Math 141a - Mathematical Logic I

Taught by Sebastien Vasey
Notes by Dongryul Kim

Fall 2018

¡+instructor+¿ ¡+meetingtimes+¿ ¡+textbook+¿ ¡+enrolled+¿ ¡+grading+¿ ¡+courseassistants+¿

Contents

1 September 7, 2018

Logic is roughly studying the foundational objects of math, for instance, sets, statements, proofs, etc.

1.1 Overview

Let me tell you few of the theorems we are going to discuss.

Theorem 1.1 (Gödel's completeness theorem). *Let T be a list of first-order axioms, and let φ be a first-order statement. Then $T \vdash \varphi$ if and only if $T \vDash \varphi$.*

The first symbol $T \vdash \varphi$ means that there is a proof of φ from the axioms in T. The second symbol $T \vDash \varphi$ means that any structure satisfying the axioms in T also satisfies φ. A proof shows that it is true for every structure, but the other direction is subtle. It means that if I can't find a unicorn everywhere, then there is a proof that show that unicorns don't exist.

Example 1.2. Let R be a binary relation, and let

$$\begin{aligned} T &= \text{“}R\text{ is an equivalence relation”} \\ &= \{\forall x R(x,x), \forall x \forall y (R(x,y) \to R(y,x)), \\ &\qquad \forall x \forall y \forall z (R(x,y) \wedge R(y,z) \to R(x,z))\}. \end{aligned}$$

So if there is a statement that is true for every equivalence relation, it has a proof. For instance,

$$\varphi = \forall x \forall y \forall z ((R(x,y) \wedge \neg R(y,z)) \to \neg R(x,z))$$

has a proof.

So it is an interesting relation between syntax and semantics. Some cool consequences include the compactness theorem.

Theorem 1.3 (compactness theorem). *Let T be a list of first-order axioms. If every finite subset of T is satisfied by some structure, then T is satisfied by a structure.*

Consider the structure of $(\mathbb{R}, +, \cdot, 0, 1)$. Let us abstractly look at all the statements that are true for the real numbers and call this set T. For instance, $\forall x \forall y (x \cdot x + y \cdot y = 0 \to x = 0 \wedge y = 0)$. Now what we can do is to consider

$$T' = T \cup \{0 < c, c < 1, c < \tfrac{1}{2}, c < \tfrac{1}{3}, \dots\}.$$

Then every finite subset of $T_0 \subseteq T'$ is a subset of $T \cup \{0 < c, c < 1, \dots, c < \frac{1}{n}\}$ for some n. This is satisfied by $(\mathbb{R}, +, \cdot, 0, 1, c = \frac{1}{n+1})$. By compactness, there is a structure satisfying this, say $\mathbb{R}^*$. One way to actually construct it is to take an ultraproduct of $\mathbb{R}$. Using this, you can do non-standard analysis.

Another application of the compactness theorem is the Ax–Grothendieck theorem.

Theorem 1.4 (Ax–Grothendieck). *If $f : \mathbb{C}^n \to \mathbb{C}^n$ is a polynomial mapping and f is injective, then f is surjective.*

Note that an injective function from a finite set to itself is automatically bijective. In this case, using the compactness theorem, you can pretend that $\mathbb{C}$ is a finite set. There are other proofs, but they are nontrivial.

We can also talk about the back and forth method. You can show that $(\mathbb{Q}, <)$ is the unique countable dense linear order without endpoints. This also shows that the first-order theory of $(\mathbb{Q}, <)$ is decidable, i.e., that is an algorithm that proves of disproves anything about $(\mathbb{Q}, <)$.

1.2 Counting

We can count pass infinity as

$$0, 1, 2, \dots, n, \dots, \omega, \omega + 1, \omega + 2, \dots, \omega + \omega, \dots, \omega \cdot \omega, \dots, \omega^{\omega}, \dots.$$

These are called **ordinals**. We define an ordinal as the set of ordinals below it, for instance as $\alpha + 1 = \alpha \cup \{\alpha\}$. They will be used to generalize induction to transfinite induction.

We can also define **cardinals**. We say that the two sets X and Y have the same cardinality if there is a bijection between them. We define the cardinality of X as the least ordinal α that has the same cardinality as X.

Proposition 1.5 (well-ordering principle). *The statement that every set has a cardinality is equivalent to the Axiom of Choice.*

2 September 10, 2018

Ordinals are like countings.

2.1 Ordinals

Definition 2.1. A **chain** is a pair $(A, <)$ where A is a set and $<$ is a binary relation on A which is:

- transitive, if $x < y$ and $y < z$ then $x < z$,
- irreflexive, $x < x$ for all $x < x$,
- total, if $x \neq y$ then either $x < y$ or $y < x$.

Example 2.2. The following are all chains: $(\mathbb{N}, <)$, $(\mathbb{Z}, <)$, $(\mathbb{Q}, <)$, $(\{0, 1\}, <)$. But $(\{\emptyset, \{0\}, \{1\}\}, \subsetneq)$ is not a chain.

For $(A, <)$ and $(B, <)$ chains, a function $f : (A, <) \to (B, <)$ is called **order-preserving** if $a_1 < a_2$ implies $f(a_1) < f(a_2)$. An isomorphism is an order-preserving bijection.

Example 2.3. The function $f : (\mathbb{N}, <) \to (\mathbb{N}, <)$ given by $n \mapsto n + 1$ is order-preserving. But $\mathbb{Z} \to \mathbb{N}$ given by $n \mapsto |n|$ is not order-preserving. In fact, there is no order-preserving map for $\mathbb{Z}$ to $\mathbb{N}$.

We can define $A + B$ for A and B chains, given by $A \amalg B$ with $a < b$ for all $a \in A$ and $b \in B$. We can also define $A \cdot B$ with the lexicographical order.

Definition 2.4. A **well-ordering** is a chain $(A, <)$ such that for every $S \subseteq A$ nonempty, there is a minimal element $x \in A$.

Any finite chain is a well-ordering, but $(\mathbb{Z}, <)$ is not.

Lemma 2.5. *If $(A, <)$ and $(B, <)$ are well-orderings, then either A is isomorphic to an initial segment of A.*

Definition 2.6. For $(A, <)$ a chain, a subset $A_0 \subseteq A$ is called an **initial segment** if for any $a < b$, $b \in A_0$ implies $a \in A_0$. That is, if $a \in A_0$ then

$$\operatorname{pred}_A(a) = \{b \in A : b < a\}$$

is in A_0.

So if you have two well-orderings, they are comparable. If $(A, <)$ is a well-ordering and $A_0 \subseteq A$ is an initial segment, then either $A_0 = A$ or $A \setminus A_0$ has a least element and

$$A_0 = \operatorname{pred}_A(a).$$

Indeed, any well-ordering is isomorphic to the set of predecessors, ordered by inclusion.

Lemma 2.7. *Let $(A,<)$ and $(B,<)$ be well-orderings. Let $f, g : (A,<) \to (B,<)$ be isomorphisms onto initial segments. Then $f = g$.*

Proof. Assume $f \neq g$, and then there exists a minimal $a \in A$ where $f(a) \neq g(a)$. Assume $f(a) < g(a)$, without loss of generality. Because $g[A]$ is an initial segment, we have $f(a) \in g[A]$. If we let $a' \in A$ be such that $g(a') = f(a)$, then $g(a') = f(a) < g(a)$ implies that $a' < a$. But $f(a') = g(a') = f(a)$ gives a contradiction. □

Now we can prove the lemma.

Proof of Lemma 2.5. We look at the set of $a \in A$ such that $\mathrm{pred}(a)$ is not isomorphic to a proper initial segment of B. If this set is nonempty, we may take a minimal a with this property. For any $a_0 < a$, we have that $\mathrm{pred}(a_0)$ is isomorphic to $\mathrm{pred}(b_{a_0})$ for some $b_{a_0} \in B$. This is moreover unique. If we let

$$f : \mathrm{pred}(a) \to B; \quad a_0 \mapsto b_{a_0},$$

this is order-preserving isomorphism onto an initial segment of B. It cannot be proper by assumption, so it is an isomorphism. Then $f^{-1} : B \to A$ shows that B is an isomorphism to an initial segment of A.

Now assume that all $\mathrm{pred}(a)$ are isomorphic to initial segments of B. If we pick $b_a \in B$ so that $\mathrm{pred}(a) \cong \mathrm{pred}(b_a)$, then

$$f : (A,<) \to (B,<); \quad a \mapsto b_a$$

is an order-preserving isomorphism to an initial segment of B. □

Ordinals are canonical representatives of well-orderings. Every ordinal will be the set of its predecessors.

Definition 2.8. An **ordinal** is a set α which is

- transitive, $x \in \alpha$ and $y \in x$ then $y \in \alpha$,
- $(\alpha, \in)$ is a well-ordering,

Examples include

$$0 = \emptyset, \quad 1 = \{\emptyset\}, \quad 2 = \{0,1\}, \quad \dots, \quad \omega = \{0,1,2,\dots\}, \quad \omega+1 = \omega \cup \{\omega\}, \dots.$$

If α is an ordinal, you can take $\alpha + 1 = \alpha \cup \{\alpha\}$, which is again an ordinal. If $(\alpha_i)_{i \in I}$ are ordinals, then

$$\alpha = \bigcup_{i \in I} \alpha_i$$

is an ordinal, called $\sup_{i \in I} \alpha_i$. For instance, $\omega = \sup_{n \in \omega} n$. If $x \in \alpha$, then

$$\mathrm{pred}_{(\alpha,\in)}(x) = x.$$

Lemma 2.9. *If α and β are isomorphic ordinals, then $\alpha = \beta$.*

Proof. Let $f : (\alpha, \in) \cong (\beta, \in)$. We claim that f is the identity. If not, there exists a minimal $a \in \alpha$ such that $f(a) \neq a$. Then

$$f(a) = \mathrm{pred}_{(\beta,\in)}(f(a)) = f[a] = a$$

because f is the identity on a. $\square$

Lemma 2.10. *Any well-ordering is uniquely isomorphic to a unique ordinal.*

Proof. We claim that if $a \in A$ has $\mathrm{pred}(a) \cong (\alpha_a, \in)$, then we can take

$$\alpha = \{\alpha_a : a \in A\}$$

and then α is an ordinal and $a \mapsto \alpha_a$ is the desired isomorphism. If there is $a \in A$ such that $\mathrm{pred}(a)$ is not isomorphic to an ordinal, we can take the minimal one. Then applying the claim gives a contradiction. $\square$

3 September 14, 2018

Last time we defined an ordinal as a transitive set such that $(\alpha, \in)$ is a well-ordering. We showed that any well-ordering is isomorphic to a unique ordinal. The intuition is that an ordinal is the set of its predecessors. For α, β ordinals, we are going to write $\alpha < \beta$ instead of $\alpha \in \beta$. In the homework, you are going to show that for α and β ordinals, either $\alpha = \beta$ or $\alpha < \beta$ or $\beta < \alpha$.

3.1 Operations on ordinals

- Given an ordinal α, we define $\alpha + 1 = \alpha \cup \{\alpha\}$.
- Given $(\alpha_i)_{i \in I}$ a set of ordinals, we define $\sup_{i \in I} \alpha_i = \bigcup_{i \in I} \alpha_i$. This is the least α such that $\alpha \geq \alpha_i$ for all $i \in I$.

Definition 3.1. For ordinals α and β, we define $\alpha + \beta$ to be the unique ordinal isomorphic to $(\alpha, \in) + (\beta, \in)$. Likewise, $\alpha \cdot \beta$ is the unique ordinal isomorphic to $(\alpha, \in)(\beta, \in)$, which is α copied β times.

On finite ordinals, these are usual addition and multiplication. We have

$$1 + \omega = \omega, \quad \omega + 1 > \omega, \quad \omega \cdot 2 = \omega + \omega, \quad 2 \cdot \omega = \omega.$$

You can do division: if α is an ordinal and $\beta > 0$, then there exist unique ordinals γ and $\delta < \beta$ such that

$$\alpha = \beta \cdot \gamma + \delta.$$

Lemma 3.2 (transfinite induction)**.** *Any nonempty collection S of ordinals has a minimal element.*

Proof. Pick $\alpha \in S$. If α is minimal, we are done. Otherwise, we can take the minimal element in $S \cap \alpha$. $\square$

Corollary 3.3. *Let $P(x)$ be a property of ordinals. Suppose that*

For any ordinal α, $P(\beta)$ for all $\beta < \alpha$ implies $P(\alpha)$.

Then $P(\alpha)$ for all ordinal α.

Proof. If not there is a minimal α such that $P(\alpha)$ is false. This contradicts our assumptions. $\square$

There are three types of ordinals. That is, for any ordinal α, exactly one of the following three is true:

- $\alpha = 0$
- $\alpha = \beta + 1$ for some β (these are called **successors**)
- $\alpha > 0$ and $\beta + 1 < \alpha$ for any $\beta < \alpha$ (these are called **limit ordinals**).

So we can we can restate transfinite induction as the following.

Corollary 3.4. *Let $P(x)$ be a property of ordinals. Suppose that*

- $P(0)$*,*
- $P(\alpha)$ *implies* $P(\alpha+1)$*,*
- $P(\beta$ *for all* $\beta < \alpha$ *implies* $P(\alpha)$*, if* α *is a limit.*

Then $P(\alpha)$ *for all ordinals* α*.*

We can also define objects by transfinite induction. We define

- $\alpha + 0 = \alpha$,
- $\alpha + (\beta + 1) = (\alpha + \beta) + 1$,
- $\alpha + \beta = \sup_{\gamma<\beta} \alpha + \gamma$ if β is a limit ordinal.

This, you can check again by induction, is equivalent to the previous definition. Similarly, we can define

- $\alpha \cdot 0 = 0$,
- $\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$,
- $\alpha \cdot \beta = \sup_{\gamma<\beta} \alpha \cdot \gamma$ if β is a limit ordinal.

We can even define exponentiation as

- $\alpha^0 = 1$,
- $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$,
- $\alpha^\beta = \sup_{\gamma<\beta} \alpha^\gamma$.

Any ordinal has a base ω representation, so we can write

$$\alpha = c_1\omega^{\beta_1} + c_2\omega^{\beta_2} + \cdots + c_n\omega^{\beta_n},$$

where $c_i < \omega$.

3.2 Cardinalities

Theorem 3.5. *For any set X, there is an ordering such that $(X, <)$ is a well-ordering.*

For instance, for $X = \mathbb{R}$, the new ordering doesn't need to have anything to do with the usual ordering. For instance, we can pick things $a_0 = 0$, $a_1 = -1$, $a_2 = \frac{1}{2}$, $a_3 = \pi$, $a_4 = \sqrt{2}$, and so on. So we keep arbitrarily picking these elements. This is not a rigorous proof, and we are going to see the rigorous proof next time.

Definition 3.6. The **cardinality** $|X|$ of a set X is the minimal ordinal α such that there is a well-ordering of X isomorphic to α.

For instance,

$$|\omega| = \omega, \quad |\omega + 1| = \omega.$$

Definition 3.7. An ordinal is a **cardinal** if $\alpha = |\alpha|$.

For example, n is a cardinal for any $n < \omega$. Although ω is a cardinal, $\omega + 1, \omega + 2, \dots, \omega + \omega, \dots, \omega \cdot \omega$ are all not cardinals. For sets X and Y, there is a bijection from X to Y if and only if $|X| = |Y|$. There is an injection from X to Y if and only if $|X| \leq |Y|$. Note that if X and Y are two sets, either $|X| < |Y|$ or $|X| > |Y|$ or $|X| = |Y|$.

Theorem 3.8 (Cantor)**.** *For any set X, $|X| < |\mathcal{P}(X)|$.*

Proof. We have $|S| \leq |\mathcal{P}(X)|$ because $x \mapsto [x]$ is injective. Suppose for a contradiction that $|X| = |\mathcal{P}(X)|$. Then there should be a bijection

$$F : X \to \mathcal{P}(X).$$

Now consider the set

$$Y = \{x \in X : x \notin F(x)\} \subseteq X.$$

Then there is a $x \in X$ such that $F(x) = Y$. If $x \in Y$, then $x \in Y = F(x)$ so $x \notin Y$. On the other hand, if $x \notin Y$ then $x \notin F(x) - Y$ implies $x \in Y$. This gives a contradiction. □

Corollary 3.9. *For any cardinal κ, there is a cardinal $\lambda > \kappa$.*

Definition 3.10. Let κ^+ be the minimal cardinal above κ.

Then we can play around with the definitions. We can define

- $\aleph_0 = \omega$,
- $\aleph_{\alpha+1} = (\aleph_\alpha)^+$,
- $\aleph_\alpha = \sup_{\beta<\alpha} \aleph_\beta$ if α is a limit.

We can think of $\aleph_\alpha$ as the αth infinite cardinal.

Theorem 3.11. *For any cardinal λ, there is α such that $\lambda = \aleph_\alpha$.*

Proof. We do this by induction on λ. Take the minimal λ where this fails. Then either $\lambda = \kappa^+$ or $\kappa^+ < \lambda$ for all $\kappa < \lambda$. Apply the induction hypothesis. □

The continuum hypothesis states that $\aleph_1 = |\mathbb{P}(\mathbb{N})| = 2^{\aleph_0}$. The generalized continuum hypothesis that $\kappa^+ = |\mathcal{P}(\kappa)|$ for every infinite cardinal κ.

4 September 17, 2018

Recall that we had this theorem.

Theorem 4.1. *Any set can be well-ordered.*

Using it, we can defined the cardinality $|X|$ as the least ordinal such that there is a well-ordering $<$ on X of type α. We also defined $\aleph_\alpha$ as the αth infinite cardinal.

4.1 Cardinal arithmetic

Given sets A and B, we can ask what

$$|A \cup B|, \quad |A \times B|, \quad |{}^B A|, \quad |\mathcal{P}(A)|,$$

and so on.

Definition 4.2. Let λ and μ be cardinals. We define

- $\lambda +^c \mu = |(\lambda \times \{1\}) \cup (\mu \times \{2\})|$,
- $\lambda \cdot^c \mu = |\lambda \times \mu|$,
- $\lambda^{c,\mu} = |{}^\mu \lambda|$.

(For today, we will drop the c.)

You can check that for finite cardinals, this agrees with the usual operations. We also have basic properties like

$$2^\lambda = |\mathcal{P}(\lambda)|, \quad (\lambda^\mu)^\kappa = \lambda^{\mu \cdot \kappa}.$$

Exponentiation is really hard; the continuum hypothesis is $2^{\aleph_0} = \aleph_1$. But we will see for infinite λ and μ, we have

$$\lambda + \mu = \lambda \cdot \mu = \max(\lambda, \mu).$$

Theorem 4.3. *For an infinite cardinal λ, we have $\lambda \cdot^c \lambda = \lambda$.*

Proof. It suffices to show that there is a well-ordering in $\lambda \times \lambda$ that has order type λ. We define the order by the lexicographical ordering on $(\max(\alpha, \beta), \alpha, \beta)$. We can check that this is a well-ordering. So it has an order type.

By induction on λ, we prove that $\lambda \times \lambda$ has order type λ. If $\lambda = \aleph_0$, we can explicitly describe this. Now assume $\lambda > \aleph_0$ and $\mu \cdot \mu = \mu$ for any infinite $\mu < \lambda$. Then for any $(\alpha, \beta) \in \lambda \times \lambda$, we have

$$|\mathrm{pred}(\alpha, \beta)| \le \max(\alpha, \beta) \cdot \max(\alpha, \beta) = \mu \cdot \mu = \mu < \lambda$$

for some $\mu < \lambda$. This implies $\lambda \cdot \lambda \le \lambda$, as needed. $\square$

Corollary 4.4. *For infinite λ and μ, we have*

$$\lambda +^c \mu = \lambda \cdot^c \mu = \max(\lambda, \mu).$$

Proof. This is commutative, so we may assume $\mu \leq \lambda$. Then

$$\lambda \leq \lambda +^c \mu \leq \lambda \cdot^c \mu \leq \lambda \cdot^c \lambda = \lambda.$$

This finishes the proof. $\square$

You can also get some other strange results.

Corollary 4.5. $(\lambda^+)^\lambda = 2^\lambda$.

Proof. We have

$$2^\lambda \leq (\lambda^+)^\lambda \leq (2^\lambda)^\lambda = 2^{\lambda \cdot^c \lambda} = 2^\lambda.$$

So we have equality. $\square$

In fact, we can prove that if we know $\lambda \mapsto 2^\lambda$, e.g.e, if we assume $2^\lambda = \lambda^+$ for any λ, then we know λ^μ for any infinite λ and μ.

4.2 Axiom of choice

What are sets? Naïvely we can say that it is some collections objects, but some collections are not sets.

Proposition 4.6 (Bural–Forb paradox)**.** *The collection* OR *of all ordinals is not a set.*

Proof. Suppose OR is a set. Then $(\mathrm{OR}, \in)$ is transitive and a well-ordering. So OR is an ordinal and so $\mathrm{OR} \in \mathrm{OR}$. This is a contradiction because $\in$ is supposed to be irreflexive. (We also just assume that there is no set that contains another.) $\square$

But we want anything that can be built out of a set to be a set.

- $\emptyset$ is a set.
- If A and B are sets, $A \cup B$, $\{A, B\}$, $A \times B$, ${}^B A$, $\mathcal{P}(A)$ are sets. (Here, if we can define $(a, b) = \{a, \{a, b\}\}$.)
- If A is a set, we can look at the set of all elements in A satisfying some property.
- If A is a set and for each $a \in A$ one defines a unique b_a, then $\{b_a : a \in A\}$ is a set.
- There is a set A such that $\emptyset \in A$ and if $x \in A$ then $x \cup \{x\} \in A$.

Definition 4.7. The **axiom of choice** says that for any set A, there is a function $f : \mathcal{P}(A) \setminus \{\emptyset\} \to A$ such that $f(A_0) \in A_0$ for all $a_0 \in \mathcal{P}(A) \setminus \emptyset$. We call such f a **choice function** on A.

For example, for $A = \{1, 2, 3\}$, we can find something like

$$f(\{i\}) = i, \quad f(\{i, j\}) = \min(i, j), \quad f(\{1, 2, 3\}) = 3.$$

The point is that the axiom of choice doesn't follow from the axioms. If A is an ordinal, we can pick a choice function

$$A_0 \mapsto \min(A_0).$$

But for other sets like $A = \mathcal{P}(\mathcal{P}(\omega))$ it is not clear how to construct this choice function.

Theorem 4.8. *The axiom of choice is equivalent to the statement that every set can be well-ordered.*

Proof. Assume every set can be well-ordered. Let A be a set, and pick a well-ordering on A. Then define

$$f : \mathcal{P}(A) \setminus \{\emptyset\} \to A; \quad A_0 \mapsto \min(A_0).$$

Now assume the axiom of choice, and assume that A cannot be well-ordered. Let $f : \mathcal{P}(A) \setminus \{\emptyset\} \to A$ be a choice function. Define by induction, a well-order on a subset of A, by putting in one of the elements that are not yet in the subset. If this process doesn't end, we get for every ordinal α, a subset of A and an order of type α. Let a_α be the element added at this step. Then we can define the union

$$A' = \{x \in A : \text{there exists } \alpha \text{ such that } a_\alpha = x\}.$$

But for each $a \in A'$, there exists a unique α such that $a = a_\alpha$. So by replacement,

$$\mathrm{OR} = \{\alpha_a : a \in A'\}$$

has to be a set. This is a contradiction. $\square$

5 September 21, 2018

We are now going to go to Chapter 1.

5.1 Cantor's theorem on chains

Theorem 5.1 (Cantor)**.** *Any two countable nonempty dense chains without endpoints are isomorphic.*

Definition 5.2. A chain $(A, <)$ is called **dense** if for any $x < y$ there is a z such that $x < z < y$.

Examples include $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$, without endpoints, and $(\mathbb{Q} \cap [0, 1], <)$, with endpoints. Endpoints are maximal or minimal elements. Any nonempty chain without endpoints is infinite, because you can always take bigger things.

Definition 5.3. A **local isomorphism** from $(A, <)$ to $(B, <)$ is an isomorphism $s : (A_0, <) \cong (B_0, <)$ where $A_0 \subseteq A$ and $B_0 \subseteq B$ are finite. We also write $A_0 = \mathrm{dom}(s)$ and $B_0 = \mathrm{im}(s)$.

This will be used in the proof. The idea is that we can build the isomorphism one by one, because the chains are dense without endpoints.

Proof. Let $(A, <)$ and $(B, <)$ be nonempty countable dense chains without endpoints. Then $|A| = |B| = \aleph_0$ by the exercise.

Now we claim the following. Suppose that s is a local isomorphism from $(A, <)$ to $(B, <)$.

- For any $a \in A$, there is a local isomorphism t such that $a \in \mathrm{dom}(t)$ and t extends s.
- For any $b \in B$, there is a local isomorphism t such that $b \in \mathrm{im}(t)$ and t extends s.

This is because both A and B are dense, and has no endpoints.

Now we alternative these two processes to inductively build a sequence. Then taking the union gives us the isomorphism. □

So for instance, $(\mathbb{Q}, <) \cong (\mathbb{Q} \cap (0, 1), <)$. However, this is no longer true for uncountable chains. The two chains $(\mathbb{R}, <)$ and $(\mathbb{R}, <) + (\mathbb{Q}, <)$ are not isomorphic. We also have that $(\mathbb{R} \setminus \{0\}, <)$ is not isomorphic to $(\mathbb{R}, <)$.

5.2 Relations

Definition 5.4. Given $1 \leq m < \omega$, an m-ary **relation** with universe E is a set $R \subseteq E^m$. If $\overline{a} = (a_1, \dots, a_n) \in R$, then we say that $\overline{a}$ satisfies R.

For m-ary relations (E, R) and (E', R'), we say that $f : (E, R) \cong (E', R')$ is an isomorphism if $f : E \to E'$ is a bijection such that $R(\overline{a})$ if and only if $R'(f(\overline{a}))$. Isomorphisms are closed under inverses and compositions.

If R is a m-ary relation with universe E, then for any subset $E' \subseteq E$ the restriction of R to E', written $R \cap E'$, is just R restricted to E'. As an abuse of language, we define the **cardinality** of R as the cardinality of E.

Definition 5.5. A **local isomorphism** from (E, R) to (E', R') is an isomorphism from a finite restriction of R to a finite restriction of R'.

We can define inductively on the ordinals α, the sets $S_\alpha(R, R')$ of local isomorphisms (called α**-isomorphisms**) from R to R' by

- $S_0(R, R')$ is the set of all local isomorphisms from R to R'.
- $S_{\alpha+1}(R, R')$ is the set of all local isomorphisms from R to R' such that
  - (i) for any $a \in E$, there is a local isomorphism such that t extends s, $a \in \mathrm{dom}(t)$, and $t \in S_\alpha(R, R')$,
  - (ii) for any $b \in E'$ there is a local isomorphism t such that t extends s, $b \in \mathrm{im}(t)$, and $t \in S_\alpha(R, R')$.
- $S_\alpha(R, R') = \bigcap_{\beta<\alpha} S_\beta(R, R')$ for α a limit.

Next time we will try to gain more intuition on what this is supposed to mean. But here are some basic properties. If s is an α-isomorphism and $\beta < \alpha$ then s is also a β-isomorphism. The class of α-isomorphisms is closed under composition, inverse, and restriction. If we started out with an honest isomorphism from R to R', then any finite restriction is an α-isomorphism for any ordinal α.

Definition 5.6. We say that s is an ∞**-isomorphism** if it is an α-isomorphism for all α.

For example, if $E = \mathbb{Q}$ and $R =<$, then any local isomorphism is an ∞-isomorphism. You can prove this by induction on α.

Definition 5.7. We say that R and R' are α**-equivalent** if $S_\alpha(R, R') \neq \emptyset$. This is equivalent to saying that the empty map is an α-isomorphism.

For special cases, we define ∞-equivalent as α-equivalent for all ordinals α. Another name for ω-equivalent is **elementarily equivalent**. For example,

$$(\mathbb{Q}, <) \sim_\infty (\mathbb{R}, <),$$

but we have

$$(\mathbb{N}, <) \sim_\omega (\mathbb{N}, <) + (\mathbb{Z}, <), \quad (\mathbb{N}, <) \not\sim_\infty (\mathbb{N}, <) + (\mathbb{Z}, <).$$

6 September 24, 2018

Last time we defined α-isomorphisms. An ω-isomorphism is also called an elementary isomorphism, and an ∞-isomorphism is an α-isomorphism for all α. This makes sense, because there is an α such that

$$S_\alpha(R, R') = S_{\alpha+1}(R, R') = \cdots .$$

Then we are saying $S_\alpha(R, R') = S_\infty(R, R')$ and any element of $S_\alpha(R, R')$ is an ∞-isomorphism.

Definition 6.1. We say that $(R, \overline{a}) \sim_\alpha (R', \overline{b})$ if there is an α-isomorphism s such that $s(a_i) = b_i$ for all i.

6.1 Hierarchy of local isomorphisms

We showed that any local isomorphism from a nonempty dense chain without endpoints to another is an ∞-isomorphism. The proof easily generalizes to the following.

Theorem 6.2 (1.14)**.** *If R and R' are countable and ∞-equivalent, they are isomorphic.*

Proof. You do the same thing, extending the maps back and forth. Then in a countable number of steps, you construct the isomorphism. □

Here are some basic observations:

- Any two relations are 0-equivalent, because the empty map is a local isomorphism.
- Assume R is a relation on E and $|E| = p$ is finite. If $R \sim_{p+1} S$, then $(E, R) \cong (E', S)$. This is because we can extend the empty $(p+1)$-isomorphism p times and get a 1-isomorphism, and then this has to be an actual isomorphism.
- Let E be an infinite set, and let $R = \emptyset$ be the empty unary relation. Let E' be another set, and let $R' = \emptyset$ be the empty relation. Then $R \sim_\infty R'$ if and only if E' is infinite. (Note that it is possible that $|E| \neq |E'|$.)

We can actually classify all unary relations up to ∞-equivalence.

Definition 6.3. We define the **character** of a unary relation R on a set E as the pair (x, y) where

$$x = \begin{cases} |R| & \text{if } |R| \text{ is finite} \\ \infty & \text{otherwise,} \end{cases} \qquad y = \begin{cases} |E \setminus R| & \text{if } |E \setminus R| \text{ is finite} \\ \infty & \text{otherwise.} \end{cases}$$

For instance, the character of (ω, odd) is (∞, ∞).

Theorem 6.4. *Two unary relations R and R' on universes E and E' are ∞-equivalent if and only if they have the same character.*

Proof. You can do this. If they don't have the same character, you can exhaust the finite ones and then we are not going to be able to extend this. If they have the same character, you can do it similarly to the previous claim. ☐

Binary relations are equivalent to m-ary relations, and so they will be hard to classify. Let us talk only about equivalence relations. These are relations that are reflexive, symmetric, and transitive.

Proposition 6.5. *If (E, R) is an equivalence relation, and $(E', R') \sim_3 (E, R)$. Then (E', R') is an equivalence relation.*

Proof. We need to prove three things. If we extend it to a size 1 empty isomorphism for some $y \in E'$, we get $y \sim y$ which is reflexivity. For transitivity, we need to extend three times. ☐

Theorem 6.6. *If (E, R) and (E', R') equivalence relations, with infinitely many classes, and all classes are infinite, then $R \sim_\infty R'$.*

Proof. The idea is that for any local isomorphism, you can extend it further by looking at the equivalence classes. ☐

6.2 Theory of discrete chains

Last time we showed that any two non-empty dense chains are ∞-equivalent.

Definition 6.7. A chain is **discrete** if any element that is not maximal has a successor and any element that is not minimal has a predecessor.

Examples include $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{Z} + \mathbb{Z}$, $\mathbb{Z} \times \mathbb{R}$, and so on. We will see that any two nonempty discrete chains without endpoints are $(\omega+1)$-equivalent. But we have

$$\mathbb{Z} \not\sim_{\omega+2} \mathbb{Z} + \mathbb{Z}.$$

To see this, we choose $(0, 1)$ and $(0, 2)$ in $\mathbb{Z} + \mathbb{Z}$ and try to map it into $\mathbb{Z}$. Then we get a map

$$(0, 1) \mapsto a, \quad (0, 2) \mapsto b$$

for some $a < b$. But this cannot be an ω-isomorphism, because there are infinitely many things between $(0, 1)$ and $(0, 2)$, but there are only finitely many things between a and b.

Lemma 6.8. *Let $(A, <)$ and $(B, <)$ be nonempty discrete chains without endpoints. Let $\overline{a} = (a_1, \ldots, a_k)$ and $\overline{b} = (b_1, \ldots, b_k)$. Then*

$$(\overline{a}, A, <) \sim_p (\overline{b}, B, <)$$

if for each $1 \leq i < k$, either $d(a_i, a_{i+1}) = d(b_i, b_{i+1})$ or $d(a_i, a_{i+1}), d(b_i, b_{i+1}) \geq 2^{p+1} - 1$.

Here, we are defining $d(a,b) = |\{x : a < x < b\}|$ where we don't distinguish between different sizes of infinity.

Proof. We do this inductively on p. If $p = 0$, this is clear. If the thing we want to extend falls in some small distance from what we already have, we extend by matching the distance. If not, we extend so that things are far away. □

Corollary 6.9. *Any two nonempty discrete chains without endpoints are* $(\omega + 1)$*-isomorphic.*

Proof. Pick $a \in A$, and extend it in any way. We now claim that this is an n-isomorphism for any $n < \omega$. This follows from the previous lemma. □

7 September 28, 2018

We were looking at chains and local isomorphisms. Intuitively, an α-isomorphism is an isomorphism that is up to some α level.

7.1 Formulas

We want to separate syntax and semantics. First we will define formulas syntactically.

Definition 7.1. The **alphabet** associated to an m-ary relation is:

- (,), ,, $\exists$, $\forall$, $\wedge$, $\vee$, $\neg$,
- $=$, r (symbol for the relation),
- variables v_0, v_1, v_2, v_3, $\dots$.

But nonsense like (($\neg$ is not a formula. So we want to talk about rules generating formulas.

Definition 7.2. The set of **formulas** (in the language associated with an m-ary relation) is defined as $F = \bigcup_{n<\omega} F_n$, where F_n is the set of formulas of complexity n.

- F_0 contains $v_i = v_j$ for $i, j < \omega$ and $r(v_{i_1}, \dots, v_{i_m})$ for $v_{i_1}, \dots, v_{i_m} < \omega$. These are also called atomic formulas.
- F_{n+1} contains $\neg(f)$, $(f) \vee (g)$, $(f) \wedge (g)$, $(\exists v_i)(f)$, $(\forall v_i)(f)$ for $f, g \in \bigcup_{i\leq n} F_i$, where at least one formula appearing as f or g is actually in F_n.

Then we form $F = \bigcup_{n<\omega} F_n$.

For example,

$$(\exists v_0)((\exists v_1)((r(v_0, v_1)) \vee (v_1 = v_2))$$

is in F_3. This "complexity" is going to be used to do induction.

Definition 7.3. We define the **quantifier rank** $\mathrm{QR}(f)$ of f as

- if f is atomic, $\mathrm{QR}(f) = 0$,
- if f is $(g) \vee (h)$ or $(g) \wedge (h)$ then $\mathrm{QR}(f) = \max(\mathrm{QR}(g), \mathrm{QR}(h))$,
- if f is $\neg(g)$ then $\mathrm{QR}(f) = \mathrm{QR}(g)$,
- if f is $(\exists v_i)(g)$ or $(\forall v_i)(g)$ then $\mathrm{QR}(f) = \mathrm{QR}(g) + 1$.

Definition 7.4. Define by induction on f, the set of **free variables** $\mathrm{FV}(f)$ of f,

- if f is atomic, then $\mathrm{FV}(f)$ is the set of variables appearing in the formula,
- if f is $(g) \vee (h)$ or $(g) \wedge (h)$ then $\mathrm{FV}(f) = \mathrm{FV}(g) \cup \mathrm{FV}(h)$,
- if f is $\neg g$ then $\mathrm{FV}(f) = \mathrm{FV}(g)$,

- if f is $(\exists v_i)(g)$ or $(\forall v_i)(g)$ then $\mathrm{FV}(f) = \mathrm{FV}(g) \setminus \{v_i\}$.

Definition 7.5. If $\mathrm{FV}(f_i) = \emptyset$, we will call f a **sentence**.

Now we are going to define what it means for a formula to be true or false. Let us write $f(\overline{x})$ with $\overline{x} = (v_{i_1}, \ldots, v_{i_n})$ to mean that f is a formula and $\mathrm{FV}(f) \subseteq \{v_{i_1}, \ldots, v_{i_n}\}$.

Definition 7.6. Assume R is an m-ary relation on E, and let $\overline{a} \in E^n$. Let $f(x_1, \ldots, x_n)$ be a formula. We are going to define what it means for R to satisfy $f(\overline{a})$ (written as $R \vDash f(\overline{a})$) by induction.

- If f is of the form $x_i = x_j$, then $R \vDash f(\overline{a})$ if and only if $a_i = a_j$.
- If f is of the form $r(x_{i_1}, \ldots, x_{i_m})$ then $R \vDash f(\overline{a})$ if and only if $(a_{i_1}, \ldots, a_{i_m}) \in R$.
- If f is $(g) \wedge (h)$ then $R \vDash f(\overline{a})$ if and only if $R \vDash g(\overline{a})$ and $R \vDash h(\overline{a})$.
- If f is $\neg(g)$ then $R \vDash f(\overline{a})$ if and only if $R \vDash g(\overline{a})$ is false.
- If f is $(\exists y)(g)$ and $\mathrm{FV}(g) \subseteq \{x_1, \ldots, x_n, y\}$, then $R \vDash f(\overline{a})$ if and only if there exists a $b \in E$ such that $R \vDash g(\overline{a}, b)$.
- If f is $(\forall y)(g)$ and $\mathrm{FV}(g) \subseteq \{x_1, \ldots, x_n, y\}$, then $R \vDash f(\overline{a})$ if and only if for all $b \in E$ we have $R \vDash g(\overline{a}, b)$.

For example,
$$(\mathbb{Q}, <) \vDash (\forall v_0)((\exists v_1)(r(v_0, v_1)))$$
because $\mathbb{Q}$ has no maximal element.

Definition 7.7. Two formulas $f(\overline{x})$ and $g(\overline{x})$ are **equivalent** if for any relation R and any $\overline{a} \in E$, we have

$$R \vDash f(\overline{a}) \quad \Leftrightarrow \quad R \vDash g(\overline{a}).$$

For instance, f should be equivalent to $\neg(\neg(f))$, and $(f) \wedge (g)$ should be equivalent to $(g) \wedge (f)$. We will use abbreviations like $f \to g$ to mean $(\neg(f)) \vee (g)$ or stuff.

Theorem 7.8 (Fraïssé)**.** *Let R and S be m-ary relations on E and E', and let $p < \omega$, $\overline{a} \in E^n$ and $\overline{b} \in (E')^n$. The following are equivalent:*

(1) $(R, \overline{a}) \sim_p (S, \overline{b})$

(2) for all formulas $f(x_1, \ldots, x_n)$ with $\mathrm{QR}(f) \leq p$, then $R \vDash f(\overline{a})$ if and only if $S \vDash f(\overline{b})$.

Proof. Let us first prove (1) $\Rightarrow$ (2). We prove it by induction on the complexity of f. Let s be the local isomorphism $s(a_i) = b_i$.

- If f is $x_i = x_j$, which has quantifier rank 0, assume $R \vDash f(\overline{a})$ then $a_i = a_j$. Then $S \vDash f(\overline{b})$. Likewise, we have the other direction.

- If f is $r(x_{i_1}, \ldots, x_{i_m})$, then this follows from s being a local isomorphism.
- If f is $(g) \wedge (h)$ or $(g) \vee (h)$ or $\neg(g)$, then this is clear.
- If f is $(\exists x)(g)$, with $g(x_1, \ldots, x_n, y)$. If $R \vDash f(\overline{a})$, then there exists a $b \in E$ such that $R \vDash g(\overline{a}, b)$. Because $\mathrm{QR}(f) = \mathrm{QR}(g) + 1$, if s is a $\mathrm{QR}(f)$-isomorphism then we can just do the back and forth on s.
- For $\forall x$, we note that $(\forall y)(g)$ is $\neg(\exists y)(\neg g)$.

We will do the other direction next time. $\square$

8 October 1, 2018

Last time we defined what formulas are. Also, we defined what p-isomorphisms are.

8.1 Fraïssé's theorem

Theorem 8.1 (Fraïssé)**.** *Let R and R' be m-ary relations on E and E'. Let $\overline{a} \in E^n$ and $\overline{a}' \in (E')^n$. Then the following are equivalent:*

(1) $(R, \overline{a}) \sim_p (R', \overline{a}')$

(2) For each formula $f(\overline{x})$ with $\mathrm{QR}(f) \leq p$, we have $R \vDash f(\overline{a})$ if and only if $R' \vDash f(\overline{a}')$.

Last time we showed (1) to (2). Today we show the other direction.

Lemma 8.2. *Fix $p, n < \omega$. Then $\sim_p$ has only a finite number $c(n,p)$ of classes, on the class of $(E, R, \overline{a})$.*

Proof. We prove by induction on p. For $p = 0$, these are zero equivalences. So we can only test at n^2 times n^m things. (We need to check if things are distinct or equal.) So $c(n, 0) \leq n^{m+2}$.

For the inductive step, observe that

$$(E, R, \overline{a}) \sim_{b+1} (E', R'\overline{a}')$$

is equivalent to that for any $b \in E$, there is a $b' \in E'$ such that $(E, R, \overline{a}, b) \sim_p (E', R', \overline{a}', b')$. So $\mathrm{iso}[(E, R, \overline{a})]_{p+1}$ is determined by $\{[E, R, \overline{a}, b]_p : b \in E\}$. This is a subset of the equivalence classes of $(k + 1)$-tuples. Therefore

$$c(n, p + 1) \leq 2^{c(n+1,p)}$$

is finite. □

Let us now prove (2) to (1).

Proof. Assume $(R, \overline{a})$ and $(R', \overline{a}')$ satisfy the same formulas of $\mathrm{QR} \leq p$. We want to show that $(R, \overline{a}) \sim_p (R', \overline{a}')$. What we will prove is that there is a formula that singles out a given equivalence class. That is, given $C = [(E, R, \overline{a})]_p$ we will show that there is a formula $f_C(\overline{x})$ with $\mathrm{QR}(f_C) \leq p$, such that

$$R' \vDash f_C(\overline{a}')$$

if and only if $(R, \overline{a}) \sim_p (R', \overline{a}')$. If we have this claim, it is clear that (2) implies (1).

We prove this by induction on p. If $p = 0$, there are only finitely many atomic formulas with variables $x_1, \dots, x_n$. Just let

$$f_C(x_1, \dots, x_n) = \bigwedge(\text{all atomic formulas (with negation) that } R \text{ satisfies}).$$

It is clear that this has quantifier rank 0. Now assume this is true for p. Let $f_1(\overline{x}, y), \ldots, f_k(\overline{x}, y)$ describing each p-equivalence classes, each of quantifier rank p. Then we let

$$f_C(x_1, \ldots, x_n) = \bigwedge((\exists y)(f_i(\overline{x}, y))) \wedge \bigwedge((\forall y)(\neg f_i(\overline{x}, y))).$$

according to the $(p+1)$-equivalence class we want to encode. $\square$

Corollary 8.3. *The following are equivalent:*

(1) $(R, \overline{a}) \sim_\omega (R', \overline{a}')$,

(2) $(R, \overline{a})$ *and* $(R', \overline{a}')$ *satisfy the same formulas.*

For instance, $(\mathbb{Z}, <)$ and $(\mathbb{Z}, <) + (\mathbb{Z}, <)$ satisfy the same formulas.

8.2 Models and theories

Definition 8.4. When f is a sentence (a closed formula) and $R \vDash f$, we say that R is a **model** of f. For A a set of sentences, we write $A \vDash f$ and say f is a **consequence** of A, if every model of A is a model of f, i.e.,

$$R \vDash A \quad \Rightarrow \quad R \vDash f.$$

We also write $f \vDash g$ if $\{f\} \vDash g$.

For example, $(\mathbb{Q}, <)$ is a model of

$$\{\forall x \exists y r(x, y), \quad \forall x \forall y \exists z (r(x, y) \to r(x, z) \wedge r(z, y))\}.$$

It is not a model of $\exists x r(x, x)$. We can also say things like

$$\{\forall x \exists y (r(x, y) \wedge x \neq y), \exists x (x = x)\} \vDash \exists x \exists y (x \neq y).$$

This is purely semantic. Two sentences f and g are equivalent if and only if $f \vDash g$ and $g \vDash f$. Also, we can write $f \vDash g$ also as $\emptyset \vDash (f \to g)$.

Definition 8.5. A set A of sentences is call **consistent** if there exists a model $R \vDash A$. We call A **inconsistent** if it is not consistent.

The set $A = \{\exists x (x \neq x)\}$ is inconsistent. Note that if A is inconsistent, then $A \vDash f$ for any f.

Definition 8.6. A **theory** E is a consistent set of sentences, closed under consequences, i.e., if $A \vDash f$ then $f \in A$.

Given a consistent set of formulas, we can close it into a theory, by

$$T_A = \{f : A \vDash f\}.$$

Definition 8.7. A theory T is said to be **complete** if for any sentence f, either $f \in T$ or $\neg f \in T$. A set A is **complete** if T_A.

Take $n = 2$. The set $A = \emptyset$ is consistence, but it is not complete, because both $A \cup \{\exists x(x = x)\}$ and $A \cup \{\neg\exists x(x = x)\}$ are consistent.

Proposition 8.8. *A consistent set of sentences is complete if and only if all its models are elementarily equivalent.*

So take A be the list of axioms, corresponding to the axiomitization of the theory of nonempty dense change without endpoints,

$$A = \{\exists x(x = x), \forall x(\neg r(x, x)), \forall x\forall y(r(x, y) \vee r(y, x) \vee x = y), \ldots\}.$$

All models are elementarily equivalent, so this theory is complete.

Here is another trivial example of a complete theory, for $m = 1$. Consider

$$\{\forall x \neg r(x), \exists x(x = x)\exists x_1\exists x_2(x_1 \neq x_2), \ldots, \exists x_1 \cdots \exists x_n \bigvee_{1\leq i<j\leq n}(x_i \neq x_j), \ldots\}.$$

Then this is complete.

Proposition 8.9. *If A is a finite set of axioms and A is complete, then there exists a program that takes as input a sentence f and outputs whether $A \vDash f$ or $A \vDash \neg f$.*

9 October 5, 2018

Recall that we were looking at formulas.

9.1 Elementary extensions

Definition 9.1. A relation R on E is a **restriction** of a relation R' on E' if $E \subseteq E'$ and for any $\overline{a}$ form E,

$$R(\overline{a}) \quad \Leftrightarrow \quad R'(\overline{a}),$$

that is, $R = R' \cap (E \times E)$. We are going to write $R = R'|_E$.

For instance, $(\mathbb{N}, <)$ is a restriction of $(\mathbb{Z}, <)$. However, this doesn't necessarily play nicely with formulas. The relations $(\mathbb{N}, <)$ and $(\mathbb{Z}, <)$ don't satisfy the same formulas, for instance,

$$(\mathbb{N}, <) \vDash \neg\forall x \exists y, r(y, x), \quad (\mathbb{Z}, <) \vDash \forall x \exists y, r(y, x).$$

Definition 9.2. Let R and R' be relations on E and E'. We say that R is an **elementary restriction** of R' (or that R' is an **elementary embedding** of R) if R is a restriction of R' and for any formula $f(\overline{x})$ and any tuple $\overline{a}$ from E, we have

$$R \vDash f(\overline{a}) \quad \Leftrightarrow \quad R' \vDash f(\overline{a}).$$

Then we write $(R, E) \preceq (R', E')$ or $R \preceq R'$.

Note that

$$(\mathbb{N}, <) \not\preceq (N \cup \{-1\}, <)$$

even though they are isomorphic, because "no x is strictly less than 0" does not evaluate to the same truth value. But we have things like

$$(\mathbb{Q}, <) \preceq (\mathbb{R}, <), \quad (\mathbb{Z}, <) \preceq (\mathbb{Z}, <) + (\mathbb{Z}, <).$$

Theorem 9.3 (Tarski's test)**.** *Assume (R, E) is a restriction of (R', E'). The following are equivalent:*

(1) $R \preceq R'$.

(2) For any formula $f(\overline{x}, y)$ and any tuple $\overline{a}$ from E, if $R' \vDash \exists y, f(\overline{a}, y)$ then there is $b \in E$ such that $R' \vDash f(\overline{a}, b)$.

So it suffices to check that any equation with coefficients in E having a solution in E', also has a solution in E.

Proof. For (1) implies (2), assume that $R \preceq R'$. Let $f(\overline{x}, y)$ be a formula, with $\overline{a}$ a tuple from E. Suppose that

$$R' \vDash \exists y f(\overline{a}, y).$$

Then R also satisfies the formula, so we get a solution in E.

For (2) implies (1), we induct this by induction on the complexity. For a given formula $f(\overline{x})$ and any $\overline{a}$ from E, we show that $R \vDash f(\overline{a})$ if and only if $R' \vDash f(\overline{a})$. If f is an atomic formula, this is clear because R is a restriction of R'. If f is $\neg g$ or $g \wedge h$ or $g \vee h$, this is just expanding the definition. So we can now think the case when f is $(\forall y)g(\overline{x}, y)$. If we assume that $R \vDash f(\overline{a})$, it also satisfies $R \vDash g(\overline{a}, b)$. By the induction hypothesis, $R' \vDash g(\overline{a}, b)$ and then we get $R' \vDash f(\overline{a})$. To do the converse direction, we use (2) and the induction hypothesis. □

So an embedding R into R' is an elementary embedding if and only if for any $\overline{a}$ from E, we have $(R, \overline{a}) \sim_\omega (R', \overline{a})$.

9.2 Löwenheim's theorem

Theorem 9.4 (Löwenheim's theorem)**.** *Any relation has a countable elementary restriction. In fact, if (E, R) is a relation and $A \subseteq E$ is countable, there is a $E_0 \subseteq E$ such that $A \subseteq E_0$ such that E_0 is countable and $R|_{E_0} \preceq R$.*

Corollary 9.5. *It is impossible to axiomatize "being uncountable". That is, any consistent set of axioms has a countable model.*

Basically, you just enlarge elements by adding solutions.

Proof. For a countable set B, let

$$F_B = \{(f(\overline{x}, y), \overline{a}) : \overline{a} \text{ is from } B \text{ and } f \text{ is a formula}\}.$$

Note that F_B is countable, because are only countably many formulas. Now fix $A \subseteq E$ countable, and define a sequence of countable sets $(A_n)_{n<\omega}$ by the following. First define

$$A_0 = A.$$

Then define A_{n+1} so that for any $(f(\overline{x}, y), \overline{a}) \in F_{A_n}$, if $R \vDash \exists y f(\overline{a}, y)$ then there exists a $b \in A_{n+1}$ such that $R \vDash f(\overline{a}, b)$. (Here, we should use some axiom of choice.) Now we take

$$E_0 = \bigcup_{n<\omega} A_n.$$

Any time you have a formula and a tuple, the tuple is in some large A_n. So we can find a solution in A_{n+1}. □

Here is a fun application, called **Skolem's paradox**. Consider the class of all sets, and consider the relation $\in$. By Löwenheim's theorem, there is a countable set $V_0 \subseteq V$ such that

$$(V_0, \in) \preceq (V, \in).$$

What will V_0 contain? Because we can write down the sentence $\exists x(\neg \exists y, r(y, x))$, and $\emptyset$ is the only set satisfying the equation. Then we see all the things we do

ωV_0 and $\mathbb{R} \in V_0$ and so on. But being countable can be encoded in set theory, so

$$V \models \text{“}\mathbb{R} \text{ is uncountable”}.$$

Then we also have

$$V_0 \models \text{“}\mathbb{R} \text{ is uncountable”}.$$

How can this be true if V_0 is countable and $\mathbb{R} \in V_0$? This is because even though $\mathbb{R} \in V_0$ we have $(\mathbb{R} \cap V_0 \notin V_0$. So it's happy with its own version of truth.

10 October 12, 2018

So far we have only looked at (E, R). Now we want to also look at things like

$$(\mathbb{R}, +, \cdot, 0, <).$$

We could think of $+$ as a ternary relation, but we want to consider it as a function, and we could think of 0 as a 0-ary function, but we will think of this as a special case.

10.1 Signatures

Definition 10.1. A **signature** (or a **similarity type**) is a set (possibly empty) containing

(1) constant symbols $c_0, c_1, c_2, \ldots$ with $(c_i)_{i<\lambda}$,

(2) function symbols $f_0, f_1, c_2, \ldots$ with $(f_i)_{i<\lambda'}$ with arities $n_0, n_1, n_2, \ldots$,

(3) relation symbols $r_0, r_1, r_2, \ldots$ with $(r_i)_{i<\lambda''}$ with arities $m_0, m_1, m_2, \ldots$.

For example, take

$$\sigma = \{f_0, f_1, r_0, c_0\}$$

with arities $n_0 = 2$, $n_1 = 2$, $m_0 = 2$, and we can abuse notation to consider it as $\{+, \cdot, <, 0\}$.

Definition 10.2. Give a signature σ, a σ**-structure** M is a set E, called the **universe** of M and denoted $E = \mathrm{univ}(M)$, and

(1) for each constant symbol $c_i \in \sigma$ an interpretation $c_i^M \in E$ for each c_i,

(2) for each function symbol $f_i \in \sigma$ a function $f_i^M : E^{n_i} \to E$,

(3) for each relation symbol r_i a subset $r_i^M \subseteq E^{m_i}$.

When we say

$$(\mathbb{R}, +, \cdot, 0, <),$$

we really mean the σ-structure on $\mathbb{R}$ where σ is the signature $\{+, \cdot, 0, <\}$. Now let us fix a signature σ. We want to define what formulas mean.

Definition 10.3. A **term** of complexity 0 is either

$$c_i \quad \text{or} \quad x_j.$$

A term of complexity $n + 1$ is a function applied to a bunch of terms,

$$f_i(t_1, t_2, \ldots, t_{n_i})$$

for f_i a function symbol of complexity of at most n, with one of them of complexity exactly n.

For example,

$$f_0(f_1(c_0, x_1), x_2)$$

is a term of complexity 2. We can also write this as $(0 \cdot x_1) + x_2$ in the example above.

Definition 10.4. A **formula** (in the language of σ) of complexity 0 is either

$$t_0 = t_1 \quad \text{or} \quad r_i(t_1, \ldots, t_{m_i})$$

for some relation r_i and terms t_i. The formulas of higher order are defined inductively in a similar way by writing things like $f \wedge g$ or $\neg g$ or $\exists x f$.

So

$$(\forall x)(\forall y)(f_0(f_1(c_0, x), y) = y)$$

is a formula. Let's now talk about semantics.

Definition 10.5. For M a σ-structure, t a term with n variables $(x_1, \ldots, x_n)$, and $\overline{a}$ an n-tuple in $\mathrm{univ}(M)$, we define

$$t^M(\overline{a})$$

by induction on the complexity of t:

(1) if t is c_i then $t^M(\overline{a}) = c_i^M$,

(2) if t is x_j then $t^M(\overline{a}) = a_j$,

(3) if t is $f_i(t_1, \ldots, t_{n_i})$, then

$$t^M(\overline{a}) = f_i^M(t_1^M(\overline{a}), t_2^M(\overline{a}), \ldots, t_{n_i}^M(\overline{a})).$$

Definition 10.6. For M a σ-structure and φ a formula with n free variables and $\overline{a}$ an n-tuple in $\mathrm{univ}(M)$, we define

$$M \vDash \varphi(\overline{a})$$

by induction on the complexity of φ:

(1) if f is atomic and is $t_1 = t_2$, then we say $M \vDash f(\overline{a})$ if and only if $t_1^M(\overline{a}) = t_2^M(\overline{a})$,

(2) if f is $r_i(t_1, \ldots, t_{m_i})$ then we say $M \vDash f(\overline{a})$ if and only if $(t_1^M(\overline{a}), \ldots, t_{m_i}^M(\overline{a})) \in r_i^M$,

(3) if f is not atomic, define $M \vDash f(\overline{a})$ exactly as before.

So $M = (\mathbb{R}; +, \cdot, 0, <)$ satisfies

$$M \vDash (\forall x \exists y)(f_0(x, y) = c_0).$$

Now we need to go to the notions we know and generalize them.

Definition 10.7. For M and N two σ-structures, we say that M is a **substructure** of N (written $M \subseteq N$) if

(1) $\mathrm{univ}(M) \subseteq \mathrm{univ}(N)$,
(2) $c_i^M = c_i^N$ for all constant symbols c_i,
(3) $f_i^M(\overline{a}) = f_i^N(\overline{a})$ for any i and n-tuple a of $\mathrm{univ}(M)$,
(4) $\overline{a} \in r_i^M$ if and only if $\overline{a} \in r_i^N$ for any i and n-tuple a of $\mathrm{univ}(M)$.

For example,

$$(\mathbb{Q}, +, \cdot, 0, <) \subseteq (\mathbb{R}, +, \cdot, 0, <).$$

Definition 10.8. We say that $M \subseteq N$ is an **elementary embedding** (and write $M \preceq N$) if

$$M \vDash \varphi(\overline{a}) \quad \Leftrightarrow \quad N \vDash \varphi(\overline{a})$$

for any formula φ and any $\overline{a}$ from $\mathrm{univ}(M)$.

We have seen that

$$(\mathbb{Q}, <) \preceq (\mathbb{R}, <),$$

but we have

$$(\mathbb{Q}, +, \cdot, 0, <) \not\preceq (\mathbb{R}, +, \cdot, 0, <)$$

because $\sqrt{2}$ does not exist in $\mathbb{Q}$.

Definition 10.9. For M and N two σ-structures, we say that M is **elementarily equivalent** to N if M and N satisfy the same sentences.

Löwenheim's theorem still goes through in this context.

Theorem 10.10 (Downward Löwenheim–Skolem theorem)**.** *Assume M is a σ-structure, and $A \subseteq \mathrm{univ}(M)$. Then there is a $M_0 \preceq \mathcal{M}$ with $A \subseteq \mathrm{univ}(M_0)$ and*

$$|\mathrm{univ}(M_0)| \leq |A| + |\sigma| + \aleph_0 = \max(|A|, |\sigma|, \aleph_0).$$

The proof is exactly the same. You close off A and iterate this.

Corollary 10.11. *If T is a theory in the language of σ, then it has a model of cardinality at most $|\sigma| + \aleph_0$.*

Example 10.12. There can be theories with no countable models. Consider $\sigma = (c_i)_{i<\aleph_1}$. Then the theory $T = \{\neg(c_i = c_j)\}$ will have at least $\aleph_1$ elements.

10.2 The upward Löwenheim–Skolem theorem

Next, we want to prove the upward Löwenheim–Skolem theorem.

Theorem 10.13 (upward Löwenheim–Skolem)**.** *If T is a theory and T has an infinite model, then for every $\lambda \geq |\sigma| + \aleph_0$, the theory T has a model with size λ.*

We will need a lot of tools to prove this. Let us take

$$M = (\mathbb{R}, +, \cdot, 0, <),$$

and take

$$T = \{\varphi : \varphi \text{ is a sentence}, M \vDash \varphi\}.$$

By the downward Löwenheim–Skolem theorem, we know that T has a countable model M_0. Then by the upward theorem, there has to be a model that is bigger, of size $2^{2^{\aleph}}$. We are going to prove this using the compactness theorem.

Theorem 10.14 (compactness theorem)**.** *If A is a set of sentences and every finite subset of A is consistent, then A is consistent.*

An application of this is non-standard analysis. We may consider

$$T_0 = \text{all sentences true in } (\mathbb{R}, +, \cdot, 0, 1)$$

and then look at

$$T = T_0 \cup \{1 < c, 1 + 1 < c, 1 + 1 + 1 < c, \ldots\}.$$

By compactness, this has to be consistent.

11 October 15, 2018

Last time we stated the compactness theorem.

Theorem 11.1 (compactness theorem)**.** *If A is a set of axioms and all finite subsets of A are consistent, then A is consistent.*

Using this, we can prove the upward Löwenheim–Skolem theorem.

Theorem 11.2 (upward Löwenheim–Skolem)**.** *If A is a set of axioms (in the language of σ) with an infinite model, then for every $\lambda \geq |\sigma| + \aleph_0$ there is a model $M \vDash A$ such that $|\mathrm{univ}(M)| = \lambda$.*

Proof. Assume that A is a set of axioms with an infinite model. Assume $\lambda \geq |\sigma| + \aleph_0$. Then we add a bunch of constants and look at

$$\sigma' = \sigma \cup \{d_i : i < \lambda\}$$

where d_i are constant symbols not in σ. Now set

$$A' = A \cup \{\neg(d_i = d_j) : i < j < \lambda\}.$$

Then every finite subset of A' has a model, so A' is consistent by compactness. Any model of A' is going to have cardinality at least λ, and so we get a model of A that has cardinality at least λ.

Now we pick any subset $A \subseteq \mathrm{univ}(N)$ with $|A| = \lambda$. Use the downward Löwenheim–Skolem to get a model $N_0 \preceq N$ with

$$\lambda \leq |\mathrm{univ}(N_0)| \leq |A| + |\sigma'| + \aleph_0 = \lambda + \lambda + \aleph_0 = \lambda.$$

So we get a model of A of size λ. □

11.1 Ultrafilters

Here is the idea of the proof of compactness. We start with A, and we know for every finite $A_0 \subseteq A$ there is a model

$$M_{A_0} \vDash A_0.$$

We somehow need a way to combine all the M_{A_0} into a model of A. So we make the universe to be something like

$$\prod_{A_0} \mathrm{univ}(M_{A_0}),$$

and then the relations will be, make the individual components vote for whether it is true. Then we need a notion of what it means for the majority to think if it is true or false.

Definition 11.3. A **filter** for a set I is a set of subsets of I satisfying

(1) $\phi \notin F$ and $I \in F$,
(2) if $A \in F$ and $A \subseteq B \subseteq I$ then $B \in F$,
(3) if $A, B \in F$ then $A \cap B \in F$.

For a fixed $A \subseteq I$, the subset

$$F_A = \{X \subseteq I : A \subseteq X\}$$

is a filter. A more interesting example is the **Fréchet filter**

$$F = \{X \subseteq I : |I - X| < \aleph_0\}$$

of cofinite sets. If I is finite, any filter looks like F_A for some A.

Definition 11.4. An **ultrafilter** on I is a filter U on I such that for any $A \subseteq I$, either $A \in U$ or $A^c \in U$.

Things of the form $F_{\{a\}}$ are ultrafilters, in fact, these are called **principal ultrafilters**. But are the non-principal ultrafilters? If I is finite, all ultrafilters are finite, but if I is infinite, this is not true.

Theorem 11.5. *For any filter F on I, there is an ultrafilter U on I such that $F \subseteq U$.*

For instance, you can take the Fréchet filter and extend it to an ultrafilter. Then this will be non-principal because it won't contain any finite set.

Proof. First we note that a filter is an ultrafilter if and only if it is maximal. This is because if a filter F is not an ultrafilter then we can just add a set A with $A, A^c \notin F$ and look at the filter generated by A and F (meaning the filter consisting of the sets containing $A \cap B$ for some $B \in F$). You can check that this is a well-defined filter that is strictly larger than F.

Now we use Zorn's lemma on the set of filters. Build $(F_\alpha)_{\alpha \in \mathrm{OR}}$ by transfinite induction by

- $F_0 = F$,
- if F_α is maximal, $F_{\alpha+1} = F_\alpha$, otherwise add one element,
- $F_\alpha = \bigcup_{\beta<\alpha} F_\beta$ for α a limit.

Then because $\mathcal{P}(I)$ is a set, we have $F_\alpha = F_{\alpha+1}$ for some α. This shows that F_α is maximal and contains F. ☐

12 October 19, 2018

Last time we defined this notion of a filter, and then there are enough ultrafilters. To be precise, we showed that every filter extends to an ultrafilter. Using this, we will construct models.

12.1 Ultraproducts

Here is the idea. Given σ-structures $(M_i)_{i\in I}$ and given an ultrafilter U on I, we want to define a σ-structure on $\prod_{i\in I} \operatorname{univ}(M_i)$. We are going to interpret everything according to what "most inI" think.

Definition 12.1. For $(a_i)_{i\in I}$ and $(b_i)_{i\in I}$ in the product $\prod_{i\in I} \operatorname{univ}(M_i)$, we define $(a_i) \sim (b_i)$ if

$$\{i \in I : a_i = b_i\} \in U.$$

(This of course depends on what the ultrafilter U is.)

This is an equivalence relation because U is an ultrafilter, so we can quotient by it.

Definition 12.2. Assume U is an ultrafilter on a set I, and let $(M_i)_{i\in I}$ be nonempty σ-structures. Then we define the **ultraproduct** of (M_i)

$$\prod_{i\in I} M_i/U$$

to be the following σ-structure on N:

- the universe of N is $\prod_{i\in I} \operatorname{univ}(M_i)/\sim$,
- for each constant symbol c, we define $c^N = [(c^{M_i})_{i\in I}]$,
- for each function symbol f of arity n, we define

$$f^N([(a_i^1)], \dots, [(a_i^n)]) = [(f^{M_i}(a_i^1, \dots, a_i^n))],$$

- for each relation symbol r of arity m, we define $([(a_i^1)_{i\in I}], \dots, [(a_i^m)_{i\in I}]) \in r^N$ if and only if

$$\{i \in I : (a_i^1, \dots, a_i^m) \in r^{M_i}\} \in U.$$

You can check this is well-defined.

Theorem 12.3 (Łoś's theorem)**.** *Assume U is an ultrafilter on a set I, let $(M_i)_{i\in I}$ be nonempty σ-structure, and let $\varphi(x_1, \dots, x_n)$ be a formula in the language σ. Let $(a_i^1)_{i\in I}, \dots, (a_i^n)_{i\in I}$ be in $\prod_{i\in I} \operatorname{univ}(M_i)$. Then*

$$\prod_{i\in I} M_i/U \vDash \varphi([(a_i^1)_{i\in I}], \dots, [(a_i^n)_{i\in I}])$$

if and only if $\{i \in I : M_i \vDash (a_i^1, \dots, a_i^n)\} \in U$.

Proof. This is true by definition if φ is an atomic formula. On the other hand, when we combine formulas, if we look at ψ is like $\neg\psi$ or $\psi\wedge\chi$ then this is doable.

Suppose φ is $(\exists y)(\psi(y,\ldots))$. If $\prod_{i\in I} M_i/U \vDash \varphi([(a_i^1)],\ldots,[(a_i^n)])$, then we can pick $[(b_i)]$ such that

$$\prod_i M_i/U \vDash \psi([(b_i)],[(a_i^1)],\ldots,[(a_i^n)]).$$

Then by the induction hypothesis, the set of $i\in I$ with $M_i \vDash (\exists y)\psi(y,a_1^i,\ldots,a_n^i)$ is in U. Conversely, if this is true, we can pick for most i a witness $b_i \in \operatorname{univ}(M_i)$ such that $M_i \vDash \psi(b_i,a_i^1,\ldots,a_i^n)$. Then the ultraproduct satisfies $\psi([(b_i)],[(a_i^1)],\ldots,[(a_i^n)])$, and hence satisfies φ. □

12.2 Proof of the compactness theorem

Recall the compactness theorem.

Theorem 12.4 (compactness thoerem)**.** *Assume A is a set of axioms. If all finite subsets of A are consistent, then A is consistent.*

Proof. Let I to be the set of all finite subsets $A_0 \subseteq A$. Then we take the filter on I generated by the sets

$$\langle A_0\rangle = \{B_0 \in I : A_0 \subseteq B_0\},$$

which can be seen to be a filter because $\langle A_0\rangle \cap \langle A_1\rangle = \langle A_0 \cup A_1\rangle$. We can then extend this filter to an ultrafilter U on I. This is what we are going to use.

Let us now look at

$$N = \prod_{A_0\in I} M_{A_0}/U.$$

We need to show that $N \vDash A$, that N satisfies each sentence separately. If φ is a sentence of A, consider the finite subset $A_0 = \{\varphi\}$. Note that by definition, if $\varphi \in B_0$ so that $A_0 \subseteq B_0$, then $M_{B_0} \vDash \varphi$. This means that $\langle A_0\rangle \subseteq \{B_0 \in I : M_{B_0} \vDash \varphi\}$, and this implies that this set is in the ultrafilter U. So by Łoś's theorem, $N \vDash \varphi$. □

Here is an equivalent form of the compactness theorem. The contrapositive would be, if a set of axioms A is inconsistent, then there is a finite subset $A_0 \subseteq A$ that is inconsistent. If we think of "inconsistency" as a "proof of a contradiction", then this says that a proof of a contradiction from an infinite set of axioms uses only a finite set of these axioms. Actually, proofs are finite, so it only should involve finitely many axioms. This strategy works, and we will see how this works next time.

13 October 22, 2018

Today we will define proofs formally. If we have this notion and if we can prove this theorem

$$A \vdash \varphi \quad \Leftrightarrow \quad A \vDash \varphi,$$

(called the completeness theorem), then we immediately get the compactness theorem. We will take a definition where it will be easy to prove things about proofs, but hard to actually prove things, for instance, using a computer.

13.1 Proof of a sentence

Let us fix σ a signature for today, and talk about everything in the language of σ. Intuitively, a propositional tautology is something like $\varphi \vee \neg\varphi$ or $\varphi \leftarrow \varphi$, things that are true in any possible model.

Definition 13.1. A **basic formula** is a formula that is not of the form $\psi \wedge \varphi$ or $\psi \vee \varphi$ or $\neg\psi$.

For example, formulas like $(\forall x)(\exists y)(r(x,y) \wedge \neg r(y,x))$ are basic.

Definition 13.2. A **truth assignment** is a function v for the set of basic formulas to $\{0,1\}$ (0 is false and 1 is true).

Given a truth assignment v, we can lift it to an assignment $\bar{v}$ on all formulas, just by inducting on the complexity of φ.

Definition 13.3. A **proposition tautology** is a formula φ such that $\bar{v}(\varphi) = 1$ for any truth assignment v.

The formula $(\forall x)(x = x)$ is not a propositional tautology even if it is true in any model. (You can't prove it only using propositional logic.) To get a propositional tautology, you need to something like $(\forall x)(x = x) \vee \neg(\forall x)(x = x)$. In general, $\varphi \vee \neg\varphi$ is a propositional tautology for any φ.

Definition 13.4. A **universal closure** of a formula φ is a sentence of the form $\forall x_0 \forall x_1 \cdots \forall x_n \varphi$ for some $n < \omega$.

Both $(\forall x_1)(x_1 = x_1)$ and $(\forall x_1 \forall x_2)(x_1 = x_1)$ are universal closures of $x_1 = x_1$.

Definition 13.5. A **logical axiom** is a universal closure of a formula of the following type:

(1) propositional tautologies,

(2) $(\forall x)(\varphi \to \psi) \to ((\forall x)\varphi \to (\forall x)\psi)$,

(3) $\varphi \to \forall x \varphi$ if x is not a free variable of φ,

(4) $(\forall x \varphi(x)) \to \varphi(t)$ for some term t, if x is a free variable of φ,

(5) $\varphi(t) \to (\exists x)(\varphi(x))$ if x is a free variable of φ,

(6) $\forall x(\neg\varphi) \leftrightarrow \neg\exists x\varphi$,

(7) $x = x$,

(8) $x = y \to y = x$,

(9) $(x = y \wedge y = z) \to x = z$,

(10) $((x_1 = y_1) \wedge \cdots \wedge (x_n = y_n)) \to (f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n))$ if f is a function symbol of arity n,

(11) $((x_1 = y_1) \wedge \cdots \wedge (x_n = y_n)) \to (r(x_1, \ldots, x_n) \leftrightarrow r(y_1, \ldots, y_n))$ if r is a relation symbol of arity n.

You can check that if φ is a logical axiom, then $M \vDash \varphi$ for any σ-structure M. (We also write this as $\vDash \varphi$.)

Definition 13.6. Assume A is a set of sentences, and let φ be a sentence. A (formal) **proof** of φ from A is a sequence $\varphi_1, \ldots, \varphi_n$ such that

(1) $\varphi_n = \varphi$,

(2) for each $i \leq n$, the sentence φ_i is in A or a logical axiom, or φ_i can be obtained from φ_j and φ_k by **modus ponens**, that is, φ_k is $\varphi_j \to \varphi_i$.

We write $A \vdash \varphi$ if there is a formal proof of φ from A.

It is really hard to give examples, but here is one example. We can show that

$$\{\varphi \wedge \psi\} \vdash \varphi.$$

This is clear semantically, and here is the formal proof:

- $\varphi_1 : \varphi \wedge \psi$ (this is in A)
- $\varphi_2 : (\varphi \wedge \psi) \to \varphi$ (this is a propositional tautology)
- $\varphi_3 : \varphi$ (this is obtained by modus ponens)

It is pretty challenging to prove anything. For instance, try to prove $\neg\forall x\varphi \leftrightarrow \exists x\neg\varphi$. Also, it is pretty clear that if $A \vDash \varphi$ then there is a finite $A_0 \subseteq A$ such that $A_0 \vdash \varphi$.

13.2 Formal properties of proofs

Lemma 13.7. *If $A \vdash \varphi$, then $A \vDash \varphi$.*

Proof. Assume $\varphi_1, \ldots, \varphi_n$ is a formal proof of φ from A. Then we prove $A \vDash \varphi_i$ by induction on i. If φ_i is in A, then this is clear. If $\varphi - i$ is a logical axiom, then this can be checked individually. If φ_i is obtained by φ_j and φ_k from modus ponens, then we know modus ponens in the real world, so we get $A \vDash \varphi_i$. $\square$

Lemma 13.8 (transitivity)**.** *If $A \vdash \psi_i$ for all $i \leq m$ and $A \cup \{\psi_1, \ldots, \psi_m\} \vdash \varphi$, then $A \vdash \varphi$.*

Proof. We just concatenate the proofs. Just list the proofs of ψ_i, and then put the proof of φ from $A \cup \{\psi_1, \dots, \psi_m\}$. $\square$

Theorem 13.9 (deduction theorem)**.** *Assume that A is a set of sentences, and assume that φ and ψ are sentences. Then $A \cup \{\varphi\} \vdash \psi$ if and only if $A \vdash (\varphi \to \psi)$.*

This is silly, but it is annoying to prove.

Proof. First suppose that $A \vdash (\varphi \to \psi)$. If we have $A \cup \{\varphi\}$ as our axiom, we can just write φ and use modus ponens.

For the other direction, assume that $A \cup \{\varphi\} \vdash \psi$. Consider a formal proof $\psi_1, \dots, \psi_n$ be a formal proof. We are going to prove by induction on $i \leq n$, that $A \vdash (\varphi \to \psi_i)$. If ψ_i is in $A \cup \{\varphi\}$ of ψ_i is a logical axiom, this is just a propositional tautology $\psi_i \to (\varphi \to \psi_i)$ used with modus ponens. If ψ_i is obtained by modus ponens from ψ_j and $\psi_k = \psi_j \to \psi_i$, we can do a similar thing. We can do

- $\varphi \to \psi_j$ (induction hypothesis)
- $\varphi \to (\psi_j \to \psi_i)$ (induction hypothesis)
- $(\varphi \to \psi_j) \to (\varphi \to (\psi_j \to \psi_i)) \to (\varphi \to \psi_i)$ (propositional tautology)
- $(\varphi \to (\psi_j \to \psi_i)) \to (\psi \to \psi_i)$ (modus ponens)
- $\varphi \to \psi_i$ (modus ponens)

and get the proof for $\varphi \to \psi_i$. $\square$

Definition 13.10. We say that A is **syntactically inconsistent** if there exists a φ such that $A \vdash \varphi$ and $A \vdash \neg\varphi$.

Lemma 13.11. *The following are equivalent:*

(1) $A \vdash \neg(\forall x(x = x))$,

(2) A is syntactically inconsistent,

(3) $A \vdash \psi$ for any sentence ψ.

Proof. (3) implies (1) is easy, and also (1) implies (2) is easy because we always have $A \vdash (\forall x)(x = x)$. For (2) implies (3), we use that $\varphi \to (\neg\varphi \to \psi)$ is a propositional tautology, and then use modus ponens. $\square$

Lemma 13.12 (proof by contradiction)**.** *For a sentence φ, we have $A \vdash \varphi$ if and only if $A \cup \{\neg\varphi\}$ is syntactically inconsistent.*

Proof. For the forward direction, if $A \vdash \varphi$, then $A \cup \{\neg\varphi\} \vdash \varphi$ and $A \cup \{\neg\varphi\} \vdash \neg\varphi$. So $A \cup \{\neg\varphi\}$ is syntactically inconsistent. For the other direction, suppose $A \cup \{\neg\varphi\}$ is inconsistent. Then $A \cup \{\neg\varphi\} \vdash \varphi$ so $A \vdash (\neg\varphi \to \varphi)$. But $(\neg\varphi \to \varphi) \to \varphi$ is a propositional tautology, so $A \vdash \varphi$. $\square$

14 October 26, 2018

Last time we defined provability and then proved some basic properties, like deduction or proof by contradiction. Here is one other property we need.

Proposition 14.1 (elimination and introduction of quantifiers). *Assume that A is a set of sentences in the language of σ, and let $\varphi(x)$ be a formula in the language of σ with free variable x. Define ψ a sentence in the language of $\sigma' = \sigma \cup \{c\}$, where c is a constant symbol.*

- *Introduction rule for $\exists$: If t is a σ-term and $A \vdash_\sigma \varphi(t)$, then $A \vdash_\sigma \exists x \varphi(x)$.*
- *Introduction rule for $\forall$: If $A \vdash_{\sigma'} \varphi(c)$, then $A \vdash_\sigma \forall x \varphi(x)$.*
- *Elimination rule for $\exists$: If $A \vdash_\sigma \exists x \varphi(x)$ and $A \cup \{\varphi(c)\} \vdash_{\sigma'} \psi$ then $A \vdash_\sigma \psi$.*
- *Elimination rule for $\forall$: If $A \vdash_\sigma \forall x \varphi(x)$ then $A \vdash_\sigma \varphi(t)$ for any σ-term t.*

Proof. The elimination rule for $\forall$ is just the logical axiom (4) with modus ponens. Similarly, the introduction rule for $\exists$ is the logical axiom (5). For the other two, we first claim that the introduction rule for $\forall$ implies the elimination rule for $\exists$. By transitivity, it suffices to show that $A \cup \{\exists x \varphi(x)\} \vdash_\sigma \psi$, where we know that $A \cup \{\varphi(c)\} \vdash_{\sigma'} \psi$. We use proof by contradiction, this is equivalent to showing that $A \cup \{\neg\psi\} \vdash_{s'} \neg\varphi(c)$. By the introduction rule for $\forall$, we have $A \cup \{\neg\psi\} \vdash \forall x \neg\varphi(x)$. By the logical axiom (6) and modus ponens, we get $A \cup \{\neg\psi\} \vdash \neg\exists x \varphi(x)$. So again by proof by contradiction, $A \cup \{\exists x \varphi(x)\} \vdash \psi$ as desired.

Now let us prove the introduction rule for $\forall$. Assume that $A \vdash_{\sigma'} \varphi(c)$. Fix a proof of $\varphi_1, \dots, \varphi_n$. Let y be a variable not appearing in $\varphi_1, \dots, \varphi_n$. We now prove by induction that

$$A \vdash \forall y \varphi_i(y),$$

where $\varphi_i(y)$ denotes φ_i with c replaced by y everywhere. If φ_i is in A, then φ_i does not contain the symbol c so this follows from the logical axiom (3) and modus ponens. If φ_i is a logical axiom in σ', then you can go through all of them and check this. If φ_i follows from φ_k and φ_j by modus ponens, $\varphi_k = (\varphi_j \to \varphi_i)$, then we can build a proof by

- $\forall y \varphi_j(y)$ (by induction hypothesis)
- $\forall y(\varphi_j(y) \to \varphi_i(y))$ (also by induction hypothesis)
- $(\forall y(\varphi_j(y) \to \varphi_i(y))) \to (\forall y \varphi_j(y) \to \forall y \varphi_i(y))$ (logical axiom (2))
- $\forall y \varphi_j(y) \to \forall y \varphi_i(y)$ (modus ponens)
- $\forall y \varphi_i(y)$ (modus ponens)

and this completes the proof. $\square$

14.1 The completeness theorem—eliminating quantifiers

We are going to prove this in a different way.

Theorem 14.2 (Model existence theorem)**.** *Any syntactically consistent set of sentence has a model.*

Using this, we immediately obtain the proof of the completeness theorem.

Proof of the completeness theorem. We only need to show that $A \vDash \varphi$ implies $A \vdash \varphi$. We are going to show the contrapositive. If we assume $A \nvdash \varphi$, then $A \cup \{\neg\varphi\} \nvdash \varphi$, by "proof by contradiction". By the model existence theorem, there is a model $M \vDash A \cup \{\neg\varphi\}$. Then $M \vDash A$ but $M \nvDash \varphi$, so $A \nvDash \varphi$. □

Here is the rough idea of the proof. Let us assume that we can eliminate all quantifiers, so A has no quantifiers. This means that we have a bunch of constants and functions and relations. For example, take $\sigma = \{+, \cdot, c_0, c_1, c_2\}$ and $A = \{c_2 = c_1 + c_1\}$. Then we can make the universe of M be just the closed terms (like $c_0, c_1, c_2, c_0 + c_1, c_1 \cdot c_2 + c_0$) and then quotient out by the provably equal terms. (We can just define $t \sim s$ if $A \vdash t = s$.) Then the universe will consist of equivalence classes of terms.

So there are two things to do: add witnesses so that formulas are without quantifiers, and show that functions are well-defined on equivalence classes.

Definition 14.3. Let A be a set of sentences. We say that A has **witnessing terms** if for any formula $\varphi(x)$ with $(\exists x)\varphi(x) \in A$, there is a closed term such that $\varphi(t) \in A$.

Lemma 14.4. *If A is a syntactically consistent set of sentences, and $(\exists x)\varphi(x) \in A$, then $A \cup \{\varphi(c)\}$ is syntactically consistent.*

Proof. This is just the introduction rule for $\forall$, along with proof by contradiction. □

Now we start with A_0, and pick $\forall x \varphi(x) \in A_0$. By this lemma, we can introduce a constant and $A_0 \cup \{\varphi(c_0)\}$ is syntactically consistent. We repeat this with A_1, if $\exists x \psi(x) \in A$ then set $A_2 = A_1 \cup \{\varphi(c_1)\}$. We repeat this, until we get $A_\omega = \bigcup_{n\in\omega} A_n$.

Lemma 14.5. *If α is a limit ordinal with $(A_i)_{i<\alpha}$ an increasing sequence of syntactically consistent sets, then $A = \bigcup_{i<\alpha} A_i$ is syntactically consistent.*

Proof. A proof of a contradiction only uses a finite number of axioms. So if A proves a contradiction, some A_i should prove a contradiction. □

So we obtain the following fact.

Proposition 14.6. *If A is synactically consistent, there is a $B \supseteq A$ (with maybe more constant symbols) that is syntactically consistent, such that for any existential formula $(\exists x)\varphi(x) \in A$ there is a term t such that $\varphi(t) \in B$.*

Proof. Keep adding witnesses, and at limits take unions. □

Definition 14.7. Let A be a syntactically consistent set A. We say that A is **maximal** if for φ, either $\varphi \in A$ or $\neg\varphi \in A$ for any sentence φ in the language of σ.

If A is a syntactically consistent set of sentence,s then $A \cup \{\varphi\}$ or $A \cup \{\neg\varphi\}$ is syntactically consistent.

Proposition 14.8. *If A is a syntactically consistent set of sentences in the language of σ, there is a $B \supseteq A$ syntactically consistent and maximal in σ.*

Theorem 14.9. *If A is syntactically consistent, then there is $B \supseteq A$ (in an expanded language) such that B is syntactically consistent, maximal, and has witnessing terms.*

Proof. Take $A_0 = A$, and given A_{2n} we build A_{2n+1} syntactically consistent so that it has witnessing terms. Given A_{2n+1} we build A_{2n+2} syntactically consistent and maximal. Then $B = \bigcup_{n<\omega} A_n$ is as desired. □

15 October 29, 2018

Last time we showed that A can be extended to a syntactically consistent $B \supseteq A$ that is maximal (either $\varphi \in B$ or $\neg\varphi \in B$) and has witnessing terms ($\exists x \varphi(x)$ in B implies that there is a term $\varphi(t) \in B$).

15.1 The completeness theorem—building the model

Note that by maximality, we have $A \vdash \varphi$ if and only if $\varphi \in A$. If $\neg\exists x(x = x) \in A$, then the empty model is a model for A so we don't have to worry about anything. So assume that $\exists x(x = x) \in A$. Since A has witnessing terms, there is a closed term t such that $(t = t) \in A$.

Let X be the set of all closed terms. Define a relation $\sim$ on X by $t \sim s$ if

$$A \vdash (t = s).$$

This is indeed an equivalence relation. Then we define a σ-structre on M as follows:

- The universe of M is $X/\sim$.
- For a constant symbol c, we define $c^M = [c]$.
- For a function symbol f of arity t, we define

$$f^M([t_1], \dots, [t_n]) = [f(t_1, \dots, t_n)].$$

- For a relation symbol r of arity r, we define

$$([t_1], \dots, [t_n]) \in r^M \quad \Leftrightarrow \quad r(t_1, \dots, t_n) \in A.$$

You can check that these functions and relations are well-defined. You can also check that if $\tau(x_1, \dots, x_n)$ are terms, then $\tau^M([t_1], \dots, [t_n]) = [\tau(t_1, \dots, t_n)]$ using induction on τ.

Proposition 15.1. *For any formula $\varphi(x_1, \dots, x_n)$ and any closed terms $t_1, \dots, t_n$, we have*

$$\varphi(t_1, \dots, t_n) \in A \quad \Leftrightarrow \quad M \vDash \varphi([t_1], \dots, [t_n]).$$

In particular $M \vDash A$.

Proof. If φ is atomic of the form $r(x_1, \dots, x_n)$, this is just true by construction. If φ is atomic of the form $\tau_1(x_1, \dots, x_n) = \tau_2(x_1, \dots, x_n)$, then we can show this using this property of $\tau^M([t_1], \dots, [t_n]) = [\tau(t_1, \dots, t_n)]$.

For the inductive step, let us just assume that $\vee$ and $\forall$ do not appear in φ. Also for simplicity, let us just assume that φ is a sentence. If φ is $\psi_1 \wedge \psi_2$, then $\varphi \in A$ if and only if $A \vdash (\psi_1 \wedge \psi_2)$ (by maximality) and this is equivalent to $A \vdash \psi_1$ and $A \vdash \psi_2$. (You can show this.) Then by the inductive hypothesis, this is equivalent to $M \vDash \psi_1$ and $M \vDash \psi_2$, and this is just $M \vDash \varphi$.

If φ is $\neg\psi$, then we see that $\varphi \in A$ is equivalent to $\psi \notin A$ (by maximality) and then this is equivalent to $M \not\vDash \psi$, which is $M \vDash \varphi$.

Finally, suppose φ is $\exists x \psi(x)$. If $\varphi \in A$, then since A has witnessing terms, there is a closed term t such that $\psi(t) \in A$. By the induction hypothesis, $M \vDash \psi([t])$, and so by the definition of $\vDash$, we see that $M \vDash \varphi$. Conversely, if $M \vDash \varphi$, then $M \vdash \psi([t])$ for some t, and by the induction hypothesis, we get $\psi(t) \in A$. So $A \vdash \psi(t)$ and the introduction rule for $\exists$ tells us that $A \vdash \exists x \psi(x)$. $\square$

So this completes the proof of the completeness theorem.

15.2 Decidability

I wanted to talk about some application of this to computer science. Let us fix a countable signature σ, codable in some way as a string of 0 and 1.

Definition 15.2. A set of formulas A is **decidable** if there is a computer program that takes as input a formula φ, and outputs yes or no depending on whether $\varphi \in A$ or not.

Note that there is no requirement about how fast this program has to be.

Example 15.3. Any finite set is decidable, because we can just hard-code the formulas into the program itself. Any set that can be written "explicitly" is decidable. For instance,

$$A = \{\exists x_1 \cdots \exists x_n (\textstyle\bigwedge_{1 \le i < j \le n} (x_i \neq x_j)) : n < \omega\}$$

is decidable. So the set of all logical axioms is decidable. For propositional tautologies, we have to check all possible assignments, but we can still code this.

On the other hand, there are sets that are not decidable. The set of all computer programs is countable, but the set of all sets of formulas is uncountable. This means that most sets are not decidable.

Definition 15.4. For a sentence φ and a set A of sentences, recall that we have defined $A \vDash \varphi$ if any model of A is a mode of φ. Then we define the **theory** of A as

$$\mathrm{Th}(A) = \{\varphi : A \vDash \varphi\}.$$

Similarly, for a model M, we define

$$\mathrm{Th}(M) = \{\varphi : M \vDash \varphi\}.$$

We say that T is a **theory** if it is a consistence set of sentences such that $T \vDash \varphi$ implies $\varphi \in T$. A theory is called **complete** if $\varphi \in T$ or $\neg\varphi \in T$, and a consistent set A is called **complete** if $\mathrm{Th}(A)$ is complete.

Note that $A \subseteq \mathrm{Th}(A)$ always, and $\mathrm{Th}(A)$ will be a theory if A is consistent. If $M \vDash A$, then $\mathrm{Th}(A) \subseteq \mathrm{Th}(M)$, and because $\mathrm{Th}(M)$ is always complete, we see that A is complete if and only if $\mathrm{Th}(A) = \mathrm{Th}(M)$.

Theorem 15.5. *If A is a (consistent) decidable and complete set, then* $\mathrm{Th}(A)$ *is decidable.*

Proof. By the completeness theorem, we can check that $A \vdash \varphi$ instead. Given an input, we can enumerate $(\bar{\psi}^n)_{n<\omega}$ all the sequences of sentences. Then for each $\bar{\psi}^n$, we can check whether $\bar{\psi}^n$ is a proof of φ or a proof of $\neg\varphi$. This is possible because both A and the set of logical axioms are decidable. This always terminates because A is a complete set. $\square$

Example 15.6. The axioms for nonempty dense chains without endpoints is finite, so decidable. Also, we showed that this is a complete set. Therefore $\mathrm{Th}(A)$ is decidable. Another algorithm for doing this is trying to eliminate quantifiers. So we can say "the theory of nonempty dense chains without endpoints is decidable".

The theory of infinite sets is decidable, and also the theory of generic graphs is decidable. We can also show that

$$\mathrm{Th}((\mathbb{C}, +, \cdot, 0, 1))$$

is decidable. We will see that this is a complete (decidable) axiomatization of the "theory of algebraically closed fields of characteristic zero". Similarly,

$$\mathrm{Th}((\mathbb{R}, +, \cdot, 0, 1))$$

is the theory of real closed fields, which is complete and decidable. The theory of discrete chains without nonempty discrete chains without endpoints,

$$\mathrm{Th}((\mathbb{Z}, <)),$$

is complete and decidable. But what about $(\mathbb{N}, +, \cdot, 0, 1)$? There are lots of unsolved problems in theory, for instance, Goldbach's conjecture.

Index