Companion to the course mathematical logic

Tapani Hyttinen

Abstract

In the lectures I will make some small changes to the definitions and proofs from J. Väänänen's lecture notes 'A short course on mathematical logic'. In these notes I will explain what the changes are and why they are made.

1. On recursive definitions

In this course almost everything is defined using recursion and proved using induction. Thus I think it is important that the student understands how and why they work. This is the purpose of this section.

Let A be a non-empty set and f an n-ary function from A to A, $n \in \mathbb{N}$ (i.e. $f : A^n \to A$). We say that $C \subseteq A$ is f-closed if for all $c_1, ..., c_n \in C$, $f(c_1, ..., c_n) \in C$. If F is a set functions from A to A of finite arity, we say that $C \subseteq A$ is F-closed if it is f-closed for all $f \in F$. For non-empty $B \subseteq A$, we write $cl_A(B, F)$ for the $\subseteq$-least subset of A that contains B and is F-closed (if such set exists).

1.1 Lemma. *$cl_A(B, F)$ exists.*

Proof. Let S be the family of all F-closed $D \subseteq A$ that contain B. Notice that $A \in S$. Let C be the intersection of all elements of S. Then $C = cl_A(B, F)$: Clearly $B \subseteq C$ and if $c_i, ..., c_n \in C$ and $f \in F$, then for all $D \in S$, $f(c_1, ..., c_n) \in D$ (since $c_1, ..., c_n \in D$ and D is f-closed) and thus $f(c_1, ..., c_n) \in C$ i.e. C is F-closed. Since every F-closed set that contains B belongs to S, C is the $\subseteq$-least such. □

1.2 Example. *Let G be a group, $f : G^2 \to G$ be the group operation and $g : G \to G$ be such that $g(a) = a^{-1}$. Then for all non-empty $B \subseteq G$, $cl_G(B, \{f, g\})$ is the subgroup generated by B.*

1.3 Lemma. *Let P be any property (expressible e.g. in set theory). Then every element of $cl_A(B,F)$ has P if the following holds:*

(i) every element of B has P,

(ii) if each of $c_1,...,c_n \in cl_A(B,F)$ has P and $f \in F$ is of arity n, then $f(c_1,...,c_n)$ has P.

Proof. Let D be the set of all $c \in cl_A(B,F)$ that has P. By (i), D contains B and by (ii), D is F-closed. Since $cl_A(B,F)$ is the $\subseteq$-least such, $cl_A(B,F) \subseteq D$ i.e. every element of $cl_A(B,F)$ has P. □

1.4 Example. *Let $B = \{0\}$ and $f : \mathbb{N} \to \mathbb{N}$ be such that $f(n) = n+1$. Then $\mathbb{N} = cl_{\mathbb{N}}(B,\{f\})$ and thus every natural number has a property P if 0 has it and if $n \in \mathbb{N}$ has P, then also $n+1$ has it.*

1.5 Exercise. *Show that for all $c \in cl_A(B,F)$ either $c \in B$ or there are $c_1,...,c_n \in cl_A(B,F)$ and $f \in F$ such that $c = f(c_1,...,c_n)$.*

1.6 Example. *Let W be the set of all finite sequences of symbols from the set $V = \{(,),+,X,1\}$ i.e. the set of all words in vocabulary V. For $u,v \in W$, we write uv for the concatenation of u and v. Let $B = \{X,1\}$ and $f : W^2 \to W$ be such that $f(u,v) = (u+v)$. We write $LP = cl_W(B,\{f\})$ (so e.g. $((X+1)+1) \in LP$, $LP =$ linear polynomials). We write also $V(u)$ for the number of left brackets in u and $O(u)$ for the number of right brackets.*

Structure trees: We will look these in lectures on blackboard.

1.7 Exercise.

(i) $O(u) = V(u)$ for all $u \in LP$.

(ii) If u and w are non-empty words and $uw \in LP$, then $V(u) > O(u)$.

We say that the triple (A,B,F) is good if for all $c \in cl_A(B,F)$ the following holds: Either $c \in B$ or there are unique $f \in F$ and $c_1,...,c_n \in cl_A(B,F)$ such that $c = f(c_1,...,c_n)$.

1.8 Exercise. *Let W, B and f be as in Example 1.6. Show that $(W,B,\{f\})$ is good. Hint: Use Exercise 1.7.*

For all $n \in \mathbb{N}$ we define $cl^n_A(B,F)$ as follows: $cl^0_A(B,F) = B$ and

$$cl^{n+1}_A(B,F) = cl^n_A(B,F) \cup \{f(c_1,...,c_n)|\ c_1,...,c_n \in cl^n_A(B,F),\ f \in F\}.$$

1.9 Lemma. $\bigcup_{n=0}^{\infty} cl^n_A(B,F) = cl_A(B,F)$.

Proof. Easy induction on n shows that for all $n \in \mathbb{N}$, $cl^n_A(B,F) \subseteq cl_A(B,F)$ and thus $\bigcup_{n=0}^{\infty} cl^n_A(B,F) \subseteq cl_A(B,F)$. On the other hand, clearly, $\bigcup_{n=0}^{\infty} cl^n_A(B,F)$ is F-closed and contains B and thus $cl_A(B,F) \subseteq \bigcup_{n=0}^{\infty} cl^n_A(B,F)$. □

1.10 Lemma. *Suppose (A,B,F) is good, R is a set, $g : B \to R$ and for all $f \in F$, $g_f : R^n \to R$, where n is the arity of f. Then there is a unique $h : cl_A(B,F) \to R$ such that $h \restriction B = g$ and for all $c \in cl_A(B,F) - B$, if $c = f(c_1,...,c_n)$, then $h(c) = g_f(h(c_1),...,h(c_n))$.*

Proof. For all $n \in \mathbb{N}$, we define function $h_n : cl_A^n(B,F) \to R$ as follows: $h_0 = g$ and $h_{n+1} : cl_A^{n+1}(B,F) \to R$ is such that $h_{n+1} \restriction cl_A^n(B,F) = h_n$ and for $c = f(c_1,...,c_n) \in cl_A^{n+1}(B,F) - cl_A^n(B,F)$, $h_{n+1}(c) = g_f(h_n(c_1),...,h_n(c_n))$. Since (A,B,F) is good, each h_n is well-defined and clearly $h = \bigcup_{n=0}^{\infty} h_n$ is the required function. The uniqueness of h follows by an easy induction. □

1.11 Example. *There is a unique function $h : LP \to \mathbb{N}$ such that $h(X) = h(1) = 0$ and $h((u+v)) = max\{h(u),h(v)\}+1$.*

Definition of h on structure trees: We will look this in lectures on blackboard.

1.12 Exercise. *Let S be the set of all (unary) functions $\mathbb{N} \to \mathbb{N}$, $f_0 \in S$ be such that $f_0(n) = n$ for all $n \in \mathbb{N}$ and $f_1 \in S$ such that $f_1(n) = 1$ for all $n \in \mathbb{N}$. Let $B = \{f_0, f_1\}$ and $F = \{f_+\}$, where $f_+ : S^2 \to S$ is such that $f_+(g_0,g_1) = g_2$ if for all $n \in \mathbb{N}$, $g_2(n) = g_0(n) + g_1(n)$. We write $g_0 + g_1$ for $f_+(g_0,g_1)$. Let $LF = cl_S(B,F)$. Show that there is no function $h : LF \to \mathbb{N}$ such that $h(f_0) = h(f_1) = 0$ and $h(g_0+g_1) = max\{h(g_0),h(g_1)\}+1$.*

1.13 Exercise. *Suppose $g : \mathbb{N} \to \mathbb{N}$. Show that $g \in LF$ iff there are $m,k \in \mathbb{N}$ such that $m+k \neq 0$ and for for all $n \in \mathbb{N}$, $g(n) = mn+k$.*

1.14 Exercise. *By Lemma 1.10, there is $Ev : LP \to LF$ (Ev = evaluation) such that $Ev(X) = f_0$, $Ev(1) = f_1$ and $Ev((u+w)) = Ev(u) + Ev(w)$. Show that Ev is a surjection.*

1.15 Exercise. *Let $n \in \mathbb{N}$. We define $Ev_n : LP \to \mathbb{N}$ as follows: $Ev_n(X) = n$, $Ev_n(1) = 1$ and $Ev_n((u+w)) = Ev_n(u) + Ev_n(w)$. Show that for all $w \in LP$, $Ev_n(w) = Ev(w)(n)$.*

So this operation $cl_A(B,F)$ gives meaning to our recursive definitions. When we use it we usually do not specify A, it is clear from the context and usually the choice of A does not matter much, and B and F are given more implicitely. E.g. the definition of LP will get the form:

(i) X and 1 are LP,

(ii) if w and u are LP, then also $(w+u)$ is.

And then inductive proofs (i.e. proofs based on Lemma 1.3) go along this form of definition. E.g. in the case of LP there are two steps: (1) $w = X$ or $w = 1$ and (2) $w = (u_1 + u_2)$ and in this latter case we assume (without mentioning it) that $u_1, u_2 \in cl_W(\{X,1\},\{f\})$ and that they satisfy the claim, this is called the induction assumption.

When a recursive definition is based on a good triple, we say that the definition is good.

2. Propositional logic

Here we follow the lecture notes. There is a typo in Definition 2.11: $v(S) = 1$ if $v(A) = 1$ for all $A \in S$. Also in the proof of Theorem 2.26 in Case 1, Theorem 2.24 should be Theorem 2.21.

The definition of propositional formulas is good. The definition of $S \vdash A$ is not good. In particular, it follows that the deductions (proofs) are not unique, there does not even exist any canonical deductions, see Problem 2.3.8 from the lecture notes.

3. Structures

Again we follow the lecture notes except that we make a small change to the notation: If $\mathcal{A} = (A, Sat_{\mathcal{A}})$ is an L-structure, we write $R^{\mathcal{A}}$ for $Sat_{\mathcal{A}}(R)$ for relation symbols $R \in L$ and similarly for function and constant symbols. Thus, e.g., if $L = \{R_0, R_1, f, c\}$, we write $\mathcal{A} = (A, R_0^{\mathcal{A}}, R_1^{\mathcal{A}}, f^{\mathcal{A}}, c^{\mathcal{A}})$ for $(A, Sat_{\mathcal{A}})$. Sometimes the symbol from the vocabulary and its interpretation are denoted by the same symbol. E.g. if $L = \{+, \times, 0, 1\}$ and $\mathcal{A}$ is the ring of integers, we write simply $\mathcal{A} = (\mathbf{Z}, +, \times, 0, 1)$ and this means that the interpretation of $+$ is the usual addition of integers etc. So e.g. $0^{\mathcal{A}} = 0$. However it is important to keep in mind that in this $0^{\mathcal{A}} = 0$, the first 0 is the symbol from the vocabulary and the second is the integer zero. Mixing these two causes trubles.

We do this to follow the usual notations from the literature. This will turn out to be beneficial accasionally.

We will also talk about congruence relations: Let $\mathcal{A} = (A, Sat_{\mathcal{A}})$ be an L-structure. We say that $E \subseteq A^2$ is a congruence relation if

(i) E is an equivalence relation,

(ii) if $R \in L$ is an $n+1$-ary relation symbol and $a_i, b_i \in \mathcal{A}$, $i \leq n$, are such that for all $i \leq n$, $(a_i, b_i) \in E$, then $(a_0, ..., a_n) \in R^{\mathcal{A}}$ iff $(b_0, ..., b_n) \in R^{\mathcal{A}}$,

(ii) if $f \in L$ is an $n+1$-ary function symbol and $a_i, b_i \in \mathcal{A}$, $i \leq n$, are such that for all $i \leq n$, $(a_i, b_i) \in E$, then $(f^{\mathcal{A}}(a_0, ..., a_n), f^{\mathcal{A}}(b_0, ..., b_n)) \in E$.

The point in these congruence relations is that they allow us to define a new structure $\mathcal{B} = \mathcal{A}/E$: We write a/E for the E-equivalence class of a and then

(i) the universe B of $\mathcal{B}$ is $\{a/E |\ a \in A\}$,

(ii) for $n+1$-ary realation symbols $R \in L$, $R^{\mathcal{B}} = \{(a_0/E, ..., a_n/E) |\ (a_0, ..., a_n) \in R^{\mathcal{A}}\}$,

(iii) for $n+1$-ary function symbols and $a_0, ..., a_n \in A$, $f^{\mathcal{B}}(a_0/E, ..., a_n/E) = f^{\mathcal{A}}(a_0, ..., a_n)/E$.

The reader may want to compare this construction to the construction of G/H, where G is a group and H is a normal subgroup of G (normality guarantees that the equivalence relation $ab^{-1} \in H$ is a congruence relation), see also Exercise 3.6.

3.1 Exercise. *Show that $\mathcal{A}/E$ is well-defined.*

3.2 Exercise. *For groups $(G, +)$ and $(H, +')$, we write $(G, +) \times (H, +')$ for the group $(F, +'')$, where $F = G \times H = \{(a, b) |\ a \in G,\ b \in H\}$ and $(a, b) +'' (c, d) = (a + c, b +' d)$. Below $\mathbf{Z} = (\mathbf{Z}, +)$ is the additive group of integers. In Exercise 3.3, $\mathbf{Z}$ is also the set of integers.*

(i) Show that $\mathbf{Z}/6\mathbf{Z}$ is isomorphic with $(\mathbf{Z}/3\mathbf{Z}) \times (\mathbf{Z}/2\mathbf{Z})$.

(ii) Show that $\mathbf{Z}/9\mathbf{Z}$ is not isomorphic with $(\mathbf{Z}/3\mathbf{Z}) \times (\mathbf{Z}/3\mathbf{Z})$.

By $(Aut(M),\circ)$ we mean the set of all automorphisms $f : M \to M$ with the composition $\circ$ of functions as the group operation.

3.3 Exercise. *Let $S : \mathbf{Z} \to \mathbf{Z}$ be such that $S(z) = z+1$ for all $z \in \mathbf{Z}$. Show that $(\mathbf{Z},+)$ is isomorphic with $(Aut((\mathbf{Z},S)),\circ)$.*

3.4 Exercise. *Suppose A is a non-empty set and f and g are bijections $A \to A$. Show that f is an automorphism of (A,g) iff g is an automorphism of (A,f).*

3.5 Exercise. *Write $\mathbf{R}^*_+ = \{x \in \mathbf{R} |\ x > 0\}$, where $\mathbf{R}$ is the set of reals. Show that $\pi : \mathbf{R}^*_+ \to \mathbf{R}^*_+$, $\pi(x) = x^2$, is an automorphism of $(\mathbf{R}^*_+, \times, <, 1)$.*

3.6 Exercise. *Let $\mathbf{R}[X] = (\mathbf{R}[X],+,\times)$ be the polynomial ring (in one indeterminate) over the reals ($+$ and $\times$ are the addition and multiplication of polynomials). Let $I \subseteq \mathbf{R}[X]$ be an ideal i.e. the zero polynomial $0 \in I$, if $P, Q \in I$, then $P+Q$ and the additive inverse $-P$ are in I (i.e. $(I,+)$ is a subgroup of $(\mathbf{R}[X],+)$) and for all $P \in I$ and $Q \in \mathbf{R}[X]$, $QP \in I$. Show that the relation $\{(P,Q)|\ P,Q \in \mathbf{R}[X],\ P-Q \in I\}$ is a congruence relation.*

4. Predicate logic

Here we will make more changes: We make small changes to the notations, we give alternative truth definition, we change a bit the definition of deduction and we will give proper definitions for concepts like FVF. We also need to make small changes to the theorems and we point out some changes to the proofs.

Already in propositional logic, when we wrote formulas, we did not follow the definition. Here we do the same for terms and formulas of predicate logic. So we do not change the definitions, we just change the way write terms and formulas. In the case of terms, we write $f(t_1,...,t_n)$ instead of $ft_1...t_n$. This is the usual way of writing functions and it induces readability. In some cases in mathematics, there are other common ways of writing these terms and in these cases we follow those. E.g. if $L = \{+,\times,0,1\}$ and we think these as addition etc., we write e.g. $(v_1(1+1)+v_2)v_3$ instead of $\times(+(\times(v_1,+(1,1)),v_2),v_3)$ and by this we mean the term $\times + \times v_1 + 11 v_2 v_3$. We need to know that the definition of a term is good in the sense of Section 1. The proof of this is a bit tricky and thus we omit it. If the one has problems with this, one can add the brackets to the definition of a term as we did above with our notations and then one can prove the goodness exactly as in Exercise 1.8. If one does this, then one must keep in mind these bracket when one works with codes for terms later in these lectures.

Similarly, we will write $t_1 = t_2$ for L-equations instead of $\approx t_1 t_2$. So following what is above, we write e.g. $(v_1(1+1)+v_2)v_3 = 1 + v_2 v_3 + v_1$ for the formula $\approx \times + \times v_1 + 11 v_2 v_3 + +1 \times v_2 v_3 v_1$. However, as above it is important to recognize when $=$ is a symbol and when it means identity: E.g. in $t_1 = t_2$ it is usually a symbol, i.e. in itself, it means nothing and so this $t_1 = t_2$ is just a string of symbols (this could also mean that the two strings t_1 and t_2 of symbols are the same but usually it does not, however, see Definition 4.1(i) below) and in $t_1^{\mathcal{A}} < s >= t_2^{\mathcal{A}} < s >$ it means identity i.e. that the two interpretations are the same. So the first is just a string of symbols and the second is a mathematical claim. One can tell the difference from the context.

We write also $R(t_1, ..., t_n)$ in place of $Rt_1...t_n$. The goodness of the definition of a formula can be proved exactly as in Exercise 1.8 (now we have the brackets). And from these $\approx t_1t_2$ and $Rt_1...t_n$ one can figure out what the terms are. Again a bit tricky but, again, by adding the brackets to the definition of a term, the problem disappears (exercise).

Now we redefine Tarski's truth definition in the form it is usually presented.

4.1 Definition. *$\mathcal{A} \models_s \phi$ is defined as follows (by recursion on the definition of a formula):*

(i) $\phi = t_1 = t_2$ (notice that the first $=$ is identity and the second $=$ is a symbol): $\mathcal{A} \models_s \phi$ if $t_1^{\mathcal{A}} < s >= t_2^{\mathcal{A}} < s >$,

(ii) $\phi = R(t_1, ..., t_n)$: $\mathcal{A} \models_s \phi$ if $(t_1^{\mathcal{A}} < s >, ..., t_n^{\mathcal{A}} < s >) \in R^{\mathcal{A}}$,

(iii) $\phi = \neg\psi$: $\mathcal{A} \models_s \phi$ if $\mathcal{A} \not\models_s \psi$,

(iv) $\phi = \psi \to \theta$: $\mathcal{A} \models_s \phi$ if $\mathcal{A} \not\models_s \psi$ or $\mathcal{A} \models_s \theta$,

(v) $\phi = \forall v_n\psi$: $\mathcal{A} \models \phi$ if for all $a \in \mathcal{A}$, $\mathcal{A} \models_{s(a/n)} \psi$.

Let us look how this definition fits to Lemma 1.10 which will guarantee that the definition is well-made: Definition 4.1 can be seen to define a valuation V in the style of propositional logic (1 true and 0 false). The definition depends on M but it is fixed, so we just remind us of M by putting it to a subscript of V, i.e. we call the valuation V_M. The definition goes by recursion on ϕ, so the domain of V_M is the set of all L-formulas. The truth value depends also on the assigment s and thus the range of V_M is the set of all functions from the set of all assignments for M to the set $\{0, 1\}$ i.e. if we write $S(M)$ for the set of all assignments for M, the range is the set of all functions $X : S(M) \to \{0, 1\}$. So now e.g. (i) can be read saying that

(a) $(V_M(t = u))(s) = 1$ iff $t^M < s >= u^M < s >$,

(iii) can be read saying that

(b) $(V_M(\phi))(s) = 1 - (V_M(\psi))(s)$

and (v) can be read saying that

(c) $(V_M(\phi))(s) = 1$ iff $(V_M(\psi))(s(a/n)) = 1$ for all $a \in M$.

These requirement determine the functions g and g_f in Lemma 1.10 (the cases (ii) and (iv) are missing from above, exercise): We let A be the set of all finite sequences of symbols that appear in L-formulas, B is the set of all atomic L-formulas and F consists of functions $f_\neg$, $f_\to$ and f_n, $n \in \mathbb{N}$, where $f_\neg(w) = \neg w$, $f_\to(w, u) = (w \to u)$ and $f_n(w) = \forall v_n w$. Then $cl_A(B, F)$ is the set of all L-formulas. Now e.g. $g(t = u)$ is the function from assignments to $\{0, 1\}$, s.t. for all assingments s,

(a') $g(t = u)(s) = 1$ iff $t^M < s >= u^M < s >$

and if X is a function from assignments to $\{0, 1\}$, then $g_{f_\neg}(X)$ is the function from assignments to $\{0, 1\}$ s.t.

(b') $(g_{f_\neg}(X))(s) = 1 - X(s)$ for all assignments s.

Then our V_M is the function h from Lemma 1.10. By our choice of g and Lemma 1.10, $V_M(t = u)(s) = g(t = u)(s) = 1$ iff $t^M < s >= u^{M<s>}$ i.e. (a) holds. Also $(V_M(\neg\psi))(s) = (g_{f_\neg}(V_M(\psi)))(s) = 1 - (V_M(\psi))(s)$ i.e. (b) holds.

And now $M \models_s \phi$ if $(V_M(\phi))(s) = 1$ and Lemma 1.10 quarantees that Tarski's truth definition is well-made. After convincing us of this, we can forget V_M and apply Tarski's truth definition simply by using the truth conditions it states.

4.2 Exercise. *Show that $\mathcal{A} \models_s \phi$ according to Definition 4.1 above iff $s \in Sat_{\mathcal{A}}(\phi)$ where $Sat_{\mathcal{A}}(\phi)$ is as in Definition 4.8 from the lecture notes. Hint: By induction on the definition of a formula, i.e. use Lemma 1.3 from above.*

In Example 4.13, there is an error: The formula for $s(0)$ is prime is incorrect. E.g. $\neg\exists v_1 \exists v_2 (v_1 \times v_2 = v_0 \wedge \neg v_1 = v_0 \wedge \neg v_1 = 1) \wedge \neg v_0 = 1$ works.

We extend Definition 4.14 from the lecture notes a bit: We write $\Sigma \models \phi$ if for all $\mathcal{A}$ and s the following holds: If $\mathcal{A} \models_s \psi$ for all $\psi \in \Sigma$, then $\mathcal{A} \models_s \phi$. If $\Sigma = \{\psi\}$, we write just $\psi \models \phi$ and if $\Sigma = \emptyset$, we write $\models \phi$ and say that ϕ is valid.

We give recursive definitions to some concepts for which semi-heuristic definitions are given in the lecture notes. It is much easier to prove theorems starting from proper definitions.

We do not need notions bound and free occurrences of variables. But we need to know which variable are free in a formula.

4.3 Definition. *The set $FV(\phi)$ of free variable of a formula ϕ is defined as follows:*
(i) ϕ atomic: $v_i \in FV(\phi)$ if v_i appears in ϕ,
(ii) $\phi = \neg\psi$: $FV(\phi) = FV(\psi)$,
(iii) $\phi = \psi \to \theta$: $FV(\phi) = FV(\psi) \cup FV(\theta)$,
(iv) $\phi = \forall v_i \psi$: $FV(\phi) = FV(\psi) - \{v_i\}$.

We give also a recursive definition for FVF.

4.4 Definition. *$FVF(t, v_n, \phi)$ is defined as follows:*
(i) ϕ atomic: $FVF(t, v_n, \phi)$ holds always,
(ii) $\phi = \neg\psi$: $FVF(t, v_n, \phi)$ holds, if $FVF(t, v_n, \psi)$ holds,
(iii) $\phi = \psi \to \theta$: $FVF(t, v_n, \phi)$ holds, if both $FVF(t, v_n, \psi)$ and $FVF(t, v_n, \theta)$ hold,
(iv) $\phi = \forall v_i \psi$: $FVF(t, v_n, \phi)$ holds, if $v_n \notin FV(\phi)$ or $FVF(t, v_n, \psi)$ holds and v_i does not appear in t.

4.5 Exercise. *Show that $FVF(t, v_n, \phi)$ holds if $t = v_n$ or variables that appear in t do not appears in ϕ (e.g. t is a constant term i.e. no variable appears in t).*

Notice that if t is a constant term, then $t^M < s >$ does not depend on the interpretation s (exercise) and thus we sometimes write just t^M for $t^M < s >$ when t is a constant term.

We give also a recursive definition for substitution.

4.6 Definition. *For terms t and u, $u(t/v_i)$ is defined as follows:*
(i) $u = v_n$: $u(t/v_i) = t$ if $n = i$ and otherwise it is u,
(ii) $u = c$: $u(t/v_i) = u$,
(iii) $u = f(u_0, ..., u_n)$: $u(t/v_i) = f(u_0(t/v_i), ..., u_n(t/v_i))$.

4.7 Definition. *$\phi(t/v_i)$ is defined as follows:*
(i) $\phi = u_0 = u_1$: $\phi(t/v_i) = u_0(t/v_i) = u_1(t/v_i)$,
(ii) $\phi = R(u_0, ..., u_n)$: $\phi(t/v_i) = R(u_0(t/v_i), ..., u_n(t/v_i)$,
(iii) $\phi = \neg\psi$: $\phi(t/v_i) = \neg(\psi(t/v_i))$ (the additional brackets are there to indicate the order in which the operations are performed),
(iv) $\phi = \psi \to \theta$: $\phi(t/v_i) = \psi(t/v_i) \to \theta(t/v_i)$.
(v) $\phi = \forall v_n\psi$: If $v_i \notin FV(\phi)$, then $\phi(t/v_i) = \phi$ and otherwise $\phi(t/v_i) = \forall v_n(\psi(t/v_i))$.

We will prove Substitution lemma in the following restricted form. This will be good enough for us and a lot easier to prove.

4.8 Theorem. *Suppose t and u are terms, ϕ is a formula, $\mathcal{A} = (A, Sat_{\mathcal{A}})$ is a structure and $s : \mathbb{N} \to A$. Let $a = t^{\mathcal{A}} < s >$. Then*

$$u^{\mathcal{A}} < s(a/i) >= (u(t/v_i))^{\mathcal{A}} < s >$$

and if $FVF(t, v_i, \phi)$ holds, then

$$\mathcal{A} \models_{s(a/i)} \phi \quad \textit{iff} \quad \mathcal{A} \models_s \phi(t/v_i).$$

There is one place where we need simultanious substitution and that is in identity axioms. So for atomic formulas ϕ, distinct natural numbers $k_1, ..., k_n$ and terms $t_1, ..., t_n$ we define $\phi(t_1/v_{k_1}, ..., t_n/v_{k_n})$ as follows: Pick any m such that (it is strictly greater than any k_j, $1 \le j \le n$ and) if v_i appears in ϕ or in any t_j, $1 \le j \le n$, then $i < m$. Then we let

$$\phi(t_1/v_{k_1}, ..., t_n/v_{k_n}) = \phi(v_{k_1+m}/v_{k_1})...(v_{k_n+m}/v_{k_n})(t_1/v_{k_1+m})...(t_n/v_{k_n+m}).$$

This definition does not depend on the choice of m (exercise).

4.9 Exercise.

(i) Suppose t is a term, $n \in \mathbb{N}$ and ϕ is a formula. Show that if $v_n \notin FV(\phi)$, then $\phi(t/v_n) = \phi$ and $FVF(t, v_n, \phi)$ holds.

(ii) Suppose ϕ is an atomic formula, $t_0, ..., t_n$ are terms and $i_0, ..., i_n$ are distinct natural numbers. Show that for all structures M and assingments s, if we write $a_k = t_k^M < s >$ for $k \le n$, then $M \models_{s(a_0/i_0)...(a_n/i_n)} \phi$ iff $M \models_s \phi(t_0/v_{i_0}, ..., t_n/v_{i_n})$.

(iii) Suppose ϕ is a formula, $i_0, ..., i_n$ are distinct natural numbers and $t_0, ..., t_n$ are terms such that for all $i \le n$, if v_k appears in t_i, then v_k does not appear in ϕ, $v_k \notin \{v_{i_0}, ..., v_{i_n}\}$ and v_k does not appear in any t_j for $j \le n$, $j \ne i$ (e.g. $t_0, ..., t_n$ are constant terms). Show that for all structures M and assigments s, if we write $a_i = t_i^M < s >$, then $M \models_s \phi(t_0/v_{i_0})...(t_n/v_{i_n})$ iff $M \models_{s(a_0/i_0)...(a_n/i_n)} \phi$.

We make a small change to the definition of definability: In addition to

(*) $M \models_s \phi$ iff $(s(o), ..., s(n-1)) \in X$

we require that $FV(\phi) \subseteq \{v_0, ..., v_{n-1}\}$. This is just for technical convenience, it does not change what is definable: E.g. if ϕ satisfies (*) but $FV(\phi) = \{v_0, ..., v_n\}$, then $\psi = \forall v_n \phi$ still satisfies (*) and now $FV(\psi) \subseteq \{v_0, ..., v_{n-1}\}$ (exercise). In the literature this form of definability is known as definability without parameters. When people talk about definability, they usually mean definability with parameters. This is a different notion but if every element of the model is definable without parameters, the two notions are the same. This is the case with our $\mathcal{N}_{exp}$, see below or the lecture notes.

4.10 Exercise. *Let $M = (P(\mathbb{N}), R^M)$, where $P(\mathbb{N})$ is the power set of $\mathbb{N}$ and $R = \{(a, b) \in P(\mathbb{N})^2 | \ a \subseteq b\}$.*

(i) Show that $f(a, b) = a \cup b$ is definable in M.

(ii) What are the definable elements of M?

(iii) Find definable $X \subseteq P(\mathbb{N})$ such that both X and $P(\mathbb{N}) - X$ are infinite.

If A is a formula of propositional logic and ϕ is a formula then by $A(\phi/p_i)$ we mean the string of symbols that we get from A by replacing each occurrence of p_i by ϕ (exercise: give a recursive definition for $A(\phi/p_i)$). We say that a formula ψ is an axiom of propositional logic if there is a formula A of propositional logic that is an axiom of propositional logic and $n \in \mathbb{N}$ and formulas ϕ_i, $i \leq n$, such that $\psi = A(\phi_0/p_0)...(\phi_n/p_n)$ (notice that from this it follows that if p_j appears in A, then $j \leq n$). We say that a formula ψ is a tautology, if there is a tautology A of propositional logic, $n \in \mathbb{N}$ and formulas ϕ_i, $i \leq n$, such that $\psi = A(\phi_0/p_0)...(\phi_n/p_n)$.

Our main change is in the definition of deduction i.e. in Definitions 4.43 and 4.45 from the lecture notes. We make this change because the definitions from the lecture notes give a notion of proof that is very difficult to code as a natural and this is what we need to do later. Our version is not without problems, but these problems do not appear in the cases we are interested in, we will look these soon.

4.11 Definition. *For fixed vocabulary L, we define $\Sigma \vdash \phi$ as in Definition 4.43 from the lecture notes except that (T4) is replaced with the following:*

(T4) if $\Sigma \vdash \psi \to \theta$ and v_j is not free in ψ nor in any formula from Σ, then $\Sigma \vdash \psi \to \forall v_j \theta$.

Few remarks: Notice that this definition is for each vocabulary L separately and thus when this vocabulary plays a role we write $\vdash_L$ for $\vdash$. Also it is not immediately clear that by extending the vocabulary the set of L-formulas provable from Σ does not increase (it does not as we will see later). Finally, notice that our 4.11 defines by recursion the set of L-formulas provable from Σ. This is not the case with Definition 4.43 from the lecture notes, it defines the set of pairs (Σ, ϕ) such that $\Sigma \vdash \phi$. So although the definitions look similar, they are fundamentally different.

4.12 Exercise.

(i) Show that $A(\phi_0/p_0)...(\phi_n/p_n)$ is a formula for all formulas A of propositional logic, $n \in \mathbb{N}$ and formulas ϕ_i, $i \leq n$, if the following holds: if p_j appears in A, then $j \leq n$.

(ii) Show that $\Sigma \vdash \psi$ for all tautologies ψ and sets Σ of formulas.

One can always use Exercise 4.12 (ii) when one shows that some formula ϕ is provable from some set Σ. This is very useful.

We define $(\phi_0, ..., \phi_n)$ is a deduction from Σ exactly as in Definition 4.45 from the lecture notes except that we replace 6 with the following:

6. there are $k < i$, L-formulas ψ and θ and a natural number j such that

(i) $\phi_k = \psi \to \theta$,

(ii) $\phi_i = \psi \to \forall v_j \theta$,

(iii) v_i is not free in ψ nor in any formula from Σ.

Notice that item 6 in Definition 4.45 do not make any sense. The correct version of item 6 (for Definition 4.43 i.e. in the context of the lecture notes) makes the definition a recursive definition. The recursion goes on the length of the deduction.

4.13 Exercise. *Show that $\Sigma \vdash \phi$ iff there is a deduction $(\phi_0, ..., \phi_n)$ from Σ such that $\phi_n = \phi$. Conclude that if $\Sigma \vdash \phi$, then there is finite $\Sigma' \subseteq \Sigma$ such that $\Sigma' \vdash \phi$ (one can prove this also by induction).*

The problem with our notion $\vdash$ is that it is not transitive i.e. it is possible that $\Sigma \vdash \psi$ for all $\psi \in \Sigma'$ and $\Sigma' \vdash \phi$ but $\Sigma \not\vdash \phi$, in fact it is possible even that $\Sigma' \subseteq \Sigma$, $\Sigma' \vdash \phi$ but $\Sigma \not\vdash \phi$. This is ugly and the reason why the more complicated proof system is used in the lecture notes. However, as we will see, this does not happen when Σ is a theory i.e. a set of sentences and we are interested only in provability from theories. For theories, the two definitions of deduction are equivalent (exercise, use completeness).

Now to the example of non-transitivity: Let $L = \{E\}$ where E is a binary relation symbol. Let Σ be the set of all $\emptyset$-identity axioms and for each $\emptyset$-formula ϕ, let ϕ^* be the L-formula we get from ϕ by replacing each atomic subformula $v_i = v_j$ by $E(v_i, v_j)$ (for all i, j, exercise: give a recursive definition for ϕ^*). Let $\Sigma^* = \{\phi^* | \ \phi \in \Sigma\}$. From the following exercise it follows that $\emptyset \vdash \forall v_1 v_1 = v_1$ but $\Sigma \not\vdash \forall v_1 v_1 = v_1$.

4.14 Exercise.

(i) Show that if $\Sigma \vdash_{\emptyset} \forall v_1 v_1 = v_1$, then $\Sigma^* \vdash_L \forall v_1 E(v_1, v_1)$.

(ii) Show that $\Sigma^* \not\vdash_L \forall v_1 E(v_1, v_1)$. Hint: Use soundness, choose the structure so that it contains at least two elements and the assignment so that it is a constant function.

(iii) Show that $\emptyset \vdash_{\emptyset} \forall v_1 v_1 = v_1$.

4.15 Exercise.

(i) Suppose $\Sigma' \vdash \phi$ and Σ is a set of sentences. Show that $\Sigma \cup \Sigma' \vdash \phi$.

(ii) Let Σ be a set of L-sentences, Σ' be a set of L-formulas and ϕ be an L-formula. Suppose $\Sigma \vdash \psi$ for every $\psi \in \Sigma'$ and $\Sigma' \vdash \phi$. Show that $\Sigma \vdash \phi$. Hint:

Start by showing that we may assume that Σ' is a singleton and then use (i) and Deduction lemma.

We proved Lemma on constants in the following form (the proof is the same as the proof of the lecture notes version):

4.16 Lemma. *Suppose Σ is a set of L-formulas, $c \notin L$ is a constant symbol and ϕ is an $L \cup \{c\}$-formula. If $\Sigma \vdash_{L\cup\{c\}} \phi$, then there is $m \in \mathbb{N}$ such that for all $k \geq m$, $\Sigma \vdash_L \phi(v_k/c)$, where $\phi(v_k/c)$ is a formula got from ϕ by replacing c by v_k everywhere.*

In Theorem 4.57 (Deduction Lemma) in 'and conversely' i.e. in the direction if $\Sigma \vdash \psi \to \phi$, then $\Sigma \cup \{\psi\} \vdash \phi$, we need to assume that ψ is a sentence.

In order to avoid the use of set theory, we prove the completeness theorem only for countable vocabularies. Thus, essentially from the beginning of Subsection 4.4 on, we assume that the vocabulary is countable.

We define that an L-theory Σ is L-inconsistent if there is an L-sentence ϕ such that $\Sigma \vdash_L \phi$ and $\Sigma \vdash_L \neg\phi$. If there is no such ϕ, we say that Σ is L-consistent.

4.17 Exercise. *Suppose that Σ is a set of L-formulas, ϕ is an L-formula and $c \notin L$ is a constant symbol. Show that if $\Sigma \vdash_{L\cup\{c\}} \phi$, then $\Sigma \vdash_L \phi$. Conclude that if Σ is $L \cup \{c\}$-inconsistent, then it is L-inconsistent. Hint: Lemma on constants.*

Now we can write the Chain lemma in the following form:

4.18 Lemma. *Suppose that for all $n \in \mathbb{N}$, Σ_n is an L_n-consistent L_n-theory. In addition suppose that for all $n \in \mathbb{N}$, $\Sigma_n \subseteq \Sigma_{n+1}$ and $L_n \subseteq L_{n+1}$. Let $\Sigma = \cup_{n=0}^{\infty}\Sigma_n$ and $L = \cup_{n=0}^{\infty}L_n$. Then Σ is L-consistent.*

Proof. For a contradiction, let $(\phi_i)_{i\leq m}$ and $(\psi_i)_{i\leq k}$ be deductions from Σ for ϕ and $\neg\phi$ for some L-sentence ϕ. It is easy to see that there is $p \in \mathbb{N}$ such that every ϕ_i is L_p-formula and if $\phi_i \in \Sigma$, then it is in Σ_p and the same for formulas ψ_i. But then Σ_p is L_p-inconsistent, a contradiction. □

Theorem 4.68: We say that an L-theory Σ is Henkin if for all L-sentences of the form $\forall v_n\psi$ there is a constant $c \in L$ such that $\psi(c/v_n) \to \forall v_n\psi$ belongs to Σ. Then we call this sentence $\psi(c/v_n) \to \forall v_n\psi$ a Henkin axiom (a bit misleading name). Then in Theorem 4.68 we assume that ϕ and ψ are sentences and add a third item:
3. If in addition Σ is Henkin and $\forall v_n\theta$ is an L-sentence, then $\Sigma \vdash \forall v_n\theta$ iff for all constant terms t, $T \vdash \theta(t/v_n)$.

To the proof of Theorem 4.70 we make many corrections:

(i) In Claim 2 we assume that $t_1, ..., t_n$ and $t'_1, ..., t'_n$ are constant terms and replace $Rt_1...t_n \in \Sigma$ and $Rt'_1...t'_n \in \Sigma$ by $\Sigma \vdash R(t_1, ..., t_n)$ and $\Sigma \vdash R(t'_1, ..., t'_n)$, respectively.

(ii) In Claim 3 we again assume that $t_1, ..., t_n$ and $t'_1, ..., t'_n$ are constant terms and replace $\approx ft_1...t_n ft'_1...t'_n \in \Sigma$ by $\Sigma \vdash f(t_1, ..., t_n) = f(t'_1, ..., t'_n)$.

(iii) Claim 4 should be: For all constant L-terms t and L-formulas ϕ the folowing holds: $t^M < s >= [t]$ for all assignments s and if $v_{k_1}, ..., v_{k_n}$ lists the free variables of ϕ (without repetition) and $t_1, ..., t_n$ are constant L-terms, then $M \models \phi(t_1/v_{k_1})...(t_n/v_{k_n})$ iff $\Sigma \vdash \phi(t_1/v_{k_1})...(t_n/v_{k_n})$.

(iv) The induction step in the proof of the first claim should be:

$$f(u_1, ..., u_m)^M < s >= f^M(u_1^M < s >, ..., u_m^M < s >) =$$

$$f^M([u_1], ..., [u_m]) = [f(u_1, ..., u_m)].$$

We recall that since, for constant terms t, $t^M < s >$ does not depend on s, we can write simply t^M for $t^M < s >$. We also notice that the formulas $\phi(t_1/v_{k_1})...(t_n/v_{k_n})$ from Claim 4 are sentences (exercise: by induction on ϕ, show that if t is a constant term then v_n is not free in $\phi(t/v_n)$).

(v) Item 1 i.e. the case when ϕ is $u_1 = u_2$: We write $u'_1 = u_1(t_1/v_{k_1})...(t_n/v_{k_n})$ and similarly for u_2. Notice that u'_1 and u'_2 are constant terms (exercise). Then

$$M \models (u_1 = u_2)(t_1/v_{k_1})...(t_n/v_{k_n}) \Leftrightarrow M \models u'_1 = u'_2 \Leftrightarrow$$

$$(u'_1)^M = (u'_2)^M \Leftrightarrow [u'_1] = [u'_2] \Leftrightarrow$$

$$u'_1 \sim u'_2 \Leftrightarrow \Sigma \vdash u'_1 = u'_2 \Leftrightarrow \Sigma \vdash (u_1 = u_2)(t_1/v_{k_1})...(t_n/v_{k_n}).$$

and similar correction to item 2.

(vi) Item 3 should be (this is the case $\phi = \neg\psi$):

$$M \models (\neg\psi)(t_1/v_{k_1})...(t_n/v_{k_n}) \leftrightarrow M \not\models \psi(t_1/v_{k_1})...(t_n/v_{k_n}) \Leftrightarrow$$

$$\Sigma \not\vdash \psi(t_1/v_{k_1})...(t_n/v_{k_n}) \Leftrightarrow \Sigma \vdash (\neg\psi)(t_1/v_{k_1})...(t_n/v_{k_n}).$$

And similar correction to item 4.

(vii) Item 5 should be (this is the case $\phi = \forall v_i \psi$): We write

$$\psi' = \psi(t_1/v_{k_1})...(t_n/v_{k_n})$$

and we may assume that $i \neq k_j$ for all $1 \leq j \leq n$. Then $M \models \phi(t_1/v_{k_1})...(t_n/v_{k_n})$ iff for all $a \in M$, $M \models_{s(a/i)} \psi'$ iff for all constant terms t, $M \models_{s([t]/i)} \psi'$ iff for all constant terms t, $M \models \psi'(t/v_i)$ iff for all constant terms t, $\Sigma \vdash \psi'(t/v_i)$ iff $\Sigma \vdash \forall v_i \psi'$ iff $\Sigma \vdash \phi(t_1/v_{k_1})...(t_n/v_{k_n})$ (the first and the last equivalences hold because $\phi(t_1/v_{k_1})...(t_n/v_{k_n}) = \forall v_i \psi'$, the third is substitution lemma, fourth is the induction assumption and the fifth is the item 3 that we added above to Theorem 4.68).

In Theorem 4.73 we need to assume that Σ is a set of L-sentences, not just L-formulas.

5. Incompleteness of number theory

Following what we have said above we write simply $+$, $\times$ and exp for $\oplus$, $\otimes$ and $\underline{exp}$. In this section, the author of the lecture notes uses exponentiation to code sequences of symbols as natural numbers. He does this to avoid a more complicated coding based on chinese remainder theorem. But this has consequences: now our standard model of number theory must contain exponentiation (or we need to show that it is definable from $+$ and $\times$, which it is by chinese remainder theorem). Because of this we can not prove the results for Peano axioms, we need to add to them axioms for exponentiation that determine it in the standard model. So our vocabulary for number theory is $L_{\text{exp}} = \{+, \times, exp, 0, 1\}$ and we add axioms:

(P8) $\forall v_0 exp(v_0, 0) = 1$

(P9) $\forall v_0 \forall v_1 exp(v_0, v_1 + 1) = exp(v_0, v_1) \times v_0$.

We write P_{exp} for the set of Peano axioms together with these two new axioms (P8) and (P9).

Below, if $n = 0$, by $f : \mathbb{N}^n \to \mathbb{N}$ we mean a '0-ary function' i.e. a natural number.

5.1 Lemma. *If $f : \mathbb{N}^n \to \mathbb{N}$ and $g : \mathbb{N}^{n+2} \to \mathbb{N}$ are functions then there is a unique function $h : \mathbb{N}^{n+1} \to \mathbb{N}$ such that*

$$h(0, x_1, ..., x_n) = f(x_1, ..., x_n)$$

and

$$h(y+1, x_1, ..., x_n) = g(y, h(y, x_1, ..., x_n), x_1, ..., x_n).$$

Proof. Let $A = \mathbb{N}^{n+2}$, $B = \{(0, x_1, ...x_n, f(x_1, ..., x_n)) | \; x_1, ..., x_n \in \mathbb{N}\}$ and $F = \{s\}$, where $s : A \to A$ is such that

$$s((y, x_1, ..., x_n, z)) = (y+1, x_1, ..., x_n, g(y, z, x_1, ..., x_n)).$$

Then for all $y, x_1, ..., x_n \in \mathbb{N}$, there is a unique $z \in \mathbb{N}$ such that $(y, x_1, ..., x_n, z) \in cl_A(B, F)$ (exercise) and so $h = cl_A(B, F)$ is a function $\mathbb{N}^{n+1} \to \mathbb{N}$. It is easy to see that h satisfies the requirements (exercise).

If also $h' : \mathbb{N}^{n+1} \to \mathbb{N}$ is such that

$$h'(0, x_1, ..., x_n) = f(x_1, ..., x_n)$$

and

$$h'(y+1, x_1, ..., x_n) = g(y, h'(y, x_1, ..., x_n), x_1, ..., x_n),$$

then an easy induction on y shows that $h' = h$ (exercise). □

In Example 5.8 we make some small (but useful) changes:

8: If $f : \mathbb{N}^{n+1} \to \mathbb{N}$ and $g : \mathbb{N}^n \to \mathbb{N}$ are p.r., then also

$$h(x_1, ..., x_n) = \Sigma_{i=0}^{g(x_1, ..., x_n)} f(i, x_1, ..., x_n)$$

and

$$h(x_1,...,x_n)=\Pi_{i=0}^{g(x_1,...,x_n)}f(i,x_1,...,x_n)$$

are p.r. functions.

9: If $R\subseteq \mathbb{N}^{n+2}$ and $g:\mathbb{N}^{n+1}\rightarrow\mathbb{N}$ are p.r., then also

$$S=\{(x_0,...,x_n)\in\mathbb{N}^{n+1}|\ \forall z\le g(x_0,...,x_n)((z,x_0,...,x_n)\in R)\}$$

and

$$S'=\{(x_0,...,x_n)\in\mathbb{N}^{n+1}|\ \exists z\le g(x_0,...,x_n)((z,x_0,...,x_n)\in R)\}$$

are p.r.

10: For $y,x_0,...,x_n\in\mathbb{N}$, we write $\mu z\le y((z,x_0,...,x_n)\in R)$ for the least $z\in\mathbb{N}$ such that $z\le y$ and $(z,x_0,...,x_n)\in R$, if there is such z and otherwise $\mu z\le y((z,x_0,...,x_n)\in R)=0$. Then if $R\subseteq\mathbb{N}^{n+2}$ and $g:\mathbb{N}^{n+1}\rightarrow\mathbb{N}$ are p.r., then the function $f:\mathbb{N}^{n+1}\rightarrow\mathbb{N}$,

$$f(x_0,...,x_n)=\mu z\le g(x_0,...,x_n)((z,x_0,...,x_n)\in R),$$

is p.r.

We make some changes to the proof of the claim from the proof of Theorem 5.15 that $f(n)=p_n$ is recursive (the proof in the lecture notes is correct but I find it hard to follow): Let $R\subseteq\mathbb{N}^2$ be such that $(z,n)\in R$ iff $z=\Pi_{i=0}^{n}p_i^{i+1}$. We show first that R is recursive: Now $(z,n)\in R$ iff the following holds (exercise):

(i) 2 divides z but 4 does not (so $z\neq 0$),

(ii) for all $p,q\le z$ if $p<q$ and both are primes then the following holds:

(a) if p^{n+1} divides z then q does not divide z

(b) if p^{n+1} does not divide z and there is no prime r such that $p<r<q$, then for all $i\le n$, p^{i+1} divides z iff q^{i+2} divides z.

Thus R is recursive. Let $g(n)=\mu z((z,n)\in R)$. Then $f(n)$ is the least prime p such that p^{n+1} divides $g(n)$ and thus f is recursive.

After Theorem 5.15, we conclude that p.r. functions are recursive.

I find proofs around Ackermann function instructing and thus we take a closer look at them.

5.2 Lemma. *Ackermann function* $A:\mathbb{N}^2\rightarrow\mathbb{N}$ *exists and is unique.*

Proof. We prove the existence first: By recursion on y we construct functions $f_y:\mathbb{N}\rightarrow\mathbb{N}$ as follows: $f_0=S$ and if we have f_y, f_{y+1} is defined by primitive recursion as follows: $f_{y+1}(0)=f_y(1)$ and $f_{y+1}(x+1)=f_y(f_{y+1}(x))$. Then $A(y,x)=f_y(x)$ satisfies the equations from the definition of Ackermann function.

Then to the uniqueness: Suppose that A' also satisfies the equations. Let $f'_y(x)=A'(y,x)$. It is enough to show that for all $y\in\mathbb{N}$, $f_y=f'_y$. We prove this by induction on y: If $y=0$, the claim is clear. So suppose that $f_y=f'_y$. By induction on $x\in\mathbb{N}$, we show that $f_{y+1}(x)=f'_{y+1}(x)$: $f_{y+1}(0)=f_y(1)=f'_y(1)=f'_{y+1}(0)$ and if $f_{y+1}(x)=f'_{y+1}(x)$, then $f_{y+1}(x+1)=f_y(f_{y+1}(x))=f'_y(f'_{y+1}(x))=f'_{y+1}(x+1)$. □

5.3 Exercise.

(i) Show that $A(1,x)=x+2$ *and* $A(2,x)=2x+3$.

(ii) Show that $y+x<A(y,x)<A(y,x+1)\leq A(y+1,x)$ *(thus* A *is increasing in both arguments).*

(iii) Show that for all $y,x\in\mathbb{N}$, *there is finite* $X_{yx}\subseteq\mathbb{N}^2$ *such that*

(a) $(y,x)\in X_{yx}$,

(b) for all $y'\in\mathbb{N}$, *if* $(y'+1,0)\in X_{yx}$, *then* $(y',1)\in X_{yx}$,

(c) for all $y',x'\in\mathbb{N}$, *if* $(y'+1,x'+1)\in X_{yx}$, *then* $(y'+1,x')\in X_{yx}$ *and* $(y',A(y'+1,x'))\in X_{yx}$.

5.4 Proposition. *A is recursive.*

Proof. Notice that for all $x,y\in\mathbb{N}$, $A(y,x)\neq 0$ by Exercise 5.3 (ii). Notice also that if $(z)_i\neq 0$, then $z\neq 0$ and that if ($z\neq 0$ and) p_i does not divide z, then $(z)_i=0$. Let $R\subseteq\mathbb{N}^3$ be such that $(z,y,x)\in R$ iff

(i) $(z)_{\pi(y,x)}\neq 0$,

(ii) for all $x'\leq z$, if $(z)_{\pi(0,x')}\neq 0$, then $(z)_{\pi(0,x')}=x'+1$,

(iii) for all $y'\leq z$, if $(z)_{\pi(y'+1,0)}\neq 0$, then $(z)_{\pi(y',1)}\neq 0$ and $(z)_{\pi(y'+1,0)}=(z)_{\pi(y',1)}$,

(iv) for all $y',x'\leq z$, if $(z)_{\pi(y'+1,x'+1)}\neq 0$, then $(z)_{\pi(y'+1,x')}\neq 0$, $(z)_{\pi(y',(z)_{\pi(y'+1,x')})}\neq 0$ and $(z)_{\pi(y'+1,x'+1)}=(z)_{\pi(y',(z)_{\pi(y'+1,x')})}$.

Clearly R is recursive and for all $y,x\in\mathbb{N}$, there is $z\in\mathbb{N}$ such that $(z,y,x)\in R$: Just let

$$z=\Pi_{(y',x')\in X_{yx}}p_{\pi(y',x')}^{A(y',x')+1},$$

where X_{yx} is as in Exercise 5.3 (iii).

Then $A(y,x)=(\mu z((z,y,x)\in R))_{\pi(y,x)}$ and thus A is recursive. □

5.5 Lemma. *Suppose* $f:\mathbb{N}^{n+1}\to\mathbb{N}$ *is p.r. Then there is* $m^*\in\mathbb{N}$ *such that for all* $x_0,...,x_n\in\mathbb{N}$, $f(x_0,...,x_n)<A(m^*,w)$ *where* $w=max\{x_0,...,x_n\}$.

Proof. By induction on the definition of p.r. functions:

(1) f is Z, S or Pr_i^{n+1}: Let $m^*=1$. Since $A(1,z)=z+2$ the claim is clear.

(2) $f=g(h_1(x_0,...,x_n),...,h_k(x_0,...,x_n))$: By induction assumption, there is m suxh that $g(y_1,...,y_k)<A(m,max\{y_1,...,y_k\})$ and for all $1\leq i\leq k$, $h_i(x_0,...,x_n)<A(m,max\{x_0,...,x_n\})$. Then

$$f(x_0,...,x_n)<A(m,A(m,max\{x_0,...,x_n\}))\leq A(m,A(m+1,max\{x_0,...,x_n\}))\leq$$

$$A(m+1,max\{x_0,...,x_n\}+1)\leq A(m+2,max\{x_0,...,x_n\}).$$

So we can let $m^*=m+2$.

(3) $f(0,x_1,...,x_n)=g(x_1,...,x_n)$ and

$$f(y+1,x_1,...,x_n)=h(y,f(y,x_1,...,x_n),x_1,...,x_n):$$

Again by induction assumption, there is m such that for all $x_1, ..., x_n$, $g(x_1, ..., x_n) < A(m, max\{x_1, ..., x_n\})$ and for all $y, z, x_1, ..., x_n$,

$$h(y, z, x_1, ..., x_n) < A(m, max\{y, z, x_1, ..., x_n\}).$$

Let $u = max\{x_1, ..., x_n\}$. We show first the following claim:

5.5.1 Claim. $f(y, x_1, ..., x_n) < A(m+1, y+u)$.

Proof. By induction on y. The case $y = 0$ is clear by the choice of m. We prove the claim for $y+1$, notice that $A(m+1, y+u) > max\{y, f(y, x_1, ..., x_n), x_1, ..., x_n\}$:

$$f(y+1, x_1, ..., x_n) = h(y, f(y, x_1, ..., x_n), x_1, ..., x_n) <$$

$$A(m, max\{y, f(y, x_1, ..., x_n), x_1, ..., x_n\}) <$$

$$A(m, A(m+1, y+u)) = A(m+1, (y+1)+u).$$

□ Claim 5.5.1.

Now with the claim and the observation that $w = max\{y, u\}$, we can proceed as follows:

$$f(y, x_1, ..., x_n) < A(m+1, y+u) \leq A(m+1, A(2, w)) \leq$$

$$A(m+1, A(m+2, w)) = A(m+2, w+1) \leq A(m+3, w).$$

So we can let $m^* = m+3$. □

5.6 Proposition. *A is not p.r.*

Proof. It is enough to prove that $A^*(x) = A(x, x)$ is not p.r. For a contradiction suppose A^* is p.r. Then by Lemma 5.5, there is m^* such that for all $x \in \mathbb{N}$, $A^*(x) < A(m^*, x)$. But now

$$A^*(m^*) < A(m^*, m^*) = A^*(m^*),$$

a contradiction. □

5.7 Exercise. *Show that $\{(y, x, w) \in \mathbb{N}^3 |\ w = A(y, x)\}$ is p.r. Hint: Find p.r. $f : \mathbb{N}^3 \to \mathbb{N}$ such that $f(y, x, A(y, x))$ is larger than z from the proof of Proposition 5.4 for (y, x).*

In the proof of Theorem 5.19, there is nothing in the definition of R that guarantees that $z+1$ is a Gödel number, it allows all codes of terms. We can e.g. define a p.r. function $G : \mathbb{N} \to \mathbb{N}$ so that

$$G(z) = \mu u \leq z(len(u) = len(z) \wedge \forall i \leq len(z)((u)_i = (z)_i))$$

(e.g. $G(2^5 5^7) = 2^5$) and by replacing R with the relation $(z,u) \in R'$ iff $(z,u) \in R \wedge G(z+1) = z+1$, we get a definition for Trm that accepts only Gödel numbers.

Theorem 5.32 should say that P_{exp} is incomplete i.e. there is an L_{exp}-sentence ϕ such that $P_{exp} \nvdash \phi$ and $P_{exp} \nvdash \neg\phi$. Also in Theorem 5.30, we let Thm be the set of all Gödel numbers of L_{exp}-sentences ϕ such that $P_{exp} \vdash \phi$. This is because in Corollary 5.31, Tr should be as in Theorem 5.23 and there it is necessary that the sentences are L_{exp}-sentences because the sentence from Gödel's fixed point theorem is an L_{exp}-sentence and not an L-sentence ($L = \{+, \times, 0, 1\}$). And now if Thm is defined as in the lecture notes, the proof of Theorem 5.32 fails because, although $Tr - Thm \neq \emptyset$, we do not know that it contains L-sentences. And if we let Thm be the set of all Gödel numbers of L_{exp}-sentences ϕ such that $P \vdash \phi$, Theorem 5.32 becomes trivial, obviously e.g. $exp(1,1) = 1$ is true but it cannot be proved from P because P says nothing about exp (this follows from soundness!).

And then, obviously, Theorem 5.33 should say that the consistency of P_{exp} is not provable from P_{exp} i.e. $Con(P_{exp})$ is a true sentence such that $P_{exp} \nvdash Con(P_{exp})$.