



## Lecture 4

# First-Order Theories



Zvonimir Rakamarić  
University of Utah

slides acknowledgements: Zohar Manna

# Announcements

- ▶ Graded homework 1
- ▶ Points lost due to incomplete/unreliable implementations
  - ▶ Budget time for thorough testing
- ▶ Many solutions do not follow the prescribed input/output specification
  - ▶ “Program input: An integer  $n > 0$  entered as a *command line argument...*”
  - ▶ Program output is also well-specified
  - ▶ I did not deduct points for this

# Last Time

- ▶ First-order logic
  - ▶ Syntax and semantics
  - ▶ Quantifiers
  - ▶ Undecidable
- ▶ Proving validity with semantic argument method

# This Time

- ▶ First-order theories
- ▶ Reading: Chapter 3

# First-Order Theories

- ▶ Software manipulates structures
  - ▶ Numbers, arrays, lists, bitvectors,...
- ▶ Software (and hardware) verification
  - ▶ Reasoning about such structures
- ▶ First-order theories
  - ▶ Formalize structures to enable reasoning about them
  - ▶ Validity is sometimes decidable

# Definition

- ▶ **First-order theory**  $T$  defined by:
  - ▶ **Signature**  $\Sigma_T$  – set of constant, function, and predicate symbols
    - ▶ Have no meaning
  - ▶ **Axioms**  $A_T$  – set of closed (no free variables)  $\Sigma_T$ -formulae
    - ▶ Provide meaning for symbols of  $\Sigma_T$

# $\Sigma_T$ -formula

- ▶  $\Sigma_T$ -formula is a formula constructed of:
  - ▶ Constants, functions, and predicate symbols from  $\Sigma_T$
  - ▶ Variables, logical connectives, and quantifiers

# $T$ -interpretation

- Interpretation  $I$  is  $T$ -interpretation if it satisfies all axioms  $A_T$  of  $T$ :

$$I \models A \text{ for every } A \in A_T$$

# Satisfiability and Validity

- ▶  $\Sigma_T$ -formula  $F$  is **satisfiable in theory  $T$**  ( $T$ -satisfiable) if there is a  $T$ -interpretation  $I$  that satisfies  $F$
- ▶  $\Sigma_T$ -formula  $F$  is **valid in theory  $T$**  ( $T$ -valid,  $T \models F$ ) if every  $T$ -interpretation  $I$  satisfies  $F$ 
  - ▶ Theory  $T$  consists of all closed  $T$ -valid formulae
- ▶ Two  $\Sigma_T$ -formulae  $F_1$  and  $F_2$  are **equivalent in  $T$**  ( $T$ -equivalent) if  $T \models F_1 \leftrightarrow F_2$

# Fragment of a Theory

- ▶ **Fragment** of theory  $T$  is a syntactically restricted subset of formulae of the theory
- ▶ Example:
  - ▶ Quantifier-free fragment of theory  $T$  is the set of formulae without quantifiers that are valid in  $T$
- ▶ Often decidable fragments for undecidable theories

# Decidability

- ▶ Theory  $T$  is **decidable** if  $T$ -validity is decidable for every  $\Sigma_T$ -formula  $F$ 
  - ▶ There is an algorithm that always terminates with “yes” if  $F$  is  $T$ -valid, and “no” if  $F$  is  $T$ -invalid
- ▶ Fragment of  $T$  is **decidable** if  $T$ -validity is decidable for every  $\Sigma_T$ -formula  $F$  in the fragment

# Common First-Order Theories

- ▶ Theory of equality
- ▶ Peano arithmetic
- ▶ Presburger arithmetic
- ▶ Linear integer arithmetic
- ▶ Reals
- ▶ Rationals
- ▶ Arrays
- ▶ Recursive data structures
- ▶ Bitvectors

# Theory of Equality $T_E$

## Signature

$$\Sigma_E : \{=, a, b, c, \dots, f, g, h, \dots, p, q, r, \dots\}$$

consists of:

- ▶ a binary predicate “=” interpreted using provided axioms
- ▶ constant, function, and predicate symbols

# Axioms of $T_E$

1.  $\forall x. x=x$  (reflexivity)
2.  $\forall x,y. x=y \rightarrow y=x$  (symmetry)
3.  $\forall x,y,z. x=y \wedge y=z \rightarrow x=z$  (transitivity)
4. for each positive int.  $n$  and  $n$ -ary function symbol  $f$ ,  
$$\forall x_1, \dots, x_n, y_1, \dots, y_n. \left( \bigwedge_{i=1}^n x_i = y_i \right) \rightarrow f(x_1, \dots, x_n) = f(y_1, \dots, y_n)$$
(function congruence)
5. for each positive int.  $n$  and  $n$ -ary predicate symbol  $p$ ,  
$$\forall x_1, \dots, x_n, y_1, \dots, y_n. \left( \bigwedge_{i=1}^n x_i = y_i \right) \rightarrow (p(x_1, \dots, x_n) \leftrightarrow p(y_1, \dots, y_n))$$
(predicate congruence)

# Decidability of $T_E$

- ▶ Bad news
  - ▶  $T_E$  is undecidable
- ▶ Good news
  - ▶ Quantifier-free fragment of  $T_E$  is decidable
  - ▶ Very efficient algorithms

## Z3 Example

$$x=y \wedge y=z \rightarrow g(f(x),y)=g(f(z),x)$$

# Arithmetic: Natural Numbers and Integers

Natural numbers  $\mathbb{N} = \{0, 1, 2, \dots\}$

Integers  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$

Three theories:

- ▶ **Peano arithmetic**  $T_{PA}$ 
  - ▶ Natural numbers with addition (+), multiplication (\*), equality (=)
- ▶ **Presburger arithmetic**  $T_{\mathbb{N}}$ 
  - ▶ Natural numbers with addition (+), equality (=)
- ▶ **Theory of integers**  $T_{\mathbb{Z}}$ 
  - ▶ Integers with addition (+), subtraction (-), comparison (>), equality (=), multiplication by constants

# Peano Arithmetic $T_{PA}$

$\Sigma_{PA} : \{0, 1, +, *, =\}$

- ▶  $T_{PA}$ -satisfiability and  $T_{PA}$ -validity are undecidable
  - ▶ Restrict the theory more

# Presburger Arithmetic $T_{\mathbb{N}}$

$\Sigma_{\mathbb{N}} : \{0, 1, +, =\}$       no multiplication!

Axioms:

1. equality axioms for =
2.  $\forall x. \neg(x+1=0)$  (zero)
3.  $\forall x, y. x+1=y+1 \rightarrow x=y$  (successor)
4.  $F[0] \wedge (\forall x. F[x] \rightarrow F[x+1]) \rightarrow \forall x. F[x]$  (induction)
5.  $\forall x. x+0=x$  (plus zero)
6.  $\forall x, y. x+(y+1)=(x+y)+1$  (plus successor)

# Decidability of $T_{\mathbb{N}}$

- ▶  $T_{\mathbb{N}}$ -satisfiability and  $T_{\mathbb{N}}$ -validity are decidable

# Theory of Integers $T_{\mathbb{Z}}$

$\Sigma_{\mathbb{Z}} : \{\dots, -2, -1, 0, 1, 2, \dots, -3^*, -2^*, 2^*, 3^*, \dots, +, -, =, >\}$

where

- ▶  $\dots, -2, -1, 0, 1, 2, \dots$  are constants
- ▶  $\dots, -3^*, -2^*, 2^*, 3^*, \dots$  are unary functions  
(intended meaning:  $2^*x$  is  $x+x$ ,  $-3^*x$  is  $-x-x-x$ )
- ▶  $+, -, >, =$  have the usual meaning
- ▶  $T_{\mathbb{N}}$  and  $T_{\mathbb{Z}}$  have the same expressiveness
  - ▶ Every  $\Sigma_{\mathbb{Z}}$ -formula can be reduced to  $\Sigma_{\mathbb{N}}$ -formula
  - ▶ Every  $\Sigma_{\mathbb{N}}$ -formula can be reduced to  $\Sigma_{\mathbb{Z}}$ -formula

# Example of $T_{\mathbb{Z}}$ to $T_{\mathbb{N}}$ Reduction

Consider  $\Sigma_{\mathbb{Z}}$ -formula

$$F_0 : \forall w, x. \exists y, z. x + 2*y - z - 13 > -3*w + 5$$

Introduce two variables  $v_p$  and  $v_n$  (range over natural numbers) for each variable  $v$  (range over integers) in  $F_0$ :

$$F_1 : \forall w_p, w_n, x_p, x_n. \exists y_p, y_n, z_p, z_n. \\ (x_p - x_n) + 2*(y_p - y_n) - (z_p - z_n) - 13 > -3*(w_p - w_n) + 5$$

Eliminate - by moving to the other side of >:

$$F_2 : \forall w_p, w_n, x_p, x_n. \exists y_p, y_n, z_p, z_n. \\ x_p + 2*y_p + z_n + 3*w_p > x_n + 2*y_n + z_p + 13 + 3*w_n + 5$$

## Example of $T_{\mathbb{Z}}$ to $T_{\mathbb{N}}$ Reduction cont.

Eliminate  $*$  and  $>$ :

$$\begin{aligned} F_3 : \forall w_p, w_n, x_p, x_n. \exists y_p, y_n, z_p, z_n. \exists u. \neg(u=0) \wedge \\ x_p + y_p + y_p + z_n + w_p + w_p + w_p \\ = x_n + y_n + y_n + z_p + w_n + w_n + w_n + u \\ + 1+1+1+1+1+1+1+1+1 \\ + 1+1+1+1+1+1+1+1+1 \end{aligned}$$

- $F_3$  is a  $\Sigma_{\mathbb{N}}$ -formula equisatisfiable to  $F_0$

# Example of $T_{\mathbb{N}}$ to $T_{\mathbb{Z}}$ Reduction

Consider  $\Sigma_{\mathbb{N}}$ -formula

$$F : \forall x. \exists y. x=y+1$$

$F$  is equisatisfiable to  $\Sigma_{\mathbb{Z}}$ -formula

$$\forall x. x > -1 \rightarrow \exists y. y > -1 \wedge x=y+1$$

# Decidability of $T_{\mathbb{Z}}$

- ▶  $T_{\mathbb{Z}}$ -satisfiability and  $T_{\mathbb{Z}}$ -validity are decidable

## Z3 Example

$$x > z \wedge y \geq 0 \rightarrow x + y > z$$

# Theory of Reals $T_{\mathbb{R}}$ and Rationals $T_{\mathbb{Q}}$

$\Sigma_{\mathbb{R}} : \{0, 1, +, -, *, =, \}$  with multiplication

$\Sigma_{\mathbb{Q}} : \{0, 1, +, -, =, \}$  without multiplication

# Decidability of $T_{\mathbb{R}}$ and $T_{\mathbb{Q}}$

- ▶ Both are decidable
  - ▶ High time complexity
- ▶ Quantifier-free fragment of  $T_{\mathbb{Q}}$  is efficiently decidable

# Theory of Arrays $T_A$

$\Sigma_A : \{select, store, =\}$

where

- ▶  $select(a, i)$  is a binary function:
  - ▶ read array  $a$  at index  $i$
- ▶  $store(a, i, v)$  is a ternary function:
  - ▶ write value  $v$  to index  $i$  of array  $a$

# Axioms of $T_A$

1.  $\forall a, i, j. i = j \rightarrow \text{select}(a, i) = \text{select}(a, j)$   
(array congruence)
2.  $\forall a, v, i, j. i = j \rightarrow \text{select}(\text{store}(a, i, v), j) = v$   
(select-store 1)
3.  $\forall a, v, i, j. i \neq j \rightarrow \text{select}(\text{store}(a, i, v), j) = \text{select}(a, j)$   
(select-store 2)

## Note about $T_A$

- ▶ Equality (=) is only defined for array elements...

- ▶ Example:

$select(a,i)=e \rightarrow \forall j. select(store(a,i,e),j)=select(a,j)$   
is  $T_A$ -valid

- ▶ ...and not for whole arrays

- ▶ Example:

$select(a,i)=e \rightarrow store(a,i,e)=a$   
is not  $T_A$ -valid

# Decidability of $T_A$

- ▶  $T_A$  is undecidable
- ▶ Quantifier-free fragment of  $T_A$  is decidable

# Theory of Arrays with Extensionality $T_A^=$

- ▶ Signature and axioms of  $T_A^=$  are the same as  $T_A$ , with one additional axiom:  
$$\forall a, b. (\forall i. \text{select}(a, i) = \text{select}(b, i)) \leftrightarrow a = b$$

(extensionality)
- ▶  $T_A^=$ -valid example  
$$\text{select}(a, i) = e \rightarrow \text{store}(a, i, e) = a$$

## Decidability of $T_A^=$

- ▶  $T_A^=$  is undecidable
- ▶ Quantifier-free fragment of  $T_A^=$  is decidable

# Summary of Decidability Results

| Theory           |                           | Quantifiers<br>Decidable | QFF<br>Decidable |
|------------------|---------------------------|--------------------------|------------------|
| $T_E$            | Equality                  | NO                       | YES              |
| $T_{PA}$         | Peano Arithmetic          | NO                       | NO               |
| $T_{\mathbb{N}}$ | Presburger Arithmetic     | YES                      | YES              |
| $T_{\mathbb{Z}}$ | Linear Integer Arithmetic | YES                      | YES              |
| $T_{\mathbb{R}}$ | Real Arithmetic           | YES                      | YES              |
| $T_{\mathbb{Q}}$ | Linear Rationals          | YES                      | YES              |
| $T_A$            | Arrays                    | NO                       | YES              |

# Summary of Complexity Results

| Theory           |                           | Quantifiers           | QF<br>Conjunctive |
|------------------|---------------------------|-----------------------|-------------------|
| PL               | Propositional Logic       | NP-complete           | $O(n)$            |
| $T_E$            | Equality                  | –                     | $O(n \log n)$     |
| $T_{\mathbb{N}}$ | Presburger Arithmetic     | $O(2^{2^{2^{(kn)}}})$ | NP-complete       |
| $T_{\mathbb{Z}}$ | Linear Integer Arithmetic | $O(2^{2^{2^{(kn)}}})$ | NP-complete       |
| $T_{\mathbb{R}}$ | Real Arithmetic           | $O(2^{2^{(kn)}})$     | $O(2^{2^{(kn)}})$ |
| $T_{\mathbb{Q}}$ | Linear Rationals          | $O(2^{2^{(kn)}})$     | PTIME             |
| $T_A$            | Arrays                    | –                     | NP-complete       |

$n$  – input formula size;  $k$  – some positive integer

# Next Time

- ▶ Z3 examples