Chapter 22. Bird Meertens Formalism (BMF) A Quick Tour

Zhenjiang Hu, Wei Zhang

School of Computer Science Peking University Email: huzj@pku.edu.cn

December 6, 2023

Outline

- 1 Running Example: Maximum Segment Sum Problem
- 2 Bird Meertens Formalism

Running Example: Maximum Segment Sum Problem

We will explain the basic concepts of BMF by demonstrating how to develop a correct linear-time program.

Maximum Segment Sum Problem

Given a list of numbers, find the maximum of sums of all *consecutive* sublists.

- $[-1,3,3,-4,-1,4,2,-1] \implies 7$
- \bullet [-1, 3, 1, -4, -1, 4, 2, -1] \implies 6
- \bullet [-1, 3, 1, -4, -1, 1, 2, -1] \implies 4

Review: Functions and Lists
Structured Recursive Computation Patterns
Horner's Rule

Outline

- 1 Running Example: Maximum Segment Sum Problem
- Bird Meertens Formalism
 - Review: Functions and Lists
 - Structured Recursive Computation Patterns
 - Horner's Rule
 - Application

Review: Functions and Lists
Structured Recursive Computation Patterns
Horner's Rule

Introduction

BMF is a calculus of functions for *people* to derive programs from specifications:

- a range of concepts and notations for defining functions;
- a set of algebraic laws for manipulating functions.

Review: Functions and Lists Structured Recursive Computation Pattern: Horner's Rule Application

Question

Consider the following simple identity:

$$(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1$$

= $((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1$

This equation generalizes in the obvious way to n variables a_1, a_2, \ldots, a_n , and we will refer to it as Horner'e rule.

Question

Consider the following simple identity:

$$(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1$$

= $((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1$

This equation generalizes in the obvious way to n variables a_1, a_2, \ldots, a_n , and we will refer to it as Horner'e rule.

ullet How many imes are used in each side?

Question

Consider the following simple identity:

$$(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1$$

= $((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1$

This equation generalizes in the obvious way to n variables a_1, a_2, \ldots, a_n , and we will refer to it as Horner'e rule.

- How many × are used in each side?
- Can we generalize \times to \otimes , + to \oplus ? What are the essential constraints for \otimes and \oplus ?

Question

Consider the following simple identity:

$$(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1$$

= $((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1$

This equation generalizes in the obvious way to n variables a_1, a_2, \ldots, a_n , and we will refer to it as Horner'e rule.

- ullet How many imes are used in each side?
- Can we generalize \times to \otimes , + to \oplus ? What are the essential constraints for \otimes and \oplus ?
- Do you have suitable notation for expressing the Horner's rule concisely?

Review: Functions

• A function f that has source type α and target type β is denoted by

$$f: \alpha \to \beta$$

We shall say that f takes arguments in α and returns results in β .

- Function application is written without brackets; thus f a means f(a). Function application is more binding than any other operation, so f $a \otimes b$ means (f $a) \otimes b$.
- Functions are curried and applications associates to the left, so f a b means (f a) b (sometimes written as f a b.

• Function composition is denoted by a centralized dot (·). We have

$$(f \cdot g) x = f(g x)$$

• Two functions f and g are equivalence iff

$$\forall x. \ f \ x = g \ x$$

• Function composition is denoted by a centralized dot (·). We have

$$(f \cdot g) x = f(g x)$$

• Two functions f and g are equivalence iff

$$\forall x. \ f \ x = g \ x$$

Exercise

Show the following equation states that functional composition is associative.

$$(f \cdot) \cdot (g \cdot) = ((f \cdot g) \cdot)$$

Review: Functions and Lists Structured Recursive Computation Patterns Horner's Rule Application

• Binary operators will be denoted by \oplus , \otimes , \odot , etc. Binary operators can be sectioned. This means that (\oplus) , $(a\oplus)$ and $(\oplus a)$ all denote functions. The definitions are:

$$(\oplus) a b = a \oplus b$$
$$(a\oplus) b = a \oplus b$$
$$(\oplus b) a = a \oplus b$$

Review: Functions and Lists
Structured Recursive Computation Patterns
Horner's Rule
Application

• Binary operators will be denoted by \oplus , \otimes , \odot , etc. Binary operators can be sectioned. This means that (\oplus) , $(a\oplus)$ and $(\oplus a)$ all denote functions. The definitions are:

$$(\oplus) a b = a \oplus b$$
$$(a\oplus) b = a \oplus b$$
$$(\oplus b) a = a \oplus b$$

Exercise

If \oplus has type \oplus : $\alpha \times \beta \to \gamma$, then what are the types for (\oplus) , $(a\oplus)$ and $(\oplus b)$ for all a in α and b in β ?

• The identity element of \oplus : $\alpha \times \alpha \to \alpha$, if it exists, will be denoted by id_{\oplus} . Thus,

$$a \oplus id_{\oplus} = id_{\oplus} \oplus a = a$$

• The constant values function $K: \alpha \to \beta \to \alpha$ is defined by the equation

$$K a b = a$$

• The identity element of \oplus : $\alpha \times \alpha \to \alpha$, if it exists, will be denoted by id_{\oplus} . Thus,

$$a \oplus id_{\oplus} = id_{\oplus} \oplus a = a$$

• The constant values function $K: \alpha \to \beta \to \alpha$ is defined by the equation

$$K a b = a$$

Exercise

What is the identity element of functional composition?

Review: Lists

- Lists are finite sequence of values of the same type. We use the notation $[\alpha]$ to describe the type of lists whose elements have type α .
 - Examples: $[1,2,1]:[\mathit{Int}] \\ [[1],[1,2],[1,2,1]]:[[\mathit{Int}]] \\ []:[\alpha]$

List Constructors

- [] : $[\alpha]$ constructs an empty list.
- [.] : $\alpha \to [\alpha]$ maps elements of α into singleton lists.

$$[.] a = [a]$$

The primitive operator on lists is concatenation (++).

$$[1] ++ [2] ++ [1] = [1, 2, 1]$$

Concatenation is associative:

$$x ++ (y ++ z) = (x ++ y) ++ z$$



Algebraic View of Lists

- $([\alpha], +, [])$ is a monoid.
- ([α], ++,[]) is a free monoid generated by α under the assignment [.] : $\alpha \to [\alpha]$.
- $([\alpha]^+, ++)$ is a semigroup.

Review: Functions and Lists
Structured Recursive Computation Patterns
Horner's Rule
Application

List Functions: Homomorphisms

A function *h* defined in the following form is called homomorphism:

$$\begin{array}{lll} h & = & id_{\oplus} \\ h & [a] & = & f a \\ h & (x++y) & = & h & x \oplus h & y \end{array}$$

It defines a map from the monoid ($[\alpha], +, []$) to the monoid ($\beta, \oplus : \beta \to \beta \to \beta, id_{\oplus} : \beta$).

List Functions: Homomorphisms

A function *h* defined in the following form is called homomorphism:

$$\begin{array}{lll} h & = & id_{\oplus} \\ h & = & f a \\ h & (x ++ y) & = & h & x \oplus h & y \end{array}$$

It defines a map from the monoid ($[\alpha], +, []$) to the monoid ($\beta, \oplus : \beta \to \beta \to \beta, id_{\oplus} : \beta$).

Property: h is uniquely determined by f and \oplus .

An Example: the function returning the length of a list.

$$#[]$$
 = 0
 $#[a]$ = 1
 $#(x++y)$ = $#x+#y$

Note that (Int, +, 0) is a monoid.

Bags and Sets

A bag is a list in which the order of the elements is ignored.
 Bags are constructed by adding the rule that # is commutative (as well as associative):

$$x +\!\!\!+ y = y +\!\!\!\!+ x$$

A set is a bag in which repetitions of elements are ignored.
 Sets are constructed by adding the rule that # is idempotent (as well as commutative and associative):

$$x ++ x = x$$

Мар

The operator * (pronounced map) takes a function on its left and a list on its right. Informally, we have

$$f * [a_1, a_2, \ldots, a_n] = [f a_1, f a_2, \ldots, f a_n]$$

Formally, (f*) (or sometimes simply written as f*) is a homomorphism:

Мар

The operator * (pronounced map) takes a function on its left and a list on its right. Informally, we have

$$f * [a_1, a_2, \ldots, a_n] = [f a_1, f a_2, \ldots, f a_n]$$

Formally, (f*) (or sometimes simply written as f*) is a homomorphism:

$$f*[] = []$$

 $f*[a] = [fa]$
 $f*(x++y) = (f*x)++(f*y)$

Мар

The operator * (pronounced map) takes a function on its left and a list on its right. Informally, we have

$$f * [a_1, a_2, \dots, a_n] = [f a_1, f a_2, \dots, f a_n]$$

Formally, (f*) (or sometimes simply written as f*) is a homomorphism:

$$f*[]$$
 = []
 $f*[a]$ = [f a]
 $f*(x++y)$ = $(f*x)++(f*y)$

Exercise

Prove the following map distributivity.

$$(f \cdot g) * = (f *) \cdot (g *)$$

Reduce

The operator / (pronounced reduce) takes an associative binary operator on its left and a list on its right. Informally, we have

$$\oplus/[a_1,a_2,\ldots,a_n]=a_1\oplus a_2\oplus\cdots\oplus a_n$$

Formally, \oplus / is a homomorphism:

Reduce

The operator / (pronounced reduce) takes an associative binary operator on its left and a list on its right. Informally, we have

$$\oplus/[a_1,a_2,\ldots,a_n]=a_1\oplus a_2\oplus\cdots\oplus a_n$$

Formally, \oplus / is a homomorphism:

$$\oplus/[]$$
 = id_{\oplus} { if id_{\oplus} exists }
 $\oplus/[a]$ = a
 $\oplus/(x++y)$ = $(\oplus/x) \oplus (\oplus/y)$

Examples:

$$\begin{array}{ll} \textit{max} & : & [\textit{Int}] \rightarrow \textit{Int} \\ \textit{max} & = & \uparrow / \\ & & \text{where } a \uparrow b = \text{if } a \leq b \text{ then } b \text{ else } a \\ \\ \textit{head} & : & [\alpha]^+ \rightarrow \alpha \\ \textit{head} & = & \lessdot / \\ & & \text{where } a \lessdot b = a \\ \\ \textit{last} & : & [\alpha]^+ \rightarrow \alpha \\ \textit{last} & = & \gtrdot / \\ & & \text{where } a \gtrdot b = b \\ \end{array}$$

Promotion

f* and \oplus / can be expressed as identities between functions.

Empty Rules

$$f*\cdot K[] = K[]$$

 $\oplus /\cdot K[] = K id_{\oplus}$

One-Point Rules

$$f*\cdot[\cdot] = [\cdot]\cdot f$$

 $\oplus/\cdot[\cdot] = id$

Join Rules

$$f* \cdot ++/ = ++/\cdot (f*)*$$

 $\oplus/\cdot ++/ = \oplus/.(\oplus/)*$

Review: Functions and Lists Structured Recursive Computation Patterns Horner's Rule Application

Exercise

Any homomorphism h can be defined in the following form:

$$h = \oplus / \cdot f*$$

for some functions \oplus and f.

An Example of Calculation

Composition of two specific homomorphisms is a homomorphism.

Directed Reductions

We introduce two more computation patterns \rightarrow (pronounced left-to-right reduce) and \leftarrow (right-to-left reduce) which are closely related to /. Informally, we have

$$\bigoplus_{e} [a_1, a_2, \dots, a_n] = ((e \oplus a_1) \oplus \dots) \oplus a_n
\oplus_{e} [a_1, a_2, \dots, a_n] = a_1 \oplus (a_2 \oplus \dots \oplus (a_n \oplus e))$$

Formally, we can define $\oplus \not\rightarrow_e$ on lists by two equations.

$$\begin{array}{lll}
\oplus \not \to_e[] & = & e \\
\oplus \not \to_e(x ++ [a]) & = & (\oplus \not \to_e x) \oplus a
\end{array}$$

Directed Reductions

We introduce two more computation patterns \rightarrow (pronounced left-to-right reduce) and \leftarrow (right-to-left reduce) which are closely related to /. Informally, we have

$$\bigoplus_{e} [a_1, a_2, \dots, a_n] = ((e \oplus a_1) \oplus \dots) \oplus a_n
\oplus_{e} [a_1, a_2, \dots, a_n] = a_1 \oplus (a_2 \oplus \dots \oplus (a_n \oplus e))$$

Formally, we can define $\oplus \not\rightarrow_e$ on lists by two equations.

$$\begin{array}{lll}
\oplus \not \to_e[] & = & e \\
\oplus \not \to_e(x ++ [a]) & = & (\oplus \not \to_e x) \oplus a
\end{array}$$

Exercise: Give a formal definition for $\oplus \not\leftarrow_e$.

Directed Reductions without Seeds

$$\begin{array}{rcl}
\oplus \not \rightarrow [a_1, a_2, \dots, a_n] &= & ((a_1 \oplus a_2) \oplus \dots) \oplus a_n \\
\oplus \not \leftarrow [a_1, a_2, \dots, a_n] &= & a_1 \oplus (a_2 \oplus \dots \oplus (a_{n-1} \oplus a_n))
\end{array}$$

Directed Reductions without Seeds

$$\begin{array}{rcl}
\oplus \not \rightarrow [a_1, a_2, \dots, a_n] &= & ((a_1 \oplus a_2) \oplus \dots) \oplus a_n \\
\oplus \not \leftarrow [a_1, a_2, \dots, a_n] &= & a_1 \oplus (a_2 \oplus \dots \oplus (a_{n-1} \oplus a_n))
\end{array}$$

Properties:

$$(\oplus \cancel{+}) \cdot ([a] ++) = \oplus \cancel{+}_a$$
$$(\oplus \cancel{+}) \cdot (++ [a]) = \oplus \cancel{+}_a$$

An Example Use of Left-Reduce

Consider the right-hand side of Horner's rule:

$$(((1 \times a_1 + 1) \times a_2 + 1) \times \cdots + 1) \times a_n + 1$$

This expression can be written using a left-reduce:

$$\odot \not\rightarrow_1[a_1, a_2, \ldots, a_n]$$

An Example Use of Left-Reduce

Consider the right-hand side of Horner's rule:

$$(((1\times a_1+1)\times a_2+1)\times \cdots +1)\times a_n+1$$

This expression can be written using a left-reduce:

$$\bigcirc \not \rightarrow_1[a_1, a_2, \dots, a_n]$$

where $a \odot b = (a \times b) + 1$

An Example Use of Left-Reduce

Consider the right-hand side of Horner's rule:

$$(((1 \times a_1 + 1) \times a_2 + 1) \times \cdots + 1) \times a_n + 1$$

This expression can be written using a left-reduce:

$$\bigcirc \not\rightarrow_1[a_1, a_2, \dots, a_n]$$
where $a \odot b = (a \times b) + 1$

Exercise

Give the definition of \ominus such that the following holds.

$$\ominus \not\rightarrow [a_1, a_2, \ldots, a_n] = (((a_1 \times a_2 + a_2) \times a_3 + a_3) \times \cdots + a_{n-1}) \times a_n + a_n$$

The Special Homework Problem

Suppose $f = \bigoplus \not \to_e = \otimes \not \leftarrow_e$.

1 Prove that f is a homomorphism, i.e., there exisits an associate operator \odot s.t.

$$f(x++y) = f xs \odot f ys.$$

② Implement in Haskell an algorithm to derive \odot from \oplus and \otimes .

Accumulations

With each form of directed reduction over lists there corresponds a form of computation called an accumulation. These forms are expressed with the operators # (pronounced left-accumulate) and # (right-accumulate) and are defined informally by

$$\bigoplus_{e}[a_1, a_2, \dots, a_n] = [e, e \oplus a_1, \dots, ((e \oplus a_1) \oplus) \dots \oplus a_n] \\
\oplus \ \#_e[a_1, a_2, \dots, a_n] = [a_1 \oplus (a_2 \oplus \dots \oplus (a_n \oplus e)), \dots, a_n \oplus e, e]$$

Formally, we can define $\oplus \not \Vdash_e$ on lists by two equations by

$$\bigoplus_{e}[]$$
 = $[e]$
 $\bigoplus_{e}([a] ++ x)$ = $[e] ++ (\bigoplus_{e \oplus a} x),$

or

$$\begin{array}{rcl} \oplus \not\!\!\!/_e[] & = & [e] \\ \oplus \not\!\!\!/_e(x +\!\!\!\!+ [a]) & = & (\oplus \not\!\!\!/_e x) +\!\!\!\!+ [b \oplus a] \\ & \text{where } b = \textit{last}(\oplus \not\!\!\!/_e x). \end{array}$$

Efficiency in Accumulate

 $\bigoplus_{e} [a_1, a_2, \ldots, a_n]$: can be evaluated with n-1 calculations of \bigoplus .

Exercise

Consider computation of first n+1 factorial numbers: $[0!,1!,\ldots,n!]$. How many calculations of \times are required for the following two programs?

- $\bullet \times \#_1[1, 2, \ldots, n]$
- 2 $fact * [0, 1, 2, \dots, n]$ where fact n = product [1..n].

Relation between Reduce and Accumulate

$$\oplus \not\rightarrow_e = \textit{last} \cdot \oplus \not \gg_e$$

$$\bigoplus_{e} = \otimes \xrightarrow{h} [e]$$
where $x \otimes a = x + + [last \ x \oplus a]$

Segments

A list y is a segment of x if there exists u and v such that

$$x = u ++ y ++ v$$
.

If u = [], then y is called an initial segment. If v = [], then y is called an final segment.

An Example:

$$segs [1, 2, 3] = [[], [1], [1, 2], [2], [1, 2, 3], [2, 3], [3]]$$

Exercise: How many segments for a list $[a_1, a_2, \ldots, a_n]$?

inits

The function inits returns the list of initial segments of a list, in increasing order of a list.

inits
$$[a_1, a_2, \dots, a_n] = [[], [a_1], [a_1, a_2], \dots, [a_1, a_2, \dots, a_n]]$$

inits

The function inits returns the list of initial segments of a list, in increasing order of a list.

inits
$$[a_1, a_2, \ldots, a_n] = [[], [a_1], [a_1, a_2], \ldots, [a_1, a_2, \ldots, a_n]]$$

$$\textit{inits} = (+\!\!\!/\!\!\!+\!\!\!/\!\!\!\!/_{[]}) \cdot [\cdot] *$$

tails

The function tails returns the list of final segments of a list, in decreasing order of a list.

tails
$$[a_1, a_2, \dots, a_n] = [[a_1, a_2, \dots, a_n], [a_2, \dots, a_n], \dots, [a_n], []]$$

tails

The function tails returns the list of final segments of a list, in decreasing order of a list.

tails
$$[a_1, a_2, \ldots, a_n] = [[a_1, a_2, \ldots, a_n], [a_2, \ldots, a_n], \ldots, [a_n], []]$$

$$\textit{tails} = (\# \#_{\text{Π}}) \cdot [\cdot] *$$

segs

$$segs = ++ / \cdot tails * \cdot inits$$

Exercise: Show the result of segs [1, 2].

Accumulation Lemma

$$(\oplus \cancel{\#}_e) = (\oplus \cancel{\Rightarrow}_e) * \cdot inits$$
$$(\oplus \cancel{\#}) = (\oplus \cancel{\Rightarrow}) * \cdot inits^+$$

The accumulation lemma is used frequently in the derivation of efficient algorithms for problems about segments.

On lists of length n, evaluation of the LHS requires O(n) computations involving \oplus , while the RHS requires $O(n^2)$ computations.

Structured Recursive Computation Par

Horner's Rule Application

The Question: Revisit

Consider the following simple identity:

$$(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1 = ((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1$$

This equation generalizes in the obvious way to n variables a_1, a_2, \ldots, a_2 , and we will refer to it as Horner'e rule.

- Can we generalize \times to \otimes , + to \oplus ? What are the essential constraints for \otimes and \oplus ?
- Do you have suitable notation for expressing the Horner's rule concisely?

Horner's Rule

The following equation

$$\oplus / \cdot \otimes / * \cdot tails = \odot \not\rightarrow_e$$
 where $e = id_{\otimes}$ $a \odot b = (a \otimes b) \oplus e$

holds, provided that \otimes distributes (backwards) over \oplus :

$$(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$$

for all a, b, and c.

Homework BMF 1-1

Prove the correctness of the Horner's rule.

Show that

$$(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$$

is equivalent to

$$(\otimes c) \cdot \oplus / = \oplus / \cdot (\otimes c) *$$
.

holds on all non-empty lists.

Show that

$$f = \oplus / \cdot \otimes / * \cdot tails$$

satisfies the equations

$$f[] = e$$

$$f(x++[a]) = f x \odot a$$

Review: Functions and Lists

Generalizations of Horner's Rule

Generalization 1:

Generalizations of Horner's Rule

Generalization 1:

Generalization 2:

The Maximum Segment Sum (mss) Problem

Compute the maximum of the sums of all segments of a given sequence of numbers, positive, negative, or zero.

$$mss [3, 1, -4, 1, 5, -9, 2] = 6$$

A Direct Solution

$$\mathit{mss} = \uparrow / \cdot + / * \cdot \mathit{segs}$$

A Direct Solution

$$\mathit{mss} = \uparrow / \cdot + / * \cdot \mathit{segs}$$

Exercise

Write a Haskell program for this direct solution.

Calculating a Linear Algorithm using Horner's Rule

```
mss
   { definition of mss }
     \uparrow / \cdot + / * \cdot segs
= { definition of segs }
     \uparrow / \cdot + / * \cdot + + / \cdot  tails * \cdot inits
= { map and reduce promotion }
     \uparrow / \cdot (\uparrow / \cdot + / * \cdot tails) * \cdot inits
= { Horner's rule with a \odot b = (a+b) \uparrow 0 }
     \uparrow / \cdot \odot \rightarrow_0 * \cdot inits
= { accumulation lemma }
     \uparrow / \cdot \odot \#_0
```

A Program in Haskell

Homework BMF 1-2

Code the derived linear algorithm for *mss* in Haskell.

Segment Decomposition

The sequence of calculation steps given in the derivation of the *mss* problem arises frequently. The essential idea can be summarized as a general theorem.

Theorem (Segment Decomposition)

Suppose S and T are defined by

$$S = \oplus / \cdot f * \cdot segs$$

 $T = \oplus / \cdot f * \cdot tails$

If T can be expressed in the form $T = h \cdot \odot \rightarrow_e$, then we have

$$S = \oplus / \cdot h * \cdot \odot / \!\!/ e$$

Homework BMF 1-3

Prove the segment decomposition theorem.