

4.4 Application: Inlining

Problem:

- global variables. The program:

```
let x = 1
    f = let x = 2
        in fun y → y + x
in f x
```

... computes something else than:

```
let x = 1
    f = let x = 2
        in fun y → y + x
in let y = x
    in y + x
```

- **recursive functions.** In the definition:

```
foo = fun y → foo y
```

foo should better not be substituted :-)

Idea 1:

- First, we introduce **unique** variable names.
- Then, we only substitute functions which are **staticly** within the scope of the **same** global variables as the application :-)
- For every expression, we determine all function definitions with this property :-)

Let $D = D[e]$ denote the set of definitions which statically arrive at e .

- If $e \equiv \text{let } x_1 = e_1 \text{ and } \dots \text{ and } x_k = e_k \text{ in } e_0$ then:

$$D[e_1] = D$$

...

$$D[e_k] = D \cup \{x_1, \dots, x_{k-1}\}$$

$$D[e_0] = D \cup \{x_1, \dots, x_k\}$$

- In all other cases, D is propagated to the sub-expressions unchanged :-)

E.g., if $e \equiv \text{fun } x \rightarrow e_1$ then:

$$D[e_1] = D$$

... in the Example:

```
let x = 1
    f = let x1 = 2
        in fun y → y + x1
in f x
```

... the application $f\ x$ is not in the scope of x_1

\implies we first duplicate the definition of x_1 :

```
let x = 1
    x1 = 2
    f = let x1 = 2
        in fun y → y + x1
in f x
```

⇒ the inner definition becomes redundant !!!

```
let x = 1
    x1 = 2
    f = fun y → y + x1
in f x
```

⇒ now we can apply inlining :

```
let x = 1
    x1 = 2
    f = fun y → y + x1
in   let y = x
      in y + x1
```

Removing **variable-variable**-assignments, we arrive at:


```
let x = 1
    x1 = 2
    f = fun y → y + x1
in x + x1
```

Idea 2:

- We apply our value analysis.
- We **ignore** global variables **:-)**
- We only substitute functions **without** free variables **:-))**

Example: The **map**-Function

```
let rec f = fun x → x · x  
    map = fun g → fun x → match x  
        with [] → []  
             | :: z → match z with (x1, x2)  
                               in x1 :: map g x2  
in map f list
```

- The **actual** parameter **f** in the application **map f** is always **fun x → x · x :-)**
- Therefore, **map f** can be specialized to a new function **h** defined by:

```

h = let g = fun x → x · x
      in fun x → match x
          with [] → []
              | :: z → match z with (x1, x2)
                  → g x1 :: map g x2

```

The inner occurrence of `map g` can be replaced with `h`

\implies **fold-Transformation** :-)

```
h = let g = fun x → x · x
     in fun x → match x
           with [] → []
                | :: z → match z with (x1, x2)
                               → g x1 :: h x2
```

Inlining the function `g` yields:

```
h = let g = fun x → x · x  
    in fun x → match x  
        with [] → []  
            | :: z → match z with (x1, x2)  
                → ( let x = x1  
                    in x * x ) :: h x2
```

Removing useless definitions and variable-variable assignments yields:

```
h = fun x → match x
      with [] → []
           | :: z → match z with (x1, x2)
                          → x1 * x1 :: h x2
```

4.5 Deforestation

- Functional programmers love to collect intermediate results in lists which are processed by higher-order functions.
- Examples of such higher-order functions are:

```
map = fun f → fun l → match l with [] → []  
    | :: z → (match z with (x, xs) →  
              f x :: map f xs)
```

```
filter = fun p → fun l → match l with [] → []  
| ::z → (match z with (x, xs) →  
           if p x then x :: filter p xs  
           else filter p xs)
```

```
foldl = fun f → fun a → fun l → match l with [] → a  
| ::z → (match z with (x, xs) →  
           foldl f (f a x) xs)
```


id = **fun** $x \rightarrow x$

comp = **fun** $f \rightarrow \mathbf{fun} \ g \rightarrow \mathbf{fun} \ x \rightarrow f (g \ x)$

comp₁ = **fun** $f \rightarrow \mathbf{fun} \ g \rightarrow \mathbf{fun} \ x_1 \rightarrow \mathbf{fun} \ x_2 \rightarrow$
 $f (g \ x_1) \ x_2$

comp₂ = **fun** $f \rightarrow \mathbf{fun} \ g \rightarrow \mathbf{fun} \ x_1 \rightarrow \mathbf{fun} \ x_2 \rightarrow$
 $f \ x_1 (g \ x_2)$

Example:

`sum` = `foldl (+) 0`

`length` = `let f = map (fun x → 1)`
`in comp sum f`

`dev` = `fun l → let s1 = sum l`
`n = length l`
`mean = s1/n`
`l1 = map (fun x → x - mean) l`
`l2 = map (fun x → x · x) l1`
`s2 = sum l2`
`in s2/n`

Observations:

- Explicit recursion does no longer occur!
- The implementation creates unnecessary intermediate data-structures!

`length` could also be implemented as:

```
length = let f = fun a → fun x → a + 1
         in foldl f 0
```

- This implementation avoids to create intermediate lists !!!

Simplification Rules:

$$\begin{aligned} \text{comp id } f &= \text{comp } f \text{ id} = f \\ \text{comp}_1 f \text{ id} &= \text{comp}_2 f \text{ id} = f \\ \text{map id} &= \text{id} \\ \text{comp } (\text{map } f) (\text{map } g) &= \text{map } (\text{comp } f g) \\ \text{comp } (\text{foldl } f a) (\text{map } g) &= \text{foldl } (\text{comp}_2 f g) a \end{aligned}$$

Simplification Rules:

`comp id f` = `comp f id` = `f`
`comp1 f id` = `comp2 f id` = `f`
`map id` = `id`
`comp (map f) (map g)` = `map (comp f g)`
`comp (foldl f a) (map g)` = `foldl (comp2 f g) a`
`comp (filter p1) (filter p2)` = `filter (fun x → if p2 x then p1 x
else false)`
`comp (foldl f a) (filter p)` = `let h = fun a → fun x → if p x then f a x
else a
in foldl h a`

Warning:

Function compositions also could occur as nested function calls ...

```
id x           = x
map id l       = l
map f (map g l) = map (comp f g) l
foldl f a (map g l) = foldl (comp2 f g) a l
filter p1 (filter p2 l) = filter (fun x → p1 x ∧ p2 x) l
foldl f a (filter p l) = let h = fun a → fun x → if p x then f a x
                        else a
                        in foldl h a l
```

Example, optimized:

`sum` = `foldl (+) 0`

`length` = `let f = comp2 (+) (fun x → 1)`
`in foldl f 0`

`dev` = `fun l → let s1 = sum l`
`n = length l`
`mean = s1/n`
`f = comp (fun x → x · x)`
`(fun x → x - mean)`
`g = comp2 (+) f`
`s2 = foldl g 0 l`
`in s2/n`

Remarks:

- All intermediate lists have disappeared :-)
- Only `foldl` remain — i.e., loops :-))
- Compositions of functions can be further simplified in the next step by `Inlining`.
- Inside `dev`, we then obtain:

$$g = \mathbf{fun} \ a \ \rightarrow \ \mathbf{fun} \ x \ \rightarrow \ \mathbf{let} \ \begin{array}{l} x_1 = x - mean \\ x_2 = x_1 \cdot x_1 \end{array} \\ \mathbf{in} \ a + x_2$$

- The result is a sequence of **let**-definitions !!!

Extension: Tabulation

If the list has been created by tabulation of a function, the creation of the list sometimes can be avoided ...

```
tabulate' = fun j → fun f → fun n →  
            if j ≥ n then []  
            else (f j) :: tabulate' (j + 1) f n  
tabulate  = tabulate' 0
```

Then we have:

$$\begin{aligned}\text{comp } (\text{map } f) (\text{tabulate } g) &= \text{tabulate } (\text{comp } f g) \\ \text{comp } (\text{foldl } f a) (\text{tabulate } g) &= \text{loop } (\text{comp}_2 f g) a\end{aligned}$$

where:

$$\begin{aligned}\text{loop}' &= \text{fun } j \rightarrow \text{fun } f \rightarrow \text{fun } a \rightarrow \text{fun } n \rightarrow \\ &\quad \text{if } j \geq n \text{ then } a \\ &\quad \text{else } \text{loop}' (j + 1) f (f a j) n \\ \text{loop} &= \text{loop}' 0\end{aligned}$$

Extension (2): List Reversals

Sometimes, the ordering of lists or arguments is reversed:

$$\begin{aligned} \text{rev}' &= \mathbf{fun} \ a \ \rightarrow \ \mathbf{fun} \ l \ \rightarrow \\ &\quad \mathbf{match} \ l \ \mathbf{with} \ [] \ \rightarrow \ a \\ &\quad | \ ::z \ \rightarrow \ (\mathbf{match} \ z \ \mathbf{with} \ (x, xs) \ \rightarrow \\ &\quad \quad \text{rev}' \ (x :: a) \ xs) \end{aligned}$$
$$\text{rev} = \text{rev}' \ []$$
$$\text{comp rev rev} = \text{id}$$
$$\text{swap} = \mathbf{fun} \ f \ \rightarrow \ \mathbf{fun} \ x \ \rightarrow \ \mathbf{fun} \ y \ \rightarrow \ f \ y \ x$$
$$\text{comp swap swap} = \text{id}$$

$$\text{foldr } f \ a \ = \ \text{comp } (\text{foldl } (\text{swap } f) \ a) \ \text{rev}$$

Discussion:

- The standard implementation of `foldr` is not tail-recursive.
- The last equation decomposes a `foldr` into two tail-recursive functions — at the price that an intermediate list is created.
- Therefore, the standard implementation is probably faster :-)
- Sometimes, the operation `rev` can also be optimized away ...

We have:

$$\begin{aligned}\text{comp rev (map } f) &= \text{comp (map } f) \text{ rev} \\ \text{comp rev (filter } p) &= \text{comp (filter } p) \text{ rev} \\ \text{comp rev (tabulate } f) &= \text{rev_tabulate } f\end{aligned}$$

Here, `rev_tabulate` tabulates in reverse ordering. This function has properties quite analogous to `tabulate`:

$$\begin{aligned}\text{comp (map } f) (\text{rev_tabulate } g) &= \text{rev_tabulate (comp}_2 f g) \\ \text{comp (foldl } f a) (\text{rev_tabulate } g) &= \text{rev_loop (comp}_2 f g) a\end{aligned}$$

Analogously, there is index-dependent accumulation:

```
foldli' = fun i → fun f → fun a → fun l →  
         match l with [] → a  
         | ::z → (match z with (x, xs) →  
                 foldli' (i + 1) f (f i a x) xs)  
foldli  = foldli' 0
```

For composition, we must take care that always the same indices are used. This is achieved by:

$$\text{compi} = \mathbf{fun} f \rightarrow \mathbf{fun} g \rightarrow \mathbf{fun} i \rightarrow \mathbf{fun} x \rightarrow f i (g i x)$$

$$\text{compi}_1 = \mathbf{fun} f \rightarrow \mathbf{fun} g \rightarrow \mathbf{fun} i \rightarrow \mathbf{fun} x_1 \rightarrow \mathbf{fun} x_2 \rightarrow \\ f i (g i x_1) x_2$$

$$\text{compi}_2 = \mathbf{fun} f \rightarrow \mathbf{fun} g \rightarrow \mathbf{fun} i \rightarrow \mathbf{fun} x_1 \rightarrow \mathbf{fun} x_2 \rightarrow \\ f i x_1 (g i x_2)$$

$$\text{cmp}_1 = \mathbf{fun} f \rightarrow \mathbf{fun} g \rightarrow \mathbf{fun} i \rightarrow \mathbf{fun} x_1 \rightarrow \mathbf{fun} x_2 \rightarrow \\ f i x_1 (g x_2)$$

$$\text{cmp}_2 = \mathbf{fun} f \rightarrow \mathbf{fun} g \rightarrow \mathbf{fun} i \rightarrow \mathbf{fun} x_1 \rightarrow \mathbf{fun} x_2 \rightarrow \\ f x_1 (g i x_2)$$

Then:

$$\begin{aligned} \text{comp } (\text{mapi } f) (\text{map } g) &= \text{mapi } (\text{comp}_2 f g) \\ \text{comp } (\text{map } f) (\text{mapi } g) &= \text{mapi } (\text{comp } f g) \\ \text{comp } (\text{mapi } f) (\text{mapi } g) &= \text{mapi } (\text{compi } f g) \\ \text{comp } (\text{foldli } f a) (\text{map } g) &= \text{foldli } (\text{cmp}_1 f g) a \\ \text{comp } (\text{foldl } f a) (\text{mapi } g) &= \text{foldli } (\text{cmp}_2 f g) a \\ \text{comp } (\text{foldli } f a) (\text{mapi } g) &= \text{foldli } (\text{compi}_2 f g) a \\ \text{comp } (\text{foldli } f a) (\text{tabulate } g) &= \text{let } h = \text{fun } a \rightarrow \text{fun } i \rightarrow \\ &\quad f i a (g i) \\ &\quad \text{in loop } h a \end{aligned}$$

Discussion:

- Warning: index-dependent transformations may not commute with `rev` or `filter`.
- All our rules can only be applied if the functions `id`, `map`, `mapi`, `foldl`, `foldli`, `filter`, `rev`, `tabulate`, `rev_tabulate`, `loop`, `rev_loop`, ... are provided by a **standard library**: Only then the algebraic properties can be guaranteed !!!
- Similar simplification rules can be derived for any kind of tree-like data-structure `tree α` .
- These also provide operations `map`, `mapi` and `foldl`, `foldli` with corresponding rules.
- Further opportunities are opened up by functions `to_list` and `from_list` ...