香港浸會大學 HONG KONG BAPTIST UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE 計算機科學系

COMP7015 Artificial Intelligence

Lecture 4: Knowledge Representation and Reasoning

Instructor: Dr. Kejing Yin

September 29, 2022

Logistics

- Written Assignment 1 dues at 23:59 pm, Oct. 5.

- Scan your solutions into a single pdf file, name it using the following format:

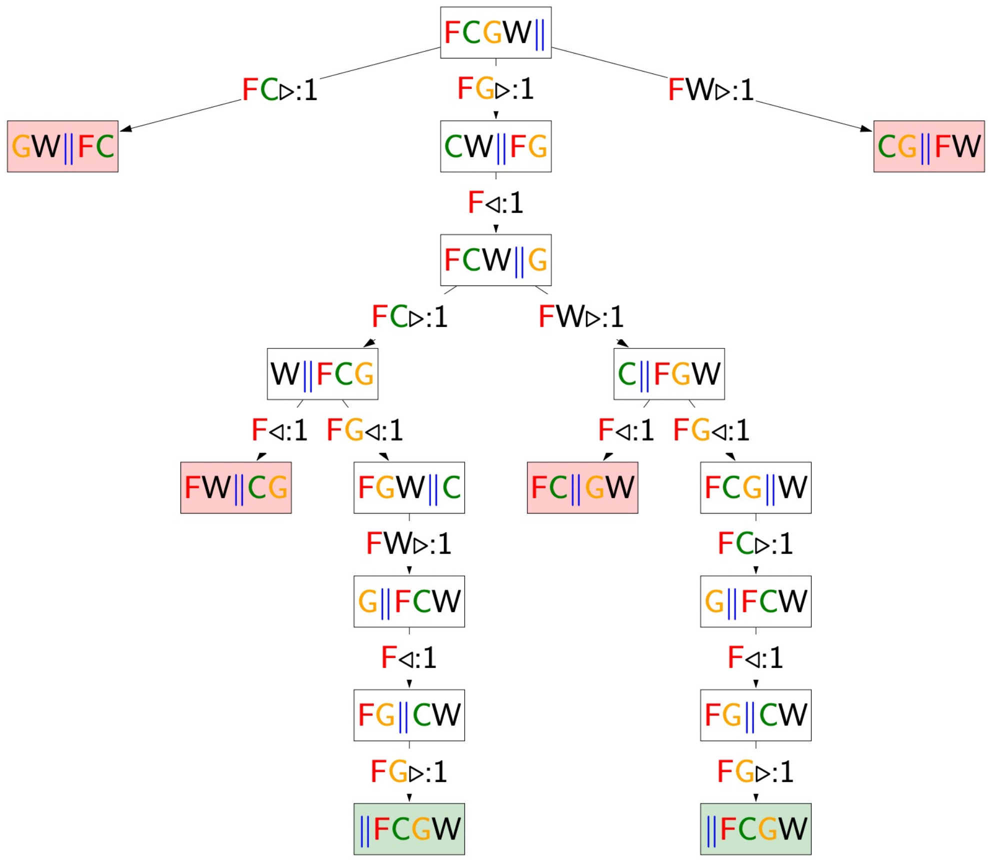
  wa1_<student id>.pdf   (*e.g.*, wa1_16483715.pdf)

- Any clarification needed?

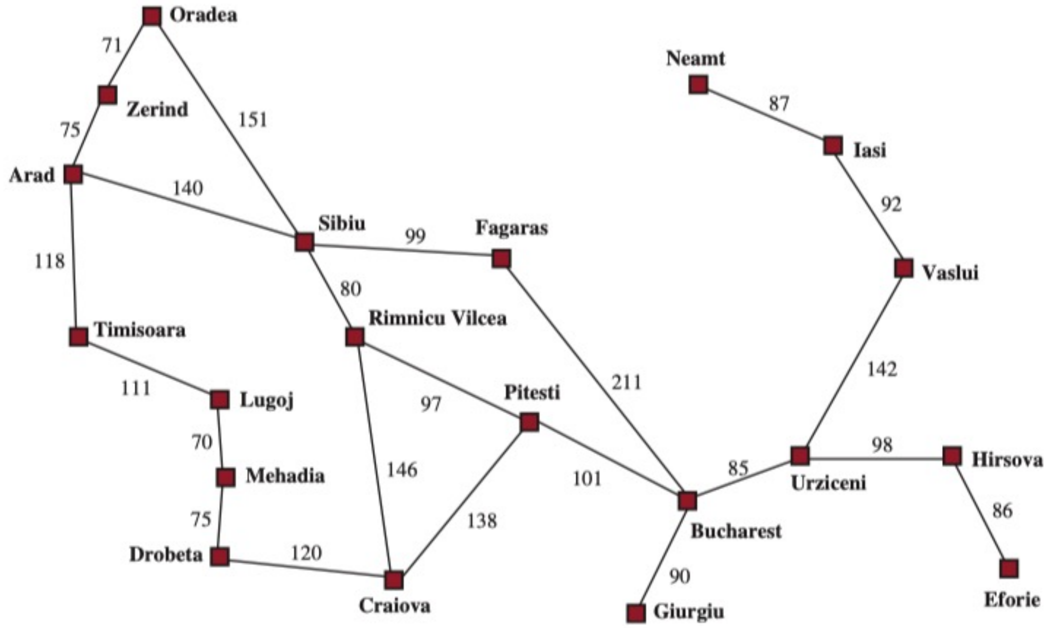
Logistics

- **Problem 1: drawing the search space**
  - Start from the initial state that you defines in Q1.
  - Draw out the resulting state of taking all possible actions. If any state violates the requirement, no need to further expand that node.

    N.B.: Nodes following the pink nodes are in the state space, but they are not “reachable”, so we can ignore them.
  - For repeated states: draw them out, or simply say that they have already been drawn.



Supplement: Level of abstraction in search problem formulation



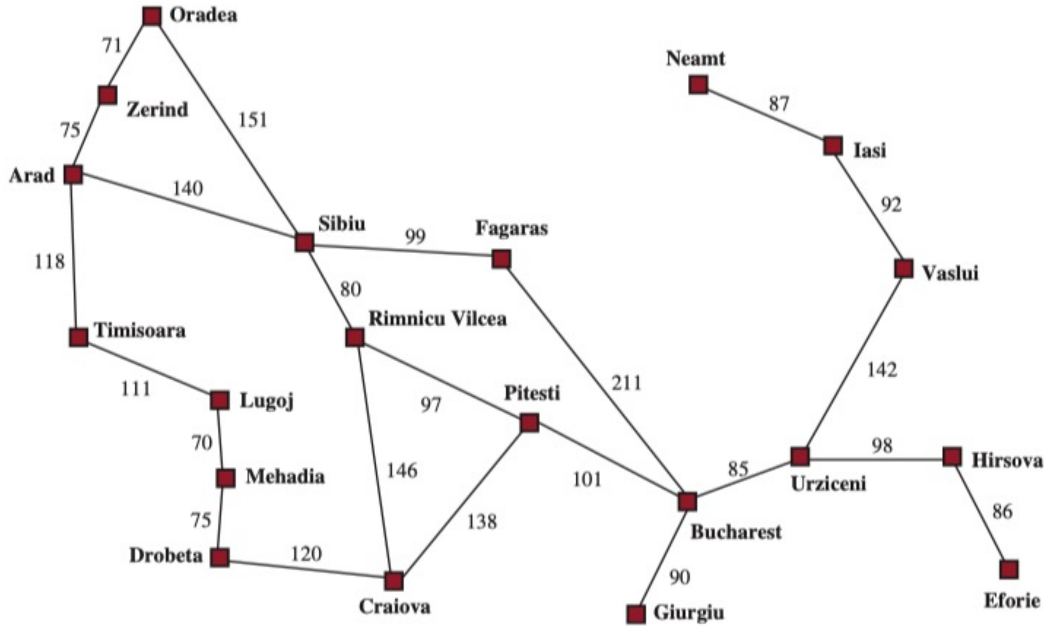
How do we formulate the path-finding problem?

e.g., states

- Option 1: *In Sibiu*, *In Fagaras*, …
- Option 2: *In Sibiu driving a red sedan*, *In Fagaras driving a white SUV with a pet,* …

Which one makes more sense to you?

Supplement: Level of abstraction in search problem formulation



How do we formulate the path-finding problem?

e.g., actions

- Option 1: *Go(Sibiu)*, *Go(Fagaras)*, …
- Option 2: *Turn on the car*, *release the brake*, *accelerate forward*, …

Which one makes more sense to you?

Supplement: Level of abstraction in search problem formulation

- **Abstraction**: the process of removing detail from a representation.

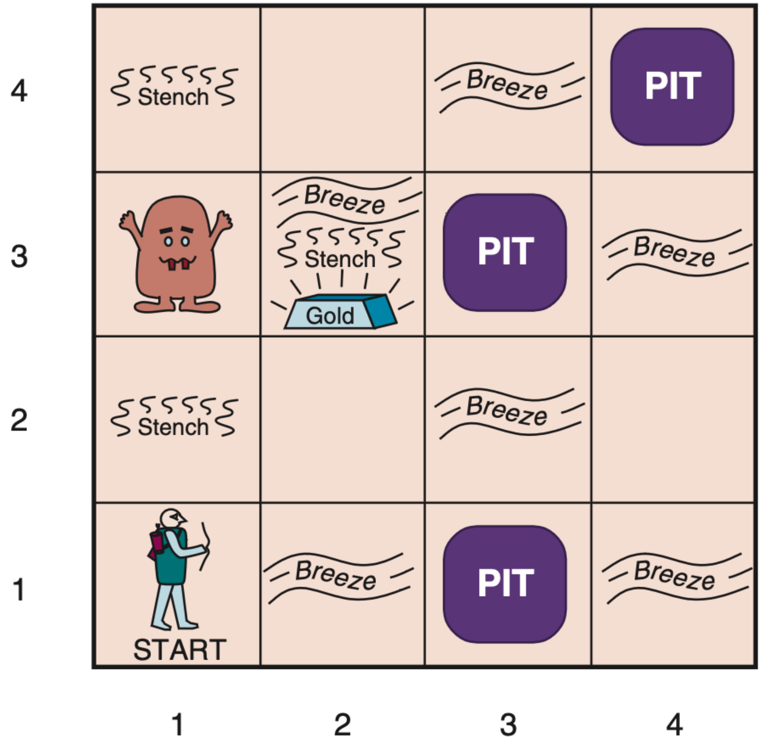
- A good problem formulation has the right level of detail. If we use option 2 to formulate the problem, we probably could never find the way out.
- Depends on the problem, e.g., in path-finding:
  - Driving a red car vs. driving a black car: no difference in general
  - Driving a car vs. taking a bus: there could be some difference

- *A rule of thumb: remove as much detail as possible and make only those distinctions necessary to ensure a valid solution.*

Let's play a game first: Wumpus World

- **Scores:**
  - +1000 for grabbing the gold;
  - -1000 for falling into a pit or being eaten by the wumpus;
  - -10 for each action taken.

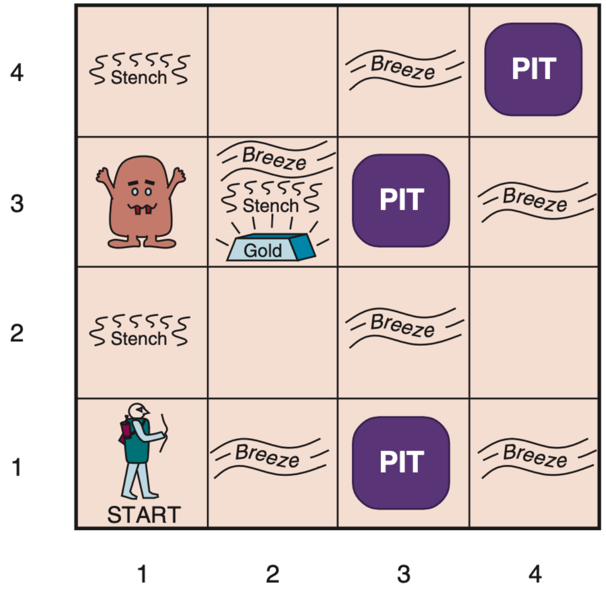
- The game **ends** when the agent either dies or climbs out of the cave.
- The agent could shoot an arrow to kill the wumpus.
- The agent can smell the stench around the wumpus.
- The agent can feel the breeze around the wumpus.

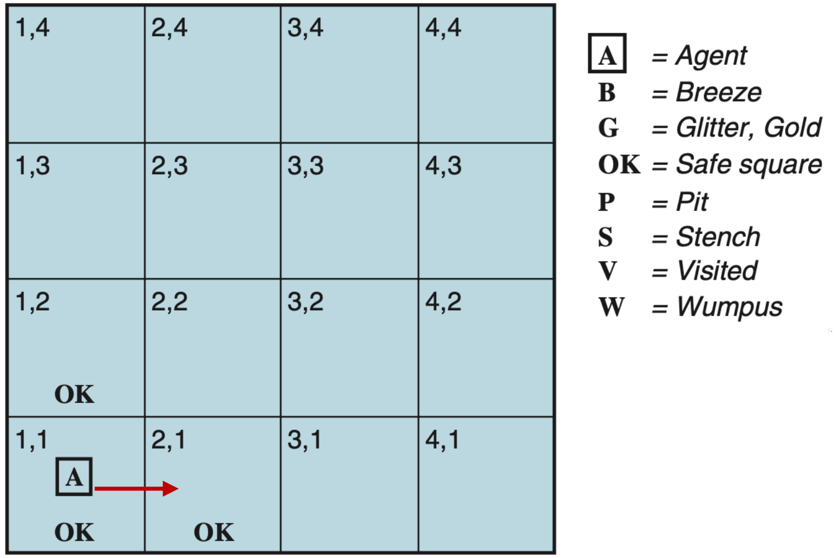


https://thiagodnf.github.io/wumpus-world-simulator/

Let’s play a game first: Wumpus World

- **How did we make decisions? Consider a simpler 4x4 case:**

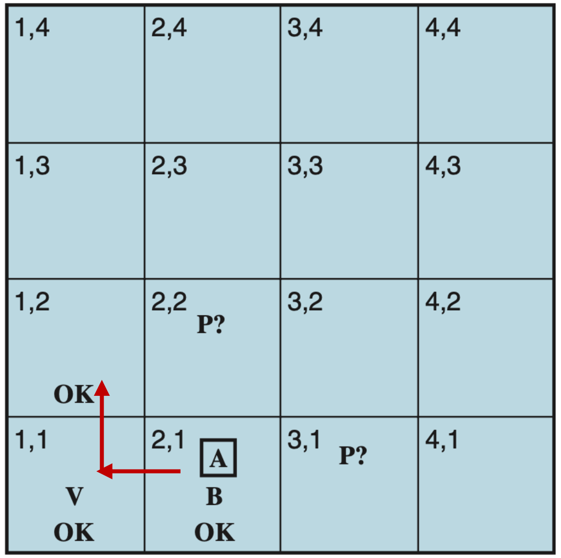




(a)

Initially at (1,1)

(1,1) is safe
→ (1,2) and (2,1) are safe



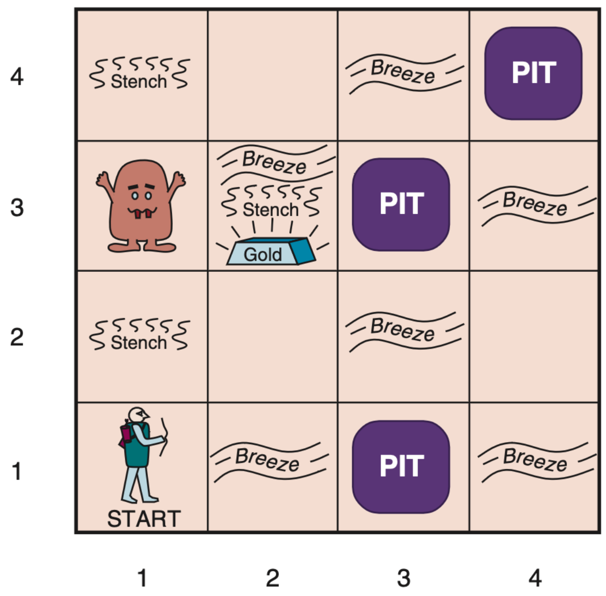
(b)

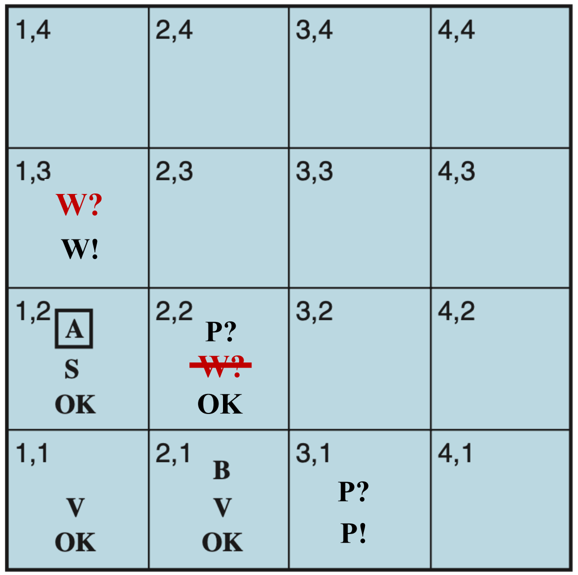
Move to (2,1)

Breeze at (2,1)
→ a pit at (2,2) and/or (3,1)

Let’s play a game first: Wumpus World

- **How did we make decisions? Consider a simpler 4x4 example:**





(c)

A = *Agent*
B = *Breeze*
G = *Glitter, Gold*
OK = *Safe square*
P = *Pit*
S = *Stench*
V = *Visited*
W = *Wumpus*

Move to (1,2)

Stench at (1,2) → Wumpus at (1,3) and/or (2,2)

No stench at (2,1) → No Wumpus at (2,2) → Wumpus at (1,3)

No breeze at (1, 2) → No pit at (2,2) → Pit at (3, 1)

Logic Reasoning

Draw conclusion from available information

The conclusion is correct if the available information is correct.

Another Motivating Example

- *Example of logic-based models: The virtual assistant*

It's raining now.
If it rains, the ground gets wet.
If the ground gets wet, it is slippery.
(knowledge)

tell

ask

Is it slippery now?

answer

Yes, it is slippery.

(inference)

Understand the information

Reason using the information

Logic Representation and Reasoning

- **Goal:** To enable the intelligent agent to represent and store information and derive conclusions from the available information.

Lecture Outline:

- Introduction to Logic
- Propositional Logic
- First-order Logic

Part I: Introduction to Logics

How do we represent knowledge?

- **Knowledge bases consist of sentences.**

Knowledge base

A dime is better than a nickel.

It it is raining, it is wet.

All students like COMP7015.

It is raining now.

If the Wumpus is at (1, 3), you can smell stench at (1, 2)

Inference:

All students like COMP7015.

Tom does not like COMP7015.

$\Rightarrow$ Tom is not a student.

How do we represent knowledge?

- **Is natural language a good choice?**

A dime is better than a nickel.
A nickel is better than a penny.
→ A dime is better than a penny.

A penny is better than nothing.
Nothing is better than world peace.
??? → A penny is better than world peace.

Natural language can be slippery

- **Logical language**: <u>precise</u> and <u>suitable to capture declarative knowledge</u>.
    - Propositional logic
    - First-order logic

Ingredients of logic: Syntax, Semantics, and Inference Rules

Syntax

Syntax defines a set of valid formulas (Formulas)

What are valid expressions in the language?

Semantics

For each formula, specify a set of **models** (assignments/ configurations of the word)

What do these expressions mean?

Inference rules

Given f, what new formulas g can be added that are guaranteed to follow?

Ingredients of logic: Syntax, Semantics, and Inference Rules

Syntax

Syntax defines a set of valid formulas (Formulas)

What are valid expressions in the language?

Examples:

- In English: "Tom ate an apple." (valid), "Tom an apple ate." (invalid)
- In arithmetic: x + y = 4 (valid), x4y+ = (invalid)
- In propositional logic: Rain ∧ Wet (valid), Rain + Wet (invalid)

Ingredients of logic: Syntax, Semantics, and Inference Rules

Semantics

Semantics defines the truth of each sentence with respect to each *possible world*.

What do these expressions mean?

Examples:

- The semantics for arithmetic specifies that the sentence "$x + y = 4$" is true in a world where x is 2 and y is 2, but false in a world where x is 1 and y is 1.
- In standard logics, every sentence must be either true or false in each possible world—there is no "in between."

Ingredients of logic: Syntax, Semantics, and Inference Rules

Inference rules

Given f, what new formulas g can be added that are guaranteed to follow?

Examples:

All students like COMP7015.

Tom does not like COMP7015.

$\Rightarrow$ Tom is not a student.

Ingredients of logic: Syntax, Semantics, and Inference Rules

Syntax

Syntax defines a set of valid formulas (Formulas)

What are valid expressions in the language?

Semantics

Semantics defines the truth of each sentence with respect to each *possible world*.

What do these expressions mean?

Inference rules

Given f, what new formulas g can be added that are guaranteed to follow?

Example: from Rain ∧ Wet, derive Rain

Logics

Higher expressivity

Proposition Logic

First-order Logic

Second-order Logic

...

Higher computational efficiency

Part II: Propositional Logics

Syntax of Propositional Logic

Building blocks: propositional symbols & connectives

- Propositional symbols (atomic formulas; atoms): A, B, C, …
- Logical connectives: $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$
- Build up formulas recursively: if A and B are formulas, so are the following:
  - Negation (not): $\neg A$
  - Conjunction (and): $A \wedge B$ *Symbol ∧ Looks like "A" for "And"*
  - Disjunction (or): $A \vee B$
  - Implication (implies): $A \Rightarrow B$
  - Biconditional (if and only if): $A \Leftrightarrow B$

Syntax of Propositional Logic

- Operator precedence: $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$
- Example: $\neg A \wedge B$ is equivalent to $(\neg A) \wedge B$ rather than $\neg(A \wedge B)$.

- When appropriate, we use <u>parentheses</u> and <u>square brackets</u> to clarify the intended sentence structure and improve readability.

- Note: They are pure symbols without any actual meaning. When we talk about syntax, <u>we are not talking about what they mean</u>. Semantics defines what the symbols mean.

Semantics of Propositional Logic

Fundamental Concept: **Models**

A model m in propositional logic is an assignment of truth values to propositional symbols.

Example:

- 3 propositional symbols: A, B, C
- $2^3 = 8$ possible models:

$$m_1 = \{A: 0, B: 0, C: 0\}$$
$$m_2 = \{A: 0, B: 0, C: 1\}$$
$$m_3 = \{A: 0, B: 1, C: 0\}$$
$$m_4 = \{A: 0, B: 1, C: 1\}$$
$$m_5 = \{A: 1, B: 0, C: 0\}$$
$$m_6 = \{A: 1, B: 0, C: 1\}$$
$$m_7 = \{A: 1, B: 1, C: 0\}$$
$$m_8 = \{A: 1, B: 1, C: 1\}$$

1: true
0: false

Semantics of Propositional Logic

Fundamental Concept: **Satisfaction**

If a sentence/formula f is true in model m, we say that m satisfies f, or we can say that m is a model of f.

We use the notation $M(f)$ to mean the set of all models of f.

Example: 3 atoms: A, B, C; 8 possible models.

$$m_1 = \{A:0, B:0, C:0\}$$
$$m_2 = \{A:0, B:0, C:1\}$$
$$m_3 = \{A:0, B:1, C:0\}$$
$$m_4 = \{A:0, B:1, C:1\}$$
$$m_5 = \{A:1, B:0, C:0\}$$
$$m_6 = \{A:1, B:0, C:1\}$$
$$m_7 = \{A:1, B:1, C:0\}$$
$$m_8 = \{A:1, B:1, C:1\}$$

1: true
0: false

f_1="A is true"

m_5 satisfies α_1;
m_6 satisfies α_1;
m_7 satisfies α_1;
m_8 satisfies α_1;

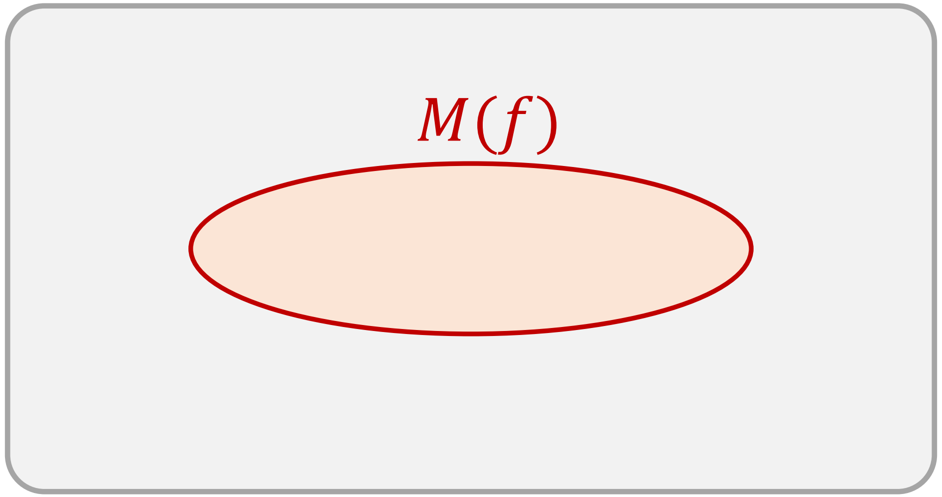
$$M(f_1) = \{m_5, m_6, m_7, m_8\}$$

Semantics of Propositional Logic

Fundamental Concept: **Satisfaction**

If a sentence/formula f is true in model m, we say that m satisfies f, or we can say that m is a model of f.

We use the notation $M(f)$ to mean the set of all models of f.



All possible models (possible worlds)

Semantics of Propositional Logic

- The semantics defines the rules for determining the truth of a sentence with respect to a particular model.

- In propositional logic, all sentences are constructed from atomic sentences and the five connectives. Therefore, we need to specify:
  1) how to compute the truth of atomic sentences and
  2) how to compute the truth of sentences formed with the connectives.

Semantics of Propositional Logic

- Atomic sentences are easy:
  - True (or 1) is true in every model.
  - False (or 0) is false in every model.

  - The truth value of every other proposition symbol must be specified directly in the model.
    E.g., in the model $m_5 = \{A: 1, B: 0, C: 0\}$*, A is true, B is false, and C is false.*

Semantics of Propositional Logic

- For complex sentences, five rules hold for any subsentences P and Q, *being them atomic or complex sentences*, in any model m.

1) $\neg P$ is true iff P is false in m.
2) $P \wedge Q$ is true iff both P and Q are true in m.
3) $\mathrm{P} \vee Q$ is true iff either P or Q is true in m.
4) $\mathrm{A} \Rightarrow B$ is true unless P is true and Q is false in m.
5) $A \Leftrightarrow B$ is true iff P and Q are both true or both false in m.

Semantics of Propositional Logic

- Truth Table

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*

Counter-intuitive: think $P \Rightarrow Q$ as saying,

"If P is true, then I am claiming that Q is true; otherwise, I am making no claim."

- "5 is even implies Sam is smart" is true, regardless of whether Sam is smart.

- Propositional logic does not require any relation of causation or relevance. "5 is odd implies Tokyo is the capital of Japan" is a true sentence of propositional logic.

Semantics of Propositional Logic

- Truth Table

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*

Bidirectional: $P \Leftrightarrow Q$ is true whenever both $P \Rightarrow Q$ and $Q \Rightarrow P$ are true.

Semantics of Propositional Logic

- Truth Table

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*

Example:

- The formula $f_2 = \neg A \wedge (B \vee C)$, evaluated in $m_2 = \{A: 0, B: 0, C: 1\}$, gives:

$$true \wedge (false \vee true) = true \wedge true = true$$

- Therefore, m_2 satisfies f_2.

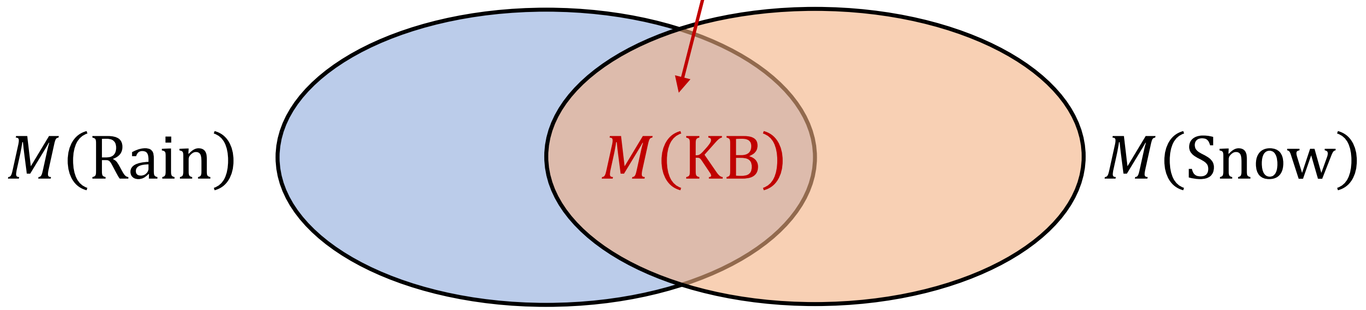
Knowledge Base

- A knowledge base KB is a set of formulas representing their intersection.

$$M(\mathrm{KB}) = \bigcap_{f \in \mathrm{KB}} M(f)$$

Example: KB = {Rain, Snow} ← *KB specifies constraints on the world.*

$M(KB)$ is the set of all worlds satisfying the constraints.



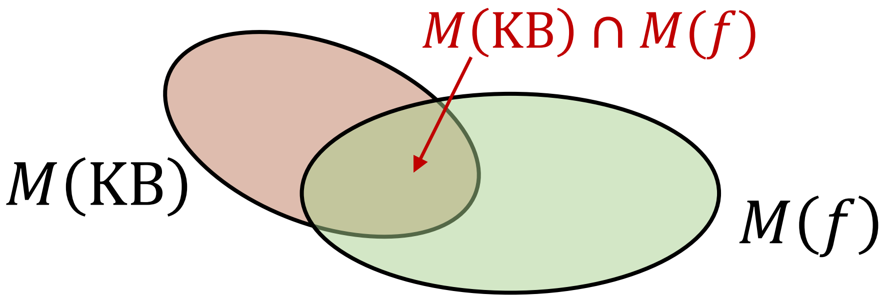
Knowledge Base: Adding knowledge

- Adding more formulas to the knowledge base:

$$\text{KB} \longrightarrow \text{KB} \cup \{f\}$$

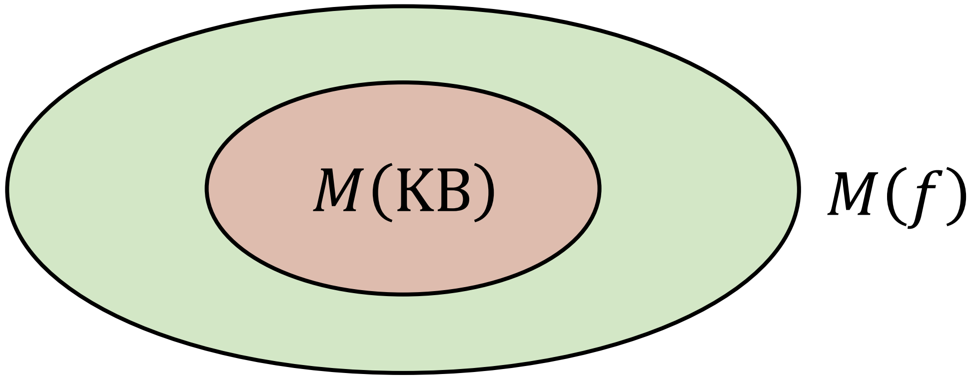
- Shrinks the set of models:

$$M(\text{KB}) \longrightarrow M(\text{KB}) \cap M(f)$$



How much does $M(\text{KB})$ shrink?

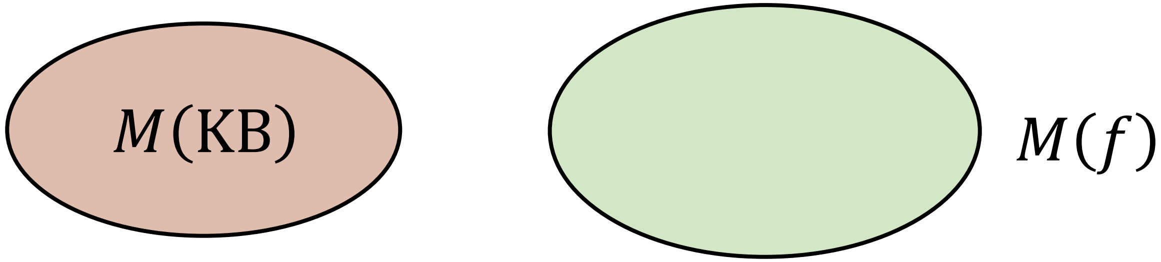
Knowledge Base: Adding knowledge (Entailment)



KB entails f (written KB $\models f$) iff $M(\mathrm{KB}) \subseteq M(f)$.

- f adds no information. It was already known.
- Example: Rain $\wedge$ Snow $\models$ Snow

Knowledge Base: Adding knowledge (Contradiction)

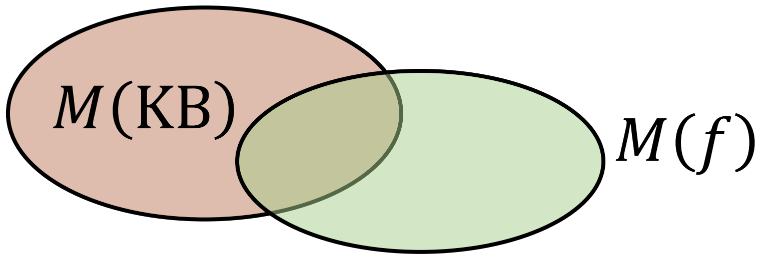


KB contradicts f iff $M(\mathrm{KB}) \cap M(f) = \emptyset$.

- f contradicts what we already know.
- Example: Rain ∧ Snow contradicts ¬Snow

Proposition: KB contradicts f iff KB entails $\neg f$.

Knowledge Base: Adding knowledge (Contingency)



$$\emptyset \subsetneq \quad M(\mathrm{KB}) \cap M(f) \quad \subsetneq M(\mathrm{KB})$$

- f adds non-trivial information to KB.
- Example: KB={Rain}, f=Snow

Knowledge Base: Tell operation

“It’s raining now.”

Tell[Rain]

?

- Possible Responses:
  - Already knew that: entailment (KB $\vDash f$)
  - Don’t believe that: contradiction (KB $\vDash \neg f$)
  - Learns something new (update KB): contingent;

Knowledge Base: Ask operation

“Is it raining now?”

Ask[Rain]

?

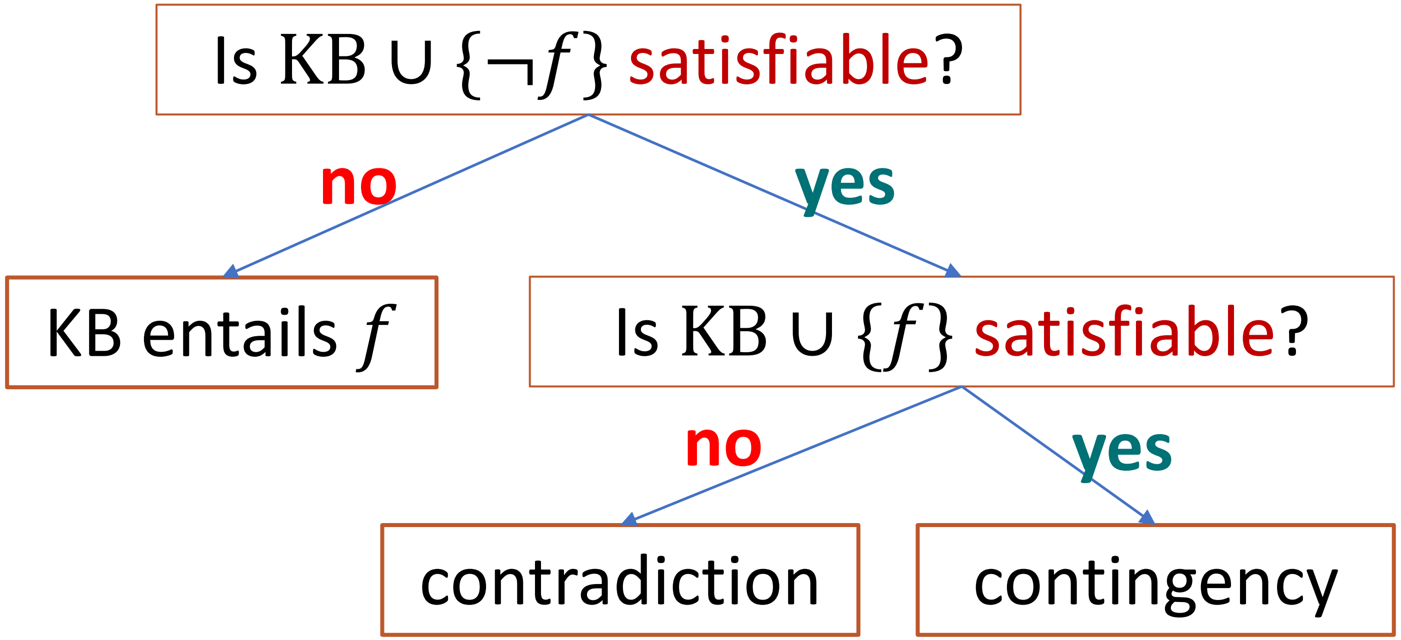
KB

- Possible Responses:
  - Yes: entailment (KB $\vDash f$)
  - No: contradiction (KB $\vDash \neg f$)
  - I don’t know: contingent;

Knowledge Base: Satisfiability

A knowledge base KB is satisfiable if $M(\mathrm{KB}) \neq \emptyset$.

- KB is satisfiable if there is some model that satisfies all formulas in KB.
- Reduce Tell[f] and Ask[f] to satisfiability:



Ingredients of logic: Syntax, Semantics, and Inference Rules

Inference rules

Given f, what new formulas g can be added that are guaranteed to follow?

Examples:

All students like COMP7015.

Tom does not like COMP7015.

→ Tom is not a student.

Formal definition:

If $f_1, \dots, f_k, g$ are formulas, then the following is an inference rule:

$$\frac{f_1, \dots, f_k, g}{g} \qquad \frac{\textit{(premises)}}{\textit{(conclusion)}}$$

Rules operate directly on syntax, not on semantics.

Inference Rules of Propositional Logic

- **Modus Ponens Inference Rule**

For any propositional symbols f and g:

$$\frac{f, \quad f \Rightarrow g}{g}$$

$$\frac{\textit{(premises)}}{\textit{(conclusion)}}$$

Example:

- It is raining (Rain)
- If it is raining, then it is wet. (Rain ⇒ Wet)
- Therefore, it is wet. (Wet)

$$\frac{\text{Rain}, \quad \text{Rain} \Rightarrow \text{Wet}}{\text{Wet}}$$

Inference Rules of Propositional Logic

- **Resolution Inference Rule**

$$\frac{f \lor g, \neg g \lor h}{f \lor h}$$

Or more generally,

$$\frac{f_1 \lor \cdots \lor f_n \lor g, \quad \neg g \lor h_1 \lor \cdots \lor h_m}{f_1 \lor \cdots \lor f_n \lor h_1 \lor \cdots \lor h_m}$$

Example:

- It is raining, or it is snowing (Rain ∨ Snow)
- It is not snowing, or there is traffic. (¬Snow ∨ Traffic)
- Therefore, it is raining, or there is traffic. (Rain ∨ Traffic)

Inference Rules of Propositional Logic

- **Modus Ponens** $\dfrac{f, \;\; f \Rightarrow g}{g}$
- **And-Elimination** $\dfrac{f_1 \wedge f_2 \wedge \cdots \wedge f_n}{f_i}$
- **And-Introduction** $\dfrac{f_1, f_2, \cdots, f_n}{f_1 \wedge f_2 \wedge \cdots \wedge f_n}$
- **Or-Introduction** $\dfrac{f_i}{f_1 \vee f_2 \vee \cdots \vee f_n}$
- **Resolution** $\dfrac{f \vee g, \neg g \vee h}{f \vee h}$
- **Double-Negation Elimination**

$$\frac{\neg\neg f}{f}$$

- **Unit Resolution**

$$\frac{f \vee g, \neg g}{f}$$