Examples and Results from a BSc-level Course on Domain Specific Languages of Mathematics

Patrik Jansson   Sólrún Halla Einarsdóttir   Cezar Ionescu

Functional Programming division, Chalmers University of Technology

WG2.1, 2018-07-05

Background to *Domain-Specific Languages of Mathematics*

- Gothenburg, 2015 (73:rd meeting): Jansson and Ionescu: "DSLM - Presenting Mathematical Analysis Using Functional Programming".
- Pedagogical project to develop the course (incl. material)
- 2015: paper at "Trends in Functional Programming in Education"
- 2016, 17, 18: Undergraduate course at Chalmers (28, 43, 39 students)
- 2018: new TFPIE paper (reported on in this talk)

Course goal and focus

Goal

Encourage students to approach mathematical domains from a functional programming perspective.

Course focus

- Make functions and types explicit
- Explicit distinction between syntax and semantics
- Types as carriers of semantic information
- Organize the types and functions in DSLs
- [New] Make variable binding and scope explicit

Lecture notes and more available at:
`https://github.com/DSLsofMath/DSLsofMath`

Example 1 - The limit of a function

We say that $f(x)$ ***approaches the limit*** L *as* x ***approaches*** a*, and we write*

$$\lim_{x \to a} f(x) = L,$$

if the following condition is satisfied:
for every number $\varepsilon > 0$ *there exists a number* $\delta > 0$*, possibly depending on* ε*, such that if* $0 < |x - a| < \delta$*, then* x *belongs to the domain of* f *and*

$$|f(x) - L| < \varepsilon$$

- Adams & Essex, Calculus - A Complete Course

Limit of a function - continued

$$\lim_{x \to a} f(x) = L,$$

if

$$\forall \varepsilon > 0$$

$$\exists \delta > 0$$

such that if

$$0 < |x - a| < \delta,$$

then

$$x \in \mathit{Dom}\, f \wedge |f(x) - L| < \varepsilon$$

Limit of a function - continued

First attempt at translation:

$$\begin{aligned} &lim\ a\ f\ L = \forall\ \epsilon > 0.\ \ \exists\ \delta > 0.\ \ P\ \epsilon\ \delta \\ &\quad \textbf{where}\ P\ \epsilon\ \delta = (0 < |x - a| < \delta) \Rightarrow \\ &\qquad\qquad\qquad\qquad (x \in Dom\ f \wedge |f\ x - L| < \epsilon) \end{aligned}$$

Limit of a function - continued

Finally (after adding a binding for x):

$$\begin{aligned} &lim\ a\ f\ L = \forall\ \epsilon > 0.\ \ \exists\ \delta > 0.\ \ P\ \epsilon\ \delta \\ &\quad \mathbf{where}\ P\ \epsilon\ \delta = \quad \forall\ x.\ \ Q\ \epsilon\ \delta\ x \\ &\qquad\qquad Q\ \epsilon\ \delta\ x = (0 < |x - a| < \delta) \Rightarrow \\ &\qquad\qquad\qquad\quad (x \in Dom\ f \wedge |f\ x - L| < \epsilon) \end{aligned}$$

Limit of a function - continued

Finally (after adding a binding for x):

$$
\begin{aligned}
&\mathit{lim}\ a\ f\ L = \forall\ \epsilon > 0.\ \ \exists\ \delta > 0.\ \ P\ \epsilon\ \delta \\
&\quad \textbf{where}\ P\ \epsilon\ \delta = \quad \forall\ x.\ \ Q\ \epsilon\ \delta\ x \\
&\qquad\qquad Q\ \epsilon\ \delta\ x = (0 < |x - a| < \delta) \Rightarrow \\
&\qquad\qquad\qquad\qquad (x \in \mathit{Dom}\ f \wedge |f\ x - L| < \epsilon)
\end{aligned}
$$

Lesson learned: be careful with scope and binding (of x in this case).

Limit of a function - continued

Finally (after adding a binding for x):

$$\begin{aligned}
&\mathit{lim}\ a\ f\ L = \forall\ \epsilon > 0.\ \ \exists\ \delta > 0.\ \ P\ \epsilon\ \delta \\
&\quad \textbf{where}\ P\ \epsilon\ \delta = \quad \forall\ x.\ \ Q\ \epsilon\ \delta\ x \\
&\qquad\qquad Q\ \epsilon\ \delta\ x = (0 < |x - a| < \delta) \Rightarrow \\
&\qquad\qquad\qquad\qquad (x \in \mathit{Dom}\ f\ \wedge |f\ x - L| < \epsilon)
\end{aligned}$$

Lesson learned: be careful with scope and binding (of x in this case).

[We will now assume limits exist and use *lim* as a function from a and f to L.]

Example 2: derivative

*The **derivative** of a function f is another function f' defined by*

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

*at all points x for which the limit exists (i.e., is a finite real number). If $f'(x)$ exists, we say that f is **differentiable** at x.*

We can write

$D\ f\ x = \mathit{lim}\ 0\ g$ **where** $g\ h = \frac{f\ (x+h) - f\ x}{h}$

Example 2: derivative

*The **derivative** of a function f is another function f' defined by*

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

*at all points x for which the limit exists (i.e., is a finite real number). If $f'(x)$ exists, we say that f is **differentiable** at x.*

We can write

$$D\ f\ x = \mathit{lim}\ 0\ g \quad \textbf{where} \quad g\ h = \frac{f\ (x+h) - f\ x}{h}$$

$$D\ f\ x = \mathit{lim}\ 0\ (\varphi\ x) \quad \textbf{where} \quad \varphi\ x\ h = \frac{f\ (x+h) - f\ x}{h}$$

Example 2: derivative

*The **derivative** of a function f is another function f' defined by*

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

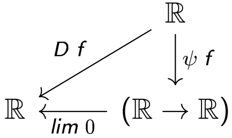
*at all points x for which the limit exists (i.e., is a finite real number). If $f'(x)$ exists, we say that f is **differentiable** at x.*

We can write

$$D\ f\ x = lim\ 0\ g \qquad \textbf{where} \qquad g\ h = \frac{f\ (x+h) - f\ x}{h}$$

$$D\ f\ x = lim\ 0\ (\varphi\ x)\ \textbf{where} \quad \varphi\ x\ h = \frac{f\ (x+h) - f\ x}{h}$$

$$D\ f \quad = lim\ 0 \circ \psi\ f\ \textbf{where}\ \psi\ f\ \ x\ h = \frac{f\ (x+h) - f\ x}{h}$$



Derivatives, cont.

Examples:

$$D : (\mathbb{R} \to \mathbb{R}) \to (\mathbb{R} \to \mathbb{R})$$

$$\begin{aligned}
sq\ x &= x \hat{}\ 2 \\
double\ x &= 2 * x \\
c_2\ x &= 2
\end{aligned}$$

$$sq' == D\ sq == D\ (\lambda x \to x \hat{}\ 2) == D\ (\hat{}\ 2) == (2*) == double$$

$$sq'' == D\ sq' == D\ double == c_2 == const\ 2$$

Note: we cannot *implement* D (of this type) in Haskell.
Given only $f : \mathbb{R} \to \mathbb{R}$ as a “black box” we cannot compute the actual derivative $f' : \mathbb{R} \to \mathbb{R}$.
We need the “source code” of f to apply rules from calculus.

Course material (chapters)

1. A DSL for arithmetic expressions and complex numbers
2. Logic and calculational proofs
3. Types in Mathematics
4. Compositional Semantics and Algebraic Structures
5. Polynomials and Power Series
6. Higher-order Derivatives and their Applications
7. Matrix algebra and linear transformations
8. Exponentials and Laplace

Course context

- Semi-compulsory course, spring of second year in CSE programme
- Students struggle with math-heavy courses in third year
- Students do well with (functional) programming
- Can a functional programming perspective help to clarify the mathematics?

CSE program

	Fall	Spring
Year 1	Compulsory courses	Compulsory courses
Year 2	Compulsory courses	DSLsofMath OR ConcProg
Year 3	TSS + Control	...

Course results

- 2016: 28 students, pass rate: 68%
- 2017: 43 students, pass rate: 58%
- 2018: 39 students, pass rate: 89%

Results in subsequent courses

	PASS	IN	OUT
TSS pass rate	77%	57%	36%
TSS mean grade	4.23	4.10	3.58
Control pass rate	68%	45%	40%
Control mean grade	3.91	3.88	3.35

Group sizes: PASS 34, IN 53, OUT 92 (145 in all)

Results in previous courses

	PASS	IN	OUT
Pass rate for first 3 semesters	97%	92%	86%
Mean grade for first 3 semesters	3.95	3.81	3.50
Math/physics pass rate	96%	91%	83%
Math/physics mean grade	4.01	3.84	3.55

Group sizes: PASS 34, IN 53, OUT 92 (145 in all)

Results

	PASS	IN	OUT
Pass rate for first 3 semesters	97%	92%	86%
Math/physics pass rate	96%	91%	83%
TSS pass rate	77%	57%	36%
Control pass rate	68%	45%	40%

Group sizes: PASS 34, IN 53, OUT 92 (145 in all)

Future work

- Working with earlier and later courses, can these ideas be useful in their curriculum?
- Better tool support in the course, proof systems?
- Polish the lecture notes into a book
- (Perhaps: more rigorous empirical evaluation of course efficacy)

Future work

- Working with earlier and later courses, can these ideas be useful in their curriculum?
- Better tool support in the course, proof systems?
- Polish the lecture notes into a book
- (Perhaps: more rigorous empirical evaluation of course efficacy)

Questions?

`https://github.com/DSLsofMath/DSLsofMath`

[Hint: There are bonus slides;-]

Example 3: Lagrangian

From [Sussman 2013, Functional Differential Geometry]:

> *A mechanical system is described by a Lagrangian function of the system state (time, coordinates, and velocities). A motion of the system is described by a path that gives the coordinates for each moment of time. A path is allowed if and only if it satisfies the Lagrange equations. Traditionally, the Lagrange equations are written*
>
> $$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$
>
> *What could this expression possibly mean?*

Lagrangian, cont.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

- The use of notation for "partial derivative", $\partial L / \partial q$, suggests that L is a function of at least a pair of arguments:

  $$L : \mathbb{R}^{\mathrm{i}} \rightarrow \mathbb{R}, i \geqslant 2$$

  This is consistent with the description: "Lagrangian function of the system state (time, coordinates, and velocities)". So, if we let "coordinates" be just one coordinate, we can take $i = 3$:

  $$L : \mathbb{R}^3 \rightarrow \mathbb{R}$$

  The "system state" here is a triple, of type $S = (T, Q, V)$, and we can call the the three components $t : T$ for time, $q : Q$ for coordinate, and $v : V$ for velocity. ($T = Q = V = \mathbb{R}$.)

Lagrangian, cont.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

- Looking again at $\partial L / \partial q$, q is the name of a variable, one of the 3 args to L. In the context, which we do not have, we would expect to find somewhere the definition of the Lagrangian as

$$\begin{aligned} L &: (T, Q, V) \to \mathbb{R} \\ L &\ (t, q, v) = \ldots \end{aligned}$$

- therefore, $\partial L / \partial q$ should also be a function of the same triple:

$$(\partial L \, / \, \partial q) : (T, Q, V) \to \mathbb{R}$$

It follows that the equation expresses a relation between *functions*, therefore the 0 on the right-hand side is *not* the real number 0, but rather the constant function *const* 0:

$$\begin{aligned} \mathit{const}\ 0 &: (T, Q, V) \to \mathbb{R} \\ \mathit{const}\ 0 &\ (t, q, v) = 0 \end{aligned}$$

Lagrangian, cont.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

- We now have a problem: $d \mathbin{/} dt$ can only be applied to functions of *one* real argument t, and the result is a function of one real argument:

  $$(d \mathbin{/} dt)\,(\partial L \mathbin{/} \partial \dot{q}) : T \to \mathbb{R}$$

  Since we subtract from this the function $\partial L/\partial q$, it follows that this, too, must be of type $T \to \mathbb{R}$. But we already typed it as $(T, Q, V) \to \mathbb{R}$, contradiction!
- The expression $\partial L/\partial \dot{q}$ appears to also be malformed. We would expect a variable name where we find $\dot{q}$, but $\dot{q}$ is the same as dq/dt, a function.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

- The only immediate candidate for an application of d/dt is "a path that gives the coordinates for each moment of time". Thus, the path is a function of time, let us say

  $w : T \to Q$ -- with T for time and Q for coords $(q : Q)$

  We can now guess that the use of the plural form "equations" might have something to do with the use of "coordinates". In an n-dim. space, a position is given by n coordinates. A path would then be

  $w : T \to Q$ -- with $Q = \mathbb{R}^n$

  which is equivalent to n functions of type $T \to \mathbb{R}$, each computing one coordinate as a function of time. We would then have an equation for each of them. We will use $n = 1$ for the rest of this example.

Lagrangian, cont.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

- Now that we have a path, the coordinates at any time are given by the path. And as the time derivative of a coordinate is a velocity, we can actually compute the trajectory of the full system state (T, Q, V) starting from just the path.

  $q : T \to Q$
  $q\ t = w\ t$ -- or, equivalently, $q = w$
  $\dot{q} : T \to V$
  $\dot{q}\ t = dw\ /\ dt$ -- or, equivalently, $\dot{q} = D\ w$

  We combine these in the "combinator" *expand*, given by

  $expand : (T \to Q) \to (T \to (T, Q, V))$
  $expand\ w\ t = (t, w\ t, D\ w\ t)$

Lagrangian, cont.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

- With *expand* in our toolbox we can fix the typing problem.

$$(\partial L \,/\, \partial q) \circ (expand\ w) : T \to \mathbb{R}$$

- We now move to using D for $d \,/\, dt$, D_2 for $\partial \,/\, \partial q$, and D_3 for $\partial \,/\, \partial \dot{q}$. In combination with *expand w* we find these type correct combinations for the two terms in the equation:

$$D\ ((D_2\ L) \circ (expand\ w)) : T \to \mathbb{R}$$
$$(D_3\ L) \circ (expand\ w) : T \to \mathbb{R}$$

The equation becomes

$$D\ ((D_3\ L) \circ (expand\ w)) - (D_2\ L) \circ (expand\ w) = const\ 0$$

or, after simplification:

$$D\ (D_3\ L \circ expand\ w) = D_2\ L \circ expand\ w$$

Example 3: Lagrangian, summary

"A path is allowed if and only if it satisfies the Lagrange equations" means that this equation is a predicate on paths:

$$Lagrange\ (L, w) = \qquad D\ (D_3\ L \circ expand\ w) == D_2\ L \circ expand\ w$$

Thus: If we can describe a mechanical system in terms of "a Lagrangian" ($L : S \to \mathbb{R}$), then we can use the predicate to check if a particular candidate path $w : T \to \mathbb{R}$ qualifies as a "motion of the system" or not. The unknown of the equation is the path w, and the equation is an example of a partial differential equation (a PDE).