# Lecture 4 First-Order Theories 

Zvonimir Rakamarić University of Utah

## Announcements

- Homework 1 is due on Thursday


## 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 / is $T$-interpretation if it satisfies all axioms $A_{T}$ of $T$ :
$I \vDash A$ 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 / that satisfies $F$
- $\Sigma_{T}$-formula $F$ is valid in theory $T$ ( $T$-valid, $T \vDash F$ )
if every $T$-interpretation / 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 $T \vDash F_{1} \leftrightarrow F_{2}$


## 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


## Theory of Equality $T_{E}$

## Signature

$$
\Sigma_{E}:\{=, a, b, c, \ldots, f, g, h, \ldots, p, q, r, \ldots\}
$$

consists of:

* a binary predicate "=" interpreted using provided axioms
- constant, function, and predicate symbols


## Axioms of $T_{E}$

1. $\forall X \cdot X=X$
2. $\forall x, y . x=y \rightarrow y=x$
3. $\forall x, y, z . ~ x=y \wedge y=z \rightarrow x=z$
4. for each positive int. n and n -ary function symbol $f$,

$$
\forall x_{1}, \ldots, x_{n}, y_{1}, \ldots, y_{n} .\left(\bigwedge_{=1}^{n} x_{i}=y_{i}\right) \rightarrow f\left(x_{1}, \ldots, x_{n}\right)=f\left(y_{1}, \ldots, y_{n}\right)
$$

(function congruence)
5. for each positive int. n and n -ary predicate symbol $p$,

$$
\begin{aligned}
\forall x_{1}, \ldots, x_{n}, y_{1}, \ldots, y_{n} .\left(\widehat{i=1}_{n} x_{i}=y_{i}\right) \rightarrow & \left(p\left(x_{1}, \ldots, x_{n}\right) \leftrightarrow p\left(y_{1}, \ldots, y_{n}\right)\right) \\
& \text { (predicate congruence) }
\end{aligned}
$$

## 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 $\mathbb{N}=\{0,1,2, \ldots\}$
Integers $\mathbb{Z}=\{\ldots,-2,-1,0,1,2, \ldots\}$
Three theories:

- Peano arithmetic $T_{P A}$
- 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 $T_{P A}$ <br> $\Sigma_{P A}:\left\{0,1,+{ }^{*},=\right\}$

- $T_{P A}$-satisfiability and $T_{P A}$-validity are undecidable

Restrict the theory more

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

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

Axioms:

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

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

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


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

$\Sigma_{\mathbb{Z}}:\left\{\ldots,-2,-1,0,1,2, \ldots,-3^{*},-2^{*}, 2^{*}, 3^{*}, \ldots,+,-,=,>\right\}$
where
( $. .,-2,-1,0,1,2, \ldots$ are constants
$\ldots,-3^{*},-2^{*}, 2^{*}, 3^{*}, \ldots$ 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 $F_{0}$ :
$F_{1}: \forall w_{p}, w_{n}, x_{p}, x_{n} . \exists y_{p}, y_{n}, z_{p}, z_{n}$.

$$
\left(x_{p}-x_{n}\right)+2^{*}\left(y_{p}-y_{n}\right)-\left(z_{p}-z_{n}\right)-13>-3^{*}\left(w_{p}-w_{n}\right)+5
$$

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 >:

$$
\begin{gathered}
F_{3}: \forall w_{p}, w_{n}, x_{p}, x_{n} \cdot \exists y_{p}, y_{n}, z_{p}, z_{n} \cdot \exists u \cdot \neg(u=0) \wedge \\
x_{p}+y_{p}+y_{p}+z_{n}+w_{p}+w_{p}+w_{p} \\
= \\
x_{n}+y_{n}+y_{n}+z_{p}+w_{n}+w_{n}+w_{n}+\mathrm{u} \\
\quad+1+1+1+1+1+1+1+1+1 \\
\quad+1+1+1+1+1+1+1+1+1
\end{gathered}
$$

- $F_{3}$ 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
$F: \forall x$. $\exists \mathrm{y} . x=y+1$
$F$ 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

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

## Theory of Reals $T_{\mathbb{R}}$ and Rationals $T_{\mathbb{Q}}$ <br> $\Sigma_{\mathbb{R}}:\left\{0,1,+,-{ }^{*},=,\right\}$ <br> 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. $\forall a, i, j . i=j \rightarrow \operatorname{select}(a, i)=\operatorname{select}(a, j)$
(array congruence)
2. $\forall a, v, i, j . i=j \rightarrow \operatorname{select}(\operatorname{store}(a, i, v), j)=v$ (select-store 1)
3. $\forall a, v, i, j . i \neq j \rightarrow \operatorname{select}(\operatorname{store}(a, i, v), j)=\operatorname{select}(a, j)$ (select-store 2)

## Note about $T_{A}$

- Equality (=) is only defined for array elements...
- Example:
$\operatorname{select}(a, i)=e \rightarrow \forall j$. select( $\operatorname{store}(a, i, e), j)=\operatorname{select}(a, j)$ is $T_{A}$-valid
...and not for whole arrays
- Example:
$\operatorname{select}(a, i)=e \rightarrow \operatorname{store}(a, i, e)=a$
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:$\forall a, b .(\forall i . \operatorname{select}(a, i)=\operatorname{select}(b, i)) \leftrightarrow a=b$
(extensionality)
$T_{A}=$-valid example
$\operatorname{select}(a, i)=e \rightarrow \operatorname{store}(a, i, e)=a$

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 |
| $T_{P A}$ | 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 | $\mathrm{O}(\mathrm{n})$ |  |
| $T_{E}$ Equality | - | $\mathrm{O}(\mathrm{n}$ log $n)$ |  |
| $T_{\mathbb{N}}$ | Presburger Arithmetic | $\mathrm{O}\left(2^{\wedge} 2^{\wedge} 2^{\wedge}(\mathrm{kn})\right)$ | NP-complete |
| $T_{\mathbb{Z}}$ Linear Integer Arithmetic | $\mathrm{O}\left(2^{\wedge} 2^{\wedge} 2^{\wedge}(\mathrm{kn})\right)$ | NP-complete |  |
| $T_{\mathbb{R}}$ | Real Arithmetic | $\mathrm{O}\left(2^{\wedge} 2^{\wedge}(\mathrm{kn})\right)$ | $\mathrm{O}\left(2^{\wedge} 2^{\wedge}(\mathrm{kn})\right)$ |
| $T_{\mathbb{Q}}$ | Linear Rationals | $\mathrm{O}\left(2^{\wedge} 2^{\wedge}(\mathrm{kn})\right)$ | PTIME |
| $T_{A}$ | Arrays | - | NP-complete |

n - input formula size; k - some positive integer

Next Time

- Z3 examples

