DEDUCTION

CALEB STANFORD

1. Natural Deduction Overview

In what follows we present a system of *natural deduction*. For a set of formulas Σ and a formula φ, we will define what it means for $\Sigma \vdash \varphi$. (Note that we are using formulas instead of sentences; the need for this will become clear when we look at deduction rules for quantifiers.)

Recall the definition of formula from before:

(1) If t_1 and t_2 are terms, then $t_1 = t_2$ is a formula.
(2) If $t_1, \ldots, t_n$ are terms, and R is an n-ary relation, then $R(t_1, \ldots, t_n)$ is a formula.
(3) If φ is a formula, then $(\neg\varphi)$ is a formula.
(4) If φ, ψ are formulas, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$ are formulas.
(5) If v_i is a variable, and φ a formula, then $\exists v_i(\varphi)$ and $\forall v_i(\varphi)$ are formulas.

We add to this one more thing:

(0) $\top$ and $\bot$ are formulas.[1]

Our goal is to say that $\Sigma \vdash \varphi$ if there is a finite sequence of steps that proves φ under the premises Σ using a number of allowed syntactic manipulations. For each logical symbol, we will have an *introduction rule* and an *elimination rule*. A formal proof basically involves constructing formulas and deconstructing formulas (respectively) using these rules.

There are many ways to define a system of deduction, and even many things that might be called "natural deduction", of which we will only go over one. It ultimately does not matter which system is used.

2. Silly Rules

Premise and Reiteration, while simple, are quite important:

(P) For any φ, $\{\varphi\} \vdash \varphi$.

(R) If $\Sigma \vdash \varphi$, and $\Sigma \subset \Gamma$, then $\Gamma \vdash \varphi$.

Then there is:

($\top$I) For any Γ, $\Gamma \vdash \top$.

($\bot$I) If $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$, then $\Gamma \vdash \bot$.

($\bot$E) For any Γ and φ, if $\Gamma \vdash \bot$ then $\Gamma \vdash \varphi$.

Date: 11 February 2015.

[1] This addition is not strictly necessary. For instance, the notion "$\Sigma \vdash \bot$" (Σ proves false, i.e. Σ is inconsistent) could be expressed instead by "$\Sigma \vdash (\varphi \wedge \neg\varphi)$ for some φ."

We can't deduce anything from $\top$, so we don't have an elimination rule for it. The $\bot$ rules are tied up with the rules for $\neg$ which we will see in a second. Instead of worrying about this sort of relationship, we will just err on the side of having too many rules.

3. Logical Connectives

The basic syntactic rules for $\wedge, \rightarrow, \vee, \neg$ are just as they were for truth-functional logic.

($\wedge$I) If $\Gamma \vdash \varphi$ and $\Gamma \vdash \psi$ then $\Gamma \vdash (\varphi \wedge \psi)$.

($\wedge$E) If $\Gamma \vdash (\varphi \wedge \psi)$ then $\Gamma \vdash \varphi$ and $\Gamma \vdash \psi$.

($\rightarrow$I) If $\Gamma \cup \{\varphi\} \vdash \psi$ then $\Gamma \vdash (\varphi \rightarrow \psi)$.

($\rightarrow$E) If $\Gamma \vdash (\varphi \rightarrow \psi)$ and $\Gamma \vdash \varphi$ then $\Gamma \vdash \psi$.

($\vee$I) If $\Gamma \vdash \varphi$ or if $\Gamma \vdash \psi$ then $\Gamma \vdash (\varphi \vee \psi)$.

($\vee$E) If $\Gamma \vdash (\varphi \vee \psi)$, $\Gamma \cup \{\varphi\} \vdash \theta$ and $\Gamma \cup \{\psi\} \vdash \theta$, then $\Gamma \vdash \theta$.

($\neg$I) If $\Gamma \cup \{\varphi\} \vdash \bot$, then $\Gamma \vdash \neg\varphi$.

($\neg$E) If $\Gamma \vdash \neg\neg\varphi$, then $\Gamma \vdash \varphi$.

Below is an example deduction. By the *scope* we mean the set of current working premises. In other words, $\Gamma \vdash \varphi$ has scope Γ and conclusion φ. Every line of a deduction can be interpreted as a true meta-statement of the form $\Gamma \vdash \varphi$.

Proof that $\{Px \vee Qx, Px \rightarrow Rx, \neg Sx \rightarrow \neg Qx\} \vdash Rx \vee Sx$

Scope	Conclusion	Rule used
1	1. $Px \vee Qx$	P
2	2. $Px \rightarrow Rx$	P
3	3. $\neg Sx \rightarrow \neg Qx$	P
4	4. Px	P
$1,2,3,4$	5. Px	R from 4
$1,2,3,4$	6. $Px \rightarrow Rx$	R from 2
$1,2,3,4$	7. Rx	$\rightarrow E$ from $5,6$
$1,2,3,4$	8. $Rx \vee Sx$	$\vee I$ from 7
9	9. Qx	P
$1,2,3,9$	10. Qx	R from 9
11	11. $\neg Sx$	P
$1,2,3,9,11$	12. $\neg Sx$	R from 11
$1,2,3,9,11$	13. $\neg Sx \rightarrow \neg Qx$	R from 3
$1,2,3,9,11$	14. $\neg Qx$	$\rightarrow E$ from $12,13$
$1,2,3,9,11$	15. Qx	R from 10
$1,2,3,9,11$	16. $\bot$	$\bot I$ from $14,15$
$1,2,3,9$	17. $\neg\neg Sx$	$\neg I$ from $12,16$
$1,2,3,9$	18. Sx	$\neg E$ from 17
$1,2,3,9$	19. $Rx \vee Sx$	$\vee I$ from 18
$1,2,3$	20. $Rx \vee Sx$	$\vee E$ from $1,8,19$

4. Equality

There are rules for equality, but we won't give any examples.

(=I) If t is a term, then $\Gamma \vdash (t = t)$.

(=E) Let φ' be obtained from φ by replacing zero or more occurences of t with t', and none of the variables of t or t' are bound. If $\Gamma \vdash (t = t')$ and $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi'$.

5. Quantifiers

($\forall$I) If $\Gamma \vdash \varphi$ and the variable v does not occur free in Γ, then $\Gamma \vdash (\forall v)\varphi$.

($\forall$E) Let φ' be obtained from φ by replacing all free instances of v in φ with a term t, such that all variables of t are free in φ'. If $\Gamma \vdash (\forall v)\varphi$ then $\Gamma \vdash \varphi'$.

($\exists$I) Let φ' be obtained from φ by replacing all free instances of v in φ with a term t, such that all variables of t are free in φ'. If $\Gamma \vdash \varphi'$, then $\Gamma \vdash (\exists v)\varphi$.

($\exists$E) If $\Gamma \vdash (\exists v)\varphi$ and $\Gamma \cup \{\varphi\} \vdash \psi$, and the variable v does not occur free in Γ or in ψ, then $\Gamma \vdash \psi$.

Here is first a bad deduction and then two good deductions using these rules.

Proof that $\{(\exists x)Px\} \vdash (\forall x)Px$

Scope	Conclusion	Rule used
1	1. $(\exists x)Px$	P
2	2. Px	P
$1,2$	3. Px	R from 2
$1,2$	4. $(\forall x)Px$	$\forall I$ from 3 (BAD)
1	5. $(\forall x)Px$	$\exists E$ from $1,4$

Line 4 is bad because the variable x occurs free in the current set of premises (1,2). All other lines are okay.

Proof that $\{(\exists x)(Px \wedge Qx), (\forall x)\neg Px\} \vdash \perp$

Scope	Conclusion	Rule used
1	1. $(\exists x)(Px \wedge Qx)$	P
2	2. $(\forall x)\neg Px$	P
$1,2$	3. $(\exists x)(Px \wedge Qx)$	R from 1
$1,2$	4. $(\forall x)\neg Px$	R from 2
$1,2$	5. $\neg Px$	$\forall E$ from 4
6	6. $Px \wedge Qx$	P
$1,2,6$	7. $Px \wedge Qx$	R from 6
$1,2,6$	8. Px	$\wedge E$ from 7
$1,2,6$	9. $\neg Px$	R from 5
$1,2,6$	10. $\perp$	$\perp I$ from $8,9$
$1,2$	11. $\perp$	$\exists E$ from $3,10$

Proof that $\{(\forall x)(\forall y)Pxy, (\exists x)(Pxx \to Qxx)\} \vdash (\exists x)Qxx$

Scope	Conclusion	Rule used
1	1. $(\forall x)(\forall y)Pxy$	P
2	2. $(\exists x)(Pxx \to Qxx)$	P
$1,2$	3. $(\forall x)(\forall y)Pxy$	R from 1
$1,2$	4. $(\exists x)(Pxx \to Qxx)$	R from 2
$1,2$	5. $(\forall y)Pxy$	$\forall E$ from 3
$1,2$	6. Pxx	$\forall E$ from 5
7	7. $Pxx \to Qxx$	P
$1,2,7$	8. $Pxx \to Qxx$	R from 7
$1,2,7$	9. Pxx	R from 6
$1,2,7$	10. Qxx	$\to E$ from $8,9$
$1,2,7$	11. $(\exists x)Qxx$	$\exists I$ from 10
$1,2$	12. $(\exists x)Qxx$	$\exists E$ from $4,11$

In all of our examples, when we use the Reiteration rule R, we have only added finitely many premises. Moreover, so far we have only added premises which were already stated via the premise rule. We could, however, add infinitely many premises, and any premises we want, though we would have to notate this a bit differently:

Proof that $\Gamma \cup \{Px\} \vdash Px$

Scope	Conclusion	Rule used
1	1. Px	P
$\Gamma, 1$	2. Px	R from 1

Here Γ could be any set. Usually, we use shorthand $1, 2, 3, \ldots$ to refer to premises by their line number instead of writing them out explicitly.

6. Consistency and Satisfiability

For most of this lecture, Γ was a set of formulas. While we could define $\Gamma \vDash \varphi$ as well for Γ any set of formulas, we only really care about theories (sets of sentences), so we now restrict to that particular case.

A theory $\mathcal{T}$ is *inconsistent* if $\mathcal{T} \vdash \bot$. Otherwise, it is *consistent*.

A theory $\mathcal{T}$ is *unsatisfiable* if it has no models. If it has at least one, it is *satisfiable*.

Soundness says that if $\mathcal{T} \vdash \varphi$ then $\mathcal{T} \vDash \varphi$.

Completeness says that if $\mathcal{T} \vDash \varphi$ then $\mathcal{T} \vdash \varphi$.

Soundness is easy to prove, but completeness is hard, due to Godel.

An elementary consequence of soundness and completeness is that a theory is satisfiable if and only if it is consistent. Another consequence is the compactness theorem, which is forthcoming.

FILTERS AND ULTRAFILTERS

CALEB STANFORD

1. Filters

Given a set X, a filter is a way of calling certain subsets of X equivalent. We will say two subsets of X are equivalent if the set of $x \in X$ where the subsets agree is in the filter. For now, let's just state the definition.

Definition 1. Let X be a set. A **filter** is a set $\mathcal{F}$ with the following properties:

(1) (nonempty) $X \in \mathcal{F}$
(2) (intersections) If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.
(3) (upper set) If $A \in \mathcal{F}$ and $B \supset A$, then $B \in \mathcal{F}$.

Given a filter $\mathcal{F}$, we say that

Definition 2. $\mathcal{F}$ is **free** if $\cap \mathcal{F} = \varnothing$.

Definition 3. $\mathcal{F}$ is **principal** if there exists a set S, $S \neq \varnothing$ and $S \subseteq X$, such that $\mathcal{F} = \{A : S \subseteq A\}$. $\mathcal{F}$ is **nonprincipal** otherwise.

Definition 4. $\mathcal{F}$ is **proper** if $\varnothing \notin \mathcal{F}$.[1]

The basic examples of a filter is a principal filter, formed by taking a set S and considering all sets containing it. Principal filters, however, are in some sense trivial and will be uninteresting to us. We are really interested in free filters (and in particular free ultrafilters); these will allow us to explicitly construct structures out of smaller structures in a very interesting way.

The canonical example of a free filter is the set of cofinite subsets of X. This is called the **Frechet filter** on X.

Here are some important propositions:

Proposition 1. *Every free filter is nonprincipal.*

Proof. Let X be a set and $\mathcal{F}$ a filter on X. Suppose by way of contrapositive that $\mathcal{F}$ is principal. Then $\mathcal{F} = \{A : A \supseteq S\}$, and in particular $S \in \mathcal{F}$, so that clearly $\cap \mathcal{F} = S$. But S is nonempty by the definition of a principal filter, so $\mathcal{F}$ is not free. □

The converse doesn't hold; nonprincipal filters exist which are not free. There are some examples of this in the next section.

Proposition 2. *Let $\mathcal{F}$ be a filter on X. TFAE:*

1. $\mathcal{F}$ is free.

Date: 9 March 2015.

[1]In the Pete Clark notes, and some other sources, filters are required to be proper in the definition. There is only one filter which isn't proper; the set of all subsets of X. Wikipedia allows this trivial filter, so I have allowed it here.

2. $\mathcal{F}$ contains the Frechet filter.

Proof. ($1 \implies 2$) Let $\mathcal{F}$ be free. For each $x \in X$, since the intersection of all sets in $\mathcal{F}$ is empty, there must be a set $A_x \in \mathcal{F}$ such that $x \notin A_x$. The set $X \setminus \{x\}$ in turn contains A_x, so $X \setminus \{x\} \in \mathcal{F}$. Any cofinite set is the intersection of sets of the form $X \setminus \{x\}$, so all cofinite sets must therefore be in $\mathcal{F}$.

($2 \implies 1$) Let F be the Frechet filter. Note that $\mathcal{F} \supset F$ implies $\cap\mathcal{F} \subset \cap F = \varnothing$, so $\cap\mathcal{F} = \varnothing$. ☐

Another thing about filters is they are really only interesting to us in the case that the underlying set is infinite.

Proposition 3. *Let X be a finite set and $\mathcal{F}$ a proper filter on X. Then $\mathcal{F}$ is principal.*

Proof. Let $S = \cap\mathcal{F}$. Since X is finite, $\mathcal{P}(X)$ is finite, so this is a finite intersection; hence $S \in \mathcal{F}$. Then $\mathcal{F} = \{A : S \subseteq A\}$. Since $\mathcal{F}$ is propert, S is nonempty. ☐

2. Examples

Here's a big list of examples. Make sure you understand the classification in each case.

X		Subset of $\mathcal{P}(X)$	Filter?	Principal?	Free?
$\{0,1\}$	1.	$\{0\}, \{0,1\}$	Y	Y	N
	2.	$\{0,1\}$	Y	Y	N
	3.	$\{0,1\}, \{0\}, \{1\}$	N		
$\mathbb{N}$	4.	$\{S : \lvert S\rvert < \infty\}$	N		
	5.	$\{S : \lvert\mathbb{N} \setminus S\rvert < \infty\}$	Y	N	Y
	6.	$\{S : \lvert\mathbb{N} \setminus S\rvert < \infty, 0 \in S\}$	Y	N	N
	7.	$\{S : \lvert\mathbb{N} \setminus S\rvert < \infty, 0 \notin S\}$	N		
	8.	$\{S : 31 \in S\}$	Y	Y	N
	9.	$\{S : 2, 4, 6, 8, \ldots \in S\}$	Y	Y	N
$\mathbb{R}$	10.	$\{S : 0 \in \mathrm{Int} S\}$	Y	N	N
	11.	$\{S : 0 \in S\}$	Y	Y	N
	12.	$\{S : \mathrm{Int} S \neq \varnothing\}$	N		
	13.	$\{S : S \text{ open}, S \neq \varnothing\}$	N		
	14.	$\{S : \mathbb{R} \setminus S \text{ is bounded}\}$	N		
	15.	$\{S : \mathbb{R} \setminus S \text{ has measure } 0\}$	Y	N	Y
	16.	$\{S : \mathbb{R} \setminus S \text{ has finite outer measure}\}$	Y	N	Y

Example 5 is the Frechet filter on $\mathbb{N}$. Examples 6 and 10 illustrate the character of nonprincipal filters which are not free: they would be free, except that a certain element (or multiple elements) are required to be in every set of the filter.

3. Ultrafilters

Definition 5. Let X be a set. An **ultrafilter** on X is a filter $\mathcal{U}$ which additionally satisfies:

(4) (proper) $\varnothing \notin \mathcal{U}$
(5) (maximal) For all $S \subseteq X$, either $S \in \mathcal{U}$ or $X \setminus S \in \mathcal{U}$.

An equivalent definition is that for all $S \subseteq X$, either $S \in \mathcal{U}$ or $X \setminus S \in \mathcal{U}$, but not both. The "but not both" implies that $\mathcal{U}$ doesn't contain the empty set, and vice versa.

Of the examples in the previous section, $1, 8$, and 11 are ultrafilters. But all of these are *principal* ultrafilters, and seem quite trivial. In fact, the following holds:

Proposition 4. *Let X be a set and $\mathcal{U}$ be an ultrafilter on X. TFAE:*

1. *$\mathcal{U}$ is not free.*
2. *$\mathcal{U}$ is principal.*
3. *$\mathcal{U} = \{S : x \in S\}$ for some $x \in X$.*

Proof. Homework. □

So, principal ultrafilters are just those generated by a single element. Why didn't we give any examples of nonprincipal ultrafilters? Because their existence turns out to be equivalent to a weaker version of the axiom of choice known as the **ultrafilter lemma**. As a result, they can't be constructed explicitly.

Theorem 1 (Ultrafilter Lemma)**.** *Let X be a set and $\mathcal{F}$ be a filter on X. If $\mathcal{F}$ is proper, then there is an ultrafilter on $\mathcal{U}$ containing X.*

Proof. Order proper filters by inclusion. Show that every chain of filters has an upper bound. Apply the lemma of Zorn and show the resulting filter is an Ultrafilter. □

Corollary 1. *Let X be an infinite set. Then there exists a nonprincipal ultrafilter on X.*

Proof. Let F be the Frechet filter on X. Since X is infinite, $\varnothing$ is not cofinite, so F is proper. By the ultrafilter lemma, there is thus an ultrafilter $\mathcal{U} \supset F$.

By proposition 2, since $\mathcal{U}$ contains the Frechet filter, it is free. By proposition 4 (or by the weaker proposition for filters), since $\mathcal{U}$ is free, it is nonprincipal. □

Note that the Ultrafilter lemma is easy in the case that the filter to be completed is not free; just take any element in the intersection of the filter, and take the principal ultrafilter generated by that element. Zorn's lemma is required only to construct nonprincipal ultrafilters.

4. Bulding structures from structures

In the past, we have used compactness and the Lowenheim-Skolem theorems to construct models for various theories. But in every case, the resulting models were not explicit. What if we want to build explicit models?

Given an index set I and $\mathcal{L}$-structures $\{X_i : i \in I\}$, one obvious thing to do is take the cartesian product, $X := \prod_{i \in I} X_i$. This is an $\mathcal{L}$-structure under the interpretation where function symbols are applied to each individual coordinate, and relation symbols hold exactly when they hold in every individual coordinate.

The problem is that, supposing $X_i \vDash \mathcal{T}$ for all i, for some theory $\mathcal{T}$, we don't necessarily have $X \vDash \mathcal{T}$, which we would like to be true. For example, this does turn out to be true in the theory of groups and the theory of rings (as the cartesian product of groups or rings is itself a group or ring), but this doesn't work when we have the theory of integral domains, or fields, or totally ordered sets.

The solution is to first form the cartesian product, and then to "mod out" by an ultrafilter. If the set I is infinite and the ultrafilter is nonprincipal, we will get (by black magic) a model which satisfies exactly the same formulas as X_i, if all X_i were the same. If the X_i are different, it is slightly more complicated, but we can still say exactly what formulas are satisfied. This result is known as Los's theorem, and is otherwise known as the fundamental theorem of ultrafilters. We will get to it next time.

LECTURE 25: ULTRAPRODUCTS

CALEB STANFORD

First, recall that given a collection of sets and an ultrafilter on the index set, we formed an ultraproduct of those sets. It is important to think of the ultraproduct as a *set-theoretic* construction rather than a *model-theoretic* construction, in the sense that it is a product of sets rather than a product of structures. I.e., if X_i are sets for $i = 1, 2, 3, \ldots$, then $\prod X_i/\mathcal{U}$ is another set. The set we use does not depend on what constant, function, and relation symbols may exist and have interpretations in X_i. (There are of course profound model-theoretic consequences of this, but the underlying construction is a way of turning a collection of sets into a new set, and doesn't make use of any notions from model theory!)

We are interested in the particular case where the index set is $\mathbb{N}$ and where there is a set X such that $X_i = X$ for all i. Then $\prod X_i/\mathcal{U}$ is written $X^{\mathbb{N}}/\mathcal{U}$, and is called the **ultrapower of $\boldsymbol{X}$ by $\boldsymbol{\mathcal{U}}$**. From now on, we will consider the ultrafilter to be a fixed nonprincipal ultrafilter, and will just consider the **ultrapower of $\boldsymbol{X}$** to be the ultrapower by this fixed ultrafilter. It doesn't matter which one we pick, in the sense that none of our results will require anything from $\mathcal{U}$ beyond its nonprincipality.

The ultrapower has two important properties. The first of these is the Transfer Principle. The second is $\aleph_0$-saturation.

1. The Transfer Principle

Let $\mathcal{L}$ be a language, X a set, and $X_{\mathcal{L}}$ an $\mathcal{L}$-structure on X. Let $\mathcal{U}$ be a nonprincipal ultrafilter on $X^{\mathbb{N}}$. Let $Y = X^{\mathbb{N}}/\mathcal{U}$. Łoś's theorem tells us that we can interpret Y as an $\mathcal{L}$-structure $Y_{\mathcal{L}}$ in a natural way, and that for any $\mathcal{L}$-sentence φ:

$$\begin{aligned} Y_{\mathcal{L}} \vDash \varphi &\iff \{i \,:\, X_i \vDash \varphi\} \in \mathcal{U} \\ &\iff \{i \,:\, X_{\mathcal{L}} \vDash \varphi\} \in \mathcal{U} \\ &\iff X_{\mathcal{L}} \vdash \varphi. \end{aligned}$$

In other words, in the Ultrapower case, Łoś's theorem collapses down to the simple statement that $Y_{\mathcal{L}}$ is elementarily equivalent to $X_{\mathcal{L}}$.

But wait! Remember that Y itself is simply a set that is a function of the set X. The above will be true for ANY $\mathcal{L}$-structure that we put on X; we will always get a corresponding $\mathcal{L}$-structure on Y that satisfies the elementary equivalence.

So, let $\mathcal{L}$ be the set of ALL functions, relations, and constants on X. This includes a symbol for every function and relation and constant you may be interested in, but it also includes symbols for everything else: many functions and relations and constants which you never thought to consider, and many which can't be written down explicitly. Then $Y_{\mathcal{L}}$ and $X_{\mathcal{L}}$ are elemantarily equivalent. This is known as the Transfer Principle. We restate it below, introducing some new notation that is often used.

Theorem 1. (Transfer Principle) Let X be a set. Let ${}^*X$ (what we have been calling Y) be an ultrapower of X. Let $c_1, c_2, c_3, \ldots$ be constant elements of X, let $R_1, R_2, R_3, \ldots$ be relations on X, and let $F_1, F_2, F_3, \ldots$ be functions on X. Let ${}^*c_1, {}^*c_2, \ldots, {}^*R_1, {}^*R_2, \ldots, {}^*F_1, {}^*F_2, \ldots$ be the corresponding constants, relations, and functions on ${}^*X$. Let φ be a first-order sentence over the language $\{c_i, R_i, F_i\}$, and let ${}^*\varphi$ be the corresponding first-order sentence over the language $\{{}^*c_i, {}^*R_i, {}^*F_i\}$. Then $X \vDash \varphi$ if and only if ${}^*X \vDash {}^*\varphi$.

Proof. From Łoś's theorem, as described above. □

Here are some examples.

Example 1. Let $X = \mathbb{R}$. In $\mathbb{R}$ it is true that $\forall x : |x| \geq 0$. Therefore, in ${}^*\mathbb{R}$, $\forall x : {}^*|x| \geq 0$.

Date: 2015-04-10.

Example 2. Let X be the set of finite binary strings. Let $\circ$ denote concatenation. In X it is the case that $\forall x \forall y \forall z : (x \circ y) \circ z = x \circ (y \circ z)$. Therefore, in ${}^*X$, $\forall x \forall y \forall z : (x \,{}^*\!\circ y) \,{}^*\!\circ z = x \,{}^*\!\circ (y \,{}^*\!\circ z)$.

First-order logic generally forces us to quantify over the entire set X and not over a subset of X. But a subset of X is actually an element of our language now (a unary relation), so we can formalize e.g. the statement "x is in the Cantor set" as Cx, where C denotes this unary relation. Thus the Transfer Principle also gives us things like:

Example 3. (open set) A set A is open in $\mathbb{R}$ if and only if

$$\forall x \in A \; \exists \epsilon \in (0, \infty) \; \forall y \in \mathbb{R} \; : \; (|x - y| < \epsilon \to y \in A)$$

Therefore, for any open set A of $\mathbb{R}$ we have

$$\forall x \in {}^*A \; \exists \epsilon \in {}^*(0, \infty) \; \forall y \in {}^*\mathbb{R} \; : \; ({}^*|x - y| \,{}^*{<} \epsilon \to y \in {}^*A).$$

Since X is embedded in ${}^*X$, in the future, we will generally think of X as being a subset of ${}^*X$. We will then generally omit the * before constants, functions, and relations, with the exception that we distinguish between a set $A \subseteq X$ and its corresponding set ${}^*A$. Just to illustrate why this is not ambiguous:

- For constants, if we say "let $c \in X$", this also implies that $c \in {}^*X$, as we are thinking of X as a subset of ${}^*X$. We are also allowed to use c in any transfer principle arguments.
- If on the other hand we say "let $c \in {}^*X$", it is perfectly clear what we mean, but we are not then allowed to apply the transfer principle to sentences involving c.
- For relations, if we say "let R be a binary relation on X", it is clear what we mean. Then R extends naturally to a binary relation on ${}^*X$, so we can compare both things in X and things in ${}^*X$ using R. Note that it is important here that ${}^*R$ agrees with R on X.
- If we say "let R be a binary relation on ${}^*X$", then R is also well-defined on all of ${}^*X$ as well as X, but we can't apply the transfer principle to a sentence involving R.
- Similarly, functions f defined on X are assumed to be extended automatically in the natural way to ${}^*X$, but functions defined on ${}^*X$ originally cannot be dealt with by the transfer principle.
- Finally, when we fix a subset A of X, this is to be thought of as different than a unary relation on X. A unary relation would extend naturally to ${}^*X$ via the transfer principle, but A is considered to be a fixed subset of X which is in turn a subset of ${}^*X$; A is the same subset of ${}^*X$ as it is of X. If we want to consider the corresponding (different) subset of ${}^*X$, we will write ${}^*A$.

In summary, we are able to keep things straight if we just remember whether the function or relation was defined originally on X, or on ${}^*X$.

2. $\aleph_0$-Saturation

Up to this point, it has seemed that ${}^*X$ is just a bigger version of X; elementarily equivalent, in fact. So what use is there in defining it? If it is so similar to X, why not just use X?

The answer is that we also have a lot more than just was in X, and we can exploit that. For instance, the archimedian property (which is not first-order) holds in $\mathbb{R}$ but not in ${}^*\mathbb{R}$, and this turns out to allow us to define calculus of $\mathbb{R}$ using infinitesimal elements of ${}^*\mathbb{R}$.

Theorem 2. ($\aleph_0$-saturation) Let $\mathcal{T} = (\varphi_1, \varphi_2, \ldots)$ be a countable set of formulas in free variables $u_1, u_2, \ldots, u_k$. Suppose that every finite subset Σ of $\mathcal{T}$ has a solution in X^k. Then $\mathcal{T}$ has a solution in $({}^*X)^k$.

Proof. Define infinite sequences $u_1, u_2, \ldots, u_k \in {}^*X$ by

$$\begin{aligned} u_1 &:= ({}_1u_1, {}_2u_1, {}_3u_1, \ldots) \\ u_2 &:= ({}_1u_2, {}_2u_2, {}_3u_2, \ldots) \\ &\cdots \\ u_k &:= ({}_1u_3, {}_2u_3, {}_3u_3, \ldots) \end{aligned}$$

such that ${}_iu_1, {}_iu_2, \ldots, {}_iu_k$ is a solution in X^k for the finite set of formulas $\Sigma_i := (\varphi_1, \varphi_2, \ldots, \varphi_i)$. Then observe that for each φ_j, there are only finitely many i such that φ_j is not in Σ_i, and hence there are only finitely many i for which ${}_iu_1, {}_iu_2, \ldots, {}_iu_k$ is not a solution to φ_j. Therefore, the set

$$\{i \; : \; X_i \vDash \varphi_j({}_iu_1, {}_iu_2, \ldots, {}_iu_k)\}$$

is cofinite, therefore being a member of our ultrafilter. This implies by the definition of satisfaction in ${}^*X$ and by Łoś's theorem that

$$ {}^*X \vDash \varphi_j(u_1, u_2, \dots, u_k) $$

This is true for any φ_j, so ${}^*X \vDash \mathcal{T}(u_1, u_2, \dots, u_k)$. ☐

Example 4. In $\mathbb{R}$, let $\mathcal{T}$ be the set of formulas

$$ \{x < 1, x < 1/2, x < 1/3, \dots\} $$

Then there is an element $x \in {}^*\mathbb{R}$ satisfying the above. This is called an **infinitesimal**.

Example 5. In $\mathbb{N}$, let $\mathcal{T}$ be the set of formulas

$$ \{1 \mid x, 2 \mid x, 3 \mid x, \dots\} $$

Then there is an element $x \in {}^*\mathbb{N}$ satisfying all the above, i.e. there is a hypterinteger divisible by every integer.

Example 6. In $\mathbb{R}$, take $\mathcal{T}$ to be

$$ \{x > 1, x > 2, x > 3, \dots\} \cup \{y > x, y > x^2, y > x^3, \dots\} $$

The result is two hyperreal numbers, x and y, both infinite, but such that y is much bigger than x.

3. Applications

3.1. **Infinitely many primes.** In class, we proved that there are infinitely many primes. The idea is to form a hyperinteger divisible by every standard prime number, then to add one. The resulting hyperinteger must be divisible by a hyperprime, but it isn't divisible by any standard primes. As a lemma, a subset $S \subseteq X$ is finite if and only if ${}^*S = S$. So the fact that there are hyperprimes which are not prime means that the set of primes is infinite.

3.2. **ZFC.** (Background: ZFC is the first-order theory of set theory. The language of ZFC consists only of the binary relation $\in$.)

Assume ZFC is consistent. Then there is a model of ZFC, call it M. Let ω be the element of M corresponding to the natural numbers, the first uncountable ordinal. Consider the set of formulas

$$ \mathcal{T} = \{x \in \omega, x \neq 1, x \neq 2, x \neq 3, \dots\} $$

Every finite subset of $\mathcal{T}$ is satisfiable in M. Therefore, by $\aleph_0$-saturation, $\mathcal{T}$ is satisfiable in ${}^*M$. That is, the set of natural numbers ω in ${}^*M$ actually contains something which is not a natural number.

Why is this a problem? Well, it is true in ZFC that every nonzero natural number has a predecessor, so x has a predecessor x_1, which has a predecessor x_2, and so on. None of these are equal to any of $1, 2, 3, \dots$, else x would be equal to one of $1, 2, 3, \dots$ just by applying taking a few successors of x_i.

So $x_1, x_2, x_3, \dots$ is an infinite decreasing chain of "natural numbers". Worse, it is an infinite decreasing chain of ordinals; each is contained in the next! And taking the set

$$ S = \{x_1, x_2, x_3, \dots\} $$

we find that S contains no element disjoint from itself, which violates one of the axioms of set theory (axiom of regularity).

What gives? Well, we can externally write down S in our meta-theory, and claim it is a set, but the bizarre model of ZFC ${}^*M$ does not know about S. Nor does this model of ZFC have any way to form the infinite decreasing chain $x_1, x_2, x_3, \dots$ and compile it into a single list. Just to illustrate this, notice that the usual definition of a countable list is a function from ω to a set; yet ω is much larger in ${}^*M$ than it is in our standard understanding of ZFC.