Corrigé

13 février 2007

Question 1 Remarquons que la règle O_1 peut aussi s'écrire $\texttt{let}\ x = b\ \texttt{in a} \hookrightarrow a[x \leftarrow b]$ si b est une valeur ou une variable.

$$\begin{array}{rcll} P & \hookrightarrow & \texttt{let}\ y = a\ \texttt{in}\ (\texttt{let}\ x = \lambda z.b\ \texttt{in}\ x\ \texttt{true}) & (O_2) \\ & \hookrightarrow & \texttt{let}\ y = a\ \texttt{in}\ ((\lambda z.b)\ \texttt{true}) & (O_1) \\ & \hookrightarrow & \texttt{let}\ y = a\ \texttt{in}\ b[z \leftarrow \texttt{true}] = P' & (O_1) \end{array}$$

Par-rapport à celle de P, l'exécution de P' économise une liaison `let`, une construction de fermeture (pour $\lambda z.b$) et une application de fonction.

Note : la constante `true` est une survivance d'une ancienne version de ce problème où l'on avait aussi des booléens. Ça marche aussi avec une constante entière.

Note : dans l'esprit du sujet, a et b étaient des termes quelconques du langage. Beaucoup d'étudiants ont interprété a et b comme des noms de variables, auquel cas P s'optimise en b. L'énoncé n'était pas assez clair, donc les deux réponses ont été considérées correctes.

Question 2 On utilise le fait que la relation $\rightarrow$ est déterministe : si $a \rightarrow a'$ et $a \rightarrow a''$, alors $a' = a''$. (Propriété énoncée dans le premier cours.)

Si a diverge : $a \rightarrow a_1 \rightarrow a_2 \rightarrow \ldots$ On a nécessairement $a_1 = a'$ par déterminisme de la relation de réduction. Donc a' diverge également : $a' \rightarrow a_2 \rightarrow \ldots$.

Si $a \xrightarrow{*} N$: on ne peut pas avoir $a = N$ car $a \rightarrow a'$. Donc, nécessairement, $a \rightarrow a_1 \xrightarrow{*} N$. Par déterminisme de la relation de réduction, $a_1 = a'$, et donc $a' \xrightarrow{*} N$.

Question 3 Par hypothèse $a \hookrightarrow a'$ par la règle O_1, on a nécessairement $a = (\lambda x.b)\ c$ et $a' = b[x \leftarrow c]$, et de plus c est une valeur ou une variable. Comme P est clos, c ne peut pas être une variable. C'est donc une valeur, et la réduction $(\lambda x.b)\ c \rightarrow b[x \leftarrow c]$ est valide. D'après le résultat de la question 2, il s'ensuit que a et a' ont le même comportement observable.

Question 4 Non. Soit $\omega = (\lambda x.\ x\ x)\ (\lambda x.\ x\ x)$ un terme qui diverge. Le programme $(\lambda x.0)\ \omega$ diverge, mais sa version "optimisée" 0 ne diverge pas.

Question 5 Remarquons que la notation `let` se réduit de la manière suivante :

$$\frac{a \rightarrow a'}{(\texttt{let}\ x = a\ \texttt{in}\ b) \rightarrow (\texttt{let}\ x = a'\ \texttt{in}\ b)} \qquad\qquad \texttt{let}\ x = v\ \texttt{in}\ b \rightarrow b[x \leftarrow v]$$

Soit $p = \texttt{let}\ x = (\texttt{let}\ y = a\ \texttt{in}\ b)\ \texttt{in}\ c$. Une séquence de réductions de p se compose :

- d'une séquence de réductions (possiblement infinie) de a dans le contexte $\texttt{let}\ x = (\texttt{let}\ y = [\,]\ \texttt{in}\ b)\ \texttt{in}\ c$;
- si cette séquence termine sur une valeur v_a, d'une séquence de réductions de $b[y \leftarrow v_a]$ dans le contexte $\texttt{let}\ x = [\,]\ \texttt{in}\ c$;
- si cette séquence termine sur une valeur v_b, d'une séquence de réductions de $c[x \leftarrow v_b]$ dans le contexte $[\,]$.

De même, une séquence de réductions de $p' = \texttt{let}\ y = a\ \texttt{in}\ (\texttt{let}\ x = b\ \texttt{in}\ c)$ se compose :

- d'une séquence de réductions (possiblement infinie) de a dans le contexte $\texttt{let}\ y = [\,]\ \texttt{in}\ (\texttt{let}\ x = b\ \texttt{in}\ c)$;
- si cette séquence termine sur une valeur v'_a, d'une séquence de réductions de $b[y \leftarrow v'_a]$ dans le contexte $\texttt{let}\ x = [\,]\ \texttt{in}\ c[y \leftarrow v'_a]$;
- si cette séquence termine sur une valeur v'_b, d'une séquence de réductions de $c[y \leftarrow v'_a][x \leftarrow v'_b]$ dans le contexte $[\,]$.

Notons que par déterminisme de la réduction, la valeur d'un terme, si elle existe, est unique, donc $v'_a = v_a$ et $v'_b = v_b$. De plus, par hypothèse, y n'est pas libre dans c, et donc $c[y \leftarrow v_a][x \leftarrow v_b] = c[x \leftarrow v_b]$. Par conséquent :

- Si a diverge, p et p' divergent.
- Si $a \xrightarrow{*} v_a$ et b diverge, p et p' divergent.
- Si $a \xrightarrow{*} v_a$ et $b[x \leftarrow v_a] \xrightarrow{*} v_b$, p et p' se comportent comme $c[x \leftarrow v_b]$.

Ceci implique que p et p' ont le même comportement.

Question 6 La preuve est par récurrence sur la dérivation de $a \to a'$.

Cas $a = (\lambda x.c)\ v \to c[x \leftarrow v] = a'$. Deux possibilités pour la dérivation de $(\lambda x.c)\ v \leadsto b$. Première possibilité : via la règle (1).

$$\frac{c[x \leftarrow v] \leadsto b}{(\lambda x.c)\ v \leadsto b}$$

Nous sommes donc dans le second cas du théorème : on a bien $a' \leadsto b$ et de plus la dérivation de ce résultat est une prémisse de celle de $a \leadsto b$; elle est donc strictement plus petite.

Seconde possibilité : via les règles (4) et (5).

$$\frac{\dfrac{c \leadsto c'}{\lambda x.c \leadsto \lambda x.c'} \qquad v \leadsto v'}{(\lambda x.c)\ v \leadsto (\lambda x.c')\ v'}$$

avec $b = (\lambda x.c')\ v'$. En utilisant le lemme de substitution, on a $c[x \leftarrow v] \leadsto c'[x \leftarrow v']$. Nous sommes donc dans le premier cas du théorème, en prenant $b' = c'[x \leftarrow v']$.

Cas $a = a_1\ a_2 \to a'_1\ a_2 = a'$ avec $a_1 \to a'_1$. Deux possibilités pour la dérivation de $a \leadsto b$. La première, via la règle (1), est impossible car a_1 ne peut pas être une λ-abstraction, cela contredirait $a_1 \to a'_1$. Nous avons donc une dérivation via la règle (5) :

$$\frac{a_1 \leadsto b_1 \qquad a_2 \leadsto b_2}{a_1\ a_2 \leadsto b_1\ b_2}$$

avec $b = b_1\ b_2$. Appliquant l'hypothèse de récurrence à la réduction $a_1 \to a'_1$ et à l'optimisation $a_1 \rightsquigarrow b_1$, nous obtenons deux sous-cas :

- Il existe b'_1 tel que $b_1 \to b'_1$ et $a'_1 \rightsquigarrow b'_1$. Prenant $b' = b'_1\ b_2$, nous avons bien $b_1\ b_2 \to b'_1\ b_2$ et $a'_1\ b_2 \rightsquigarrow b'_1\ b_2$.
- $a'_1 \rightsquigarrow b_1$ avec une dérivation strictement plus petite que celle de $a_1 \rightsquigarrow b_1$. Nous pouvons donc dériver $a'_1\ a_2 \rightsquigarrow b_1\ b_2$ avec une dérivation strictement plus petite que celle de $a_1\ a_2 \rightsquigarrow b_1\ b_2$.

Question 7 Considérons une séquence de réductions issue de a : $a = a_0 \to a_1 \to a_2 \to a_3 \ldots$ En itérant le résultat de la question 6, nous obtenons une suite de termes $b = b_0, b_1, b_2, b_3, \ldots$ tels que, pour tout i, $a_i \rightsquigarrow b_i$ et ou bien $b_i \to b_{i+1}$ ou bien $b_i = b_{i+1}$. Remarquons que $b \xrightarrow{*} b_i$ pour tout i.

Si la réduction de a s'arrête sur un entier $N = a_n$ pour un certain n, alors le terme b_n est tel que $N \rightsquigarrow b_n$. Par examen des règles définissant $\rightsquigarrow$, nous avons nécessairement $b_n = N$. Par conséquent, $b \xrightarrow{*} N$.

Si la réduction issue de a est infinie, deux possibilités. Ou bien la suite infinie (b_i) contient une infinité de réductions $b_i \to b_{i+1}$. Dans ce cas, b diverge. Ou bien (b_i) contient un nombre fini de réductions $b_i \to b_{i+1}$. Dans ce cas, à partir d'un certain rang j, la suite des b_i est stationnaire et n'effectue plus aucune réduction. Mais dans ce cas, la taille des dérivations de $a_i \rightsquigarrow b_i$ formerait une suite strictement décroissante à partir du rang j. C'est bien sûr impossible car c'est une suite d'entiers positifs.

Question 8 On remarque d'abord que la relation $\rightsquigarrow$ est réflexive : $a \rightsquigarrow a$ pour tout terme a. (Récurrence sur a et utilisation des règles (2) à (5).) Donc, si $a = (\lambda x.c)\ d \hookrightarrow b = c[x \leftarrow d]$ par la règle O_1, nous pouvons dériver $a \rightsquigarrow b$ par la règle (1) :

$$\frac{c[x \leftarrow d] \rightsquigarrow c[x \leftarrow d] \qquad d \text{ est une valeur ou une variable}}{(\lambda x.c)\ d \rightsquigarrow c[x \leftarrow d]}$$

On en déduit que $C[a] \rightsquigarrow C[b]$ par récurrence sur le contexte C, en utilisant les règles (4) et (5), et la réflexivité de $\rightsquigarrow$.

Le résultat de la question 7 montre alors que $C[a]$ et $C[b]$ ont le même comportement observable.

Question 9 Pour éviter de se perdre dans un gros calcul, remarquons avant tout que la transformation CPS de $\texttt{let}\ x = a\ \texttt{in}\ b$ est telle que :

$$\begin{array}{rcl}
[\![(\lambda x.b)\ a]\!] & = & \lambda k.\ (\lambda k'.\ k'\ (\lambda x.\ [\![b]\!]))\ (\lambda y.\ [\![a]\!]\ (\lambda z.\ y\ z\ k)) \\
& \hookrightarrow & \lambda k.\ (\lambda y.\ [\![a]\!]\ (\lambda z.\ y\ z\ k))\ (\lambda x.\ [\![b]\!]) \\
& \hookrightarrow & \lambda k.\ [\![a]\!]\ (\lambda z.\ (\lambda x.\ [\![b]\!])\ z\ k) \\
& \hookrightarrow & \lambda k.\ [\![a]\!]\ (\lambda x.\ [\![b]\!]\ k)
\end{array}$$

En d'autres termes, la relation $\hookrightarrow$ effectue l'élimination des redex administratifs mentionnée en cours. On a donc :

$$\begin{array}{rcl}
[\![\texttt{let}\ x = (\texttt{let}\ y = a\ \texttt{in}\ b)\ \texttt{in}\ c]\!] & \stackrel{*}{\hookrightarrow} & \lambda k.\ [\![\texttt{let}\ y = a\ \texttt{in}\ b]\!]\ (\lambda x.\ [\![c]\!]\ k) \\
& \stackrel{*}{\hookrightarrow} & \lambda k.\ (\lambda k'.\ [\![a]\!]\ (\lambda y.\ [\![b]\!]\ k'))\ (\lambda x.\ [\![c]\!]\ k) \\
& \hookrightarrow & \lambda k.\ \ [\![a]\!]\ (\lambda y.\ [\![b]\!]\ (\lambda x.\ [\![c]\!]\ k))
\end{array}$$

Note : la variable y étant non libre dans c, elle est non libre dans $[\![c]\!]$ également, et donc on peut faire "rentrer" $[\![c]\!]$ sous le $\lambda y \ldots$ sans risque de capture de variable. D'autre part, nous avons :

$$\begin{array}{rcl}
[\![\texttt{let}\ y = a\ \texttt{in}\ (\texttt{let}\ x = b\ \texttt{in}\ c)]\!] & \stackrel{*}{\hookrightarrow} & \lambda k.\ [\![a]\!]\ (\lambda y.\ [\![\texttt{let}\ x = b\ \texttt{in}\ c]\!]\ k) \\
& \stackrel{*}{\hookrightarrow} & \lambda k.\ [\![a]\!]\ (\lambda y.\ (\lambda k'.\ [\![b]\!]\ (\lambda x.\ [\![c]\!]\ k'))\ k) \\
& \hookrightarrow & \lambda k.\ [\![a]\!]\ (\lambda y.\ [\![b]\!]\ (\lambda x.\ [\![c]\!]\ k))
\end{array}$$

CQFD.