A. Miller
M776
January 1992

First Order Logic

Let $\emptyset$ denote the empty set ($\{\}$). The finite ordinals are written

$$0 = \emptyset, 1 = \{\emptyset\}, 2 = \{\emptyset, \{\emptyset\}\}, \ldots, n = \{0, 1, \ldots, n-1\}, \cdots$$

The first infinite ordinal is written $\omega = \{0, 1, 2, \ldots\}$. For ordinals α and β we write $\alpha < \beta$ for $\alpha \in \beta$.

We begin with the syntax of first order logic. The logical symbols are $\vee, \neg, \exists, =$ and for each $n \in \omega$ a variable symbol x_n. There are also grammatical symbols such as paranthesis and commas that we use to parse things correctly but have no meaning. For clarity we usually use x, y, z, u, v, etc. to refer to arbitrary variables. The nonlogical symbols consist of a given set L that may include operation symbols, predicate symbols, and constant symbols. The case where L is empty is referred to as the language of pure equality. Each symbol $s \in L$ has a nonnegative integer $\#(s)$ called its arity assigned to it. If $\#(s) = 0$, then s is a constant symbol. If f is an operation symbol and $\#(f) = n$ then f is an n-ary operation symbol. Similarly if R is a predicate symbol and $\#(R) = n$ then R is an n-ary predicate symbol. In addition we always have that "=" is a logical binary predicate symbol.

For the theory of groups the appropriate language is $L = \{e, \cdot, ^{-1}\}$ where "e" is a constant symbol, so $\#(e) = 0$, "$\cdot$" is a binary operation symbol, so $\#(\cdot) = 2$, and "$^{-1}$" is a unary operation symbol, so $\#(^{-1}) = 1)$. For the theory of partially order sets we have that $L = \{\leq\}$ where $\leq$ is a binary relation symbol, so $\#(\leq) = 2$.

Our next goal is to define what it means to be a formula of first order logic. Let L be a fixed language. An expression is a finite string of symbols that are either logical symbols or symbols from L.

The set of terms of L is the smallest set of expressions that contain the variables, constant symbols of L (if any), and is closed under the formation rule: if $t_1, t_2, \ldots, t_n$ are terms of L and f is an n-ary operation symbol of L, then $t = f(t_1, t_2, \ldots, t_n)$ is a term of L. If L has no function symbols then the only terms of L are the variables and constant symbols. So for example if c is a constant symbol, f is a 3-ary operation symbol, g is a binary operation symbol, and h is a unary operation symbol, then $h(f(g(x, h(y)), y, h(c)))$ is a term.

The set of atomic formulas of L is the set of all expressions of the form $R(t_1, t_2, \ldots, t_n)$, where $t_1, t_2, \ldots, t_n$ are terms of L and R is a n-ary predicate symbol of L. Since we always have equality as a binary relation we always have atomic formulas of the form $t_1 = t_2$.

The set of formulas of L is the smallest set of expressions that includes the atomic formulas and is closed under the formation rule: if θ and ψ are L formulas and x is any variable, then $(\theta \vee \psi)$, $\neg\theta$, and $\exists x\, \theta$ are L formulas. We think of other logical connectives as being abbreviations, e.g. $(\theta \rightarrow \psi)$ abbreviates $(\neg\theta \vee \psi)$, and $\forall x\, \theta$ abbreviates $\neg\exists x \neg\theta$. We often add and sometimes drop parentheses to improve readability. Also we write $x \neq y$ for the formally correct but harder to read $\neg x = y$.

It is common practice to write symbols not only in prefix form as above but also in postfix and infix forms. For example in our example of group theory instead of writing the term $\cdot(x, y)$ we usually write it in infix form $x \cdot y$, and ${}^{-1}(x)$ is usually written in postfix form x^{-1}. Similarly in language of partially ordered sets we usually write $x \leq y$ instead of the prefix form $\leq (x, y)$. Binary relations such as partial orders and equivalence relations are most often written in infix form. We regard the more natural forms we write as abbreviations of the more formally correct prefix notation.

Next we want to describe the syntactical concept of substitution. To do so we must first describe what it means for an occurrence of a variable x in a formula θ to be free. If an occurrence of a variable x in a formula θ is not free it is said to be bound. Example:

$$(\exists x\ x = y \vee x = f(y))$$

Both occurrences of y are free, the first two occur of x are bound, and the last occurrence of x is free. In the formula:

$$\exists x\ (x = y \vee x = f(y))$$

all three occurrence of x are bound.

Formally we proceed as follows. All occurrences of variables in an atomic formula are free. The free occurrences of x in $\neg\theta$ are exactly the free occurrences of x in θ. The free occurrences of x in $(\theta \vee \psi)$ are exactly the free occurrences of x in θ and in ψ. If x and y are distinct variables, then the free occurrences of x in $\exists y\ \theta$ are exactly the free occurrences of x in θ. And finally no occurrence of x in $\exists x\ \theta$ is free. This gives the inductive definition of free and bound variables.

We show that x might occur freely in θ by writing $\theta(x)$. If c is a constant symbol the formula $\theta(c)$ is gotten by substituting c for all free occurrences (if any) of x in θ. For example: if $\theta(x)$ is $\exists y\ (y = x \wedge \forall x\ x = y)$, then $\theta(c)$ is $\exists y\ (y = c \wedge \forall x\ x = y)$.

We usually write $\theta(x_1, x_2, \ldots, x_n)$ to indicate that the free variables of θ are amongst the $x_1, x_2, \ldots, x_n$. A formula is called a sentence if no variable occurs freely in it.

Our next goal is to describe the semantics of first order logic. A structure $\mathfrak{A}$ for the language L is a pair consisting of a set A called the universe of $\mathfrak{A}$ and an assignment or interpretation function from the nonlogical symbols of L to individuals, relations , and functions on A. Thus for each constant symbol c in L we have an assignment $c^{\mathfrak{A}} \in A$, for each n-ary operation symbol f in L we have a function $f^{\mathfrak{A}} : A^n \to A$, and for each n-ary predicate symbol R we have a relation $R^{\mathfrak{A}} \subseteq A^n$. The symbol $=$ is always interpreted as the binary relation of equality, which is why we consider it a logical symbol, ie. for any structure $\mathfrak{A}$ we have $=^{\mathfrak{A}}$ is $\{(x, x) : x \in A\}$. We use the word structure and model interchangeably.

For example, suppose L is the language of group theory. One structure for this theory is $\mathfrak{Q} = (\mathbb{Q}, +, -x, 0)$ where the universe is the rationals, $\cdot^{\mathfrak{Q}}$ is ordinary addition of rationals, ${}^{-1^{\mathfrak{Q}}}$ is the function taking each rational r to $-r$, and $e^{\mathfrak{Q}} = 0$. Another structure in this language is $\mathfrak{R} = (\mathbb{R}^+, \times, \frac{1}{x}, 1)$ where the universe is the set of positive real numbers, $\cdot^{\mathfrak{R}}$ is multiplication, ${}^{-1^{\mathfrak{R}}}$ is the function x goes to $\frac{1}{x}$, and $e^{\mathfrak{R}} = 1$. Another example is the group S_n of permutations. Here $\cdot^{S_n}$ is composition of functions, ${}^{-1^{S_n}}$ is the functional which takes each permutation to its inverse, and e^{S_n} is the identity permutation. Of course there are many examples of structures in this language which are not groups.

For another example, the language of partially ordered sets is $L = \{\leq\}$ where $\leq$ is a binary relation symbol. Then the following are all L-structures which happen to be partial orders: $(\mathbb{R}, \{(x, y) \in \mathbb{R}^2 : x \leq y\})$, $(\mathbb{Q}, \{(x, y) \in \mathbb{Q}^2 : x \geq y\})$, and $(\mathbb{N}, \{(x, y) \in \mathbb{N}^2 : x \text{ divides } y\})$. For any nonempty set A and $R \subseteq A^2$, (A, R) is an L-structure. If in addition the relation R is transitive, reflexive, and antisymmetric, then (A, R) is a partial order.

Next we define what it means for an L structure $\mathfrak{A}$ to model or satisfy an L sentence θ, this is written $\mathfrak{A} \models \theta$. For example, $(\mathbb{Q}, +, 0) \models \forall x \exists y\, x \cdot y = e$, because for all $p \in \mathbb{Q}$ there exists $q \in \mathbb{Q}$ such that $p + q = 0$.

Usually it is not the case that every element of a model has a constant symbol which names it. Let $L_{\mathfrak{A}} = L \cup \{c_a : a \in A\}$ where each c_a is a new constant symbol. Let $(\mathfrak{A}, a)_{a\in A}$ be the $L_{\mathfrak{A}}$ structure gotten by augmenting the structure $\mathfrak{A}$ by interpreting each symbol c_a as the element a. We begin by defining $(\mathfrak{A}, a)_{a\in A} \models \theta$ where θ is any $L_{\mathfrak{A}}$ sentence.

The interpretation function can be extended to the variable free terms of $L_{\mathfrak{A}}$ by the rule:

$$(f(t_1, t_2, \ldots, t_n))^{(\mathfrak{A},a)_{a\in A}} = f^{(\mathfrak{A},a)_{a\in A}}(t_1^{(\mathfrak{A},a)_{a\in A}}, t_2^{(\mathfrak{A},a)_{a\in A}}, \ldots, t_n^{(\mathfrak{A},a)_{a\in A}})$$

Hence for each variable free term t we get an interpretation $t^{(\mathfrak{A},a)_{a\in A}} \in A$. For example if $L = \{S, c\}$ where S is a unary operation symbol and c is a constant symbol, and $\mathfrak{Z}$ is the L-structure with universe $\mathbb{Z}$ and where $S^{\mathfrak{Z}}(x) = x + 1$ and $c^{\mathfrak{Z}} = 0$, then $S(S(S(S(c))))^{\mathfrak{Z}} = 4$.

Our definition of $\models$ is by induction on the logical complexity of the sentence θ, i.e. the number of logical symbols in θ.

1. $(\mathfrak{A}, a)_{a\in A} \models R(t_1, \ldots, t_n)$ iff $(t_1^{(\mathfrak{A},a)_{a\in A}}, \ldots, t_n^{(\mathfrak{A},a)_{a\in A}}) \in R^{(\mathfrak{A},a)_{a\in A}}$.
2. $(\mathfrak{A}, a)_{a\in A} \models \neg\theta$ iff not $(\mathfrak{A}, a)_{a\in A} \models \theta$.
3. $(\mathfrak{A}, a)_{a\in A} \models (\theta \vee \psi)$ iff $(\mathfrak{A}, a)_{a\in A} \models \theta$ or $(\mathfrak{A}, a)_{a\in A} \models \psi$.
4. $(\mathfrak{A}, a)_{a\in A} \models \exists x\, \theta(x)$ iff there exist an b in the universe A such that $(\mathfrak{A}, a)_{a\in A} \models \theta(c_b)$.

If θ is an L-sentence and $\mathfrak{A}$ is and L-structure, then we define $\mathfrak{A} \models \theta$ iff θ is true in the augmented structure, ie. $(\mathfrak{A}, a)_{a\in A} \models \theta$.

If $L_1 \subseteq L_2$ and $\mathfrak{A}$ is a L_2 structure, then the reduct of $\mathfrak{A}$ to L_1, written $\mathfrak{A} \upharpoonright L_1$, is the L_1 structure with the same universe as $\mathfrak{A}$ and same relations, operations, and constants as $\mathfrak{A}$ for the symbols of L_1.

Lemma 1 *Let $L_1 \subseteq L_2$ and $\mathfrak{A}$ be an L_2 structure. Then for any θ an L_1 sentence,*

$$\mathfrak{A} \models \theta \textit{ iff } \mathfrak{A} \upharpoonright L_1 \models \theta$$

proof:

Prove by induction on the number of logical symbols in the sentence that for any $L_{1\mathfrak{A}}$ sentence θ:

$$(\mathfrak{A}, a)_{a\in A} \models \theta \text{ iff } (\mathfrak{A}, a)_{a\in A} \restriction L_{1\mathfrak{A}} \models \theta$$

Let $\mathfrak{A}_2 = (\mathfrak{A}, a)_{a\in A}$ and $\mathfrak{A}_1 = (\mathfrak{A}, a)_{a\in A} \restriction L_{1\mathfrak{A}}$.

atomic sentences: By induction on the size of the term, for any $L_{1\mathfrak{A}}$ variable free term t we have that $t^{\mathfrak{A}_1} = t^{\mathfrak{A}_2}$. For any n-ary relation symbol R in L_1 we have $R^{\mathfrak{A}_1} = R^{\mathfrak{A}_2}$ (since $\mathfrak{A}_1$ is a reduct of $\mathfrak{A}_2$). Hence for any atomic $L_{1\mathfrak{A}}$-sentence $R(t_1, \ldots, t_n)$,

$$\mathfrak{A}_1 \models R(t_1, \ldots, t_n) \text{ iff } \langle t_1^{\mathfrak{A}_1}, \ldots t_n^{\mathfrak{A}_1}\rangle \in R^{\mathfrak{A}_1} \text{ iff } \langle t_1^{\mathfrak{A}_2}, \ldots t_n^{\mathfrak{A}_2}\rangle \in R^{\mathfrak{A}_2}$$

$$\text{iff } \mathfrak{A}_2 \models R(t_1, \ldots, t_n)$$

negation: $\mathfrak{A}_1 \models \neg\theta$ iff not $\mathfrak{A}_1 \models \theta$ iff (by induction) not $\mathfrak{A}_2 \models \theta$ iff $\mathfrak{A}_2 \models \neg\theta$.

disjunction: $\mathfrak{A}_1 \models (\theta \vee \rho)$ iff $\mathfrak{A}_1 \models \theta$ or $\mathfrak{A}_1 \models \rho$) iff (by induction) $\mathfrak{A}_2 \models \theta$ or $\mathfrak{A}_2 \models \rho$) iff $\mathfrak{A}_2 \models (\theta \vee \rho)$.

existential quantifier: $\mathfrak{A}_1 \models \exists x\theta(x)$ iff there exists $a \in A$ such that $\mathfrak{A}_1 \models \theta(c_a)$ iff there exists $a \in A$ such that $\mathfrak{A}_2 \models \theta(c_a)$ iff $\mathfrak{A}_2 \models \exists x\theta(x)$.

■

Compactness Theorem

The compactness theorem (for countable languages) was proved by Kurt Gödel in 1930. Malcev extended it to uncountable languages in 1936. The proof we give here was found by Henkin in 1949.

We say that a set of L sentences Σ is finitely satisfiable iff every finite subset of Σ has a model. Σ is complete iff for every L sentence θ either θ is in Σ or $\neg\theta$ is in Σ.

Lemma 2 *For every finitely satisfiable set of L sentences Σ there is a complete finitely satisfiable set of L sentences $\Sigma' \supseteq \Sigma$.*

proof:

Let $B = \{Q : Q \supseteq \Sigma \text{ is finitely satisfiable}\}$. B is closed under unions of chains, because if $C \subseteq B$ is a chain, and $F \subseteq \cup C$ is finite then there exists $Q \in C$ with $F \subseteq Q$, hence F has a model. By the maximum principal, there

exist $\Sigma' \in B$ maximal. But for every L sentence θ either $\Sigma' \cup \{\theta\}$ is finitely satisfiable or $\Sigma' \cup \{\neg\theta\}$ is finitely satisfiable. Otherwise there exists finite $F_0, F_1 \subseteq \Sigma'$ such that $F_0 \cup \{\theta\}$ has no model and $F_1 \cup \{\neg\theta\}$ has no model. But $F_0 \cup F_1$ has a model $\mathfrak{A}$ since Σ is finitely satisfiable. Either $\mathfrak{A} \models \theta$ or $\mathfrak{A} \models \neg\theta$. This is a contradiction.
■

Lemma 3 *If Σ is a finitely satisfiable set of L sentences, and $\theta(x)$ is an L formula with one free variable x, and c a new constant symbol (not in L), then $\Sigma \cup \{(\exists x\, \theta(x)) \rightarrow \theta(c)\}$ is finitely satisfiable*

proof:

This new sentence is called a Henkin sentence and c is called the Henkin constant. Suppose it is not finitely satisfiable, then there exists $F \subseteq \Sigma$ finite such that $F \cup \{(\exists x\, \theta(x)) \rightarrow \theta(c)\}$ has no model. Let $\mathfrak{A}$ be an L-structure modeling F. Since the constant c is not in the language L we are free to interpret it any way we like. If $\mathfrak{A} \models \exists x\, \theta(x)$ choose $c \in A$ so that $(\mathfrak{A}, c) \models \theta(c)$, otherwise choose $c \in A$ arbitrarily. In either case $(\mathfrak{A}, c)$ models F and the Henkin sentence.
■

We say that a set of L sentences Σ is Henkin iff for every L formula $\theta(x)$ with one free variable x, there is a constant symbol c in L such that $(\exists x\, \theta(x)) \rightarrow \theta(c) \in \Sigma$.

Lemma 4 *If Σ is a finitely satisfiable set of L sentences, then there exists $\Sigma' \supseteq \Sigma$ with $L' \supseteq L$ and Σ' a finitely satisfiable Henkin set of L' sentences.*

proof:

For any set of Σ of L sentence, let

$$\Sigma^* = \Sigma \cup \{(\exists x\, \theta(x)) \rightarrow \theta(c_\theta) : \theta(x) \text{ an } L \text{ formula with one free variable}\}$$

The language of Σ^* contains a new constant symbol c_θ for each L formula $\theta(x)$. Σ^* is finitely satisfiable, since any finite subset of it is contained in a set of the form

$$F \cup \{(\exists x\, \theta_1(x)) \rightarrow \theta(c_{\theta_1}), \ldots, (\exists x\, \theta_n(x)) \rightarrow \theta(c_{\theta_n})\}$$

where $F \subseteq \Sigma$ is finite. To prove this set has a model use induction on n and note that from the point of view of

$$\Sigma \cup \{(\exists x\, \theta_1(x)) \to \theta(c_{\theta_1}), \ldots, (\exists x\, \theta_{n-1}(x)) \to \theta(c_{\theta_{n-1}})\}$$

c_{θ_n} is a new constant symbol, so we can apply the last lemma.

Now let $\Sigma_0 = \Sigma$ and let $\Sigma_{m+1} = \Sigma_m^*$. Then $\Sigma' = \cup_{m=0,1,\ldots}\Sigma_m$ is Henkin and finitely satisfiable, since it is the union of a chain of finitely satisfiable sets.

■

If Σ is a set of L sentences, then the canonical structure $\mathfrak{A}$ built from Σ is the following. Let X be the set of all variable free terms of L. For $t_1, t_2 \in X$ define $t_1 \sim t_2$ iff $(t_1 = t_2) \in \Sigma$. Assuming that $\sim$ is an equivalence relation let $[t]$ be the equivalence class of $t \in X$. The universe of the canonical model $\mathfrak{A}$ is the set of equivalence classes of $\sim$. For any n-ary relation symbol R we define $([t_1], [t_2], \ldots, [t_n]) \in R^{\mathfrak{A}}$ iff $R(t_1, t_2, \ldots, t_n) \in \Sigma$. We define $f^{\mathfrak{A}}([t_1], [t_2], \ldots, [t_n]) = [f(t_1, t_2, \ldots, t_n)]$. and $c^{\mathfrak{A}} = [c]$ for n-ary operation symbols f and constant symbols c.

Lemma 5 *If Σ is a finitely satisfiable complete Henkin set of L sentences, then the canonical model $\mathfrak{A}$ built from Σ is well defined and for every L sentence θ, $\mathfrak{A} \models \theta$ iff $\theta \in \Sigma$.*

proof:

First we show that $\sim$ is an equivalence relation. Suppose t, t_1, t_2, t_3 are variable free terms.

$t \sim t$: If $t = t \notin \Sigma$ then, since Σ is complete we have that $\neg t = t \in \Sigma$. But clearly $\neg t = t$ has no models and so Σ is not finitely satisfiable.

$t_1 \sim t_2$ implies $t_2 \sim t_1$: If not by completeness of Σ we must have that $t_1 = t_2$ and $\neg t_2 = t_1$ are both in Σ. But then Σ was not finitely satisfiable.

($t_1 \sim t_2$ and $t_2 \sim t_3$) implies $t_1 \sim t_3$: If not by completeness of Σ we must have that $t_1 = t_2$, $t_2 = t_3$, and $\neg t_1 = t_3$ are all in Σ. But then Σ was not finitely satisfiable.

So $\sim$ is an equivalence relation. Next we show that it is a congruence relation.

Suppose $t_1, \ldots, t_n, t_1', \ldots, t_n'$ are variable free terms and f is an n-ary operation symbol. Then if $t_1 \sim t_1', \ldots, t_n \sim t_n'$ then $f(t_1, \ldots, t_n) \sim f(t_1', \ldots, t_n')$.

This amounts to saying if $\{t_1 = t'_1, \dots, t_n = t'_n\} \subseteq \Sigma$, then $f(t_1, \dots, t_n) = f(t'_1, \dots, t'_n) \in \Sigma$. But again since Σ is complete we would have

$$f(t_1, \dots, t_n) \neq f(t'_1, \dots, t'_n) \in \Sigma$$

but $\{t_1 = t'_1, \dots, t_n = t'_n, f(t_1, \dots, t_n) \neq f(t'_1, \dots, t'_n)\}$ has no models and so Σ wouldn't be finitely satisfiable.

By a completely similar argument: Suppose $t_1, \dots, t_n, t'_1, \dots, t'_n$ are variable free terms and R is an n-ary operation symbol. Then if $t_1 \sim t'_1, \dots, t_n \sim t'_n$ then $R(t_1, \dots, t_n) \in \Sigma$ iff $R(t'_1, \dots, t'_n) \in \Sigma$.

This shows the canonical model is well defined.

Now we prove by induction on the number of logical symbols for any L sentence θ

$$\mathfrak{A} \models \theta \iff \theta \in \Sigma$$

The atomic formula case is by definition.

$\neg$: $\mathfrak{A} \models \neg\theta$ iff not $\mathfrak{A} \models \theta$ iff(by induction) not $\theta \in \Sigma$ iff(by completeness) $\neg\theta \in \Sigma$.

$\vee$:$\mathfrak{A} \models (\theta \vee \rho)$ iff $\mathfrak{A} \models \theta$ or $\mathfrak{A} \models \rho$ iff(by induction) $\theta \in \Sigma$ or $\rho \in \Sigma$ iff $(\theta \vee \rho) \in \Sigma$. This last "iff" uses completeness and finite satisfiability of Σ. For left to right, if $(\theta \vee \rho) \notin \Sigma$ then by completeness $\neg(\theta \vee \rho) \in \Sigma$ but $\{\theta, \rho, \neg(\theta \vee \rho)\}$ has no model. For right to left, if $\theta \notin \Sigma$ and $\rho \notin \Sigma$, then by completeness $\neg\theta \in \Sigma$ and $\neg\rho \in \Sigma$, but $\{(\theta \vee \rho), \neg\theta, \neg\rho\}$ has no model.

$\exists$: $\mathfrak{A} \models \exists x\theta(x)$ implies there exists $a \in A$ such that $\mathfrak{A} \models \theta(a)$ implies(by induction) $\theta(a) \in \Sigma$ implies $\exists x\theta(x) \in \Sigma$, (since otherwise $\neg\exists x\theta(x) \in \Sigma$ but $\{\neg\exists x\theta(x), \theta(a)\}$ has no model. Hence $\mathfrak{A} \models \exists x\theta(x)$ implies $\exists x\theta(x) \in \Sigma$. For the other direction suppose $\exists x\theta(x) \in \Sigma$, then since Σ is Henkin for some constant symbol c we have $(\exists x\theta(x)) \rightarrow \theta(c) \in \Sigma$. Using completeness and finitely satisfiable we must have $\theta(c) \in \Sigma$. By induction $\mathfrak{A} \models \theta(c)$ hence $\mathfrak{A} \models \exists x\theta(x)$.

■

Compactness Theorem. For any language L and set of L sentences Σ,

Σ has a model iff every finite subset of Σ has a model.

proof:

First Henkinize Σ, then complete it. Then take its canonical model.

■

Some naive set theory

Define $|X| = |Y|$ iff there is a one-to-one onto map from X to Y. We say X and Y have the same cardinality. Define $|X| \leq |Y|$ iff there is a one-to-one map from X to Y. We may think of the cardinal numbers as these equivalence classes which we will denote with greek letters κ or γ. The axiom of choice implies that for any two sets X and Y either $|X| \leq |Y|$ or $|Y| \leq |X|$. The Cantor-Bendixson Theorem says that if $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$.

The first infinite cardinal is written either ω or $\aleph_0$. A set is countable iff it is either finite or of the same cardinality as ω. A set is uncountable iff it is not countable. $\mathbb{R}$ is the set of real numbers and we use $\mathfrak{c} = |\mathbb{R}|$ to denote its cardinality which is also called the cardinality of the continuum. The first uncountable cardinal is written either $\aleph_1$ or ω_1. Cantor's continuum hypothesis is the assertion that $\mathfrak{c} = \aleph_1$.

A basic fact of the cardinality is that the union of κ many sets of cardinality κ has cardinal κ, eg the countable union of countable sets is countable. A more precisely statement would be: for any infinite cardinal κ and if $\{X_i : i \in I\}$ is family of sets such that $|X_i| \leq \kappa$ for each $i \in I$ and $|I| \leq \kappa$, then $|\bigcup\{X_i : i \in I\}| \leq \kappa$.

For cardinals κ and γ we define $\kappa + \gamma$ to be the cardinality of the union of A and B where A and B are disjoint and $|A| = \kappa$ and $|B| = \gamma$. Note that if at least one of κ and γ is infinite, then $\kappa + \gamma = max(\kappa, \gamma)$.

For κ a cardinal $\kappa^{<\omega}$ is the set of all finite sequences of elements of κ. It is a fact that if κ is an infinite cardinal, then $|\kappa^{<\omega}| = \kappa$. The reason is that $\kappa \times \kappa$ has cardinality κ since it the κ union of the sets $\alpha \times \kappa$ for $\alpha \in \kappa$. Similarly $\kappa \times \kappa \times \kappa = \kappa^3$ has cardinality κ. But then $|\kappa^{<\omega}| = \kappa$, since $\kappa^{<\omega} = \bigcup\{\kappa^n : n \in \omega\}$ is a countable union of sets of cardinality κ.

Lowenheim-Skolem Theorems

The first version of the Lowenheim-Skolem Theorem was proved in 1915. The final version that is presented here was developed by Tarski in the 1950's.

Lemma 6 *The number of L formulas is $|L| + \aleph_0$.*

proof:

There are only countably many logical symbols. Hence if $\kappa = |L| + \aleph_0$ then every formula may be regarded as an element of $\kappa^{<\omega}$ and we know $|\kappa^{<\omega}| = \kappa$.

■

For any theory T in a language L if T has a model, then it has one of cardinality less than or equal to $|L| + \aleph_0$. This is because the canonical model of the completion of Henkinization of T has cardinality $\leq |L| + \aleph_0$.

$\mathfrak{A} \subseteq \mathfrak{B}$ means that $\mathfrak{A}$ is a substructure of $\mathfrak{B}$, equivalently $\mathfrak{B}$ is an extension or superstructure of $\mathfrak{A}$. This means that both structures are in the same language L, $A \subseteq B$, for every n-ary relation symbol R of L, $R^{\mathfrak{A}} = R^{\mathfrak{B}} \cap A^n$, for every n-ary function symbol f of L, $f^{\mathfrak{A}} = f^{\mathfrak{B}} \upharpoonright A^n$, and for every constant symbol c of L, $c^{\mathfrak{A}} = c^{\mathfrak{B}}$.

$\mathfrak{A} \preceq \mathfrak{B}$ means that $\mathfrak{A}$ is an elementary substructure of $\mathfrak{B}$, equivalently $\mathfrak{B}$ is an elementary extension of $\mathfrak{A}$. This means that $\mathfrak{A} \subseteq \mathfrak{B}$ and for every formula $\theta(x_1, x_2, \ldots, x_n)$ of the language L and for every $a_1, a_2, \ldots, a_n \in A$ we have

$$(\mathfrak{A}, a)_{a \in A} \models \theta(c_{a_1}, c_{a_2}, \ldots, c_{a_n}) \text{ iff } (\mathfrak{B}, a)_{a \in A} \models \theta(c_{a_1}, c_{a_2}, \ldots, c_{a_n})$$

To ease the notational complexity we will write $\mathfrak{A} \models \theta(a_1, \ldots, a_n)$ instead of $(\mathfrak{A}, a)_{a \in A} \models \theta(c_{a_1}, c_{a_2}, \ldots, c_{a_n})$. It should be kept in mind that the language L may have no constant symbols in it.

Problem: Let $\mathfrak{A} = (\omega, <)$ and let $\mathfrak{B} = (Evens, <)$. Show that $\mathfrak{B} \subseteq \mathfrak{A}$ but $\mathfrak{B}$ is not an elementary substructure of $\mathfrak{A}$.

Lemma 7 *(Tarski-Vaught criterion) Suppose* $\mathfrak{A} \subseteq \mathfrak{B}$ *are* L *structures and for every* L *formula* $\theta(x, y_1, y_2, \ldots, y_n)$, *for all* $a_1, a_2, \ldots, a_n \in A$, *and* $b \in B$

$$\mathfrak{B} \models \theta(b, a_1, a_2, \ldots, a_n) \textit{ implies there exists } a \in A \;\; \mathfrak{B} \models \theta(a, a_1, a_2, \ldots, a_n)$$

then $\mathfrak{A} \preceq \mathfrak{B}$.

proof:

The proof is by induction on the number of logical symbols in the formula θ. The atomic formula case is trivial because $\mathfrak{A}$ is a substructure of $\mathfrak{B}$.

$\neg$: $\mathfrak{A} \models \neg\theta$ iff not $\mathfrak{A} \models \theta$ iff(by induction) not $\mathfrak{B} \models \theta$ iff $\mathfrak{B} \models \neg\theta$.

$\vee$: $\mathfrak{A} \models (\theta \vee \rho)$ iff $\mathfrak{A} \models \theta$ or $\mathfrak{A} \models \rho$ iff(by induction) $\mathfrak{B} \models \theta$ or $\mathfrak{B} \models \rho$ iff $\mathfrak{B} \models (\theta \vee \rho)$.

$\exists$: $\mathfrak{A} \models \exists x\theta(x, a_1, \ldots, a_n)$ implies there exists $a \in A$ such that

$$\mathfrak{A} \models \theta(a, a_1, \ldots, a_n)$$

which implies (by induction) $\mathfrak{B} \models \theta(a, a_1, \ldots, a_n)$. For the other direction we use the criterion.

$\mathfrak{B} \models \exists x\theta(x, a_1, \ldots, a_n)$ implies there exists $a \in A$ such that

$$\mathfrak{B} \models \theta(a, a_1, \ldots, a_n).$$

Hence by induction $\mathfrak{A} \models \theta(a, a_1, \ldots, a_n)$ and so $\mathfrak{A} \models \exists x\theta(x, a_1, \ldots, a_n)$.
■

Lemma 8 *Suppose* $X \subseteq B$ *and* $|X| = \kappa \geq |L| + \aleph_0$ *where* $\mathfrak{B}$ *is an* L *structure. Then there exists* $X^* \supseteq X, |X^*| = \kappa$*, and for every formula* $\theta(x, y_1, y_2, \ldots, y_n)$*, for all* $a_1, a_2, \ldots, a_n \in X$*, and* $b \in B$

$$\mathfrak{B} \models \theta(b, a_1, a_2, \ldots, a_n) \rightarrow \exists a \in X^* \; \mathfrak{B} \models \theta(a, a_1, a_2, \ldots, a_n)$$

proof:

A linear order $(L, \leq)$ is a wellorder iff for every nonempty $X \subseteq L$ there exists $x \in X$ such that for every $y \in X$ $x \leq y$. The axiom of choice implies that every set can be wellordered. Fix $\leq$ a wellordering of B. For any L formula $\theta(x, y_1, \ldots, y_n)$ and $a_1, \ldots, a_n \in B$ define $a_{\theta(x,a_1,\ldots,a_n)} \in B$ to be the $\leq$ least element b of B such that $\mathfrak{B} \models \theta(b, a_1, \ldots, a_n)$ if one exists otherwise let it be arbitrary. Now let $X_0 = X$, $L_0 = L$, and for any $m < \omega$ let $X_{m+1} =$

$$\{a_{\theta(x,a_1,\ldots,a_n)} : \theta(x, y_1, \ldots, y_n) \text{ is an } L_m \text{ formula and } \{a_1, \ldots, a_n\} \subseteq X_m\}$$

and let $L_{m+1} \supset L_m$ be the language with all these new constant symbols adjoined. Clearly if X_m and L_m have cardinality κ then so do X_{m+1} and L_{m+1}, since $|\kappa^{<\omega}| = \kappa$. Let $X^* = \cup_{m<\omega} X_m$, then it has cardinality κ since it is the countable union of sets of cardinality κ. For every formula $\theta(x, a_1, a_2, \ldots, a_n)$ there exist some $m < \omega$ such that $\{a_1, a_2, \ldots, a_n\} \subseteq X_m$ and so the criterion for $\theta(x, a_1, a_2, \ldots, a_n)$ is satisfied at stage $m + 1$.
■

Definition: $|\mathfrak{A}|$ is the cardinality of the universe A of $\mathfrak{A}$.

Downward Lowenheim-Skolem Theorem. Suppose $\mathfrak{B}$ is an infinite structure in the language L, κ is a cardinal such that $\aleph_0 + |L| \leq \kappa \leq |\mathfrak{B}|$, and $X \subseteq B$ such that $|X| \leq \kappa$. Then there is a structure $\mathfrak{A}$ such that $\mathfrak{A} \preceq \mathfrak{B}$, $X \subseteq A$, and $|\mathfrak{A}| = \kappa$.

proof:

By the lemma there exists $X^* \supseteq X$ of cardinality κ satisfying the criterion. But note that the criterion implies that X^* is closed under the operations of $\mathfrak{B}$. (Just look at the sentence $\exists x \;\; x = f(a_1, \ldots, a_n)$.) Consequently there is a substructure $\mathfrak{A}$ of $\mathfrak{B}$ with universe $A = X^*$. By the Tarski-Vaught criterion $\mathfrak{A} \preceq \mathfrak{B}$.

■

$\mathfrak{A} \simeq \mathfrak{B}$ means that $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic, that is, there is a bijection $j : A \to B$ such that for every n-ary relation symbol R and for every $a_1, a_2, \ldots, a_n \in A$,

$$\langle a_1, a_2, \ldots, a_n \rangle \in R^{\mathfrak{A}} \text{ iff } \langle j(a_1), j(a_2), \ldots, j(a_n) \rangle \in R^{\mathfrak{B}}$$

and for every n-ary function symbol f and for every $a_1, a_2, \ldots, a_n \in A$,

$$j(f^{\mathfrak{A}}(a_1, a_2, \ldots, a_n)) = f^{\mathfrak{B}}(j(a_1), j(a_2), \ldots, j(a_n))$$

(for n=0 this means that for every constant symbol c, $j(c^{\mathfrak{A}}) = c^{\mathfrak{B}}$.)

Lemma 9 *Suppose j is an isomorphism between the L structures $\mathfrak{A}$ and $\mathfrak{B}$. Then for any L formula $\theta(x_1, x_2, \ldots, x_n)$ and any $a_1, a_2, \ldots, a_n \in A$,*

$$\mathfrak{A} \models \theta(a_1, a_2, \ldots, a_n) \textit{ iff } \mathfrak{B} \models \theta(j(a_1), j(a_2), \ldots, j(a_n))$$

proof:

First show by induction that for every L-term $\tau(x_1, \ldots, x_n)$ and sequence $a_1, \ldots, a_n \in A$ that $j(\tau^{\mathfrak{A}}(a_1, \ldots, a_n))) = \tau^{\mathfrak{B}}(j(a_1), \ldots, j(a_n)$. The proof of the lemma is by induction on the logical complexity of θ. For atomic formula it follows from the definition of isomorphism. The propositional steps are easy and the quantifier step is handled by using that j is onto.

■

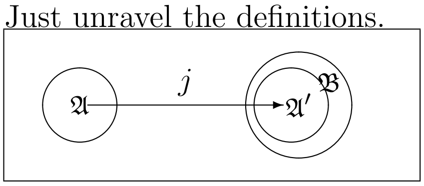
A map $j : \mathfrak{A} \to \mathfrak{B}$ is an elementary embedding iff it is an isomorphism of $\mathfrak{A}$ with an elementary substructure of $\mathfrak{B}$, we write $\mathfrak{A} \preceq^j \mathfrak{B}$.

Lemma 10 *A map $j : \mathfrak{A} \to \mathfrak{B}$ is an elementary embedding iff for any formula $\theta(x_1, x_2, \ldots, x_n)$ and any $a_1, a_2, \ldots, a_n \in A$,*

$$\mathfrak{A} \models \theta(a_1, a_2, \ldots, a_n) \text{ iff } \mathfrak{B} \models \theta(j(a_1), j(a_2), \ldots, j(a_n))$$

proof:

Just unravel the definitions.



■

Lemma 11 *Suppose that $\mathfrak{A} \preceq^j \mathfrak{B}$, then there exists a structure $\mathfrak{B}'$ isomorphic to $\mathfrak{B}$ such that $\mathfrak{A} \preceq \mathfrak{B}'$. Furthermore j is the restriction of this isomorphism to A.*

proof:

Let B' be a superset of A such that the map j can be extended to a bijection $j : B' \to B$, (which we also will call j). Now define $f^{\mathfrak{B}'}$ and $R^{\mathfrak{B}'}$ in such away as to make j an isomorphism. This means that

$$R^{\mathfrak{B}'} = \{(b_1, \ldots, b_n) : (j(b_1), \ldots, j(b_n)) \in R^{\mathfrak{B}}\}$$

$$f^{\mathfrak{B}'}(b_1, \ldots, b_n) = j^{-1}(f^{\mathfrak{B}}(j(b_1), \ldots, j(b_n))$$

Now check that $\mathfrak{B}'$ works.

■

The elementary diagram of $\mathfrak{A}$ is defined as follows:

$$D(\mathfrak{A}) = \{\theta : \theta \text{ is an } L_{\mathfrak{A}} \text{ sentence and } (\mathfrak{A}, a)_{a \in A} \models \theta\}$$

This means that $D(\mathfrak{A})$ is the theory of $\mathfrak{A}$ with constants adjoined for each element of the universe.

Lemma 12 *If $\mathfrak{A}$ is an L structure and $\mathfrak{B}$ is an $L_{\mathfrak{A}}$ structure such that $\mathfrak{B} \models D(\mathfrak{A})$, then there is a j such that $\mathfrak{A} \preceq^j \mathfrak{B} \restriction L$.*

proof:

Define $j : A \to B$ by $j(a) = c_a^{\mathfrak{B}}$.

■

Upward Lowenheim-Skolem Theorem. For any infinite structure $\mathfrak{A}$ in the language L and cardinal κ such that $|\mathfrak{A}| + |L| \leq \kappa$, there is a structure $\mathfrak{B}$ such that $\mathfrak{A} \preceq \mathfrak{B}$ and $|\mathfrak{B}|{=}\kappa$.

proof:

Let $\Sigma = D(\mathfrak{A}) \cup \{c_\alpha \neq c_\beta : \alpha, \beta < \kappa, \alpha \neq \beta\}$ where the c_α are totally new constant symbols.

Σ is finitely satisfiable. To see this let $F \subseteq \Sigma$ be finite. Then there exists a finite $G \subseteq \kappa$ such that $F \subseteq D(\mathfrak{A}) \cup \{c_\alpha \neq c_\beta : \alpha, \beta \in G, \alpha \neq \beta\}$. Since the model $\mathfrak{A}$ is infinite we can choose distinct elements of $a_\alpha \in A$ for $\alpha \in G$ and then $(\mathfrak{A}, a_\alpha)_{\alpha \in G} \models F$.

It follows from the compactness theorem that Σ has a model. Since the language of Σ has cardinality κ by the downward Lowenheim Skolem theorem Σ has a model $\mathfrak{C}$ of cardinality κ. By the lemma there exists j such that $\mathfrak{A} \preceq^j \mathfrak{C} \restriction L$. By the other lemma $\mathfrak{C} \restriction L$ is isomorphic to a model $\mathfrak{B}$ such that $\mathfrak{A} \preceq \mathfrak{B}$.

■