Resource Constrained Programming with Full Dependent Types

Robert Atkey Strathclyde University, Glasgow robert.atkey@strath.ac.uk

CCS Colloquium, Augusta University 20th November 2020

- Programming Language
- Proof Language

Programming Language

Proof Language

So we can write programs

- Programming Language
- Proof Language

So we can write programs

and reason about them

- Programming Language
- Proof Language

So we can write programs

and reason about them but only the "extensional behaviour"

Having predicates for complexity won't work:

 $Ptime : (Nat \rightarrow Nat) \rightarrow Set$

Allows the theory to distinguish extensionally equivalent functions.

Two ideas:

- Implicit: all functions are in a fixed complexity class (e.g., Ртіме)
- Explicit: types tell us what the complexity is.

This talk

- Implicit and explicit typed complexity analysis for Dependent Type Theory

Challenges

- Nice systems for implicit and explicit complexity
- Integrating them with dependent types

Two Implicit PTIME systems

Requirements

- Extension of typed λ -calculus; *higher order*
- No impredicative polymorphism (no Church encodings)
- Proper datatypes (definitely no Church encodings)

Requirements

- Extension of typed λ -calculus; *higher order*
- No impredicative polymorphism (no Church encodings)
- Proper datatypes (definitely no Church encodings)

Forget dependent types for now

- Simply typed λ -calculus
- A natural number type NAT, zero, suc with an iterator

 $\frac{\Gamma \vdash M_z : A \quad \Gamma, x : A \vdash M_s : A \quad \Gamma \vdash N : \text{NAT}}{\Gamma \vdash \text{iter}(M_z, x. M_s, N) : A}$

Easily yields exponential time:

iter(suc, f. λx . f(f(x)), N) zero : NAT

computes 2^N

Easily yields exponential time:

iter(suc, f. λx . f(f(x)), N) zero : NAT

computes 2^N

Culprits

- Duplication of the higher order value f
- Construction of new numbers

Disallows:

iter(suc, f. λx . f(f(x)), N) zero : NAT

because f is used twice.

Disallows:

iter(suc, f. λx . f(f(x)), N) zero : NAT

because f is used twice.

But

Can write:

dup : NAT \multimap NAT \otimes NAT dup x = iter((zero, zero), (m, n).(suc m, suc n), x)

Disallows:

iter(suc, f. λx . f(f(x)), N) zero : NAT

because f is used twice.

But

Can write:

dup : NAT
$$\multimap$$
 NAT \otimes NAT
dup x = iter((zero, zero), (m, n).(suc m, suc n), x)

- add : NAT \multimap NAT \multimap NAT is linear.

Disallows:

iter(suc, f. λx . f(f(x)), N) zero : NAT

because f is used twice.

But

Can write:

dup : NAT
$$\multimap$$
 NAT \otimes NAT
dup x = iter((zero, zero), (m , n).(suc m , suc n), x)

- mul : NAT \neg NAT \neg NAT can be written using dup, add.

Disallows:

iter(suc, f. λx . f(f(x)), N) zero : NAT

because f is used twice.

But

Can write:

dup : NAT
$$\neg o$$
 NAT \otimes NAT
dup x = iter((zero, zero), (m, n).(suc m, suc n), x)

- add : NAT \multimap NAT \multimap NAT is linear.
- mul : NAT \neg NAT \neg NAT can be written using dup, add.
- $exp : NAT \multimap NAT \multimap NAT$ can be written using dup, mul.

Disallows:

iter(suc, f. λx . f(f(x)), N) zero : NAT

because f is used twice.

But

Can write:

dup : NAT
$$\neg o$$
 NAT \otimes NAT
dup x = iter((zero, zero), (m, n).(suc m, suc n), x)

- add : NAT \multimap NAT \multimap NAT is linear.
- mul : NAT \multimap NAT \multimap NAT can be written using dup, add.
- exp : NAT NAT NAT can be written using dup, mul.
- Get exponential time.

Linearity + No constructors

- Can't write dup or add (or mul or exp)

Linearity + No constructors

- Can't write dup or add (or mul or exp)
- Iterable NAT:
 - Not constructible
 - Has an iterator

Linearity + No constructors

- Can't write dup or add (or mul or exp)
- Iterable NAT:
 - Not constructible
 - Has an iterator
- Non-iterable NAT°:
 - Constructible
 - Case analysis

 $\frac{\Gamma_1 \vdash M_z : A \quad \Gamma_2, x : \operatorname{Nat}^{\circ} \vdash M_s : A \quad \Gamma_3 \vdash N : \operatorname{Nat}^{\circ}}{\Gamma_1, \Gamma_2, \Gamma_3 \vdash \operatorname{case}(M_z, x. M_s, N) : A}$

Is this enough?

- Only source of iterable NAT is the input
- So only linear time in the size of the NAT "fuel" provided
- To get polytime, allow duplication of variables of type NAT.

Is this enough?

- Only source of iterable NAT is the input
- So only linear time in the size of the NAT "fuel" provided
- To get polytime, allow duplication of variables of type NAT.

Completeness

- Given a step function s: TAPE \multimap TAPE, and a \mathbb{N} -polynomial $p(n) = \sum a_i n^i$
- *n* iterations: iter $(\lambda x. x, f. \lambda x. s(fx), n)$: TAPE \multimap TAPE
- n^2 iterations: iter($\lambda x.x, f.\lambda x.$ iter($\lambda x.x, f.\lambda x. s(fx), n$), n) : TAPE \multimap TAPE
- n^i iterations...
- Addition by composition

Recovering Constructibility?

- This system works, but is restricted to everything being driven by NAT-iteration
- Some programs are more easily expressible by iteration over trees, etc.

Martin Hofmann's LFPL: principle of "conservation of iterability"

- A special type \diamond , representing a chunk of iterability
- Required for construction:

 $zero : \diamondsuit \multimap Nat$ $suc : \diamondsuit \multimap Nat \multimap Nat$

Recovered on iteration:

 $\frac{\Gamma_1, d: \diamond \vdash M_z : A \qquad d: \diamond, x : A \vdash M_s : A \qquad \Gamma_2 \vdash N : \text{NAT}}{\Gamma_1, \Gamma_2 \vdash \text{iter}(d. M_z, d x. M_s, N) : A}$

Extends easily to other datatypes

Iterating a step function

- Assume we have a function *s* : TAPE → TAPE one step of a Turing machine
- Linear $\binom{n}{1}$ iterations:

$$I_1 = \lambda(n, t).iter(d. (zero(d), t), : NAT \otimes TAPE \multimap NAT \otimes TAPE$$
$$d(n, t). (suc(d, n), s t),$$
$$n)$$

 $-\binom{n}{2}$ iterations:

$$\begin{split} I_2 &= \lambda(n, t). \texttt{iter}(d. \ (\texttt{zero}(d), t), \\ & d \ (n, t). \ \texttt{let} \ (n, t) = I_1(n, t) \ \texttt{in} \ (\texttt{suc}(d, n), s(t)), \\ n) \end{split}$$
: NAT \otimes TAPE \multimap NAT \otimes TAPE \rightarrow

 $-\binom{n}{3}$ iterations: Iterate the above

Iterating a step function

- Obtain a $\binom{n}{k}$ iterator for any k
- And get the original number back as an output
- Chain them together to get any polynomial:

$$p(n) = \sum_{i=0}^{k} p_i \binom{n}{k}$$

So we get polytime completeness

Explicit Complexity

Amortised Resource Analysis -

- Reinterpret \diamond as the cost of a step of iteration
- Inspired by Tarjan's *amortised complexity analyis*
 - storing potential inside data structures
- Building a NAT still requires \diamond s:

 $\mathsf{zero}: \diamondsuit \multimap \mathsf{Nat} \qquad \mathsf{suc}: \diamondsuit \multimap \mathsf{Nat} \multimap \mathsf{Nat}$

- But iteration no longer gives you them back:

 $\frac{\Gamma_1 \vdash M_z : A \qquad x : A \vdash M_s : A \qquad \Gamma_2 \vdash N : \text{NAT}}{\Gamma_1, \Gamma_2 \vdash \text{iter}_A(M_z, x.M_s, N) : A}$

– Back to linear time...

More flexibility

- Annotate data structures with number of \diamond s per constructor

 Nat^p

– Duplication:

 $\operatorname{Nat}^{p_1+p_2} \multimap \operatorname{Nat}^{p_1} \otimes \operatorname{Nat}^{p_2}$

- Hofmann & Jost (2001) used linear programming to infer the *p*s

Regaining polynomial time –

(Hoffmann & Hofmann, ESOP 2010)

- Annotate with sequences of naturals:

 $\operatorname{Nat}^{(p_1,\ldots,p_k)}$

Interpretation is that

 $\sum_{i=1}^{k} p_i \binom{n}{i}$

is the number of \diamond s is attached to a natural *n*.

Regaining polynomial time –

(Hoffmann & Hofmann, ESOP 2010)

- Annotate with sequences of naturals:

 $\operatorname{Nat}^{(p_1,\ldots,p_k)}$

– Interpretation is that

 $\sum_{i=1}^{k} p_i \binom{n}{i}$

is the number of \diamond s is attached to a natural *n*.

– Iterator:

$$\begin{split} &\Gamma_1 \vdash M_z : A \\ &n : \operatorname{NAT}^{(p_1 + p_2, p_2 + p_3, \dots, p_k)}, d : \diamondsuit^{p_1}, x : A \vdash M_s : A \\ &\Gamma_2 \vdash N : \operatorname{NAT}^{(p_1 + 1, \dots, p_k)} \end{split}$$

 $\Gamma_1, \Gamma_2 \vdash iter(M_z, n \ d \ x.M_s, N) : A$

Adapting these systems to dependent types

Dependency and Accountancy

In Martin-Löf Type Theory

$x_1: S_1, \ldots, x_n: S_n \vdash M: T$

In Martin-Löf Type Theory

$$x_1:S_1,\ldots,x_n:S_n \vdash M:T$$

variables x_1, \ldots, x_n are mixed usage

x is used *computationally*

x is used *computationally*

n is used *logically*

In Linear Logic

$x_1: X_1, \ldots, x_n: X_n \vdash M: Y$

In Linear Logic

$x_1: X_1, \ldots, x_n: X_n \vdash M: Y$

the presence of a variable x records its usage each x_i must be "used" by M exactly once

In Linear Logic

$x_1: X_1, \ldots, x_n: X_n \vdash M: Y$

the presence of a variable x records its usage each x_i must be "used" by M exactly once

Enables:

- 1. Insight into computational behaviour
- **2.** e.g., time complexity

Can we read this judgement linearly?

Can we read this judgement linearly?

▶ *n* appears in the context, but is not used computationally

Can we read this judgement linearly?

▶ *n* appears in the context, but is not used computationally

 \triangleright *n* appears *twice* in types

Can we read this judgement linearly?

▶ *n* appears in the context, but is not used computationally

 \triangleright *n* appears *twice* in types

Is *n* even used at all?

Separate intuitionistic / unrestricted uses and linear uses

Separate intuitionistic / unrestricted uses and linear uses

▷ Types can depend on intuitionistic data, but not linear data

▷ Separate *intuitionistic / unrestricted* uses and *linear* uses

▷ Types can depend on intuitionistic data, but not linear data

(Barber, 1996) (Cervesato and Pfenning, 2002) (Krishnaswami, Pradic, and Benton, 2015) (Vákár, 2015) Quantitative Coeffect calculi:

$$x_1 \stackrel{\rho_1}{:} S_1, \ldots, x_n \stackrel{\rho_n}{:} S_n \vdash M : T$$

Quantitative Coeffect calculi:

$$x_1 \stackrel{\rho_1}{:} S_1, \ldots, x_n \stackrel{\rho_n}{:} S_n \vdash M : T$$

- ▷ The ρ_i record usage from some semiring *R*
 - . $1 \in R a$ use
 - . $0 \in R$ not used
 - . ρ_1 + ρ_2 adding up uses (e.g., in an application)
 - . $\rho_1\rho_2-{\rm nested}$ uses

Quantitative Coeffect calculi:

$$x_1 \stackrel{\rho_1}{:} S_1, \ldots, x_n \stackrel{\rho_n}{:} S_n \vdash M : T$$

▷ The
$$\rho_i$$
 record usage from some semiring *R*

- . $1 \in R a$ use
- . $0 \in R \text{not used}$
- . $\rho_1 + \rho_2$ adding up uses (e.g., in an application)
- . $\rho_1\rho_2-{\rm nested}$ uses

(Petricek, Orchard, and Mycroft, 2014) (Brunel, Gaboardi, Mazza, and Zdancewic, 2014) (Ghica and Smith, 2014)

McBride's idea:

▷ allow 0-usage data to appear in types. (McBride, 2016)

McBride's idea:

▷ allow 0-usage data to appear in types. (McBride, 2016)

$$x_1 \stackrel{\rho_1}{:} S_1, \ldots, x_n \stackrel{\rho_n}{:} S_n \vdash M \stackrel{\sigma}{:} T$$

where $\sigma \in \{0, 1\}$. $\triangleright \sigma = 1 - \text{the "real" computational world}$ $\triangleright \sigma = 0 - \text{the types world}$

(allowing arbitrary ρ yields a system where substitution is inadmissible (Atkey, 2018))

McBride's idea:

▷ allow 0-usage data to appear in types. (McBride, 2016)

$$x_1 \stackrel{\rho_1}{:} S_1, \ldots, x_n \stackrel{\rho_n}{:} S_n \vdash M \stackrel{\sigma}{:} T$$

where $\sigma \in \{0, 1\}$. $\triangleright \sigma = 1 - \text{the "real" computational world}$ $\triangleright \sigma = 0 - \text{the types world}$

(allowing arbitrary ρ yields a system where substitution is inadmissible (Atkey, 2018))

Zero-ing is an admissible rule:

$$\frac{\Gamma \vdash M \stackrel{!}{:} T}{0\Gamma \vdash M \stackrel{0}{:} T}$$
 allowing promotion to the type world

Zero-ing is admissible

$$\frac{\Gamma \vdash M \stackrel{1}{:} T}{0\Gamma \vdash M \stackrel{0}{:} T}$$

means that every linear term has an "extensional" counterpart (or constitutent)

which can be used at type checking time to construct types

has the effect of making the linear system a restriction of the intuitionistic

A suitable semiring for affine linearity?

- Carrier: $\{0, 1, \omega\}$
- Ordered: $\omega < 1 < 0$
- Operations:

		1				1	
		1				0	
		ω		1	0	1	ω
ω	ω	ω	ω	ω	0	ω	ω

- Would admit an unrestricted ! modality.

Strict resource counting

- Carrier: \mathbb{N}
- Ordered: $\cdots < 2 < 1 < 0$
- Operations: normal operations on $\mathbb N$

Diamonds

$$\frac{\Gamma \vdash}{0\Gamma \vdash \diamondsuit} \text{Ty-Dia} \qquad \qquad \frac{0\Gamma \vdash}{0\Gamma \vdash * \overset{0}{:} \diamondsuit} \text{Tm-Dia}$$

- In the $\sigma = 0$ fragment, $\diamond s$ are free.

LFPL

Natural number introduction

 $\frac{\Gamma \vdash d \stackrel{\sigma}{:} \diamondsuit}{\Gamma \vdash \operatorname{zero}(d) \stackrel{\sigma}{:} \operatorname{Nat}}$

 $\frac{\Gamma \vdash d \stackrel{\sigma}{:} \diamondsuit}{\Gamma \vdash \operatorname{succ}(d, n) \stackrel{\sigma}{:} \operatorname{Nat}}$

LFPL

- Natural number elimination ($\sigma = 1$ case)

$$0\Gamma, x: \operatorname{NAT} \vdash A$$

$$\Gamma_1, d \stackrel{1}{:} \diamond \vdash M_z \stackrel{1}{:} A\{\operatorname{zero}(*)/x\}$$

$$d \stackrel{1}{:} \diamond, n \stackrel{0}{:} \operatorname{NAT}, r \stackrel{1}{:} A\{n/x\} \vdash M_s \stackrel{1}{:} A\{\operatorname{succ}(*, n)/x\}$$

$$\Gamma_2 \vdash N \stackrel{1}{:} \operatorname{NAT}$$

$$\Gamma_1 + \Gamma_2 = \Gamma$$

 $\overline{\Gamma} \vdash \operatorname{iter}(x.A, d.M_z, d \ n \ r.M_s, N) \stackrel{1}{:} A\{N/x\}$

- Crucial: n is not available for computational use in M_s .

Encoding lists

– Define (in $\sigma = 0$ fragment):

 $\operatorname{Vec} A:\operatorname{Nat}\to\operatorname{Set}$

by iteration on the natural number.

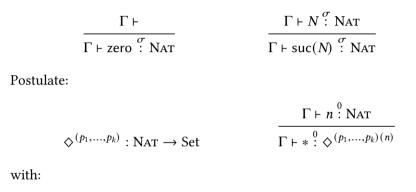
– Lists:

List $A = (n \stackrel{1}{:} NAT) \otimes Vec A n$

Amortised Analysis

_

- Unrestricted introduction rules for natural numbers:



split :
$$(n \stackrel{0}{:} \operatorname{NAT}) \to \diamond^{(p_1+p'_1,\dots,p_k+p'_k)}(n) \multimap \diamond^{(p_1,\dots,p_k)}(n) \otimes \diamond^{(p'_1,\dots,p'_k)}(n)$$

join : $(n \stackrel{0}{:} \operatorname{NAT}) \to \diamond^{(p_1,\dots,p_k)}(n) \otimes \diamond^{(p'_1,\dots,p'_k)}(n) \multimap \diamond^{(p_1+p'_1,\dots,p_k+p'_k)}(n)$
shift : $(n \stackrel{0}{:} \operatorname{NAT}) \to \diamond^{(p_1,\dots,p_k)}(\operatorname{suc}(n)) \multimap \diamond^{(p_1+p_2,\dots,p_k)}(n)$

Amortised Analysis

- Natural number elimination ($\sigma = 1$ case)

$$0\Gamma, x \stackrel{0}{:} \operatorname{NAT} \vdash A$$

$$\Gamma_{1} \vdash M_{z} \stackrel{1}{:} A\{\operatorname{zero}/x\}$$

$$n \stackrel{1}{:} \operatorname{NAT}, r \stackrel{1}{:} A\{n/x\} \vdash M_{s} \stackrel{1}{:} A\{\operatorname{succ}(n)/x\}$$

$$\Gamma_{2} \vdash N \stackrel{1}{:} \operatorname{NAT}$$

$$\Gamma_{3} \vdash D \stackrel{1}{:} \diamondsuit^{(1)}(N)$$

$$\Gamma_{1} + \Gamma_{2} + \Gamma_{3} = \Gamma$$

$$\overline{\Gamma} \vdash \operatorname{iter}(x.A, M_{z}, n r.M_{s}, N, D) : A\{N/x\}$$

- *n* is available for use in M_s
- Pay up front for the iteration with *D*
- Get nested iteration by passing in enough \diamondsuit{s} to pay for it

$$A[n] = \diamondsuit^{(p_1, \dots, p_k)}(n) \multimap B[n]$$

Semantic Interpretation : Soundness

Resource monoids

- Let $\mathbb{N}_{-\infty}$ be category with objects $\mathbb{N} \cup \{-\infty\}$ and $m \to n$ if $m \le n$, with $-\infty \le n$
 - Strict symmetric monoidal category with (+, 0)
- − A resource monoid *M* is a $\mathbb{N}_{-\infty}$ -enriched strict symmetric monoidal category.

Resource monoids

- Let $\mathbb{N}_{-\infty}$ be category with objects $\mathbb{N} \cup \{-\infty\}$ and $m \to n$ if $m \le n$, with $-\infty \le n$
 - Strict symmetric monoidal category with (+, 0)
- A resource monoid *M* is a $\mathbb{N}_{-\infty}$ -enriched strict symmetric monoidal category.
- (*M*, +, 0) is a commutative monoid
- $\quad 0 \le M(\alpha, \alpha)$
- *−* $M(\alpha, \beta) \in \mathbb{N}_{-\infty}$ is the difference between *α* and *β*
- $M(\alpha,\beta) + M(\beta,\gamma) \le M(\alpha,\gamma)$
- $\quad M(\alpha,\beta) \leq M(\alpha+\gamma,\beta+\gamma)$

Resource monoids

Linear time:

- $M = \mathbb{N}$
- Differencing:

$$M(n, m) = \begin{cases} m - n & n \le m \\ -\infty & \text{otherwise} \end{cases}$$

- Wrinkle: counts recursion steps, not the actual number of steps.

Resource Monoids: Polynomial time (for LFPL)

- $M \ni (n, p)$, where
 - − $n \in \mathbb{N}$ is the amount of iterability (number of \diamond s)
 - *p* is a polynomial with \mathbb{N} coefficients

$$- (n, p) + (m, q) = (n + m, p + q).$$

Cost differencing:

$$M((n, p), (m, q)) = \begin{cases} q(m) - p(m) & n \le m \text{ and } (q - p) \text{ is non-negative} \\ & \text{and non-decreasing } \ge m \\ -\infty & \text{otherwise} \end{cases}$$

Resource Monoids: Polynomial time (for Constructor-free System)

- $M \ni (n, p)$, where
 - − $n \in \mathbb{N}$ is the amount of iterability (number of \diamond s)
 - *p* is a polynomial with \mathbb{N} coefficients

$$- (n, p) + (m, q) = (\max n \, m, p + q).$$

Cost differencing:

$$M((n, p), (m, q)) = \begin{cases} q(m) - p(m) & n \le m \text{ and } (q - p) \text{ is non-negative} \\ & \text{and non-decreasing } \ge m \\ -\infty & \text{otherwise} \end{cases}$$

- Hofmann and Dal Lago used this resource monoid for Lafont's *Soft Linear Logic*.

Cost model

- Assume a model of computation with a cost model:

 $e,\eta \Downarrow_k v$

step count *k*, expressions $e \in \mathcal{E}$, values $v \in \mathcal{V}$.

Interpretation of Types and Terms

- Types are interpreted by $(|X|, \models_X)$ where:
 - |X| is a set
 - $\models_X \subseteq (M \times \mathcal{V}) \times |X|.$
- Functions $f: X \to Y$:
 - $\quad f\colon |X|\to |Y|$
 - exists $e \in \mathcal{E}$, $\gamma \in M$, such that
 - for all α , ν , x.

 $(\alpha, \nu) \models_X x \text{ implies} \\ \text{exists } \beta, k, \nu' \text{ s.t.} \\ e, [\nu] \downarrow_k \nu', \\ (\beta, \nu') \models_Y f(x), \\ k \le M(\alpha + \gamma, \beta)$

Some types

In the amortised system:

$$- \quad \diamondsuit = (\{*\}, (n, *) \models_{\diamondsuit} * \Leftrightarrow n \ge 1)$$

In LFPL:

$$- \quad \diamondsuit = (\{*\}, ((n, p), * \models_{\diamondsuit} * \Leftrightarrow n \ge 1, p \ge 0)$$

- NAT =
$$(\mathbb{N}, ((n, p), n \models_{\diamond} \underline{m}) \Leftrightarrow n \ge m, p \ge 0)$$

In the constructor free system:

- NAT =
$$(\mathbb{N}, ((n, p), n \models_{\diamond} \underline{m}) \Leftrightarrow n \ge m, p \ge 0)$$

Summary

- Quantitative Type Theory for Complexity Analysis
- ▷ Careful combination of dependency and linearity
- Dependent Types for reasoning about programs
- ▷ Dependent Types for reasoning about complexity (in the explicit system)

▷ Quantitative Type Theory for Complexity Analysis

▷ Careful combination of dependency and linearity

Dependent Types for reasoning about programs

▷ Dependent Types for reasoning about complexity (in the explicit system)

Related Work

Sized types

Used for controlling well foundedness For complexity analysis require "tick" monads

- Gaboardi and Dal Lago: Linear Dependent Types for ICC Dependent Types only for counting time
- ► Future:
 - ► LAL, EAL, BLL, Logspace, ...
 - Polytime mathematics?