First-Order Logic: Soundness and Completeness

Yuting Wang

John Hopcroft Center for Computer Science
Shanghai Jiao Tong University

Dec 11, 2023

First-Order Logic: Soundness and Completeness

Reading

- Enderton, Chapters 2.4, 2.5

What is the equivalent definition of "tautological implication" in first-order logic?

Logical Implication

Definition

Let Σ be a set of wffs and φ a wff. Σ logically implies φ, written as

$$\Sigma \vDash \varphi$$

if for every structure $\mathfrak{A}$ and every assignment $s : V \to |\mathfrak{A}|$,

if $\mathfrak{A}$ satisfies Σ with s, then $\mathfrak{A}$ satisfies φ with s.

Remark

$\Sigma \vDash \varphi$ is also read as:

- φ is a logical consequence of Σ, or
- Σ semantically implies φ, or
- φ is a semantic consequence of Σ.

Logical Implication for Sentences

Theorem

For a set of sentences Σ and a sentence σ, $\Sigma \models \sigma$ iff for every model $\mathfrak{A}$ of Σ, $\mathfrak{A}$ is a model of σ.

Scorates Again

Question

Assume the following premises:

- All men are mortal.
- Socrates is a man.

We can derive the conclusion:

- Socrates is mortal.

How do we express this reasoning using logical implication?

Answer

Let $\mathbb{L}$ be the first-order language with 1-ary predicate symbols:

- P for asserting a being is a man;
- Q for asserting a being is mortal;

and a constant symbol c denoting Socrates.
Let $\Sigma = \{\forall x(P(x) \to Q(x)), P(c)\}$. Then

$$\Sigma \models Q(c).$$

Logical Equivalence

As before, we write $\alpha \models \beta$ for $\{\alpha\} \models \beta$.

Definition

α and β are logically equivalent, written as $\alpha \mathrel{\models\!\!\!\dashv} \beta$, if $\alpha \models \beta$ and $\beta \models \alpha$.

Example

$$\forall x \forall y (P(x) \rightarrow \neg Q(y)) \mathrel{\models\!\!\!\dashv} \forall x \forall y (\neg(P(x) \wedge Q(y)))$$

Relations to Valid Wffs

Theorem

Let φ be a wff in the language $\mathbb{L}$. φ is valid if $\emptyset \vDash \varphi$.

Remark

We shall write $\emptyset \vDash \varphi$ as $\vDash \varphi$.

Soundness and Completeness

Our goal is to show the following are equivalent for any language $\mathbb{L}$:

- The set of provable wffs of $\mathbb{L}$;
- The set of valid wffs of $\mathbb{L}$.

This is accomplished by proving the following two theorems:

Theorem (Soundness)

Every provable wff is valid. That is, given any φ, $\vdash \varphi$ implies $\vDash \varphi$.

Theorem (Completeness)

Every valid wff is provable. That is, given any φ, $\vDash \varphi$ implies $\vdash \varphi$.

Soundness of First-Order Logic

Substitution Lemma

Lemma

Given a first-order language $\mathbb{L}$, let

- $\mathfrak{A}$ be a structure for $\mathbb{L}$ and s be an assignment function for $\mathfrak{A}$;
- u and t be terms and x be a variable.

Then

$$\bar{s}(u^x_t) = \overline{s(x|\bar{s}(t))}(u)$$

Lemma (Substitution Lemma)

Let s be an assignment function for $\mathfrak{A}$. If t is substitutable for x in α then

$$\models_{\mathfrak{A}} \alpha^x_t[s] \iff \models_{\mathfrak{A}} \alpha[s(x|\bar{s}(t))]$$

Soundness Theorem

A general form of soundness:

Theorem

If $\Sigma \vdash \alpha$, then $\Sigma \vDash \alpha$.

Proof.

By induction on the proof trees of $\Sigma \vdash \alpha$. Substitution lemma is needed when the last rule is a quantifier rule (i.e., $\forall$-*I*, $\forall$-*E*, $\exists$-*I*, and $\exists$-*E*). $\square$

Corollary

If $\vdash \alpha$, then $\vDash \alpha$.

Remark

Soundness is useful for showing certain wff is not provable. That is, if $\nvDash \varphi$, then φ is not provable in natural deduction.

Soundness and Logical Equivalence

Corollary

If $\vdash \varphi \leftrightarrow \psi$ then φ and ψ are logically equivalent.

Proof.

Show $\varphi \vDash \psi$ and $\psi \vDash \varphi$. $\square$

Consistency

Definition

- Σ is inconsistent if there is some wff α such that $\Sigma \vdash \alpha$ and $\Sigma \vdash \neg\alpha$.
- Σ is consistent if it is not inconsistent.

Properties of consistency:

Proposition

- If Σ is inconsistent then for every β, $\Sigma \vdash \beta$;
- $\Sigma \nvdash \alpha$ iff $\Sigma; \neg\alpha$ is consistent;
- Σ is consistent iff every finite subset of Σ is consistent;
- If Σ is consistent then for every α, either $\Sigma; \alpha$ is consistent or $\Sigma; \neg\alpha$ is consistent.

Alternative Statement of Soundness

From soundness we can derive the following property:

Corollary

If Σ is satisfiable then Σ is consistent.

In fact, soundness is equivalent to the above statement:

Theorem

The following to statements are equivalent:

- For any Σ and α, if $\Sigma \vdash \alpha$, then $\Sigma \vDash \alpha$;
- For any Σ, if Σ is satisfiable then Σ is consistent.

Completeness of First-Order Logic

Completeness

Theorem (Gödel Extended Completeness Theorem)

If $\Sigma \models \alpha$, then $\Sigma \vdash \alpha$.

An immediate consequence is

Corollary (Gödel Completeness Theorem)

If $\models \alpha$, then $\vdash \alpha$.

There is no easy inductive proof of completeness.

Alternative Statement of Completeness

Similar to Soundness, Completeness has an equivalent expression:

Theorem

The following to statements are equivalent:

- For any Σ and α, if $\Sigma \vDash \alpha$, then $\Sigma \vdash \alpha$;
- For any Σ, if Σ is consistent then Σ is satisfiable.

That is, to prove completeness, its suffices to show any consistent set of wffs is satisfiable!

A Hilbert-Style Deduction Calculus

A Hilbert-Style Deduction Calculus

We shall use a deduction calculus equivalent to natural deduction in Hilbert's style.

Let $\mathbb{L}$ be a first-order language. The calculus contains

- A set Λ of wffs called logical axioms and
- A single rule of inference for forming a new wff from a pairs of wffs.

We then systematically generate a set of wffs from the logical axioms by using the rule of inference. They are called provable wffs.

Logical Symbols

We only have two logical connectives: $\rightarrow$ and $\neg$.

Other connectives are obtained from the following abbreviations:

- $(\alpha \vee \beta)$ abbreviates $((\neg\alpha) \rightarrow \beta)$;
- $(\alpha \wedge \beta)$ abbreviates $(\neg(\alpha \rightarrow (\neg\beta)))$;
- $(\alpha \leftrightarrow \beta)$ abbreviates $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$;
- $\exists x\alpha$ abbreviates $(\neg\forall x(\neg\alpha))$.

Generalizations

Definition

A generalization of the wff α is any wff obtained by putting zero or more universal quantifiers in front of α.

Example

- $\forall x \forall x \forall y\ \alpha$ is a generalization of α;
- Every wff is a generalization of itself.

The Logical Axioms

Definition

Let $\mathbb{L}$ be a first-order language. The set Λ of logical axioms of $\mathbb{L}$ consists of all generalizations of the wffs in the following groups.

1. Instances of tautologies;
2. Wffs of the form $\forall x\, \alpha \to \alpha^x_t$ such that the term t is substitutable for x in α;
3. Wffs of the form $\forall x(\alpha \to \beta) \to (\forall x\, \alpha \to \forall x\, \beta)$;
4. Wffs of the form $\alpha \to \forall x\, \alpha$ such that x does not occur *free* in α;
5. Wffs of the form $x \doteq x$;
6. Wffs of the form $x \doteq y \to (\alpha \to \alpha')$ such that α is *atomic* and α' is obtained from α by replacing zero or more *occurrences* of x in α by y.

Instances of Wffs of Sentential Logic

Definition

Let

- $\alpha_1, \ldots, \alpha_n, \ldots$ be an infinite sequence of wffs of the first-order language $\mathbb{L}$;
- φ be a wff of sentential logic with just the connective symbols $\neg$ and $\rightarrow$;
- φ^* the wff of $\mathbb{L}$ obtained by replacing in φ, (for each n) every occurrence of the sentence symbol A_n by α_n.

We say that φ^* is an instance of φ.

Example

Example

Let $\alpha_1 = \forall y \neg P(y)$ and $\alpha_2 = P(x)$.

- Let $\varphi = (A_1 \to A_2) \to \neg A_2$.

$$\varphi^* = (\forall y \neg P(y) \to P(x)) \to \neg P(x).$$

- Let $\varphi = (A_1 \to \neg A_2) \to (A_2 \to \neg A_1)$.

$$\varphi^* = (\forall y \neg P(y) \to \neg P(x)) \to (P(x) \to \neg \forall y \neg P(y)).$$

A Rule of Inference

In Hilbert-style system, there is only one rule of inference:

Definition

Given any wffs α and β, the rule of modus ponens provides the operation for deriving β from $\alpha \to \beta$ and α.

We often say that β is inferred from α and $\alpha \to \beta$ by modus ponens.

Remark

The rule of modus ponens is a template for certain derivations.

Deductions

Definition

Let Σ be a set of wffs of $\mathbb{L}$. A deduction from Σ is a finite sequence

$$\alpha_0, \ldots, \alpha_n$$

of wffs such that every $\alpha_i (0 \leq i \leq n)$ is either

- in Σ, or
- in Λ, or
- is inferred by modus ponens from two wffs α_j and $\alpha_k = \alpha_j \to \alpha_i$ such that $j, k < i$.

Definition

$\Sigma \vdash \alpha$ (α is a theorem of Σ) if there is a deduction $\alpha_0, \ldots, \alpha_n$ from Σ such that $\alpha = \alpha_n$. (We write $\vdash \alpha$ for $\emptyset \vdash \alpha$.)

Example

- Show $\vdash P(x) \rightarrow \exists y P(y)$:

$$P(x) \rightarrow \neg\forall y \neg P(y)$$

$$\forall y \neg P(y) \rightarrow \neg P(x) \qquad\qquad (\forall y \neg P(y) \rightarrow \neg P(x)) \rightarrow (P(x) \rightarrow \neg\forall y \neg P(y))$$

(in axiom group 2) (in axiom group 1)

Example

- Show $\vdash \forall x(P(x) \to \exists y P(y))$.

$$\forall x(P(x) \to \neg\forall y \neg P(y))$$

$$\forall x(\forall y \neg P(y) \to \neg P(x)) \qquad \forall x(\forall y \neg P(y) \to \neg P(x)) \to \forall x(P(x) \to \neg\forall y \neg P(y))$$

(in axiom group 2)

$$\forall x((\forall y \neg P(y) \to \neg P(x)) \to (P(x) \to \neg\forall y \neg P(y)))$$ (a member of axiom group 3)

(in axiom group 1)

Equivalence to Natural Deduction

Let $\vdash_N$ denotes provability in natural deduction and $\vdash_H$ denotes provability in the Hilbert-style deduction calculus.

Theorem

Given any Σ and φ, $\Sigma \vdash_N \varphi$ iff $\Sigma \vdash_H \varphi$.

Proof.

Proof by structural induction in both directions. $\square$

Proof of Completeness

Proof of Completeness

Theorem

If Σ is consistent then Σ is satisfiable.

The proof is similar to that for compactness:

- Extend Σ to $\Delta \supseteq \Sigma$ such that Δ is consistent and maximal (i.e., for any α, either $\alpha \in \Delta$ or $\neg\alpha \in \Delta$).
- Define a structure $\mathfrak{A}$ and an assignment s for $\mathfrak{A}$ such that $\mathfrak{A}$ satisfies Δ with s.

The actual proof is more complex because the need to deal with $\doteq$.

Step One: Expanding Language with Constants

Step 1: Let Σ be a consistent set of wffs in a countable language. Expand the language with a countably infinite set of new constant symbols $c_1, \ldots, c_n, \ldots$.

Remark

Σ is consistent in the new language.

Step Two: Preparing for Satisfiability of Quantified Wffs

Step 2: In the new language, for any pair of wff φ and variable x, introduce a formula

$$\neg\forall x\varphi \to \neg\varphi^x_c$$

where c is a new constant symbol. Let Θ be the set of all these formulas.

Remark

- c identifies a counter example for φ;
- $\Sigma \cup \Theta$ is consistent.

Step Three: Get Maximally Consistent Set

Step 3: Extend $\Sigma \cup \Theta$ to a set Δ of wffs such that

- Δ is consistent, and
- for any wff α, $\alpha \in \Delta$ or $\neg\alpha \in \Delta$, but not both.

Remark

$\Sigma \cup \Theta$ is consistent implies that there is a truth assignment v that satisfies $\Sigma \cup \Theta \cup \Lambda$ in the sense of sentential logic. Pick

$$\Delta = \{\varphi \mid \bar{v}(\varphi) = T\}.$$

Step Four: Make a Structure for the New Language

Step 4: Make a structure $\mathfrak{A}$ from Δ for the new language where $\doteq$ is replaced by a 2-ary symbol E:

- $|\mathfrak{A}| =$ the set of all terms in the new language;
- $(u, t) \in E^{\mathfrak{A}} \iff u \doteq t \in \Delta$;
- For any n-ary predicate symbol:
$$(t_1, \ldots, t_n) \in P^{\mathfrak{A}} \iff P(t_1, \ldots, t_n) \in \Delta;$$
- For any n-ary function symbol:
$$f^{\mathfrak{A}}(t_1, \ldots, t_n) = f(t_1, \ldots, t_n);$$
- For any constant symbol c, $c^{\mathfrak{A}} = c$.

Step Four: Make a Structure for the New Language

Step 4 (Cont'd): Make an assignment function $s : V \to |\mathfrak{A}|$

$$s(x) = x.$$

Then $\overline{s}(t) = t$. For any wff φ, let φ^* be φ with $\doteq$ replaced by E. We have

$$\models_{\mathfrak{A}} \varphi^*[s] \iff \varphi \in \Delta.$$

If there is no $\doteq$ in our language, then we are done. Δ, hence Σ is satisified by $\mathfrak{A}$ with s. However, there is more to do if $\doteq$ is present.

Step Five: Deal with Equality

Question

If there are two constant c and d, and Σ contains $c \doteq d$. We need to $c^{\mathfrak{A}} = d^{\mathfrak{A}}$ for $\mathfrak{A}$ to satisfy Σ. This may not be true.

Step 5: Construct a quotient structure $\mathfrak{A}/E$ by identifying $E^{\mathfrak{A}}$ as a congruence relation for $\mathfrak{A}$.

Step Five: Deal with Equality

Proposition

$E^{\mathfrak{A}}$ as a congruence relation for $\mathfrak{A}$ in the following senses:

- $E^{\mathfrak{A}}$ is an equivalence relation on $|\mathfrak{A}|$;
- For each n-ary predicate symbol P, $P^{\mathfrak{A}}$ is compatible with $E^{\mathfrak{A}}$, meaning:

$$\text{If } (t_1, \ldots, t_n) \in P^{\mathfrak{A}} \text{ and } (t_i, t_i') \in E^{\mathfrak{A}} \text{ for } (1 \leq i \leq n)$$
$$\text{then } (t_1', \ldots, t_n') \in P^{\mathfrak{A}};$$

- For each n-ary function symbol f, $f^{\mathfrak{A}}$ is compatible with $E^{\mathfrak{A}}$, meaning:

$$\text{If } (t_i, t_i') \in E^{\mathfrak{A}} \text{ for } (1 \leq i \leq n)$$
$$\text{then } (f^{\mathfrak{A}}(t_1, \ldots, t_n), f^{\mathfrak{A}}(t_1', \ldots, t_n')) \in P^{\mathfrak{A}}.$$

Step Five: Deal with Equality

Definition

The quotient structure $\mathfrak{A}/E$ is defined as follows:

- $|\mathfrak{A}/E|$ is the set of all equivalent classes of members of $|\mathfrak{A}|$;
- For each n-ary predicate symbol P,

$$([t_1], \ldots, [t_n]) \in P^{\mathfrak{A}/E} \iff (t_1, \ldots, t_n) \in P^{\mathfrak{A}};$$

- For each n-ary function symbol f,

$$f^{\mathfrak{A}/E}([t_1], \ldots, [t_n]) = [f^{\mathfrak{A}}(t_1, \ldots, t_n)];$$

- For each constant symbol c,

$$c^{\mathfrak{A}/E} = [c^{\mathfrak{A}}].$$

Note that $\mathfrak{A}/E$ is well-defined because $E^{\mathfrak{A}}$ is a congruence relation.

Step Five: Deal with Equality

Proposition

Let $h : |\mathfrak{A}| \to |\mathfrak{A}/E|$ be
$$h(t) = [t].$$
Then $\mathfrak{A}/E$ satisifies Δ with $h \circ s$.

Proof.

By definition, we have the following facts:

(a) h is a homomorphism of $\mathfrak{A}$ into $\mathfrak{A}/E$;

(b) $E^{\mathfrak{A}/E}$ is the equality relation on $|\mathfrak{A}/E|$.

Therefore, we have

$$\begin{aligned}
\varphi \in \Delta &\iff \models_{\mathfrak{A}} \varphi^*[s] \\
&\iff \models_{\mathfrak{A}/E} \varphi^*[h \circ s] \qquad \text{by (a)} \\
&\iff \models_{\mathfrak{A}/E} \varphi[h \circ s] \qquad \text{by (b)}
\end{aligned}$$

□

Step Six: Wrapping Up

Step 6: Restrict $\mathfrak{A}/E$ to the original language, then it also satisifies Δ (hence Σ) with $h \circ s$.

We have now proved that Σ is satisfiable if it is consistent! Done!

Compactness Theorem

Compactness of First-Order Logic is a corollary of its Completeness. It has the following to equivalent statements:

Theorem (Compactness Theorem)

- If $\Sigma \vDash \varphi$, then there exists a finite subset Σ' of Σ such that $\Sigma' \vDash \varphi$;
- If every finite subset Σ' of Σ is satisfiable, then Σ is satisfiable.