ART Memo 7 (Version 0)
A Semantics for Redex Trails

Olaf Chitil

July 4, 2000

1 Introduction

1.1 Redex Trails (from Memo 2)

A redex trail is a directed graph with an application at every node. An application is either a redex or a value (constant, constructor application, partial application). The redex trail shows the computation history of an application. A redex is the parent of an application, if it created that application. An application has exactly one parent (except for root redexes), but every of its subexpressions also has a parent which may be distinct. The redex trail of an application is the graph consisting of its parent and the parents of its subexpressions and recursively the redex trails of these parents and the subexpressions of these parents. Although the sequence of parents can be seen as the main redex trail of an application, the whole trail is a complex branching structure.

The trace of a program run is the redex trail of its result. This trace contains the redex trails of all intermediate values and also of redexes that were created but have never been evaluated during the computation.

Because an application consists of values and sometimes redexes, the graph structure of the trace is complex. A redex is the source of a single arrow and the target of (possibly several) arrows. A value or a subexpressions of an application is the source of a single arrow but never the target of an arrow. Note that a redex is a subexpression of an application only if this redex was never evaluated. Hence it cannot be the target of an arrow.

1.2 Definition of Redex Trails

The current definition of redex trails in [SR97b, SR97a] is insufficient. Redex trails are introduced informally only at the hand of a few examples. The only formal definition of redex traces is given through the transformation that enables the creation of the trace. However, this definition is inadequate. The fact that the description of the transformation is incomplete and contains errors is only a minor problem.

To obtain the redex trail of a program, the program has to be transformed as described and then be executed to create the trail. This has the following disadvantages:

- It is not easy to determine the redex trail of even a simple program by hand.
- The second stage of obtaining a redex trail is based on the semantics of the *transformed* program. This is rather confusing, because the purpose of a redex tail is to give an abstract view of the computation of the *original* program.
- Although it is the purpose of a redex tail to give an abstract view of a computation, a description of the computation of a program is not part of the definition. It remains unclear which assumptions about the operational semantics are made, both in that a redex trace shall be an abstraction of a computation and that the execution of the transformed program shall yield the desired redex trail.

So the transformation defines a redex trail only in a very indirect manner, as depicted here:

$$\text{computation} \overset{semantics}{\longleftarrow} \text{program 1} \overset{semantics}{\longrightarrow} \text{program 2} \overset{resultsemantics}{\longrightarrow} \text{redex trail}$$

We conclude that the transformation is a good definition of an implementation of redex trails but not a good definition of redex trails themselves.

Hence we give a direct definition of redex trails here. Because a redex trail shall give an abstract view of a computation, we start with a high-level semantics of a functional language and then extend it to describe

redex trails beside the normal semantics. This semantics will not have the disadvantage listed before and we can furthermore use it for the following purposes:

- Inconsistencies between the two definitions of redex trails will improve our understanding of them. The final consistency of the two definitions of redex trails will give us confidence that we have defined what we had in mind.
- To prove the correctness of the transformation that implements redex trails.
- To extend redex trails systematically to language constructs for which they are not yet (properly) defined, for example list comprehensions and IO.
- In [SR97a] it is claimed that the redex trail of a program is complete, that "every reduction occurring in the computation can be observed in the trail". It would be nice to prove this property.

2 A High-Level Operational Call-by-Need Semantics

Although a redex trail does not fully describe a computation, it relates redexes and describes sharing properties of call-by-need evaluation. Hence we have to start with an operational semantics. However, we do not want to tie ourselves to a specific abstract machine. Hence Launchbury's natural semantics for call-by-need [Lau93] is the obvious choice for an operational semantics. However, we modify it in several aspects. Some modifications are necessary to reach our goal of defining redex trails, some are more for convenience.

A natural semantics only describes computations of expressions that have a value, that is, a weak head normal form. However, we are also interested in obtaining a redex trail for expressions that have value $\bot$, that is, whose computation leads to an error, a black hole, or an infinite computation that has to be interrupted. To be able to obtain a computation in this case we use a small step variant of Launchbury's semantics (cf. [WP99]).

Maybe Sestoft's Mark 1 machine [Ses97] would be better to describe interrupted computations???

In our first definition of the semantics we still restrict ourselves to a small language. We discuss further language constructs in subsequent sections. Especially we forget about IO for the moment. In our semantics an expression just evaluates to a value.

To achieve sharing easily, the Launchbury requires that arguments of functions (and later also data constructors) are variables. So a preliminary translation is needed that introduces `let` constructs to remove non-variable arguments. This translation will not effect the redex trail, because the reduction of a `let` construct is not registered in the redex trail, a `let` construct is not considered as a redex. We restrict the language further so that only a definition body in a `let` construct can be an application. Thus we assure that a redex will never be a proper subexpression of an expression in a heap cell, but that every redex will have its own heap cell. We will need this property later when we extend the semantics for redex trails. The syntax is defined in Figure 1. We give a semantics to an expression by defining to which value it evaluates.

To express sharing the semantics operates on graphs instead of expressions. For graphs a machine-oriented representation is chosen. A graph $\Gamma : M$ consists of a heap Γ and an expression M. The expression component represents the top of the graph and the heap contains possibly shared parts. A heap, denoted by Γ or Δ, is a partial mapping from variables to expressions. We write it as a set of tuples $x \mapsto M$. By domain of a heap we mean the domain of the mapping. The operator $+$ combines two heaps under the assumption that their domains are disjunct. We only consider closed graphs, that is, graphs where the free variables of the term and of all terms in the heap are a subset of the domain of the heap. So a closed graph represents a closed term with shared subterms.

In Figure 2 we define a reduction relation on graphs, where $\Gamma : M \Rightarrow \Delta : N$ means that the graph $\Gamma : M$ reduces to the graph $\Delta : N$. A computation is a finite or infinite sequence of graphs $G_1, G_2, \ldots$ with $G_i \Rightarrow G_{i+1}$ for all graphs in the sequence. Note that the reduction relation is actually a function, that is, the second graph is determined by the first. The computation of a expression M is the computation beginning with the graph $\emptyset : M$. M evaluates to a value, if its computation is finite with the last graph of the form $\Delta : V$. We write $\emptyset : M \stackrel{*}{\Rightarrow} \Delta : V$.

Our reduction rules differ in one important point from those given in [WP99]. In this paper no relation similar to our value lookup relation is used. Instead, the rules stay closer to the spirit of Launchbury's natural semantics and include a rule

$$\frac{}{\Gamma + \{x \mapsto V\} : x \Rightarrow \Gamma + [x \mapsto V] : V}$$

to give values upwards. We cannot simply give values upwards, because this would conflict with the construction of redex trails as we will see in the subsequent section.

Note that the APPLICATION REDUCTION rule does not require a context $\mathbb{C}$, because a redex is always either the top level expression or the content of a heap cell (the latter being reduced with the help of the FOLLOW VARIABLE rule). It is never a proper subexpression of the content of a heap cell.

variable	x, y, z			
expression	M, N	$:=$	$\hat{M}$	restricted expression
		$\mid$	$\hat{M}\, x$	application
		$\mid$	$\texttt{let}\ \{x_i = M_i\}_{i=1}^{k}\ \texttt{in}\ N$	let
restricted expression	$\hat{M}, \hat{N}$	$:=$	x	variable
		$\mid$	$\lambda x.\, M$	abstraction
		$\mid$	$\texttt{let}\ \{x_i = M_i\}_{i=1}^{k}\ \texttt{in}\ \hat{N}$	let
value	V	$:=$	$\lambda x.\, M$	abstraction
leftmost outermost context	$\mathbb{C}$	$:=$	$[\cdot]$	hole
		$\mid$	$\mathbb{C}x$	application context

Figure 1: Syntax

$$\frac{M \equiv_\Gamma (\lambda x.\, N)\, y}{\Gamma : M \Rightarrow \Gamma : N[y/x]}\ \text{APPLICATION REDUCTION}$$

$$\frac{y_i \notin \mathrm{dom}(\Gamma)}{\Gamma : \mathbb{C}[\texttt{let}\ \{x_i = M_i\}_{i=1}^{k}\ \texttt{in}\ N] \Rightarrow \Gamma + \{y_i \mapsto M_i[y_1/x_1, \ldots, y_k/x_k]\} : \mathbb{C}[N[y_1/x_1, \ldots, y_k/x_k]]}\ \texttt{LET}\ \text{REDUCTION}$$

$$\frac{\Gamma : \Gamma(x) \Rightarrow \Delta : M}{\Gamma : \mathbb{C}[x] \Rightarrow \Delta[x \mapsto M] : \mathbb{C}[x]}\ \text{FOLLOW VARIABLE}$$

Figure 2: Small step reduction rules

$$\frac{}{M \equiv_\Gamma M} \qquad \frac{M \equiv_\Gamma N}{N \equiv_\Gamma M} \qquad \frac{\Gamma(x) \equiv_\Gamma V}{\mathbb{C}[x] \equiv_\Gamma \mathbb{C}[V]}$$

Figure 3: Value lookup

3 Extension for Redex Trails

We allocate the redexes which form the redex trails in the heap as well. It is not possible to completely separate redex trails from the expression that is under evaluation anyway, because the argument of a redex may be evaluated later and the redex trail shall hold the most evaluated form.

To every heap cell we add a variable which points to the redex that created this heap cell. We also add

such a variable for the top level expression. Figure 4 shows the extended semantics. To increase readability we use s and t for these special variables.

Every redex reduces to an expression with its own history. Hence every redex needs its own heap cell and we restricted the syntax to reach this goal. An alternative would be to add trail variables to subexpressions. This would, however, also require a modification of syntax and make expressions far less readable.

So for a (possible incomplete) computation $\emptyset : \text{Root} \cdot M \stackrel{*}{\Rightarrow} \Delta : t \cdot N$ the result is $\Delta : N$ and the redex trail is $\Delta : t$.

$$\frac{M \equiv_\Gamma (\lambda x.\, N)\, y \quad t \notin \text{dom}(\Gamma)}{\Gamma : s \cdot M \;\Rightarrow\; \Gamma + \{t \mapsto s \cdot M\} : t \cdot N[y/x]} \text{ APPLICATION REDUCTION}$$

$$\frac{y_i \notin \text{dom}(\Gamma)}{\Gamma : t \cdot \mathbb{C}[\texttt{let}\ \ \{x_i = M_i\}_{i=1}^k\ \ \texttt{in}\ \ N] \Rightarrow \Gamma + \{y_i \mapsto t \cdot M_i[y_1/x_1, \ldots, y_k/x_k]\} : t \cdot \mathbb{C}[N[y_1/x_1, \ldots, y_k/x_k]]} \texttt{LET} \text{ REDUCTION}$$

$$\frac{\Gamma : \Gamma(x) \Rightarrow \Delta : s \cdot M}{\Gamma : t \cdot \mathbb{C}[x] \;\Rightarrow \Delta[x \mapsto s \cdot M] : t \cdot \mathbb{C}[x]} \text{ FOLLOW VARIABLE}$$

Figure 4: Small step reduction rules for redex trails

The additional trail variables are reminiscence of the cost centers in [SP97].

Note that the definition of contexts $\mathbb{C}$ has no influence on how redex trails are constructed. It should be possible to make even clearer, that evaluation order has no influence on the redex trail structure, only on how far arguments in redexes are evaluated.

4 An Example

For a readable example we need some simple primitive type. We extend our system by natural numbers with addition as given in Figure 5[1].

$$\begin{array}{rrcll} \text{expression} & M, N & := \ldots & | \quad M + M & \text{addition} \\ \text{restricted expression} & \hat{M}, \hat{N} & := \ldots & | \quad n & (n \in I\!N) \\ \text{value} & V & := \ldots & | \quad n & \text{numbers} \\ \text{leftmost outermost context} & \mathbb{C} & := \ldots & | \quad \mathbb{C} + \mathbb{C} & \end{array}$$

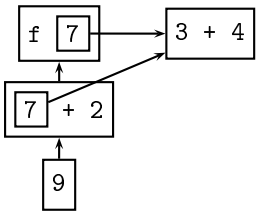
$$\frac{M \equiv_\Gamma n_1 + n_2 \quad t \notin \text{dom}(\Gamma)}{\Gamma : s \cdot M \;\Rightarrow\; \Gamma + \{t \mapsto s \cdot M\} : t \cdot n_1 \underline{+} n_2}$$

Figure 5: Extension for numbers

Now we evaluate the expression $\texttt{let}\ \{f = \lambda x.\, x + 2, a = 3 + 4\}\ \texttt{in}\ f\, a$.

$$\begin{array}{ll} & \emptyset : \text{Root} \cdot \texttt{let}\ \{f = \lambda x.\, x + 2, a = 3 + 4\}\ \texttt{in}\ f\, a \\ \Rightarrow & \{f \mapsto \text{Root} \cdot \lambda x.\, x + 2, a \mapsto \text{Root} \cdot 3 + 4\} : \text{Root} \cdot f\, a \\ \Rightarrow & \{\ldots\} + \{t \mapsto \text{Root} \cdot f\, a\} : t \cdot a + 2 \\ \Rightarrow & \{\ldots, a \mapsto s \cdot 7, \ldots\} + \{s \mapsto \text{Root} \cdot 3 + 4\} : t \cdot a + 2 \\ \Rightarrow & \{\ldots\} + \{r \mapsto t \cdot a + 2\} : r \cdot 9 \\ = & \{f \mapsto \text{Root} \cdot \lambda x.\, x + 2, a \mapsto s \cdot 7, t \mapsto \text{Root} \cdot f\, a, s \mapsto \text{Root} \cdot 3 + 4, r \mapsto t \cdot a + 2\} : r \cdot 9 \end{array}$$

[1] The new context makes reduction non-deterministic; however, this context needed for value lookup.



Note that the parent of $3 + 4$ is Root, not a.

5 Functions with Arity

Until now a redex is an application of a λ-abstraction to one argument. This leads to a large number of redexes and reductions. We want a function value to have an arity. Only the application of a function value of arity n to n arguments is a redex. Figure 6 shows that syntax and semantics can easily be modified to cater for functions with arity. Note that application is still binary, because the arity of function component is not known statically. Note that the definition of values includes abstractions without any arguments.

---

expression	M, N, O, P	$:=$	x	variable
		$\mid$	$\lambda x_1 \ldots x_n . M$	abstraction of arity n
		$\mid$	$\texttt{let}\ \{x_i = \hat{M}_i\}_{i=1}^{k}\ \texttt{in}\ N$	`let`
value	V	$:=$	$(\lambda x_1 \ldots x_n . M)\, y_1 \ldots y_m$ with $m < n$	partial application

$$\frac{M \equiv_\Gamma (\lambda x_1 \ldots x_n . N)\, y_1 \ldots y_n \quad t \notin \mathrm{dom}(\Gamma)}{\Gamma : s \cdot M \Rightarrow \Gamma[t \mapsto s \cdot M] : t \cdot N[y_1/x_1, \ldots, y_n/x_n]}\ \text{APPLICATION REDUCTION}$$

Figure 6: Changes in syntax and semantics for redex trails for functions with arity

---

6 Named Functions

λ-abstractions are very difficult to read. Usually functions have names and we prefer to show these names. Although we do not support[2] the syntax $f x_1 \ldots x_n = M$ we can name a function nonetheless by binding a λ-abstraction to a variable in a `let` construct: $\texttt{let}\ \{f = \lambda x_1 \ldots x_n . M\}\ \texttt{in}\ \ldots$. When we show a redex trail we would like to be able to show this variable instead of the λ-abstraction. We achieve this by adding yet another variable to every heap cell. Not that this variable only serves as a name and may not be in the domain of the heap. We should probably distinguish between variables occurring in the original program and those introduced to fit our syntax (arguments must be variables). The later should not be shown in the redex trail.

7 Pattern Matching

Lauchbury only gives the semantics for a core language. The idea is that standard functional language can be translated into this core language, for example by transforming pattern matching into nested case constructs. For our purposes such a preliminary translation is undesirable, because the close link to the original program would be lost. Although we will obviously not give a semantics for full Haskell, we at least outline how it could be done without requiring a translation into the core language.

[2] It simplifies the semantics that a λ-abstraction is just an expression. For example, we can thus easily store it in the heap.

8 IO

9 Variations

In Launchbury's semantics the term to be reduced is represented by a tuple of a heap and a term, the latter of which represents the top of the term and has variables pointing into the heap. However, we do not use the top term component to give values upwards because we have to use the value lookup relation instead. Hence we wonder if we should not simply move the top term into the heap. So a term would be represented as a heap plus an address variable pointing to its top. Because this address variable would be the same for the input and output of a reduction, a reduction would be triple $\Gamma \stackrel{z}{\Rightarrow} \Delta$.

References

[Lau93] John Launchbury. A natural semantics for lazy evaluation. In *Conference Record of the Twentieth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 144–154. ACM Press, 1993.

[Ses97] Peter Sestoft. Deriving a lazy abstract machine. *Journal of Functional Programming*, 7(3):231–264, May 1997.

[SP97] Patrick M. Sansom and Simon L. Peyton Jones. Formally based profiling for higher-order functional languages. *ACM Transactions on Programming Languages and Systems*, 19(2):334–385, March 1997.

[SR97a] Jan Sparud and Colin Runciman. Complete and partial redex trails of functional computations. In C. Clack, K. Hammond, and T. Davie, editors, *Selected papers from 9th Intl. Workshop on the Implementation of Functional Languages (IFL'97)*, pages 160–177. Springer LNCS Vol. 1467, September 1997.

[SR97b] Jan Sparud and Colin Runciman. Tracing lazy functional computations using redex trails. In H. Glaser, P. Hartel, and H. Kuchen, editors, *Proc. 9th Intl. Symposium on Programming Languages, Implementations, Logics and Programs (PLILP'97)*, pages 291–308. Springer LNCS Vol. 1292, September 1997.

[WP99] Keith Wansbrough and Simon Peyton Jones. Once upon a polymorphic type. In *Conference Record of POPL '99: The* 26$^{\text{nd}}$ *ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, 1999.