M2 LMFI – Quantification du second-ordre et points fixes en logique

Introduction to system F

Introduction to system F

1 Definition of System F

$\forall X.F$

$X(t_1 \ldots t_n)$.

$n = 0$.

Définition 1.1 (*Types of system* **F**)

Let us consider an infinte, countable set of type variables (or second-order type variables), $\mathcal{V}_{\mathsf{F}}$. System F *types are given by the grammar:*

$$T, U, V ::= X \mid U \to V \mid \forall X.T$$

$\forall X.T$ binds X in T, so that types are considered with the expected notions of free and bound (type) variables, capture-free substitition of an F-type for a variable (written $T\{U/X\}$).

$A\{B/X\}$: $X\{B/X\} = B$

$\forall X.T =_\alpha \forall Y.T\{Y/X\}$

Définition 1.2 (*Church-style System* **F**)

One considers, for each type T of F*, an infinite countable set of variables for this type, $\mathcal{V}^T$.*

Church-style System F *terms are the least set such that:*

- *For any variable x in $\mathcal{V}^T$, x^T is a term of type T (with free variables $\{x\}$);*
- *For any term v of type V and any variable x in $\mathcal{V}^{\mathcal{U}}$, $\lambda x^U.v$ is a term of type $U \to V$ (with free variables $fv(t) \setminus \{x\}$);*
- *For any terms t and u of respective types $U \to T$ and U, $(t)u$ is a term of type T (with free variables $fv(t) \cup fv(u)$);*
- *For any type variable X and any term t of type T, $\Lambda X.t$ is a term of type $\forall X.T$* **under the condition that, for any free variable libre x of t, X does not occur free in the type of** *x (with free variables $fv(t)$);*
- *For any term t of type $\forall X.T$ and any type U of* F*, $(t)U$ is a term of type $T\{U/X\}$ (with free variables $fv(t)$).*

Curry-style System F

As for the simply-typed case, Curry-style system F consists in the pure λ-calculus together with a typing relation given by a type system, inspired by second order natural deduction. That is simply the type system for simply-typed λ-calculus extended with the two following typing rules:

$$\frac{\Gamma \vdash t : T}{\Gamma \vdash t : \forall X.T}\ \forall i\ (*)$$

$$\frac{\Gamma \vdash t : \forall X.T}{\Gamma \vdash t : T\{U/X\}}\ \forall e(U)$$

$(*)$ The inference rule $\forall i$ can only be applied if X does not occur free in the type of (the variables in) Γ.

The judgement that a λ-term t appear in conclusion of a typing derivation with type T under context Γ will be written $\Gamma \vdash_{\mathsf{F}} t : T$.

Définition 1.2 (*Church-style System* F)

One considers, for each type T of F*, an infinite countable set of variables for this type, $\mathcal{V}^T$.*

Church-style System F *terms are the least set such that:*

- *For any variable x in $\mathcal{V}^T$, x^T is a term of type T (with free variables $\{x\}$);*
- *For any term v of type V and any variable x in $\mathcal{V}^U$, $\lambda x^U.v$ is a term of type $U \to V$ (with free variables $fv(t) \setminus \{x\}$);*
- *For any terms t and u of respective types $U \to T$ and U, $(t)u$ is a term of type T (with free variables $fv(t) \cup fv(u)$);*
- *For any type variable X and any term t of type T, $\Lambda X.t$ is a term of type $\forall X.T$* **under the condition that, for any free variable libre x of t, X does not occur free in the type of x** *(with free variables $fv(t)$);*
- *For any term t of type $\forall X.T$ and any type U of* F*, $(t)U$ is a term of type $T\{U/X\}$ (with free variables $fv(t)$).*

If $X \neq Y$
Y not free in U.

$(\Lambda X.\Lambda Y.t)UV$

$\rightarrow_{\forall} (\Lambda Y.t\{U/X\})V$

$(\lambda x.t)u : T$

$\rightarrow_{\beta} t\{u/x\} : T$

Proposition 1.3

If t is an F*-term of type T, x is a variable of type U and u is an* F*-term of type U, then $t\{u/x\}$ is an* F*-term of type T, having the free variables as $(fv(t) \setminus \{x\}) \cup fv(u)$.*

Proposition 1.4

If t is an F *term of type T and if X is a type variable that does not occur free in the free variables of t, then for any* F*-type U, $t\{U/X\}$ is an* F*-term of type $T\{U/X\}$, having the same free variables as t.*

The previous propositions ensure that the following definition is meaningful:

$(\Lambda X.t)U : T\{U/X\}$

$\rightarrow_{\forall} t\{U/X\} : T\{U/X\}$

Définition 1.5 (*Dynamics of* F)

Church-style F*-terms are equipped with two reduction rules:*

$$\begin{array}{rcl}(\lambda x.t)u & \longrightarrow_\beta & t\{u/x\} \\ (\Lambda X.t)U & \longrightarrow_\forall & t\{U/X\}\end{array}$$

The first rule is called β-reduction, as usual, while the second is called **universal reduction**.

$s(t)$:

$s(x) = 1$

$s(\lambda x . t) = s(t) + 1$

$s((t)u) = s(t) + s(u)$.

$s(\Lambda X . t) = s(t) + 1$

$s((t)U) = s(t) + 1$.

$s((\Lambda X . t)U) = s(t\{U/X\}) + 2$.

$\lambda x . (x)x$

$((\Lambda X . t)U)u$

$\lambda x . t \quad \forall$

$(\lambda x . t\{U/X\})u$.

Définition 1.2 (*Church-style System* F)

One considers, for each type T of F*, an infinite countable set of variables for this type, $\mathcal{V}^T$.*

Church-style System F *terms are the least set such that:*

- *For any variable x in $\mathcal{V}^T$, x^T is a term of type T (with free variables $\{x\}$);*
- *For any term v of type V and any variable x in $\mathcal{V}^U$, $\lambda x^U.v$ is a term of type $U \to V$ (with free variables $fv(t) \setminus \{x\}$);*
- *For any terms t and u of respective types $U \to T$ and U, $(t)u$ is a term of type T (with free variables $fv(t) \cup fv(u)$);*
- *For any type variable X and any term t of type T, $\Lambda X.t$ is a term of type $\forall X.T$* **under the condition that, for any free variable libre x of t, X does not occur free in the type of x** *(with free variables $fv(t)$);*
- *For any term t of type $\forall X.T$ and any type U of* F*, $(t)U$ is a term of type $T\{U/X\}$ (with free variables $fv(t)$).*

$\vdash$ a Typing context, Γ
F a ~~term~~ term t
a type T

$\boxed{\Gamma \vdash_F t : T}$

Proposition 1.3

If t is an F*-term of type T, x is a variable of type U and u is an* F*-term of type U, then $t\{u/x\}$ is an* F*-term of type T, having the free variables as $(fv(t) \setminus \{x\}) \cup fv(u)$.*

Proposition 1.4

If t is an F *term of type T and if X is a type variable that does not occur free in the free variables of t, then for any* F*-type U, $t\{U/X\}$ is an* F*-term of type $T\{U/X\}$, having the same free variables as t.*

The previous propositions ensure that the following definition is meaningful:

Définition 1.5 (*Dynamics of* F)

Church-style F*-terms are equipped with two reduction rules:*

$$\begin{array}{rcl}(\lambda x.t)u & \longrightarrow_\beta & t\{u/x\} \\ (\Lambda X.t)U & \longrightarrow_\forall & t\{U/X\}\end{array}$$

The first rule is called β-reduction, as usual, while the second is called **universal reduction**.

Curry-style System F

$$\boxed{\Lambda X.t \qquad (t)U}$$

As for the simply-typed case, Curry-style system F consists in the pure λ-calculus together with a typing relation given by a type system, inspired by second-order natural deduction. That is simply the type system for simply-typed λ-calculus exgtended with the two following typing rules:

$$\frac{\delta : \vdash \lambda x.x : \forall X.X \to X}{\vdash \lambda x.x : U \to U}\ \forall e(U)$$

$$\frac{\Gamma \vdash t : T}{\Gamma \vdash t : \forall X.T}\ \forall i \quad (\star)$$

$$\cancel{\frac{\dfrac{}{x : X \vdash x : X}\ \mathrm{Var}}{x : X \vdash x : \forall X.X}\ \forall i}$$

$$\frac{\Gamma \vdash t : \forall X.T}{\Gamma \vdash t : T\{U/X\}}\ \forall e(U)$$

($\star$) The inference rule $\forall i$ can only be applied if X does not occur free in the type of (the variables in) Γ.

The judgement that a λ-term t appear in conclusion of a typing derivation with type T under context Γ will be written $\Gamma \vdash_{\mathsf{F}} t : T$.

$$\boxed{\Gamma \vdash_F t : T}$$

$$\frac{(x:T) \in \Gamma}{\Gamma \vdash_F x : T}\ \mathrm{Var} \qquad \frac{\Gamma, x : U \vdash_F t : V}{\Gamma \vdash_F \lambda x.t : U \to V}\ \mathrm{Abs.}$$

$$\delta = \frac{\dfrac{\dfrac{}{x : X \vdash_F x : X}\ \mathrm{Var}}{\vdash \lambda x.x : X \to X}\ \mathrm{Abs.}}{\vdash \lambda x.x : \forall X.X \to X}\ \forall i$$

$$\frac{\Gamma \vdash_F t : U \to V \qquad \Gamma \vdash_F u : U}{\Gamma \vdash (t)u : V}\ \mathrm{App.}$$

Prop: Γ context, $t \in \Lambda$, T a type if $\Gamma \vdash t : T$ and if $t \to_\beta u \in \Lambda$ then $\Gamma \vdash_F u : T$

Relations between the two presentations

Définition 1.6 (*forgetful map*)

One defines inductively a type-forgetting map from Church-style F*-terms to pure* λ*-terms:*

- $[x^T]^- = x$;
- $[\lambda x^T.\, t]^- = \lambda x [t]^-$.
- $[(t)\, u]^- = ([t]^-)\, [u]^-$;
- $[\Lambda X.\, t]^- = [t]^-$;
- $[(t)\, T]^- = [t]^-$.

Proposition 1.7

Let $t : T$ be a Church-style F*-term with free variables libres among $(x_i^{T_i})_{1 \le i \le n}$. Then $x_1 : T_1, \ldots x_n : T_n \vdash_{\mathsf{F}} [t]^- : T$ is derivable in (Curry-style) F.*

Proposition 1.8

A type derivation δ for a judgment $\Gamma \vdash_{\mathsf{F}} t : T$ is isomorphic to a Church-style term $u : T$ the free variables of which are among the variables of Γ (and typed according to Γ. In addition, $[u]^- = t$.

Lemme 1.9

Universal reduction is strongly normalizing in (Church-style) system F.

Reductions in Church-style and Curry-style system F can also be compared:

Proposition 1.10

1. *The type-erasure of a (Church-style) normal form is a normal form.*
2. *If t reduces to u with a universal-step, then $[t]^- = [u]^-$.*
3. *If t reduces to u with a β-step, then $[t]^-$ reduces to $[u]^-$ with a β-step.*
4. *The previous statements show that if $[t]^-$ is normal, reductions from t contain only universal steps.*
5. *If $[t]^-$ reduces in one step to u, then t reduces in at least one step to some v such that $[v]^- = u$.*

Finally, one can state the result we expected:

Théorème 1.11

Weak (resp. strong) normalization of Church-style system F *is equivalent to the weak (resp. strong) normalization of Curry-style system* F.

Finally, one can state the result we expected:

Théorème 1.11

Weak (resp. strong) normalization of Church-style system F *is equivalent to the weak (resp. strong) normalization of Curry-style system* F.

Démonstration : For weak normalization, the reasoning is direct and sumple: Assume F is weakly normalizing and let t be a λ-term which is typable in Curry-style F. We know that there exists u such that $[u]^- = t$ and that by hypothesis, u has a normal form v. The previous propositions ensure that $t \longrightarrow^\star [v]^-$ which is normal. In the other direction, if every λ-term which is typable in Curry-style F normalizes, let us consider some term t of Curry-style F. We know that $[t]^-$ normalizes to u and that there exists v in F such that $t \longrightarrow^\star v$ and $[v]^- = u$. Let us remark that v is not necessarily normal but that, by the previous propositions, all its reductions are universal, which we know to be strongly normalizing.

For strong normalization, the reasoning is slightly less immediate. Assume that there exists a term t typable in Curry-style system F from which an infinite reduction sequence can be drawn. We know that there exists some (Church-style) F-term u such that $[u]^- = t$ and the previous propositions ensure that from this term, an infinite reduction sequence can also be build. By contraposition, strong normalization of F ensures that every pure terme that is typable in (Curry-style) system F strongly normalizes. Assume now that there exists some F-term t having an infinite reduction ρ. By strong normalization of universal reduction, we know that ρ contains infinitely many β-reductions from which one conclude that $[t]^-$ also has an infinite reduction.

□

Définition 5.1 (*Some data types in System* F)

Let us consider:

- $\bot = \forall X.X$;
- $1 = \mathsf{ID} = \forall X.(X \to X)$;
- $\mathsf{Bool} = \forall X.(X \to (X \to X))$;
- $\mathsf{Nat} = \forall X.(X \to (X \to X) \to X)$;
- $T \times U = \forall X.(U \to V \to X) \to X)$;
- $T + U = \forall X.(T \to X) \to (U \to X) \to X)$;
- $\mathsf{DNE} = \forall X.((X \to \bot) \to \bot) \to X$;
- $\mathsf{List}(T) = \forall X.X \to (T \to (X \to X)) \to X$;
- $\mathsf{List} = \forall Y.\forall X.X \to (Y \to (X \to X)) \to X$;
- $\mathsf{Tree}(T) = \forall X.X \to ((T \to X) \to X) \to X$;
- $\mathsf{Tree} = \forall Y.\forall X.X \to ((Y \to X) \to X) \to X$.

Curry:
$$\frac{\vdash_F t : \bot}{\vdash t : X\{U/X\} = U}\ \forall_e(U)$$

Church:
$$\frac{t : \bot}{(t)U : U}$$

$\forall X.(X \to X) \to (X \to X)$.

$\vdash_F \lambda x.x : \mathsf{Nat}$

$\cong \lambda x.\lambda y.(x)y$

$\|$

$[1]$

Proposition 5.2

The is no closed term t such that $\vdash_{\mathsf{F}} t : \bot$.

Proposition 5.3

If $\vdash_{\mathsf{F}} t : \mathsf{ID}$, then $t \longrightarrow^{\star}_{\beta} \lambda x.\,x$.

Proposition 5.4

If $\vdash_{\mathsf{F}} t : \mathsf{Bool}$, then $t \longrightarrow^{\star}_{\beta} \lambda x.\,\lambda y.\,x$ or $t \longrightarrow^{\star}_{\beta} \lambda x.\,\lambda y.\,y$.

Proposition 5.5

If $\vdash_{\mathsf{F}} t : \mathsf{Nat}$, then there exists an natural n such that $t \longrightarrow^{\star}_{\beta} \lambda z.\,\lambda s.\,(s)^n z$.

Proposition 5.10

Let T be a type, let $\vdash_{\mathsf{F}} t : \mathsf{List}(T)$. Assuming that $v_{::}$ and $v_{[]}$ are two variables of system F*, there exists $n \geq 0$ and closed terms $a_1, \ldots, a_n$ such that $\vdash_{\mathsf{F}} a_i : T$ for $1 \leq i \leq n$ such that $t \longrightarrow^{\star}_{\beta} \lambda v_{[]}.\,\lambda v_{::}.\,((v_{::})\,a_1((v_{::})\,a_2 \ldots ((v_{::})\,a_n v_{[]})))$.*

2 Weak normalization of system F

2.1 Introducing reducibility candidates

In order to extend the reducibility technique from simple types to System F, one faces a difficulty.

— $\mathsf{RED}^{\mathsf{WN}}(X) = \mathsf{Norm}(X)$;

— $\mathsf{RED}^{\mathsf{WN}}(U \to V) = \{t : U \to V ; \forall u \in \mathsf{RED}^{\mathsf{WN}}(U), (t)\, u \in \mathsf{RED}^{\mathsf{WN}}(V)\}$.

the natural extension of reducibility would be:

$$\mathsf{RED}^{\mathsf{WN}}(\forall X.U) = \{t : \forall X.U \mid \text{for all type } V, (t)\, V \in \mathsf{RED}^{\mathsf{WN}}(U\,\{V/X\})\}$$

But this would be an ill-formed definition

To solve this problem, the idea will consist in avoiding to define reducible terms for each type but, rather to define an abstract notion of set of reducible terms, from basic properties it should satisfy, and to axiomatise the notion of reducibility in some sense.

From there, one will define a notion of parametric reducibility from valuations which will associate *reducibility candidates* to each type variable.

To proceed, one shall therefore identify the characteristic of candidates which are given by the properties that we need for the theorem. There are actually two properties which are crucial to prove weak normalization in the simply typed case:

- the sets $\mathsf{RED}^{\mathsf{WN}}(T)$ are ***closed by β-expansion***;
- the sets $\mathsf{RED}^{\mathsf{WN}}(T)$ are ***adapted***.

Définition 2.1 (*Reducibility candidate*)

A **reducibility candidate** *of type T is a set R of λ-terms of type T which satisfies the following two conditions:*

$\forall$ if $t : T$ is such that $t \in R$ and if $u : T$ st. $u \to t$ then $u \in R$

- *(CR1) R is closed by β-expansion;*
- *(CR2)* $\mathsf{Neut}(T) \subseteq R \subseteq \mathsf{Norm}(T)$.

One shall denote by $\mathsf{CR}(T)$ the set of all reducibility candidates of type T. Note that for any T, $\mathsf{Norm}(T) \in \mathsf{CR}(T)$ so $\mathsf{CR}(T)$ is never empty.

2 Weak normalization of system F

Définition 2.1 (*Reducibility candidate*)

A **reducibility candidate** *of type T is a set R of λ-terms of type T which satisfies the following two conditions:*

- *(CR1) R is closed by β-expansion;*
- *(CR2) $\mathsf{Neut}(T) \subseteq R \subseteq \mathsf{Norm}(T)$.*

One shall denote by $\mathsf{CR}(T)$ the set of all reducibility candidates of type T. Note that for any T, $\mathsf{Norm}(T) \in \mathsf{CR}(T)$ so $\mathsf{CR}(T)$ is never empty.

$$\mathcal{C} = \bigcup_{T\ \text{type}} CR(T).$$

If R is a reducibility candidate, there is a unique type T such that $R \in \mathsf{CR}(T)$, this type will be noted $\mathsf{Type}(R)$.

A valuation ρ is a partial function from type variables to reducibility candidates the domain of which, noted $\mathsf{dom}(\rho)$, is finite. One shall write $\rho[X := R]$ for the valuation ρ' of domain $\mathsf{dom}(\rho) \cup \{X\}$ such that:

$$\rho'(X) = R \qquad \rho'(Y) = \rho(Y) \text{ si } Y \neq X.$$

One shall say that a valuation ρ covers a type T (resp. a term $t : T$) if its domain contains all the free type variables of T (resp. of T and of the types of the free variables of t).

Définition 2.3 (***Initial valuation, ρ_Γ***)

Let $\Gamma = X_1, \ldots, X_n$, one defines the **initial valuation** *(or default valuation) on Γ, ρ_Γ as the valuation dedfined on Γ such that $\rho_\Gamma(X_i) = \mathsf{Norm}(X_i)$.*

A valuation induces a type substitution:

Définition 2.4 (T^ρ, t^ρ)

Let ρ be a valuation, one defines the type substitution:

- $X_i^\rho = X_i$ *if* $X \notin dom\rho$; U *if* $X \in dom\rho$ *and* $\rho(X) \in CR(U)$
- $T^\rho = T\{X_i^\rho / X_i, 1 \leq i \leq n\}$ *if the free variables of* T *are the* $X_1, \ldots, X_n$.

One also defines a substitution on terms:

$$t^\rho = t\{X_i^\rho / X_i, 1 \leq i \leq n\}.$$

Remarque 2.5

In particular, if the free type variables of T and t are all in Γ, one has $T^{\rho_\Gamma} = T$ and $t^{\rho_\Gamma} = t$.

Définition 2.6 ($\mathsf{RED^{WN}}_\rho(T)$)

$\mathsf{RED^{WN}}_\rho(T)$ *is defined by induction on T (if ρ covers T):*

- $\mathsf{RED^{WN}}_\rho(X) = \rho(X)$;
- $\mathsf{RED^{WN}}_\rho(U \to V) = \{t : (U \to V)^\rho / \forall u \in \mathsf{RED^{WN}}_\rho(U), (t)u \in \mathsf{RED^{WN}}_\rho(V)\}$;
- $\mathsf{RED^{WN}}_\rho(\forall X.U) = \{t : (\forall X.U)^\rho / \forall V \in Type, \forall R \in CR(V), (t)V \in \mathsf{RED^{WN}}_{\rho[X:=R]}(U)\}$

2.3 Proof of the normalization theorem

The normalization theorem relies on the following two lemmas:

Lemme 2.7

For any type T and any valuation ρ covering T, $\mathsf{RED^{WN}}_\rho(T) \in \mathsf{CR}(T)$.

Lemme 2.8 (*Adequation lemma*)

Let $t : T$ be a Church-style F-term of free variables $(x_i^{U_i})_{1\leq i\leq n}$, then for any valuation ρ covering t and for any $u_i \in \mathsf{RED^{WN}}_\rho(U_i)$, $1 \leq i \leq n$, one has: $t^\rho \{u_i/x_i, 1 \leq i \leq n\} \in \mathsf{RED^{WN}}_\rho(T)$.

Théorème 2.9

Every term of (Church-style) System F (weakly) normalises.

Démonstration of the normalization theorem : The proof is identical to the simply-typed case.

Let $t : T$ be of free variables $(x_i^{T_i})_{1\leq i\leq n}$ and let Γ be such that is contains all the free variables occurring in T and in the T_i. One has $T^{\rho_\Gamma} = T$ and $T_i^{\rho_\Gamma} = T_i$. Since the $\mathsf{RED^{WN}}_\rho(U)$ are candidates, one has by (CR2) that $1 \leq i \leq n$, $x_i^{T_i} \in \mathsf{RED^{WN}}_\rho(T_i)$ since (term) variables are neutral.

Adequation lemma ensures that $t\left\{x_i^{T_i}/x_i, 1 \leq i \leq n\right\} = t \in \mathsf{RED^{WN}}_\rho(T) \subseteq \mathsf{Norm}(T)$.

$\square$

Lemme 2.10

If $t : \forall X.T$, and U is a type, then $(t)U$ normalizable implies that t is itself normalizable.

Lemme 2.11 (*substitution*)

Let V, W be types and ρ be a valuation covering V, W. One has:

$$\mathsf{RED}^{\mathsf{WN}}{}_{\rho}(V\{W/Y\}) = \mathsf{RED}^{\mathsf{WN}}{}_{\rho[Y:=\mathsf{RED}^{\mathsf{WN}}{}_{\rho}(W)]}(V).$$

Démonstration of adequation lemma : The lemma is proved by induction on the structure of term, as usual. The case for variable, lambda-abstraction and application are treated as in the simply typed case. One only details the constructions which are specific to F:

- If $t = \Lambda X.u$, $U = \forall X.U'$ and $u : U'$. One can assume that X is not in the domain of the valuation, renaming the bound variable if needed. Let V be a type and $R \in CR(V)$ and let $u' = u^{\rho[X:=R]}\{u_i/x_i\}$. By induction hypothesis, one has $u' \in \mathsf{RED}^{\mathsf{WN}}{}_{\rho[X:=R]}(T)$ for any $R \in CR(V)$ and because $(t^{\rho}\{u_i/x_i\})V \to u^{\rho}\{u_i/x_i\}\{V/X\} = u^{\rho[X:=R]}\{u_i/x_i\} = u' \in \mathsf{RED}^{\mathsf{WN}}{}_{\rho[X:=R]}(T)$. Finally, closure of candidates by anti-reduction ensures that $(t^{\rho}\{u_i/x_i\})V \in \mathsf{RED}^{\mathsf{WN}}{}_{\rho[X:=R]}(T)$.

- If $t = (u)V$, with $u : \forall X.U$. Let us set $U' = U\{V/X\}$ and $u' = u^{\rho}\{u_i/x_i\}$. We have, by induction hypothesis that $u' \in \mathsf{RED}^{\mathsf{WN}}{}_{\rho}(\forall X.U)$ so that $(u')V^{\rho} \in \mathsf{RED}^{\mathsf{WN}}{}_{\rho[X:=R]}(U)$. But we also have $(u')V^{\rho} = t^{\rho}\{u_i/x_i\}$ and one can conclude, by using the lemma on substitutions, that $t^{\rho}\{u_i/x_i\} \in \mathsf{RED}^{\mathsf{WN}}{}_{\rho}(U')$.

$\square$