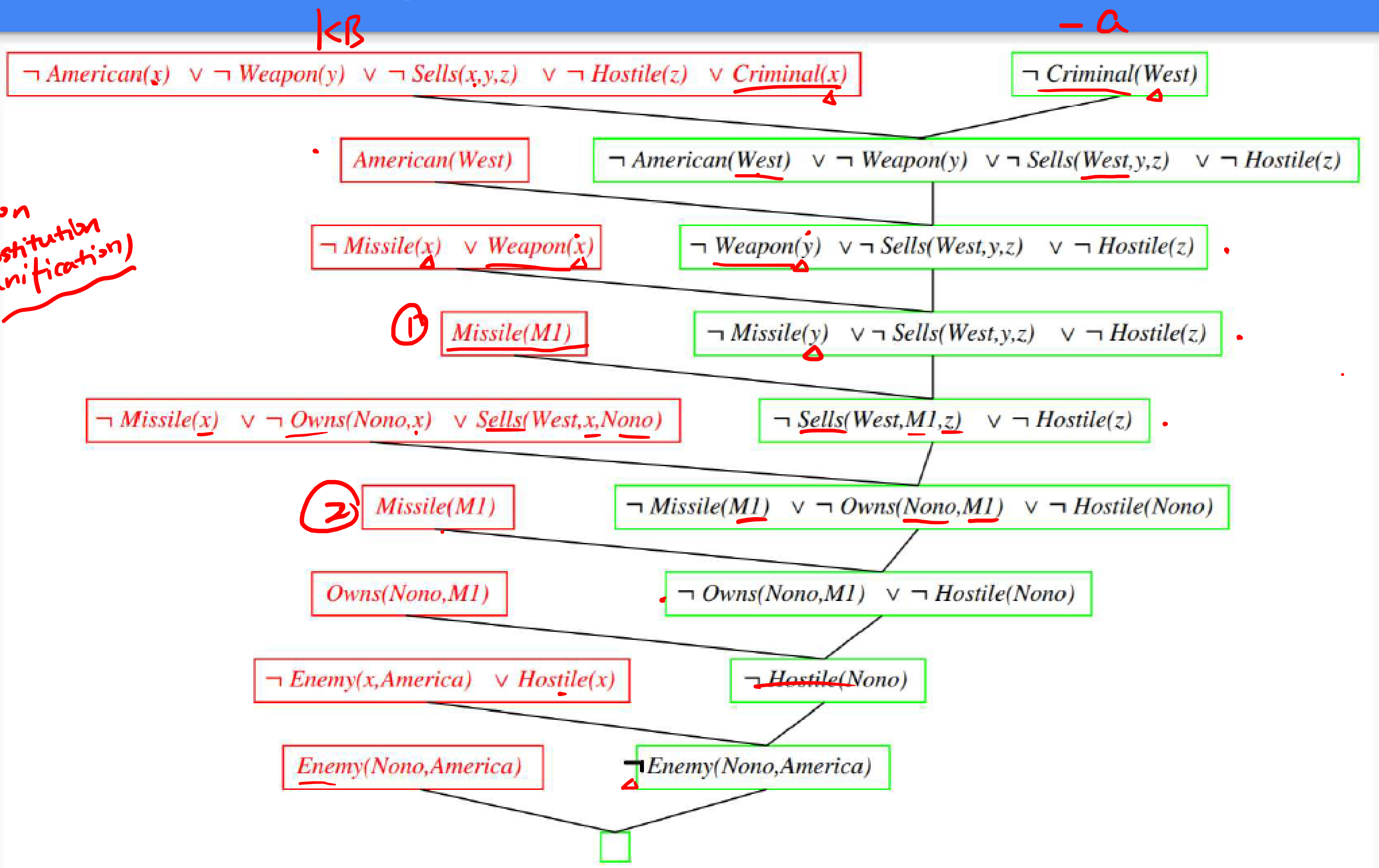
- Apply resolution steps to $CNF(\underline{KB \land \neg\alpha})$

unsatisfiable
(empty clause)

- Show $KB \land \neg Criminal(West)$ is unsatisfiable!

Resolution in FOL – Example 2

Resolution Steps:



resolution with substitution (unification)

Summary of Propositional Logic & FOL

- Propositional Logic
  - Inference: *Depth-first enumeration, Deduction Theorem, Modus Ponens, And Introduction, Or Introduction, And Elimination, Resolution*
  - Inference and Proof (of entailment): *Model Checking, Use Inference Rules such as Resolution (proof with refutation + resolution + CNF), SAT solvers (search), tractable NNF circuits*
- First-order Logic
  - Syntax and Semantics: *basic element, sentences, model, KB*
  - Quantifier: *$\forall$ and $\exists$, nesting quantifier, duality*
  - Translating Between English Sentences and FOL
- Inference First-order Logic
  - Reduce FOL to PL: instantiation, grounding
  - Definite clause KB: *Forward and Backward Chaining (proof with modus ponens)*
  - Directly apply rules:
    - Unification
    - Resolution: turn text into KB, unification, conversion to CNF, resolution + refutation

Today's Topics

- Midterm
- First Order Logic Inference
  - Universal and Existential Instantiation
  - Reduction to Propositional Logic
  - Definite Clauses, Forward Chaining and Backward Chaining
  - Resolution in FOL
- **Reasoning Under Uncertainty**
  - Terms and Properties
  - Conditional Probability and Independence
  - Probability Inference – Bayes Rule
- **Hints for HW7**

Problems of Logical Inference

Consider the scenario:

Q: Let action A_t = leave for airport t minutes before flight, will A_t get me there on time?

A: If we use a purely logical approach:

- "A_{25} will get me there on time", may be wrong!
- "A_{25} will get me there on time if there's no accident on the bridge and it doesn't rain and my tires remain intact, etc.",

→ There exists uncertainty! Logical inference can't handle uncertainty!

How to handle uncertainty?

Problems of Logical Inference

Probability helps to handle uncertainty!

Probabilistic assertions summarize effects of

- Laziness: failure to enumerate exceptions, qualifications, etc.
- Ignorance: lack of relevant facts, initial conditions, etc.

- **Subjective** or **Bayesian** probability:
  - Probabilities relate propositions to one's own state of knowledge
    - E.g., P(A25|no reported accidents) = 0.06
  - Probabilities of propositions can change with new evidence:
    - E.g., P(A25|no reported accidents, 5 a.m.) = 0.15

Term and Properties

- Sample Space: $\boldsymbol{\Omega}$
    - E.g. 6 possible rolls of a die (1, 2, 3, 4, 5, 6)
- Sample Point/Possible World/Atomic Event: $\omega \in \boldsymbol{\Omega}$
    - E.g. 1 is a sample point
- Probability Space/Probability Model:
    - A sample space with an assignment $P(\omega)$ for every $\omega \in \boldsymbol{\Omega}$ s.t. $\Sigma P(\omega)=1$ and $0 =< P(\omega) <= 1$
        - E.g. $P(1) = P(2) = \ldots = P(6) = \frac{1}{6}$
- Event:
    - a subset $A \in \boldsymbol{\Omega}$, $P(A) = \Sigma\_\{\omega \in A\} P(\omega)$
        - E.g. $P(\text{die roll} < 4) = P(1)+P(2)+P(3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$
- Random Variable:
    - Variable in probability theory
    - Each random variable has a domain (set of possible values it can take on), and is associate with a probability distribution

Term and Properties

- Prior/Unconditional Probability
  - Belief prior to arrival of any evidence
    - E.g. P(weather = sunny) = 0.8
- Probability distribution
  - Gives values for all possible assignments:
    - E.g. P(Weather) = {0.8, 0.09, 0.1, 0.01} (**normalized**, i.e., sums to 1)
- Joint Probability Distribution
  - Gives the probability of every atomic event on a set of random variables
    - E.g. P(weather, event) = 4 x 2 matrix of values:

Weather =	sunny	rainy	cloudy	snow
Event = true	0.7	0.04	0.09	0.002
Event = false	0.1	0.05	0.01	0.008
	0.8	0.09	0.1	0.01 = P(weather)

- Conditional/Posterior Probability
  - A measure of the probability of an event occurring given that another event has (by assumption, presumption, assertion or evidence) occurred

Conditional Probability

Conditional Probability:

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \text{ if } P(b) \neq 0$$

- Product rule:

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

- Chain Rule (general product rule): conditional

$$\begin{aligned}\mathbf{P}(X_1, \ldots, X_n) &= \mathbf{P}(X_1, \ldots, X_{n-1})\ \mathbf{P}(X_n|X_1, \ldots, X_{n-1}) \\ &= \mathbf{P}(X_1, \ldots, X_{n-2})\ \mathbf{P}(X_{n-1}|X_1, \ldots, X_{n-2})\ \mathbf{P}(X_n|X_1, \ldots, X_{n-1}) \\ &= \ldots \\ &= \prod_{i=1}^{n} \mathbf{P}(X_i|X_1, \ldots, X_{i-1})\end{aligned}$$

Σ, $\prod_{i=1}^{n}$

$X_1 + X_2 + \cdots$ $X_1 \cdot X_2 \cdot X_3 \cdots$

← Note that:
$P(X_1,\ldots,X_n) = P(X_1 \wedge X_2 \wedge \cdots \wedge X_n)$

Conditional Probability - Example

Example:

Q: Given a propositional sentence $\alpha: A \vee B$, what's the probability P(A,B | α)?

A	B	P(A,B)
T	F	0.2
F	T	0.1
T	T	0.5
F	F	0.2

$\Big\}$ 0.8

A:

A	B	P(A,B | a)
T	F	0.2/0.8=0.25
F	T	0.1/0.8
T	T	0.5/0.8
F	F	0

$\Big\}$ 1

① find the worlds/rows s.t. a is true.
② find conditional probability based on sum of prob. in ①.

Independence

Independence:

- Independence: $A \perp B$
  - P(A|B) = P(A), or P(B|A) = P(B), or P(A,B) = P(A)P(B)
  - E.g. $$\mathbf{P}(Toothache, Catch, Cavity, Weather) \\ = \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$$
- Conditional independence: $A \perp B|C$
  - P(A,B|C) = P(A|C)P(B|C), or P(A|B,C) = P(A|C)
  - This indicates: A and B are independent when the value of C is known and fixed
  - E.g. $Catch$ is **conditionally independent** of $Toothache$ given $Cavity$:

    $$\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$$

    $$\begin{aligned} &\mathbf{P}(Toothache, Catch, Cavity) \\ &= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch, Cavity) \\ &= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity) \\ &= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity) \end{aligned}$$

Probability Inference – Bayes Rule

Probability inference: inference the probability of one event

How to do it?

- Inference by enumeration
- Inference rule:
    - Bayes' Rule:

Bayes' Rule can be used in probablity inference when we have $P(b|a)$ but not $P(a|b)$.

Product rule $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$

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

or in distribution form

$$\mathbf{P}(Y|X) = \frac{\mathbf{P}(X|Y)\mathbf{P}(Y)}{\mathbf{P}(X)} = \alpha \mathbf{P}(X|Y)\mathbf{P}(Y)$$

Bayes Rule

Useful for assessing **diagnostic** probability from **causal** probability:

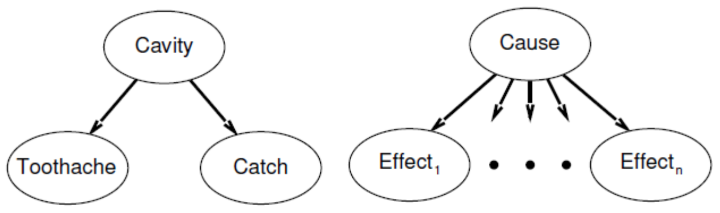
$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

- **Naïve Bayes** model:

$$\mathbf{P}(Cause, Effect_1, \ldots, Effect_n) = \mathbf{P}(Cause)\prod_i \mathbf{P}(Effect_i|Cause)$$

Example:

$$\begin{aligned}
&\mathbf{P}(Cavity|toothache \wedge catch) \\
&= \alpha\,\mathbf{P}(toothache \wedge catch|Cavity)\mathbf{P}(Cavity) \\
&= \alpha\,\mathbf{P}(toothache|Cavity)\mathbf{P}(catch|Cavity)\mathbf{P}(Cavity)
\end{aligned}$$



Questions?

My slides take the following materials as references:

- Prof. Darwiche's lecture video
- Shirley Chen's slides
- Yewen Wang's (Winter 2020's TA) slides
- Prof. Quanquan Gu's (Winter 2020) slides

Thank you!