Functional Programming: Functional Programming
6. Folds, and Fold-Fusion

Shin-Cheng Mu

Spring 2022

1 Folds On Lists

A Common Pattern We've Seen Many Times…

$$
\begin{aligned}
&\mathit{sum}\ [\,] &&= 0\\
&\mathit{sum}\ (x : xs) &&= x + \mathit{sum}\ xs
\end{aligned}
$$

$$
\begin{aligned}
&\mathit{length}\ [\,] &&= 0\\
&\mathit{length}\ (x : xs) &&= 1 + \mathit{length}\ xs
\end{aligned}
$$

$$
\begin{aligned}
&\mathit{map}\ f\ [\,] &&= [\,]\\
&\mathit{map}\ f\ (x : xs) &&= f\ x : \mathit{map}\ f\ xs
\end{aligned}
$$

This pattern is extracted and called foldr:

$$
\begin{aligned}
&\mathit{foldr}\ f\ e\ [\,] &&= e,\\
&\mathit{foldr}\ f\ e\ (x : xs) &&= f\ x\ (\mathit{foldr}\ f\ e\ xs).
\end{aligned}
$$

For easy reference, we sometimes call e the "base value" and f the "step function."

1.1 The Ubiquitous foldr

Replacing Constructors

$$
\begin{aligned}
&\mathit{foldr}\ f\ e\ [\,] &&= e\\
&\mathit{foldr}\ f\ e\ (x : xs) &&= f\ x\ (\mathit{foldr}\ f\ e\ xs)
\end{aligned}
$$

- One way to look at $\mathit{foldr}\ (\oplus)\ e$ is that it replaces $[\,]$ with e and $(:)$ with $(\oplus)$:

$$
\begin{aligned}
&\ \ \mathit{foldr}\ (\oplus)\ e\ [1, 2, 3, 4]\\
= &\ \ \mathit{foldr}\ (\oplus)\ e\ (1 : (2 : (3 : (4 : [\,]))))\\
= &\ \ 1 \oplus (2 \oplus (3 \oplus (4 \oplus e))).
\end{aligned}
$$

- $\mathit{sum} = \mathit{foldr}\ (+)\ 0$.
- $\mathit{length} = \mathit{foldr}\ (\lambda x\ n.1 + n)\ 0$.
- $\mathit{map}\ f = \mathit{foldr}\ (\lambda x\ xs.f\ x : xs)\ [\,]$.
- One can see that $\mathit{id} = \mathit{foldr}\ (:)\ [\,]$.

Some Trivial Folds on Lists

- Function max returns the maximum element in a list:

$$
\begin{aligned}
&\mathit{max}\ [\,] &&= \text{-}\infty,\\
&\mathit{max}\ (x : xs) &&= x \uparrow \mathit{max}\ xs.
\end{aligned}
$$

  $\mathit{max} = \mathit{foldr}\ (\uparrow)\ \text{-}\infty$.

- This function is actually called $\mathit{maximum}$ in the standard Haskell Prelude, while max returns the maximum between its two arguments. For brevity, we denote the former by max and the latter by $(\uparrow)$.
- Function prod returns the product of a list:

$$
\begin{aligned}
&\mathit{prod}\ [\,] &&= 1,\\
&\mathit{prod}\ (x : xs) &&= x \times \mathit{prod}\ xs.
\end{aligned}
$$

  $\mathit{prod} = \mathit{foldr}\ (\times)\ 1$.

- Function and returns the conjunction of a list:

$$
\begin{aligned}
&\mathit{and}\ [\,] &&= \mathit{true},\\
&\mathit{and}\ (x : xs) &&= x \wedge \mathit{and}\ xs.
\end{aligned}
$$

  $\mathit{and} = \mathit{foldr}\ (\wedge)\ \mathit{true}$.

- Lets emphasise again that id on lists is a fold:

$$
\begin{aligned}
&\mathit{id}\ [\,] &&= [\,],\\
&\mathit{id}\ (x : xs) &&= x : \mathit{id}\ xs.
\end{aligned}
$$

  $\mathit{id} = \mathit{foldr}\ (:)\ [\,]$.

Some Functions We Have Seen…

- $({+\!\!+}\ ys) = \mathit{foldr}\ (:)\ ys$.

$$
\begin{aligned}
&({+\!\!+}) &&:: \mathit{List}\ a \to \mathit{List}\ a \to \mathit{List}\ a\\
&[\,] {+\!\!+}\ ys &&= ys\\
&(x : xs) {+\!\!+}\ ys &&= x : (xs {+\!\!+}\ ys)\ \ .
\end{aligned}
$$

- $\mathit{concat} = \mathit{foldr}\ ({+\!\!+})\ [\,]$.

$$
\begin{aligned}
&\mathit{concat} &&:: \mathit{List}\ (\mathit{List}\ a) \to \mathit{List}\ a\\
&\mathit{concat}\ [\,] &&= [\,]\\
&\mathit{concat}\ (xs : xss) &&= xs {+\!\!+}\ \mathit{concat}\ xss\ \ .
\end{aligned}
$$

Replacing Constructors

- Understanding *foldr* from its type. Recall

$$\mathbf{data}\ \mathit{List}\ a\ =\ [\,]\ \mid\ a : \mathit{List}\ a\ \ .$$

- Types of the two constructors: $[\,] :: \mathit{List}\ a$, and $(:) :: a \to \mathit{List}\ a \to \mathit{List}\ a$.

- *foldr* replaces the constructors:

$$\begin{array}{lcl} \mathit{foldr} & :: & (a \to b \to b) \to b \to \mathit{List}\ a \to b \\ \mathit{foldr}\ f\ e\ [\,] & = & e \\ \mathit{foldr}\ f\ e\ (x : xs) & = & f\ x\ (\mathit{foldr}\ f\ e\ xs)\ \ . \end{array}$$

Functions on Lists That Are Not *foldr*

- A function f is a *foldr* if in $f\ (x : xs) = ...f\ xs..$, the argument xs does not appear outside of the recursive call.

- Not all functions taking a list as input is a *foldr*.

- The canonical example is perhaps $\mathit{tail} :: \mathit{List}\ a \to \mathit{List}\ a$.

  - $\mathit{tail}(x : xs) = ...\mathit{tail}\ xs..??$

  - *tail* dropped too much information, which cannot be recovered.

- Another example is $\mathit{dropWhile}\ p :: \mathit{List}\ a \to \mathit{List}\ a$.

Longest Prefix

- The function call $\mathit{takeWhile}\ p\ xs$ returns the longest prefix of xs that satisfies p:

$$\begin{array}{lcl} \mathit{takeWhile}\ p\ [\,] & = & [\,] \\ \mathit{takeWhile}\ p\ (x : xs) & = & \\ \quad \mathbf{if}\ p\ x\ \mathbf{then}\ x : \mathit{takeWhile}\ p\ xs & & \\ \quad \mathbf{else}\ [\,]\ \ . & & \end{array}$$

- E.g. $\mathit{takeWhile}\ (\leq 3)\ [1, 2, 3, 4, 5] = [1, 2, 3]$.

- It can be defined by a fold:

$$\begin{array}{l} \mathit{takeWhile}\ p \\ \quad \mathit{foldr}\ (\lambda\, x\, xs \to \mathbf{if}\ p\ x\ \mathbf{then}\ x : xs\ \mathbf{else}\ [\,])\ [\,]. \end{array}$$

All Prefixes

- The function *inits* returns the list of all prefixes of the input list:

$$\begin{array}{lcl} \mathit{inits}\ [\,] & = & [[\,]], \\ \mathit{inits}\ (x : xs) & = & [\,] : \mathit{map}\ (x\ :)\ (\mathit{inits}\ xs). \end{array}$$

- E.g. $\mathit{inits}\ [1, 2, 3] = [[\,], [1], [1, 2], [1, 2, 3]]$.

- It can be defined by a fold:

$$\mathit{inits}\ =\ \mathit{foldr}\ (\lambda\, x\, xss \to [\,] : \mathit{map}\ (x\ :)\ xss)\ [[\,]].$$

All Suffixes

- The function *tails* returns the list of all suffixes of the input list:

$$\begin{array}{lcl} \mathit{tails}\ [\,] & = & [[\,]], \\ \mathit{tails}\ (x : xs) & = & (x : xs) : \mathit{tails}\ xs. \end{array}$$

  - It appears that *tails* is not a *foldr*!

- Luckily, we have $\mathit{head}\ (\mathit{tails}\ xs) = xs$. Therefore,

$$\begin{array}{lcl} \mathit{tails}\ (x : xs) & = & \mathbf{let}\ yss\ =\ \mathit{tails}\ xs \\ & & \mathbf{in}\ (x : \mathit{head}\ yss) : yss. \end{array}$$

- The function *tails* may thus be defined by a fold:

$$\begin{array}{l} \mathit{tails}\ =\ \mathit{foldr}\ (\lambda x\ yss \to \\ \qquad\qquad (x : \mathit{head}\ yss) : yss)\ [[\,]]. \end{array}$$

1.2 The Fold-Fusion Theorem

Why Folds?

- "What are the three most important factors in a programming language?" Abstraction, abstraction, and abstraction!

- Control abstraction, procedure abstraction, data abstraction,…can programming patterns be abstracted too?

- Program structure becomes an entity we can talk about, reason about, and reuse.

  - We can describe algorithms in terms of fold, unfold, and other recognised patterns.
  - We can prove properties about folds,
  - and apply the proved theorems to all programs that are folds, either for compiler optimisation, or for mathematical reasoning.

- Among the theorems about folds, the most important is probably the *fold-fusion* theorem.

The Fold-Fusion Theorem

The theorem is about when the composition of a function and a fold can be expressed as a fold.

Theorem 1 ($foldr$-Fusion)**.** *Given* $f :: a \to b \to b$*,* $e :: b$*,* $h :: b \to c$*, and* $g :: a \to c \to c$*, we have:*

$$h \cdot foldr\ f\ e\ =\ foldr\ g\ (h\ e)\ \ ,$$

if $h\ (f\ x\ y) = g\ x\ (h\ y)$ *for all* x *and* y*.*

For program derivation, we are usually given h, f, and e, from which we have to construct g.

Tracing an Example

Let us try to get an intuitive understand of the theorem:

$$\begin{array}{cl}
& h\ (foldr\ f\ e\ [a, b, c]) \\
= & \{\ \text{definition of } foldr\ \} \\
& h\ (f\ a\ (f\ b\ (f\ c\ e))) \\
= & \{\ \text{since } h\ (f\ x\ y) = g\ x\ (h\ y)\ \} \\
& g\ a\ (h\ (f\ b\ (f\ c\ e))) \\
= & \{\ \text{since } h\ (f\ x\ y) = g\ x\ (h\ y)\ \} \\
& g\ a\ (g\ b\ (h\ (f\ c\ e))) \\
= & \{\ \text{since } h\ (f\ x\ y) = g\ x\ (h\ y)\ \} \\
& g\ a\ (g\ b\ (g\ c\ (h\ e))) \\
= & \{\ \text{definition of } foldr\ \} \\
& foldr\ g\ (h\ e)\ [a, b, c]\ \ .
\end{array}$$

Sum of Squares, Again

- Consider $sum \cdot map\ square$ again. This time we use the fact that $map\ f = foldr\ (mf\ f)\ [\,]$, where $mf\ f\ x\ xs = f\ x : xs$.
- $sum \cdot map\ square$ is a fold, if we can find a ssq such that $sum\ (mf\ square\ x\ xs) = ssq\ x\ (sum\ xs)$. Let us try:

$$\begin{array}{cl}
& sum\ (mf\ square\ x\ xs) \\
= & \{\ \text{definition of } mf\ \} \\
& sum\ (square\ x : xs) \\
= & \{\ \text{definition of } sum\ \} \\
& square\ x + sum\ xs \\
= & \{\ \text{let } ssq\ x\ y = square\ x + y\ \} \\
& ssq\ x\ (sum\ xs)\ \ .
\end{array}$$

  Therefore, $sum \cdot map\ square = foldr\ ssq\ 0$.

Sum of Squares, without Folds

Recall that this is how we derived the inductive case of $sumsq$ yesterday:

$$\begin{array}{cl}
& sumsq\ (x : xs) \\
= & \{\ \text{definition of } sumsq\ \} \\
& sum\ (map\ square\ (x : xs)) \\
= & \{\ \text{definition of } map\ \} \\
& sum\ (square\ x : map\ square\ xs) \\
= & \{\ \text{definition of } sum\ \} \\
& square\ x + sum\ (map\ square\ xs) \\
= & \{\ \text{definition of } sumsq\ \} \\
& square\ x + sumsq\ xs\ \ .
\end{array}$$

Comparing the two derivations, by using fold-fusion we supply only the “important” part.

More on Folds and Fold-fusion

- Compare the proof with the one yesterday. They are essentially the same proof.
- Fold-fusion theorem abstracts away the common parts in this kind of inductive proofs, so that we need to supply only the “important" parts.

Scan

- The following function $scanr$ computes $foldr$ for every suffix of the given list:

$$\begin{array}{l}
scanr :: (a \to b \to b) \to b \to List\ a \to List\ b \\
scanr\ f\ e = map\ (foldr\ f\ e) \cdot tails\ \ .
\end{array}$$

- E.g. computing the running sum of a list:

$$\begin{array}{cl}
& scanr\ (+)\ 0\ [8, 1, 3] \\
= & map\ sum\ (tails\ [8, 1, 3]) \\
= & map\ sum\ [[8, 1, 3], [1, 3], [3], [\,]] \\
= & [12, 4, 3, 0].
\end{array}$$

- Surely there is a quicker way to compute $scanr$, right?

Scan

- Recall that $tails$ is a $foldr$:

$$\begin{array}{l}
tails\ =\ foldr\ (\lambda x\ yss \to \\
\qquad\qquad (x : head\ yss) : yss)\ [[\,]]\ .
\end{array}$$

- By *foldr*-fusion we get:

$$scanr\ f\ e\ =\ foldr\ (\lambda\, x\, ys \to f\ x\ (head\ ys) : ys)\ [e]\ ,$$

- which is equivalent to this inductive definition:

$$\begin{array}{l} scanr\ f\ e\ [\,] = [e] \\ scanr\ f\ e\ (x : xs) = f\ x\ (head\ ys) : ys\ , \\ \quad where\ ys = scanr\ f\ e\ xs\ . \end{array}$$

Tupling as Fold-fusion

- Tupling can be seen as a kind of fold-fusion. The derivation of $steepsum$, for example, can be seen as fusing:

$$steepsum \cdot id\ =\ steepsum \cdot foldr\ (:)\ [\,].$$

- Recall that $steepsum\ xs = (steep\ xs, sum\ xs)$. Reformulating $steepsum$ into a fold allows us to compute it in one traversal.

Accumulating Parameter as Fold-Fusion

- We also note that introducing an accumulating parameter can often be seen as fusing a higher-order function with a $foldr$.
- Recall the function $reverse$. Observe that

$$reverse = foldr\ (\lambda\, x\, xs \to xs +\!\!+ [x])\ [\,]\ \ .$$

- Recall $revcat\ xs\ ys = reverse\ xs +\!\!+ ys$. It is equivalent to

$$revcat = (+\!\!+) \cdot reverse\ \ .$$

- Deriving $revcat$ is performing a fusion!

2 Folds on Other Algebraic Datatypes

- Folds are a specialised form of induction.
- Inductive datatypes: types on which you can perform induction.
- Every inductive datatype give rise to its fold.
- In fact, an inductive type can be defined by its fold.

Fold on Natural Numbers

- Recall the definition:

$$\mathbf{data}\ Nat\ =\ 0\ \mid\ \mathbf{1}_+\ Nat\ \ .$$

- Constructors: $0 :: Nat$, $(\mathbf{1}_+) :: Nat \to Nat$.
- What is the fold on Nat?

$$\begin{array}{lcl} foldN & :: & (a \to a) \to a \to Nat \to a \\ foldN\ f\ e\ 0 & = & e \\ foldN\ f\ e\ (\mathbf{1}_+\ n) & = & f\ (foldN\ f\ e\ n)\ \ . \end{array}$$

Examples of $foldN$

- $(+n) = foldN\ (\mathbf{1}_+)\ n$.

$$\begin{array}{lcl} 0 + n & = & n \\ (\mathbf{1}_+\ m) + n & = & \mathbf{1}_+\ (m + n)\ \ . \end{array}$$

- $(\times n) = foldN\ (n+)\ 0$.

$$\begin{array}{lcl} 0 \times n & = & 0 \\ (\mathbf{1}_+\ m) \times n & = & n + (m \times n)\ \ . \end{array}$$

- $even = foldN\ not\ True$.

$$\begin{array}{lcl} even\ 0 & = & True \\ even\ (\mathbf{1}_+\ n) & = & not\ (even\ n)\ \ . \end{array}$$

Fold-Fusion for Natural Numbers

Theorem 2 ($foldN$-Fusion). *Given* $f :: a \to a$, $e :: a$, $h :: a \to b$, *and* $g :: b \to b$, *we have:*

$$h \cdot foldN\ f\ e\ =\ foldN\ g\ (h\ e)\ \ ,$$

if $h\ (f\ x) = g\ (h\ x)$ *for all* x.

Exercise: fuse $even$ into $(+)$?

Folds on Trees

- Recall some datatypes for trees:

$$\begin{array}{lcl} \mathbf{data}\ ITree\ a & = & \mathsf{Null}\ \mid\ \mathsf{Node}\ a\ (ITree\ a)\ (ITree\ a)\ \ , \\ \mathbf{data}\ ETree\ a & = & \mathsf{Tip}\ a\ \mid\ \mathsf{Bin}\ (ETree\ a)\ (ETree\ a)\ \ . \end{array}$$

- The fold for $ITree$, for example, is defined by:

$$\begin{array}{lcl} foldIT & :: & (a \to b \to b \to b) \to b \to ITree\ a \to b \\ foldIT\ f\ e\ \mathsf{Null} & = & e \\ foldIT\ f\ e\ (\mathsf{Node}\ a\ t\ u) & = & f\ a\ (foldIT\ f\ e\ t)\ (foldIT\ f\ e\ u)\ \ . \end{array}$$

- The fold for $ETree$, is given by:

$$\begin{array}{lcl} foldET & :: & (b \to b \to b) \to (a \to b) \to ETree\ a \to b \\ foldET\ f\ k\ (\mathsf{Tip}\ x) & = & k\ x \\ foldET\ f\ k\ (\mathsf{Bin}\ t\ u) & = & f\ (foldET\ f\ k\ t)\ (foldET\ f\ k\ u)\ \ . \end{array}$$

Some Simple Functions on Trees

- To compute the size of an $ITree$:

$$sizeITree \;=\; foldIT\ (\lambda x\ m\ n \rightarrow \mathbf{1}_{+}\ (m+n))\ 0\ \ .$$

- To sum up labels in an $ETree$:

$$sumETree \;=\; foldET\ (+)\ id.$$

- To compute a list of all labels in an $ITree$ and an $ETree$:

$$\begin{aligned} flattenIT\ \ &= foldIT\ (\lambda x\ xs\ ys \rightarrow xs \mathbin{+\!\!+} [x] \mathbin{+\!\!+} ys)\ [\,],\\ flattenET\ \ &= foldET\ (\mathbin{+\!\!+})\ (\lambda x \rightarrow [x]). \end{aligned}$$

- **Exercise**: what are the fusion theorems for $foldIT$ and $foldET$?

3 Finally, Solving Maximum Segment Sum

Specifying Maximum Segment Sum

- Finally we have introduced enough concepts to tackle the maximum segment sum problem!
- A segment can be seen as a prefix of a suffix.
- The function $segs$ computes the list of all the segments.

$$segs \;=\; concat \cdot map\ inits \cdot tails.$$

- Therefore, mss is specified by:

$$mss \;=\; max \cdot map\ sum \cdot segs.$$

The Derivation!

We reason:

$$\begin{aligned} & max \cdot map\ sum \cdot concat \cdot map\ inits \cdot tails\\ =\ & \{\ \text{since } map\ f \cdot concat = concat \cdot map\ (map\ f)\ \}\\ & max \cdot concat \cdot map\ (map\ sum) \cdot map\ inits \cdot tails\\ =\ & \{\ \text{since } max \cdot concat = max \cdot map\ max\ \}\\ & max \cdot map\ max \cdot map\ (map\ sum) \cdot map\ inits \cdot tails\\ =\ & \{\ \text{since } map\ f \cdot map\ g = map\ (f.g)\ \}\\ & max \cdot map\ (max \cdot map\ sum \cdot inits) \cdot tails\ \ . \end{aligned}$$

Recall the definition $scanr\ f\ e = map\ (foldr\ f\ e) \cdot tails$. If we can transform $max \cdot map\ sum \cdot inits$ into a fold, we can turn the algorithm into a $scanr$, which has a faster implementation.

Maximum Prefix Sum

Concentrate on $max \cdot map\ sum \cdot inits$ (let $ini\ x\ xss = [\,] : map\ (x\,:)\ xss$):

$$\begin{aligned} & max \cdot map\ sum \cdot inits\\ =\ & \{\ \text{definition of } init,\ ini\ x\ xss = [\,] : map\ (x\,:)\ xss\ \}\\ & max \cdot map\ sum \cdot foldr\ ini\ [[\,]]\\ =\ & \{\ \text{fold fusion, see below}\ \}\\ & max \cdot foldr\ zplus\ [0]\ \ . \end{aligned}$$

The fold fusion works because:

$$\begin{aligned} & map\ sum\ (ini\ x\ xss)\\ =\ & map\ sum\ ([\,] : map\ (x\,:)\ xss)\\ =\ & 0 : map\ (sum \cdot (x\,:))\ xss\\ =\ & 0 : map\ (x+)\ (map\ sum\ xss)\ \ . \end{aligned}$$

Define $zplus\ x\ yss = 0 : map\ (x+)\ yss$.

Maximum Prefix Sum, 2nd Fold Fusion

Concentrate on $max \cdot map\ sum \cdot inits$:

$$\begin{aligned} & max \cdot map\ sum \cdot inits\\ =\ & \{\ \text{definition of } init,\ ini\ x\ xss = [\,] : map\ (x\,:)\ xss\ \}\\ & max \cdot map\ sum \cdot foldr\ ini\ [[\,]]\\ =\ & \{\ \text{fold fusion, } zplus\ x\ xss = 0 : map\ (x+)\ xss\ \}\\ & max \cdot foldr\ zplus\ [0]\\ =\ & \{\ \text{fold fusion, let } zmax\ x\ y = 0 \uparrow (x+y)\ \}\\ & foldr\ zmax\ 0\ \ . \end{aligned}$$

The fold fusion works because $\uparrow$ distributes into $(+)$:

$$\begin{aligned} & max\ (0 : map\ (x+)\ xs)\\ =\ & 0 \uparrow max\ (map\ (x+)\ xs)\\ =\ & 0 \uparrow (x + max\ xs)\ \ . \end{aligned}$$

Back to Maximum Segment Sum

We reason:

$$\begin{aligned} & max \cdot map\ sum \cdot concat \cdot map\ inits \cdot tails\\ =\ & \{\ \text{since } map\ f \cdot concat = concat \cdot map\ (map\ f)\ \}\\ & max \cdot concat \cdot map\ (map\ sum) \cdot map\ inits \cdot tails\\ =\ & \{\ \text{since } max \cdot concat = max \cdot map\ max\ \}\\ & max \cdot map\ max \cdot map\ (map\ sum) \cdot map\ inits \cdot tails\\ =\ & \{\ \text{since } map\ f \cdot map\ g = map\ (f.g)\ \}\\ & max \cdot map\ (max \cdot map\ sum \cdot inits) \cdot tails\\ =\ & \{\ \text{reasoning in the previous slides}\ \}\\ & max \cdot map\ (foldr\ zmax\ 0) \cdot tails\\ =\ & \{\ \text{introducing } scanr\ \}\\ & max \cdot scanr\ zmax\ 0\ \ . \end{aligned}$$

Maximum Segment Sum in Linear Time!

- We have derived $mss = max \cdot scanr\ zmax\ 0$, where $zmax\ x\ y = 0 \uparrow (x + y)$.
- The algorithm runs in linear time, but takes linear space.
- A tupling transformation eliminates the need for linear space.

  $$mss \ = \ fst \cdot maxhd \cdot scanr\ zmax\ 0$$

  where $maxhd\ xs \ = \ (max\ xs, head\ xs)$. We omit this last step in the lecture.
- The final program is $mss = fst \cdot foldr\ step\ (0, 0)$, where $step\ x\ (m, y) \ = \ ((0 \uparrow (x + y)) \uparrow m, 0 \uparrow (x + y))$.