CS 540 Introduction to Artificial Intelligence

Logic

University of Wisconsin-Madison

Spring 2026 Sections 1 & 2

Announcements

- **Homework:**
  - HW 1 due on Wednesday **February 4 at 11:59PM**
  - HW2 will be released on Wednesday
- Class roadmap:

Logic
NLP
Machine Learning: Introduction
Machine Learning: Unsupervised Learning I

Mostly Foundations

Outline

- Review PCA
- Introduction to logic
  - Arguments, validity, soundness
- Propositional logic
  - Sentences, semantics, inference
- First order logic (FOL)
  - Predicates, objects, formulas, quantifiers

BEGRIFFSSCHRIFT,

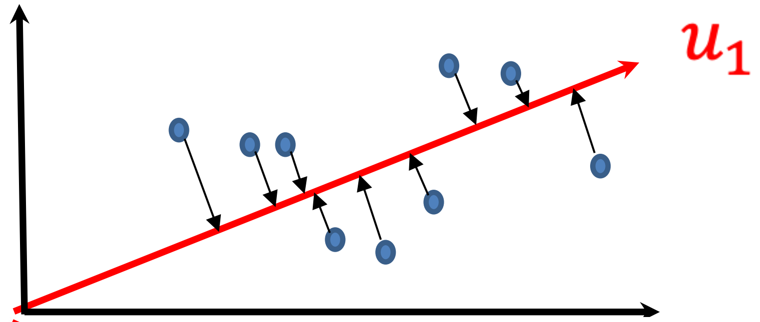
EINE DER ARITHMETISCHEN NACHGEBILDETE

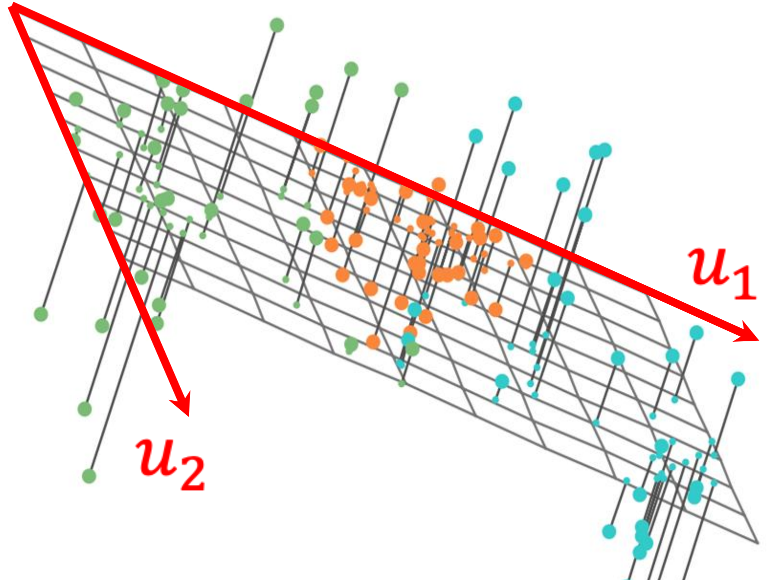
FORMELSPRACHE

DES REINEN DENKENS.

Principal Components Analysis (**PCA**)

- Find axes $u_1, u_2, \ldots, u_n \in \mathbb{R}^d$ of a subspace
  - Will project to this subspace
- These vectors are the **principal components**
- Want to preserve data
  - Minimize projection error

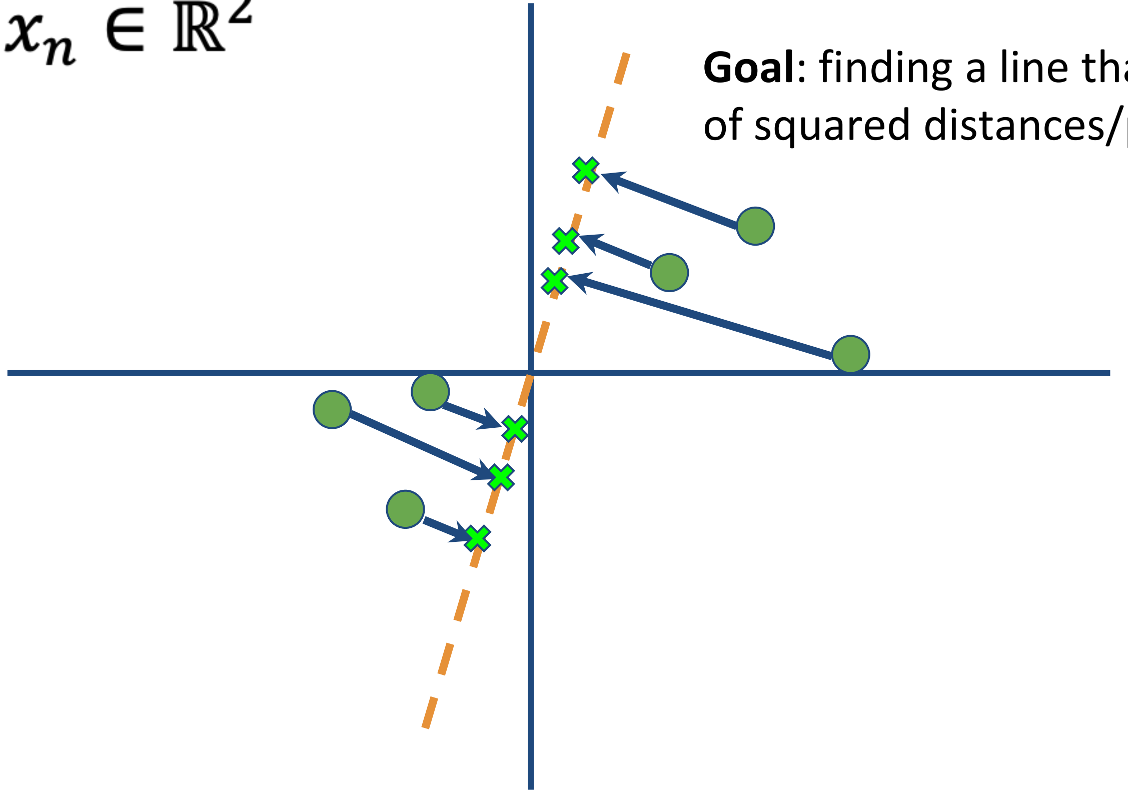




Projection: An Example

$x_1, x_2, \ldots, x_n \in \mathbb{R}^2$

Goal: finding a line that minimizes the sum of squared distances/projection errors.

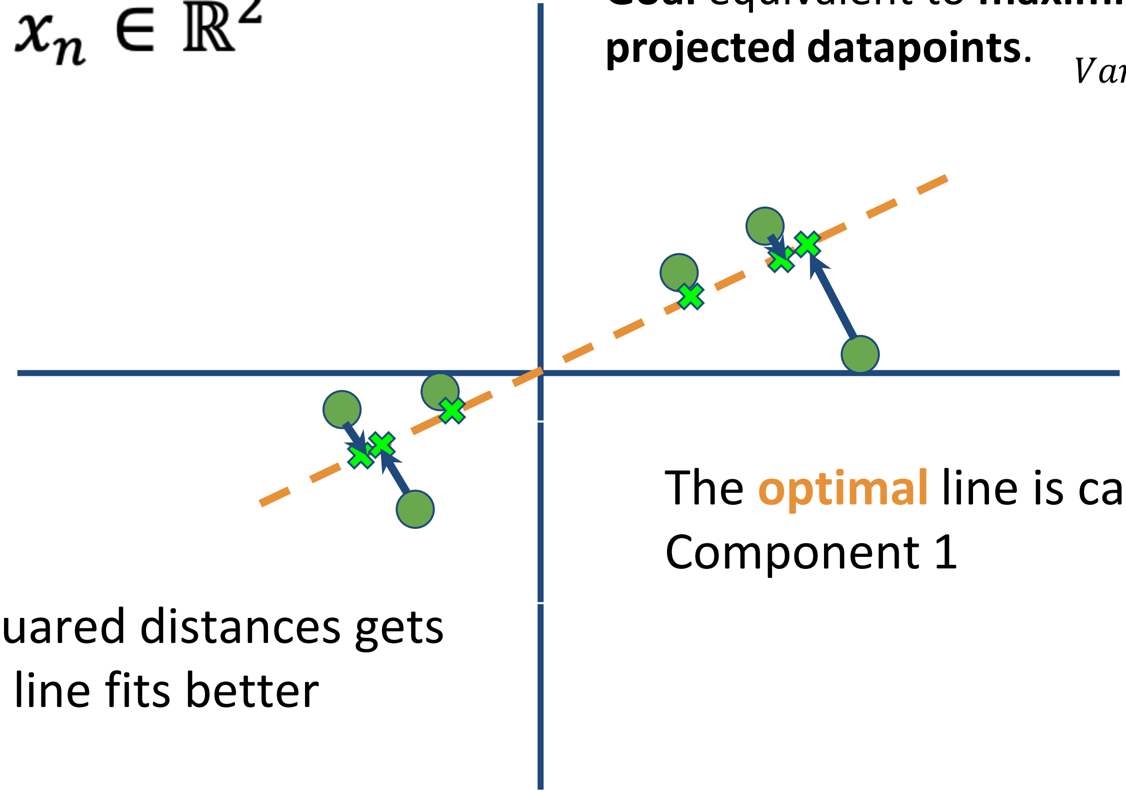


Projection: An Example

$x_1, x_2, \dots, x_n \in \mathbb{R}^2$

Goal equivalent to **maximizing the variance of projected datapoints**.

$$Variance = \frac{1}{n-1}\sum_{i=1}^{n}\langle u, x_i\rangle^2$$



The **optimal** line is called Principal Component 1

The sum of squared distances gets smaller as the line fits better

PCA Procedure

Input: data $x_1, x_2, \ldots \ldots x_n \in \mathbb{R}^d$
Output: principal components $u_1, u_2, \ldots \ldots u_m \in \mathbb{R}^d$

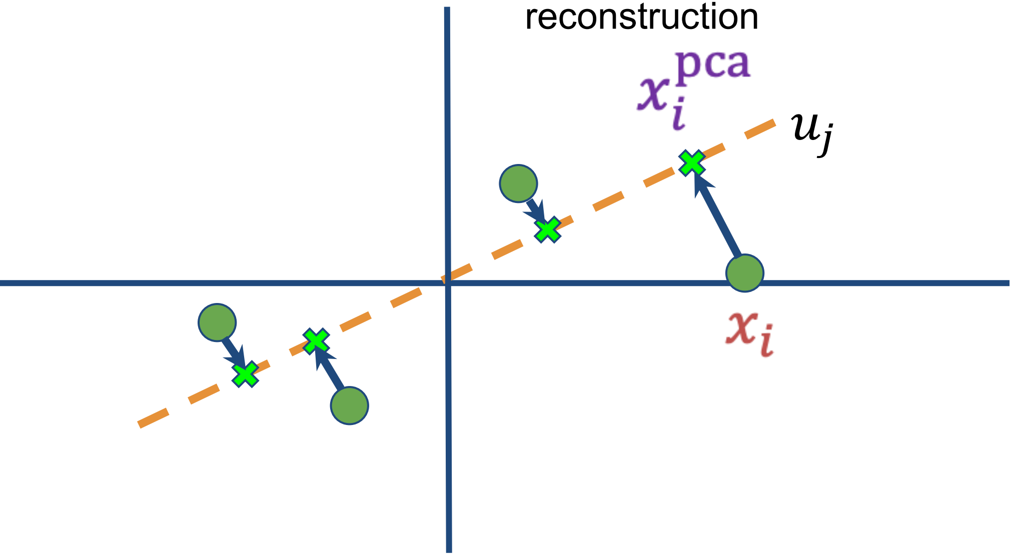
Step 1: Center data at the origin so that $\hat{\mu} = 0$
Step 2: Compute **covariance matrix** ($S = \frac{1}{n-1} \sum_{i=1}^{n} x_i x_i^T$)
Step 3: Compute **eigenvectors** and **eigenvalues** of covariance matrix
Step 4: **Sort eigenvalues** from high to low
Step 5: Select **top-m eigenvectors,** these are the **principal components**

PCA Projection and Reconstruction

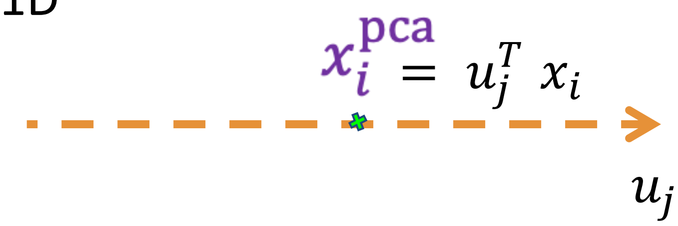
projection

For each x_i: $x_i^{pca} = \sum_{j=1}^{m} (u_j^T x_i) u_j$

reconstruction



2D → 1D



Break & Quiz

Q 1.1 Let $x_1, \dots x_n \in \mathbb{R}^3$ be a centered dataset and let S be the sample covariance matrix. Suppose S has the following eigenvectors with eigenvalues of 3, 9, and 6, respectively. What is the second principal component?

$$v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

A. $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ B. $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ C. $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ D. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Break & Quiz

Q 1.1 Let $x_1, \dots x_n \in \mathbb{R}^3$ be a centered dataset and let S be the sample covariance matrix. Suppose S has the following eigenvectors with eigenvalues of 3, 9, and 6, respectively. What is the second principal component?

$$v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

A. $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ B. $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ C. $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ D. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Break & Quiz

Q 1.1 Let $x_1, \dots x_n \in \mathbb{R}^3$ be a centered dataset and let S be the sample covariance matrix. Suppose S has the following eigenvectors with eigenvalues of 3, 9, and 6, respectively. What is the second principal component?

$$v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

A. $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ B. $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$ C. $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ D. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Principal components are ordered by the magnitude of their corresponding eigenvalues (from largest to smallest).

1. The eigenvalue 9 is the largest, so its eigenvector

is the first principal component.

2. The eigenvalue 6 is the second largest, so its eigenvector is the second principal component.

3. The eigenvalue 3 is the smallest.

Therefore, the correct answer is the vector corresponding to eigenvalue 6.

Break & Quiz

Q 1.2 Suppose we perform PCA to compress a dataset of points in $\mathbb{R}^4$ into a dataset of points in $\mathbb{R}^2$. We find that the first principal component is $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ and the second principal component is $\begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ 0 \\ -\frac{1}{\sqrt{3}} \end{pmatrix}$. If $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ is a point in the original dataset, what is the compressed version $\bar{x} \in \mathbb{R}^2$

A. $\begin{pmatrix} 1 \\ -\frac{1}{\sqrt{3}} \end{pmatrix}$ B. $\begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}$ C. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ D. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Break & Quiz

Q 1.2 Suppose we perform PCA to compress a dataset of points $\mathbb{R}^4$ into a dataset of points in $\mathbb{R}^2$. We find that the first principal component is $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ and the second principal component is $\begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ 0 \\ -\frac{1}{\sqrt{3}} \end{pmatrix}$. If $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ is a point in the original dataset, what is the compressed version $\bar{x} \in \mathbb{R}^2$

The compressed representation is $\begin{pmatrix} u_1^\top x \\ u_2^\top x \end{pmatrix}$

A. $\begin{pmatrix} 1 \\ -\frac{1}{\sqrt{3}} \end{pmatrix}$ B. $\begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}$ C. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ D. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Logic

Logic & AI

Why are we studying logic?

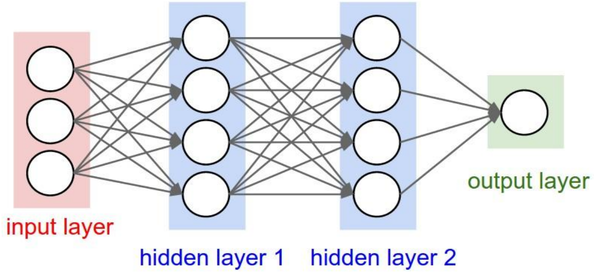
- **Traditional** approach to AI ('50s-'80s)
  - "Symbolic AI"
  - The Logic Theorist – 1956
    - Proved a bunch of theorems!

- Logic also the language of:
  - Knowledge representation, databases, etc.

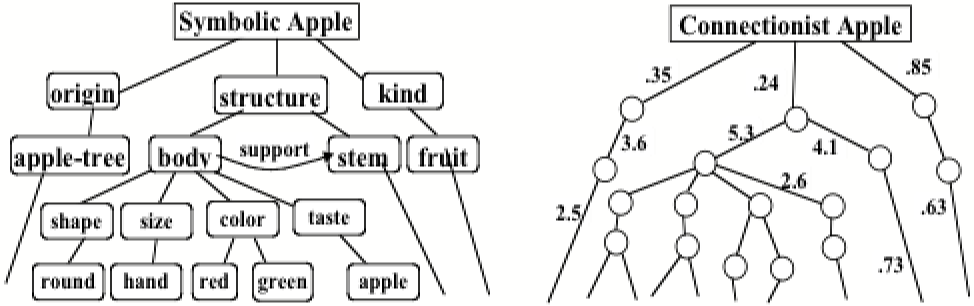
Symbolic vs Connectionist

Rival approach: **connectionist**

- Probabilistic models
- Neural networks
- **Extremely popular** last 20 years



Stanford CS231n

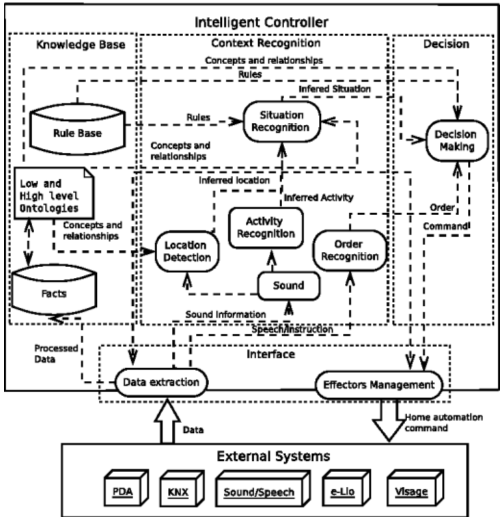


M. Minsky

Symbolic vs Connectionist

Which is better?

- Future: combination; best-of-both-worlds.
  - "Neurosymbolic AI"
  - **Example**: Markov Logic Networks



Logic, AI, and the Future of Math

Tools of logic might allow AI to write new, formally verifiable proofs

JUNE 8, 2024 | 12 MIN READ

AI Will Become Mathematicians' 'Co-Pilot'

Fields Medalist Terence Tao explains how proof checkers and AI programs are dramatically changing mathematics

BY CHRISTOPH DRÖSSER EDITED BY GARY STIX

Just_Super/Getty Images

AI-Driven Formal Theorem Proving

Code | Docs | Models | Dataset (Lean 3) | Dataset (Lean 4)

The Grand Challenge

The integration of artificial intelligence with formal mathematics presents a critical research challenge in bridging two fundamentally different computational paradigms. Large Language Models demonstrate remarkable capabilities in mathematical reasoning and proof generation, yet suffer from inconsistencies and hallucinations that compromise logical reliability. Formal proof assistants such as Lean provide absolute verification through mechanized type theory, ensuring every mathematical statement is rigorously validated by a trusted kernel. **Our central ambition** is to combine the power of LLMs with Lean to produce more verifiable mathematics, code, and scientific reasoning for a wide range of downstream applications.

Our Research Program

Our laboratory has developed several different research projects that systematically explores different facets of AI-assisted theorem proving. This work is primarily driven by researchers at Caltech, under the leadership of Professor Anima Anandkumar.

https://www.scientificamerican.com/article/ai-will-become-mathematicians-co-pilot/
https://leandojo.org/

Basics of Logic

- Arguments, premises, conclusions
  - Argument: a set of sentences (premises) + a sentence (a conclusion)
  - **Validity:** argument is valid iff it's necessary that if all premises are true, the conclusion is true
  - **Soundness:** argument is sound iff valid & premises true
  - **Entailment:** when valid arguments, premises entail conclusion

Propositional Logic Basics

Logic Vocabulary:

- Sentences, symbols, connectives, parentheses
  - Symbols: P, Q, R, … (**atomic** sentences)
  - Connectives:

∧	and	[conjunction]
∨	or	[disjunction]
⇒	implies	[implication]
⇔	is equivalent	[biconditional]
¬	not	[negation]

  - Literal: P or negation ¬P

Propositional Logic Basics

Examples:

- $(P \lor Q) \Rightarrow S$
  - "If it is cold or it is raining, then I need a jacket"
- $Q \Rightarrow P$
  - "If it is raining, then it is cold"
- $\neg R$
  - "It is not hot"

Propositional Logic Basics

Several rules in place

- Precedence: $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$
- Use parentheses when needed
- Sentences: **well-formed** or not well-formed:
  - $P \Rightarrow Q \Rightarrow S$ **not well-formed (not associative!)**

Sentences & Semantics

- Sentences: built up from symbols with connectives

  - **Interpretation**: assigning True / False to symbols (a row in truth table)
  - **Semantics**: interpretations for which sentence evaluates to True
  - **Model**: (of a set of sentences) interpretation for which all sentences are True

Another kind of model :)

Evaluating a Sentence

- Example:

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
$false$	$false$	$true$	$false$	$false$	$true$	$true$
$false$	$true$	$true$	$false$	$true$	$true$	$false$
$true$	$false$	$false$	$false$	$true$	$false$	$false$
$true$	$true$	$false$	$true$	$true$	$true$	$true$

- Note:
  - P ⇒ Q equivalent to ¬PVQ
  - If P is false, P ⇒ Q is true regardless of Q ("5 is even implies 6 is odd" is True!)
  - Causality not needed: "5 is odd implies the Sun is a star" is True!)

Evaluating a Sentence: Truth Table

- **Ex**:

P	Q	R	¬P	Q∧R	¬P∨Q∧R	¬P∨Q∧R⇒Q
0	0	0	1	0	1	0
0	0	1	1	0	1	0
0	1	0	1	0	1	1
0	1	1	1	1	1	1
1	0	0	0	0	0	1
1	0	1	0	0	0	1
1	1	0	0	0	0	1
1	1	1	0	1	1	1

- **Satisfiable**
  - There exists some interpretation where the sentence is true.

Break & Quiz

Q 1.1: Suppose P is false, Q is true, and R is true. Does this assignment satisfy

(i) $\neg(\neg p \rightarrow \neg q) \wedge r$

(ii) $(\neg p \vee \neg q) \rightarrow (p \vee \neg r)$

- A. Both
- B. Neither
- C. Just (i)
- D. Just (ii)

Break & Quiz

Q 1.1: Suppose P is false, Q is true, and R is true. Does this assignment satisfy

(i) $\neg(\neg p \rightarrow \neg q) \wedge r$

(ii) $(\neg p \vee \neg q) \rightarrow (p \vee \neg r)$

- A. Both
- B. Neither
- **C. Just (i)**
- D. Just (ii)

Break & Quiz

Q 1.1: Suppose P is false, Q is true, and R is true. Does this assignment satisfy

(i) $\neg(\neg p \rightarrow \neg q) \wedge r$

(ii) $(\neg p \vee \neg q) \rightarrow (p \vee \neg r)$

- A. Both
- B. Neither
- **C. Just (i)**
- D. Just (ii)

Plug interpretation into each sentence.

For (i): $(\neg p \rightarrow \neg q)$ will be false so $\neg(\neg p \rightarrow \neg q)$ will be true and r is true by assignment.

For (ii): $(\neg p \vee \neg q)$ is true and $(p \vee \neg r)$ is false which makes the implication false.

Break & Quiz

Q 1.2: Let A =“Aldo is Italian” and B =“Bob is English”. Formalize “Aldo is Italian or if Aldo isn’t Italian then Bob is English”.

- a. $A \lor (\neg A \rightarrow B)$
- b. $A \lor B$
- c. $A \lor (A \rightarrow B)$
- d. $A \rightarrow B$

Break & Quiz

Q 1.2: Let A =“Aldo is Italian” and B =“Bob is English”. Formalize “Aldo is Italian or if Aldo isn’t Italian then Bob is English”.

- **a. A V (¬A → B)**
- **b. A V B (equivalent!)**
- c. A V (A → B)
- d. A → B

Break & Quiz

Q 1.2: Let A =“Aldo is Italian” and B =“Bob is English”. Formalize “Aldo is Italian or if Aldo isn’t Italian then Bob is English”.

- **a. A V (¬A → B)**
- **b. A V B (equivalent!)**
- c. A V (A → B)
- d. A → B

Answer a. is the exact translation of the English sentence into a logic sentence. You can see that answer b. is also correct by writing out the truth table for all answers and seeing that a and b have the same truth tables.

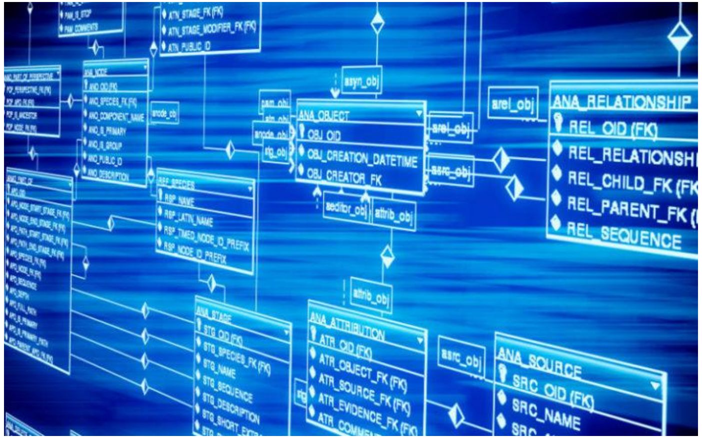
Or you can use the fact that $\neg A \rightarrow B = A \vee B$ and that $A \vee A \vee B = A \vee B$ to prove equivalence.

Knowledge Bases

- **Knowledge Base (KB)**: A set of sentences $\{A_1, \ldots, A_n\}$
  - Like a long sentence, connect with conjunction
  - KB: $A_1 \land A_2 \land \ldots \land A_n$

Model of a KB: interpretations where all sentences are True

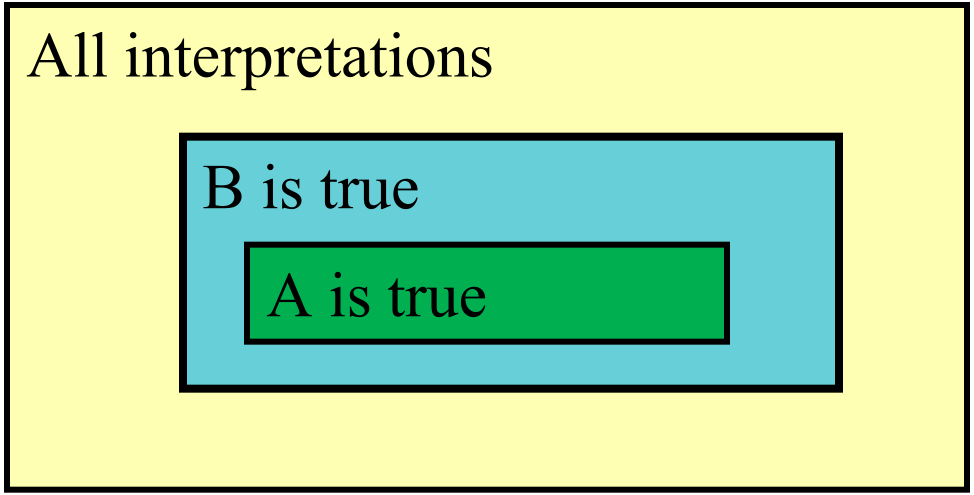
Goal: inference to discover new sentences



Entailment

Entailment: a sentence B logically follows from A

- Write A ⊨ B
- A ⊨ B iff in every interpretation where A is true, B is also true



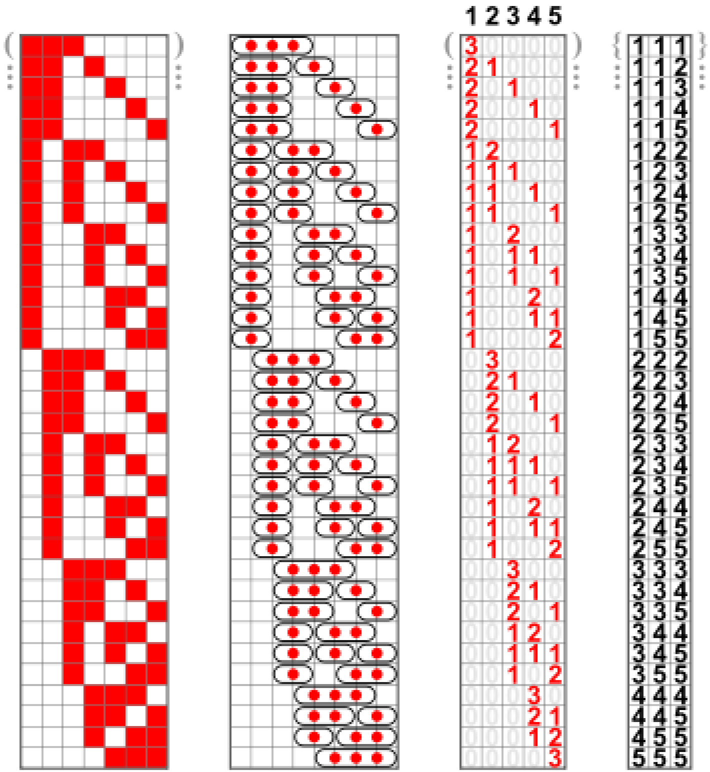
Inference

- Given a set of sentences (a KB), **logical inference** creates new sentences
  - Goal: Does KB entail sentence B ?
  - Compare to prob. inference!
- **Challenges**:
  - Soundness
  - Completeness
  - Efficiency

Methods of Inference: **1. Enumeration**

- Enumerate all interpretations; look at the truth table
  - "Model checking"

- Downside: 2^n interpretations for n symbols



Wiki

Methods of Inference: **2. Using Rules**

- *Modus Ponens*: $(A \Rightarrow B, A) \vDash B$
- And-elimination
- Other rules on the next page
  - Commutativity, associativity, de Morgan's laws, distribution for conjunction/disjunction

Logical equivalences

$$
\begin{aligned}
(\alpha \wedge \beta) &\equiv (\beta \wedge \alpha) && \text{commutativity of } \wedge \\
(\alpha \vee \beta) &\equiv (\beta \vee \alpha) && \text{commutativity of } \vee \\
((\alpha \wedge \beta) \wedge \gamma) &\equiv (\alpha \wedge (\beta \wedge \gamma)) && \text{associativity of } \wedge \\
((\alpha \vee \beta) \vee \gamma) &\equiv (\alpha \vee (\beta \vee \gamma)) && \text{associativity of } \vee \\
\neg(\neg\alpha) &\equiv \alpha && \text{double-negation elimination} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\beta \Rightarrow \neg\alpha) && \text{contraposition} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\alpha \vee \beta) && \text{implication elimination} \\
(\alpha \Leftrightarrow \beta) &\equiv ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)) && \text{biconditional elimination} \\
\neg(\alpha \wedge \beta) &\equiv (\neg\alpha \vee \neg\beta) && \text{de Morgan} \\
\neg(\alpha \vee \beta) &\equiv (\neg\alpha \wedge \neg\beta) && \text{de Morgan} \\
(\alpha \wedge (\beta \vee \gamma)) &\equiv ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma)) && \text{distributivity of } \wedge \text{ over } \vee \\
(\alpha \vee (\beta \wedge \gamma)) &\equiv ((\alpha \vee \beta) \wedge (\alpha \vee \gamma)) && \text{distributivity of } \vee \text{ over } \wedge
\end{aligned}
$$

You can use these equivalences to modify sentences.

Break & Quiz

Q 2.1: Which has more rows: a truth table on *n* symbols, or a joint distribution table on *n* binary random variables?

- A. Truth table
- B. Distribution
- C. Same size
- D. It depends

Break & Quiz

Q 2.1: Which has more rows: a truth table on *n* symbols, or a joint distribution table on *n* binary random variables?

- A. Truth table
- B. Distribution
- **C. Same size**
- D. It depends

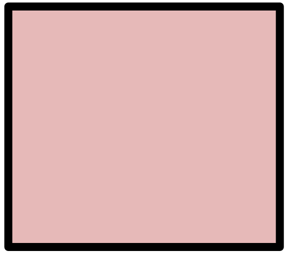
First Order Logic (FOL)

Propositional logic has some limitations

- Ex: how to say “all squares have four sides”
- No context, hard to generalize; express facts

FOL is a more expressive logic; works over

- Facts, Objects, Relations, Functions



First Order Logic Syntax

- **Term**: an object in the world
  - **Constant**: Alice, 2, Madison, Green, …
  - **Variables**: x, y, a, b, c, …
  - **Function**($term_1$, …, $term_n$)
    - Sqrt(9), Distance(Madison, Chicago)
    - Maps one or more objects to another object
    - Can refer to an unnamed object: LeftLeg(John)
    - Represents a user defined functional relation

- A **ground term** is a term without variables.
  - Constants or functions of constants

FOL Syntax

- **Atom**: smallest T/F expression
  - **Predicate**($term_1$, …, $term_n$)
    - Teacher(Blerina, you), Bigger(sqrt(2), x)
    - Convention: read “Blerina (is)Teacher(of) you”
    - Maps one or more objects to a truth value
    - Represents a user defined relation
  - $\mathbf{term_1 = term_2}$
    - Radius(Earth)=6400km, 1=2
    - Represents the equality relation when two terms refer to the same object

FOL Syntax

- **Sentence**: T/F expression
  - Atom
  - Complex sentence using connectives: $\land \lor \neg \Rightarrow \Leftrightarrow$
    - $Less(x,22) \land Less(y,33)$
  - Complex sentence using quantifiers $\forall, \exists$
- Sentences are evaluated under an interpretation

FOL Quantifiers

- Universal quantifier: ∀
- Sentence is true **for all** values of x in the domain of variable x.

- Main connective typically is ⇒
  - Forms if-then rules
  - “all humans are mammals”

$$\forall x\ \mathrm{human}(x) \Rightarrow \mathrm{mammal}(x)$$

  - Means if x is a human, then x is a mammal

FOL Quantifiers

- Existential quantifier: ∃
- Sentence is true **for some** value of x in the domain of variable x.

- Main connective typically is ∧
  - "some humans are male"

$$\exists x \; human(x) \; \wedge \; male(x)$$

  - Means there is an x who is a human and is a male

Break & Quiz

Q 2.1: How many entries does a truth table have for a FOL sentence with k variables where each variable can take on n values?

- A. Truth tables are not applicable to FOL.
- B. 2^k
- C. n^k
- D. It depends

Break & Quiz

Q 2.1: How many entries does a truth table have for a FOL sentence with k variables where each variable can take on n values?

- A. Truth tables are not applicable to FOL.
- B. 2^k
- C. $\boldsymbol{n^k}$
- D. It depends

Break & Quiz

Q 2.1: How many entries does a truth table have for a FOL sentence with k variables where each variable can take on n values?

- A. Truth tables are not applicable to FOL.
- B. 2^k
- C. $\boldsymbol{n^k}$
- D. It depends

Must have one entry for every possible assignment of values to variables. That number is (C).

Suggested Readings

- Textbook: *Artificial Intelligence: A Modern Approach (4th edition). Stuart Russell and Peter Norvig. Pearson, 2020.*
  - Chapters 7-9