A nominal syntax for internal parametricity

Thorsten Altenkirch, Ambrus Kaposi

University of Nottingham

TYPES Meeting, Tallinn
18-21 May 2015

Parametricity and univalence

- Parametricity says that terms respect logical relations.
  - Two functions are related if they map related inputs to related outputs.
  - Two pairs are related if they are componentwise related.
  - Two types are related if there is a relation between them.
- In homotopy type theory, terms respect equality.
  - Two functions are equal if they map equal inputs to equal outputs (function extensionality).
  - Two pairs are equal if they are componentwise equal.
  - Two types are equal if there is a relation between them and this relation is the graph of an equivalence (univalence).
- Our goal is to replace intensional equality in type theory by one which is defined recursively over the type structure.
- Inspiration:
  - Bernardy-Moulin: Internal parametricity, 2012
  - Bezem-Coquand-Huber: The cubical sets model of type theory, 2013
  - Altenkirch-McBride-Swierstra: Observational type theory, 2007
  - Martin-Löf: An intuitionistic theory of types, 1972

Parametricity

- Parametricity says that terms respect logical relations.

$$A : \mathsf{U}, u : A, s : A \to A \vdash t : A$$

$$\rho_0 \equiv (A \mapsto \mathbb{N}, u \mapsto \mathsf{zero}, s \mapsto \mathsf{suc})$$
$$\rho_1 \equiv (A \mapsto \mathsf{Bool}, u \mapsto \mathsf{true}, s \mapsto \mathsf{not})$$

- If ρ_0 and ρ_1 are related, then $t[\rho_0]$ is related to $t[\rho_1]$.
- It can't happen that $t[\rho_0] \equiv \mathsf{suc}\,(\mathsf{suc}\,\mathsf{zero})$ and $t[\rho_1] \equiv \mathsf{false}$.

$$\sim_{\mathsf{A}} : \mathbb{N} \to \mathsf{Bool} \to \mathsf{U}$$
$$x \sim_A b :\equiv (x \text{ even}) \text{ is } b$$

- A simpler example:

$$A : \mathsf{U}, x : A \vdash t : A$$

Specifying a logical relation

- The logical relation for a type A:

$$\frac{\Gamma \vdash A : \mathsf{U}}{\Gamma^{=} \vdash \sim_A : A[0] \to A[1] \to \mathsf{U}} \qquad \frac{\Gamma \vdash}{\Gamma^{=} \vdash} \qquad 0_\Gamma, 1_\Gamma : \Gamma^{=} \Rightarrow \Gamma$$

- The context of related elements:

$$\begin{aligned} \cdot^{=} &\equiv \cdot \\ (\Gamma, x : A)^{=} &\equiv \Gamma^{=}, x_0 : A[0], x_1 : A[1], x_2 : x_0 \sim_A x_1 \end{aligned}$$

- Substitutions 0, 1 project out the corresponding components:

$$\begin{aligned} i_{\cdot} &\equiv () &&: \cdot \Rightarrow \emptyset \\ i_{\Gamma, x:A} &\equiv (i_\Gamma, x \mapsto x_i) &&: (\Gamma, x : A)^{=} \Rightarrow \Gamma, x : A \end{aligned}$$

Defining the logical relation

$$\frac{\Gamma \vdash A : \mathsf{U}}{\Gamma^{=} \vdash \sim_A : A[0] \to A[1] \to \mathsf{U}}$$

$$f_0 \sim_{\Pi(x:A).B} f_1 \equiv \Pi(x_0 : A[0], x_1 : A[1], x_2 : x_0 \sim_A x_1).f_0\, x_0 \sim_B f_1\, x_1$$

$$(a, b) \sim_{\Sigma(x:A).B} (a', b') \equiv \Sigma(x_2 : a \sim_A a').b \sim_B [x_0 \mapsto a, x_1 \mapsto a']\, b'$$

$$A \sim_{\mathsf{U}} B \equiv A \to B \to \mathsf{U} \text{ (parametricity)}$$

$$A \sim_{\mathsf{U}} B \equiv A \simeq B \text{ (later)}$$

Parametricity

$$\frac{\Gamma \vdash t : A}{\Gamma^= \vdash t^= : t[0] \sim_A t[1]}$$

This can be proven by induction on the term t:

$$\begin{aligned}
(f\ u)^= &\equiv f^=\ u[0]\ u[1]\ u^= \\
(\lambda x.t)^= &\equiv \lambda x_0, x_1, x_2\ .\ t^= \\
x^= &\equiv x_2 \\
\mathsf{U}^= &\equiv \sim_{\mathsf{U}} \\
(t[\rho])^= &\equiv t^=[\rho^=] \qquad\qquad (\rho^= \text{ is pointwise})
\end{aligned}$$

The previous example:

$$(A : \mathsf{U}, u : A, s : A \to A)^= \vdash t^= : t[0] \sim_A t[1]$$

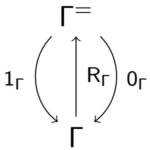
Internalisation

- The type of param is not well-formed:

$$\cdot \vdash \mathsf{param} : \Pi(A : \mathsf{U}, t : A).t \sim_A t$$

- We need a substitution from $(A : \mathsf{U}, t : A)$ to $(A : \mathsf{U}, t : A)^=$.
- We define $\mathsf{R}_\Gamma : \Gamma \Rightarrow \Gamma^=$:

$$\mathsf{R}_\cdot \equiv ()$$
$$\mathsf{R}_{\Gamma.x:A} \equiv (\mathsf{R}_\Gamma, x, x, \mathsf{refl}\, x)$$



$$\frac{\Gamma \vdash a : A}{\Gamma \vdash \mathsf{refl}\, a \equiv (a^=)[\mathsf{R}_\Gamma] : a \sim_A [\mathsf{R}_\Gamma]\, a}$$

- Now we can define param:

$$\cdot \vdash \lambda A, t.\mathsf{refl}\, t : \Pi(A : \mathsf{U}, t : A).t \sim_A [\mathsf{R}_{A:U}]\, t$$

- Before adding refl we didn't extend the theory. Now $(\mathsf{refl}\, x)$ is a new normal form if x is a variable.

Parametricity of parametricity

- Given $(\mathsf{refl}\,x)$ as a new normal form, we need to say what $(\mathsf{refl}\,x)^{=}$ is, i.e. $\mathsf{refl}^{=}\,x_0\,x_1\,x_2$.
- We could add a new term former $\mathsf{refl}^{=}$:

$$\frac{\Gamma \vdash A : \mathsf{U} \qquad \Gamma^{=} \vdash a_0 : A[0] \quad \Gamma^{=} \vdash a_1 : A[1] \quad \Gamma^{=} \vdash a_2 : a_0 \sim_A a_1}{\Gamma^{=} \vdash \mathsf{refl}^{=}\,a_0\,a_1\,a_2 : A^{==}[\mathsf{R}^{=}_{\Gamma}]\,a_0\,a_1\,a_2\,a_0\,a_1\,a_2\,(\mathsf{refl}\,a_0)\,(\mathsf{refl}\,a_1)}$$

$$\Gamma^{=} \vdash \mathsf{refl}\,a_2 : A^{==}[\mathsf{R}_{\Gamma^{=}}]\,a_0\,a_0\,(\mathsf{refl}\,a_0)\,a_1\,a_1\,(\mathsf{refl}\,a_1)\,a_2\,a_2$$

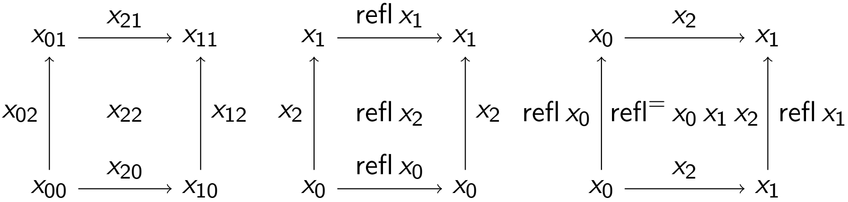
- refl and $\mathsf{refl}^{=}$ ways correspond to the two ways of degenerating a line into a square.
- We have $\mathsf{refl}\,a \equiv a^{=}[\mathsf{R}]$.
- We don't know how to compute $\mathsf{refl}^{=}\,a_0\,a_1\,a_2$.

Higher dimensions

$(x : A)^{==}$ can be viewed as a context of squares:

$$
\begin{aligned}
& (x : A)^{==} \\
\equiv & (x_0 : A[0] \quad .x_1 : A[1] \quad .x_2 : x_0 \sim_A x_1)^{=} \\
\equiv & \; x_{00} : A[00] \quad .x_{01} : A[01] \quad .x_{02} : x_{00} \sim_{A[0]} x_{01} \\
& .x_{10} : A[10] \quad .x_{11} : A[11] \quad .x_{12} : x_{10} \sim_{A[1]} x_{11} \\
& .x_{20} : x_{00} \sim_A [0]\, x_{10} \, .x_{21} : x_{01} \sim_A [1]\, x_{11} \, .x_{22} : x_{20} \sim_{x_0 \sim_A x_1} x_{21}
\end{aligned}
$$

This can be drawn as: The result of $\mathrm{R}_{(x:A)^=}$ and $(\mathrm{R}_{x:A})^=$:



Adding dimension names

First we add more information to the contexts: for each usage of $-^{=}$ we will use a new dimension name, so $(x : A)^{==}$ will become $(x : A)^{ij}$:

$$(\Gamma, x : A)^i \equiv \Gamma^i, x_{i0} : A[0_i], x_{i1} : A[1_i], x_{i2} : A^i\, x_{i0}\, x_{i1}$$

This refines the types of $\mathsf{R}_{(x:A)^=}$ and $\mathsf{R}^{=}_{x:A}$:

$$\begin{aligned} \mathsf{R}_{i(x:A)^j} \quad &: (x : A)^j \Rightarrow (x : A)^{ji} \\ (\mathsf{R}_{i(x:A)})^j &: (x : A)^j \Rightarrow (x : A)^{ij} \end{aligned}$$

Their targets are different contexts, with dimensions swapped.

A definitional quotient

We would like to equate $(x : A)^{ij}$ and $(x : A)^{ji}$. These have different variable names, but they contain the same information: eg. x_{i1j2} corresponds to x_{j2i1}. We add the following quotients:

$$\Gamma^{ij} \equiv \Gamma^{ji} \qquad \frac{\rho : \Delta \Rightarrow \Gamma}{\rho^{ij} \equiv \rho^{ji} : \Delta^{ij} \Rightarrow \Gamma^{ij}} \qquad \frac{\Gamma \vdash t : A}{\Gamma^{ij} \vdash t^{ij} \equiv t^{ji} : A^{ij}\{x \mapsto t\}_{ij}}$$

$\{x \mapsto t\}_{ij}$ is $(x \mapsto t)^{ij}$ omitting the last element.
Every term former needs to be symmetric, eg. Π types need to know which argument corresponds to which index, because we need

$$(\Pi(x : A).B)^{ij} \equiv (\Pi(x : A).B)^{ji}.$$

i.e.

$$\Pi(x_{i0j0} : A[0_i 0_j], x_{i0j1} : A[0_i 1_j], x_{i0j1} : A[0_i]^j\, x_{i0j0}\, x_{i0j1}, \dots).B^{ij} \dots$$
$$\equiv \Pi(x_{i0j0} : A[0_i 0_j], x_{i0j1} : A[0_i 1_j], \dots, x_{j1i0} : (A^j[0_i]\, x_{j0i0}\, x_{j1i0}).B^{ji} \dots.$$

New rules for Π types

New rules for full Π types and relations (I is a set of dimension names):

$$\frac{\xi : \Gamma \Rightarrow (X : \mathsf{U})^I \quad \Gamma.(x : X)^I[\xi] \vdash B : \mathsf{U}}{\Gamma \vdash \Pi(x : X)^I[\xi].B : \mathsf{U}} \qquad \frac{\xi : \Gamma \Rightarrow \{X : \mathsf{U}\}_I}{\Gamma \vdash \Pi\{x : X\}_I[\xi].\mathsf{U} : \mathsf{U}}$$

$$\frac{\Gamma \mathbin{+\!\!+} \{\!\!\{x : X\}\!\!\}_I[\xi] \vdash t : B}{\Gamma \vdash \lambda\{\!\!\{x : X\}\!\!\}_I[\xi].t : \Pi\{\!\!\{x : X\}\!\!\}_I[\xi].B}$$

$$\frac{\Gamma \vdash f : \Pi\{\!\!\{x : X\}\!\!\}_I[\xi].B \quad \Gamma \vdash \omega : \{\!\!\{y : X\}\!\!\}_I[\xi]}{\Gamma \vdash \mathsf{app}_I(f, \omega) : B[\mathsf{id} \mathbin{+\!\!+} \omega]}$$

The computation rule of $\mathsf{refl}^=$

We generalise the rule for parametricity from adding one dimension to adding a set of dimensions at a time. The parametricity rule becomes:

$$\frac{\Gamma \vdash t : A}{\Gamma^I \vdash t^I : \mathsf{app}_I(A^I, \{x \mapsto t\}_I)}$$

The order of dimensions does not matter.
Lifting of the universe becomes:

$$\mathsf{U}^I \equiv \lambda\{X : \mathsf{U}\}_I.\Pi\{x : X\}_I.\mathsf{U}$$

Now we have $(\mathsf{refl}_i\ a)^j \equiv (a^i[\mathsf{R}_{i\Gamma}])^j \equiv a^{ij}[\mathsf{R}_{i\Gamma^j}] \equiv a^{ji}[\mathsf{R}_{i\Gamma^j}] \equiv \mathsf{refl}_i\ a^j$, so the problem with the computation rule of $\mathsf{refl}^=$ disappears.

Operational semantics

To rigorously define the theory, we need telescope contexts and substitutions.
We defined a call-by-name operational semantics for this theory where the weak-head normal forms are:

t, A	$::= \dots$	terms
ν	$::= () \mid (\nu, x \mapsto g) \mid (\nu, x \mapsto t[\nu])$	environments
v	$::= \mathsf{U} \mid (\Pi \{\!\!\{ x : X \}\!\!\} \imath[\rho].A)[\nu] \mid (\lambda \{\!\!\{ x \}\!\!\} \imath.t)[\nu] \mid n$	values (whnfs)
n	$::= g \mid \mathsf{app}\imath(n, \nu)$	neutral values
g	$::= x \mid \mathsf{refl}_i\, g$	generic values

Capturing univalence

Now we would like to replace the definition of U^i to be a relation *which is a graph of an equivalence*.
We use the following notion of equivalence:

$$\begin{aligned}(A \sim_{\mathsf{U}^i} B) = \Sigma(\sim : A \to B \to U)&\\ .\Pi(x : A).\mathrm{isContr}(\Sigma(y : B).x \sim y)&\\ \times \Pi(y : B).\mathrm{isContr}(\Sigma(x : A).x \sim y)&.\end{aligned}$$

The new rule for parametricity expresses that terms respect (cubical) equality (we need to project out the relation):

$$\frac{\Gamma \vdash t : A}{\Gamma^I \vdash t^I : \mathrm{app}_I(\sim_{A^I}, \{x \mapsto t\}_I)} \qquad \frac{\Gamma \vdash A : \mathsf{U}}{\Gamma^i \vdash A^i : A[0_i] \sim_{\mathsf{U}^i} A[1_i]}$$

However we don't have $\mathsf{U}^{ij} \equiv \mathsf{U}^{ji}$, so we need to refine this definition.

Conclusions

- We defined a theory for internal parametricity with an operational semantics.
- We still need to prove termination and completeness of the semantics.
- It's not clear which definition of equivalence to use to capture univalence in a symmetric way.
- The final goal is a type theory with internal parametricity and univalence where some dimensions are parametricity dimensions and some dimensions are equalities (1-to-1 relations, having Kan fillers).

Thank you for your attention!