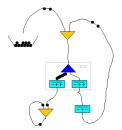
# Productivity

introduction, history pure stream format, pebbleflow nets, decidability

Jörg Endrullis Clemens Grabmayer Dimitri Hendriks

Vrije Universiteit Amsterdam - Universiteit Utrecht

ISR 2010, Utrecht, July 7



### course overview

#### today:

- 1. introduction, history (D)
- 2. the pure stream format, pebbleflow nets, decidability (D)
- 3. extended formats (C)

#### tomorrow:

- data-oblivious productivity (C)
- 5. productivity of infinite data structures via termination (J)
- 6. complexity and variants of productivity (C)
- 7. practicum: defining streams (you)

### outline parts 1 & 2

- 1. introduction
- 2. history
- 3. the pure stream format
- 4. modelling with nets
- 5. deciding productivity
- 6. conclusion

## what is productivity?

a finite expression is productive if it

- represents a unique infinite object, and
- allows for the construction of this infinite object

in a slogan:

productivity = constructive well-definedness

#### in this talk:

- infinite objects: streams (infinite contructor normal forms)
- finite expressions: functional programs / term rewriting systems
- construction: evaluation / term rewriting

# why study productivity?

productivity is a crucial property for correctness of programs dealing with infinite data structures:

- productivity ensures unlimited progress
- a productive program is produces values indefinitely
- a productive program is immune to starvation

#### streams

- ▶ a stream over A is an infinite sequence of elements from A
- we write streams as

$$a_0 : a_1 : a_2 : \dots$$

where ':' is the stream constructor symbol

### example

```
zeros = 0:0:0:...

fib = 0:1:1:2:3:5:8:...

morse = 0:1:1:0:1:0:0:1:...

facts = 1:1:2:6:24:120:...
```

## productivity of stream specifications

#### definition

a stream specification is an orthogonal,  $\{S, D\}$ -sorted constructor TRS

#### definition

a stream specification is productive for a term *t* if outermost-fair evaluation results in an infinite constructor normal form:

```
t - a_0 : a_1 : a_2 : \dots
```

productivity of stream specifications is undecidable in general

```
example (productive? yes!) {\sf zeros} \to 0 : {\sf zeros}
```

The rule for zeros produces 0's indefinitely:

```
zeros \rightarrow 0: zeros \rightarrow 0:0: zeros \rightarrow 0:0:0:...
```

```
example (productive? no!) N \to 0 : tail(N) tail(x : xs) \to xs
```

#### we cannot make any progress:

```
N \rightarrow 0: tail(N)
 \rightarrow 0: tail(0: tail(N))
 \rightarrow 0: tail(N)
 \rightarrow \dots
```

```
example (productive? yes!)
A \rightarrow 0 : read1(A)
read1(x : xs) \rightarrow x : read1(xs)
```

```
A \rightarrow 0 : read1(A)

\rightarrow 0 : read1(0 : read1(A))

\rightarrow 0 : 0 : read1(read1(A))

\rightarrow 0 : 0 : 0 : ...
```

```
example (productive? no!)
B \rightarrow 0 : read2(B)
read2(x : y : xs) \rightarrow x : y : read2(xs)
```

The rule for read2 can never be applied:

```
\begin{split} \textbf{B} &\rightarrow 0 : \text{read2(B)} \\ &\rightarrow 0 : \text{read2(0: read2(B))} \\ &\rightarrow 0 : \text{read2(0: read2(0: read2(...)))} \end{split}
```

### operational versus extensional

```
read1(x:xs) \rightarrow x: read1(xs)read2(x:y:xs) \rightarrow x: y: read2(xs)
```

read1 and read2 are extensionally equal, but intensionally/operationally they are not!

## productivity versus unique solvability

### example (productive? no!)

$$Z = h(Z)$$
$$h(x : xs) = 0 : h(xs)$$

- Z has a unique solution 0:0:0:...
- but this solution cannot be found by evaluating the specification:

$$Z = h(Z) = h(h(Z)) = \dots$$

### example (productive? yes!)

```
morse \rightarrow 0 : 1 : h(tail(morse))
h(x : xs) \rightarrow x : not(x) : h(xs)
tail(x : xs) \rightarrow xs
not(0) \rightarrow 1
not(1) \rightarrow 0
```

# productivity of stream specifications

and hence:

```
J - 0:1:0:0:even^{\omega}
```

J is strongly normalizing, but not productive

▶ question: how many bits do we have to add to make J productive, i.e. for which n is  $J \rightarrow a_0 : a_1 : ... : a_n : even(J)$  productive?

## Productivity recognition: previous approaches

- ► Wadge (1981): 'cyclic sum test' (limited, computable criterion)
- ➤ Sijtsma (1989): mathematical theory of productivity based on 'production moduli' (not directly computable criteria)
- Coquand (1994): 'guardedness' as a syntactic criterion for productivity (automatable, but restrictive criterion)
- Hughes, Pareto, and Sabry (1996): introduce a type system for proving productivity (automatable criterion)
- ► Telford and Turner (1997): extend the notion of guardedness by a method in the flavour of Wadge
- Buchholz (2004): type system for proving productivity (automatable for a restricted subsystem)

all use a data-oblivious analysis, a 'quantitative' analysis where the knowledge about concrete values of data elements is ignored

### Production moduli

#### definition

let  $f:(A^\omega)^r\to A^\omega$  be a stream function. a production modulus for f is a function  $\nu_f:(\overline{\mathbb{N}})^r\to \overline{\mathbb{N}}$  such that the first  $\nu_f(t_1,\ldots,t_r)$  elements of  $f(t_1,\ldots,t_r)$  can be computed whenever the first  $n_i$  elements of  $t_i$  are defined.

### example

$$ail(x:xs) o xs \qquad \qquad 
u_{ ext{tail}}(n) = n \dot{-} 1 \ ext{dup}(x:xs) o x:x: ext{dup}(xs) \qquad \qquad 

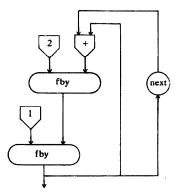
 $u_{ ext{dup}}(n) = 2n \ ext{odd}(x:y:xs) o y: ext{odd}(xs) \qquad \qquad 

 $u_{ ext{odd}}(n) = \lfloor \frac{n}{2} \rfloor \ ext{add}(x:xs,y:ys) o (x+y): ext{add}(xs,ys) \qquad 

 $u_{ ext{add}}(n,m) = \min(n,m)$$$$$

### the cyclic sum test of Wadge

- a network is a device for computing the least fixed point of a system of equations (Kahn 1974)
- ► Wadge 1981 studies deadlock in dataflow networks. (free of deadlock ≈ productive).



- a loop means that some node is consuming its own output
- ▶ a node might starve, when it is waiting for itself to produce data

## the cyclic sum test of Wadge

- associate with each of the arguments of the operations an integer which says how far the output leads (or lags) that argument;
- network passes the test if, for every cycle, the sum of associated numbers is positive
- a network passing the test is guaranteed to be immune to deadlock

### the cyclic sum test of Wadge

- by using numbers it can only be expressed that consumption and production differs by a constant value,
- with constant functions as production moduli one cannot express that production depends on the number of elements consumed
- for example, to the argument of dup

$$dup(x:xs) \rightarrow x:x:dup(xs)$$

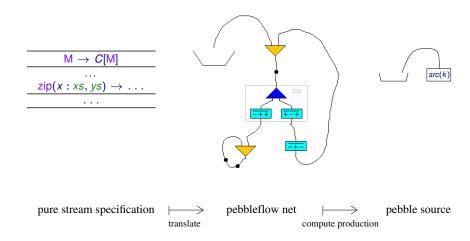
one has to associate 0 to its argument, for it first has to consume a stream element before it can produce one (two)

but then e.g.:

$$D \rightarrow 1 : tail(dup(D))$$

is not recognized to be productive

# deciding productivity via pebbleflow



## desiderata specification format

we want a syntactic format of stream specifications such that:

- the exact production moduli of stream functions can be computed
- the set of moduli is closed under composition and infimum
- least fixed points of production moduli can be computed

# the pure stream format (PSF)

### example

stream constants
stream functions

a pure stream specification is a 3-layered,  $\{S, D\}$ -sorted orthogonal constructor TRS:

- data layer: stream independent, terminating TRS
- stream function layer: no pattern matching on data, no nesting of stream functions
- stream constant layer: no restrictions

## no nesting in stream function rules in PSF

- ▶ in PSF no nesting is allowed in the stream function layer
- ► (nesting in the stream constant layer *is* allowed!)

### example (productive, but not pure)

```
zeros \rightarrow 0 : log(exp(zeros))

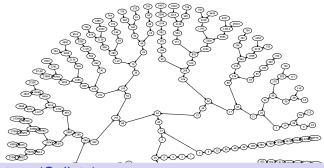
exp(x : xs) \rightarrow x : dup(exp(xs))

log(x : xs) \rightarrow x : log(odd(xs))
```

### no stream dependent data functions in PSF

- in PSF data terms cannot be built using stream terms
- stream dependent data functions possibly create 'look-ahead'
- productivity is undecidable for specification including stream dependent data functions

```
example (productive? iff n is even)
S \to 0: S(n): S
head(x:xs) \to x
tail(x:xs) \to xs
where <math>S(n) := head(tail^n(S))
```



### conjecture (Collatz)

 $(\forall n \geq 1) (\exists i \in \mathbb{N}) (f^i(n) = 1)$ , where f is defined for all  $n \geq 1$  by:

$$f(n) := \begin{cases} \frac{n}{2} & n \text{ is even} \\ 3n+1 & n \text{ is odd} \end{cases}$$



writing '•' for successful termination, and dividing 3n + 1 immediately by 2 in case n is odd, the Collatz conjecture can be reformulated as:

$$(\forall n \geq 1) (\exists i \in \mathbb{N}) (F^i(n) = \bullet)$$

for  $F : \mathbb{N} \to \mathbb{N} \cup \{\bullet\}$  defined by:

$$F(1) = \bullet$$

$$F(2n) = n \qquad (n \ge 1)$$

$$F(2n+1) = 3n+2 \qquad (n \ge 1)$$

let

$$C \rightarrow \bullet : zip(C, third(tail^4(C)))$$
  
 $zip(xs, ys) \rightarrow head(xs) : zip(ys, tail(xs))$   
 $third(xs) \rightarrow head(xs) : third(tail^3(xs))$ 

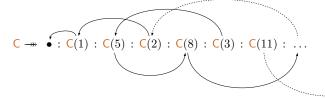
then

$$C \twoheadrightarrow \bullet : C(1) : C(5) : C(2) : C(8) : C(3) : C(11) : \dots$$

resembling Collatz function F written as a stream:

$$F = \bullet : 1 : 5 : 2 : 8 : 3 : 11 : \dots$$

picturing the 'runs' through C, we get



$$C \rightarrow \bullet : zip(C, third(tail^4(C)))$$

### proposition

Collatz conjecture is true

 $\iff$  the specification for C is productive:

C ---- •:•:•:•:...

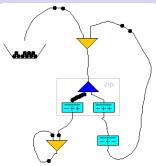
## the stream of positive rational numbers

yet another pure stream specification:

```
rats \rightarrow mkpairs(aux)
                    aux \rightarrow 0 : aux'
                   aux' \rightarrow 1 : zip(aux', add(aux', tail(aux')))
           tail(x:xs) \rightarrow xs
\mathsf{mkpairs}(x:y:xs) \to \langle x,y \rangle : \mathsf{mkpairs}(y:xs)
      zip(x:xs,ys) \rightarrow x:zip(ys,xs)
 add(x:xs,y:ys) \rightarrow (x+y):add(xs,ys)
                 0 + v \rightarrow v
             s(x) + y \rightarrow s(x + y)
```

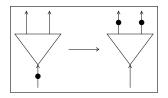
rats  $\twoheadrightarrow$   $\langle 0, 1 \rangle : \langle 1, 1 \rangle : \langle 1, 2 \rangle : \langle 2, 1 \rangle : \langle 1, 3 \rangle : \langle 3, 2 \rangle : \langle 2, 3 \rangle : \langle 3, 1 \rangle : \langle 1, 4 \rangle \dots$ 

# modelling stream specifications by pebbleflow nets



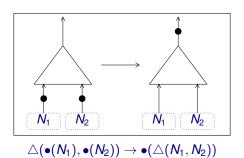
- ▶ tool for visualizing the production of a stream specification
- stream elements are abstracted from in favour of 'pebbles' •
- evaluation is modelled by the flow of pebbles through the net
- pebbleflow nets can be implemented by interaction nets
- a pure stream specification is productive if and only if its net generates an infinite chain of pebbles
- a pebbleflow net is a network built of pebble processing units (fans, boxes, meets, sources) connected by wires

#### fan

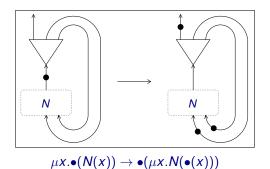


- duplicates an incoming pebble along its output ports
- explicit sharing device
- enables construction of cyclic nets
- used to implement recursion, in particular feedback

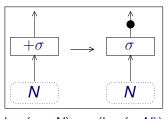
### meet

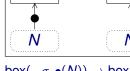


### recursion/feedback



#### box





 $\mathsf{box}(+\sigma, N) \to \bullet(\mathsf{box}(\sigma, N))$ 

 $\mathsf{box}(-\sigma, \bullet(N)) \to \mathsf{box}(\sigma, N)$ 

 $\sigma$ 

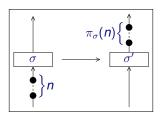
- ▶ boxes contain I/O sequences: infinite sequences over {-,+}
- ▶ I/O sequences model production moduli of stream functions
- ▶ + : a ready state for an output pebble
- ▶ : a requirement for an input pebble

# translating unary stream functions to I/O sequences

```
Example  \begin{aligned} \operatorname{dup}(x:xs) &\to x:x:\operatorname{dup}(xs) \\ \operatorname{even}(x:xs) &\to x:\operatorname{odd}(xs) \\ \operatorname{odd}(x:xs) &\to \operatorname{even}(xs) \end{aligned}   \begin{aligned} [\operatorname{dup}] &= -++[\operatorname{dup}] & [\operatorname{dup}] &= \overline{-++} \\ [\operatorname{even}] &= -+[\operatorname{odd}] & [\operatorname{even}] &= \overline{-+-} \\ [\operatorname{odd}] &= -[\operatorname{even}] & [\operatorname{odd}] &= \overline{--+} \end{aligned}
```

• every pure stream function f is mapped to an eventually periodic I/O sequence [f], corresponding to an eventually periodically increasing function  $\pi_{[\![f]\!]}$ , which forms the exact production modulus of f

# from I/O sequences to production moduli



▶ the production function  $\pi_{\sigma} : \overline{\mathbb{N}} \to \overline{\mathbb{N}}$  of  $\sigma \in \pm^{\omega}$  is defined by  $\pi_{\sigma}(n) := \pi(\sigma, n)$ :

$$\pi(+\sigma, n) = \pi(\sigma, n) + 1$$
  

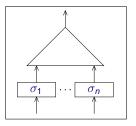
$$\pi(-\sigma, 0) = 0$$
  

$$\pi(-\sigma, n + 1) = \pi(\sigma, n)$$

• example: 
$$\pi_{\texttt{[dup]}}(n) = \pi_{-++}(n) = 2n$$
 
$$\pi_{\texttt{[odd]}}(n) = \pi_{--+}(n) = \lfloor \frac{n}{2} \rfloor$$

## translation of stream functions into gates

gates are used to model stream functions



a gate with *n* input ports

### example

$$\begin{aligned} \operatorname{zip}(x:xs,ys) &= x: \operatorname{zip}(ys,xs) \\ [\![\operatorname{zip}]\!]_1 &= -+[\![\operatorname{zip}]\!]_2 \\ [\![\operatorname{zip}]\!]_2 &= +[\![\operatorname{zip}]\!]_1 \end{aligned} \qquad [\![\operatorname{zip}]\!]_2 &= \overline{+-+} \end{aligned}$$

## pebbleflow rewrite system

▶ terms for pebbleflow nets  $(k \in \mathbb{N} \cup \{\infty\}, x \in VAR, \sigma \in \pm^{\omega})$ :

$$N ::= \operatorname{src}(k) \mid x \mid \bullet(N) \mid \operatorname{box}(\sigma, N) \mid \mu x. N \mid \triangle(N, N)$$

pebbleflow rewrite rules:

$$\triangle(\bullet(N_1), \bullet(N_2)) \rightarrow \bullet(\triangle(N_1, N_2))$$

$$\mu x. \bullet(N(x)) \rightarrow \bullet(\mu x. N(\bullet(x)))$$

$$box(+\sigma, N) \rightarrow \bullet(box(\sigma, N))$$

$$box(-\sigma, \bullet(N)) \rightarrow box(\sigma, N)$$

$$src(1 + k) \rightarrow \bullet(src(k))$$

### pebbleflow tool

- net visualization tool written by Ariya Isihara,
- available via: http://infinity.few.vu.nl/productivity

```
J \rightarrow 0:1:even(J) [\![J]\!] = \mu x. \bullet (\bullet (box(\overline{-+-},x))) rats \rightarrow mkpairs(aux) aux \rightarrow 0: aux' aux' \rightarrow 1: zip(aux', add(aux', tail(aux')))
```

$$\llbracket \mathsf{rats} \rrbracket = -\overline{-+} ( \bullet (\mu \mathsf{X}. \bullet (\triangle (\overline{-++}(\mathsf{X}), \overline{+-+}(\triangle (\overline{-+}(\mathsf{X}), \overline{-+}(-\overline{-+}(\mathsf{X}))))))) )$$
 (where box $(\sigma, \mathsf{N})$  is abbreviated to  $\sigma(\mathsf{N})$ )

## translation preserves production

▶ for a specification  $\mathcal{R} = \langle \Sigma, R \rangle$  the production  $\Pi_{\mathcal{R}}(t)$  of a stream term t is defined by:

$$\Pi_{\mathcal{R}}(\mathsf{t}) := \sup\{ \, n \in \mathbb{N} \mid \mathsf{t} \twoheadrightarrow \mathsf{d}_1 : \ldots : \mathsf{d}_n : \mathsf{t}' \, \}$$

▶ the production  $\Pi_{\bullet}(N)$  of a net N is defined by:

$$\Pi_{\bullet}(N) := \sup\{ n \in \mathbb{N} \mid N \twoheadrightarrow \bullet^{n}(N') \}$$

translation of pure stream specifications to pebbleflow nets preserves production:

$$\Pi_{\bullet}([\![M]\!])=\Pi_{\mathcal{R}}(M)$$

pure stream specifications are translated into rational nets, i.e. nets with eventually periodic I/O sequences only

### net reduction

```
\bullet(N) \rightarrow box(+\overline{-+}, N)
box(\sigma_1, box(\sigma_2, N)) \rightarrow box(\sigma_1 \circ \sigma_2, N)
   box(\sigma, \triangle(N_1, N_2)) \rightarrow \triangle(box(\sigma, N_1), box(\sigma, N_2))
           \mu x.\triangle(N_1,N_2) \rightarrow \triangle(\mu x.N_1,\mu x.N_2)
                            \mux.N \rightarrow N
                                                                                                         if x \notin FV(N)
             \mu x.\mathsf{box}(\sigma,x) \to \mathsf{src}(\mathsf{fix}(\sigma))
 \triangle(\operatorname{src}(k_1),\operatorname{src}(k_2))\to\operatorname{src}(\min(k_1,k_2))
           box(\sigma, src(k)) \rightarrow src(\pi_{\sigma}(k))
                             \mu x.x \rightarrow src(0)
```

# box composition & fixed point computation

composition 
$$\circ: \pm^\omega \times \pm^\omega \to \pm^\omega$$
 is defined by: 
$$+\sigma \circ \tau = +(\sigma \circ \tau) \\ -\sigma \circ +\tau = \sigma \circ \tau \\ -\sigma \circ -\tau = -(-\sigma \circ \tau)$$

• associative, preserves periodicity, and implements composition of the production functions:  $\pi_{\sigma\circ\tau}=\pi_\sigma\circ\pi_\tau$ 

the operation fixed point fix  $: \pm^{\omega} \to \overline{\mathbb{N}}$  is defined by:

$$\begin{aligned} \operatorname{fix}(+\sigma) &= 1 + \operatorname{fix}(\delta(\sigma)) & \delta(+\sigma) &= +\delta(\sigma) \\ \operatorname{fix}(-\sigma) &= 0 & \delta(-\sigma) &= \sigma \end{aligned}$$

•  $fix(\sigma)$  is the least fixed point of  $\pi_{\sigma}$ 

## properties of net reduction

net reduction is production preserving:

$$N woheadrightarrow N'$$
 implies  $\Pi_{\bullet}(N) = \Pi_{\bullet}(N')$ 

- net reduction is terminating and confluent, hence: every closed net normalises to a unique normal form
- ▶ normal forms are of the form src(k), a source of k pebbles, with  $k \in \mathbb{N}$  or  $k = \infty$
- normal forms of rational nets can be computed

### **ProPro**

- ProPro: a tool for proving productivity
- available via: http://infinity.few.vu.nl/productivity

```
\begin{split} J &\to 0:1: even(J) \\ aux &\to 0: aux' \\ aux' &\to 1: zip(aux', add(aux', tail(aux'))) \\ rats &\to mkpairs(aux) \end{split}
```

## productivity of PSF is decidable

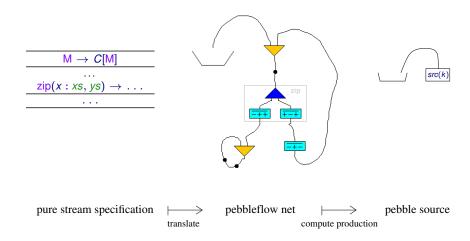
#### theorem

productivity of pure stream specifications is decidable

decision algorithm for stream constant M in pure stream specification  $\mathcal{R}$ :

- ► translate M to the rational net M
- ► reduce [M] to a source src(k)
- $(recall \Pi_{\mathcal{R}}(\mathsf{M}) = \Pi_{\bullet}(\llbracket \mathsf{M} \rrbracket) = k)$
- ▶ if  $k = \infty$ , output:  $\mathcal{R}$  is productive for M
- ▶ if  $k \in \mathbb{N}$ , output:  $\mathcal{R}$  is not productive for M

# deciding productivity via pebbleflow



### conclusion

- previous approaches: sufficient conditions for productivity, not automatable or only for a limited subclass
- pebbleflow approach: decision algorithm for productivity of a rich class of stream specifications, only stream function layer is restricted

#### coming up next:

 extended formats of stream specifications: stream functions defined using pattern matching on data