Quelques modèles de graphes du λ-calcul

1 Les modèles de graphes

Les modèles de graphes du λ-calcul sont ainsi nommés parce qu'ils codent les fonctions comme certains sous-ensembles de leur graphe. Par exemple $\mathbb{P}\omega$ est un modèle de graphes.

Plutôt que de parler de cpos comme dans les notes de cours, on va utiliser les dcpos, comme pendant les séances de cours. Une famille F d'un ensemble ordonné est *dirigée* si et seulement si toute partie finie de F est majorée dans F. Si on écrit $F = (x_i)_{i \in I}$, ceci revient à dire que I est non vide (autrement dit, F est non vide) et pour tous $i, j \in I$ il existe $k \in I$ tel que $x_i, x_j \leq x_k$. Un *dcpo* est un ensemble ordonné dans lequel toute partie dirigée a une borne supérieure (on dira aussi supremum ou sup).

Les notes de cours parlent de cpos, où seulement les chaînes — des parties dirigées particulières — ont un sup.

Une fonction f d'un dcpo vers un autre est *Scott-continue* si et seulement si elle est *monotone* ($x \leq y$ implique $f(x) \leq f(y)$) et préserve les sups dirigés (si x est le sup de la famille dirigée $(x_i)_{i \in I}$, alors $f(x)$ est le sup de la famille, nécessairement dirigée, $(f(x_i))_{i \in I}$). C'est une notion légèrement différente de celle de fonction continue des notes de cours.

On note $\mathbb{P}(D)$ l'ensemble des parties de D. C'est un treillis complet, donc un dcpo, pour l'ordre d'inclusion $\subseteq$. On note $\mathbb{P}_{\text{fin}}(D)$ le sous-ensemble des parties finies de D. Pour tous dcpos X et Y, on notera $[X \to Y]$ l'ensemble des fonctions (totales) Scott-continues de X vers Y, ordonné par l'ordre point à point : $f \leq g$ ssi $f(x) \leq g(x)$ pour tout $x \in X$. $[X \to Y]$ est encore un dcpo. Tout cela sera admis.

En général, on appelle *modèle de graphes* un couple $(D, \to_D)$, où D est un ensemble non vide et $\to_D$ est une application (fonction totale) injective de $\mathbb{P}_{\text{fin}}(D) \times D$ dans D. On notera $E \to_D d$ au lieu de $\to_D (E, d)$. Comme pour les types, on estimera que $\to_D$ associe à droite, c'est-à-dire que $E_1 \to_D E_2 \to_D \cdots \to_D E \to_D d$ signifie $E_1 \to_D (E_2 \to_D \cdots \to_D (E \to_D d) \cdots)$.

On note r_D la fonction de $[\mathbb{P}(D) \to \mathbb{P}(D)]$ vers $\mathbb{P}(D)$ définie par :

$$r_D(f) = \{E \to_D d \mid E \in \mathbb{P}_{\mathsf{fin}}(D), d \in f(E)\}. \tag{1}$$

On notera bien que $f(E)$ est la valeur de la fonction f sur l'argument E, et pas l'image $\{f(x) \mid x \in E\}$ de l'ensemble E par f, ce qui n'aurait de toute façon aucun sens.

Réciproquement, on définit $i_D \colon \mathbb{P}(D) \to [\mathbb{P}(D) \to \mathbb{P}(D)]$ par :

$$i_D(A)(B) = \{d \in D \mid \exists E \in \mathbb{P}_{\mathsf{fin}}(D). E \subseteq B \text{ et } (E \to_D d) \in A\}. \tag{2}$$

(J'utiliserai en général E pour une partie finie de D, et A, B, pour des parties arbitraires. Je vous incite fortement à obéir à cette convention, utilisée par souci de lisibilité.)

1. Montrer que si E est une partie finie de D, et si B est le supremum (l'union) d'une famille dirigée $(B_i)_{i \in I}$ de $\mathbb{P}(D)$, avec $E \subseteq B$, alors il existe un $i \in I$ tel que $E \subseteq B_i$.
2. En déduire que, pour n'importe quel $A \in \mathbb{P}(D)$, $i_D(A)$ est bien une fonction Scott-continue, c'est-à-dire un élément de $[\mathbb{P}(D) \to \mathbb{P}(D)]$.

On admettra dans la suite les résultats suivants :

— i_D elle-même est Scott-continue de $\mathbb{P}(D)$ vers $[\mathbb{P}(D) \to \mathbb{P}(D)]$;

— r_D est Scott-continue ;

— $i_D \circ r_D = \mathrm{id}_{\mathbb{P}(D)}$;

— D est un modèle du λ-calcul, au sens où la fonction de sémantique :

  — $D\llbracket x \rrbracket \rho = \rho(x)$,

  — $D\llbracket uv \rrbracket \rho = i_D(D\llbracket u \rrbracket \rho)(D\llbracket v \rrbracket \rho)$,

  — $D\llbracket \lambda x \cdot u \rrbracket \rho = r_D(A \mapsto D\llbracket u \rrbracket (\rho[x := A]))$

  a la propriété que pour tout couple de λ-termes β-équivalents u et v, pour tout environnement ρ, $D\llbracket u \rrbracket \rho = D\llbracket v \rrbracket \rho$.

— La fonction qui à $A \in \mathbb{P}_{\mathsf{fin}}(D)$ associe $D\llbracket u \rrbracket (\rho[x := A])$, pour n'importe quel terme u et n'importe quel environnement ρ, est Scott-continue, et en particulier monotone.

Pour un terme clos u, $D\llbracket u \rrbracket \rho$ ne dépend pas de ρ, et on notera donc cette valeur simplement $D\llbracket u \rrbracket$.

3. Montrer qu'aucun modèle de graphe $(D, \to_D)$ ne valide la η-règle. Précisément, montrer que, pour deux variables distinctes x et y, $D\llbracket \lambda x \cdot x \rrbracket \neq D\llbracket \lambda x, y \cdot xy \rrbracket$. Je veux l'argument le plus simple possible.
4. Soit $\Omega = \delta\delta$, où $\delta = \lambda x.xx$. Montrer que si $d \in D\llbracket \Omega \rrbracket$, alors il existe un $E \in \mathbb{P}_{\mathsf{fin}}(D)$ tel que $(E \to_D d) \in E$.

2 Normalisation de tête et modèles de graphes

On rappelle que $\mathbf{V} = \lambda x, y \cdot x$ et $\mathbf{F} = \lambda x, y \cdot y$.

Le modèle d'Engeler $\mathcal{E}_C$ au-dessus d'un ensemble non vide C, vu en TD, est le modèle de graphes où :

— D est l'ensemble des arbres finis dont les feuilles sont des éléments de C, et dont les nœuds internes ont deux fils : le fils gauche consistant en un ensemble fini de sous-arbres, le fils gauche étant un unique sous-arbre ;

— $E \to_{\mathcal{E}_C} d$ est l'arbre donc le fils gauche est l'ensemble fini E, et le fils droit est d. On le notera simplement $E \to d$.

5. Montrer que $\mathcal{E}_C [\![\Omega]\!] = \emptyset$.
6. Un *test de normalisabilité de tête* est un λ-terme clos t tel que, pour tout terme u qui a une forme normale de tête, $tu =_\beta \mathbf{V}$, et pour tout terme u qui n'a pas de forme normale de tête, $tu =_\beta \mathbf{F}$. On rappelle que pour tout $A \in \mathbb{P}(\mathcal{E}_C)$, $i_{\mathcal{E}_C}(A)$ est une fonction Scott-continue, donc monotone. En déduire ainsi que de la question précédente que, s'il existe un test de normalisabilité de tête t, alors $\mathcal{E}_C [\![\mathbf{F}]\!] \subseteq \mathcal{E}_C [\![\mathbf{V}]\!]$.
7. En déduire qu'il n'existe pas de test de normalisabilité de tête. On pourra comparer $\mathcal{E}_C [\![\mathbf{F}xy]\!] \rho$ et $\mathcal{E}_C [\![\mathbf{V}xy]\!] \rho$, où x et y sont deux variables distinctes.

3 Résolubilité

Si $\vec{u}$ est une suite finie de termes $u_1, u_2, \ldots, u_n$, on notera $t\vec{u}$ pour $tu_1u_2 \cdots u_n$. (On peut avoir $n = 0$.)

Un terme clos t est *résoluble* si et seulement s'il existe une telle suite $\vec{u}$ telle que $t\vec{u} =_\beta \mathbf{I}$, où $\mathbf{I}$ est par définition le terme $\lambda x \cdot x$.

Un terme t, non nécessairement clos, est résoluble si et seulement si le terme $\lambda x_1, \cdots, x_m \cdot t$, où $x_1, \ldots, x_m$ sont les variables libres de t, est résoluble. (Le terme en question est clos, et l'on applique donc la définition précédente de la résolubilité.)

8. Montrer que tout terme t qui a une forme normale de tête (au sens où t se β-réduit en un nombre quelconque de β-réductions, non nécessairement en tête, vers une forme normale de tête) est résoluble.
9. Montrer que, quel que soit le modèle de graphes $(D, \to_D)$, pour tout terme clos résoluble t, $D [\![t]\!] \neq \emptyset$.
10. Donner un exemple de terme non résoluble. Justifier. On n'utilisera pas les résultats des questions qui suivent.

4 K-candidats

On rappelle la notion suivante, vue en cours dans le cadre de la preuve de forte normalisation du λ-calcul simplement typé : pour deux ensembles S et S' de λ-termes,

$$S \Rightarrow S' = \{u \mid \forall v \in S.uv \in S'\}.$$

On dira qu'un ensemble S de λ-termes est *saturé* si et seulement s'il est clos par réduction de tête faible inverse, autrement dit si $u[x := t]v_1v_2 \cdots v_n \in S$ implique $(\lambda x \cdot u)tv_1v_2 \cdots v_n$.

11. Montrer que si S' est saturé, alors $S \Rightarrow S'$ est saturé pour tout ensemble S.

On note Λ l'ensemble de tous les λ-termes.

On note S_h l'ensemble des termes t dont la réduction de tête termine, autrement dit qui n'a pas de réduction infinie de la forme $t \rightarrow_t t_1 \rightarrow_t t_2 \rightarrow_t \cdots$.

On note S_0 l'ensemble des termes de la forme $x\vec{u}$ où x est une variable et $\vec{u}$ une suite quelconque de λ-termes. S_0 est trivialement saturé et inclus dans S_h.

On dira que S est un *K-candidat* (pour « candidat de Krivine ») si et seulement si S est saturé et $S_0 \subseteq S \subseteq S_h$.

12. Soient t un λ-terme, et x une variable. Supposons que la réduction de tête de t ne termine pas. Montrer que celle de tx ne termine pas non plus.
13. Montrer que pour tout K-candidat S', et pour tout ensemble S de λ-termes contenant au moins une variable, $S \Rightarrow S'$ est un K-candidat. Où utilise-t-on l'hypothèse que S contient une variable ?

On en déduit que $\Rightarrow$ est une opération qui envoie tout couple de K-candidats vers un K-candidat (puisque tout K-candidat contient S_0, donc toutes les variables, donc au moins une), et aussi que $\Lambda \Rightarrow S$ est un K-candidat pour tout K-candidat S. (Noter que Λ n'est pas un K-candidat.)

On notera aussi que S_h est un K-candidat (le plus grand), et que toute intersection d'une famille non vide de K-candidats est un K-candidat. Pour toute famille $(S_i)_{i \in I}$ de K-candidats, on notera $\bigwedge_{i \in I} S_i$ l'intersection $\bigcap_{i \in I} S_i$ si I est non vide, et S_h si $I = \emptyset$. C'est la borne inférieure de la famille dans le treillis complet des K-candidats.

Pour chaque arbre $d \in \mathcal{E}_C$, on définit un K-candidat $I(d)$ par récurrence sur la taille de d comme suit :

$$\begin{aligned} I(c) &= S_h \text{ pour toute feuille } c \in C \\ I(E \rightarrow d) &= I(E) \Rightarrow I(d), \end{aligned}$$

où l'on pose pour tout ensemble (fini ou non) A, $I(A) = \bigwedge_{d' \in A} I(d')$; ce qui permet notamment de définir $I(E)$ dans la définition de $I(E \rightarrow d)$ ci-dessus.

14. Montrer que pour tout λ-terme t, pour tous ρ, θ, d satisfaisant les hypothèses suivantes :
    (a) ρ est un environnement associant à toute variable un élément de $\mathbb{P}(\mathcal{E}_C)$,
    (b) $\theta = [x_1 := v_1, \cdots, x_n := v_n]$ est une substitution de domaine $\{x_1, \cdots, x_n\}$ contenant au moins les variables libres de t,
    (c) pour tout i tel que $1 \leq i \leq n$, $v_i \in I(\rho(x_i))$,
    (d) $d \in \mathcal{E}_C [\![t]\!] \rho$,

    alors $t\theta \in I(d)$.
15. En déduire que pour tout terme t tel que $\mathcal{E}_C [\![t]\!] \rho \neq \emptyset$ (pour au moins un environnement ρ), la réduction de tête partant de t termine. Montrez clairement où vous utilisez les questions précédentes, et la définition des K-candidats.

En mettant ensemble les résultats des questions 8, 9, et 15, plus le fait évident que tout terme dont la réduction de tête termine a une forme normale de tête, on en déduit que les notions suivantes sont équivalentes pour un terme clos t :

a la réduction de tête de t termine ;

b t a une forme normale de tête ;

c t est résoluble ;

d $\mathcal{E}_C [\![t]\!] \neq \emptyset$.

5 Normalisation faible

Cette partie a pour but de redémontrer, en particulier, le théorème de standardisation. Il est donc interdit d'y faire appel ! (Même indirectement, mais aucun résultat du cours autre que le théorème de standardisation ne l'utilise.)

On dira qu'un terme est *normalisable par la gauche* si et seulement si sa réduction gauche (leftmost-outermost) termine. On note S_ℓ l'ensemble des termes normalisables par la gauche. Il est facile de voir que S_ℓ est saturé.

On propose de suivre une démonstration similaire à celle de la section précédente. On redéfinit pour cela certaines des notions :

— S'_0 est l'ensemble des termes de la forme $xu_1u_2 \cdots u_n$, $n \geq 0$, où tous les u_i sont normalisables par la gauche ;

— un K'-candidat est un ensemble S saturé tel que $S'_0 \subseteq S \subseteq S_\ell$;

— les éléments *stricts* de $\mathcal{E}_C$ sont les arbres dont aucun nœud interne n'a un fils gauche vide ; autrement dit, ils sont définis par récurrence comme ceux de la forme c, $c \in C$, ou bien $E \to d$ où E est *non vide*, d est strict, et tous les éléments de E sont stricts.

16. Montrer, en adaptant les techniques des parties précédentes, que pour un terme clos, les trois affirmations suivantes sont équivalentes :

    a t a une forme normale ;

    b $\mathcal{E}_C \llbracket t \rrbracket$ contient un élément strict ;

    c t est normalisable par la gauche.

6 Bonus

Vous avez dû voir avec Guillaume Bury, en TD, des théorèmes comme celui-ci, portant sur des systèmes de typage *conjonctifs* : les λ-termes fortement normalisants sont exactement ceux qui sont typables dans le système $\mathcal{D}$; les λ-termes normalisables sont exactement ceux qui sont typables dans le système $\mathcal{D}_\Omega$ avec des types où Ω n'intervient qu'en position négative.

17. Formaliser la relation entre système $\mathcal{D}_\Omega$ et modèle d'Engeler, et traduire les conditions « $\mathcal{E}_C \llbracket t \rrbracket \neq \emptyset$ » et « $\mathcal{E}_C \llbracket t \rrbracket$ contient un élément strict » en conditions de typage. Qu'observez-vous ?