First Order Logic

Based on slides by Thom Fruewirth

Foundations from Logic

> *Good, too, Logic, of course; in itself, but not in fine weather.*

Arthur Hugh Clough, 1819-1861

> *Die Logik muß für sich selber sorgen.*

Ludwig Wittgenstein, 1889-1951

First-Order Logic

Syntax – Language

- Alphabet
- Well-formed Expressions

Semantics – Meaning

- Interpretation
- Logical Consequence
- Herbrand models*

Calculi – Derivation

- Refutation Inference Rule
    - Clausal normal form for FOL
    - Substitutions
    - Unification

Why First-Order Logic

Propositional Logic is lacking structure:

- P = Every man is mortal
- S = Socrate is a man
- Q = Socrate is mortal

In propositional logic Q is not a logic consequence of P and S, but we would like to express this relationship.

Syntax of First-Order Logic

Signature $\Sigma = (\mathcal{P}, \mathcal{F})$ of a first-order language

- $\mathcal{P}$: finite set of *predicate symbols*, each with *arity* $n \in \mathbb{N}$
- $\mathcal{F}$: finite set of *function symbols*, each with *arity* $n \in \mathbb{N}$

Naming conventions:

- *nullary, unary, binary, ternary* for arities $0, 1, 2, 3$
- *constants*: nullary function symbols
- *propositions*: nullary predicate symbols

Syntax of First-Order Logic (ctnd)

Alphabet

- $\mathcal{P}$: predicate symbols: $p, q, r, \ldots$
- $\mathcal{F}$: function symbols: $a, b, c, \ldots, f, g, h, \ldots$
- $\mathcal{V}$: countably infinite set of variables: $X, Y, Z, \ldots$
- logic symbols:
  - truth symbols: $\perp$ (false), $\top$ (true)
  - logical connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$
  - quantors: $\forall$, $\exists$
  - syntactic symbols: "(", ")", ","

Term

Set of *terms* $\mathcal{T}(\Sigma, \mathcal{V})$:

- a *variable* from $\mathcal{V}$, or
- a *function term* $f(\overline{t})$, where f is an n-ary function symbol from Σ and the *arguments* $\overline{t}$ are terms ($n \geq 0$).

Examples ($a/0$, $f/1$, $g/2$):

- X
- a
- $f(X)$
- $g(f(X), g(Y, f(a)))$

Well-Formed Expressions (ctnd)

Well-Formed Formula

Set of *(well-formed) formulae* $\mathcal{F}(\Sigma, \mathcal{V}) = \{A, B, C, \ldots, F, G, \ldots\}$:

- an *atomic formula* (*atom*) $p(\overline{t})$, where p is an n-ary predicate symbol from Σ and the *arguments* $\overline{t}$ are terms, or
  $\perp$, or
  $\top$, or
  $t_1 \doteq t_2$ for terms t_1 and t_2 (i.e., $\doteq(t_1, t_2)$), or
- the *negation* $\neg F$ of a formula F, or
- the *conjunction* $(F \wedge F')$, the *disjunction* $(F \vee F')$, or the *implication* $(F \rightarrow F')$ between two formulae F and F', or
- a *universally quantified formula* $\forall X F$, or an *existentially quantified formula* $\exists X F$, where X is a variable and F is a formula.

Example – Terms

$\mathcal{P} = \{\texttt{mortal/1}\}$, $\mathcal{F} = \{\texttt{socrates/0}, \texttt{father/1}\}$, $\mathcal{V} = \{X, \dots\}$

- *Terms*:
  X, `socrates`, `father(socrates)`,
  `father(father(socrates))`,
  but not: $\texttt{father}(X, \texttt{socrates})$

Example – Formulae

- *Atomic Formulae*:
  $\texttt{mortal}(X)$, $\texttt{mortal(socrates)}$,
  $\texttt{mortal(father(socrates))}$
  but not: $\texttt{mortal(mortal(socrates))}$
- *Non-Atomic Formulae*:
  $\texttt{mortal(socrates)} \ \wedge \ \texttt{mortal(father(socrates))}$
  $\forall X.\texttt{mortal}(X) \rightarrow \texttt{mortal(father}(X))$
  $\exists X.X \doteq \texttt{socrates}$

Free Variables

- quantified formula $\forall X F$ or $\exists X F$ *binds* variable X within *scope* F
- set $Fv(F)$ of free (not bounded) variables of a formula F:

$$
\begin{aligned}
Fv(t_1 \doteq t_2) &:= vars(t_1) \cup vars(t_2) \\
Fv(p(\overline{t})) &:= \cup vars(\overline{t}) \\
Fv(\top) := Fv(\bot) &:= \emptyset \\
Fv(\neg F) &:= Fv(F) \text{ for a formula } F \\
Fv(F * F') &:= Fv(F) \cup Fv(F') \text{ for formulae } F \text{ and } F' \\
& \qquad\qquad \text{and } * \in \{\wedge, \vee, \rightarrow\} \\
Fv(\forall X F) := Fv(\exists X F) &:= Fv(F) \setminus \{X\} \text{ for a variable } X \\
& \qquad\qquad \text{and a formula } F
\end{aligned}
$$

Example – Free Variables

Give the set of free variables for each braced part.

$$\underbrace{p(X) \ \wedge \ \overbrace{\exists X p(X)}}$$

$$\overbrace{\underbrace{(\forall X \, \underbrace{p(X, Y)})} \ \vee \ \underbrace{q(X)}}$$

Universal and Existential Closure of F

- *universal closure* $\forall F$ of F:

$$\forall X_1 \forall X_2 \dots \forall X_n F$$

- *existential closure* $\exists F$ of F:

$$\exists X_1 \exists X_2 \dots \exists X_n F$$

where $X_1, X_2, \dots, X_n$ are all free variables of F

Naming Conventions:

- *closed formula* or *sentence*: does not contain free variables
- *theory*: set of sentences
- *ground term or formula*: does not contain any variables

Semantics of FOL – Meaning

Semantics of First-Order Logic

- We want now to understand which is the *meaning* of a formula written in FOL
- The presence of variables makes this task more difficult than in the case of Propositional Logic: id we write $X \geq 1$ of course the truth of this formula depends on the value of X and on the meaning that we give to the predicate $\geq$ and to the function (constant) 1
- We need to fix:
  - A *universe*, i.e. a domain from which we take the values (numbers, graphs, apples, animals ....)
  - An *interpretation* of function and predicate symbols (on the given Universe)
  - A *valuation* for variables.

Semantics of First-Order Logic

Interpretation I of Σ

Consists of:

- A *universe* U, that is a non-empty set
- A function $I(f) : U^n \rightarrow U$ for every n-ary function symbol f of Σ
- A relation $I(p) \subseteq U^n$ for every n-ary predicate symbol p of Σ

Variable Valuation for $\mathcal{V}$ w.r.t. I

- $\eta : \mathcal{V} \rightarrow U$: for every variable X of $\mathcal{V}$ into the universe U of I

(Σ signature of a first-order language, $\mathcal{V}$ set of variables)

Interpretation of Terms

Given Σ signature, I interpretation with universe U, $\eta : V \to U$ variable valuation, the function

$$\eta^I : \mathcal{T}(\Sigma, V) \to U$$

defined as follows provides the interperation of terms:

$$\eta^I(X) := \eta(X) \text{ for a variable } X$$

$$\eta^I(f(t_1, \ldots, t_n))) := I(f)(\eta^I(t_1), \ldots, \eta^I(t_n))$$

$$\text{for an } n\text{-ary function symbol } f \text{ and terms } \mathrm{t}_1, \ldots, t_n$$

Example – Interpretation of Terms

- signature $\Sigma = (\emptyset, \{1/0, */2, +/2\})$
- universe $U = \mathbb{N}$
- interpretation I
  $I(1) := 1$
  $I(A{+}B) := A + B$, $I(A{*}B) := AB$ for $A, B \in U$
- variable valuation η
  $\eta(X) := 3$, $\eta(Y) := 5$

$$\begin{aligned}
\eta^I\big(X{*}(Y{+}1)\big) &= I(*)\big(\eta^I(X), \eta^I(Y{+}1)\big) \\
&= I(*)(3, I(+)\big(\eta^I(Y), \eta^I(1)\big) \\
&= \cdots = 3 \cdot (5 + 1) = 18
\end{aligned}$$

Interpretation of Formulae – Preliminaries

For a function g, the function $g[Y \mapsto a]$ is

$$g[Y \mapsto a](X) := \begin{cases} g(X) & \text{if } X \neq Y, \\ a & \text{if } X = Y, \end{cases}$$

where X and Y are variables.

Given an interpretation of I of Σ, a variable valuation η, we must define when I and η make a formula F true or, to be more precise, when I and η *satisfy* a formula F, written $I, \eta \models F$. This is defined in the next slide,

Interpretation of Formulae

Given I interpretation of Σ, η variable valuation, $I, \eta \models F$ (I, η satisfies F) is defined as follows:

- $I, \eta \models \top$ and $I, \eta \not\models \bot$
- $I, \eta \models s \doteq t$ iff $\eta^I(s) = \eta^I(t)$
- $I, \eta \models p(t_1, \ldots, t_n)$ iff $(\eta^I(t_1), \ldots, \eta^I(t_n)) \in I(p)$
- $I, \eta \models \neg F$ iff $I, \eta \not\models F$
- $I, \eta \models F \wedge F'$ iff $I, \eta \models F$ and $I, \eta \models F'$
- $I, \eta \models F \vee F'$ iff $I, \eta \models F$ or $I, \eta \models F'$
- $I, \eta \models F \rightarrow F'$ iff $I, \eta \not\models F$ or $I, \eta \models F'$
- $I, \eta \models \forall X F$ iff $I, \eta[X \mapsto u] \models F$ for all $u \in U$
- $I, \eta \models \exists X F$ iff $I, \eta[X \mapsto u] \models F$ for some $u \in U$

Example – Interpretation

$$\boxed{\forall X . p(X, a, b) \rightarrow q(b, X)}$$

.	$I_1(.)$	$I_2(.)$
U	real things	natural numbers
a	"food"	5
b	"Fitz the cat"	10
p	"_ gives _ _"	$\_ + \_ > \_$
q	"_ loves _"	$\_ < \_$

I_1: "Fitz the cat loves everybody who gives him food."
I_2: "10 is less than any X if $X + 5 > 10$."
(counterexample: $X = 6$)

Model of F, Validity

Given a signature Σ, an interpretation I and a formula F:

- We say that I *model of* F or I *satisfies* F, written $I \models F$ when: $I, \eta \models F$ for every variable valuation η
- I a *model* of theory T when I is a model of each formula in T

Sentence S is

- *valid*: satisfied by every interpretation, i.e., $I \models S$ for every I
- *satisfiable*: satisfied by some interpretation, i.e., $I \models S$ for some I
- *falsifiable*: not satisfied by some interpretation, i.e., $I \not\models S$ for some I
- *unsatisfiable*: not satisfied by any interpretation, i.e., $I \not\models S$ for every I

Example – Validity, Satisfiability, Falsifiable, Unsatisfiability (1)

	valid	satisfiable	falsifiable	unsatisfiable
$A \vee \neg A$				
$A \wedge \neg A$				
$A \rightarrow \neg A$				
$A \rightarrow (B \rightarrow A)$				
$A \rightarrow (A \rightarrow B)$				
$A \leftrightarrow \neg A$				

(A, B formulae)

Example – Validity, Satisfiability, Unsatisfiability (2)

	correct/counter example
If F is valid, then F is satisfiable.	
If F is satisfiable, then $\neg F$ is unsatisfiable.	
If F is valid, then $\neg F$ is unsatisfiable.	
If F is unsatisfiable, then $\neg F$ is valid.	

(F formula)

Example – Interpretation and Model

- signature $\Sigma = (\{\texttt{pair/2}\}, \{\texttt{next/1}\})$
- universe $U = \{Mon, Tue, Wed, Thu, Fri, Sat, Sun\}$
- interpretation I

$$I(\texttt{pair}) := \left\{ \begin{array}{cccc} (Mon, Tue), & (Mon, Wed), & \ldots, & (Mon, Sun), \\ & (Tue, Wed), & \ldots, & (Tue, Sun), \\ & & \ddots & \vdots \\ & & & (Sun, Sun) \end{array} \right\}$$

$I(\texttt{next}) : U \to U$, next day function
$Mon \mapsto Tue, \ldots, Sun \mapsto Mon$

Example – Interpretation and Model (cont)

- (one possible) valuation η
  $\eta(X) := \mathit{Sun}$, $\eta(Y) := \mathit{Tue}$
  - $\eta^I(\mathtt{next}(X)) = \mathit{Mon}$
  - $\eta^I(\mathtt{next}(\mathtt{next}(Y))) = \mathit{Thu}$
  - $I, \eta \models \mathtt{pair}(\mathtt{next}(X), Y)$
- model relationship ("for all variable valuations")
  - $I \not\models \forall X \forall Y.\mathtt{pair}(\mathtt{next}(X), Y)$
  - $I \models \forall X.\mathtt{pair}(\mathtt{next}(X), \mathit{Sun})$
  - $I \models \forall X \exists Y.\mathtt{pair}(X, Y)$

Logical Consequence

- A sentence/theory T_1 is a *logical consequence* of a sentence/theory T_2, written $T_2 \models T_1$, if every model of T_2 is also a model of T_1, i.e. $I \models T_2$ implies $I \models T_1$.
- Two sentences or theories are *equivalent* ($\Leftrightarrow$) if they are logical consequences of each other.
- Theorem. It is undecidable whether a first order logic formula F is true under alla possible interpretations, i.e. if $\models F$. [Church]

Example:
$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$ (de Morgan)

Example – Tautology Laws (1)

Dual laws hold for $\wedge$ and $\vee$ exchanged.

- $A \Leftrightarrow \neg\neg A$ (double negation)
- $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$ (de Morgan)
- $A \wedge A \Leftrightarrow A$ (idempotence)
- $A \wedge (A \vee B) \Leftrightarrow A$ (absorption)
- $A \wedge B \Leftrightarrow B \wedge A$ (commutativity)
- $A \wedge (B \wedge C) \Leftrightarrow (A \wedge B) \wedge C$ (associativity)
- $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge B)$ (distributivity)

Example – Tautology Laws (2)

- $A \rightarrow B \Leftrightarrow \neg A \vee B$ (implication)
- $A \rightarrow B \Leftrightarrow \neg B \rightarrow \neg A$ (contraposition)
- $(A \rightarrow (B \rightarrow C)) \Leftrightarrow (A \wedge B) \rightarrow C$
- $\neg \forall X A \Leftrightarrow \exists X \neg A$
- $\neg \exists X A \Leftrightarrow \forall X \neg A$
- $\forall X (A \wedge B) \Leftrightarrow \forall X A \wedge \forall X B$
- $\exists X (A \vee B) \Leftrightarrow \exists X A \vee \exists X B$
- $\forall X B \Leftrightarrow B \Leftrightarrow \exists X B$ (with X not free in B)

Example – Logical Consequence

F	G	$F \models G$ or $F \not\models G$
A	$A \vee B$	
A	$A \wedge B$	
A, B	$A \vee B$	
A, B	$A \wedge B$	
$A \wedge B$	A	
$A \vee B$	A	
$A, (A \rightarrow B)$	B	

Note: I is a model of $\{A, B\}$, iff $I \models A$ and $I \models B$.

Example – Logical Consequence, Validity, Unsatisfiability

The following statements are equivalent:

1. $F_1, \ldots, F_k \models G$
   (G is a logical consequence of $F_1, \ldots, F_k$)
2. $\left(\bigwedge_{i=1}^{k} F_i\right) \rightarrow G$ is valid.
3. $\left(\bigwedge_{i=1}^{k} F_i\right) \wedge \neg G$ is unsatisfiable.

Note: In general, $F \not\models G$ does *not* imply $F \models \neg G$.

Herbrand models

Herbrand Interpretation – Motivation

The formula A is valid in I, $I \models A$, if $I, \eta \models A$ for every valuation η. This requires to fix a universe U as both I and η use U.

Jacques Herbrand (1908-1931) discovered that there is a *universal* domain together with a *universal* interpretation, s.t. that any *universally* valid formula is valid in *any* interpretation.

Therefore, only interpretations in the *Herbrand universe* need to be checked, provided the Herbrand universe is infinite.

Herbrand Interpretation

- *Herbrand universe*: set $\mathcal{T}(\Sigma, \emptyset)$ of ground terms
- for every n-ary function symbol f of Σ, the assigned function $I(f)$ maps a tuple $(\bar{t})$ of ground terms to the ground term "$f(\bar{t})$".
- *Herbrand base* for signature Σ: set of ground atoms in $\mathcal{F}(\Sigma, \emptyset)$, i.e., $\{p(\bar{t}) \mid p$ is an n-ary predicate symbol of Σ and $\bar{t} \in \mathcal{T}(\Sigma, \emptyset)\}$.
- *Herbrand model* of sentence/theory: Herbrand interpretation satisfying sentence/theory.

Example – Herbrand Interpretation

For the formula $F \equiv \exists X \exists Y . p(X, a) \wedge \neg p(Y, a)$ the Herbrand universe is $\{a\}$ and F is unsatisfiable in the Herbrand universe as $p(a, a) \wedge \neg p(a, a)$ is false, i.e., there is no Herbrand model. However, if we add (another element) b we have $p(a, a) \wedge \neg p(b, a)$ so F is valid for any interpretation whose universe' cardinality is greater than 1.

For the formula $\forall X \forall Y . p(X, a) \wedge q(X, f(Y))$ the (infinite, as there is a constant and a function symbol) Herbrand domain is $\{a, f(a), f(f(a)), f(f(f(a))), \ldots\}$.

Herbrand theorem (simple version)

Let P be a set of universal sentences. The following are equivalent:

1. P has an Herbrand model
2. P has a model
3. ground(P) is satisfiable

Proof of Herbrand theorem*

(1) $\Rightarrow$ (2) Obvious
(2) $\Rightarrow$ (3) Every sentence in ground(P) is a logical consequence of P (proof as exercise). Hence every model of P is a model of ground(P).
(3) $\Rightarrow$ (1) If ground(P) is satisfiable then ground(P) has an Herbrand model **A**. If fact let M be a model of ground(P). Then we can define **A** in the usual way for function symbols, while for atomic fromulas we can define $\mathbf{A} \models p(t_1, ..., t_n)$ iff $\mathbf{M} \models p(t_1, ..., t_n)$ and then inductively for arbitrary formulas (details for exercise).
Now we have that **A** is also a model of P. In fact, assume that $A \models ground(\forall\phi)$ where ϕ is quantifier free and $Var(\phi) = \{x_1, \ldots, x_n\}$. Let $t_1, \ldots, t_n$ be n be n arbitrary ground terms and define $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$. Since $A \models ground(\forall\phi)$ we have that $A \models \phi\theta$. Since the terms t_i are generic elements of the domain of **A** this means that $A \models \forall\phi$ and conludes the proof.

Calculus for FOL: Resolution

We have seen the notion of calculus:

Calculus

- *axioms*: given formulae, elementary tautologies and contradictions which cannot be derived within the calculus
- *inference rules*: allow to derive new formulae from given formulae
- *derivation* $\phi \vdash \psi$: a sequence of inference rule applications starting with formula ϕ and ending in formula ψ

$\models$ and $\vdash$ should coincide

- *Soundness*: $\phi \vdash \rho$ implies $\phi \models \rho$
- *Completeness*: $\phi \models \rho$ implies $\phi \vdash \rho$

The question is: can we define a correct and complete calculus for FOL? And can we use resolution?

Resolution calculus for FOL

The idea is the same as for propositional logic: Theory united with negated consequence must be unsatisfiable.

However we have first to transform FOL formulae in clauses, which is more complicated

Then we must define the resolution inference rule for FOL: this requires substitusions and unifcation

Negation Normal Form of Formula F

F in *negation normal form* F_{neg}:

- no sub-formula of the form $F \rightarrow F'$
- in every sub-formula of the form $\neg F'$ the formula F' is atomic

Normal forms and clauses for FOL

Negation Normal Form – Computation

For every sentence F, there is an equivalent sentence F_{neg} in negated normal form (apply "tautologies" from left to right.):

Negation

$$\neg\bot \Leftrightarrow \top \qquad \neg\top \Leftrightarrow \bot$$

$$\neg\neg F \Leftrightarrow F \quad F \text{ is atomic}$$

$$\neg(F \wedge F') \Leftrightarrow \neg F \vee \neg F' \qquad \neg(F \vee F') \Leftrightarrow \neg F \wedge \neg F'$$

$$\neg\forall X F \Leftrightarrow \exists X \neg F \qquad \neg\exists X F \Leftrightarrow \forall X \neg F \qquad \neg(F \to F') \Leftrightarrow F \wedge \neg F'$$

Implication

$$F \to F' \Leftrightarrow \neg F \vee F'$$

Skolemization of Formula F

- $F = F_{\text{neg}} \in \mathcal{F}(\Sigma, \mathcal{V})$ in negation normal form
- occurrence of sub formula $\exists X G$ with free variables $\overline{V}$
- f/n function symbol not occurring in Σ

Compute F' by replacing $\exists X G$ with $G[X \mapsto f(\overline{V})]$ in F (all occurrences of X in G are replaced by $f(\overline{V})$).

Naming conventions:

- *Skolemized form* of F: F'
- *Skolem function*: f/n

Equivalence of Skolemized form

- A formula F is satisfiable iff its skolemized form is satisfiable
- A formula F and its skolemized form are not logically equivalent

Consider $F = \forall X.\exists Y.p(a, Y, a, b)$ and consider an interpretation I where f(a)=b (f skolem function) and only p(a,a,a,b) holds.

I is a model of F, but I is not a model of the skolemized form of F

Prenex Form of Formula F

$F = Q_1 X_1 \dots Q_n X_n G$

- Q_i quantifiers
- X_i variables
- G formula without quantifiers

- *quantifier prefix*: $Q_1 X_1 \dots Q_n X_n$
- *matrix*: G
- For every sentence F, there is an equivalent sentence in prenex form and it is possible to compute such a sentence from F by applying tautology laws to push the quantifieres outwards.

Example – Prenex Normal Form

$$\begin{aligned}\neg\exists x p(x) \vee \forall x r(x) &\Leftrightarrow \forall x \neg p(x) \vee \forall x r(x)\\ &\Leftrightarrow \forall x \neg p(x) \vee \forall y r(y)\\ &\Leftrightarrow \forall x \left(\neg p(x) \vee \forall y r(y)\right)\\ &\Leftrightarrow \forall x \forall y \left(\neg p(x) \vee r(y)\right)\end{aligned}$$

Example – Skolemization

$$\begin{aligned}&\forall z \exists x \forall y \left(p(x,z) \wedge q(g(x,y),x,z)\right)\\ &\qquad \Leftrightarrow \forall z \forall y \left(p(f(z),z) \wedge q(g(f(z),y),f(z),z)\right)\end{aligned}$$

With new function symbol $f/1$.

Clauses and Literals

- *literal*: atom (*positive* literal) or negation of atom (*negative* literal)
- *complementary* literals: positive literal L and its negation $\neg L$
- *clause (in disjunctive normal form)*: formula of the form $\bigvee_{i=1}^{n} L_i$ where L_i are literals.
  - *empty clause (empty disjunction)*: $n = 0$: $\perp$

Clauses and Literals (cont)

- *implication form* of the clause:

$$F = \underbrace{\bigwedge_{j=1}^{n} B_j}_{body} \rightarrow \underbrace{\bigvee_{k=1}^{m} H_k}_{head}$$

  for

$$F = \bigvee_{i=1}^{n+m} L_i \text{ with } L_i = \begin{cases} \neg B_i & \text{for } i = 1, \ldots, n \\ H_{i-n} & \text{for } i = n+1, \ldots, n+m \end{cases}$$

  for atoms B_j and H_k

- *closed clause*: sentence $\forall \overline{x} C$ with C clause
- *clausal form* of theory: consists of closed clauses

Normalization steps

An arbitrary theory T can be transformed into clausal form as follows

- Convert every formula in the theory into an equivalent formula in negation normal form.
- Perform Skolemization in order to eliminate all existential quantifiers.
- Convert the resulting theory, which is still in negation normal form, into an equivalent theory in clausal form: Move conjunctions and universal quantifiers outwards.

Substitutions and unification

Substitution

- A substitution σ is a mapping $\sigma : \mathcal{V} \to \mathcal{T}(\Sigma, \mathcal{V}')$ which modifies a *finite* number of variables, written as $\{X_1 \mapsto t_1, \ldots, X_n \mapsto t_n\}$ where X_i are distinct variables and t_i are terms.
- The *identity substitution* is indicated by $\epsilon = \emptyset$
- Often written as postfix operators, application from left to right in composition
- On terms a substitution $\sigma : \mathcal{T}(\Sigma, \mathcal{V}) \to \mathcal{T}(\Sigma, \mathcal{V}')$ is the implicit homomorphic extension, i.e., $f(\bar{t})\sigma := f(t_1\sigma, \ldots, t_n\sigma)$.

Example:
$\sigma = \{X \mapsto 2, Y \mapsto 5\}$: $(X * (Y + 1))\sigma = 2 * (5 + 1)$

Substitution applied to a Formula

Homomorphic extension

- $p(\bar{t})\sigma := p(t_1\sigma, \ldots, t_n\sigma)$
- $(s \doteq t)\sigma := (s\sigma) \doteq (t\sigma)$
- $\bot\sigma := \bot$ and $\top\sigma := \top$
- $(\neg F)\sigma := \neg(F\sigma)$
- $(F * F')\sigma := (F\sigma) * (F'\sigma)$ for $* \in \{\wedge, \vee, \rightarrow\}$

Except

- $(\forall X F)\sigma := \forall X'(F\sigma[X \mapsto X'])$
- $(\exists X F)\sigma := \exists X'(F\sigma[X \mapsto X'])$

where X' is a fresh variable.

Example – Application of substitution

- $\sigma = \{X \mapsto Y, Z \mapsto 5\}$: $(X * (Z + 1))\sigma = Y * (5 + 1)$
- $\sigma = \{X \mapsto Y, Y \mapsto Z\}$: $p(X)\sigma = p(Y) \neq p(X)\sigma\sigma = p(Z)$
- $\sigma = \{X \mapsto Y\}, \tau = \{Y \mapsto 2\}$
  - $(X * (Y + 1))\sigma\tau = (Y * (Y + 1))\tau = (2 * (2 + 1))$
  - $(X * (Y + 1))\tau\sigma = (X * (2 + 1))\sigma = (Y * (2 + 1))$
- $\sigma = \{X \mapsto Y\}$: $(\forall X p(3))\sigma = \forall X' p(3)$
- $\sigma = \{X \mapsto Y\}$: $(\forall X p(X))\sigma = \forall X' p(X')$, $(\forall X p(Y))\sigma = \forall X' p(Y)$
- $\sigma = \{Y \mapsto X\}$: $(\forall X p(X))\sigma = \forall X' p(X')$, $(\forall X p(Y))\sigma = \forall X' p(X)$

Logical Expression over $\mathcal{V}$

- *term* with variables in $\mathcal{V}$,
- *formula* with free variables in $\mathcal{V}$,
- *substitution* from an arbitrary set of variables into $\mathcal{T}(\Sigma, \mathcal{V})$, or
- *tuple* of logical expressions over $\mathcal{V}$.

A logical expression is a *simple expression* if it does not contain quantifiers.

Examples:
$\mathcal{V} = \{X\}$: $f(X)$, $\forall Y.p(X) \wedge q(Y)$, $\{X \mapsto a\}$, $\langle p(X), \sigma \rangle$

Instance, Variable Renaming, Variants

- e *instance of* e': $e = e'\sigma$
- e' *more general than* e: e is instance of e'
- *variable renaming* for e: substitution σ
  - σ injective
  - $X\sigma \in \mathcal{V}$ for all $X \in \mathcal{V}$
  - $X\sigma$ does not occur in e for free variables X of e
- e and e' *variants* (identical modulo variable renaming): $e = e'\sigma$ and $e' = e\tau$

Examples:
variable renaming $(\forall X p(X_1))\{X_1 \mapsto X_2\} = \forall X' p(X_2)$
but not $(\forall X p(X_1) \wedge q(X_2))\{X_1 \mapsto X_2\} = \forall X'.p(X_2) \wedge q(X_2)$
(e, e' logical expressions, σ, τ substitutions)

Unifier and m.g.u.

- σ is a *unifier* for $e_1, \ldots, e_n$ if $e_1\sigma = \cdots = e_n\sigma$
- $e_1, \ldots, e_n$ *unifiable* if unifier exists
- σ is *most general unifier (mgu)* for $e_1, \ldots, e_n$ if every unifier τ for $\bar{e}$ is instance of σ, i.e., $\tau = \sigma\rho$ for some ρ

($e_1, \ldots, e_n$ simple expressions, σ, τ, ρ, σ_i substitutions)

Example – Most General Unifier

$$f(X, a) \quad \doteq \quad f(g(U), Y) \quad \doteq \quad Z$$

MGU:

$$\sigma = \{X \mapsto g(U), Y \mapsto a, Z \mapsto f(g(U), a)\}$$

Proof: $f(X, a)\sigma = f(g(U), Y)\sigma = Z\sigma = f(g(U), a)$ one element.

Unifier, but not MGU:

$$\sigma' = \{X \mapsto g(h(b)), U \mapsto h(b), Y \mapsto a, Z \mapsto f(g(h(b)), a)\}$$

Proof: $\sigma' = \sigma\{U \mapsto h(b)\}$.

Most General Unifier by Hand

- *unbound variable*: there is no substitution for it
- Start with ϵ
- scan terms simultaneously from left to right according to their structure
- check the syntactic equivalence of the symbols encountered repeat
  - *different function symbols*: halt with failure
  - *identical*: continue
  - *one is unbound variable and other term*:
    - variable *occurs* in other term: halt with failure
    - apply the new substitution to the logical expressions

    add corresponding substitution
  - *variable is not unbound*: replace it by applying substitution

Example – Most General Unifier

to unify	current substitution, remarks
$p(X, f(a)) \doteq p(a, f(X))$	ϵ, start
$X \doteq a$	$\{X \mapsto a\}$, substitution added
$f(a) \doteq f(X)$	continue
$a \doteq X$	$\{X \mapsto a\}$, variable is not unbound
$a \doteq a$	continue

MGU is $\{X \mapsto a\}$

What about $p(X, f(b)) = p(a, f(X))$?

Example – Most General Unifier

s	t	
f	g	failure
a	a	ϵ
X	a	$\{X \mapsto a\}$
X	Y	$\{X \mapsto Y\}$, but also $\{Y \mapsto X\}$
$f(a, X)$	$f(Y, b)$	$\{Y \mapsto a, X \mapsto b\}$
$f(g(a, X), Y)$	$f(c, X)$	failure
$f(g(a, X), h(c))$	$f(g(a, b), Y)$	$\{X \mapsto b, Y \mapsto h(c)\}$
$f(g(a, X), h(Y))$	$f(g(a, b), Y)$	failure

Resolution inference rule for FOL

Resolution Calculus – Inference Rules

Works by contradiction: Theory united with negated consequence must be unsatisfiable ("derive empty clause").

Axiom

empty clause (i.e. the elementary contradiction)

Resolution

$$\frac{R \vee A \qquad R' \vee \neg A'}{(R \vee R')\sigma}$$

σ is a most general unifier for the atoms A and A'

Factoring

$$\frac{R \vee L \vee L'}{(R \vee L)\sigma}$$

σ is a most general unifier for the literals L and L'

$R \vee A$ and $R' \vee \neg A'$ must have different variables: rename variables apart

Resolution – Remarks

- *resolution rule*:
  - two clauses C and C' instantiated s.t. literal from C and literal from C' complementary
  - two instantiated clauses are combined into a new clause
  - *resolvent* added
- *factoring rule*:
  - clause C instantiated, s.t. two literals become equal
  - remove one literal
  - *factor* added

Example – Resolution Calculus

Resolution:

$$\frac{p(a, X) \vee q(X) \qquad \neg p(a, b) \vee r(X)}{(q(X) \vee r(X))\{X \mapsto b\}}$$

Factoring:

$$\frac{p(X) \vee p(b)}{p(X)\{X \mapsto b\}}$$

Refutation completeness of resolution

Theorem

Assume that $\vdash$ denotes resolution and F is a set of clauses. Then

$$F \vdash \bot \text{ iff } F \models \bot$$

.

Completeness of FOL

Gödel completeness theorem

$$\text{If } \models F \text{ then } \vdash F$$

(in the original proof $\vdash$ refers to the Hilbert-Ackermann proof system).

More general form

$$\Gamma \models F \text{ IFF } \Gamma \vdash F$$

(in the original proof $\vdash$ refers to the Hilbert-Ackermann proof system).

Undecidability of FOL

Theorem

It is undecidable whether a first order logic formula is provable (or true under all possible interpretations).

FO arithmetic

- First-order *a*rithmetic is a language of terms and formulas. Terms or (positive) polynomials are built from variables X,Y,Z,..., the constants 0 and 1 and the operators + and x of addition and multiplication. The multiplication operator is normally suppressed in writing. The simplest formulas are the equations, obtained by writing an = between two terms, for instance $X + Y^2 = 2Z^3$ which is an abbreviation for $X + YY = (1 + 1)ZZZ$. More complicated formulas can be built from equations by using the usual FOL connectives and quantifiers.
- Arithmetic is interpreted in terms of the natural numbers. Every formula is either true or false (if there are free variables a formula is considered equivalent to its universal closure).

More undecidability results

Theorem

It is undecidable whether an arithmetical formula is true

Theorem

The set of true arithmetical formulas is not even semi-decidable.

Gödel incompleteness theorem

If a proof system for arithmetic is sound (meaning that only true formulas are provable) then there must be a true formula that is not provable.