

Item #201

The New Simplification Algorithm

John H. Williams, Edward L. Wimmers
IBM Research - Almaden Research Center
K53-803

650 Harry Road
San Jose, CA 95120-6099

February 8, 1990

1 Justification for the Simplification Algorithm

An obscure note: $\text{Func}(e)$ is an expression that is guaranteed to be a function that is SEF. Our names are divided into two classes: the function names and the names that can stand for any value. The names that can stand for any value will usually be denoted by x . During the Func processing, Func may assume that a function name is a function but may only assume that one of the other names is a SEF expression.

Thruout this note, it assumed that the expression $\lambda(x).e$ does NOT coerce the e to $\text{Func}(e)$. This language restriction (when adopted) will be taken care of by the parser. This is a technical advantage to the simplifier since it will not have to explicitly do the coercion. For example, not having this restriction allows the simplifier to recurse more easily since not all subexpressions of a function expression need be functions (e.g. $\lambda x. 3$).

Theorem 1.1

Assume that pattern P makes the definitions $\{\text{def } p_1 \equiv e_1 \quad \dots \quad \text{def } p_n \equiv e_n\}$. Then
 $\lambda(P)E \equiv (\lambda(x).E[p_1: x/p_1, \dots, p_n: x/p_n])$ where
 $\{\text{def } p_1 \equiv \gamma:\text{Func}(e_1) \quad \dots \quad \text{def } p_n \equiv \gamma:\text{Func}(e_n)\}$

Proof: This seems likely to be true. $\square$

Note: it is assumed that $U: f: \tau$ produces $\perp_{\text{err}}$ if τ is not a pair.

Recall that $C = GC:2$.

Lemma 1.2

Let $F_1, \dots, F_m, G_1, \dots, G_n$ be arbitrary expressions with all the expressions $F_1, \dots, F_m$ having no occurrences of any one of the function names $g_1, \dots, g_n$. Then

$$\begin{aligned} & E \text{ where } \{\text{def } f_1 \equiv F_1 \dots \text{def } f_m \equiv F_m \text{ def } g_1 \equiv G_1 \dots \text{def } g_n \equiv G_n\} \\ \equiv & (E \text{ where } \{\text{def } g_1 \equiv G_1 \dots \text{def } g_n \equiv G_n\}) \text{ where } \{\text{def } f_1 \equiv F_1 \dots \text{def } f_m \equiv F_m\} \end{aligned}$$

Proof: This follows easily by computing the meanings of the two expressions. $\square$

Lemma 1.3

Let $F_1, \dots, F_m$ be arbitrary expressions none of which have any occurrence of the name x . Then

$$\begin{aligned} & \lambda(x).(E \text{ where } \{\text{def } f_1 \equiv F_1 \dots \text{def } f_m \equiv F_m\}) \\ \equiv & (\lambda(x).E) \text{ where } \{\text{def } f_1 \equiv F_1 \dots \text{def } f_m \equiv F_m\} \end{aligned}$$

Proof: This follows easily by computing the meanings of the two expressions. $\square$

Lemma 1.4

Let $G_1, \dots, G_n$ be arbitrary expressions. Then

$$\begin{aligned} & \lambda(x).(E \text{ where } \{\text{def } g_1 \equiv G_1 \dots \text{def } g_n \equiv G_n\}) \\ \equiv & \left(\lambda(x).E[g_1:x/g_1, \dots, g_n:x/g_n] \right) \quad \text{where} \\ & \{ \\ & \quad \text{def } g_1 \equiv \text{delay}:(\lambda(x).(\text{Func}(G_1)[g_1:x/g_1, \dots, g_n:x/g_n])) \\ & \quad \dots \\ & \quad \text{def } g_n \equiv \text{delay}:(\lambda(x).(\text{Func}(G_n)[g_1:x/g_1, \dots, g_n:x/g_n])) \\ & \} \end{aligned}$$

Furthermore, in the RHS expression, note that $g_i:x$ is an FSEF for any SEF x provided this does not cause `delay` to be optimized away.

Proof: This lemma is too complicated to possibly be false. $\square$

As Alex noted, the `delay` is really needed in the previous lemma. Here is a counterexample if the `delay` is not included.

$\lambda(x).(g \text{ where } \{\text{def } g \equiv g\}) \equiv \text{nt}$
 but $(\lambda(x).(g:x)) \text{ where } \{\text{def } g \equiv \lambda(x).(g:x)\} \equiv \text{nt}.$

Theorem 1.5

Let $F_1, \dots, F_m, G_1, \dots, G_n$ be arbitrary expressions with all the expressions $F_1, \dots, F_m$ having no occurrences of any one of the function names $x, g_1, \dots, g_n$. Then

$$\begin{aligned}
 & \lambda(x). (E \text{ where} \\
 & \quad \{ \\
 & \quad \quad \text{def } f_1 \equiv F_1 \quad \dots \quad \text{def } f_m \equiv F_m \\
 & \quad \quad \text{def } g_1 \equiv G_1 \quad \dots \quad \text{def } g_n \equiv G_n \\
 & \quad \}) \\
 & \equiv (\lambda(x). E[g_1: x/g_1, \dots, g_n: x/g_n]) \quad \text{where} \\
 & \quad \{ \\
 & \quad \quad \text{def } f_1 \equiv F_1 \quad \dots \quad \text{def } f_m \equiv F_m \\
 & \quad \quad \text{def } g_1 \equiv \text{delay}:(\lambda(x). (\text{Func}(G_1)[g_1: x/g_1, \dots, g_n: x/g_n])) \\
 & \quad \quad \dots \\
 & \quad \quad \text{def } g_n \equiv \text{delay}:(\lambda(x). (\text{Func}(G_n)[g_1: x/g_1, \dots, g_n: x/g_n])) \\
 & \quad \}
 \end{aligned}$$

Furthermore, in the RHS expression, note that $g_i: x$ is an PSEF for any SEF x provided this does not cause delay to be optimized away.

Proof: Easily follows from the previous Lemmas. $\square$

The following conventions are used

a is a SEF and that does not depend on x

b is a SEF for any SEF x

f is a FSEF and that does not depend on x

g is a FSEF for any SEF x

X is a pattern giving names to x

y is an arbitrary expression (not even assumed to be SEF)

$\lambda_s(X)y$ produces $\perp_{err}$ if the pattern X produces a non-true answer

Here are some definitions. It is assumed that $\perp_{err}$ is produced if the argument does not have the indicated form.

$$U: g \langle t_1, t_2 \rangle \equiv g: t_1: t_2$$

$$GU: i: g \langle t_1, \dots, t_i \rangle \equiv g \langle t_1, \dots, t_{i-1} \rangle: t_i$$

$$GC: i: g: t_1: \dots: t_i \equiv g \langle t_1, \dots, t_i \rangle \text{ for } i \geq 1$$

$$C \equiv GC: 2$$

$$ra: i: \langle t_1, \dots, t_i \rangle \equiv \langle \langle t_1, \dots, t_{i-1} \rangle, t_i \rangle$$

Definition 1.6

A function f is n -distributive iff

$$f: \langle g_1, \dots, g_n \rangle \circ Safe: g_0 \equiv f: \langle g_1 \circ Safe: g_0, \dots, g_n \circ Safe: g_0 \rangle \text{ for all functions}$$

$$g_1, \dots, g_n, g_0.$$

A function f is distributive iff it is n -distributive for all n .

Note that `cons` is distributive, `cond` is 3-distributive, and f' is distributive for all functions f .

Here are some interesting facts.

Lemma 1.7

1. $U:(C:g) \vdash g$
2. $C:(U:g) \equiv \text{delay}:g$
3. $GU:i:(C:g) \equiv g \circ \text{ra}:i$
4. $GC:i:(C:g) \equiv GC:(i+1):(g \circ \text{ra}:(i+1))$ for $i \geq 2$
5. $U:g \circ \text{ra}:i \equiv GU:i:g$
6. $(\lambda_s(\llbracket X_1, \dots, X_i \rrbracket)y) \circ \text{Safe}:s1 \circ \text{ra}:(i+1) \equiv \lambda_s(\llbracket X_1, \dots, X_i, tt \rrbracket)y$
7. $(\lambda_s(\llbracket X_1, \dots, X_{i-1} \rrbracket, X_i \rrbracket)y) \circ \text{ra}:i \equiv \lambda_s(\llbracket X_1, \dots, X_{i-1}, X_i \rrbracket)y$
8. $U:(GC:i:g) \vdash GC:(i-1):(g \circ \text{Safe}:(a1 \circ \llbracket s1, s2 \rrbracket, t1 \circ t1)))$ for $i \geq 3$
9. if f is n -distributive, then

$$\lambda_s(X)(f:(g_1, \dots, g_n)) \vdash C:(f:(U:(\lambda_s(X)g_1), \dots, U:(\lambda_s(X)g_n)))$$
10. $\lambda_s(X)(f \circ g) \vdash C:(f \circ U:(\lambda_s(X)g))$
11. $\lambda_s(X)(g_1 \circ g_2) \vdash C:(((\lambda_s(X)g_1) \circ \text{Safe}:s1) : ' (U:(\lambda_s(X)g_2)))$
12. $\lambda_s(X)(K:b) \vdash C:(\lambda_s(X)b \circ (\text{Safe}:s1))$

Proof: All cases except case 9 are left to the reader. Note that since g_1 is FSEF, it follows that $\lambda_s(X)g_1 \vdash \text{delay}:(\lambda_s(X)g_1)$. Thus, the following holds for any value t .

$$\begin{aligned}
 & (\lambda_s(X)(f:(g_1, \dots, g_n))):t \\
 & \equiv f:(\lambda_s(X)g_1):t, \dots, (\lambda_s(X)g_n):t \\
 & \vdash f:(\text{delay}:(\lambda_s(X)g_1):t, \dots, \text{delay}:(\lambda_s(X)g_n):t) \\
 & \equiv f:(C:(U:(\lambda_s(X)g_1))):t, \dots, C:(U:(\lambda_s(X)g_n))):t \\
 & \equiv f:(U:(\lambda_s(X)g_1) \circ [\neg t, \text{id}], \dots, U:(\lambda_s(X)g_n) \circ [\neg t, \text{id}]) \\
 & \equiv f:(U:(\lambda_s(X)g_1), \dots, U:(\lambda_s(X)g_n)) \circ [\neg t, \text{id}] \\
 & \equiv C:(f:(U:(\lambda_s(X)g_1), \dots, U:(\lambda_s(X)g_n))):t
 \end{aligned}$$

From this it follows that $\lambda_s(X)(f:(g_1, \dots, g_n)) \equiv C:(f:(U:(\lambda_s(X)g_1), \dots, U:(\lambda_s(X)g_n)))$. $\square$

Here are some more interesting facts, mostly applicable to the general case.

1. $(\lambda_s(\mathbf{I})y) \circ \text{Safe:si} \circ \text{Safe}:(\Delta:(\text{islen: } j)) \equiv \lambda_s([\text{tt}, \dots, \text{tt}, \mathbf{I}, \text{tt}, \dots, \text{tt}])y$
where $\mathbf{I}$ occurs in position i
2. $U:(K:g) \equiv g \circ \text{Safe:s2} \circ \text{Safe}:(\Delta:\text{ispair})$
3. $GU:i:(K:g) \equiv g \circ \text{Safe:si} \circ \text{Safe}:(\Delta:(\text{islen: } i))$
4. $U:(g_1:g_2) \equiv (g_1 \circ \text{Safe:s1} \circ \text{Safe}:(\Delta:\text{ispair}))$
 $\quad :' (g_2 \circ \text{Safe:s1} \circ \text{Safe}:(\Delta:\text{ispair})) :' \text{Safe:s2}$
5. $GU:i:(g_1:g_2) \equiv (g_1 \circ \text{Safe:s1} \circ \text{ra:i}) :' (g_2 \circ \text{Safe:s1} \circ \text{ra:i}) :' \text{Safe:si}$

Here are some preliminary results that help in the Abstractor.

1. $\lambda'_s(\mathbf{I}_1) \dots \lambda'_s(\mathbf{I}_{n-1}) \lambda_s(\mathbf{I}_n)y \vdash GC:n:(\lambda_s([\mathbf{I}_1, \dots, \mathbf{I}_n])y) \quad \text{for } n \geq 2$
2. $\lambda'_s(\mathbf{I}_1) \dots \lambda'_s(\mathbf{I}_{n-1})(\text{delay}:(\lambda_s(\mathbf{I}_n)y))$
 $\quad \vdash GC:(n+1):(GU:(n+1):(\lambda_s([\mathbf{I}_1, \dots, \mathbf{I}_n])y)) \quad \text{for } n \geq 2$
3. $\text{delay}:(\lambda_s(\mathbf{I})y) \equiv C:(U:(\lambda_s(\mathbf{I})y))$
4. $\lambda_s(\mathbf{I})(y_1:y_2) \equiv (\lambda_s(\mathbf{I})y_1):'(\lambda_s(\mathbf{I})y_2)$
5. $\lambda_s(\mathbf{I})(\langle y_1, \dots, y_n \rangle) \equiv [\lambda_s(\mathbf{I})y_1, \dots, \lambda_s(\mathbf{I})y_n]$
6. $\lambda_s(\mathbf{I})(a) \vdash K:a$
7. $\lambda_s([\mathbf{I}_1, \dots, \mathbf{I}_n])(x) \vdash (\lambda_s(\mathbf{I}_1)(x)) \circ \text{Safe:si} \quad \text{if } \mathbf{I}_1 \text{ gives a definition to } x$
8. $\lambda_s(x).x \equiv \text{id}$

Using Lemma 1.7, the following facts can be established for all $i \geq 2$.

if $f = \text{cond}$ and $n = 3$ or $f = \text{cons}$ or f has the form $\text{lift}:f_0$, then

$$\begin{aligned} \lambda_n(X)(f:\langle g_1, \dots, g_n \rangle) &\vdash C:(f:\langle U:(\lambda_n(X)g_1), \dots, U:(\lambda_n(X)g_n) \rangle) \\ GC:i:(\lambda_n(X)(f:\langle g_1, \dots, g_n \rangle)) &\vdash \\ &\quad GC:(i+1):(f:\langle GU:(i+1):(\lambda_n(X)g_1), \dots, GU:(i+1):(\lambda_n(X)g_n) \rangle) \\ U:(\lambda_n(X)(f:\langle g_1, \dots, g_n \rangle)) &\vdash f:\langle U:(\lambda_n(X)g_1), \dots, U:(\lambda_n(X)g_n) \rangle \\ GU:i:(\lambda_n(X)(f:\langle g_1, \dots, g_n \rangle)) &\vdash f:\langle GU:i:(\lambda_n(X)g_1), \dots, GU:i:(\lambda_n(X)g_n) \rangle \end{aligned}$$

The special case for composition is as follows:

$$\begin{aligned} \lambda_n(X)(f \circ g) &\vdash C:(f \circ U:(\lambda_n(X)g)) \\ GC:i:(\lambda_n(X)(f \circ g)) &\vdash GC:(i+1):(f \circ GU:(i+1):(\lambda_n(X)g)) \\ U:(\lambda_n(X)(f \circ g)) &\vdash f \circ U:(\lambda_n(X)g) \\ GU:i:(\lambda_n(X)(f \circ g)) &\vdash f \circ GU:i:(\lambda_n(X)g) \end{aligned}$$

The general case for composition is as follows:

$$\begin{aligned} \lambda_n(X)(g_1 \circ g_2) &\vdash C:(\lambda_n([X, \tau\tau])g_1 :' (U:(\lambda_n(X)g_2))) \\ GC:i:(\lambda_n([X_1, \dots, X_i])(g_1 \circ g_2)) &\vdash GC:(i+1):(\lambda_n([X_1, \dots, X_i, \tau\tau])g_1 :' (GU:(i+1):(\lambda_n(X)g_2))) \\ U:(\lambda_n(X)(g_1 \circ g_2)) &\vdash (\lambda_n([X, \tau\tau])g_1 :' (U:(\lambda_n(X)g_2))) \\ GU:i:(\lambda_n([X_1, \dots, X_{i-1}])g_1 \circ g_2) &\vdash (\lambda_n([X_1, \dots, X_{i-1}, \tau\tau])g_1 :' (GU:i:(\lambda_n([X_1, \dots, X_{i-1}])g_2))) \end{aligned}$$

Here is the general case for K

$$\begin{aligned} \lambda_n(X)(K:b) &\vdash C:(\lambda_n([X, \tau\tau])b) \\ GC:i:(\lambda_n([X_1, \dots, X_i])(K:b)) &\vdash GC:(i+1):(\lambda_n([X_1, \dots, X_i, \tau\tau])b) \\ U:(\lambda_n(X)(K:b)) &\vdash \lambda_n([X, \tau\tau])b \\ GU:i:(\lambda_n([X_1, \dots, X_{i-1}])(K:b)) &\vdash \lambda_n([X_1, \dots, X_{i-1}, \tau\tau])b \end{aligned}$$

Here is how the general case for handling the Right-hand-sides. In these rules, $\lambda'_s(\mathbf{x})y$ stands for either $\lambda_s(\mathbf{x})y$ or $\text{delay}:(\lambda_s(\mathbf{x})y)$.

$$\lambda'_s(\mathbf{x}_1) \dots \lambda'_s(\mathbf{x}_{n-1}) \lambda_s(\mathbf{x}_n) y \vdash \text{GC:n}:(\lambda_s([\mathbf{x}_1, \dots, \mathbf{x}_n])y) \quad \text{for } n \geq 2$$

$$\begin{aligned} \lambda'_s(\mathbf{x}_1) \dots \lambda'_s(\mathbf{x}_{n-1}) (\text{delay}:(\lambda_s(\mathbf{x}_n)y)) \\ \vdash \text{GC:(n+1)}:(\text{GU:(n+1)}:(\lambda_s([\mathbf{x}_1, \dots, \mathbf{x}_n])y)) \quad \text{for } n \geq 2 \end{aligned}$$

$$\text{delay}:(\lambda_s(\mathbf{x})y) \equiv \text{C:U}:(\lambda_s(\mathbf{x})y)$$

Here is general case for the Abstractor for $i \geq 2$.

Lambda

$$\lambda_a(X_1)\lambda_a(X_2)y \equiv C:(\lambda_a([X_1, X_2])y)$$

$$GC:i:(\lambda_a([X_1, \dots, X_i])\lambda_a(X_0)y) \equiv GC:(i+1):(\lambda_a([X_1, \dots, X_i, X_0])y)$$

$$U:(\lambda_a(X_1)\lambda_a(X_2)y) \equiv \lambda_a([X_1, X_2])y$$

$$GU:i:(\lambda_a([X_1, \dots, X_{i-1}])\lambda_a(X_0)y) \equiv \lambda_a([X_1, \dots, X_{i-1}, X_0])y$$

Application (no special optimization for GC)

$$\lambda_a(X)(y_1:y_2) \equiv (\lambda_a(X)y_1):'(\lambda_a(X)y_2)$$

$$U:(\lambda_a(X)(y_1:y_2)) \equiv (\lambda_a([X, tt])y_1) :' (\lambda_a([X, tt])y_2) :' Safe:s2$$

$$GU:i:(\lambda_a([X_1, \dots, X_{i-1}])y_1:y_2) \equiv (\lambda_a([X_1, \dots, X_{i-1}, tt])y_1) :' (\lambda_a([X_1, \dots, X_{i-1}, tt])y_2) :' Safe:si$$

Sequences (No special optimizations)

$$\lambda_a(X)(\langle y_1, \dots, y_n \rangle) \equiv [\lambda_a(X)y_1, \dots, \lambda_a(X)y_n]$$

Atoms – no special optimizations. Free functions have some optimizations (but none for GC).

$$\lambda_a(X)(a) \vdash K:a$$

$$U:(\lambda_a(X)(f)) \vdash f \circ Safe:s2$$

$$GU:i:(\lambda_a(X)(f)) \vdash f \circ Safe:si$$

Functions symbols g that depend on $x_1, \dots, x_n$ with $n \geq 1$ have no special optimization for GC.

$$\lambda_a(X)g \equiv (g \circ \lambda_a(X)x_1) :' \lambda_a(X)x_2 :' \dots :' \lambda_a(X)x_n$$

$$U:(\lambda_a(X)g) \equiv (g \circ \lambda_a([X, tt])x_1) :' (\lambda_a([X, tt])x_2) :' \dots :' (\lambda_a([X, tt])x_n) :' Safe:s2$$

$$GU:i:(\lambda_a([X_1, \dots, X_{i-1}])g) \equiv (g \circ \lambda_a([X_1, \dots, X_{i-1}, tt])x_1) :' (\lambda_a([X_1, \dots, X_{i-1}, tt])x_2) :' \dots :' (\lambda_a([X_1, \dots, X_{i-1}, tt])x_n) :' Safe:si$$

$$\lambda_a([X_1, \dots, X_n])(x) \vdash (\lambda_a(X_i)(x)) \circ Safe:si \quad \text{if } X_i \text{ gives a definition to } x$$

$$\lambda_a(x).x \equiv id$$

Recall that $\text{split}:i:\langle x_1, \dots, x_j \rangle \equiv \langle \langle x_1, \dots, x_i \rangle, \langle x_{i+1}, \dots, x_j \rangle \rangle$ for $i \leq j$.

Here are some rewrite rules to use in the rewriter:

1. $U:\text{id} \vdash \text{apply}$
2. $U:g \vdash (g \circ \text{Safe}:s1) : (\text{Safe}:s2)$
3. $GC:i:(GC:j:g) \equiv GC:(i+j-1):(g \circ \text{Safe}:(\text{al} \circ \text{split}:i))$
4. $\text{Safe}:(s1 \circ s1 \circ \text{split}:j) \vdash \text{Safe}:s1$
5. $\text{Safe}:(s1 \circ s2 \circ \text{split}:j) \vdash \text{Safe}:(s:(i+j))$
6. $C:g:b \equiv g \circ [\text{b}, \text{id}]$
7. $GC:i:g:b \equiv GC:(i-1):(g \circ \text{Safe}:(\text{al} \circ [\text{b}, \text{id}]))$ for $i \geq 3$
8. $C:(U:g) \equiv \text{delay}:g$

Example 1.8

In the course of this example, the effects of Theorem 1.1 are ignored. Given the program

$$\lambda(x_1.p_1)\lambda(x_2.p_2)E$$

the parser produces

$$p_1 \rightarrow \lambda_s(x_1.)(p_2 \rightarrow \lambda_s(x_2.)E; e_2); e_1$$

Simplification proceeds on the true-arm as follows:

$$\begin{aligned} & \lambda_s(x_1.)(p_2 \rightarrow \lambda_s(x_2.)E; e_2) \\ \equiv & C : (U : (\lambda_s(x_1.)p_2) \rightarrow U : (\lambda_s(x_1.)\lambda_s(x_2.)E); U : (\lambda_s(x_1.)e_2)) \\ \equiv & C : (p_2 \circ \text{Safe} : s_2 \rightarrow U : (\lambda_s(x_1.)\lambda_s(x_2.)E); e_2 \circ \text{Safe} : s_2) \\ \equiv & C : (p_2 \circ \text{Safe} : s_2 \rightarrow \lambda_s([x_1., x_2.])E; e_2 \circ \text{Safe} : s_2) \end{aligned}$$

Thus the original program

$$\lambda(x_1.p_1)\lambda(x_2.p_2)E$$

is turned into the program

$$p_1 \rightarrow C : (p_2 \circ \text{Safe} : s_2 \rightarrow \lambda_s([x_1., x_2.])E; e_2 \circ \text{Safe} : s_2); e_1$$

whereas the old simplification algorithm produced

$$p_1 \rightarrow (\tilde{p}_2 \rightarrow' \lambda_s(x_1.)\lambda_s(x_2.)E; \tilde{e}_2); e_1$$

The new version is simpler in two ways. First of all, the new version uses the Curry and selection functions whereas the old version uses the K function and the prime operator. Thus, the new version localizes the higher-order effects to the occurrences of the Curry function in contrast to the old version which tended to put a prime or tilde (and sometimes more than one) on practically everything. The second improvement is perhaps even more important. Lambda appears in the new version as the single lambda $\lambda_s([x_1., x_2.])$ whereas it appears as the double lambda $\lambda_s(x_1.)\lambda_s(x_2.)$ in the old version. This double (and, in general, multiple) occurrence of lambda in the old version meant there was a potential for a large blow-up in the size of the final program. The single (albeit more complex) occurrence of lambda in the new version greatly limits the potential for blow-up.

Notice that the new form rewrites easily. For example, if $p_2 \equiv tt$, it is an easy task (two steps) for the rewriter to simplify

$p_1 \rightarrow C : (p_2 \circ \text{Safe} : s_2 \rightarrow \lambda_s([x_1., x_2.])E; e_2 \circ \text{Safe} : s_2); e_1$ to $p_1 \rightarrow C : (\lambda_s([x_1., x_2.])E); e_1$ using already existing rules. However, the old form does not rewrite so easily. For example, either a new rule to express $\rightarrow'$ as a Curried function or a new rule such as $(\tilde{tt} \rightarrow' g_1; g_2) \equiv g_1$ would have to be added, and this rule is valid only if g_1, g_2 are function-valued functions.