

Chapter 22. Bird Meertens Formalism (BMF)

A Quick Tour

Zhenjiang Hu, Wei Zhang

Department of Computer Science and Technology
Peking University
Email: huzj@pku.edu.cn

Peking, 2022

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.

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

Outline

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

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.

Question

Consider the following simple identity:

$$\begin{aligned}(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1 \\ = ((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1\end{aligned}$$

This equation generalizes in the obvious way to n variables a_1, a_2, \dots, a_n , and we will refer to it as [Horner's rule](#).

Question

Consider the following simple identity:

$$\begin{aligned}(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1 \\ = ((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1\end{aligned}$$

This equation generalizes in the obvious way to n variables a_1, a_2, \dots, a_n , and we will refer to it as [Horner's rule](#).

- How many \times are used in each side?

Question

Consider the following simple identity:

$$\begin{aligned}(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1 \\ = ((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1\end{aligned}$$

This equation generalizes in the obvious way to n variables a_1, a_2, \dots, a_n , and we will refer to it as [Horner's rule](#).

- How many \times 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:

$$\begin{aligned}(a_1 \times a_2 \times a_3) + (a_2 \times a_3) + a_3 + 1 \\ = ((1 \times a_1 + 1) \times a_2 + 1) \times a_3 + 1\end{aligned}$$

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

- How many \times 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

$$f : \alpha \rightarrow \beta$$

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 (\cdot). 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 (\cdot). 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.

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

- 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:

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

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

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

- 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 $\oplus : \alpha \times \beta \rightarrow \gamma$, 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 $\oplus : \alpha \times \alpha \rightarrow \alpha$, if it exists, will be denoted by id_{\oplus} . Thus,

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

- The constant values function $K : \alpha \rightarrow \beta \rightarrow \alpha$ is defined by the equation

$$K a b = a$$

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

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

- The constant values function $K : \alpha \rightarrow \beta \rightarrow \alpha$ 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:
 $[1, 2, 1] : [Int]$
 $[[1], [1, 2], [1, 2, 1]] : [[Int]]$
 $[] : [\alpha]$

List Constructors

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

$$[\cdot] a = [a]$$

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

List Functions: Homomorphisms

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

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

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

List Functions: Homomorphisms

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

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

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

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

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

$$\begin{aligned}\# [] &= 0 \\ \# [a] &= 1 \\ \# (x ++ y) &= \# x + \# y\end{aligned}$$

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):

$$x ++ x = x$$

Map

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:

Map

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:

$$\begin{aligned} f * [] &= [] \\ f * [a] &= [f a] \\ f * (x ++ y) &= (f * x) ++ (f * y) \end{aligned}$$

Map

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:

$$\begin{aligned} f * [] &= [] \\ f * [a] &= [f a] \\ f * (x ++ y) &= (f * x) ++ (f * y) \end{aligned}$$

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

$$\oplus/[a_1, a_2, \dots, a_n] = a_1 \oplus a_2 \oplus \dots \oplus a_n$$

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

$$\oplus/[a_1, a_2, \dots, a_n] = a_1 \oplus a_2 \oplus \dots \oplus a_n$$

Formally, $\oplus/$ is a homomorphism:

$$\begin{aligned}\oplus/[] &= id_{\oplus} && \{ \text{if } id_{\oplus} \text{ exists} \} \\ \oplus/[a] &= a \\ \oplus/(x ++ y) &= (\oplus/x) \oplus (\oplus/y)\end{aligned}$$

Examples:

max : $[Int] \rightarrow Int$
 max = $\uparrow /$
 where $a \uparrow b = \text{if } a \leq b \text{ then } b \text{ else } a$

$head$: $[\alpha]^+ \rightarrow \alpha$
 $head$ = $\ll /$
 where $a \ll b = a$

$last$: $[\alpha]^+ \rightarrow \alpha$
 $last$ = $\gg /$
 where $a \gg b = b$

Promotion

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

Empty Rules

$$\begin{aligned} f^* \cdot K [] &= K [] \\ \oplus/ \cdot K [] &= K id_{\oplus} \end{aligned}$$

One-Point Rules

$$\begin{aligned} f^* \cdot [\cdot] &= [\cdot] \cdot f \\ \oplus/ \cdot [\cdot] &= id \end{aligned}$$

Join Rules

$$\begin{aligned} f^* \cdot ++ / &= ++ / \cdot (f^*)^* \\ \oplus/ \cdot ++ / &= \oplus/ \cdot (\oplus/)^* \end{aligned}$$

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.

$$\begin{aligned}
 & \oplus / \cdot f * \cdot ++ / \cdot g * \\
 = & \quad \{ \text{map promotion} \} \\
 & \oplus / \cdot ++ / \cdot f * * \cdot g * \\
 = & \quad \{ \text{reduce promotion} \} \\
 & \oplus / \cdot (\oplus /) * \cdot f * * \cdot g * \\
 = & \quad \{ \text{map distribution} \} \\
 & \oplus / \cdot (\oplus / \cdot f * \cdot g) *
 \end{aligned}$$

Directed Reductions

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

$$\begin{aligned}\oplus \rightarrow_e [a_1, a_2, \dots, a_n] &= ((e \oplus a_1) \oplus \dots) \oplus a_n \\ \oplus \leftarrow_e [a_1, a_2, \dots, a_n] &= a_1 \oplus (a_2 \oplus \dots \oplus (a_n \oplus e))\end{aligned}$$

Formally, we can define $\oplus \rightarrow_e$ on lists by two equations.

$$\begin{aligned}\oplus \rightarrow_e [] &= e \\ \oplus \rightarrow_e (x ++ [a]) &= (\oplus \rightarrow_e x) \oplus a\end{aligned}$$

Directed Reductions

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

$$\begin{aligned}\oplus \rightarrow_e [a_1, a_2, \dots, a_n] &= ((e \oplus a_1) \oplus \dots) \oplus a_n \\ \oplus \leftarrow_e [a_1, a_2, \dots, a_n] &= a_1 \oplus (a_2 \oplus \dots \oplus (a_n \oplus e))\end{aligned}$$

Formally, we can define $\oplus \rightarrow_e$ on lists by two equations.

$$\begin{aligned}\oplus \rightarrow_e [] &= e \\ \oplus \rightarrow_e (x ++ [a]) &= (\oplus \rightarrow_e x) \oplus a\end{aligned}$$

Exercise: Give a formal definition for $\oplus \leftarrow_e$.

Directed Reductions without Seeds

$$\begin{aligned}\oplus \rightarrow [a_1, a_2, \dots, a_n] &= ((a_1 \oplus a_2) \oplus \dots) \oplus a_n \\ \oplus \leftarrow [a_1, a_2, \dots, a_n] &= a_1 \oplus (a_2 \oplus \dots \oplus (a_{n-1} \oplus a_n))\end{aligned}$$

Directed Reductions without Seeds

$$\begin{aligned} \oplus \nearrow [a_1, a_2, \dots, a_n] &= ((a_1 \oplus a_2) \oplus \dots) \oplus a_n \\ \oplus \nwarrow [a_1, a_2, \dots, a_n] &= a_1 \oplus (a_2 \oplus \dots \oplus (a_{n-1} \oplus a_n)) \end{aligned}$$

Properties:

$$\begin{aligned} (\oplus \nearrow) \cdot ([a] ++) &= \oplus \nearrow a \\ (\oplus \nwarrow) \cdot (++ [a]) &= \oplus \nwarrow a \end{aligned}$$

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:

$$\odot \leftarrow_1 [a_1, a_2, \dots, a_n]$$

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:

$$\odot \leftarrow_1 [a_1, a_2, \dots, a_n]$$

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:

$$\odot \leftarrow_1 [a_1, a_2, \dots, a_n]$$

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

Exercise

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

$$\ominus \leftarrow [a_1, a_2, \dots, a_n] = (((a_1 \times a_2 + a_2) \times a_3 + a_3) \times \cdots + a_{n-1}) \times a_n + a_n$$

The Special Homework Problem

Suppose $f = \oplus \not\leftarrow_e = \otimes \not\leftarrow_e$.

- 1 Prove that f is a homomorphism, i.e., there exists 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 $\mathop{\oplus}\limits_{\rightarrow}$ (pronounced **left-accumulate**) and $\mathop{\oplus}\limits_{\leftarrow}$ (**right-accumulate**) and are defined informally by

$$\begin{aligned}\oplus\limits_{\rightarrow} e [a_1, a_2, \dots, a_n] &= [e, e \oplus a_1, \dots, ((e \oplus a_1) \oplus) \cdots \oplus a_n] \\ \oplus\limits_{\leftarrow} e [a_1, a_2, \dots, a_n] &= [a_1 \oplus (a_2 \oplus \cdots \oplus (a_n \oplus e)), \dots, a_n \oplus e, e]\end{aligned}$$

Formally, we can define $\oplus \# \#_e$ on lists by two equations by

$$\begin{aligned}\oplus \# \#_e [] &= [e] \\ \oplus \# \#_e ([a] ++ x) &= [e] ++ (\oplus \# \#_{e \oplus a} x),\end{aligned}$$

or

$$\begin{aligned}\oplus \# \#_e [] &= [e] \\ \oplus \# \#_e (x ++ [a]) &= (\oplus \# \#_e x) ++ [b \oplus a] \\ &\text{where } b = \text{last}(\oplus \# \#_e x).\end{aligned}$$

Efficiency in Accumulate

$\oplus \#_e [a_1, a_2, \dots, a_n]$: can be evaluated with $n - 1$ calculations of \oplus .

Exercise

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

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

Relation between Reduce and Accumulate

$$\oplus \nearrow e = \textit{last} \cdot \oplus \nearrow e$$

$$\oplus \nearrow e = \otimes \nearrow [e]$$

$$\text{where } x \otimes a = x \textit{ ++ } [\textit{last } x \oplus a]$$

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:

$$\text{segs } [1, 2, 3] = [], [1], [1, 2], [2], [1, 2, 3], [2, 3], [3]$$

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

inits

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

$$\mathit{inits} [a_1, a_2, \dots, a_n] = [[], [a_1], [a_1, a_2], \dots, [a_1, a_2, \dots, a_n]]$$

inits

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

$$\mathit{inits} [a_1, a_2, \dots, a_n] = [[], [a_1], [a_1, a_2], \dots, [a_1, a_2, \dots, a_n]]$$

$$\mathit{inits} = (\# \# \# \#) \cdot [\cdot]^*$$

tails

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

$$\text{tails } [a_1, a_2, \dots, a_n] = [[a_1, a_2, \dots, a_n], [a_2, \dots, a_n], \dots, [a_n], []]$$

tails

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

$$\text{tails } [a_1, a_2, \dots, a_n] = [[a_1, a_2, \dots, a_n], [a_2, \dots, a_n], \dots, [a_n], []]$$

$$\text{tails} = (\text{++} \leftarrow \# \text{ []}) \cdot [\cdot]^*$$

segs

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

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

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, \dots, a_n , and we will refer to it as [Horner's 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

$$\oplus / \cdot \otimes / * \cdot \text{tails} = \odot \dashv e$$

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

$$f = \oplus / \cdot \otimes / * \cdot \text{tails}$$

satisfies the equations

$$\begin{aligned} f [] &= e \\ f (x ++ [a]) &= f x \odot a \end{aligned}$$

Generalizations of Horner's Rule

Generalization 1:

$$\oplus / \cdot \otimes / * \cdot \textit{tails}^+ = \odot \dashrightarrow$$

where

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

Generalizations of Horner's Rule

Generalization 1:

$$\oplus / \cdot \otimes / * \cdot \textit{tails}^+ = \odot \dashv$$

where

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

Generalization 2:

$$\oplus / \cdot (\otimes / \cdot f *) * \cdot \textit{tails} = \odot \dashv_e$$

where

$$e = id_{\otimes}$$

$$a \odot b = (a \otimes f b) \oplus e$$

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.

$$\text{mss } [3, 1, -4, 1, 5, -9, 2] = 6$$

A Direct Solution

$$mss = \uparrow / \cdot + / * \cdot segs$$

A Direct Solution

$$mss = \uparrow / \cdot + / * \cdot segs$$

Exercise

Write a Haskell program for this direct solution.

Calculating a Linear Algorithm using Horner's Rule

$$\begin{aligned}
 & mss \\
 = & \quad \{ \text{definition of } mss \} \\
 & \uparrow / \cdot + / * \cdot \text{segs} \\
 = & \quad \{ \text{definition of segs} \} \\
 & \uparrow / \cdot + / * \cdot ++ / \cdot \text{tails} * \cdot \text{inits} \\
 = & \quad \{ \text{map and reduce promotion} \} \\
 & \uparrow / \cdot (\uparrow / \cdot + / * \cdot \text{tails}) * \cdot \text{inits} \\
 = & \quad \{ \text{Horner's rule with } a \odot b = (a + b) \uparrow 0 \} \\
 & \uparrow / \cdot \odot \not\rightarrow_0 * \cdot \text{inits} \\
 = & \quad \{ \text{accumulation lemma} \} \\
 & \uparrow / \cdot \odot \not\rightarrow_0
 \end{aligned}$$

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

$$S = \oplus / \cdot f * \cdot \text{segs}$$

$$T = \oplus / \cdot f * \cdot \text{tails}$$

If T can be expressed in the form $T = h \cdot \odot \not\rightarrow_e$, then we have

$$S = \oplus / \cdot h * \cdot \odot \not\rightarrow_e$$

Homework BMF 1-3

Prove the segment decomposition theorem.