Un modèle ensembliste simple de la théorie de Martin-Löf

Benjamin Werner

novembre 2000

1 La théorie des types de Martin-Löf

Comme dans le reste du cours, on considère le “noyau dur” commun aux divers formalismes proposés par er Martin-Löf. C’est-à-dire un système de types dépendants, prédicatif, muni en particulier des entiers naturels, de la récurrence structurelle, de la somme disjointe et de Σ-types. On commence par rappeller la définition précise.

1.1 Les termes

Comme c’est courant pour les systèmes à types dépendants, on utilise une syntaxe unique pour les types/propositions et les termes/preuves. Elle est donnée par la grammaire suivante :

$$\begin{array}{lll} t & ::= & x \mid \lambda x : t.t \mid (t\ t) \mid \Pi x : t.t \mid \Sigma x : t.t \mid (t,t)^{\Sigma x:t.t} \mid \pi_1(t) \mid \pi_2(t) \mid \\ & & t+t \mid i(t)^{t+t} \mid j(t)^{t+t} \mid \delta^{t,t,t}(t,t,t) \mid 0 \mid S(t) \mid \mathrm{Rec}^t \mid \ =_t \ \mid \mathrm{refl}(t) \mid L(t,t,t,t,t)(\mid \\ & & \bot \mid \mathrm{efq}(t,t) \mid \mathrm{Type} \mid \mathrm{Kind}. \end{array}$$

Comme d’habitude, on utilisera la notation infoxe pour l’application de l’égalité $=_T$.

1.2 Réduction

La réduction à la racine d’un terme, notée $\rightarrow$ est la plus petite relation définie par les clauses suivantes :

$$\begin{array}{rcl} (\lambda x : A.t\ u) & \rightarrow & t[x \setminus u] \\ \pi_1(a,b)^{\Sigma x:A.B} & \rightarrow & a \\ \pi_2(a,b)^{\Sigma x:A.B} & \rightarrow & b \\ \delta^{A,B,C}(i(a)^{A'+B'}, f, g) & \rightarrow & (f\ a) \\ \delta^{A,B,C}(j(b)^{A'+B'}, f, g) & \rightarrow & (g\ b) \\ (\mathrm{Rec}^P\ 0\ p_0\ p_S) & \rightarrow & p_0 \end{array}$$

$$\begin{array}{rcl}(\mathrm{Rec}^P\ S(n)\ p_0\ p_S) & \to & (p_S\ n\ (\mathrm{Rec}^P\ n\ p_0\ p_S))\\ L(P,a,b,p_a,\mathrm{refl}(a')) & \to & p_a\end{array}$$

On note $\triangleright$ la congruence engendrée par cette réduction (c'est-à-dire en autorisant les réductions sur des sous-termes), $\triangleright^*$ la clôture transitive-reflexive de $\triangleright$ et $=_\beta$ la clôture transitive-réflexive-symétrique.

Théorème 1.1 *La relation $\triangleright^*$ vérifie la propriété de Church-Rosser.*

La preuve suit la technique habituelle de Tait et Martin-Löf. Le détail est juste un petit peu plus long à cause des diverses réductions.

1.3 Les règles

La relation de typage $\Gamma \vdash t : T$ et la proprieté de bonne formation des contexte Γ wf sont définies par les règles d'inférence de la figure 1.3.

2 Un modèle ensembliste

Le principe de la construction de ce modèle est extrêmement simple :
- Un type A est interprété par un ensemble $|A|$.
- Un terme preuve de type A est interprété par un élément de l'interprétation $|A|$ de son type (donc la relation "être de type de..." est traduite par l'appartenance dans le modèle).
- Un type fonction $A \to B$ est interprété par l'ensemble de toutes les fonctions totales de $|A|$ vers $|B|$.

Il s'agit donc quasiment d'une traduction de la théorie des types vers la théorie des ensembles.

Définition 2.1 *Soit un famille d'ensembles $(B_i)_{i\in A}$. On définit les ensembles suivant de la manière usuelle :*

$$\begin{array}{rcl}\Sigma_{i\in A}B_i & \equiv & \{(a,b), a\in A \wedge b \in B_i\}\\ \Pi_{i\in A}B_i & \equiv & \{f \in A \to (\bigcup_{i\in A} B_i), \forall a \in A.f(a)\in B_a\}\end{array}$$

où $X \to Y$ désigne l'espace des applications de X vers Y.

Définition 2.2 *Soient deux ensembles A et B. On définit*

$$A + B \equiv \{(0,a), a \in A\} \cup \{(1,b), b\in B\}.$$

Un univers ensembliste

Pour simplifier la construction du modèle, on suppose l'existence d'un "gros" ensemble $\mathcal{U}$ qui vérifie les conditions de clôture suivantes :
- $\forall A \in \mathcal{U}.(\forall \alpha \in A.B_\alpha \in \mathcal{U}) \Rightarrow \Pi_{\alpha\in A}B_\alpha \in \mathcal{U}$.
- $\forall A \in \mathcal{U}.(\forall \alpha \in A.B_\alpha \in \mathcal{U}) \Rightarrow \Sigma_{\alpha\in A}B_\alpha \in \mathcal{U}$.

$$(\textsc{empty})\ [\,]\ \mathrm{wf} \qquad (\textsc{Type})\ \frac{\Gamma\ \mathrm{wf}}{\Gamma \vdash \mathrm{Type} : \mathrm{Kind}}$$

$$(\textsc{var})\ \frac{\Gamma\ \mathrm{wf}}{\Gamma \vdash x : T}\ (\mathrm{si}\ (x : T) \in \Gamma) \qquad (\textsc{weak})\ \frac{\Gamma \vdash T : s}{\Gamma(x : T)\ \mathrm{wf}}\ (s \in \{\textsc{Type}, \textsc{Kind}\})$$

$$(\textsc{Pi})\ \frac{\Gamma(x : A) \vdash B : s' \qquad \Gamma \vdash A : s}{\Gamma \vdash \Pi x : A.B : s'}\ (\textsc{si}\ (s, s') \in \{(\textsc{Type}, \textsc{Type}), (\textsc{Type}, \textsc{Kind})\})$$

$$(\textsc{app})\ \frac{\Gamma \vdash t : \Pi x : A.B \qquad \Gamma \vdash u : A}{\Gamma \vdash (t\ u) : B[x \setminus u]} (\textsc{lam})\ \frac{\Gamma(x : A) \vdash t : B \qquad \Gamma \vdash \Pi x : A.b : s}{\Gamma \vdash \lambda x : A.t : \Pi x : A.B}$$

$$(\textsc{sigma})\ \frac{\Gamma \vdash A : \mathrm{Type} \qquad \Gamma(x : A) \vdash B : \mathrm{Type}}{\Gamma \vdash \Sigma x : A.B : \mathrm{Type}}$$

$$(\textsc{paire})\ \frac{\Gamma \vdash a : A \qquad \Gamma \vdash b : B[x \setminus a] \qquad \Gamma \vdash \Sigma x : A.B : \mathrm{Type}}{\Gamma \vdash (a, b)^{\Sigma x:A.B} : \Sigma x : A.B}$$

$$(\textsc{disj})\ \frac{\Gamma \vdash A : \mathrm{Type} \qquad \Gamma \vdash B : \mathrm{Type}}{\Gamma \vdash A + B : \mathrm{Type}}$$

$$(\textsc{di-g})\ \frac{\Gamma \vdash t : A \qquad \Gamma \vdash A + B : \mathrm{Type}}{\Gamma \vdash i(t)^{A+B} : A + B} \qquad (\textsc{di-d})\ \frac{\Gamma \vdash t : d \qquad \Gamma \vdash A + B : \mathrm{Type}}{\Gamma \vdash j(t)^{A+B} : A + B}$$

$$(\textsc{di-el})\ \frac{\Gamma \vdash t : A + B \qquad \Gamma \vdash C : \mathrm{Type} \qquad \Gamma \vdash f : A \to C \qquad \Gamma \vdash g : B \to C}{\Gamma \vdash \delta^{A,B,C}(t, f, g) : C}$$

$$(\textsc{nat})\ \frac{\Gamma\ \mathrm{wf}}{\Gamma \vdash \mathrm{nat} : \mathrm{Type}} \qquad (\textsc{zero})\ \frac{\Gamma\ \mathrm{wf}}{\Gamma \vdash 0 : \mathrm{nat}} \qquad (\textsc{succ})\ \frac{\Gamma \vdash t : \mathrm{nat}}{\Gamma \vdash S(t) : \mathrm{nat}}$$

$$(\textsc{rec})\ \frac{\Gamma \vdash P : \mathrm{nat} \to \mathrm{Type}}{\Gamma \vdash \mathrm{Rec}^P : (P\ 0) \to (\Pi n : \mathrm{nat}.(P\ n) \to (P\ S(n))) \to \Pi m : \mathrm{nat}.(P\ m)}$$

$$(\textsc{eq})\ \frac{\Gamma \vdash A : \mathrm{Type}}{\Gamma \vdash =_A : A \to A \to \mathrm{Type}} \qquad (\textsc{refl})\ \frac{\Gamma \vdash t : A \qquad \Gamma \vdash A : \mathrm{Type}}{\Gamma \vdash \mathrm{refl}(t) : t =_A t}$$

$$(\textsc{Leibniz})\ \frac{\Gamma \vdash A : \mathrm{Type} \quad \Gamma \vdash P : A \to \mathrm{Type} \quad \Gamma \vdash a : A \quad \Gamma \vdash b : A \quad \Gamma \vdash p_a : (P\ a) \quad \Gamma \vdash e : a =_A b}{\Gamma \vdash L(P, a, b, p_a, e) : (P\ b)}$$

$$(\textsc{bot})\ \frac{\Gamma \vdash t : \bot \qquad \Gamma \vdash A : \mathrm{Type}}{\Gamma \vdash \mathrm{efq}(t, A) : A}$$

– $\mathbb{N} \in \mathcal{U}$.
– $\forall(\alpha, \beta) \in \mathcal{U}^2.(\alpha, \beta) \in \mathcal{U}$.

Formellement, cette hypothèse est assez forte. Il est toutefois possible de s'en passer au prix d'une legère complication des définitions suivantes.

2.1 La construction du modèle

Pour interpréter un terme de la théorie des types, il faut se donner une interprétation de ses variables. On se donne donc une application $\mathcal{I}$ qui associe un élément $\mathcal{I}(x)$ de $\mathcal{U}$ à chaque variable x. Pour chaque terme t on définit alors $|t|_\mathcal{I}$, également dans $\mathcal{U}$ par les clauses suivantes. Remarquez qu'il s'agit, a priori, d'une définition partielle. On montrera ensuite qu'elle est en fait totale sur les termes bien typés.

$$\begin{array}{rcl}
|x|_\mathcal{I} & \equiv & \mathcal{I}(x) \\
|\lambda x : A.t|_\mathcal{I} & \equiv & \alpha \in |A|_I \mapsto |t|_{\mathcal{I};x\leftarrow\alpha} \\
|(t\ u)|_\mathcal{I} & \equiv & |t|_\mathcal{I}(|u|_\mathcal{I}) \\
|\Pi x : A.B|_\mathcal{I} & \equiv & \Pi_{\alpha\in|A|_\mathcal{I}}|B|_{\mathcal{I};x\leftarrow\alpha} \\
|\Sigma x : A.B|_\mathcal{I} & \equiv & \Sigma_{\alpha\in|A|_\mathcal{I}}|B|_{\mathcal{I};x\leftarrow\alpha} \\
|A + B|_\mathcal{I} & \equiv & |A|_\mathcal{I} + |B|_\mathcal{I} \\
|(a, b)^{\Sigma x:A.B}|_\mathcal{I} & \equiv & (|a|_\mathcal{I}, |b|_\mathcal{I}) \\
|i(a)^{A+B}|_\mathcal{I} & \equiv & (0, |a|_\mathcal{I}) \\
|j(b)^{A+B}|_\mathcal{I} & \equiv & (1, |b|_\mathcal{I}) \\
|0|_\mathcal{I} & \equiv & 0 \\
|S(t)|_\mathcal{I} & \equiv & |t|_\mathcal{I} + 1 \\
|\bot|_\mathcal{I} & \equiv & \emptyset \\
|=_A|_\mathcal{I} & \equiv & \alpha \in |A|_\mathcal{I}, \beta \in |A|_\mathcal{I} \mapsto \begin{array}{l} \{\emptyset\} \text{ si } \alpha = \beta \\ \emptyset \text{ sinon} \end{array} \\
|\mathrm{refl}(t)|_\mathcal{I} & \equiv & \emptyset
\end{array}$$

L'interprétation des opérateurs d'élimination n'est définie que dans certains cas précis :

$$\begin{array}{rcl}
|\pi_1(t)|_\mathcal{I} & \equiv & \alpha \text{ si } |t|_\mathcal{I} \text{ est de la forme } (\alpha, \beta) \\
|\pi_2(t)|_\mathcal{I} & \equiv & \beta \text{ si } |t|_\mathcal{I} \text{ est de la forme } (\alpha, \beta) \\
|\delta^{A,B,C}(t, f, g)|_\mathcal{I} & \equiv & \begin{array}{l} |f|_\mathcal{I}(\alpha) \text{ si } |t|_\mathcal{I} \text{ est de la forme } (0, \alpha) \\ |g|_\mathcal{I}(\alpha) \text{ si } |t|_\mathcal{I} \text{ est de la forme } (1, \alpha) \end{array} \\
|L(P, a, b, p_a, e)|_\mathcal{I} & \equiv & \begin{array}{l} |p_a|_\mathcal{I} \text{ si } |a|_\mathcal{I} = |b|_\mathcal{I} \\ \text{pas défini sinon} \end{array}
\end{array}$$

A noter qu'il n'est pas nécessaire de définir $|\mathrm{efq}(t, A)|_\mathcal{I}$.

3 Propriétés de l'interprétation

3.1 Correction vis-à-vis de la réduction

Théorème 3.1 (substitution) *Soient $\mathcal{I}$, une variable x et deux termes t et u. Si $|u|_\mathcal{I}$ et $|t|_{\mathcal{I};x\leftarrow|u|_\mathcal{I}}$ sont définis, alors $|t[x \setminus u]_\mathcal{I}$ est défini et*

$$|t|_{\mathcal{I};x\leftarrow|u|_\mathcal{I}} = |t[x \setminus u]_\mathcal{I}.$$

Théorème 3.2 (subject reduction) *Soit $\mathcal{I}$ et un term t. Si $|t|_\mathcal{I}$ est défini, alors, pour tout term t', si $t \triangleright_\beta t'$, alors $|t'|_\mathcal{I}$ est défini et $|t|_\mathcal{I} = |t'|_\mathcal{I}$.*

Preuve Les deux théoremes se prouvent par récurrence sur la structure de t. C'est un bon exercice que de vérifier quelques cas. ■

Le lemme précédent sert surtout à assurer la correction de l'interprétation vis-a-vis de la règle de conversion ; plus précisément :

Théorème 3.3 *Soient $\mathcal{I}$ et deux termes A et B. Si $A =_\beta B$ et $|A|_\mathcal{I}$ et $|B|_\mathcal{I}$ sont définis, alors $|A|_\mathcal{I} = |B|_\mathcal{I}$.*

Preuve Facile a partir du lemme précédent et de la proprieté de Church-Rosser. ■

3.2 Interprétations des contextes

Etant donné $\mathcal{I}$ et un contexte Γ, on définit la propriété "$\mathcal{I}$ convient à Γ", notée $\mathcal{I} \in |\Gamma|$ par les clauses suivantes :

- tout $\mathcal{I}$ convient à []
- $\mathcal{I} \in \Gamma(x : A)$ si et seulement si $\mathcal{I} \in |\Gamma|$, $|A|_\mathcal{I}$ est défini et $\mathcal{I}(x) \in |A|_\mathcal{I}$.

On dit qu'un contexte Γ est cohérent si $|\Gamma|$ est "non vide" (cad il existe $\mathcal{I} \in \||\Gamma|)$.

3.3 Correction de l'interprétation

On peut alors énoncer le résultat principal :

Théorème 3.4 *Si le jugement $\Gamma \vdash t : T$ est dérivable dans la théorie des types de Martin-Löf, alors pour tout $\mathcal{I}$ qui convient à Γ, $|t|_\mathcal{I}$ et $|T|_\mathcal{I}$ sont définis et $|t|_\mathcal{I} \in |T|_\mathcal{I}$.*

Preuve Par récurrence sur la dérivation de $\Gamma \vdash t : T$. Encore une fois, on peut traiter quelques cas à titre d'exercice. ■

4 Conséquence de la correction

Le résultat principal est que le formalisme est cohérent.

Théorème 4.1 *Si le contexte Γ est cohérent, il n'y a pas de jugement dérivable de la forme $\Gamma \vdash t : \bot$. En particulier, $[] \vdash t : \bot$ n'est pas dérivable.*

Preuve En effet, soit $\mathcal{I}$ qui convienne à Γ. On a $|\bot|_{\mathcal{I}}$ qui est défini et vide par définition. Il est donc impossible que $|t|_{\mathcal{I}}$ soit élément de $\bot$. ■

On peut également se servir de ce modèle pour vérifier que certains axiomes ne rendent pas le formalisme incohérent.

Théorème 4.2 *Soit* Γ *un contexte cohérent, et* A *tel que* $\Gamma \vdash A : Type$ *soit dérivable. Alors* $\Gamma(em : A + (A \rightarrow \bot))$ *est cohérent.*

Preuve Deux cas possibles :
- Si $|A|_{\mathcal{I}}$ est non vide, on choisit $\alpha \in |A|_{\mathcal{I}}$. Alors $\mathcal{I}; em \leftarrow (0, alpha)$ convient à $\Gamma(em : A + (A \rightarrow \bot))$.
- Si $|A|_{\mathcal{I}}$ est vide, appellons $f_{\emptyset}$ l'unique application de l'ensemble vide vers lui-même. Alors $|A \rightarrow \bot|_{\mathcal{I}}$ est égal au singleton $\{f_{\emptyset}\}$. Alors $\mathcal{I}; em \leftarrow (1, f_{\emptyset})$ convient à $\Gamma(em : A + (A \rightarrow \bot))$.

■