6.825 Techniques in Artificial Intelligence

Logic

- When we have too many states, we want a convenient way of dealing with sets of states.
- The sentence "It's raining" stands for all the states of the world in which it is raining.
- Logic provides a way of manipulating big collections of sets by manipulating short descriptions instead.
- Instead of thinking about all the ways a world could be, we're going to work in the a language of expressions that describe those sets.

What is a logic?

- A formal language
  - Syntax – what expressions are legal
  - Semantics – what legal expressions mean
  - Proof system – a way of manipulating syntactic expressions to get other syntactic expressions (which will tell us something new)
- Why proofs? Two kinds of inferences an agent might want to make:
  - Multiple percepts => conclusions about the world
  - Current state & operator => properties of next state

Propositional Logic Syntax

Syntax: what you're allowed to write

- **for (thing t = fizz; t == fuzz; t++){ … }**
- *Colorless green ideas sleep furiously*.

Sentences (wffs: well formed formulas)

- <u>true</u> and <u>false</u> are sentences
- Propositional variables are sentences: P,Q,R,Z
- If ϕ and ψ are sentences, then so are (ϕ), $\neg\phi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \rightarrow \psi$, $\phi \leftrightarrow \psi$
- Nothing else is a sentence

Precedence

Operator	
$\neg$	highest
$\wedge$	
$\vee$	
$\rightarrow$	
$\leftrightarrow$	lowest

$A \vee B \wedge C$	$A \vee (B \wedge C)$
$A \wedge B \rightarrow C \vee D$	$(A \wedge B) \rightarrow (C \vee D)$
$A \rightarrow B \vee C \leftrightarrow D$	$(A \rightarrow (B \vee C)) \leftrightarrow D$

- Precedence rules enable "shorthand" form of sentences, but formally only the fully parenthesized form is legal.
- Syntactically ambiguous forms allowed in shorthand only when semantically equivalent: $A \wedge B \wedge C$ is equivalent to $(A \wedge B) \wedge C$ and $A \wedge (B \wedge C)$

Recitation Exercises: Part 1

Which of these are legal sentences?

$$P \rightarrow Q \rightarrow R$$
$$P, R \rightarrow Q$$
$$A \wedge (B \vee C \vee \neg D) \leftrightarrow \neg\neg Z$$
$$\neg P(Q)$$

Give fully parenthesized expressions for the legal sentences. (If there is more than one solution, just pick any one).

Semantics

- Meaning of a sentence is truth value {**t, f**}
- Interpretation is an assignment of truth values to the propositional variables
- $\models_i \phi$ [Sentence ϕ is **t** in interpretation i]
- $\not\models_i \phi$ [Sentence ϕ is **f** in interpretation i]

Semantic Rules

- $\models_i$ <u>true</u> for all i
- $\not\models_i$ <u>false</u> for all i [the sentence <u>false</u> has truth value **f** in all interpret.]
- $\models_i \neg\phi$ if and only if $\not\models_i \phi$
- $\models_i \phi \wedge \psi$ if and only if $\models_i \phi$ and $\models_i \psi$ [conjunction]
- $\models_i \phi \vee \psi$ if and only if $\models_i \phi$ or $\models_i \psi$ [disjunction]
- $\models_i P$ iff i(P) = **t**

Some important shorthand

Some important shorthand

- $\phi \rightarrow \psi \equiv \neg \phi \vee \psi$ [conditional, implication]

  antecedent → consequent
- $\phi \leftrightarrow \psi \equiv (\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$ [biconditional, equivalence]

Truth Tables

P	Q	¬ P	P ∧ Q	P ∨ Q	P → Q	Q → P	P ↔ Q
f	**f**	**t**	**f**	**f**	**t**	**t**	**t**
f	**t**	**t**	**f**	**t**	**t**	**f**	**f**
t	**f**	**f**	**f**	**t**	**f**	**t**	**f**
t	**t**	**f**	**t**	**t**	**t**	**t**	**t**

Note that implication is not "causality", if P is **f** then P → Q is **t**

Terminology

- A sentence is valid iff its truth value is **t** in all interpretations ($\models \phi$)

  Valid sentences: <u>true</u>, ¬ <u>false</u>, P ∨ ¬ P
- A sentence is satisfiable iff its truth value is **t** in at least one interpretation

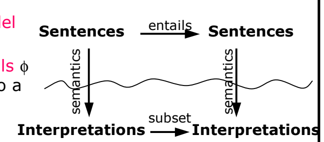
  Satisfiable sentences: P, <u>true</u>, ¬ P
- A sentence is unsatisfiable iff its truth value is **f** in all interpretations

  Unsatisfiable sentences: P ∧ ¬ P, <u>false</u>, ¬ <u>true</u>

All are finitely decidable.

Models and Entailment

- An interpretation i is a model of a sentence ϕ iff $\models_i \phi$
- A set of sentences KB entails ϕ iff every model of KB is also a model of ϕ

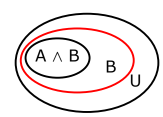


$KB \models \phi$ iff $\models KB \rightarrow \phi$

KB entails ϕ if and only if (KB → ϕ) is valid

KB = A ∧ B
ϕ = B

A ∧ B ⊨ B



Examples

Sentence	Valid?	Interpretation that make sentence's truth value = **f**
smoke → smoke smoke ∨ ¬smoke	valid	
smoke → fire	satisfiable, not valid	smoke = **t**, fire = **f**
(s → f) → (¬ s → ¬ f)	satisfiable, not valid	s = **f**, f = **t** s → f = **t**, ¬ s → ¬ f = **f**
contrapositive (s → f) → (¬ f → ¬ s)	valid	
b ∨ d ∨ (b → d) b ∨ d ∨ ¬ b ∨ d	valid	

Recitation Exercises: Part II

For each of the following sentences, say whether it is valid, unsatisfiable, or satisfiable but not valid. If it is neither valid nor unsatisfiable, provide an interpretation in which it is true and another in which it is false.

$(P \rightarrow Q) \rightarrow ((Q \rightarrow R) \rightarrow (P \rightarrow R))$

$(P \vee Q) \wedge (\neg P \vee R) \wedge (\neg R \vee Q)$

$\neg P \rightarrow (P \rightarrow Q)$

$(P \rightarrow Q) \wedge (Q \rightarrow P)$

$(P \rightarrow (Q \rightarrow R)) \rightarrow (P \rightarrow (R \rightarrow Q))$

Russell & Norvig problems 6.5 and 6.7