Foundations of Artificial Intelligence

29. Propositional Logic: Basics

Thomas Keller and Florian Pommerening

University of Basel

April 19, 2023

Classification

classification:

Propositional Logic

environment:

- static vs. dynamic
- deterministic vs. non-deterministic vs. stochastic
- fully vs. partially vs. not observable
- discrete vs. continuous
- single-agent vs. multi-agent

problem solving method:

- problem-specific vs. general vs. learning

(applications also in more complex environments)

Propositional Logic: Overview

Chapter overview: propositional logic

- 29. Basics
- 30. Reasoning and Resolution
- 31. DPLL Algorithm
- 32. Local Search and Outlook

Motivation

Propositional Logic: Motivation

propositional logic

- modeling and representing problems and knowledge
- basics for general problem descriptions and solving strategies ($\rightsquigarrow$ automated planning $\rightsquigarrow$ later in this course)
- allows for automated reasoning

Relationship to CSPs

- previous topic: constraint satisfaction problems
- satisfiability problem in propositional logic can be viewed as non-binary CSP over $\{\mathbf{F}, \mathbf{T}\}$
- formula encodes constraints
- solution: satisfying assignment of values to variables
- SAT algorithms for this problem: $\rightsquigarrow$ DPLL (next week)

Propositional Logic: Description of State Spaces

propositional variables for missionaries and cannibals problem:

```
two-missionaries-are-on-left-shore
one-cannibal-is-on-left-shore
boat-is-on-left-shore
...
```

- problem description for general problem solvers
- states represented as truth values of atomic propositions

Propositional Logic: Intuition

propositions: atomic statements over the world
that cannot be divided further

Propositions with logical connectives like
“and”, “or” and “not” form the propositional formulas.

Syntax and Semantics

Today, we define syntax and semantics of propositional logic.

Syntax

- defines which symbols are allowed in formulas
  $(,), \aleph, \wedge, A, B, C, X, \heartsuit, \rightarrow, \nearrow, \ldots ?$
- ... and which sequences of these symbols are correct formuals
  $(A \wedge B), \quad ((A) \wedge B), \quad \wedge)A(B, \ldots ?$

Semantics

- defines the meaning of formulas
- uses interpretations to describe a possible world
  $I_1 = \{A \mapsto \mathbf{T}, B \mapsto \mathbf{F}\}$
- defines which formulas are true in which worlds

Syntax

Syntax

Σ alphabet of propositions
(e.g., $\{P, Q, R, \dots\}$ or $\{X_1, X_2, X_3, \dots\}$).

Definition (propositional formula)

- $\top$ and $\bot$ are formulas.
- Every proposition in Σ is an (atomic) formula.
- If φ is a formula, then $\neg\varphi$ is a formula (negation).
- If φ and ψ are formulas, then so are
  - $(\varphi \land \psi)$ (conjunction)
  - $(\varphi \lor \psi)$ (disjunction)
  - $(\varphi \rightarrow \psi)$ (implication)

binding strength: $(\neg) > (\land) > (\lor) > (\rightarrow)$
$\rightsquigarrow$ (may omit redundant parentheses, use responsibly)
note: minor differences to Discrete Mathematics course

Semantics

Semantics

A formula is true or false,
depending on the interpretation of the propositions.

Semantics: Intuition

- A proposition p is either true or false.
  The truth value of p is determined by an interpretation.
- The truth value of a formula follows from the truth values of the propositions.

Example

$\varphi = (P \vee Q) \wedge R$

- If P and Q are false, then φ is false (independent of the truth value of R).
- If P and R are true, then φ is true (independent of the truth value of Q).

Semantics: Formally

- defined over interpretation $I : \Sigma \to \{\mathbf{T}, \mathbf{F}\}$
- interpretation I: assignment of propositions in Σ
- When is a formula φ true under interpretation I?
  symbolically: When does $I \models \varphi$ hold?

Semantics: Formally

Definition ($I \models \varphi$, read: "I satisfies φ" or "φ holds under I")

Let φ and ψ be propositional formulas over Σ.

- $I \models \top$ and $I \not\models \bot$
- $I \models P$ iff $I(P) = \mathbf{T}$ for $P \in \Sigma$
- $I \models \neg\varphi$ iff $I \not\models \varphi$
- $I \models (\varphi \wedge \psi)$ iff $I \models \varphi$ and $I \models \psi$
- $I \models (\varphi \vee \psi)$ iff $I \models \varphi$ or $I \models \psi$
- $I \models (\varphi \rightarrow \psi)$ iff $I \not\models \varphi$ or $I \models \psi$

Definition ($I \models \Phi$)

Let Φ be a set of propositional formulas

- $I \models \Phi$ iff $I \models \varphi$ for all $\varphi \in \Phi$

Examples

Example (Interpretation I)

$I = \{P \mapsto \mathbf{T}, Q \mapsto \mathbf{T}, R \mapsto \mathbf{F}, S \mapsto \mathbf{F}\}$

Which formulas are true under I?

- $\varphi_1 = \neg(P \wedge Q) \wedge (R \wedge \neg S)$. Does $I \models \varphi_1$ hold?
- $\varphi_2 = (P \wedge Q) \wedge \neg(R \wedge \neg S)$. Does $I \models \varphi_2$ hold?
- $\varphi_3 = (R \rightarrow P)$. Does $I \models \varphi_3$ hold?

Terminology

Definition (model)

An interpretation I is called a model of φ if $I \models \varphi$.

Definition (satisfiable etc.)

A formula φ is called

- satisfiable if there is an interpretation I such that $I \models \varphi$.
- unsatisfiable if φ is not satisfiable.
- falsifiable if there is an interpretation I such that $I \not\models \varphi$.
- valid (= a tautology) if $I \models \varphi$ for all interpretations I.

Terminology (Side Note)

What does "φ is true" mean?

- not formally defined
- implicit missing interpretation
  - could be meant as "in the obvious interpretation"
  - or as "in all possible interpretations" (tautology)
- imprecise language $\rightsquigarrow$ avoid

Terminology

Definition (logical equivalence)

Formulas φ and ψ are called logically equivalent ($\varphi \equiv \psi$) if for all interpretations I: $I \models \varphi$ iff $I \models \psi$.

Equivalences

Logical Equivalences

Let φ, ψ, and η be formulas.

- $(\varphi \wedge \psi) \equiv (\psi \wedge \varphi)$ and $(\varphi \vee \psi) \equiv (\psi \vee \varphi)$ (commutativity)
- $((\varphi \wedge \psi) \wedge \eta) \equiv (\varphi \wedge (\psi \wedge \eta))$ and $((\varphi \vee \psi) \vee \eta) \equiv (\varphi \vee (\psi \vee \eta))$ (associativity)
- $((\varphi \wedge \psi) \vee \eta) \equiv ((\varphi \vee \eta) \wedge (\psi \vee \eta))$ (distributivity)
- $(\varphi \rightarrow \psi) \equiv (\neg\varphi \vee \psi)$ (($\rightarrow$)-elimination)
- $\neg(\varphi \wedge \psi) \equiv (\neg\varphi \vee \neg\psi)$ (De Morgan)
- $\neg(\varphi \vee \psi) \equiv (\neg\varphi \wedge \neg\psi)$ (De Morgan)
- $\neg\neg\varphi \equiv \varphi$ (double negation)

Commutativity and associativity are often used implicitly
$\rightsquigarrow$ We write $(X_1 \wedge X_2 \wedge X_3 \wedge X_4)$ instead of $(X_1 \wedge (X_2 \wedge (X_3 \wedge X_4)))$

Truth Tables

Truth Tables

How to determine automatically if a given formula is (un)satisfiable, falsifiable, valid?

$\rightsquigarrow$ simple method: truth tables

example: Is $\varphi = ((P \lor H) \land \neg H) \rightarrow P$ valid?

P	H	$P \lor H$	$((P \lor H) \land \neg H)$	$((P \lor H) \land \neg H) \rightarrow P$
F	**F**	**F**	**F**	**T**
F	**T**	**T**	**F**	**T**
T	**F**	**T**	**T**	**T**
T	**T**	**T**	**F**	**T**

$I \models \varphi$ for all interpretations $I \rightsquigarrow \varphi$ is valid.

- satisfiability, falsifiability, unsatisfiability?

Normal Forms

Normal Forms: Terminology

Definition (literal)

If $P \in \Sigma$, then the formulas P and $\neg P$ are called literals.

P is called positive literal, $\neg P$ is called negative literal.

The complementary literal to P is $\neg P$ and vice versa.
For a literal ℓ, the complementary literal to ℓ is denoted with $\bar{\ell}$.

Question: What is the difference between $\bar{\ell}$ and $\neg\ell$?

Normal Forms: Terminology

Definition (clause)

A disjunction of 0 or more literals is called a clause.
The empty clause $\bot$ is also written as $\square$.
Clauses consisting of only one literal are called unit clauses.

Definition (monomial)

A conjunction of 0 or more literals is called a monomial.

Normal Forms

Definition (normal forms)

A formula φ is in conjunctive normal form (CNF, clause form) if φ is a conjunction of 0 or more clauses:

$$\varphi = \bigwedge_{i=1}^{n} \left(\bigvee_{j=1}^{m_i} \ell_{i,j} \right)$$

A formula φ is in disjunctive normal form (DNF) if φ is a disjunction of 0 or more monomials:

$$\varphi = \bigvee_{i=1}^{n} \left(\bigwedge_{j=1}^{m_i} \ell_{i,j} \right)$$

Normal Forms

For every propositional formula, there exists a logically equivalent propositional formula in CNF and in DNF.

Conversion to CNF with equivalences

1. eliminate implications
   $(\varphi \rightarrow \psi) \equiv (\neg\varphi \lor \psi)$ (($\rightarrow$)-elimination)
2. move negations inside
   $\neg(\varphi \land \psi) \equiv (\neg\varphi \lor \neg\psi)$ (De Morgan)
   $\neg(\varphi \lor \psi) \equiv (\neg\varphi \land \neg\psi)$ (De Morgan)
   $\neg\neg\varphi \equiv \varphi$ (double negation)
3. distribute $\lor$ over $\land$
   $((\varphi \land \psi) \lor \eta) \equiv ((\varphi \lor \eta) \land (\psi \lor \eta))$ (distributivity)
4. simplify trivial subformulas $(\top, \bot)$

There are formulas φ for which every logically equivalent formula in CNF and DNF is exponentially longer than φ.

Summary

Summary (1)

- Propositional logic forms the basis for a general representation of problems and knowledge.
- Propositions (atomic formulas) are statements over the world which cannot be divided further.
- Propositional formulas combine atomic formulas with $\neg$, $\wedge$, $\vee$, $\rightarrow$ to more complex statements.
- Interpretations determine which atomic formulas are true and which ones are false.

Summary (2)

- important terminology:
  - model
  - satisfiable, unsatisfiable, falsifiable, valid
  - logically equivalent
- different kinds of formulas:
  - atomic formulas and literals
  - clauses and monomials
  - conjunctive normal form and disjunctive normal form