Système F_ω

Correction.

Documents autorisés (en particulier le poly).

Nous allons étudier le système F_ω, déjà vu rapidement en cours. Alors que le système F est une forme simplifiée de logique d'ordre 2, F_ω est essentiellement la même chose que la logique d'ordre supérieur (d'ordre ω).

On rappelle que F_ω est un système à trois niveaux:

1. les termes de preuve $s, t, u, v, \ldots$, qui ont pour types les formules $F, G, \ldots$;
2. les expressions logiques $S, T, \ldots$, ont pour types les sortes $K, L, \ldots$; parmi les expressions logiques, les formules sont celles qui sont de sorte $Prop$;
3. les sortes, parmi lesquelles la sorte $Prop$ des formules, et la sorte ι des individus (termes du premier ordre).

De façon formelle, les *sortes* sont définies par la grammaire:

$$K \quad ::= \quad Prop \mid \iota \mid K \to K$$

Les *pré-expressions logiques* sont définies par la grammaire:

$$S, T, \ldots \quad ::= \quad X_K \mid \lambda X_K \cdot T \mid S\,T \mid \mathrm{Imp} \mid \mathrm{All}_K$$

où la notation X_K dénote une variable de sorte K, et l'on suppose qu'il existe une infinité dénombrable de variables de sorte K, pour chaque sorte K. Il y a d'autre part une constante All_K pour chaque sorte K. On notera en abrégé:

$$\begin{aligned} F \Rightarrow G &= \mathrm{Imp}\ F\ G \\ \forall X : K \cdot F &= \mathrm{All}_K(\lambda X_K \cdot F) \end{aligned}$$

Une *expression logique* S de sorte K est une pré-expression logique telle que l'on puisse déduire $\triangleright S : K$ dans le système suivant:

$$\frac{}{\triangleright X_K : K}\,(KVar) \qquad \frac{}{\triangleright \mathrm{Imp} : Prop \to Prop \to Prop}\,(\mathrm{Imp}) \qquad \frac{}{\triangleright \mathrm{All}_K : (K \to Prop) \to Prop}\,(\mathrm{All}_K)$$

$$\frac{\triangleright S : K \to K' \quad \triangleright T : K}{\triangleright S\,T : K'}\,(KApp) \qquad \frac{\triangleright T : K'}{\triangleright \lambda X_K \cdot T : K \to K'}\,(KAbs)$$

Il s'agit donc d'un typage simple du λ-calcul formé des pré-expressions logiques, en style de Church, c'est-à-dire où toutes les variables sont annotées par le type (la sorte) qu'elle doivent avoir. On n'a donc pas besoin de contexte de typage pour définir le type des variables. Ce calcul, muni de la règle de réduction $(\lambda X_K \cdot S)T \to S[X_K := T]$, est confluent et fortement normalisant.

On observera, ce que l'on n'a pas besoin de démontrer, que pour toute pré-expression logique T, il existe au plus une sorte K telle que $\triangleright T : K$ soit dérivable. K sera *la sorte* de T, et l'on notera parfois $T : K$ pour le rappeler.

Les *termes de preuve* sont juste les λ-termes usuels:

$$s, t, u, v, \ldots \quad ::= \quad x \mid \lambda x \cdot s \mid st$$

Nous utiliserons les règles de réduction suivantes, collectivement nommées *β-réduction*:

$$\begin{array}{rcl}(\lambda x \cdot u)v & \to & u[x := v] \\ (\lambda X_K \cdot S)T & \to & S[X_K := T]\end{array}$$

Nous noterons encore $\to$ la relation de réécriture par β-réduction, $\to^*$ sa clôture réflexive-transitive, $=_\beta$ sa clôture réflexive-symétrique-transitive (β-conversion).

Les règles de typage de F_ω sont formées à l'aide de jugements $\Gamma \vdash u : F$, où Γ est un ensemble de couples $y : G$, avec y des variables distinctes deux à deux et G une formule (c'est-à-dire que l'on peut dériver $\triangleright G : Prop$). On notera que ce ne sont pas exactement les règles de typage du système F_ω du poly.

$$\frac{\triangleright F_1 : Prop \quad \ldots \triangleright F_n : Prop}{x_1 : F_1, \ldots, x_n : F_n \vdash x_i : F_i}\ (Ax) \qquad \frac{\Gamma \vdash u : F_1 \quad \triangleright F_2 : Prop \quad F_1 =_\beta^* F_2}{\Gamma \vdash u : F_2}\ (=_\beta)$$

$$\frac{\Gamma \vdash u : F_1 \Rightarrow F_2 \quad \Gamma \vdash v : F_1}{\Gamma \vdash uv : F_2}\ (\Rightarrow E) \qquad \frac{\Gamma, x : F_1 \vdash u : F_2}{\Gamma \vdash \lambda x \cdot u : F_1 \Rightarrow F_2}\ (\Rightarrow I)$$

$$\frac{\Gamma \vdash u : \forall X : K \cdot F \quad \triangleright S : K}{\Gamma \vdash u : F[X_K := S]}\ (\forall E) \qquad \frac{\Gamma \vdash u : F}{\Gamma \vdash u : \forall X : K \cdot F}\ (\forall I)$$

(X_K non libre dans aucune formule de Γ)

On observera, ce qui ne nécessitera pas de démonstration, que, si l'on peut dériver $\Gamma \vdash u : F$, alors on peut dériver $\triangleright F : Prop$; de plus, si Γ s'écrit $x_1 : F_1, \ldots, x_n : F_n$, alors on peut dériver $\triangleright F_i : Prop$ pour tout i, $1 \le i \le n$.

On rappelle qu'un *candidat* C est un ensemble de λ-termes vérifiant:

CR1 tout élément de C est dans SN, l'ensemble des λ-termes fortement normalisants;

CR2 si $u \in C$ et $u \to u'$, alors $u' \in C$;

CR3 si u est neutre, et si (pour tout u' tel que $u \to u'$, $u' \in C$), alors $u \in C$; (parenthèses ajoutées pour bien montrer la portée du quantificateur universel sur u'.)

Un terme est neutre s'il ne commence pas par λ. On note $\mathcal{CR}$ l'ensemble de tous les candidats, et Λ l'ensemble des λ-termes (les termes de preuve).

On définit une interprétation des sortes et des expressions logiques comme suit.

Sortes $[\![Prop]\!] = \mathcal{CR}$, $[\![\iota]\!] = SN$, $[\![K \to K']\!]$ est l'ensemble de toutes les fonctions (totales) de $[\![K]\!]$ vers $[\![K']\!]$;

Expressions logiques Une expression logique S de sorte K est interprétée dans un *contexte* ρ, qui à chaque variable $X_{K'}$ associe un élément de $[\![K']\!]$. Formellement, on définit une interprétation d'une expression logique $S : K$ dans ρ, notée $[\![S : K]\!]\rho$. On pose:

$$\begin{array}{rcl}
[\![X_K : K]\!]\rho & = & \rho(X_K) \\
[\![\mathrm{Imp} : Prop \to Prop \to Prop]\!]\rho & = & (C_1 \in \mathcal{CR} \mapsto (C_2 \in \mathcal{CR} \mapsto C_1 \Rrightarrow C_2)) \\
\text{où } C_1 \Rrightarrow C_2 & = & \{u \in \Lambda \mid \forall v \in C_1 \cdot uv \in C_2\} \\
[\![\mathrm{All}_K : (K \to Prop) \to Prop]\!]\rho & = & (f \in [\![K \to Prop]\!] \mapsto \bigcap_{a \in [\![K]\!]} f(a)) \\
\text{où } \bigcap_{a \in A} f(a) & = & \{u \in \Lambda \mid \forall a \in A \cdot u \in f(a)\} \\
[\![ST : K']\!]\rho & = & [\![S : K \to K']\!]\rho([\![T : K]\!]\rho) \\
[\![\lambda X_K \cdot T : K \to K']\!]\rho & = & (a \in [\![K]\!] \mapsto [\![T : K']\!](\rho[X_K \mapsto a]))
\end{array}$$

où la notation $(a \in A \mapsto f(a))$ dénote la fonction qui à tout élément a de A associe la quantité $f(a)$, et $\rho[X_K \mapsto a]$ est le contexte qui à X_K associe a, et à toute autre variable Y associe $\rho(Y)$. Dans le cas de $ST : K'$, K est l'unique sorte telle que $\triangleright T : K$ soit dérivable.

On notera que l'on a:

$$[\![F_1 \Rightarrow F_2]\!]\rho \quad = \quad \{u \in \Lambda \mid \forall v \in [\![F_1]\!]\rho \cdot uv \in [\![F_2]\!]\rho\}$$

C'est une vérification élémentaire, qui ne dépend que du fait que $[\![F_1]\!]\rho$ et $[\![F_2]\!]\rho$ soient effectivement des candidats (ce que l'on demandera de montrer plus bas). On rapprochera cette formule de la formule définissant l'ensemble $RED^{\rho}_{F_1 \Rightarrow F_2}$ dans le cours, où ρ était un contexte de candidats: $RED^{\rho}_{F_1 \Rightarrow F_2} = \{u \in \Lambda \mid \forall v \in RED^{\rho}_{F_1} \cdot uv \in RED^{\rho}_{F_2}\}$.

On notera que l'on a aussi:

$$[\![\forall X_K \cdot F]\!]\rho \quad = \quad \{u \in \Lambda \mid \forall a \in [\![K]\!] \cdot u \in [\![F]\!](\rho[X_K \mapsto a])\}$$

dès que l'on peut assurer que $[\![F]\!](\rho[X_K \mapsto a])$ est bien dans $[\![Prop]\!] = \mathcal{CR}$ pour tout $a \in [\![K]\!]$. En particulier,

$$[\![\forall X : Prop \cdot F]\!]\rho \quad = \quad \{u \in \Lambda \mid \forall C \in \mathcal{CR} \cdot u \in [\![F]\!](\rho[X_{Prop} \mapsto C)\}$$

ce que l'on rapprochera de la formule du cours $RED^{\rho}_{\forall\alpha\cdot F} = \{u \in \Lambda \mid \forall C \in \mathcal{CR} \cdot u \in RED^{\rho[\alpha\mapsto C]}_F\}$.

1. Montrer que $[\![K]\!]$ n'est pas vide, pour aucune sorte K.

   Par récurrence structurelle sur K. C'est évident lorsque $K = \iota$ ou K est une sorte flèche. Dans le cas où $K = Prop$, $[\![K]\!] = \mathcal{CR}$, et il s'agit de montrer qu'il existe un candidat. Or on a vu en cours que SN lui-même était un candidat.

2. Montrer que pour toute expression logique $T : K$, pour tout contexte ρ, $[\![T : K]\!]\rho$ est un élément de $[\![K]\!]$. (Indication: attention, lorsque A est vide, on a $\bigcap_{a \in A} f(a) = \Lambda$. Si vous ne voyez pas pourquoi je dis ça à la fin de la question, vous avez raté un point important.)

   On procède bien sûr par récurrence structurelle sur T. C'est évident si T est une variable X_K, si T est une application, ou une λ-abstraction. Restent les cas des constantes Imp et All_K, qui sont très similaires aux cas déjà vus en cours pour le système F.

   Imp. On doit montrer que pour tous candidats C_1 et C_2, $C_1 \Rrightarrow C_2$ est encore un candidat. C'est essentiellement l'argument usuel du cours montrant **CR1**, **CR2**, **CR3** *sur les types implication. Voici donc l'argument complet:*

   CR1 *Soit $u \in C_1 \Rrightarrow C_2$. Puisque C_1 est un candidat, par la propriété* **CR3** *sur C_1, tous les termes normaux, donc en particulier les variables x, sont dans C_1. Donc ux est dans C_2. Par* **CR1** *sur C_2, ux est fortement normalisant, donc u aussi.*

   CR2 *Soit $u \in C_1 \Rrightarrow C_2$, et supposons $u \to u'$. Pour tout $v \in C_1$, $uv \in C_2$ par définition, et $uv \to u'v$, donc par* **CR2** *sur C_2, on a $u'v \in C_2$. Comme v est arbitraire, u' est dans $C_1 \Rrightarrow C_2$.*

   CR3 *Soit u un terme neutre dont tous les réduits en une étape sont dans $C_1 \Rrightarrow C_2$. Montrons que pour tout $v \in C_1$, uv est dans C_2, ce qui établira que u est aussi dans $C_1 \Rrightarrow C_2$. Nous pouvons de plus le montrer par récurrence sur la longueur de la plus grande réduction partant de v, puisque par* **CR1** *sur C_1, v est fortement normalisant. Enfin, uv étant neutre, il suffit de regarder tous les réduits en une étape de uv et d'établir qu'ils sont tous dans C_2: c'est* **CR3** *appliqué à C_2. Or, les réduits en une étape de uv sont soit de la forme $u'v$ avec $u \to u'$ soit de la forme uv' avec $v \to v'$. Il n'y a pas d'autre cas car u est neutre. Dans le premier cas, on conclut par le fait que $u' \in C_1 \Rrightarrow C_2$ par hypothèse. Dans le second cas, on conclut par hypothèse de récurrence, sachant que v' est encore dans C_2, par* **CR2** *sur C_2.*

   All_K. Il s'agit de montrer que, pour toute fonction f de $[\![K]\!]$ vers $\mathcal{CR}$, $\bigcap_{a \in [\![K]\!]} f(a)$ est encore un candidat. L'indication montre que si $[\![K]\!]$ était vide, nous devrions montrer que Λ est un candidat. Or c'est clairement faux, puisque par exemple $(\lambda x \cdot xx)(\lambda x \cdot xx)$ est dans Λ mais pas dans SN, donc dans aucun candidat par **CR1**. *Il s'agit donc d'abord d'observer que $[\![K]\!]$* ne peut pas *être vide, par la question 1. Le reste est maintenant facile, et découle du fait que toute intersection $\bigcap_{a \in A} f(a)$ de candidats $f(a)$ est encore un candidat, dès que A est non vide:*

CR1 *Soit $u \in \bigcap_{a\in A} f(a)$. Comme A est non vide, on peut choisir un $a \in A$. Alors $u \in f(a)$. Comme $f(a)$ est un candidat, par* **CR1** *sur ce candidat on obtient $u \in SN$.*

CR2 *Soit $u \in \bigcap_{a\in A} f(a)$, et supposons $u \to u'$. Pour chaque $u \in A$, on a $u \in f(a)$, et par* **CR2** *sur le candidat $f(a)$, on en déduit $u' \in f(a)$. Donc $u' \in \bigcap_{a\in A} f(a)$.*

CR3 *Soit u neutre, et supposons que tous ses réduits sont dans $\bigcap_{a\in A} f(a)$. Ils sont donc dans $f(a)$, pour chaque $a \in A$. Par* **CR3** *sur le candidat $f(a)$, u est donc dans $f(a)$. Comme a est arbitraire, u est dans $\bigcap_{a\in A} f(a)$.*

3. Montrer que, pour tous candidats C_1 et C_2, pour tout λ-terme u tel que (pour tout $v \in C_1$, $u[x := v]$ est dans C_2), alors $\lambda x \cdot u$ est dans $C_1 \Rrightarrow C_2$.

   C'est une démonstration que nous avons déjà faite à de nombreuses reprises... avec des notations différentes. Par **CR3**, *$v = x$ est dans C_1. Donc par hypothèse $u = u[x := x]$ est dans C_2. Par* **CR1**, *u est donc fortement normalisant; par* **CR1** *encore, tout $v \in C_1$ est fortement normalisant. Montrons que $(\lambda x \cdot u)v$ est dans C_2 pour tout u vérifiant l'hypothèse (pour tout $v \in C_1$, $u[x := v]$ est dans C_2) et pour tout $v \in C_1$, par récurrence sur la somme des longueurs de réduction partant de u et de v respectivement. Comme $(\lambda x \cdot u)v$ est neutre, il suffit d'appliquer* **CR3** *et de vérifier que tous ses réduits en une étape sont dans C_2. Ces réduits en une étape sont de trois formes possibles:*

   - *$u[x := v]$, qui est dans C_2 par hypothèse;*
   - *$(\lambda x \cdot u')v$ avec $u \to u'$: mais alors u' vérifie encore l'hypothèse (pour tout $v \in C_1$, $u'[x := v]$ est dans C_2) par* **CR2** *sur C_2, et l'on conclut par hypothèse de récurrence;*
   - *$(\lambda x \cdot u)v'$ avec $v \to v'$: mais alors v' est encore dans C_1 par* **CR2** *sur C_1, et l'on conclut encore par hypothèse de récurrence.*

4. On admettra le lemme de substitution classique: si $\triangleright\forall X : K \cdot F : Prop$ et $\triangleright S : K$ sont dérivables, alors $[\![F[X_K := S] : Prop]\!]\rho = [\![F : Prop]\!](\rho[X_K \mapsto [\![S]\!]\rho])$. On admettra aussi que $[\![F : Prop]\!]\rho$ ne dépend que des variables libres de F, autrement dit que si $\rho(X_K) = \rho'(X_K)$ pour toute variable X_K libre dans F, alors $[\![F : Prop]\!]\rho = [\![F : Prop]\!]\rho'$.

   Démontrer que, si $x_1 : G_1, \ldots, x_n : G_n \vdash u : F$ est dérivable, alors pour tout contexte ρ, pour tous termes $v_1 \in [\![G_1 : Prop]\!]\rho, \ldots, v_n \in [\![G_n : Prop]\!]\rho$, on a $u[x_1 := v_1, \ldots, x_n := v_n] \in [\![F : Prop]\!]\rho$.

   C'est encore comme dans le cours. On effectue une récurrence sur la structure de la dérivation de $\Gamma \vdash u : F$, où $\Gamma = x_1 : G_1, \ldots, x_n : G_n$. C'est évident si la dernière règle appliquée est (Ax).

   Dans le cas de $(=_\beta)$, ceci découle du fait que, par une remarque faite plus haut, comme $\Gamma \vdash u : F_1$, $\triangleright F_1 : Prop$ aussi; par hypothèse, $\triangleright F_2 : Prop$. Or il est facile de montrer que, pour toutes expressions logiques $S, T : K$, si $S \to T$ alors $[\![S : K]\!]\rho = [\![T : K]\!]\rho$. On en déduit que si $S =_\beta T$ alors $[\![S : K]\!]\rho = [\![T : K]\!]\rho$ par récurrence sur le nombre de réductions et réductions inverses entre S et T. Dans le cas $S = F_1$, $T = F_2$, $K = Prop$, on en déduit $[\![F_1 : Prop]\!]\rho = [\![F_2 : Prop]\!]\rho$. Or par hypothèse de récurrence, pour tous $v_1 \in [\![G_1 : Prop]\!], \ldots, v_n \in [\![G_n : Prop]\!]$, on a $u[x_1 := v_1, \ldots, x_n := v_n] \in [\![F_1 : Prop]\!]$. On a donc aussi $u[x_1 := v_1, \ldots, x_n := v_n] \in [\![F_2 : Prop]\!]$.

   Dans le cas de $(\Rightarrow E)$, c'est par définition de $[\![F_1 \Rightarrow F_2]\!]\rho = [\![F_1]\!]\rho \Rrightarrow [\![F_2]\!]\rho$. Dans le cas de $(\Rightarrow I)$, c'est par la question précédente. Les cas de $(\forall E)$ et $(\forall I)$ se règlent en utilisant les résultats admis.

5. En déduire que tout terme typable en F_ω est fortement normalisant. Autrement dit, si $\Gamma \vdash u : F$ est dérivable en F_ω, alors u est fortement normalisant.

   En utilisant la question précédente. Par **CR3**, *on peut prendre $v_i = x_i$, qui est bien dans $[\![G_i : Prop]\!]\rho$— et ceci a une sens par la remarque liminaire selon laquelle, si $\Gamma \vdash u : F$ est dérivable, avec $\Gamma = x_1 : G_1, \ldots, x_n : G_n$, alors $\triangleright G_i : Prop$ aussi est dérivable. On en déduit que $u \in [\![F : Prop]\!]\rho$, ce qui de nouveau a un sens car $\triangleright F : Prop$ est lui aussi dérivable. Par* **CR1**, *on en déduit que $u \in SN$.*

6. En déduire que F_ω est cohérente, au sens où il n'existe pas de terme de preuve u tel que $\vdash u : \bot$, où le faux logique, $\bot$, est défini comme $\forall X : Prop \cdot X$.

Par la question précédente, et la propriété d'auto-réduction, qui se démontre facilement comme en cours, si un tel u existait, on pourrait supposer qu'il est en forme normale. La dérivation de $\vdash u : \bot$ est alors de la forme:

$$\begin{array}{c} \vdots\ \pi \\ \vdash u : F \\ \vdots \\ \vdash u : \bot \end{array}$$

où le passage de $\vdash u : F$ à $\vdash u : \bot$ est une suite d'applications des règles $(=_\beta)$, $(\forall_E)$, et $(\forall_I)$, et la dernière règle de π n'est pas l'une de ces règles. Écrivons F sous la forme $\forall X_1, \ldots, X_k \cdot G$, où G ne commence pas par $\forall$, et montrons que G n'est pas non plus de la forme $G_1 \Rightarrow G_2$ pour aucunes formules G_1, G_2.

Pour ceci, montrons par récurrence sur le nombre n de règles appliquées entre $\vdash u : F$ et $\vdash u : \bot$ que F n'est pas β-convertible à une formule de la forme $\forall X_1, \ldots, X_k \cdot G_1 \Rightarrow G_2$.

Si $n = 0$, $F = \bot = \forall X_{Prop} \cdot X$ n'est pas convertible à une formule de la forme $\forall X_1, \ldots, X_k \cdot G_1 \Rightarrow G_2$. En effet, le λ-calcul des expressions logiques étant confluent, ceci impliquerait que $\forall X_{Prop} \cdot X$ serait la forme normale de $\forall X_1, \ldots, X_k Imp\ G_1\ G_2$. Mais celle-ci est nécessairement de la forme $\forall X_1, \ldots, X_k Imp\ G'_1\ G'_2$, avec G'_1 la forme normale de G_1 et G'_2 celle de G_2; ce qui est impossible.

Sinon, on déduit un jugement de la forme $\vdash u : F'$ de $\vdash u : F$ par l'une des règles $(=_\beta)$, $(\forall_E)$, ou $(\forall_I)$, sachant par hypothèse de récurrence que F' n'est pas convertible à une formule de la forme $\forall X_1, \ldots, X_k \cdot G_1 \Rightarrow G_2$. Si la règle est $(=_\beta)$, F non plus.

Si cette règle est $(\forall_E)$, alors F serait de la forme $\forall Y_K \cdot F_1(Y_K)$, avec $F' = F_1(S)$ pour un certain $S : K$. Si F était convertible à une forme de la forme $\forall X_1, \ldots, X_k \cdot G_1 \Rightarrow G_2$, elle se réduirait par confluence à une formule de cette forme, car une expression logique de la forme $\forall Z \cdot H$ ne peut se réduire qu'à une expression de la forme $\forall Z \cdot H'$ avec H' un réduit de H. Par ce même argument, on en déduit que $F_1(X_1)$ se réduirait à $\forall X_2, \ldots, X_k \cdot G_1 \Rightarrow G_2$, donc $F' = F_1(S)$ se réduirait à $\forall X_2, \ldots, X_k \cdot G_1 \Rightarrow G_2[X_1 := S]$, ce qui est impossible.

Enfin, si la règle est $(\forall I)$, alors F' serait de la forme $\forall X_1 \cdot F$. Comme F' n'est pas convertible à une forme de la forme $\forall X_1, \ldots, X_k \cdot G_1 \Rightarrow G_2$, c'est-à-dire, par confluence, que sa forme normale n'est pas de cette forme, il en est de même de F.

Une autre stratégie pour montrer ce fait était de normaliser la dérivation de u en éliminant les détours $(\forall_I)/(\forall_E)$, et de forcer à prendre les β-formes normales de toutes les formules systématiquement. Ceci implique juste d'utiliser $(=_\beta)$ juste derrière $(\forall_I)$ systématiquement, pour calculer la β-forme normale de $F[X_K := S]$. Ceci permet ensuite de simplifier l'étude du passage de $\vdash u : F$ à $\vdash u : \bot$.

Nous savons donc maintenant que π est une dérivation de $\vdash u : F$, où F est de la forme $\forall X_1, \ldots, X_k \cdot G$, et où G n'est ni une implication ni une quantification universelle. Écrivons u sous forme normale $\lambda x_1, \ldots, x_m \cdot x u_1 \ldots u_n$. La dernière règle de π ne peut alors pas être (Ax), le contexte étant vide, ni $(\Rightarrow I)$, sinon F serait une implication. Donc $m = 0$. Or x est alors libre dans u, mais le contexte de typage est vide, ce qui est impossible. Il est en effet facile de voir par récurrence sur la dérivation de $\Gamma \vdash v : G$ que toute variable libre de u est nécessairement dans le domaine de Γ.