First order theories

(Chapter 1, Sections 1.4 – 1.5)

First order logic

- A first order theory consists of
  - Variables
  - Logical symbols: $\wedge$ $\vee$ $\neg$ $\forall$ $\exists$ `(' `)'
  - Non-logical Symbols Σ: Constants, predicate and function symbols
  - Syntax

Examples

- $\Sigma = \{0, 1, \text{'+'}, \text{'>'}\}$
  - '0','1' are constant symbols
  - '+' is a binary function symbol
  - '>' is a binary predicate symbol

- An example of a Σ-formula:

$$\exists y\ \forall x.\ x > y$$

Examples

- $\Sigma = \{1, \text{'>'}, \text{'<'}, \text{'isprime'}\}$
  - '1' is a constant symbol
  - '>', '<' are binary predicates symbols
  - 'isprime' is a unary predicate symbol
- An example Σ-formula:

$$\forall n\ \exists p.\ n > 1 \rightarrow \text{isprime}(p) \wedge n < p < 2n.$$

- Are these formulas valid ?
- So far these are only symbols, strings. No meaning yet.

Interpretations

- Let Σ = {0,1, '+', '='} where 0,1 are constants, '+' is a binary function symbol and '=' a predicate symbol.
- Let $\phi = \exists x.\ x + 0 = 1$
- Q: Is ϕ true in $\mathcal{N}_0$?
- A: Depends on the interpretation!

Structures

- A structure is given by:
  1. A domain
  2. An interpretation of the nonlogical symbols: i.e.,
     - Maps each predicate symbol to a predicate of the same arity
     - Maps each function symbol to a function of the same arity
     - Maps each constant symbol to a domain element
  3. An assignment of a domain element to each free (unquantified) variable

Structures

- Remember $\phi = \exists x.\ x + 0 = 1$
- Consider the structure S:
  - Domain: $\mathcal{N}_0$
  - Interpretation:
    - '0' and '1' are mapped to 0 and 1 in $\mathcal{N}_0$
    - '=' $\mapsto$ = (equality)
    - '+' $\mapsto$ * (multiplication)
- Now, is ϕ true in S ?

Satisfying structures

- Definition: A formula is satisfiable if there exists a structure that satisfies it

- Example: $\phi = \exists x.\ x + 0 = 1$ is satisfiable
- Consider the structure S':
  - Domain: $\mathcal{N}_0$
  - Interpretation:
    - '0' and '1' are mapped to 0 and 1 in $\mathcal{N}_0$
    - '=' $\mapsto$ = (equality)
    - '+' $\mapsto$ + (addition)
- S' satisfies ϕ. S' is said to be a model of ϕ.

First-order theories

- First-order logic is a framework.
- It gives us a generic syntax and building blocks for constructing restrictions thereof.
- Each such restriction is called a first-order theory.

- A theory defines
  - the signature Σ (the set of nonlogical symbols) and
  - the interpretations that we can give them.

Definitions

- Σ – the signature. This is a set of nonlogical symbols.
- Σ-formula: a formula over Σ symbols + logical symbols.
- A variable is free if it is not bound by a quantifier.
- A sentence is a formula without free variables.
- A Σ-theory T is defined by a set of Σ-sentences.

Sentences are the axioms capturing the properties of the theory

Definitions…

- Let T be a Σ-theory
- A Σ-formula ϕ is T-satisfiable if there exists a structure that satisfies both ϕ and the sentences defining T.
- A Σ-formula ϕ is T-valid if all structures that satisfy the sentences defining T also satisfy ϕ.

Example

- Let $\Sigma = \{0,1, \text{'+'}, \text{'='}\}$
- Recall $\phi = \exists x.\ x + 0 = 1$
- ϕ is a Σ-formula.
- We now define the following Σ-theory:
  - $\forall x.\ x = x$ // '=' must be reflexive
  - $\forall x,y.\ x + y = y + x$ // '+' must be commutative
- Not enough to prove the validity of ϕ !

Theories through axioms

- The number of sentences that are necessary for defining a theory may be large or infinite.
- Instead, it is common to define a theory through a set of axioms.
- The theory is defined by these axioms and everything that can be inferred from them by a sound inference system.

Example 1

- Let $\Sigma = \{‘=’\}$
  - An example Σ-formula is $\phi = ((x = y) \wedge \neg (y = z)) \rightarrow \neg(x = z)$
- We would now like to define a Σ-theory T that will limit the interpretation of ‘=‘ to equality.
- We will do so with the equality axioms:
  - $\forall x.\ x = x$ (reflexivity)
  - $\forall x,y.\ x = y \rightarrow y = x$ (symmetry)
  - $\forall x,y,z.\ x = y \wedge y = z \rightarrow x = z$ (transitivity)
- Every structure that satisfies these axioms also satisfies ϕ above.
- Hence ϕ is T-valid.

Example 2

- Let Σ = {'<'}
- Consider the Σ-formula ϕ: $\forall x\ \exists y.\ y < x$
- Consider the theory T:
  - $\forall x,y,z.\ x < y \wedge y < z \rightarrow x < z$ (transitivity)
  - $\forall x,y.\ x < y \rightarrow \neg(y < x)$ (anti-symmetry)

Example 2 (cont'd)

- Recall: ϕ: $\forall x\ \exists y.\ y < x$
- Is ϕ T-satisfiable?
- We will show a model for it.
  - Domain: $\mathcal{Z}$
  - '<' $\mapsto$ <
- Is ϕ T-valid ?
- We will show a structure to the contrary
  - Domain: $\mathcal{N}_0$
  - '<' $\mapsto$ <

Fragments

a theory restricts only the nonlogical symbols. If we want to restrict the set of logical symbols or the grammar, we need to speak about fragments of the logic.

Fragments

- So far we only restricted the nonlogical symbols.
- Sometimes we want to restrict the grammar and the logical symbols that we can use as well.
- These are called logic fragments.
- Examples:
  - The quantifier-free fragment over Σ = {‘=’, ‘+’,0,1}
  - The conjunctive fragment over Σ = {‘=’, ‘+’,0,1}

What about restricting the interpretation of the logical symbols?

The axioms that restrict the interpretation of the logical symbols, called the logical axioms, are assumed to be "built in", i.e., they are common to all first-order theories.

Fragments

- Let Σ = {}
  - (T must be empty: no nonlogical symbols to interpret)
- Q: What is the quantifier-free fragment of T ?
- A: propositional logic
- Thus, propositional logic is also a first-order theory.
  - A very degenerate one.

Theories

- Let $\Sigma = \{\}$
  - (T must be empty: no nonlogical symbols to interpret)
- Q: What is T ?
- A: Quantified Boolean Formulas (QBF)

- Example:
  - $\forall x_1 \, \exists x_2 \, \forall x_3 . \; x_1 \rightarrow (x_2 \vee x_3)$

Some famous theories

- Presburger arithmetic: $\Sigma = \{0, 1, \text{'+'}, \text{'='}\}$
- Peano arithmetic: $\Sigma = \{0, 1, \text{'+'}, \text{'*'}, \text{'='}\}$
- Theory of reals
- Theory of integers
- Theory of arrays
- Theory of pointers
- Theory of sets
- Theory of recursive data structures
- …

The algorithmic point of view...

- It is also common to present theories NOT through the axioms that define them.
- The interpretation of symbols is fixed to their common use.
  - Thus '+' is plus, ...
- The fragment is defined via grammar rules rather than restrictions on the generic first-order grammar.

The algorithmic point of view...

- Example: equality logic (= "the theory of equality")
- *Grammar:*

formula : *formula* $\vee$ *formula* | $\neg$ *formula* | *atom*

atom : term-variable = term-variable
| term-variable = constant | Boolean-variable

- Interpretation:
  '=' is equality.

The algorithmic point of view...

- This simpler way of presenting theories is all that is needed when our focus is on decision procedures specific for the given theory.

- The traditional way of presenting theories is useful when discussing generic methods (for any decidable theory T)
  - Example 1: algorithms for combining two or more theories
  - Example 2: generic SAT-based decision procedure given a decision procedure for the conjunctive fragment of T.

Expressiveness of a theory

trade-of between what a theory can express and how hard it is to decide

- Each formula defines a language:
  the set of satisfying assignments ('models') are the words accepted by this language.

- Consider the fragment '2-CNF'

  formula : (*literal* $\vee$ *literal*) | *formula* $\wedge$ *formula*
  literal: Boolean-variable | $\neg$Boolean-variable

$$(x_1 \vee \neg x_2) \wedge (\neg x_3 \vee x_2)$$

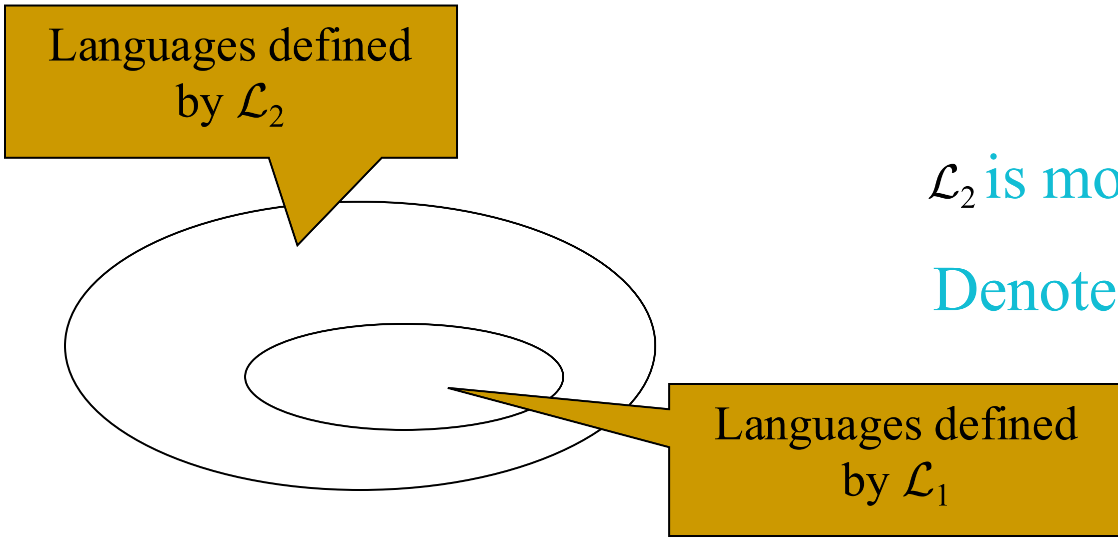
Expressiveness of a theory

- Now consider a Propositional Logic formula

  ϕ: $(x_1 \vee x_2 \vee x_3)$.
- Q: Can we express this language with 2-CNF?
- i.e., CNF in which each clause has at most two literals.
- A: No.

  Proof:
  - The language accepted by ϕ has 7 words: all assignments other than $x_1 = x_2 = x_3 =$ F.
  - The first 2-CNF clause removes ¼ of the assignments, which leaves us with 6 accepted words. Additional clauses only remove more assignments.

Expressiveness of a theory



$\mathcal{L}_2$ is more expressive than $\mathcal{L}_1$.

Denote: $\mathcal{L}_1 \prec \mathcal{L}_2$

- *Claim*: 2-CNF $\prec$ Propositional Logic
- Generally there is only a partial order between theories.

The tradeoff

- So we see that theories can have different expressive power.
- Q: why would we want to restrict ourselves to a theory or a fragment ? why not take some ‘maximal theory’…
- A: Adding axioms to the theory may make it harder to decide or even undecidable.

Example: Hilbert axiom system ($\mathcal{H}$)

- Let $\mathcal{H}$ be (M.P) + the following axiom schemas:

$$\frac{}{A \rightarrow (B \rightarrow A)} \quad (H1)$$

$$\frac{}{((A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))} \quad (H2)$$

$$\frac{}{(\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)} \quad (H3)$$

- $\mathcal{H}$ is sound and complete
- This means that with $\mathcal{H}$ we can prove any valid propositional formula, and only such formulas. The proof is finite.

Example

- But there exists first order theories defined by axioms which are not sufficient for proving all T-valid formulas.

Example: First Order Peano Arithmetic

- $\Sigma = \{0, 1, \text{'+'}, \text{'*'}, \text{'='}\}$
- Domain: Natural numbers
- Axioms ("semantics"):
  1. $\forall x : (0 \neq x + 1)$
  2. $\forall x : \forall y : (x \neq y) \rightarrow (x + 1 \neq y + 1)$
  3. Induction
  4. $\forall x : x + 0 = x$
  5. $\forall x : \forall y : (x + y) + 1 = x + (y + 1)$
  6. $\forall x : x * 0 = 0$
  7. $\forall x \forall y : x * (y + 1) = x * y + x$

+ (axioms 4–5)

* (axioms 6–7)

These axioms define the semantics of '+'

Undecidable!

Example: First Order Presburger Arithmetic

- $\Sigma = \{0, 1, `+', \cancel{`*'}, `='\}$
- Domain: Natural numbers

- Axioms ("semantics"):

decidable!

1. $\forall x : (0 \neq x + 1)$
2. $\forall x : \forall y : (x \neq y) \rightarrow (x + 1 \neq y + 1)$
3. Induction
4. $\forall x : x + 0 = x$
5. $\forall x : \forall y : (x + y) + 1 = x + (y + 1)$
6. ~~$\forall x : x * 0 = 0$~~
7. ~~$\forall x \forall y : x * (y + 1) = x * y + x$~~

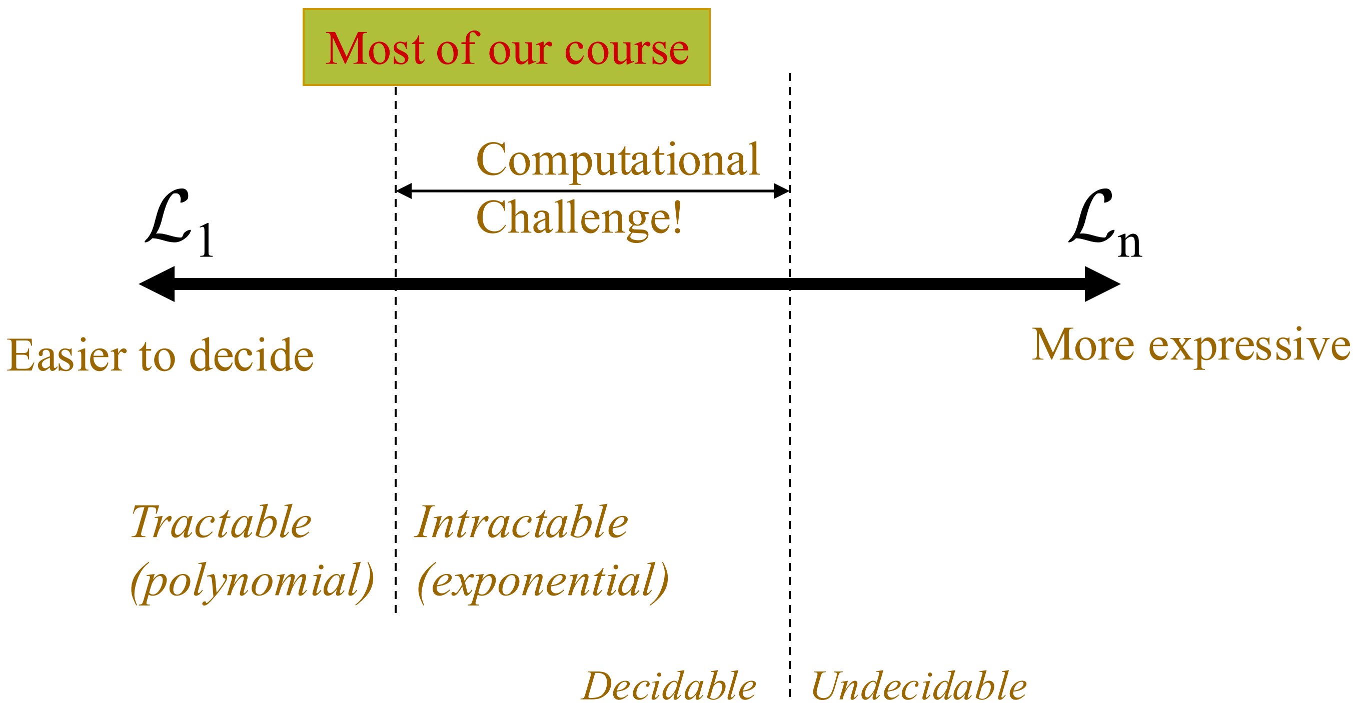
+ (axioms 4–5)

* (axioms 6–7)

These axioms define the semantics of '+'

Tradeoff: expressiveness/computational hardness.

- Assume we are given theories $\mathcal{L}_1 \prec \ldots \prec \mathcal{L}_n$



When is a specific theory useful?

1. Expressible enough to state something interesting.
2. Decidable (or semi-decidable) and more efficiently solvable than richer theories.
3. More expressible, or more natural for expressing some models in comparison to 'leaner' theories.

In the following classes, we will see decision-procedures for some of these theories and their fragments

Expressiveness and complexity : Homework and Discussion

- Q1: Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two theories whose satisfiability problem is decidable and in the same complexity class. Is the satisfiability problem of an $\mathcal{L}_1$ formula reducible to a satisfiability problem of an $\mathcal{L}_2$ formula?

- Q2: Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two theories whose satisfiability problems are reducible to one another.
  Are $\mathcal{L}_1$ and $\mathcal{L}_2$ in the same complexity class ?