First Order Logic (FOL) [1]
http://lcs.ios.ac.cn/~znj/DM2017

Naijun Zhan

March 19, 2017

[1] Special thanks to Profs Hanpin Wang (PKU) and Lijun Zhang (ISCAS) for their courtesy of the slides on this course.

Outline

Why FOL

- Propositional logic is a coarse language, which only concerns about propositions and boolean connectives. Practically, this logic is not powerful enough to describe important properties we are interested in.

Example (Syllogism of Aristotle)

Consider the following assertions:

1. All men are mortal.
2. Socrates is a man.
3. So Socrates would die.

$$\forall x(Man(x) \to Mortal(x))$$

Difference between FOL and PL

First order logic is an extension of proposition logic:

1. To accept parameters, it generalized propositions to predicates.
2. To designate elements in the domain, it is equipped with functions and constants.
3. It also involves quantifiers to capture infinite conjunction and disjunction.

Signature

- We are given:
  - an arbitrary set of variable symbols $VS = \{x, y, x_1, \ldots\}$;
  - an arbitrary set (maybe empty) of function symbols $FS = \{f, g, f_1, \ldots\}$, where each symbol has an arity;
  - an arbitrary set (maybe empty) of predicate symbols $PS = \{P, Q, P_1, \ldots\}$, where each symbol has an arity;
  - an equality symbol set ES which is either empty or one element set containing $\{\approx\}$.

- Let $L = VS \cup \{(,), \rightarrow, \neg, \forall\} \cup FS \cup PS \cup ES$. Here $VS \cup \{(,), \rightarrow, \neg, \forall\}$ are referred to as *logical symbols*, and $FS \cup PS \cup ES$ are referred to as *non-logical symbols*.

- We often make use of the
  - set of constant symbols, denoted by $CS = \{a, b, a_1, \ldots\} \subseteq FS$, which consist of function symbols with arity 0;
  - set of propositional symbols, denoted by $PCS = \{p, q, p_1, \ldots\} \subseteq PS$, which consist of predicate symbols with arity 0.

FOL terms

The terms of the first order logic are constructed according to the following grammar:

$$t ::= x \mid ft_1 \ldots t_n$$

where $x \in VS$, and $f \in FS$ has arity n.
Accordingly, the set T of terms is the smallest set satisfying the following conditions:

- each variable $x \in VS$ is a term.
- Compound terms: $ft_1 \ldots t_n$ is a term (thus in T), provided that f is a n-arity function symbol, and $t_1, \ldots, t_n \in \mathbf{T}$. Particularly, $a \in CS$ is a term.

We often write $f(t_1, \ldots, t_n)$ for the compound terms.

FOL formulas

The well-formed formulas of the first order logic are constructed according to the following grammar:

$$\varphi ::= Pt_1 \ldots t_n \mid \neg\varphi \mid \varphi \rightarrow \varphi \mid \forall x\varphi$$

where $t_1, \ldots, t_n$ are terms, $P \in PS$ has arity n, and $x \in VS$.
We often write $P(t_1, \ldots, t_n)$ for clarity. Accordingly, the set *FOL* of first order formulas is the smallest set satisfying:

- $P(t_1, \ldots, t_n) \in FOL$ is a formula, referred to as the atomic formula.
- Compound formulas: $(\neg\varphi)$ (negation), $(\varphi \rightarrow \psi)$ (implication), and $(\forall x\varphi)$ (universal quantification) are formulas (thus in *FOL*), provided that $\varphi, \psi \in FOL$.

We omit parentheses if it is clear from the context.
As syntactic sugar, we can define $\exists x\varphi$ as $\exists x\varphi := \neg\forall x\neg\varphi$. We assume that $\forall$ and $\exists$ have higher precedence than all logical operators.

Examples of first-order logics

Mathematical theories

- **Presburger Arithmetic** $\langle \mathbb{N}, 0, 1, +, =, < \rangle$.
- **Peano Arithmetic** $\langle \mathbb{N}, 0, S, +, \cdot, =, < \rangle$
- **Tarski Algebra** $\langle \mathbb{R}, 0, +, \cdot, =, < \rangle$
- **Group** $\langle e, +, = \rangle$.
- **Equivalence** $\langle R \rangle$.

Example

- Write "every son of my father is my brother" in predicate logic.
- Let *me* denote "me", $S(x, y)$ (x is a son of y), $F(x; y)$ (x is the father of y), and $B(x; y)$ (x is a brother of y) be predicate symbols of arity 2. Consider

$$\forall x \forall y (F(x; me) \wedge S(y; x) \rightarrow B(y; me)).$$

- Alternatively, let f ($f(x)$ is the father of x) be a unary function symbol. Consider

$$\forall x (S(x; f(me)) \rightarrow B(x; me)).$$

- **Translating an English sentence into predicate logic can be tricky.**

Sub-formulas

For a formula φ, we define the sub-formula function $Sf : FOL \to 2^{FOL}$ as follows:

$$\begin{aligned}
Sf(P(t_1, \ldots, t_n)) &= \{P(t_1, \ldots, t_n)\} \\
Sf(\neg\varphi) &= \{\neg\varphi\} \cup Sf(\varphi) \\
Sf(\varphi \to \psi) &= \{\varphi \to \psi\} \cup Sf(\varphi) \cup Sf(\psi) \\
Sf(\forall x\varphi) &= \{\forall x\varphi\} \cup Sf(\varphi) \\
Sf(\exists x\varphi) &= \{\exists x\varphi\} \cup Sf(\varphi)
\end{aligned}$$

Scope

The part of a logical expression to which a quantifier is applied is called the scope of this quantifier. Formally, each sub-formula of the form $Qx\psi \in Sf(\varphi)$, the scope of the corresponding quantifier Qx is ψ. Here $Q \in \{\forall, \exists\}$.

Sentence

We say an occurrence of x in φ is free if it is not in scope of any quantifiers $\forall x$ (or $\exists x$). Otherwise, we say that this occurrence is a bound occurrence. If a variable φ has no free variables, it is called a *closed formula*, or a *sentence*.

Substitution

Substitution

The substitution of x with t within φ, denoted as $S^x_t\varphi$, is obtained from φ by replacing each free occurrence of x with t.

- We would extend this notation to $S^{x_1,\ldots,x_n}_{t_1,\ldots,t_n}\varphi$.

Remark 1

It is important to remark that $S^{x_1,\ldots,x_n}_{t_1,\ldots,t_n}\varphi$ is not the same as $S^{x_1}_{t_1}\ldots S^{x_n}_{t_n}\varphi$: the former performs a simultaneous substitution.
For example, consider the formula $P(x,y)$: the subsitution $S^{x,y}_{y,x}P(x,y)$ gives $S^{x,y}_{y,x}P(x,y) = P(y,x)$ while the substitutions $S^x_y S^y_x P(x,y)$ give $S^x_y S^y_x P(x,y) = S^x_y P(x,x) = P(y,y)$.

Remark 2

Consider $\varphi = \exists y(x < y)$ in the number theory. What is $S^x_t\varphi$ for the special case of $t = y$?

Substitution

Substitutable on Terms

We say that t is substitutable for x within φ iff for each variable y occurring in t, there is no free occurrence of x in scope of $\forall y/\exists y$ in φ.

α-β condition

If the formula φ and the variables x and y fulfill:

1. y has no free occurrence in φ, and
2. y is substitutable for x within φ,

then we say that φ, x and y meet the α-β condition, denoted as $C(\varphi, x, y)$.

Lemma

If $C(\varphi, x, y)$, then $S_x^y S_y^x \varphi = \varphi$.

Outline

Axioms

As for propositional logic, also FOL can be axiomatized.

Axioms

A1 $\varphi \to (\psi \to \varphi)$

A2 $(\varphi \to (\psi \to \eta)) \to ((\varphi \to \psi) \to (\varphi \to \eta))$

A3 $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$

A4 $\forall x \varphi \to S^x_t \varphi$
if t is substitutable for x within φ

A5 $\forall x(\varphi \to \psi) \to (\forall x \varphi \to \forall x \psi)$

A6 $\varphi \to \forall x \varphi$
if x is not free in φ

A7 $\forall x_1 \ldots \forall x_n \varphi$
if φ is an instance of (one of) the above axioms

MP Rule

$$\frac{\varphi \to \psi \quad \varphi}{\psi}$$

Deduction Theorem

Deductive sequence

Given a formula set Γ, a deductive sequence of φ from Γ is a sequence

$$\varphi_0, \varphi_1, \dots, \varphi_n = \varphi$$

where each φ_i should be one of the following cases:

1. $\varphi_i \in \Gamma$.
2. φ_i is an instance of some axiom.
3. There exists some $j, k < i$, such that $\varphi_k = \varphi_j \to \varphi_i$.

And, we denote by $\Gamma \vdash \varphi$ if there exists such deductive sequence. We write $\Gamma, \psi \vdash \varphi$ for $\Gamma \cup \{\psi\} \vdash \varphi$.

Theorem (Deduction theorem)

$\Gamma, \varphi \vdash \psi$ *if and only if* $\Gamma \vdash \varphi \to \psi$.

Generalization Theorem

Syntactical Equivalence

We say φ and ψ are syntactically equivalent iff $\varphi \vdash \psi$ and $\psi \vdash \varphi$.

Theorem

(Gen): If x has no free occurrence in Γ, then $\Gamma \vdash \varphi$ implies $\Gamma \vdash \forall x \varphi$.

Solution

Suppose that $\varphi_0, \varphi_1, \ldots, \varphi_n = \varphi$ is the deductive sequence of φ from Γ.

- *If φ_i is an instance of some axiom, then according to (A7), $\forall x \varphi_i$ is also an axiom.*
- *If $\varphi_i \in \Gamma$, since x is not free in Γ, we have $\vdash \varphi_i \to \forall x \varphi_i$ according to (A6). Therefore, we have $\Gamma \vdash \forall x \varphi_i$ in this case.*
- *If φ_i is obtained by applying (MP) to some φ_j and $\varphi_k = \varphi_j \to \varphi_i$. By induction, we have $\Gamma \vdash \forall x \varphi_j$ and $\Gamma \vdash \forall x(\varphi_j \to \varphi_i)$. With (A5) and (MP), we also have $\Gamma \vdash \forall x \varphi_i$ in this case.*

Thus, we have $\Gamma \vdash \forall x \varphi_n$, i.e., $\Gamma \vdash \forall x \varphi$.

Examples and proof techniques

Eg 1. Prove that

1. $\forall x(\varphi \to \psi) \vdash \forall x(\neg\psi \to \neg\varphi)$,
2. $\forall x(\varphi \to \psi) \vdash \exists x\varphi \to \exists x\psi$.

Eg 2. Prove that

1. $\forall x\forall y\varphi \vdash \forall y\forall x\varphi$,
2. $\exists x\forall y\varphi \vdash \forall y\exists x\varphi$.

Eg 3. Prove that

1. If $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\psi$, then $\Gamma \vdash \neg(\varphi \to \psi)$,
2. $\forall x\neg(\varphi \to \psi) \vdash \neg(\varphi \to \exists x\psi)$.

Proof techniques

- **By contradiction:** In order to prove $\Gamma \vdash \varphi$, we only need to prove $\Gamma, \neg\varphi \vdash F$.
- **By assumption:** Assume $S^x_{x_0}\varphi$, where x_0 is a fresh variable, once we have $\Gamma, S^x_{x_0}\varphi \vdash \psi$, then $\Gamma \vdash \exists x.\varphi \to \psi$.

Lemmas and theorems

Lemma

(Ren): *If* $C(\varphi, x, y)$, *then* $\forall x\varphi$ *and* $\forall y S^x_y \varphi$ *are syntactical equivalent. That is,*

1. $\forall x\varphi \vdash \forall y S^x_y \varphi$.
2. $\forall y S^x_y \varphi \vdash \forall x\varphi$.

Lemma

(RS): *Let* η^{φ}_{ψ} *denote the formula obtained by replacing (some or all)* φ *inside* η *by* ψ. *If* $\varphi \vdash \psi$ *and* $\psi \vdash \varphi$ *then* $\eta \vdash \eta^{\varphi}_{\psi}$ *and* $\eta^{\varphi}_{\psi} \vdash \eta$.

Lemma

If $C(\varphi, x, y)$ *and* $\Gamma \vdash \psi$, *then* $\Gamma \vdash \psi^{\forall x\varphi}_{\forall y S^x_y \varphi}$.

Theorem

(GenC) *If* $\Gamma \vdash S^x_a \varphi$ *where* a *does not occur in* $\Gamma \cup \{\varphi\}$, *then* $\Gamma \vdash \forall x\varphi$.

Outline

Tarski structure

To give semantics of terms/formulas of first order logic, we need an appropriate structure in which interpret the functions and predicates of FOL.

Tarski structure

A Tarski structure is a pair $\mathscr{I} = \langle \mathcal{D}, \mathcal{I} \rangle$, where:

- $\mathcal{D}$ is a non-empty set, called the domain.
- For each n-ary function f, we have $\mathcal{I}(f) \in \mathcal{D}^n \to \mathcal{D}$.
- For each n-ary predicate P, we have $\mathcal{I}(P) \in \mathcal{D}^n \to \{0, 1\}$.

Thus, for each constant a, we have $\mathcal{I}(a) \in \mathcal{D}$.

Assignment

Given a Tarski structure $\mathscr{I} = \langle \mathcal{D}, \mathcal{I} \rangle$, an assignment σ under $\mathscr{I}$ is a mapping $\sigma \colon VS \to \mathcal{D}$.
We use $\Sigma_{\mathscr{I}}$ to denote the set consisting of assignments under $\mathscr{I}$.

Formal semantics–Interpretation of terms

Let $\mathscr{I} = \langle \mathcal{D}, \mathcal{I} \rangle$ and $\sigma \in \Sigma_{\mathscr{I}}$.
Each term t is interpreted to an element $\mathscr{I}(t)(\sigma)$ belonging to $\mathcal{D}$:

- If $t = x$ is a variable, then $\mathscr{I}(t)(\sigma) = \sigma(x)$.
- If $t = f(t_1, \ldots, t_n)$ where f is an n-ary function, then $\mathscr{I}(t)(\sigma) = \mathcal{I}(f)(\mathscr{I}(t_1)(\sigma), \ldots, \mathscr{I}(t_n)(\sigma))$.

Thus, if $t = a$ is a constant, then $\mathscr{I}(t)(\sigma) = \mathcal{I}(a)$.

Formal semantics–Interpretation of formulas

Each formula φ has a truth value $\mathscr{I}(\varphi)(\sigma) \in \{0, 1\}$:

- If $\varphi = P(t_1, \ldots, t_n)$, where P is an n-ary predicate, then $\mathscr{I}(\varphi)(\sigma) = \mathcal{I}(P)(\mathscr{I}(t_1)(\sigma), \ldots, \mathscr{I}(t_n)(\sigma))$.
- If $\varphi = \neg\psi$, then $\mathscr{I}(\varphi)(\sigma) = 1 - \mathscr{I}(\psi)(\sigma)$.
- If $\varphi = \psi \rightarrow \eta$, then

$$\mathscr{I}(\varphi)(\sigma) = \begin{cases} 1 & \text{if } \mathscr{I}(\psi)(\sigma) = 0 \text{ or } \mathscr{I}(\eta)(\sigma) = 1, \\ 0 & \text{if } \mathscr{I}(\psi)(\sigma) = 1 \text{ and } \mathscr{I}(\eta)(\sigma) = 0. \end{cases}$$

- If $\varphi = \forall x \psi$, then

$$\mathscr{I}(\varphi)(\sigma) = \begin{cases} 1 & \text{if } \mathscr{I}(\psi)(\sigma[x/d]) = 1 \text{ for each } d \in \mathcal{D}, \\ 0 & \text{if } \mathscr{I}(\psi)(\sigma[x/d]) = 0 \text{ for some } d \in \mathcal{D} \end{cases}$$

  where $\sigma[x/d]$ is a new assignment defined as

$$\sigma[x/d](y) = \begin{cases} \sigma(y) & \text{if } y \neq x, \\ d & \text{if } y = x. \end{cases}$$

We write $(\mathscr{I}, \sigma) \Vdash \varphi$ if $\mathscr{I}(\varphi)(\sigma) = 1$.

Theorem of Substitution

Theorem of Substitution

Suppose that t is substitutable for x within φ, then

$$(\mathscr{I}, \sigma) \Vdash S_t^x \varphi \text{ if and only if } (\mathscr{I}, \sigma[x/\mathscr{I}(t)(\sigma)]) \Vdash \varphi.$$

We say that $\mathscr{I}$ is a model of φ, denoted as $\mathscr{I} \Vdash \varphi$, if $(\mathscr{I}, \sigma) \Vdash \varphi$ for each $\sigma \in \Sigma_{\mathscr{I}}$.
In particular, we say that $\mathscr{I} = \langle \mathcal{D}, \mathcal{I} \rangle$ is a frugal model of φ if $|\mathcal{D}|$ is not more than the cardinality of the language.
Recall that φ is a sentence, if there is no free variable occurring in φ.

Theorem

If φ is a sentence, then

- $\mathscr{I} \Vdash \varphi$ iff $(\mathscr{I}, \sigma) \Vdash \varphi$ for some $\sigma \in \Sigma_{\mathscr{I}}$.

Satisfiability and validity

Let φ, ψ be FOL formulas and Γ be a set of FOL formulas. Then we define:

- $(\mathscr{I}, \sigma) \Vdash \Gamma$ if for each $\eta \in \Gamma$, $(\mathscr{I}, \sigma) \Vdash \eta$;
- $\Gamma \models \varphi$ if for each $\mathscr{I}$ and $\sigma \in \Sigma_{\mathscr{I}}$, $(\mathscr{I}, \sigma) \Vdash \Gamma$ implies $(\mathscr{I}, \sigma) \Vdash \varphi$;
- φ and ψ are equivalent if $\{\varphi\} \models \psi$ and $\{\psi\} \models \varphi$;
- φ is valid if $\emptyset \models \varphi$.

Tautology for FOL

For a formula $\varphi \in FOL$, we construct φ' as follows:

- for each sub-formula ψ of φ which is either an atomic formula, or a formula of the form $\forall x \eta$, we replace it with a corresponding propositional variable p_{ψ}.

If φ' is a tautology in propositional logic, then we say φ is a tautology for FOL.

Prenex Normal Form (PNF)

A formula is in prenex normal form if and only if it is of the form $Q_1 x_1 Q_2 x_2 \ldots Q_k x_k P(x_1, x_2, \ldots, x_k)$, where each $Q_i, i = 1, 2, \ldots , k$ is either the existential quantifier or the universal quantifier, and $P(x_1, \ldots, x_k)$ is a predicate involving no quantifiers.

Question: can we transform a formula into an equivalent PNF form?

Outline

Outline

Soundness

Similarly to propositional logic, for FOL we have the soundness property:

Theorem

If $\Gamma \vdash \varphi$, *then* $\Gamma \models \varphi$.

Hint.

For proving the theorem, show and make use of the following results:

- $\{\forall x(\varphi \to \psi), \forall x \varphi\} \models \forall x \psi$;
- if x is not free in φ, then $\vdash \varphi \to \forall x \varphi$. □

Corollary

If $\vdash \varphi$, *then* $\models \varphi$.

Outline

A Hintikka set Γ is a set of FOL formulas fulfilling the following properties:

1. For each atomic formula φ (i.e, $\varphi = P(t_1, \dots, t_n)$, where $n \geq 0$), either $\varphi \notin \Gamma$ or $\neg\varphi \notin \Gamma$.
2. $\varphi \to \psi \in \Gamma$ implies that either $\neg\varphi \in \Gamma$ or $\psi \in \Gamma$.
3. $\neg\neg\varphi \in \Gamma$ implies that $\varphi \in \Gamma$.
4. $\neg(\varphi \to \psi) \in \Gamma$ implies that $\varphi \in \Gamma$ and $\neg\psi \in \Gamma$.
5. $\forall x \varphi \in \Gamma$ implies that $S^x_t \varphi \in \Gamma$ for each t which is substitutable for x within φ.
6. $\neg\forall x \varphi \in \Gamma$ implies that there is some t with $C(\varphi, x, t)$ such that $\neg S^x_t \varphi \in \Gamma$.

Note: $C(\varphi, x, t)$ iff $C(\varphi, x, y)$ for all y occurring in t.

Completeness (cont'd)

Lemma

A Hintikka set Γ *is consistent, and moreover, for each formula* φ*, either* $\varphi \notin \Gamma$*, or* $\neg\varphi \notin \Gamma$.

Theorem

A Hintikka set Γ *is satisfiable, i.e, there is some interpretation* $\mathscr{I}$ *and some* $\sigma \in \Sigma_{\mathscr{I}}$ *such that* $(\mathscr{I}, \sigma) \Vdash \varphi$ *for each* $\varphi \in \Gamma$.

Completeness (cont'd)

Theorem

If Γ is a set of FOL formulas, then "Γ is consistent" implies that "Γ is satisfiable". Particularly, if Γ consists only of sentences, then Γ has a frugal model.

Proof.

Let us enumerate[a] the formulas as $\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots$, and subsequently define a series of formula sets as follows. Let $\Gamma_0 = \Gamma$, and

$$\Gamma_{i+1} = \begin{cases} \Gamma_i \cup \{\neg\varphi_i\} & \text{if } \Gamma_i \vdash \neg\varphi_i \\ \Gamma_i \cup \{\varphi_i\} & \text{if } \Gamma_i \nvdash \neg\varphi_i \text{ and } \varphi_i \neq \neg\forall x\psi \\ \Gamma_i \cup \{\varphi_i, \neg S^x_a\psi\} & \text{if } \Gamma_i \nvdash \neg\varphi_i\text{, and } \varphi_i = \neg\forall x\psi \end{cases}$$

Above, for each formula $\forall x\psi$, we pick and fix the constant a which does not occur in $\Gamma_i \cup \{\varphi_i\}$. Finally let $\Gamma^* = \lim_{i\to\infty} \Gamma_i$.

If Γ is consistent, the set Γ^* is maximal and consistent, and is referred to as the Henkin set. Thus, a Henkin set is also a Hintikka set. □

[a] We assume the language to be countable, yet the result can be extended to languages with arbitrary cardinality.

Completeness (cont'd)

Theorem

If $\Gamma \models \varphi$, *then* $\Gamma \vdash \varphi$.

Corollary

If $\models \varphi$, *then* $\vdash \varphi$.

Theorem

Γ *is consistent iff each of its finite subset is consistent. Moreover,* Γ *is satisfiable iff each of its finite subsets is satisfiable.*

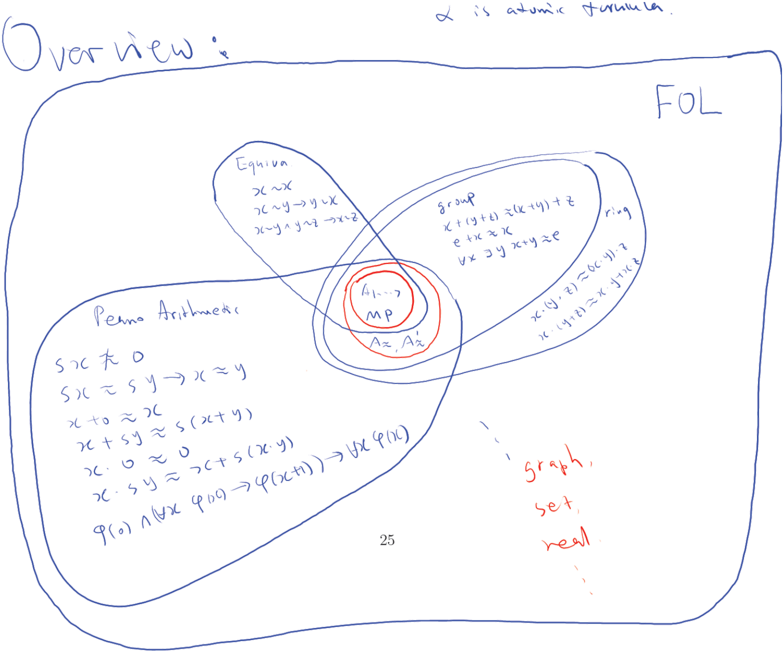
Outline

The axiomatization based on the Hilbert's systems seen in the previous section can be extended to the case of first order logic with the equality $\approx$. To do this, two additional axioms have to be included in the Hilbert's system:

$A_{\approx}$**:** $x \approx x$;

$A'_{\approx}$**:** $(x \approx y) \rightarrow (\alpha \rightarrow \alpha^x_y)$, where α is an atomic formula.

The soundness and completeness results can be proved similarly in the extended Hilbert's system; note that for the completeness one, a variation of the Tarski structure is required, namely, the domain considered in the construction modulo the relation $\approx$. This allows us so manage correctly the formulas that are equivalent under $\approx$.
The actual details about the above construction are omitted; the interested reader is invited to formalize them.


α is atomic formula.
Overview:
FOL
Equiva.
x ∼ x
x ∼ y → y ∼ x
x ∼ y ∧ y ∼ z → x ∼ z
group
x + (y + z) ≈ (x + y) + z
e + x ≈ x
∀x ∃y x + y ≈ e
ring
x · (y · z) ≈ (x · y) · z
x · (y + z) ≈ x · y + x · z
A1, ..., 7
MP
A≈, A'≈
Peano Arithmetic
S x ≉ 0
S x ≈ S y → x ≈ y
x + 0 ≈ x
x + S y ≈ S(x + y)
x · 0 ≈ 0
x · S y ≈ x + S(x · y)
φ(0) ∧ (∀x φ(x) → φ(x+1)) → ∀x φ(x)
graph,
set,
real

25