Programmation, partiel: sémantique de Lisp

Correction.

À rendre pour le vendredi 05 décembre 2003
en version papier : Jean Goubault-Larrecq, LSV ;
ou électronique : `goubault@lsv.ens-cachan.fr`

Recommandations. Votre copie (papier ou électronique) devra être lisible et bien structurée. La note tiendra compte autant du fond que de la présentation. D'autre part, vous devrez justifier toutes vos affirmations, soit par une preuve soit par un contre-exemple. (Ces recommandations sont bien sûr valables dans tout devoir.)

Contexte. Dans le devoir, il sera fait plusieurs fois références à la sémantique de mini-Caml. On pourra se référer à `http://www.lsv.ens-cachan.fr/~goubault/cours.html`.

Lisp est l'un des premiers langages fonctionnel. Nous en considérons une version simplifiée et adaptée. La syntaxe des expressions est :

e	::=	x	variables
	$\mid$	$(e\ e)$	application
	$\mid$	$(\texttt{lambda}\ (x)\ e)$	fonction
	$\mid$	$(\texttt{let}\ (x\ e)\ e)$	abréviation
	$\mid$	$e; e$	séquence
	$\mid$	$(\texttt{setq}\ x\ e)$	affectation

En voici la sémantique naturelle, qui permet de dériver des jugements de la forme $\rho \vdash e \Rightarrow \rho', V$, où ρ et ρ' sont des *environnements*, c'est-à-dire des applications qui à chaque variable associent des valeurs, et où V est une *valeur*, définie par la grammaire :

V	::=	n	entier (dans $\mathbb{Z}$)
	$\mid$	$\langle x, e\rangle$	fonction définie

Les règles de la sémantique naturelle sont les suivantes :

$$\frac{}{\rho \vdash x \Rightarrow \rho, \rho(x)}\,(Var)$$

$$\frac{\rho \vdash e_1 \Rightarrow \rho_1, \langle x, e_3\rangle \quad \rho_1 \vdash e_2 \Rightarrow \rho_2, V_2 \quad \rho_2[x := V_2] \vdash e_3 \Rightarrow \rho', V}{\rho \vdash (e_1\ e_2) \Rightarrow \rho'[x := \rho_2(x)], V}\,(App)$$

$$\frac{}{\rho \vdash (\texttt{lambda}\ (x)\ e) \Rightarrow \rho, \langle x, e\rangle}\,(Fun)$$

$$\frac{\rho \vdash e_1 \Rightarrow \rho_1, V_1 \quad \rho_1[x := V_1] \vdash e_2 \Rightarrow \rho_2, V_2}{\rho \vdash (\texttt{let}\ (x\ e_1)\ e_2) \Rightarrow \rho_2[x := \rho_1(x)], V_2}\,(Let)$$

$$\frac{\rho \vdash e_1 \Rightarrow \rho_1, V_1 \quad \rho_1 \vdash e_2 \Rightarrow \rho_2, V_2}{\rho \vdash e_1; e_2 \Rightarrow \rho_2, V_2}\,(Seq)$$

$$\frac{\rho \vdash e \Rightarrow \rho', V}{\rho \vdash (\texttt{setq}\ x\ e) \Rightarrow \rho'[x := V], V}\,(Assign)$$

1. Écrire la dérivation de sémantique naturelle la plus générale (dans un sens intuitif) pour les expressions :

   $$(a)\ (\texttt{setq}\ x\ (\texttt{lambda}\ (y)\ y)); (x\ z) \qquad (b)\ (\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (x\ z))$$

   Quelle est la différence entre les deux expressions ?

   Pour (a), posons $\rho' = \rho[x := \langle y, y\rangle]$. Alors :

   $$\frac{\dfrac{\dfrac{}{\rho \vdash (\texttt{lambda}\ (y)\ y) \Rightarrow \rho, \langle y, y\rangle}\,(Fun)}{\rho \vdash (\texttt{setq}\ x\ (\texttt{lambda}\ (y)\ y)) \Rightarrow \rho', \langle y, y\rangle}\,(Assign) \quad \dfrac{\dfrac{}{\rho' \vdash x \Rightarrow \rho', \langle y, y\rangle}\,(Var) \quad \dfrac{}{\rho'[y := \rho(z)] \vdash y \Rightarrow \rho'[y := \rho(z)], \rho(z)}\,(Var)}{\rho' \vdash (x\ z) \Rightarrow \rho', \rho(z)}\,(App)}{\rho \vdash (\texttt{setq}\ x\ (\texttt{lambda}\ (y)\ y)); (x\ z) \Rightarrow \rho', \rho(z)}$$

   Pour (b) :

   $$\frac{\dfrac{}{\rho \vdash (\texttt{lambda}\ (y)\ y) \Rightarrow \rho, \langle y, y\rangle} \quad \dfrac{\dfrac{}{\rho' \vdash x \Rightarrow \rho', \langle y, y\rangle}\,(Var) \quad \dfrac{}{\rho'[y := \rho(z)] \vdash y \Rightarrow \rho'[y := \rho(z)], \rho(z)}\,(Var)}{\rho' \vdash (x\ z) \Rightarrow \rho', \rho(z)}\,(App)}{\rho \vdash (\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (x\ z)) \Rightarrow \rho, \rho(z)}\,(Let)$$

   La différence est que l'environnement final est ρ' dans le cas (a), ρ dans le cas (b).

2. Les constructions $(\texttt{let}\ (x\ e)\ e')$ et $((\texttt{lambda}\ (x)\ e')\ e)$ sont-elles équivalentes ? On dira que deux expressions e et e' sont *équivalentes* si et seulement si : $\rho \vdash e \Rightarrow \rho', V$ est dérivable si et seulement si $\rho \vdash e' \Rightarrow \rho', V$ l'est. Justifier.

   Elles le sont. En effet, les dérivations possibles de la première construction sont :

   $$\frac{\begin{array}{cc}\vdots & \vdots \\ \rho \vdash e \Rightarrow \rho_1, V_1 & \rho_1[x := V_1] \vdash e' \Rightarrow \rho_2, V_2\end{array}}{\rho \vdash (\texttt{let}\ (x\ e)\ e') \Rightarrow \rho_2[x := \rho_1(x)], V_2}\,(Let)$$

alors que celles de la seconde sont :

$$\dfrac{\dfrac{}{\rho \vdash (\texttt{lambda}\ (x)\ e') \Rightarrow \rho, \langle x, e' \rangle}\,(Fun) \quad \begin{array}{c}\vdots\\ \rho \vdash e \Rightarrow \rho_1, V_1\end{array} \quad \begin{array}{c}\vdots\\ \rho_1[x := V_1] \vdash e' \Rightarrow \rho_2, V_2\end{array}}{\rho \vdash ((\texttt{lambda}\ (x)\ e')\ e) \Rightarrow \rho_2[x := \rho_1(x)], V_2}\,(App)$$

3. On ajoute au langage une construction `letrec`. La construction :

$$(\texttt{letrec}\ (x\ e)\ e')$$

est vue comme synonyme de

$$(\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e); e')$$

Donner une règle de sémantique naturelle équivalente directe pour `letrec`. Prouvez l'équivalence dans le cas où e est de la forme $(\texttt{lambda}\ (z)\ e_1)$. Précisément, vous devrez prouver qu'un jugement $\rho \vdash (\texttt{letrec}\ (x\ e)\ e') \Rightarrow \rho', V$ est dérivable par votre règle si et seulement si $\rho \vdash (\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e); e') \Rightarrow \rho', V$ l'est.

Il suffit d'écrire la dérivation la plus générale de $(\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e); e')$. *Posons* $\rho' = \rho[x := \langle y, y \rangle]$ *:*

$$\dfrac{\dfrac{}{\rho \vdash (\texttt{lambda}\ (y)\ y) \Rightarrow \rho, \langle y, y \rangle}\,(Fun) \quad \dfrac{\begin{array}{c}\vdots\\ \rho' \vdash e \Rightarrow \rho_1, V_1\end{array}}{\rho' \vdash (\texttt{setq}\ x\ e) \Rightarrow \rho_1[x := V_1], V_1}\,(Assign) \quad \begin{array}{c}\vdots\\ \rho_1[x := V_1] \vdash e' \Rightarrow \rho_2, V_2\end{array}}{\rho' \vdash (\texttt{setq}\ x\ e); e' \Rightarrow \rho_2, V_2}$$

$$\rho \vdash (\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e); e') \Rightarrow \rho_2[x := \rho(x)], V_2$$

La règle souhaitée est donc :

$$\dfrac{\rho[x := \langle y, y \rangle] \vdash e \Rightarrow \rho_1, V_1 \quad \rho_1[x := V_1] \vdash e' \Rightarrow \rho_2, V_2}{\rho \vdash (\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e); e') \Rightarrow \rho_2[x := \rho(x)], V_2}$$

qui est correcte par construction.
Il est ici à noter que $(\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e); e')$ *veut bien sûr dire la même chose que* $(\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ ((\texttt{setq}\ x\ e); e'))$, *et non que* $(\texttt{let}\ (x\ (\texttt{lambda}\ (y)\ y))\ (\texttt{setq}\ x\ e)); e'$.

4. Écrire la dérivation de sémantique naturelle la plus générale pour l'expression

$$\begin{array}{l}(\texttt{setq}\ f\ (\texttt{lambda}\ (x)\ (g\ x)));\\ (\texttt{let}\ (g\ (\texttt{lambda}\ (y)\ \texttt{deux}))\\ \quad (f\ \texttt{un}))\end{array}$$

dans un environnement initial ρ qui à `un` associe la constante $1 \in \mathbb{Z}$ et à `deux` associe $2 \in \mathbb{Z}$.

Posons π_1 *la dérivation :*

$$\dfrac{\dfrac{}{\rho \vdash (\texttt{lambda}\ (x)\ (g\ x)) \Rightarrow \rho, \langle x, (g\ x) \rangle}\,(Fun)}{\rho \vdash (\texttt{setq}\ f\ (\texttt{lambda}\ (x)\ (g\ x))) \Rightarrow \rho[f := \langle x, (g\ x) \rangle], \langle x, (g\ x) \rangle}\,(Assign)$$

Posons d'autre part $A = (\texttt{setq}\ f\ (\texttt{lambda}\ (x)\ (g\ x)))$, *et* $\rho_f = \rho[f := \langle x, (g\ x) \rangle]$.
Posons maintenant $\rho_{fg} = \rho_f[g := \langle y, \texttt{deux} \rangle]$, *et* π_2 *la dérivation :*

$$\dfrac{\dfrac{}{\rho_{fg}[x := 1] \vdash g \Rightarrow \rho_{fg}[x := 1], \langle y, \texttt{deux} \rangle}\,(Var) \quad \dfrac{}{\rho_{fg}[x := 1] \vdash x \Rightarrow \rho_{fg}[x := 1], 1}\,(Var) \quad \dfrac{}{\rho_{fg}[x := 1, y := 1] \vdash \texttt{deux} \Rightarrow \rho_{fg}[x := 1, y := 1], 2}\,(Var)}{\rho_{fg}[x := 1] \vdash (g\ x) \Rightarrow \rho_{fg}[x := 1], 2}\,(App)$$

Posons maintenant π_3 la dérivation :

$$\dfrac{\dfrac{}{\rho_f \vdash (\texttt{lambda}\ (y)\ \texttt{deux}) \Rightarrow \rho_f, \langle y, \texttt{deux}\rangle}\,(Fun) \quad \dfrac{\dfrac{}{\rho_{fg} \vdash f \Rightarrow \rho_{fg}, \langle x, (g\ x)\rangle}\,(Var) \quad \dfrac{\vdots\ \pi_2}{\rho_{fg}[x := 1] \vdash (g\ x) \Rightarrow \rho_{fg}[x := 1], 2} \quad \dfrac{}{\rho_{fg} \vdash \texttt{un} \Rightarrow \rho_{fg}, 1}\,(Var)}{\rho_{fg} \vdash (f\ \texttt{un}) \Rightarrow \rho_{fg}, 2}\,(App)}{\rho_f \vdash (\texttt{let}\ (g\ (\texttt{lambda}\ (y)\ \texttt{deux}))(f\ \texttt{un})) \Rightarrow \rho_f, 2}\,(Let)$$

Soit $B = (\texttt{let}\ (g\ (\texttt{lambda}\ (y)\ \texttt{deux}))(f\ \texttt{un}))$. L'expression dont on cherche à trouver une dérivation de sémantique est $A; B$. La dérivation souhaitée est alors :

$$\dfrac{\vdots\ \pi_1 \quad \vdots\ \pi_3}{\dfrac{\rho \vdash A \Rightarrow \rho_f, \langle x, (g\ x)\rangle \quad \rho_f \vdash B \Rightarrow \rho_f, 2}{\rho \vdash A; B \Rightarrow \rho_f, 2}\,(Seq)}$$

5. Donner une définition intuitive de l'ensemble des variables libres d'une expression Lisp. Vous pourrez vous inspirer de la définition des variables libres d'une expression mini-Caml.

$$\begin{aligned}
\mathrm{fv}(x) &= \{x\} \\
\mathrm{fv}((e_1\ e_2)) &= \mathrm{fv}(e_1) \cup \mathrm{fv}(e_2) \\
\mathrm{fv}((\texttt{lambda}\ (x)\ e)) &= \mathrm{fv}(e) \setminus \{x\} \\
\mathrm{fv}((\texttt{let}\ (x\ e)\ e')) &= \mathrm{fv}(e) \cup (\mathrm{fv}(e') \setminus \{x\}) \\
\mathrm{fv}(e; e') &= \mathrm{fv}(e) \cup \mathrm{fv}(e') \\
\mathrm{fv}((\texttt{setq}\ x\ e)) &= \mathrm{fv}(e)
\end{aligned}$$

Il y a bien entendu différentes possibilités pour la dernière ligne.

6. Pourquoi, contrairement à mini-Caml, la sémantique d'une expression Lisp ne dépend-elle pas que de ses variables libres ?

L'unique problème est que la sémantique d'une expression ayant une variable f libre peut aussi dépendre, si f est une fonction, des variables libres dans la fonction, valeur de f, et ainsi de suite récursivement. Par exemple, si $\rho(f) = \langle x, (g\ x)\rangle$, alors la sémantique de $(f\ e)$ dépend non seulement de la valeur de la variable f (soit $\langle x, (g\ x)\rangle$) et de celles des variables libres dans e, mais aussi de la valeur de g, notamment.

7. Sous quelles hypothèses sur l'environnement ρ la sémantique d'une expression e (à savoir l'ensemble des ρ', V tels que $\rho \vdash e \Rightarrow \rho', V$ est dérivable) ne dépend-elle que des variables libres de e ?

Il suffit de supposer que pour tout x dans le domaine de ρ tel que $\rho(x)$ est de la forme $\langle y, e\rangle$, alors $\mathrm{fv}(e) \subseteq \{y\}$, par exemple. Cette contrainte est cependant difficilement réalisable dans ce langage...

8. La question 4 illustre le fait que Lisp est un langage à *liaison dynamique* : la valeur des variables dépend du moment où l'on en est dans l'exécution du programme. À l'opposé, dans un langage à *portée statique* comme Caml, la valeur des variables ne dépend que de l'endroit dans le texte où elles sont définies.

   On souhaite définir un langage dont les expressions sont les expressions Lisp, mais qui soit à portée statique comme Caml. Noter qu'une différence importante par rapport à Caml est que l'on peut effectuer des effets de bord sur les variables. Donnons quelques exemples de comportements souhaités, en restant informel ; ce sera votre rôle de formaliser la suite. D'abord, l'exemple de la question 4

devrait maintenant retourner comme valeur $G(1)$, si $\rho(g)$ est la fonction G, et $\rho(\texttt{un}) = 1$. Ensuite, on souhaite que les valeurs des fonctions soient des clôtures, mais que les clôtures soient modifiables : après avoir effectué $(\texttt{setq}\ x\ 0); (\texttt{setq}\ (f\ (\texttt{lambda}\ (y)\ (\texttt{setq}\ x\ (+\ x\ 1)); x)))$ (après avoir étendu le langage de sorte que 0 dénote zéro, 1 un et $(+\ e\ e')$ l'addition de e et de e' évaluée de gauche à droite), on souhaite que le premier appel à f retourne 1, le deuxième retourne 2, etc.

Donner une nouvelle sémantique pour Lisp qui tienne compte de ces besoins. Elle devra rester la plus simple et la plus élégante possible. Indication : vous pouvez imaginer une traduction de Lisp vers mini-Caml et déduire la nouvelle sémantique de Lisp à partir de celle de mini-Caml.

Une façon simple est de définir une sémantique à la mini-Caml. Au lieu qu'une variable Lisp x contienne directement une valeur, on va s'arranger pour qu'elle contienne une référence r qui contient la valeur. On reprend des jugements à la mini-Caml, de la forme $\rho, \mu \vdash e \Rightarrow \mu', V$, où $\mu, \mu' \in Mem = Addr \to Val$ et $\rho \in Env = Var \to Addr$. Noter que les environnements envoient des variables vers des addresses, et non des valeurs.

Les valeurs de plus sont comme en mini-Caml, avec de vraies clôtures :

$$\begin{array}{rcll} V & ::= & n & \textit{entier (dans } \mathbb{Z}\textit{)} \\ & | & \langle x, e, \rho \rangle & \textit{clôture} \end{array}$$

La première règle est celle de la lecture des variables. On a maintenant deux indirections : on doit lire l'adresse $\rho(x)$ où est stockée x, puis lire le contenu de cette adresse :

$$\frac{x \in \mathrm{dom}\rho, \rho(x) \in \mathrm{dom}\mu}{\rho, \mu \vdash x \Rightarrow \mu, \mu(\rho(x))}\ (Var')$$

La règle d'application est pratiquement celle de mini-Caml :

$$\frac{\begin{array}{c}\rho, \mu \vdash e_1 \Rightarrow \mu_1, \langle x, e_3, \rho_1 \rangle \quad \rho, \mu_1 \vdash e_2 \Rightarrow \mu_2, V_2 \\ a \in Addr \setminus \mathrm{dom}\mu_2 \quad \rho_1[x := a], \mu_2[a := V_2] \vdash e_3 \Rightarrow \mu_3, V\end{array}}{\rho, \mu \vdash (e_1\ e_2) \Rightarrow \mu_3, V}\ (App')$$

sauf que l'on alloue une adresse a pour stocker le contenu du paramètre x (en pratique, cette adresse est alloué en pile). La sémantique des fonctions est exactement celle de mini-Caml :

$$\frac{}{\rho, \mu \vdash (\texttt{lambda}\ (x)\ e) \Rightarrow \mu, \langle x, e, \rho \rangle}\ (Fun')$$

Pour les `let`*, il suffit de définir* $(\texttt{let}\ (x\ e_1)\ e_2)$ *comme équivalent à* $((\texttt{lambda}\ (x)\ e_2)\ e_1)$ *:*

$$\frac{\rho, \mu \vdash e_1 \Rightarrow \mu_1, V_1 \quad a \in Addr \setminus \mathrm{dom}\mu_1 \quad \rho[x := a], \mu_1[a := V_1] \vdash e_2 \Rightarrow \mu_2, V}{\rho, \mu \vdash (\texttt{let}\ (x\ e_1)\ e_2) \Rightarrow \mu_2, V}\ (Let')$$

On a ensuite la règle évidente :

$$\frac{\rho, \mu \vdash e_1 \Rightarrow \mu_1, V_1 \quad \rho, \mu_1 \vdash e_2 \Rightarrow \mu_2, V_2}{\rho, \mu \vdash e_1; e_2 \Rightarrow \mu_2, V_2}\ (Seq')$$

et enfin :

$$\frac{x \in \mathrm{dom}\rho \quad \rho, \mu \vdash e \Rightarrow \mu_1, V}{\rho, \mu \vdash (\texttt{setq}\ x\ e) \Rightarrow \mu_1[\rho(x) := V], V}\ (Assign')$$

Ce qui est décrit ici est la sémantique du langage Scheme.