CSE4051: Program Verification

First-Order Theories

2025 Fall

Woosuk Lee

Review: First-Order Logic

- FOL is an extension of PL with predicates, functions, and quantifiers.
- The semantics is determined by an interpretation.
- An interpretation I consists of a domain (D_I) and an assignment (A_I) for free variables and **nonlogical symbols** (functions, predicates, and constants).
  - $\exists x \,.\, x + 0 = 1$ is true under the conventional interpretation but false if we interpret $+$ as multiplication.

First-Order Theories

- In practice, we are NOT interested in pure logical validity (i.e., valid in all interpretations) of FOL formulas but in validity in a specific class of interpretations.
  - $\exists x \,.\, x + 0 = 1$ under interpretations where + is treated as addition
  - In many cases, we have a particular meaning of functions/predicates in mind.
- First-order logic is a general framework for building a specific, restricted logic called **theories**.

First-Order Theories

- The restrictions are made on nonlogical symbols and interpretations. For instance, in the theory of integers, only + and − are allowed for function symbols with their conventional interpretations.
- One natural way for restricting interpretations is to provide a set of **axioms**; we only consider interpretations that satisfy the axioms.

First-Order Theories

- A first-order theory T is defined by the two components:
  - **Signature**: a set of nonlogical symbols. Given a signature Σ, a Σ-formula is one whose nonlogical symbols are from Σ. Signature restricts the syntax.
  - **Axioms**: A set of closed FOL formulas whose nonlogical symbols are from Σ. Axioms restrict the interpretations.

Example: The Theory of Heights

- A theory with a predicate $taller$
- Signature $\Sigma_H = \{taller\}$
- Axioms A_H provide the meaning of the symbols in Σ_H (i.e., $taller$)
  - $\forall x, y \,.\, taller(x, y) \implies \neg taller(y, x)$
- An interpretation $I = (D_I, A_I)$ where $D_I = \{A, B\}$ and

  $A_I(taller) = \{(A, B) \mapsto \text{true}, (B, A) \mapsto \text{true}\}$ does not satisfy the axiom.
- An interpretation $I = (D_I, A_I)$ where $D_I = \{A, B\}$ and

  $A_I(taller) = \{(A, B) \mapsto \text{true}\}$ satisfies the axiom.

Example: The Theory of Equality

- A theory with a fixed interpretation for =. For example, the formula must be valid according to the conventional interpretation of =:

$$\forall x, y, z \,.\, (((x = y) \wedge \neg(y = z)) \implies \neg(x = z))$$

- To fix this interpretation, it is sufficient to enforce the following axioms:
  - Reflexivity: $\forall x \,.\, x = x$
  - Symmetry: $\forall x, y \,.\, x = y \implies y = x$
  - Transitivity: $\forall x, y, z \,.\, x = y \wedge y = z \implies x = z$

Satisfiability and Validity

- Instead of pure logical satisfiability / validity under any interpretation, we focus on satisfiability / validity under interpretations of interest.
- Given a theory T with signature Σ and axioms A, an interpretation I is called T-interpretation if
  - $I \models a$ for every $a \in A$ (every axiom in A is valid under I)
- A Σ-formula F is T-satisfiable (or satisfiable modulo T) if there exists a T -interpretation that satisfies F.
- A Σ-formula F is T-valid (or valid modulo T) if every T-interpretation satisfies F (we write $T \models F$).
- The theory T consists of all closed formulae that are T-valid.

Decidability and Completeness

- A theory T is **decidable** if there exists a procedure that for any Σ-formula (formula consisting of symbols in Σ) F, (1) eventually halts and (2) answers yes if F is T-valid and no otherwise.
- A theory T is **complete** if for every closed Σ-formula F, $T \models F$ or $T \models \neg F$.

Fragments of Theories

- A theory restricts only the nonlogical symbols. Restrictions on the logical symbols or the grammar are done by defining fragments of the logic. Two popular fragments:
  - **Quantifier-free fragment**: the set of Σ-formulas without quantifiers.
  - **Conjunctive fragment**: the set of formulas where the only boolean connective that is allowed is conjunction.
- Many first-order theories are undecidable while their quantifier-free fragments are decidable. In practice, we are mostly interested in the satisfiability problem of the quantifier-free fragment of first-order theories.

Example

Recall the theory of heights T_H.

- An interpretation $I = (D_I, A_I)$ where $D_I = \{A, B\}$ and

  $A_I(taller) = \{(A, B) \mapsto \text{true}, (B, A) \mapsto \text{true}\}$ is NOT a T_H-interpretation.
- An interpretation $I = (D_I, A_I)$ where $D_I = \{A, B\}$ and

  $A_I(taller) = \{(A, B) \mapsto \text{true}\}$ is a T_H-interpretation.
- The following formula is T_H-valid.

$$\forall x \,.\, \neg taller(x, x)$$

First-Order Theories for Programs

- When reasoning in SW, we have particular structures in mind (e.g., numbers, lists, arrays, …)
- First-order theories formalize these structures to enable reasoning about them.
- These theories include a theory of
  - Equality
  - Integers
  - Rationals and reals
  - Arrays
  - Bitvectors
  - …

Theory of Equality with Uninterpreted Functions (T_E)

- The simplest first-order theory
- Signature Σ_E consisting of
  - = (equality), a binary predicate,
  - and all other symbols (constant, function, and predicate symbols)
- Equality = is **interpreted** predicate symbol: its meaning will be defined via axioms.
- The other functions, predicates, and constants are left unspecified (**uninterpreted**)
- Axioms A_E :
  - Reflexivity: $\forall x \,.\, x = x$
  - Symmetry: $\forall x, y \,.\, x = y \implies y = x$
  - Transitivity: $\forall x, y, z \,.\, x = y \wedge y = z \implies x = z$
  - …

Theory of Equality with Uninterpreted Functions (T_E)

- …
- Function congruence (for each positive integer n and n-ary function symbol f):

$\overline{x}$: list of variables
$x_1, \ldots, x_n$

$$\forall \overline{x}, \overline{y}. \left(\bigwedge_{i=1}^{n} x_i = y_i \right) \rightarrow f(\overline{x}) = f(\overline{y})$$

- Predicate congruence (for each positive integer n and n-ary predicate symbol p):

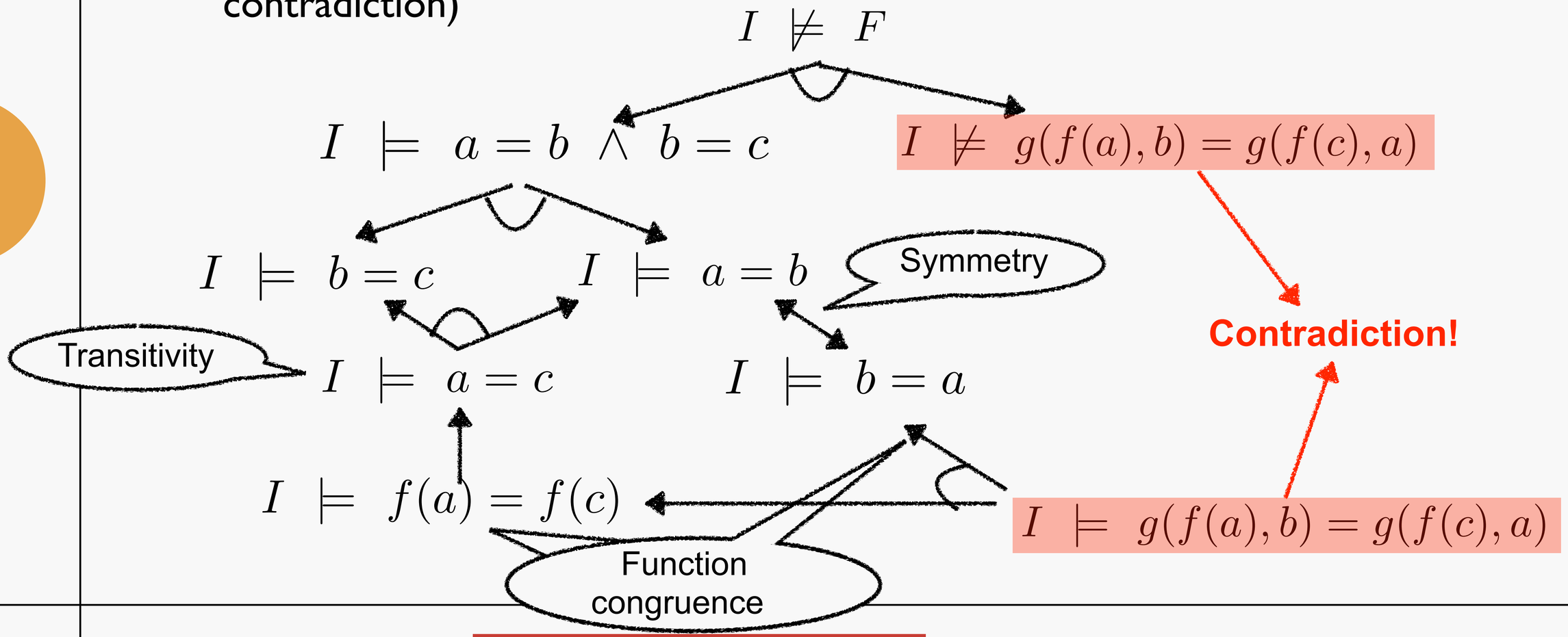
$$\forall \overline{x}, \overline{y}. \left(\bigwedge_{i=1}^{n} x_i = y_i \right) \rightarrow (p(\overline{x}) \leftrightarrow p(\overline{y}))$$

$\leftrightarrow$: $\Rightarrow$ and $\Leftarrow$

- Meaning: no matter what functions and predicates are used, if the inputs are the same, the outcomes are also the same.

Example

- Prove that $F: \; a = b \; \wedge \; b = c \; \rightarrow \; g(f(a), b) = g(f(c), a)$ is T_E -valid (proof by contradiction)



Uninterpreted Functions

- In T_E, function symbols are uninterpreted since the axioms do not assign meaning to them.
- The only thing we know about them is that they are functions.
- A main use of uninterpreted functions is to abstract complex formulas that are otherwise difficult to automatically reason about.

Uninterpreted Functions for Program Equivalence

- Suppose we want to prove equivalence of the following two programs:

```
int fun1(int y) {
  int x, z, w;
  z = y;
  w = x;
  x = z;
  return x*x;
}
```

```
int fun2(int y) {
 return y*y;
}
```

- Let r_1, r_2 be return values of `fun1` and `fun2` respectively.
- We want to prove unsatisfiability of

$$z = y \wedge w = x \wedge x = z \wedge r_1 = x \times x \wedge r_2 = y \times y \wedge \neg(r_1 = r_2)$$

Uninterpreted Functions for Program Equivalence

- We can solve it by reducing the problem into a SAT problem by treating variables x,z,w,y as 32-bit bit vectors.
- But a SAT solver fails to solve within 5 minutes because multiplication makes the problem hard.
- Using an uninterpreted function *sqr*, we can rewrite the formula as

$$z = y \wedge w = x \wedge x = z \wedge r_1 = sqr(x) \wedge r_2 = sqr(y) \wedge \neg(r_1 = r_2)$$

  which is UNSAT in the theory of equality with uninterpreted functions.
- Therefore, the two programs are equal (why?)

Theory of Peano Arithmetic

- The theory of Peano arithmetic T_{PA} has signature $\Sigma_{PA} = \{0, 1, +, \cdot, =\}$.
  - 0, 1 : constants, $+, \cdot$ (addition & multiplication) : binary functions
  - = : binary predicate
- Axioms A_{PA} define addition, multiplication and equality over natural numbers.

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

$F[x]$: When F has one free variable, a formula obtained by replacing the free variable by x

Example

- The formula $3x + 5 = 2y$ can be written using Σ_{PA} as

$$x + x + x + 1 + 1 + 1 + 1 + 1 = y + y$$

or

$$(1 + 1 + 1) \cdot x + 1 + 1 + 1 + 1 + 1 = (1 + 1) \cdot y$$

- The formula $3x + 5 > 2y$ can be written as

$$\exists z \,.\, \neg(z = 0) \wedge 3x + 5 = 2y + z$$

- The formula $3x + 5 \geq 2y$ can be written as

$$\exists z . 3x + 5 = 2y + z$$

Example

- Pythagorean Theorem:

$$\exists x, y, z.\ x \neq 0 \ \wedge \ y \neq 0 \ \wedge \ z \neq 0 \ \wedge \ xx + yy = zz$$

- Fermat's Last Theorem:

$$\{\forall x, y, z.\ x \neq 0 \ \wedge \ y \neq 0 \ \wedge \ z \neq 0 \ \rightarrow \ x^n + y^n \neq z^n \ : \ n > 2 \ \wedge \ n \in \mathbb{Z}\}$$

Decidability and Completeness

- Validity in T_{PA} is NOT **decidable**: there does NOT exist a procedure that for any Σ_{PA} formula F, (1) eventually halts and (2) answers "yes" if F is T_{PA}-valid and answers "no" otherwise.
- Validity in even the quantifier-free fragment T_{PA} (i.e., T_{PA} without quantifiers) is not decidable.
- T_{PA} is **incomplete**: Not all valid Σ_{PA}-formulae can be proved to be valid using the axioms A_{PA}.
- To be decidable and complete, we need to **drop multiplication.**

Theory of Presburger Arithmetic

- The theory of Presburger arithmetic $T_{\mathbb{N}}$ has signature $\Sigma_{\mathbb{N}} = \{0,1, +, = \}$.
  - 0, 1 : constants, + (addition) : binary function
  - = : binary predicate
- Axioms $A_{\mathbb{N}}$ define addition, multiplication and equality over natural numbers.

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

- $T_{\mathbb{N}}$ is both complete and decidable.

Encoding Negative Numbers

- How can we reason about all integers $\mathbb{Z}$ (including negative numbers?)
- Consider $F_0 : \ \forall w, x.\ \exists y, z.\ x + 2y - z - 13 > -3w + 5$

  where $-$ is meant to be subtraction, and all variables are intended to range over $\mathbb{Z}$. The formula

$$F_1 : \begin{array}{l} \forall w_p, w_n, x_p, x_n.\ \exists y_p, y_n, z_p, z_n. \\ \quad (x_p - x_n) + 2(y_p - y_n) - (z_p - z_n) - 13 > -3(w_p - w_n) + 5 \end{array}$$

  introduces two variables, v_p and v_n for each variable v in F_0 (each v_p and v_n can only range over $\mathbb{N}$, $v_p - v_n$ should range over $\mathbb{Z}$. Then, how is $-$ interpreted?

Encoding Negative Numbers

- Moving negated terms to the other side eliminates $-$:

$$F_2: \quad \begin{aligned} &\forall w_p, w_n, x_p, x_n.\ \exists y_p, y_n, z_p, z_n. \\ &\quad x_p + 2y_p + z_n + 3w_p > x_n + 2y_n + z_p + 13 + 3w_n + 5\ . \end{aligned}$$

- The final transformation eliminates constant coefficients and strict inequality:

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

Theory of Integers

- Although integer reasoning can be done with natural numbers, it is convenient to have a theory of integers.
- The theory of integers $T_{\mathbb{Z}}$ has signature

$$\Sigma_{\mathbb{Z}}: \{\ldots, -2, -1, 0, 1, 2, \ldots, -3\cdot, -2\cdot, 2\cdot, 3\cdot, \ldots, +, -, =, >\}$$

where

$\ldots, -2, -1, 0, 1, 2, \ldots$ are integer constants,

$\ldots, -3\cdot, -2\cdot, 2\cdot, 3\cdot, \cdots$ are unary functions representing constant coefficients (e.g., $2 \cdot x$, abbreviated $2x$)

$+, -$ are binary functions and $=, >$ are binary predicates over $\mathbb{Z}$

Theory of Rationals

- The theory of rationals $T_{\mathbb{R}}$ has signature $\Sigma_{\mathbb{Q}}$

$$\Sigma_{\mathbb{Q}} : \{0,\ 1,\ +,\ -,\ =,\ \geq\}$$

- Axioms $A_{\mathbb{Q}}$

1. $\forall x, y.\ x \geq y\ \wedge\ y \geq x\ \rightarrow\ x = y$	(antisymmetry)
2. $\forall x, y, z.\ x \geq y\ \wedge\ y \geq z\ \rightarrow\ x \geq z$	(transitivity)
3. $\forall x, y.\ x \geq y\ \vee\ y \geq x$	(totality)
4. $\forall x, y, z.\ (x + y) + z = x + (y + z)$	(+ associativity)
5. $\forall x.\ x + 0 = x$	(+ identity)
6. $\forall x.\ x + (-x) = 0$	(+ inverse)
7. $\forall x, y.\ x + y = y + x$	(+ commutativity)
8. $\forall x, y, z.\ x \geq y\ \rightarrow\ x + z \geq y + z$	(+ ordered)

...

Theory of Rationals

- Axioms $A_{\mathbb{Q}}$

  ...

  9. for each positive integer n,

     $$\forall x.\ nx = 0 \ \rightarrow \ x = 0$$ (torsion-free)

  10. for each positive integer n,

      $$\forall x.\ \exists y.\ x = ny$$ (divisible)

Theory of Rationals

- Strict inequality is simple to express in $T_{\mathbb{Q}}$. Write

$$\forall x, y.\ \exists z.\ x + y > z$$

as $\Sigma_{\mathbb{Q}}$-formula

$$\forall x, y.\ \exists z.\ \neg(x + y = z)\ \wedge\ x + y \geq z\ .$$

- Rational coefficients are also simple to express. Write

$$\frac{1}{2}x + \frac{2}{3}y \geq 4$$

as

$$3x + 4y \geq 24.$$

Theory of Rationals vs. Presburger arithmetic

- Rational numbers do not satisfy $T_{\mathbb{Z}}$ axioms but they satisfy $T_{\mathbb{Q}}$ axioms.
- $\exists x.2x = 3$ is $T_{\mathbb{Z}}$-invalid. However, assigning x to $\frac{3}{2}$ satisfies it, so satisfiable in the theory of rationals.
- Every formula valid in $T_{\mathbb{Z}}$ is valid in $T_{\mathbb{Q}}$, but not vice versa.
  - Therefore, deciding $T_{\mathbb{Z}}$-validity is more difficult than $T_{\mathbb{Q}}$-validity.
- Both theories (full and quantifier-free) are decidable.

Theory of Lists

- The theory of lists T_{cons} has signature

  $\Sigma_{\text{cons}} : \{\text{cons}, \text{car}, \text{cdr}, \text{atom}, =\}$

  where

  - cons is a binary function, called the constructor: cons(a, b) represents the list constructed by concatenating a to b;
  - car is a unary function, called the left projector: car(cons(a, b)) = a
  - cdr is a unary function, called the right projector: cdr(cons(a, b)) = b
  - atom is a unary predicate: atom(x) is true iff x is a single-element list,
  - and =

Theory of Lists

- Examples
  - cons(a,cons(b,c)) is a list of three elements: a, b, and c.
  - atom(a) is true, atom(cons(a,cons(b,c))) is false
  - car(cons(a, cons(b, c))) = a
  - cdr(cons(a, cons(b, c))) = cons(b, c)
  - cdr(cdr(cons(a, cons(b, c)))) = c

Theory of Lists

- Axioms A_{cons}:
  - The axioms of reflexivity, symmetry, and transitivity of T_E
  - $\forall x_1, x_2, y_1, y_2.\ x_1 = x_2 \ \wedge\ y_1 = y_2 \ \rightarrow\ \mathsf{cons}(x_1, y_1) = \mathsf{cons}(x_2, y_2)$
  - $\forall x, y.\ x = y \ \rightarrow\ \mathsf{car}(x) = \mathsf{car}(y)$
  - $\forall x, y.\ x = y \ \rightarrow\ \mathsf{cdr}(x) = \mathsf{cdr}(y)$
  - $\forall x, y.\ x = y \ \rightarrow\ (\mathsf{atom}(x) \leftrightarrow \mathsf{atom}(y))$
  - $\forall x, y.\ \mathsf{car}(\mathsf{cons}(x, y)) = x$
  - $\forall x, y.\ \mathsf{cdr}(\mathsf{cons}(x, y)) = y$
  - $\forall x.\ \neg\mathsf{atom}(x) \ \rightarrow\ \mathsf{cons}(\mathsf{car}(x), \mathsf{cdr}(x)) = x$
  - $\forall x, y.\ \neg\mathsf{atom}(\mathsf{cons}(x, y))$

Theory of Arrays

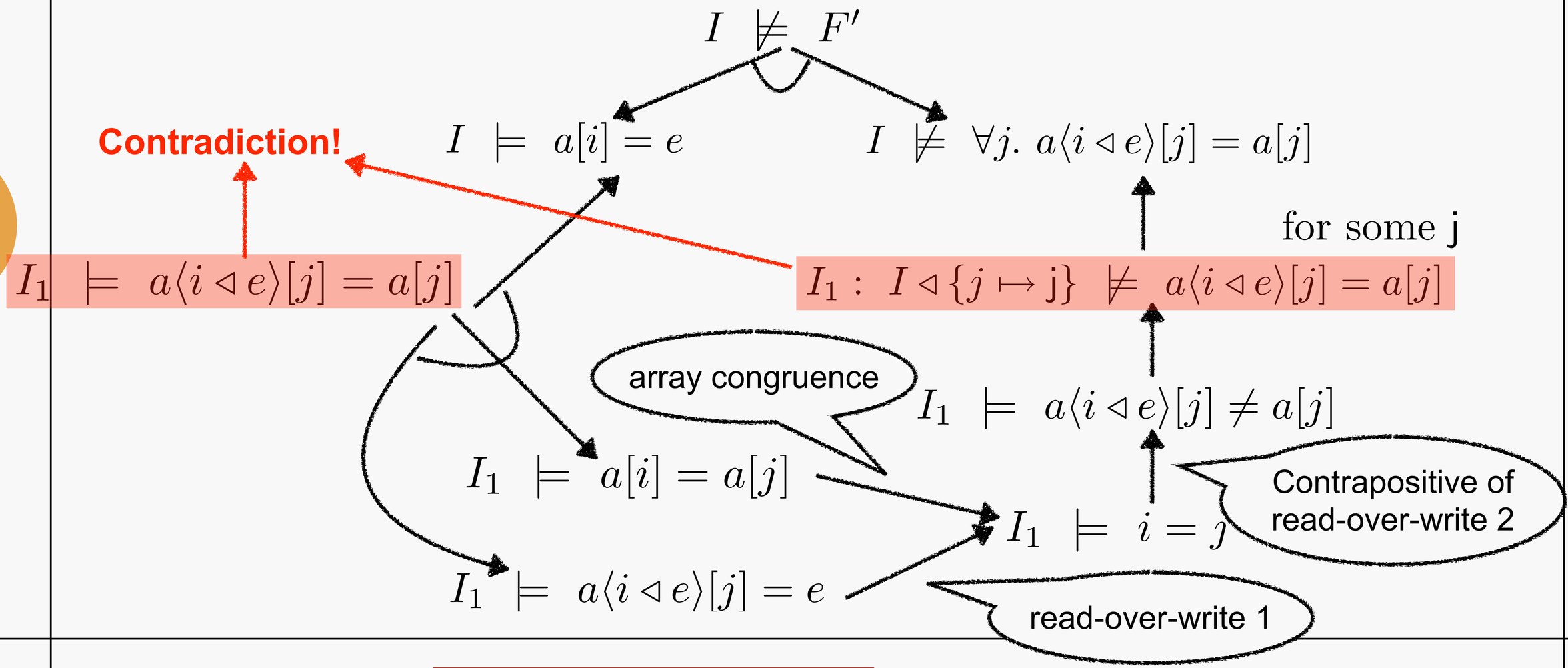
- Arrays are similar to the uninterpreted functions of T_E except they can be modified.
- The theory of arrays T_A has signature $\Sigma_A : \{\cdot[\cdot],\ \cdot\langle\cdot \lhd \cdot\rangle,\ =\}$ where
  - $\cdot[\cdot]$ (read) is a binary function: $a[i]$ represents the value of array a at position i
  - $\cdot\langle\cdot \lhd \cdot\rangle$ (write) is a ternary function: $a\langle i \lhd v\rangle$ represents the modified array a in which position i has value v

Theory of Arrays

- Axioms A_A
  - The axioms of reflexivity, symmetry, and transitivity of T_E
  - $\forall a, i, j.\ i = j \ \rightarrow\ a[i] = a[j]$ (array congruence)
  - $\forall a, v, i, j.\ i = j \ \rightarrow\ a\langle i \triangleleft v\rangle[j] = v$ (read-over-write 1)
  - $\forall a, v, i, j.\ i \neq j \ \rightarrow\ a\langle i \triangleleft v\rangle[j] = a[j]$ (read-over-write 2)
- The equality predicate = is only defined for array "elements" (equality between arrays is not allowed).

Example

- Prove $F' : \; a[i] = e \;\rightarrow\; \forall j.\; a\langle i \triangleleft e\rangle[j] = a[j]$ T_A-vaid.



Theory of Fixed-Width Bitvectors

- The theory of fixed-width bitvectors has signature
  - constants
  - fixed-width words (modeling machine ints, longs, etc.)
  - arithmetic operations (+, -, *, /, etc.)
  - bitwise operations (&, |, ^, etc.)
  - comparison operators (<, >, etc.)
  - =
- With many axioms

Decidability of theories and quantifier-free fragments

Theory	Description	Full	QFF
T_{E}	equality	no	yes
T_{PA}	Peano arithmetic	no	no
$T_{\mathbb{N}}$	Presburger arithmetic	yes	yes
$T_{\mathbb{Z}}$	linear integers	yes	yes
$T_{\mathbb{R}}$	reals (with $\cdot$)	yes	yes
$T_{\mathbb{Q}}$	rationals (without $\cdot$)	yes	yes
T_{RDS}	recursive data structures	no	yes
T^{+}_{RDS}	acyclic recursive data structures	yes	yes
T_{A}	arrays	no	yes
$T^{=}_{\mathsf{A}}$	arrays with extensionality	no	yes

T_A with axiom

$\forall a, b.\ (\forall i.\ a[i] = b[i]) \ \leftrightarrow\ a = b$

Summary

- First-order theories
- Signature, axioms
- Decidability