# Proofs as Programs in Classical Logic Notes

Alexander Pluska

November 8, 2021

## 1 Type theory

As we are interested in a proof-as-programs correspondence for first order logic our underlying type theory has to be of dependant nature, meaning that types can and will depend on terms.

One interesting aspect of choosing a type theory is how to handle equality. There are many different ways to do this but for our purposes the relatively simple model in agda [7] is sufficient (?). Namely the only equality inherent to the type theory is definitional equality, i.e. equality induced by equivalence under  $\beta$ -reductions and  $\eta$ -equivalences whereas propositional equality is itself an inductive type (or rather inductive schema) with a single constructor

$$\operatorname{refl}_x : x =_A x$$

(the type is parametrized by a type A and x:A), i.e. for all types A and x:A we have the rule

$$\frac{}{\operatorname{refl}_x : x =_A x} = -\operatorname{intro}$$

Many of the expected properties of equality hold. For instance we can derive the term

subst : 
$$\Pi_A \Pi_{x,y:A} \Pi_{P:A \to Type} \Pi_{e:x=Ay} \Pi_{p:Px} Py$$
  
subst  $A x y P \operatorname{refl}_x px = px$ 

by pattern matching, i.e. matching refl the type checker can unify x and y.

However there are some issues with this handling of equality. For example it is impossible to define quotient structures: Let's say we prove prime decomposition for PIDs and want to apply it to some ring. Then the ring will be most likely defined as a setoid-like structure with the equality being a more coarse relation on the set of terms and in particular not equal to the equality above. Our proof-as-programs translation will have to take this into account.

# 2 Proof translation for saturation-based theorem proving with induction

Most modern theorem provers for first order logic utilize resolution and saturation based techniques. The standard intuitionistic proofs-as-programs correspondence fails as even the very method itself crucially depends on double negation elimination. As a first step we work out a way to convert proofs of ∀∃-proofs based on the calculus presented in [2] into programs. An advantage of the calculus as presented in [2] is that it can (at least in theory) handle general induction for inductive datatype. Many advances have recently been made for incorporating induction in existing theorem provers [4][5] however all of the calculi fail to support induction in the general case, in particular sentences involving existential quantifiers and requiring induction depth greater than 1 are problematic.

The approach in [2] seems promising as, despite being able to handle general induction, there is no explosion in the number of clauses due to a clever use of so-called *constrained clauses*, special symbols  $\mathbf{T}_{\alpha}^{\prec}$  for induction and a special activation mechanism for potential inductive invariants, although how it performs in practice is yet to be seen.

Program translation for classical proofs has been outlined in [3]. Based on this [1] and [?] have been developed. Our a approach is a combination of both. At the moment it is quite crude and is mostly based on the exception method from [1]. We aim to later translate this to the CPS approach presented in the same paper to better see how it relates to the normal proofs-as-program correspondence known from intuitionistic logic.

#### 2.1 The Calculus

As explained above our calculus is based on [2]. We take the core rules, which are not mentioned in [2], from [6].

The first five inferences are standard in resolution based theorem proving. We assume some simplification ordering  $\succeq^*$ .

Resolution.

$$\frac{A \vee C_1 \qquad \neg B \vee C_2}{(C_1 \vee C_2)\theta}$$

where  $\theta$  is a mgu of A and  $\neg B$ .

Factoring.

$$\frac{A \vee \neg B \vee C}{(A \vee C)\theta}$$

where theta is a mgu of A and B.

Superposition.

$$\frac{l=r\vee C_1}{(L[r]\vee C_1\vee C_2)\theta}\frac{l=r\vee C_1}{(t[r]=t'\vee C_1\vee C_2)\theta}\frac{l=r\vee C_1}{(t[r]\neq t'\vee C_1\vee C_2)\theta}\frac{l=r\vee C_1}{(t[r]\neq t'\vee C_1\vee C_2)\theta}$$

where  $\theta$  is a mgu of l and s, s is not a variable,  $r\theta \not\succeq^* l\theta$ , L[s] is not an equality literal, and  $t'\theta \succeq^* t[s]\theta$ .

Equality Resolution.

$$\frac{s \neq t \vee C}{C\theta}$$

where  $\theta$  is a mgu of s and t.

Equality Factoring.

$$\frac{s = t \vee s' = t' \vee C}{(s = t \vee t = t' \vee C)\theta}$$

where  $\theta$  is an mgu of s and s',  $t\theta \succeq^* s\theta$ , and  $t'\theta \succeq^* t\theta$ .

In addition to regular clauses and inference rules we introduce c-clauses and special rules for them, i.e. (all of this will be made more precise later)

**Definition 2.1.** A constrained clause (c-clause) is an expression of the form  $[\![C \mid \mathcal{X}]\!]$ , where

- C is a clause,
- $\mathcal{X}$  is a conjunction of the form  $\bigwedge_{i=1}^n f_i(\mathbf{u_i}) \simeq v_i \wedge \bigwedge_{i=1}^m \mathbf{T}_{\alpha_i}^{\prec}(z_i, \mathbf{w}_i)$ , where  $n, m \geq 0$  and  $\forall i \in \{1 \dots n\}$ ,  $f_i$  are uninterpreted function symbols (possibly nullary).

The semantics of  $[\![C \mid \mathcal{X}]\!]$  is more or less "if all constraints in  $\mathcal{X}$  hold, then C holds" and of  $\mathbf{T}_{\alpha_i}^{\prec}(z, \mathbf{w})$  "for all  $z' \prec z$   $\alpha(z', \mathbf{w})$  holds" where  $\prec$  is the standard wfo of an inductive type.

All of the above inference rules extend naturally to constrained clauses, i.e. if  $\frac{H_1, \dots, H_n}{C\theta}$  is an inference rule for regular clauses, then

$$\frac{\llbracket H_1 \mid \mathcal{X}_1 \rrbracket, \dots, \llbracket H_1 \mid \mathcal{X}_1 \rrbracket}{\llbracket C \mid \mathcal{X}_1 \wedge \dots \wedge \mathcal{X}_n \rrbracket \theta}$$

is the corresponding inference rule for c-clauses.

Furthermore we have some special rules for c-clauses:

Constraint Factorization.

$$\frac{\llbracket C \mid l_1 \wedge l_2 \wedge \mathcal{X} \rrbracket}{\llbracket C \mid l_1 \wedge \mathcal{X} \rrbracket \theta}$$

where  $\theta$  is an mgu of  $l_1$  and  $l_2$ .

Abstraction.

$$\frac{ \llbracket C[f(\mathbf{t})]_p \mid \mathcal{X} \rrbracket }{ \llbracket C[x]_p \mid f(\mathbf{t}) \simeq x \wedge \mathcal{X} \rrbracket }$$

where f is a uninterpreted function symbol, x is a fresh variable.

Instantiation.

$$\frac{C \mid \mathcal{X} \wedge t \simeq s]}{\llbracket C \mid \mathcal{X} \rrbracket \theta}$$

where  $\theta$  is a mgu of t and s.

In addition there are some rules that generate candidates for an inductive invariant in [2]. We will leave those out for now an choose our variant "magically" by hand. For every formula  $\alpha[z, \mathbf{w}]$  we than have axiom

$$\Gamma_{\mathbf{T}_{\alpha}} := [\![z' \not\prec z \lor \alpha(z', \mathbf{w}) \mid \mathbf{T}_{\alpha}^{\prec}(z, \mathbf{w})]\!]$$

and  $\prec$  is axiomatized by

$$\Gamma_{\prec}^s := [x_i \prec s(x_1, \dots, x_n) \mid \top] \qquad \Gamma_{\prec} := [x \not\prec y \lor y \not\prec z \lor x \prec z \mid \top]$$

for every n-ary inductive constructor  $s, i \in \{1 \dots n\}$ .

Finally we have additional rules for inductive types. Since we are (for now) not interested in how to generate inductive invariants it is possible to present just one combination rule for Domain Decomposition and Induction.

#### Induction.

$$\frac{\llbracket C_1 \mid \beta_1(t_1) \wedge \mathcal{X}_1 \rrbracket \dots \llbracket C_n \mid \beta_n(t_n) \wedge \mathcal{X}_n \rrbracket}{\llbracket C_1 \vee \dots \vee C_n \vee \alpha(x, \mathbf{w}_1) \mid \mathcal{X}_1 \setminus \{\mathbf{T}_{\alpha}^{\prec}(t_1, \mathbf{w}_1)\} \wedge \dots \wedge \mathcal{X}_n \setminus \{\mathbf{T}_{\alpha}^{\prec}(t_n, \mathbf{w}_n)\} \wedge \beta_1(x) \rrbracket \theta}$$

where  $\theta$  is the most general idempotent substitution such that  $\beta_j \theta \subseteq \beta_1 \theta$  for  $n \in \{2...n\}$  and a mgu for  $\mathbf{w}_1, ..., \mathbf{w}_n$ , the variables occurring in  $t_1, ..., t_n$  do not occur in  $C_i, \beta_i, \mathbf{w}_i$  or  $\mathcal{X}_i$ .  $\{t_1...t_n\}$  is covering, i.e. the whole type can be constructed using  $\{t_1...t_n\}$ , for instance for nat both  $\{z, s(x)\}$  and  $\{z, s(z), s(s(x))\}$  are covering and x is some fresh variable.

Note that setting  $\alpha = \bot$  and  $\mathbf{w_i} = \emptyset$  we obtain the regular Domain Decomposition Rule from [2].

### 2.2 Type system

Since we want to apply our method to first-order resolution proofs we will require a dependant type system. For this purpose we shall use the practical type system of agda as described in [7].

### 2.3 Examples and Translations

#### Example 2.2.

```
1. \forall x_0 : Nat(z + x_0 = x_0)
                                                                                                              Axiom
 2.\forall x_0x_1: Nat(s(x_0+x_1)=s(x_0)+x_1)
                                                                                                              Axiom
 3. \neg \forall x_0 : Nat \exists x_1 : Nat(x_1 + z = x_0)
                                                                                                     negated claim
 4.(\forall x_0 : Nat \exists x_1 : Nat, x_1 + z = x_0) \lor (\forall x_0 : Nat, z \neq x_0 + z) \lor
   \exists x_0 : Nat(\forall x_1 : Nat, x_1 + z \neq s(x_0) \land \exists x_1 : Nat, x_1 + z = x_0)
                                                                                      Induction axiom (ENNF)
 5.\exists x_0 : Nat \forall x_1 : Nat(x_1 + z \neq x_0)
                                                                                        ENNF\ transformation\ 3
 6.\forall x_1: Nat(x_1+z\neq sk_0)
                                                                                                   skolemization 5
 7.(\forall x_0 : Nat, sk_1(x_0) + z = x_0) \lor (\forall x_0 : Nat, z \neq x_0 + z) \lor
   (\forall x_1 : Nat, x_1 + z \neq s(sk_2) \land sk_3 + z = sk_2)
                                                                                                   skolemization 4
 8.x_1 + z \neq sk_0
                                                                                                                cnf 6
 9.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor sk_2 = sk_3 + z
                                                                                                                cnf 7
10.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor x_2 + z \neq s(sk_2)
                                                                                                                cnf 7
11.s(x_0 + x_1) = s(x_0) + x_1
                                                                                                                cnf 2
12.z + x_0 = x_0
                                                                                                                cnf 1
13.sk_1(x_0) + z = x_0 \lor z \neq z \lor sk_2 = sk_3 + z
                                                                                              superposition 9, 12
14.sk_1(x_0) + z = x_0 \lor sk_2 = sk_3 + z
                                                                                   trivial inequality removal 13
15.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor s(x_3 + z) \neq s(sk_2)
                                                                                            supserposition 10, 11
16.sk_0 \neq x_0 \lor sk_2 = sk_3 + z
                                                                                              superposition 8, 14
17.sk_2 = sk_3 + z
                                                                                            equality resolution 16
18.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor s(sk_2) \neq s(sk_2)
                                                                                             superposition 15, 17
19.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z
                                                                                   trivial inequality removal 18
20.sk_1(x_0) + z = x_0 \lor z \neq z
                                                                                             superposition 12, 19
21.sk_1(x_0) + z = x_0
                                                                                   trivial inequality removal 20
22.sk_0 \neq x_0
                                                                                               superoisition 21, 8
23.\bot
                                                                                            equality resolution 22
```

#### Example 2.3.

```
1. \vdash \lambda f. f(\lambda x. id_x) : Cont \Pi_{x_0:Nat}(z + x_0 = x_0)
                                                                                                                                     Axiom
 2. \vdash \lambda f. f(\lambda x_0 x_1. id_{s(x_0 + x_1)}) : Cont \Pi_{x_0: Nat} \Pi_{x_1: Nat} (s(x_0 + x_1) = s(x_0) + x_1)
                                                                                                                                     Axiom
 3. \vdash f: (\prod_{x_0: Nat} \Sigma_{x_1: Nat} x_1 + z = x_0) \rightarrow \bot
                                                                                                                           negated claim
 4. \vdash \lambda f. : Cont (\Pi_{x_0:Nat} \Sigma_{x_1:Nat} x_1 + z = x_0) \lor (\Pi_{x_0:Nat} z \neq x_0 + z) \lor
   \sum_{x_0: Nat} (\prod_{x_1: Nat} x_1 + z \neq s(x_0) \wedge \sum_{x_1: Nat} x_1 + z = x_0)
                                                                                                           Induction axiom (ENNF)
 5.\exists x_0 : Nat \forall x_1 : Nat(x_1 + z \neq x_0)
                                                                                                              ENNF transformation 3
 6.\forall x_1: Nat(x_1+z\neq sk_0)
                                                                                                                         skolemization 5
 7.(\forall x_0 : Nat, sk_1(x_0) + z = x_0) \lor (\forall x_0 : Nat, z \neq x_0 + z) \lor
   (\forall x_1 : Nat, x_1 + z \neq s(sk_2) \land sk_3 + z = sk_2)
                                                                                                                         skolemization 4
 8.x_1 + z \neq sk_0
                                                                                                                                       cnf 6
 9.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor sk_2 = sk_3 + z
                                                                                                                                       cnf 7
10.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor x_2 + z \neq s(sk_2)
                                                                                                                                       cnf 7
11.s(x_0 + x_1) = s(x_0) + x_1
                                                                                                                                       cnf 2
12.z + x_0 = x_0
                                                                                                                                       cnf 1
13.sk_1(x_0) + z = x_0 \lor z \neq z \lor sk_2 = sk_3 + z
                                                                                                                    superposition 9, 12
14.sk_1(x_0) + z = x_0 \lor sk_2 = sk_3 + z
                                                                                                        trivial inequality removal 13
15.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor s(x_3 + z) \neq s(sk_2)
                                                                                                                 supserposition 10, 11
16.sk_0 \neq x_0 \lor sk_2 = sk_3 + z
                                                                                                                    superposition 8, 14
17.sk_2 = sk_3 + z
                                                                                                                  equality resolution 16
18.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z \lor s(sk_2) \neq s(sk_2)
                                                                                                                   superposition 15, 17
19.sk_1(x_0) + z = x_0 \lor z \neq x_1 + z
                                                                                                        trivial inequality removal 18
20.sk_1(x_0) + z = x_0 \lor z \neq z
                                                                                                                   superposition 12, 19
21.sk_1(x_0) + z = x_0
                                                                                                        trivial inequality removal 20
22.sk_0 \neq x_0
                                                                                                                     superoisition 21, 8
23.\bot
                                                                                                                 equality resolution 22
```

For now we will consider two minimal examples. The first doesn't use induction. There are two constant symbols a, b and a binary proposition variable p(-,-). The only axiom is  $\forall x \ p(x,a) \lor p(x,b)$ . We seek to prove  $\forall x \exists y p(x,y)$ . It is immediately clear that unless the axiom has computational content, our extracted program cannot yield a value. On the other hand we will see that we can still create a well-typed program. And if the axiom contains computational content it will indeed yield a result.

#### Example 2.4.

$$p(x_1, a) \lor p(x_1, b) \qquad Axiom \qquad (1)$$

$$\llbracket \neg p(x_2, y_2) | x \simeq x_2 \rrbracket \qquad negated \ Hypothesis \qquad (2)$$

$$\llbracket p(x_1, b) | x \simeq x_1 \rrbracket \qquad res. \ (1)(2)[x_1/x_2, a/y_2] \qquad (3)$$

$$\llbracket \Box | x \simeq x_1 \rrbracket \qquad res. \ (3)(2)[x_1/x_2, b/y_2] \qquad (4)$$

$$\Box \qquad Instantiation \ (4) \qquad (5)$$

Now for our computational translation we introduce a special exception terms  $e_{\alpha}(t_1,\ldots,t_n)$  and  $\overline{e}_{\alpha}(t_1,\ldots,t_n)$  for every atom  $\alpha[t_1,\ldots,t_n]$  with n term-variables, which intuitively indicate that  $\alpha[t_1,\ldots,t_n]$  is not false. They are usually introduced in the manner of the last derivation of definition 3.4 in [1]. If one were to assign types we would have  $e_{\alpha}(t_1,\ldots,t_n):(\alpha(t_1,\ldots,t_n)\to \bot)\to \bot$  and  $\overline{e}_{\alpha}(t_1,\ldots,t_n):((\alpha(t_1,\ldots,t_n)\to \bot)\to \bot)\to \bot$ . However as we have the (classical) identification, we can simplify  $\overline{e}_{\alpha}(t_1,\ldots,t_n):\alpha(t_1,\ldots,t_n)\to \bot$ . Having this in mind both  $e_{\alpha}(t_1,\ldots,t_n)=(t_1,\ldots,t_n)$  and  $\overline{e}_{\alpha}(t_1,\ldots,t_n)=(t_1,\ldots,t_n)$ .  $t_1$  make "type-sense" in the exception sense of [1]. We will use this in our program. Let us now translate the above proof. Note that the terms for axioms are given. For clarity we will use braces for function application. For convenience we will write  $\alpha_i$  for the i-the generated term.

#### Example 2.5.

$$\begin{array}{lll} \alpha_1 & (p(x_1,a) \rightarrow \bot) \rightarrow (p(x_1,b) \rightarrow \bot) \rightarrow \bot & (1) \\ \overline{e}_p(x_2,y_2) & p(x_2,y_2) \rightarrow \bot & (2) \\ \lambda t.\mathbf{let}\ v: \neg p(x_1,a) & (p(x_1,b) \rightarrow \bot) \rightarrow \bot & (3) \\ & \mathbf{in}\ \alpha_1(v,t) & \\ & \mathbf{handle}\ v(w) \Rightarrow \overline{e}_p(x_1,a)(w) & [=\alpha_2[x_1/x_2,a/y_2](w)] \\ & \mathbf{let}\ v: \neg p(x_1,b) & \bot & (4) \\ & \mathbf{in}\ \alpha_3(v) & \\ & \mathbf{handle}\ v(w) \Rightarrow \overline{e}_p(x_1,b)(w) & [=\alpha_2[x_1/x_2,b/y_2](w)] \\ \lambda x_1.\alpha_4 & (x:X) \rightarrow \bot & (5) \end{array}$$

Now to retrieve a value we just need to handle the final exception, i.e.

$$\lambda x_1.\alpha_4$$
 handle  $\overline{e}_p(x_2,y_2)(w) \Rightarrow y_2$ 

Of course this will not yield an actual value unless  $\alpha_1$  yields.

We will now consider a minimal inductive example.

Consider the inductive type nat with constructors z: nat and  $s: nat \to nat$  and some predicate  $p(x,y) \subseteq nat^2$ . We are going to prove  $\forall x \exists y \ p(x,y)$  from p(z,z) and  $\forall n,m: nat(p(n,m) \Rightarrow p(s(n),s(s(m))))$ .

# Example 2.6.

p(z,z)	Axiom	(6)
$\neg p(n_2, m_2) \lor p(s(n_2), s(s(m_2)))$	Axiom	(7)
$\llbracket \neg p(n_3, m_3)   x \simeq n_3 \rrbracket$	$negated\ Hypothesis$	(8)
$\llbracket\Box x\simeq z\rrbracket$	res. $(1)(3)[z/n_3, z/m_3]$	(9)
$\llbracket \neg p(n_2, m_2)   x \simeq s(n_2) \rrbracket$	$res.(2)(3)[s(n_2)/n_3], s(s(m_2))/m_3]$	(10)
$Activate \ \alpha := p(a,b)[a,b]$	Trigger (4)	(11)
$[n_7 \not\prec n_7' \lor p(n_7, m_7)   \mathbf{T}_{\alpha}^{\prec}(n_7', m_7)]$	$\Gamma_{\mathbf{T}_{m{lpha}}}$	(12)
$[n_7 \not\prec n_7'   x \simeq s(n_7) \wedge \mathbf{T}_{\alpha}^{\prec}(n_7', m_7)]$	$res.(5)(7)[n_7/n_2, m_7/m_2]$	(13)
$[n_7 \prec s(n_7) \top]$	$\Gamma^s_{\prec}$	(14)
$\llbracket \Box   x \simeq s(n_7) \wedge \mathbf{T}_{\alpha}^{\prec}(s(n_7), m_7) \rrbracket$	$res.(8)(9)[s(n_7)/n_7']$	(15)
$\llbracket\Box x\simeq n_{10}\rrbracket$	Induction $(4)(9)$	(16)
	Instantiation (10)	(17)

## 2.4 Intuitionistic Example

For simplicity we will use a single sort (0-level type).

Function symbols: f(-)

Predicate symbols: P(-), Q(-, -)

Axioms:  $\forall x P(x), \ \forall x \exists y x = f(y), \ \forall x (P(x) \supset Q(f(x), x))$ 

To Show:  $\forall x \exists y Q(x, y)$ 

Below is a ND-style proof and translation based on the calculus and correspondence presented in [8].

We can reduce the last term to  $\lambda x. \langle \alpha(x)_1, \operatorname{subst}(f(\alpha(x)_1), x, \Pi_t Q(t, \alpha(x)_1), \alpha(x)_2, \beta(\alpha(x)_1)(\gamma(\alpha(x)_1)) \rangle$ . Finally replacing the proof terms by units we obtain  $\lambda x. \langle \alpha(x)_1, () \rangle : \Pi_x \Sigma_y \top$  from which we may further extract  $\lambda x. \alpha(x)_1$ .

# References

- [1] DE GROOTE, P. A simple calculus of exception handling. In *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 1995, pp. 201–215.
- [2] ECHENIM, M., AND PELTIER, N. Combining induction and saturation-based theorem proving. *Journal of Automated Reasoning 64*, 2 (mar 2019), 253–294.
- [3] Griffin, T. G. A formulae-as-type notion of control. In *Proceedings of the 17th ACM SIGPLAN-SIGACT symposium on Principles of programming languages POPL '90* (1990), ACM Press.
- [4] Hajdu, M. Automating inductive reasoning with recursive functions, 2021.
- [5] HAJDÚ, M., HOZZOVÁ, P., KOVÁCS, L., SCHOISSWOHL, J., AND VORONKOV, A. Induction with generalization in superposition reasoning. In *Intelligent Computer Mathematics* (Cham, 2020), C. Benzmüller and B. Miller, Eds., Springer International Publishing, pp. 123–137.
- [6] KOVÁCS, L., AND VORONKOV, A. First-order theorem proving and vampire. In Computer Aided Verification. Springer Berlin Heidelberg, 2013, pp. 1–35.
- [7] NORELL, U. Towards a practical programming language based on dependent type theory, vol. 32. Citeseer, 2007.
- [8] Peenning, F. Constructive logic, 2000.