Note 03

October 9, 2021

1 First-order languages

Definition 1.1. The alphabet of a first-order language consists of the following groups of symbols:

- Parenthesis: (and)
- Connectives: [$\neg$,negation,not], [$\wedge$, conjunction,and], and [$\vee$, disjunction,or]];
- Quantifiers: [$\forall$,for all] and [$\exists$, there exists];
- A denumerably infinite list of variables: $v_0, v_1, ..., v_n, ...$;
- $=$;
- A set of constant symbols $\mathcal{C}$;
- A set of function symbols $\mathcal{F}$, and positive integers n_f for each $f \in \mathcal{F}$, which is referred to as the arity of the function;
- A set of relation symbols $\mathcal{R}$, and positive integers n_R for each $R \in \mathcal{R}$, which is referred to as the arity of the relation.

In this chapter, we only consider the language has a unique m-ary relation r.

Definition 1.2. We define the sets $F_0, ..., F_n$ by induction on n as follows:

- F_0, called the set of atomic formulas or formulas of complexity 0, consists of all the words of the form $x_1 = x_2$ and $r(x_1, ..., x_m)$, where $x_1, ..., x_m$ are variables, not necessarily distinct.
- F_{n+1}, called the set of formulas of complexity $n+1$, consists of all words of the form

  $$\neg(f),\ (f) \wedge (g),\ (f) \vee (g),\ (\exists x)(f),\ \text{or}\ (\forall x)(f),$$

  where x is a variable and $f, g \in F_0 \cup ... \cup F_n$.

The union of all the F_n is called the set F of formulas.

Definition 1.3. Let f be a formula, then the set $S(f)$ of subformulas of f is defined by induction on the complexity of f:

- If f is atomic, then $S(f) = \{f\}$;
- If f is $\neg(g)$, or $(\exists x)(g)$, or $(\forall x)(g)$, then $S(f) = S(g) \cup \{f\}$;
- If f is $(g) \wedge (h)$, or $(g) \vee (h)$, then $S(f) = S(g) \cup S(h) \cup \{f\}$.

Definition 1.4. Let f be a formula, then the quantifier rank of f, denoted by $QR(f)$, is defined by induction on the complexity of f:

- If f is atomic, then $QR(f) = 0$;
- If f is $\neg(g)$, then $QR(f) = QR(g)$;
- If f is $(g) \wedge (h)$, or $(g) \vee (h)$, then $QR(f) = \max\{QF(g),\ QF(h)\}$;
- If f is $(\exists x)(g)$, or $(\forall x)(g)$, then $QR(f) = QR(g) + 1$.

The formulas of quantifier rank 0 are call quantifier-free formulas, which are exactly the Boolean combinations of the atomic formulas.

Definition 1.5. Let f be a formula, then we define the set $FV(f)$ of the free variables of f, as follows:

- If f is atomic, then $FV(f) =$ all variables occurring in f;
- If f is $\neg(g)$, then $FV(f) = FV(g)$;
- If f is $(g) \wedge (h)$, or $(g) \vee (h)$, then $FV(f) = FV(g) \cup QF(h)$;
- If f is $(\exists x)(g)$, or $(\forall x)(g)$, then $FV(f) = FV(g) \backslash \{x\}$.

If $FV(f) = \emptyset$, then we call f a closed formula or sentence.

Definition 1.6.
- When we write a formula $f(\bar{x})$, where $\bar{x}$ is an n-tuple of variables $(x_1, ..., x_n)$, we understand that all free variables of f are contained among $x_1, ..., x_n$. Namely,
$$\{x_1, ..., x_n\} \subseteq FV(f)$$
- Let (M, R) be an m-ary relation, and $\bar{a} = (a_1, ..., a_n) \in M^n$;
- We will define, by induction on the complexity of f, what it means for R to satisfy $f(\bar{a})$, or equivalently for $f(\bar{a})$ to be true for R. We write
$$(M, R) \models f(\bar{a})$$
to mean (M, R) satisfies $f(\bar{a})$, where $f(a)$ is not a formula in our language, but rather what we get from the formula $f(\bar{x})$ by replacing free occurrences of $x_1, ..., x_n$ by $a_1, ..., a_n$, respectively.

- If f is of the form $x = y$, then $(M, R) \models a = b$ iff a and b are identical;
- If f is of the form $r(x_1, ..., x_n)$, then $(M, R) \models r(a_1, ..., a_n)$ iff $(a_1, ..., a_n) \in R$;
- $(M, R) \models \neg(f)(\bar{a})$ iff (M, R) does not satisfy $f(\bar{a})$;
- $(M, R) \models (f) \vee (g)(\bar{a})$ iff (M, R) satisfies $f(\bar{a})$ or (M, R) satisfies $g(\bar{a})$;
- $(M, R) \models (f) \wedge (g)(\bar{a})$ iff (M, R) satisfies $f(\bar{a})$ and (M, R) satisfies $g(\bar{a})$;
- $(M, R) \models (\exists x)(f)(\bar{a}, x)$ iff there exists $b \in M$ such that R satisfies $f(\bar{a}, b)$;
- $(M, R) \models (\forall x)(f)(\bar{a}, x)$ iff for all $b \in M$, R satisfies $f(\bar{a}, b)$.

We will assume that the universe of every relation is not empty.

Definition 1.7. Let $f(\bar{x})$ and $g(\bar{x})$, we say that f and g are equivalent if for any n-tuple $\bar{a}$ and any relation (M, R),

$$(M, R) \models f(\bar{a}) \iff (M, R) \models g(\bar{x}).$$

$(\forall x)(f)$ is equivalent to $\neg(\exists x)\neg(f)$

Definition 1.8. A formula is said to be in *prenex form* if all its quantifiers occur at the beginning.

Lemma 1.9. *Every formula has an equivalent prenex form.*

2 Connections to Back-and-Forth Technique

Theorem 2.1. *(Frässé's Theorem) Let (M, R) and (N, S) be m-ary relations, let $\bar{a} \in M^n$ and $\bar{b} \in N^n$. Then $\bar{a}$ and $\bar{b}$ are p-equivalent iff*

$$(M, R) \models f(\bar{a}) \iff (N, R) \models f(\bar{b})$$

for any formula $f(\bar{x})$ with quantifier rank at most p.

Proof. $\Rightarrow$: By induction on p.

- If $\bar{a} \sim_0 \bar{b}$;
- Then, by definition, they satisfy the same atomic formulas;
- Therefore, they satisfy the same quantifier-free formulas.
- Suppose that $\bar{a} \sim_{p+1} \bar{b}$;
- The formula $f := (\exists y) g(\bar{x}, y)$ has quantifier rank at most $p + 1$;

- So $g(\bar{x}, y)$ is a formula of quantifier rank at most p;
- $(M, R) \models f(\bar{a})$ iff there is $c \in M$ such that $(M, R) \models g(\bar{a}, c)$;
- there is $d \in N$ such that $\bar{a}c \sim_p \bar{b}d$;
- by induction hypothesis, $(N, S) \models g(\bar{b}, d)$, and thus $\models (\exists y)g(\bar{b}, y)$;
- Similarly, $(N, S) \models (\exists y)g(\bar{b}, y) \implies (M, R) \models (\exists y)g(\bar{a}, y)$.

To prove the converse, we need the following lemma: $\square$

Lemma 2.2. *If the arity m of a relation, and the integers n and p, are fixed, there is only finite number $C(n, p)$ of p-equivalence classes of n-typles.*

Proof. Induction on p. If $p = 0$, then

- Consider a set of symbols $X = \{x_1, ..., x_n\}$;
- There are at most finitely many m-ary relations defined on X;
- Also, there are at most finitely many ways to interpret the relation "=" on X;
- Let (M, R) and (N, S) be m-ary relations, $\bar{a} \in M^n$ and $\bar{b} \in N^n$;
- Let $A = \{a_1, ..., a_n\}$ and $B = \{b_1, ..., b_n\}$;
- Let $R_A = R \cap A^m$, the restriction of R on A;
- Let $S_B = S \cap B^m$, the restriction of S on B;
- If $p = 0$, $\bar{a} \sim_0 \bar{b}$ iff R_A is isomorphic to R_B via $a_i \mapsto b_i$, $i = 1, ..., n$;
- So there are at most finitely many 0-equivalence classes of n-tuples;

From p to $p + 1$:

- by induction hypothesis,
  - there exists relations $\{(M_k, R_k)|\ k \leq C(n + 1, p)\}$, and
  - $\{\bar{d}_k \in M_k^{n+1}|\ k \leq C(n + 1, p)\}$
- such that each $n + 1$-tuple is p-equivalent to some $\bar{d}_k$;
- Now consider an arbitrary relation (M, R) and an n-tuple $\bar{a}$;
- We define $[\bar{a}] = \{k|\ \exists c \in M(\bar{a}c \sim_p \bar{d}_k)\}$
- For any relation (N, S) and $\bar{b} \in N^n$;

- It is easy to see that $\bar{a} \sim_{p+1} \bar{b} \iff [\bar{a}] = [\bar{b}]$;
- So $C(n, p+1)$ is bounded by $2^{C(n+1,p)}$.

$\square$

Proof of Theorem 2.1, Part 2: We now show that if $\bar{a}$ and $\bar{b}$ satisfy the same formulas of QR at most p, then $\bar{a} \sim_p \bar{b}$.

We Claim that for each p-equivalence class C, there is a formula f_C of QR p such that the tuples in C are exactly those satisfy f_C.

- Induction on p.
- If $p = 0$:
- Given an n-tuple $\bar{a}$;
- There are only finitely many atomic formulas with variables $x_1, ..., x_n$;
- Let f_C be the conjunction of those satisfied by $\bar{a}$ and negation of those not satisfied by $\bar{a}$.
- Then f_C characterizes the 0-equivalence class of $\bar{a}$.

From p to $p + 1$:

- Let $\bar{a}$ be an n-tuple of (M, R);
- Let $f_1(\bar{x}, y), ..., f_k(\bar{x}, y)$ characterize the p-equivalence classes C_1,...,C_k, on $n+1$-tuples, respectively;
- Let $\langle\bar{a}\rangle = \{i \leq k | \ (M, R) \models (\exists y) f_i(\bar{a}, y)\}$;
- it is easy to see that $\langle\bar{a}\rangle = [a]$ if we list $C_1, ..., C_k$ as in the previous lemma.
- Let $f_C(\bar{x}) = \bigwedge_{i \in \langle\bar{a}\rangle} (\exists y) f_i(\bar{x}, y) \wedge \bigwedge_{i \notin \langle\bar{a}\rangle} \neg(\exists y) f_i(\bar{x}, y)$;
- It is easy to see that $\bar{b} \sim_{p+1} \bar{a}$ iff $[\bar{a}] = [\bar{b}]$ iff $\langle\bar{a}\rangle = \langle\bar{b}\rangle$ iff $f_C(\bar{b})$ holds.

$\square$

Remark 2.3. Recall that two formula $f(\bar{x})$ and $g(\bar{x})$ are equivalent if for all $\bar{a}$ and R,

$$R \models f(\bar{a}) \iff R \models g(\bar{a}).$$

- Let $f(\bar{x})$ have quantifier rank p and free variables $x_1, ..., x_n$;
- Suppose that $\bar{a} \in C$ and $R \models f(\bar{a})$;

- Then by Frässé's Theorem, $R' \models f(\bar{a}')$ for each $\bar{a}' \in C$;
- So f is equivalent to a disjunction of some of the f_C;
- There is finitely many f_C's.
- Up to equivalence, there are finitely many formulas of quantifier rank p and free variables $x_1, ..., x_n$;

3 Models and Theories

Definition 3.1. Let R be a relation, f a sentence, and A a set of sentence.

- We say that R is a model of f if $R \models f$;
- R is a model of A if it is a model of every member of A;
- We say that A is consistent if it has a model, otherwise, A is inconsistent/contradictory;
- A sentence f is a consequence of A if it satisfied by every model of A, written $A \models f$;
- The sentences that are consequences of $\emptyset$ are called theses/ theorems;
- An antithesis is the negation of a thesis;
- By a theory we mean a consistent set of sentences containing all its own consequences.
- If A is consistent, the set T_A of consequences of A is a theory, theory generated by A;
- We also call A a set of axioms for T_A or an axiomatization of T_A;
- A maximal theory is called a complete theory;

Remark 3.2.
- If A is inconsistent, then every sentence is a consequence of A;
- A consistent set A can not have both f and $\neg f$ as consequences.
- $g \models f \iff \emptyset \models g \to f$;
- For any relation (M, R), the set T_M of sentences satisfied by M is a complete theory;
- A theory T is complete iff any two of its models are elementarily equivalent;

Example 3.3. Consider the following axioms:

$$A_p : \ (\exists x_1)...(\exists x_p) \bigwedge_{1 \le i < j \le p} (x_i \neq x_j).$$

- Let R be a relation defined on p elements $a_1, ..., a_p$;

- Let $f(\bar{x})$ be the conjunction of atomic formulas and negations of atomic formulas satisfied by the tuple $(a_1, ..., a_p)$;
- The theory of R is axiomatized by the following single axiom:
$$((\exists x_1)...(\exists x_p)f(x_1, ..., x_p)) \wedge \neg A_{p+1}.$$
- The theory of R characterizes R up to isomorphism.

Example 3.4. To axiomatize the theory of the empty unary relation on infinitely many elements, we use
$$\{(\forall x)\neg R(x)\} \cup \{A_1, A_2, ..., A_n, ...\}.$$

Example 3.5. To express that R is an equivalence relation:

- $(\forall x)R(x, x)$;
- $(\forall x)(\forall y)(R(x, y) \leftrightarrow R(y, x))$;
- $(\forall x)(\forall y)(\forall z)(R(x, y) \wedge R(y, z) \rightarrow R(x, z))$.

Example 3.6. To express that an equivalence relation R has infinitely many classes, we need the following infinite list of axioms, one for each $n \in \mathbb{N}$:
$$(\exists x_1)...(\exists x_n) \bigwedge_{1 \le i < j \le n} \neg R(x_i, x_j),$$
and likewise an infinite list of axioms to express that each class is infinite
$$(\forall x)(\forall y_1)...(\forall y_n)(\exists z)(R(x, z) \wedge \bigwedge_{1 \le i \le n} y_i \neq z).$$

Example 3.7. To express the relation is a linear order we need the following axioms:

- $(\forall x)x \le x$;
- $(\forall x)(\forall y)(x \le y \wedge y \le x \rightarrow x = y)$;
- $(\forall x)(\forall y)(\forall z)(x \le y \wedge y \le z \rightarrow x \le z)$;
- $(\forall x)(\forall y)(x \le y \vee y \le x)$.

Let $x < y$ be an abbreviation for $x \le y \wedge x \neq y$. To express that there is no greatest element and no least element:

- $(\forall x)(\exists y)(x < y)$;
- $(\forall x)(\exists y)(y < x)$.

To express that it is nonempty and dense:

- $(\exists x)(x = x)$;
- $(\forall x)(\forall y)(\exists z)(x < y \rightarrow (x < z \wedge z < y))$.

This finite list of axioms axiomatize the complete theory of DLO without endpoints.

4 Elementary Extensions:Tarski's Test, Löwenhenheim's Theorem

Definition 4.1. Let (N, S) be an extension of the relation (M, R). We say that S is an elementary extension of R and write $R \prec S$ if for any tuple $(a_1, ..., a_n) \in M^n$ and any formula $f(x_1, ..., x_n)$

$$R \models f(a_1, ..., a_n) \iff S \models f(a_1, ..., a_n)$$

Theorem 4.2. *(Tarski's Test). If (N, S) is an extension of the relation (M, R), this extension is elementary iff for every $\bar{a} \in M^n$ and every formula $f(\bar{x}, y)$, if $S \models (\exists y) f(\bar{a}, y)$, then there is $b \in M$ such that $S \models f(\bar{a}, b)$.*

Proof. $\Rightarrow$:

- If $R \prec S$, then $S \models (\exists y) f(\bar{a}, y) \implies R \models (\exists y) f(\bar{a}, y)$
- $R \models (\exists y) f(\bar{a}, y) \implies$ there exists $b \in M$ such that $R \models f(\bar{a}, b) \implies S \models f(\bar{a}, b)$.

$\Leftarrow$: Induction on the QR p of $f(\bar{x})$ that $R \models f(\bar{a})$ iff $S \models f(\bar{a})$ for each $\bar{a}$ from M.

- If $p = 0$, f is a Boolean combination of the atomic formulas. It is trivial by the definition of extension.
- From p to $p + 1$;
- Suppose that f has the form $(\exists y) g(\bar{x}, y)$ and g has QR p;
- If $R \models f(\bar{a})$, then there is $b \in M$ such that $R \models g(\bar{a}, b)$.
- By induction hypothesis, $S \models g(\bar{a}, b)$, thus, $S \models (\exists y) g(\bar{a}, y)$;
- If $S \models f(\bar{a})$, then, by the hypothesis, there is $b \in M$ such that $S \models g(\bar{a}, b)$.
- By induction hypothesis, $R \models g(\bar{a}, b)$, thus, $R \models (\exists y) g(\bar{a}, y)$.

$\square$

Theorem 4.3. *(Löwenheim Theorem). Every relation R has a finite or denumerable elementary restriction; more precisely, if A is a infinite subset of the domain of R, then we can find an elementary restriction of R whose domain contains A and has the same cardinarlity as A.*

Proof.

- We enumerate all formulas $f(\bar{a}, y)$ with parameters $\bar{a}$ in A;
- There are denumerably many formulas $f(\bar{x}, y)$;
- There are $Card(A)$ n-tuples from A;
- There are $Card(A)$ formulas with parameters from A;

- For each of these formulas for which $R \models (\exists y) f(\bar{a}, y)$, add to A an element $b_{f,\bar{a}}$ such that $R \models f(\bar{a}, b_{f,\bar{a}})$;
- There are at most $Card(A)$ elements to add;
- Let $A_1 = A \cup \{b_{f,\bar{a}} | \ R \models f(\bar{a}, b_{f,\bar{a}})\}$, then $Card(A) = Card(A_1)$;
- Repeat the operation that replace $A_0 = A$ By A_1, get A_2,...
- Let $B = \bigcup_{n\in} A_n$, then B satisfies the hypothesis of Tarski's test.

$\square$