Proof-relevant parametricity*

Neil Ghani, Fredrik Nordvall Forsberg, and Federico Orsanigo

University of Strathclyde, UK
{neil.ghani, fredrik.nordvall-forsberg, federico.orsanigo}@strath.ac.uk

Abstract. Parametricity is one of the foundational principles which underpin our understanding of modern programming languages. Roughly speaking, parametricity expresses the hidden invariants that programs satisfy by formalising the intuition that programs map related inputs to related outputs. Traditionally parametricity is formulated with proof-irrelevant relations but programming in Type Theory requires an extension to proof-relevant relations. But then one might ask: can our proofs that polymorphic functions are parametric be parametric themselves? This paper shows how this can be done and, excitingly, our answer requires a trip into the world of higher dimensional parametricity.

1 Introduction

According to Strachey [2000], a polymorphic program is *parametric* if it applies the same uniform algorithm at all instantiations of its type parameters. Reynolds [1983] proposed *relational parametricity* as a mathematical model of parametric polymorphism. Phil Wadler, with his characteristic ability to turn deep mathematical insight into practical gains for programmers, showed how Reynolds' relational parametricity has strong consequences [Wadler, 1989, 2007]: it implies equivalences of different encodings of type constructors, abstraction properties for datatypes, and famously, it allows properties of programs to be derived "for free" purely from their types.

Within relational parametricity, types containing free type variables map not only sets to sets, but also relations to relations. A relation R between sets A and B is a subset $R \subseteq A \times B$. We call these proof-irrelevant relations as, given $a \in A$ and $b \in B$, the only information R conveys is whether a is related to b and not, for example, *how* a is related to b. However, the development of dependently type programming languages, constructive logics and proof assistants means such relations are insufficient in a number of settings. For example, it is often natural to consider a relation R between sets A and B to be a function $R : A \times B \to \mathsf{Set}$, where we think of $R(a, b)$ as the set of proofs that R relates a and b. Such *proof-relevant* relations are needed if one wants to work in the pure Calculus of Constructions [Coquand and Huet, 1988] without assuming additional axioms

* This work is partially supported by SICSA, and EPSRC grants EP/K023837/1 and EP/M016951/1.

(in contrast, Atkey [2012] formalised (proof-irrelevant) relational parametricity in Coq, an implementation of the Calculus of Constructions, by assuming the axiom of Propositional Extensionality). This paper asks the fundamental question:

> *Does the relational model of parametric polymorphism extend from proof-irrelevant relations to proof-relevant relations?*

At first sight, one might hope for a straightforward answer. Many properties in the proof-irrelevant world have clear analogues as proof-relevant constructions. Indeed, as we shall see, this approach gives a satisfactory treatment of the function space. However, universally quantified types pose a much more significant challenge; it is insufficient to simply take the uniformity condition inherent within a proof-irrelevant parametric polymorphic function and replace it with a function acting on proofs; this causes the Identity Extension Lemma to fail. Instead, to prove this lemma in a proof-relevant setting, we need to strengthen the uniformity condition on parametric polymorphic functions by requiring it to itself be parametric.

Proof-relevant parametricity thus entails adding a second layer of parametricity to ensure that the proofs that functions are parametric are themselves parametric. This takes us into the world of 2-dimensional parametricity where type constructors now act upon sets, relations and 2-relations. But, there are actually surprisingly many choices as to what a 2-relation is! Further, at higher dimensions, there are a number of potential equality relations and it is not *a priori* clear which of these need to be preserved and which do not. Relations are naturally organised in a cubical or simplicial manner, and so this will not surprise those familiar with simplicial and cubical methods, where there is an analogous choice of which face maps and degeneracies to consider. For example, do connections [Brown et al., 2011] have a role to play in proof-relevant parametricity? These questions are not at all obvious — we went down many false routes before finding the right answer.

The paper is structured as follows: in Section 2, we introduce the preliminaries we need, while Sections 3 and 4 introduce proof-relevant relations and 2-relations. Section 5 constructs a 2-dimensional model of System F, and proves it correct by establishing 2-dimensional analogues of the Identity Extension Lemma and the Abstraction Theorem. We present a proof-of-concept application in Section 6, where we generalise the usual proof that parametricity implies naturality to the 2-dimensional setting. Section 7 concludes, with plans for future work including higher-dimensional logical relations, and the relationship with the cubical sets model of HoTT.

2 Impredicative Type Theory and the Identity Type

In order to make proof-relevant relations precise, we work in the constructive framework of impredicative Martin-Löf Type Theory [Coquand and Huet, 1988; Martin-Löf, 1972]. Impredicativity allows us to quantify over all types of sort

Type in order to construct a new object of sort Type.[1] Following Atkey [2012], we will use impredicative quantification in the meta-theory to interpret impredicative quantification in the object theory. This simplifies the presentation, and allows us to focus on the proof-relevant aspects of the logical relations.

Apart from impredicativity, the type theory we employ is standard; we make use of dependent function types $(\Pi x : A)B(x)$ and dependent pair types $(\Sigma x : A)B(x)$ with the usual introduction and elimination rules. We write $A \to B$ for $(\Pi x : A)B$ and $A \times B$ for $(\Sigma x : A)B$ when B does not depend on $x : A$. Crucial for our development will be Martin-Löf's *identity type*, given by the following rules:

$$\frac{A : \mathsf{Type} \qquad a, b : A}{\mathsf{Id}_A(a, b) : \mathsf{Type}} \qquad \frac{a : A}{\mathsf{refl}(a) : \mathsf{Id}_A(a, a)}$$

$$\frac{C : (\Pi x, y : A)(\mathsf{Id}_A(x, y) \to \mathsf{Type}) \qquad d : (\Pi x : A)C(x, x, \mathsf{refl}(x))}{J(C, d) : (\Pi x, y : A)(\Pi p : \mathsf{Id}_A(x, y))C(x, y, p)}$$

In the language of the HoTT book [The Univalent Foundations Program, 2013], the elimination rule J is called *path induction*. We stress that we are *not* assuming Uniqueness of Identity Proofs, as that would in effect result in proof-irrelevance once again. In this paper, we will however restrict attention to types where identity proofs of identity proofs are unique, i.e. to types A where $\mathsf{Id}_{\mathsf{Id}_A(x,y)}(p, q)$ is trivial. Garner [2009] has investigated the semantics of Type Theory where all types are of this form. For our purposes, it is enough to work with a *subuniverse* of such types. To make this precise, define

$$\begin{aligned} \mathsf{isProp}(A) &:= (\Pi x, y : A)\mathsf{Id}_A(x, y) & \mathsf{Prop} &:= (\Sigma X : \mathsf{Type})(\mathsf{isProp}(X)) \\ \mathsf{isSet}(A) &:= (\Pi x, y : A)\mathsf{isProp}(\mathsf{Id}_A(x, y)) & \mathsf{Set} &:= (\Sigma X : \mathsf{Type})(\mathsf{isSet}(X)) \\ \mathsf{is\text{-}1\text{-}Type}(A) &:= (\Pi x, y : A)\mathsf{isSet}(\mathsf{Id}_A(x, y)) & \mathsf{1\text{-}Type} &:= (\Sigma X : \mathsf{Type})(\mathsf{is\text{-}1\text{-}Type}(X)) \end{aligned}$$

Here Prop is the subuniverse of *propositions*, i.e. types with at most one inhabitant up to identity, while Set is the subuniverse of *sets*, i.e. types whose identity types in turn are propositional. Finally, we are interested in the subuniverse $\mathsf{1\text{-}Type}$ of 1-types, i.e. types whose identity types are sets. Subuniverses of an impredicative universe are also impredicative. Furthermore, all three of Prop, Set and $\mathsf{1\text{-}Type}$ are closed under Π- and Σ-types. The witness that a type is in a subuniverse is itself a proposition, and so we will abuse notation and leave it implicit — if there is a proof, it is unique up to identity.

We denote by $a \equiv_A b$ the existence of a proof $p : \mathsf{Id}_A(a, b)$. We often leave out the subscript if it can be inferred from context. A function $f : A \to B$ is said to be an *equivalence* if it has a left and a right inverse, and if there exists an equivalence $A \to B$, we write $A \cong B$. If $P : A \to \mathsf{Prop}$, then we write $\{x : A \,|\, P(x)\}$ for $(\Sigma x : A)P(x)$. Since $P(x)$ is a proposition for each $x : A$, we have that

[1] In Coq, this feature can be turned on by means of the command line option `-impredicative-set`.

$\mathsf{Id}_{\{x:A \,|\, P(x)\}}((a,p),(b,q)) \cong \mathsf{Id}_A(a,b)$. For this reason, we will often leave the proof $p : P(a)$ implicit when talking about an element (a,p) of $\{x : A \,|\, P(x)\}$. We also suggestively write $a \in P$ for $P(a)$. The identity type has a rich structure. In order to introduce notation, we list some basic facts here, and refer to the HoTT book [The Univalent Foundations Program, 2013] for more information.

Lemma 1 (Structure on $\mathsf{Id}_A(a,b)$).

(i) *For any $p : \mathsf{Id}_A(a,b)$ there is $p^{-1} : \mathsf{Id}_A(b,a)$.*

(ii) *For any $p : \mathsf{Id}_A(a,b)$ and $q : \mathsf{Id}_A(b,c)$ there is $p \cdot q : \mathsf{Id}_A(a,c)$, and $\mathsf{refl}(a)$, $-^{-1}$ and $- \cdot -$ satisfy the laws of a (higher) groupoid.*

(iii) *All functions $f : A \to B$ are functorial in Id_A, i.e. there is a term $\mathsf{ap}(f) : \mathsf{Id}_A(a,b) \to \mathsf{Id}_B(f(a), f(b))$.*

(iv) *All type families respect Id_A, i.e. there is a function*

$$\mathsf{tr} : (P : A \to \mathsf{Type}) \to \mathsf{Id}_A(a,b) \to P(a) \to P(b) \ . \qquad \square$$

We frequently use the following characterisation of equality in Σ-types:

Lemma 2. $\mathsf{Id}_{(\Sigma x:A)B(x)}((x,y),(x',y')) \cong (\Sigma p : \mathsf{Id}_A(x,x'))\mathsf{Id}_{B(x')}(\mathsf{tr}(B,p,y),y')$.

For function types, the corresponding statement is not provable, so we rely on the following axiom:

Axiom 3 (Function extensionality). *The function*

$$\mathsf{happly} : \mathsf{Id}_{(\Pi x:A)B(x)}(f,g) \to (\Pi x : A)\mathsf{Id}_{B(x)}(f(x),g(x))$$

defined using J in the obvious way is an equivalence. In particular, we have an inverse

$$\mathsf{ext} : (\Pi x : A)\mathsf{Id}_{B(x)}(f(x),g(x)) \to \mathsf{Id}_{(\Pi x:A)B(x)}(f,g)$$

This axiom is justified by models of impredicative Type Theory in intuitionistic set theory. It also follows from Voevodsky's Univalence Axiom [Voevodsky, 2010], which we do not assume in this paper. We will use function extensionality in order to derive the Identity Extension Lemma for arrow types, as in e.g. Wadler [2007].

3 Proof-relevant relations

We now define proof-relevant relations:

Definition 4. *The collection of proof-relevant relations is denoted* Rel *and consists of triples* (A, B, R)*, where* $A, B : 1\text{-}\mathsf{Type}$ *and* $R : A \times B \to \mathsf{Set}$*. The* 1*–type of morphisms from* (A, B, R) *to* (A', B', R') *is*

$$(\Sigma f : A \to A')(\Sigma g : B \to B')(\Pi x : A, y : B) R(a, b) \to R'(fa, gb)$$

In the rest of this paper we take relation to mean proof-relevant relation. The above definition means morphisms between relations have a proof-relevant equality and, thus, showing morphisms are equal involves constructing explicit proofs to that effect. Indeed, the equality of morphisms is given by

$$\mathsf{Id}((f, g, p), (f', g', p')) \cong (\Sigma \phi : \mathsf{Id}(f, f'), \psi : \mathsf{Id}(g, g'))\ \mathsf{Id}(\mathsf{tr}(\phi, \psi)p, p')$$

However, since $R : A \times B \to \mathsf{Set}$ has codomain Set, while A and B are 1–types, the complexity of R compared to A and B has decreased. This means relations between proof-relevant relations are in fact proof-irrelevant (see Section 4).

Given a relation (A, B, R), we often denote A by R_0 and B by R_1, write $R : \mathsf{Rel}(R_0, R_1)$, or $R : R_0 \leftrightarrow R_1$, and call R a relation between A and B. Similarly, given a morphism (f, g, p), we denote f by p_0, g by p_1 and write $p : (R_0 \to R_1)(p_0, p_1)$. If $R : \mathsf{Rel}(R_0, R_1)$ and $P : \mathsf{Rel}(P_0, P_1)$, then we have $\mathbf{1} : \mathsf{Rel}(\mathbf{1}, \mathbf{1})$, $P \times R : \mathsf{Rel}(P_0 \times R_0, P_1 \times R_1)$ and $R \Rightarrow P : \mathsf{Rel}(R_0 \to P_0, R_1 \to P_1)$ defined by

$$\begin{aligned} \mathbf{1}(x, y) &:= \mathbf{1} \\ (R \times P)((x, y), (x', y')) &:= R(x, x') \times P(y, y') \\ (R \Rightarrow P)(f, g) &:= (\Pi x : R_0, y : R_1)(R(x, y) \to P(fx, gy)) \end{aligned}$$

Interpreting abstraction and application requires the following functions:

Lemma 5. *Let* $R : \mathsf{Rel}(A, B)$*,* $R' : \mathsf{Rel}(A', B')$*, and* $R'' : \mathsf{Rel}(A'', B'')$*. There is an equivalance* $\mathsf{abs} : (R \times R' \to R'') \to (R \to (R' \Rightarrow R''))$ *with inverse* $\mathsf{app} : (R \to (R' \Rightarrow R'')) \to (R \times R' \to R'')$*.* □

We will also make use of the *equality relation* $\mathsf{Eq}(A)$ for each 1-type A:

Definition 6. *Equality* $\mathsf{Eq} : 1\text{-}\mathsf{Type} \to \mathsf{Rel}$ *is defined by* $\mathsf{Eq}(A) = (A, A, \mathsf{Id}_A)$ *on objects and* $\mathsf{Eq}(f) = (f, f, \mathsf{ap}(f))$ *on morphisms.*

Proposition 7. Eq *is full and faithful in that* $(\mathsf{Eq}X \to \mathsf{Eq}Y) \cong X \to Y$*.*

Proof. By function extensionality and contractability of singletons, we have

$$\begin{aligned} (\mathsf{Eq}X \to \mathsf{Eq}Y) &= (\Sigma f : X \to Y)(\Sigma g : X \to Y)(\Pi x x')\mathsf{Id}_X(x, x') \to \mathsf{Id}_Y(fx, gx') \\ &\cong (\Sigma f : X \to Y)(\Sigma g : X \to Y)\mathsf{Id}_{X \to Y}(f, g) \\ &\cong (\Sigma f : X \to Y)\mathbf{1} \cong X \to Y \ . \end{aligned}$$ □

Similarly, the exponential of equality relations is an equality relation. Here, we abuse notation and use the same symbol for equivalence of types and isomorphisms of relations:

Proposition 8. *For all* $X, Y : 1\text{-}\mathsf{Type}$*, we have* $(\mathsf{Eq}X \Rightarrow \mathsf{Eq}Y) \cong \mathsf{Eq}(X \to Y)$.

Proof. By extensionality it is enough to show

$$((\Pi x, x' : X)\mathsf{Id}(x, x') \to \mathsf{Id}(fx, gx')) \cong (\Pi x : X)\mathsf{Id}(fx, gx)$$

for every $f, g : X \to Y$. Functions can easily be constructed in both directions and proved inverse using extensionality and path induction. □

4 Relations between relations

Intuitively, 2-relations should relate proofs of relatedness in proof-relevant relations. Although conceptually simple, formalising 2-relations is non-trivial as various choices arise. For instance, if R and R' are proof-relevant relations, one may consider 2-relations between them as being given by functions

$$Q : (\Pi a : R_0, a' : R'_0, b : R_1, b' : R'_1)\ (R(a, b) \times R'(a', b')) \to \mathsf{Prop}$$

with the intuition of $(p, p') \in Q(a, a', b, b')$ being that Q relates the proof p to the proof p'. However, the natural arrow type of such 2-relations does not preserve equality. The problem is that, while a is related to b, and a' is related to b', there is no relationship between a and a' and b and b'. Thus, we were led to the following definition which seems to originate with Grandis (see e.g. Grandis [2009]):

Definition 9. *A 2-relation consists of the following 1-types and proof-relevant relations between them*

$$\begin{array}{ccc} Q_{00} & \xleftrightarrow{Q_{r0}} & Q_{10} \\ {\scriptstyle Q_{0r}}\updownarrow & & \updownarrow{\scriptstyle Q_{1r}} \\ Q_{01} & \xleftrightarrow[Q_{r1}]{} & Q_{11} \end{array}$$

together with a predicate

$$Q : (\Pi a : Q_{00}, b : Q_{10}, c : Q_{01}, d : Q_{11})$$
$$Q_{r0}(a, b) \times Q_{0r}(a, c) \times Q_{r1}(c, d) \times Q_{1r}(b, d) \to \mathsf{Prop}$$

A morphism of 2-relations consists of 4 functions between each corresponding node, 4 maps of relations such that each is over the appropriate pair of morphisms of 1-types, and a predicate stating that proofs related in one 2-relation are mapped to proofs which are related in the other 2-relation.

Thus a 2-relation is a 9-tuple and, even worse, a morphism of 2-relations is a 27-tuple! This combinatorial complexity is enough to scupper any noble mathematical intentions. We therefore develop a more abstract treatment beginning with the indices in a 2-relation. This extends the notion of reflexive graphs [Robinson and Rosolini, 1994; O'Hearn and Tennent, 1995; Dunphy and Reddy, 2004] to a second level of 2-relations; this notion, in turn, is just the first few levels of the notion of a cubical set [Brown and Higgins, 1981].

Definition 10. *Let I_0 be the type with elements $\{00, 01, 10, 11\}$ of indices for 1-types, and I_1 the type with elements $\{0r, r0, 1r, r1\}$ of indices for proof-relevant relations. Define the source and target function $@ : I_1 \times \mathsf{Bool} \to I_0$ where $w@i$ replaces the occurrence of r in w by i. We write $w@i$ as wi.*

I_0**-types:** Next we develop algebra for the types contained in 2-relations.

Definition 11. *An I_0-type is a function $X : I_0 \to 1\text{-}\mathsf{Type}$. To increase legibility we write X_w for Xw. The collection of maps between two I_0-types is defined by*

$$X \to X' \coloneqq (\Pi w : I_0) X_w \to X'_w$$

We define the following operations on I_0-types:

$$\begin{aligned} \mathbf{1} &\coloneqq \lambda w.\mathbf{1} \\ X \times X' &\coloneqq \lambda w.X_w \times X'_w \\ X \Rightarrow X' &\coloneqq \lambda w.X_w \to X'_w \end{aligned}$$

If X is an I_0-type, define its elements $\mathsf{El}X = (\Pi w : I_0)X_w$. The natural extension of this action to morphisms $f : X \to X'$ is denoted $\mathsf{El}\, f : \mathsf{El}X \to \mathsf{El}X'$.

Note that elements deserve that name as $\mathsf{El}X \cong \mathbf{1} \to X$. The construction of elements preserves structure as the following lemma shows:

Lemma 12. *Let X and X' be I_0-types. Then*

$$\begin{aligned} \mathsf{El}\ \mathbf{1} &\cong \mathbf{1} \\ \mathsf{El}(X \times X') &\cong \mathsf{El}X \times \mathsf{El}X' \\ \mathsf{El}(X \Rightarrow X') &= (\Pi w : I_0)X_w \to X'_w \end{aligned}$$

□

Finally, we show how to interpret abstraction and application over I_0-types:

Lemma 13. *Let X, X' and X'' be I_0-types. The function*

$$\mathsf{abs} = \lambda w.\, \lambda f.\, \lambda x.\, \lambda x'.\, f(x, x') : (X \times X' \to X'') \to (X \to (X' \Rightarrow X''))$$

is an equivalence with inverse $\mathsf{app} = \lambda w.\, \lambda f.\, \lambda y.\, f\,(\pi_0 y)\,(\pi_1 y)$. □

I_1**-Relations:** Next we develop algebra for the relations contained in 2-relations.

Definition 14. *An I_1-relation is a pair (X, R) of an I_0-type X and a function $R : (\Pi w : I_1)\mathsf{Rel}(X_{w0}, X_{w1})$. The collection of maps between two I_1-relations is defined by*

$$(X, R) \to (X', R') :\equiv (\Sigma f : X \to X')(\Pi w : I_1)(R_w \to R'_w)(f_{w0}, f_{w1})$$

We define the following operations on I_1-relations:

$$\begin{aligned} \mathbf{1} &:\equiv (\mathbf{1}, \lambda w.\mathbf{1}) \\ (X, R) \times (X', R') &:\equiv (X \times X', \lambda w.R_w \times R'_w) \\ (X, R) \Rightarrow (X', R') &:\equiv (X \Rightarrow X', \lambda w.R_w \Rightarrow R'_w) \end{aligned}$$

If (X, R) is an I_1-relation, define its elements

$$\mathsf{El}(X, R) = (\Sigma x : \mathsf{El} X)(\Pi w : I_1)R_w(x_{w0}, x_{w1})$$

The natural extension of El to morphisms $(f, g) : (X, R) \to (X', R')$ is denoted $\mathsf{El}(f, g) : \mathsf{El}(X, R) \to \mathsf{El}(X', R')$.

Note that elements deserve that name as $\mathsf{El}(X, R) \cong \mathbf{1} \to (X, R)$. The construction of elements preserves structure as the following lemma shows:

Lemma 15. *Let (X, R) and (X', R') be I_1-relations. Then*

$$\begin{aligned} \mathsf{El}\ \mathbf{1} &\cong \mathbf{1} \\ \mathsf{El}((X, R) \times (X', R')) &\cong \mathsf{El}(X, R) \times \mathsf{El}(X', R') \\ \mathsf{El}((X, R) \Rightarrow (X', R')) &= (\Sigma f : \mathsf{El}(X \Rightarrow X'))(\Pi w : I_1)(R_w \Rightarrow R'_w)(f_{w0}, f_{w1}) \end{aligned}$$

□

Finally, we show how to interpret abstraction and application over I_0-types:

Lemma 16. *Let $(X, R), (X', R')$ and (X'', R'') be I_1-relations. There is an equivalence $\mathsf{abs} : ((X, R) \times (X', R') \to (X'', R'')) \to ((X, R) \to ((X', R') \Rightarrow (X'', R'')))$ with inverse*

$$\mathsf{app} : ((X, R) \to ((X', R') \Rightarrow (X'', R''))) \to ((X, R) \times (X', R') \to (X'', R''))$$

Proof. The proof is similar to the proof of Lemma 5, but rests crucially on the fact that $R \Rightarrow P : \mathsf{Rel}(R_0 \to P_0, R_1 \to P_1)$. □

2-Relations: Finally, we develop the same algebra for 2-relations.

Definition 17. *An 2-relation is a pair consisting of an I_1-relation (X, R) and a function $Q : \mathsf{El}(X, R) \to \mathsf{Prop}$. The collection of maps between two 2-relations is defined by*

$$((X, R), Q) \to ((X', R'), Q') \coloneqq (\Sigma(f, g) : (X, R) \to (X', R'))$$
$$(\Pi(x, p) : \mathsf{El}(X, R))p \in Q(x) \Rightarrow (\mathsf{El}\ g\ p) \in Q'(\mathsf{El}\ f\ x)$$

We define the following operations on 2-relations

$$\begin{aligned} \mathbf{1} &= (\mathbf{1}, \lambda\_.\mathbf{1}) \\ ((X, R), Q) \times ((X', R)', Q') &= ((X, R) \times (X', R'), \\ &\qquad \lambda(x, y)\lambda(p, q).p \in Q(x) \wedge q \in Q'(y)) \\ ((X, R), Q) \Rightarrow ((X', R'), Q') &= ((X, R) \Rightarrow (X', R'), \\ &\qquad \lambda(f, g).(\Pi(x, p) : \mathsf{El}(X, R))p \in Q(x) \Rightarrow (\mathsf{El}\ g\ p) \in Q'(\mathsf{El}\ f\ x)) \end{aligned}$$

Lemma 18. *Let $((X, R), Q), ((X', R'), Q')$ and $((X'', R''), Q'')$ be 2-relations. There is an equivalence*

$$\mathsf{abs} : (((X, R), Q) \times ((X', R'), Q') \to ((X'', R''), Q'')) \cong$$
$$(((X, R), Q) \to (((X', R'), Q') \Rightarrow ((X'', R''), Q'')))$$

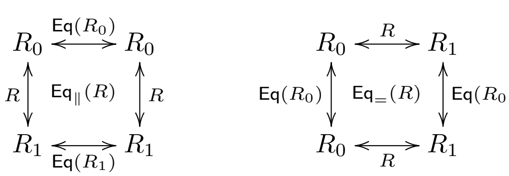
with inverse app.

Proof. Note that if X, X' and X'' are I_0-types, and if $f : X \times X' \to X''$ and $\mathsf{abs} f : X \to (X' \Rightarrow X'')$, then for any $x : \mathsf{El}X, x' : \mathsf{El}X'$, and $w : I_0$

$$(\mathsf{El} f)\ (x, x')\ w \equiv (\mathsf{El}\ (\mathsf{abs} f)\ x\ w)(x'\ w)$$

Similar results hold for app and for the analogous lemmas for I_1-sets. This, together with Lemma 16, extensionality and direct calculation gives the result. □

As in cubical and simplicial settings, there is more than one "degenerate" relation in higher dimensional relations. For example, we can duplicate a relation vertically or horizontally giving two functors $\mathsf{Eq}_{||}, \mathsf{Eq}_{=} : \mathsf{Rel} \to 2\mathsf{Rel}$ sending a relation R to the 2-relation indexed, repectively, by



where $(p, q, p', q') \in \mathsf{Eq}_{||}(R)(a, b, c, d)$ if and only if $\mathsf{tr}(p, p')q \equiv_{R(b,d)} q'$, while $(p, q, p', q') \in \mathsf{Eq}_{=}(R)(a, b, c, d)$ if and only if $\mathsf{tr}(q, q')p \equiv_{R(c,d)} p'$. Note that both the compositions $\mathsf{Eq}_{||} \circ \mathsf{Eq}$ and $\mathsf{Eq}_{=} \circ \mathsf{Eq}$ define the same functor which we denote

Eq_2. Another degeneracy, called a connection [Brown et al., 2011], is defined by a functor $\mathbf{C}\colon \mathsf{Rel} \to \mathsf{2Rel}$ which maps a relation R to the 2-relation indexed by

$$\begin{array}{ccc} R_0 & \xleftrightarrow{\mathsf{Eq}(R_0)} & R_0 \\ {\scriptstyle \mathsf{Eq}(R_0)}\updownarrow & \mathbf{C}R & \updownarrow{\scriptstyle R} \\ R_0 & \xleftrightarrow[R]{} & R_1 \end{array}$$

and with $(p, q, p', q') \in \mathbf{C}(R)(a, b, c, d)$ if and only if $\mathsf{tr}(q^{-1} \cdot p)p' \equiv_{R(b,d)} q'$ (there is of course also a symmetric version which swaps the role of $\mathsf{Eq}(R_0)$ and R, but we will not make us of this in the current paper). Again $\mathbf{C} \circ \mathsf{Eq}$ gives Eq_2.

Proposition 19. *The functor* $\mathsf{Eq}_{||}$ *is full and faithful.*

Proof. Similar to the proof of Proposition 7. □

Again, we can prove that exponentiation preserves all the degeneracies and the connection:

Proposition 20. *For all* $R, R' : \mathsf{Rel}$*, we have*

(i) an equivalence $\mathsf{Eq}_{||}R \Rightarrow \mathsf{Eq}_{||}R' \cong \mathsf{Eq}_{||}(R \to R')$

(ii) an equivalence $\mathsf{Eq}_{=}R \Rightarrow \mathsf{Eq}_{=}R' \cong \mathsf{Eq}_{=}(R \to R')$

(iii) an equivalence $\mathbf{C}R \Rightarrow \mathbf{C}R' \cong \mathbf{C}(R \to R')$*.* □

5 Proof-relevant two-dimensional parametricity

We now have the structure needed to define a 2-dimensional, proof-relevant model of System F. We recall the rules of System F in Fig. 1. Each type judgement $\Gamma \vdash T$ type, with $|\Gamma| = n$, will be interpreted in the semantics as

$$\begin{aligned} [\![T]\!]_0 &: |\mathsf{1\text{-}Type}|^n \to \mathsf{1\text{-}Type} \\ [\![T]\!]_1 &: |\mathsf{Rel}|^n \to \mathsf{Rel} \\ [\![T]\!]_2 &: |\mathsf{2Rel}|^n \to \mathsf{2Rel} \end{aligned}$$

by induction on type judgements with $[\![T]\!]_1$ over $[\![T]\!]_0 \times [\![T]\!]_0$, and $[\![T]\!]_2$ over $[\![T]\!]_1 \times [\![T]\!]_1 \times [\![T]\!]_1 \times [\![T]\!]_1$. This is similar to our previous work on bifibrational functorial models of (proof-irrelevant) parametricity [Ghani et al., 2015a,b], but with an additional 2-relational level.

Type formation rules:

$$\frac{}{\Gamma \vdash X \text{ type}}\ (X \in \Gamma) \qquad \frac{\Gamma \vdash A \text{ type} \qquad \Gamma \vdash B \text{ type}}{\Gamma \vdash A \to B \text{ type}} \qquad \frac{\Gamma, X \vdash A \text{ type}}{\Gamma \vdash \forall X.A \text{ type}}$$

Term typing rules:

$$\frac{}{\Gamma; \Delta \vdash x : A}\ (x : A \in \Delta) \quad \frac{\Gamma; \Delta, x : A \vdash t : B}{\Gamma; \Delta \vdash \lambda x.\, t : A \to B} \quad \frac{\Gamma; \Delta \vdash s : A \to B \qquad \Gamma; \Delta \vdash t : A}{\Gamma; \Delta \vdash s\, t : B}$$

$$\frac{\Gamma, X; \Delta \vdash t : A}{\Gamma; \Delta \vdash \Lambda X.t : \forall X.A}\ (X \notin FV(\Delta)) \qquad \frac{\Gamma; \Delta \vdash t : \forall X.A \qquad \Gamma; \Delta \vdash B \text{ type}}{\Gamma; \Delta \vdash t[B] : A[X \mapsto B]}$$

Judgemental equality:

$$\frac{}{\Gamma; \Delta \vdash (\lambda x.\, t)u = t[x \mapsto u] : B} \qquad \frac{}{\Gamma; \Delta \vdash t = \lambda x.\, tx : A \to B}\ (x : A \notin \Delta)$$

$$\frac{}{\Gamma; \Delta \vdash (\Lambda X.t)[B] = t[X \mapsto B] : A[X \mapsto B]} \qquad \frac{}{\Gamma; \Delta \vdash t = \Lambda X.t[X] : \forall X.A}\ (X \notin \Gamma)$$

Fig. 1: Typing rules for System F

5.1 Interpretation of types

The full interpretation can be found in Fig. 2. For type variables and arrow types, we just use projections and exponentials at each level. Elements of $[\![\forall X.T]\!]_0 \bar{A}$ consist of an ad-hoc polymorphic function f_0, a proof f_1 that f_0 is suitably uniform, and finally a (unique) proof (A0) that also the proof f_1 is parametric. Similarly, elements of $([\![\forall X.T]\!]_1 \bar{R})(f, g)$ are proofs ϕ that are suitably parametric in relation to f and g, both with respect to equalities (conditions A1.1 and A1.2) and connections (condition A1.3). We have not included uniformity also with respect to the "symmetric" connection since it is not needed for our applications, and we wish to keep the logical relation minimal.

Using Lemma 2 and function extensionality, we can characterise equality in the interpretation of $\forall$-types in the following way (note that $\mathsf{Id}_{([\![\forall X.T]\!]_2 \bar{Q})\vec{f}}(\vec{\phi}, \vec{\psi})$ is trivial by assumption, since $([\![\forall X.T]\!]_2 \bar{Q})\vec{f}$ is a proposition):

Lemma 21. *For all* $f, g : [\![\forall X.T]\!]_0 \bar{A}$*,*

$$\begin{aligned} \mathsf{Id}_{[\![\forall X.T]\!]_0 \bar{A}}(f, g) \cong \{\, \tau : (\Pi A : \mathit{1}\text{-}\mathsf{Type})\mathsf{Id}_{[\![T]\!]_0(\bar{A}, A)}(f_0 A, g_0 A) \mid \\ (\forall R : \mathsf{Rel}) \quad (f_1 R, \tau R_0, g_1 R, \tau R_1) \in \\ \mathsf{Eq}_{=}([\![T]\!]_1(\mathsf{Eq}(\bar{A}), R))(f_0 R_0, f_0 R_1, g_0 R_0, g_0 R_1)\} \end{aligned}$$

This can be used to prove a generalised version of the Identity Extension Lemma:

$$[\![X_0,\ldots,X_n \vdash X_k \text{ type}]\!]_i\vec{Y} = Y_k \qquad [\![S \to T]\!]_i\vec{Y} = [\![S]\!]_i\vec{Y} \Rightarrow_i [\![T]\!]_i\vec{Y}$$

$$\begin{aligned}
[\![\forall X.T]\!]_0\bar{A} = \{\, & f_0 : (\Pi A : \text{1-Type})[\![T]\!]_0(\bar{A}, A), \\
& f_1 : (\Pi R : \mathsf{Rel})[\![T]\!]_1(\mathbf{Eq}(\bar{A}), R)(f_0R_0, f_0R_1) \mid \\
& (\forall Q \colon \mathsf{2Rel})\ (f_1Q_{r0}, f_1Q_{0r}, f_1Q_{r1}, f_1Q_{1r}) \in \\
& [\![T]\!]_2(\mathsf{Eq}_2(\bar{A}), Q)(f_0Q_{00}, f_0Q_{10}, f_0Q_{01}, f_0Q_{11})\} \qquad \text{(A0)}
\end{aligned}$$

$$\begin{aligned}
([\![\forall X.T]\!]_1\bar{R})((f_0,f_1),(g_0,g_1)) = \{\, & \phi : (\Pi R : \mathsf{Rel})[\![T]\!]_1(\bar{R}, R)(f_0R_0, g_0R_1) \mid \\
& (\forall Q \colon \mathsf{2Rel}) \\
& \big((f_1Q_{r0}, \phi Q_{0r}, g_1Q_{r1}, \phi Q_{1r}) \in \\
& \quad [\![T]\!]_2(\mathsf{Eq}_{\|}(\bar{R}), Q)(f_0Q_{00}, f_0Q_{10}, g_0Q_{01}, g_0Q_{11}) \qquad \text{(A1.1)} \\
& \wedge (\phi Q_{r0}, f_1Q_{0r}, \phi Q_{r1}, g_1Q_{1r}) \in \\
& \quad [\![T]\!]_2(\mathsf{Eq}_{=}(\bar{R}), Q)(f_0Q_{00}, f_0Q_{10}, g_0Q_{01}, g_0Q_{11}) \qquad \text{(A1.2)} \\
& \wedge (f_1Q_{r0}, f_1Q_{0r}, \phi Q_{r1}, \phi Q_{1r}) \in \\
& \quad [\![T]\!]_2(\mathbf{C}\bar{R}, Q)(f_0Q_{00}, f_0Q_{10}, f_0Q_{01}, g_0Q_{11})\big)\} \qquad \text{(A1.3)}
\end{aligned}$$

$$\begin{aligned}
& (\phi_0,\phi_1,\phi_2,\phi_3) \in ([\![\forall X.T]\!]_2\bar{Q})(f,g,h,l) \text{ iff} \\
& \quad (\forall Q \colon \mathsf{2Rel})\ \ (\phi_0Q_{r0}, \phi_1Q_{0r}, \phi_2Q_{r1}, \phi_3Q_{1r}) \in \\
& \qquad [\![T]\!]_2(\bar{Q}, Q)(f_0Q_{00}, g_0Q_{10}, h_0Q_{01}, l_0Q_{11})
\end{aligned}$$

Fig. 2: Interpretation of types

Theorem 22 (IEL). *For every type judgement $\Gamma \vdash T$ type, we have*

(i) an equivalence $\Theta_{T,0} : [\![T]\!]_1 \circ \mathsf{Eq} \cong \mathsf{Eq} \circ [\![T]\!]_0$,

(ii) an equivalence $\Theta_{T,\|} : [\![T]\!]_2 \circ \mathsf{Eq}_{\|} \cong \mathsf{Eq}_{\|} \circ [\![T]\!]_1$ *over* $\Theta_{T,0}$,

(iii) an equivalence $\Theta_{T,=} : [\![T]\!]_2 \circ \mathsf{Eq}_{=} \cong \mathsf{Eq}_{=} \circ [\![T]\!]_1$ *over* $\Theta_{T,0}$, *and*

(iv) an equivalence $\Theta_{T,\mathbf{C}} : [\![T]\!]_2 \circ \mathbf{C} \cong \mathbf{C} \circ [\![T]\!]_1$ *over* $\Theta_{T,0}$. □

Proof. We prove (i) for ∀-types, since it is useful in order to understand the logical relations in Fig. 2. We refer to the appendix for the rest of the proof.

We define maps

$$\Theta_{\forall X.T,0} : [\![\forall X.T]\!]_1\mathsf{Eq}(\bar{A})(f,g) \to \mathsf{Eq}([\![\forall X.T]\!]_0\bar{A})(f,g)$$
$$\Theta^{-1}_{\forall X.T,0} : \mathsf{Eq}([\![\forall X.T]\!]_0\bar{A})(f,g) \to [\![\forall X.T]\!]_1\mathsf{Eq}(\bar{A})(f,g)$$

for all f, g and show that they are inverses — this does not come for free, as in the proof-irrelevant setting, but will be considerably easier since we are considering 1-types only, and not arbitrary types. We first define $\Theta_{\forall X.T,0}(\rho) \coloneqq \varphi(\lambda(A : \text{1-Type}).\, \rho\, \mathsf{Eq}(A))$, where φ is part of the equivalence given by Lemma 21. The condition from Lemma 21 is satisfied by (A1.1) together with the induction hypothesis.

For $\Theta^{-1}_{\forall X.T,0}$, we define $\Theta^{-1}_{\forall X.T,0}(\tau) \coloneqq \lambda R \colon \mathsf{Rel}.\, \mathsf{tr}(f_1 \mathsf{Eq}(R_0), \tau R_1) f_1 R$. We need to check that conditions (A1.1), (A1.2) and (A1.3) are satisfied — we verify (A1.1) in detail here, (A1.2) and (A1.3) follow analogously. Let $Q : \mathsf{2Rel}$. By (A0), we have

$$(f_1 Q_{r0}, f_1 Q_{0r}, f_1 Q_{r1}, f_1 Q_{1r}) \in [\![T]\!]_2(\mathsf{Eq}_2(\bar{A}), Q)(f_0 Q_{00}, f_0 Q_{10}, f_0 Q_{01}, f_0 Q_{11})$$

while we want to prove

$$\begin{aligned}(f_1 Q_{r0}, \mathsf{tr}(f_1 \mathsf{Eq}(Q_{00}), \tau Q_{01}) f_1 Q_{0r}, g_1 Q_{r1}, \mathsf{tr}(f_1 \mathsf{Eq}(Q_{10}), \tau Q_{11}) f_1 Q_{1r})\\ \in [\![T]\!]_2(\mathsf{Eq}_{||}(\mathsf{Eq}(\bar{A})), Q)(f_0 Q_{00}, f_0 Q_{10}, g_0 Q_{01}, g_0 Q_{11}) \ .\end{aligned}$$

Since $\mathsf{Eq}_{||}(\mathsf{Eq}(\bar{A})) \cong \mathsf{Eq}_2(\bar{A})$, we only need to prove

$$p : (f_0 Q_{00}, f_0 Q_{10}, f_0 Q_{01}, f_0 Q_{11}) \equiv (f_0 Q_{00}, f_0 Q_{10}, g_0 Q_{01}, g_0 Q_{11})$$

and

$$\begin{aligned}q : \mathsf{tr}(p)(f_1 Q_{r0}, f_1 Q_{0r}, f_1 Q_{r1}, f_1 Q_{1r}) \equiv\\ (f_1 Q_{r0}, \mathsf{tr}(f_1 \mathsf{Eq}(Q_{00}), \tau Q_{01}) f_1 Q_{0r}, g_1 Q_{r1}, \mathsf{tr}(f_1 \mathsf{Eq}(Q_{10}), \tau Q_{11}) f_1 Q_{1r})\end{aligned}$$

and transport along $\mathsf{pair}_{=}(p, q)$. We use $p = \mathsf{pair}_{=}(f_1 \mathsf{Eq}(Q_{00}), f_1 \mathsf{Eq}(Q_{10}), \tau Q_{01}, \tau Q_{11})$ and q given by conditions (A1.1), (A1.3), (A0) for f_1 and the condition from Lemma 21.

We now check that $\Theta_{\forall X.T,0} \circ \Theta^{-1}_{\forall X.T,0} = \mathsf{id}$ and $\Theta^{-1}_{\forall X.T,0} \circ \Theta_{\forall X.T,0} = \mathsf{id}$. One way round

$$\Theta_{\forall X.T,0}(\Theta^{-1}_{\forall X.T,0}(\tau))(A) = \mathsf{tr}(f_1 \mathsf{Eq}(A), \tau A) f_1 \mathsf{Eq}(A) \equiv (f_1 \mathsf{Eq}(A))^{-1} \bullet f_1 \mathsf{Eq}(A) \bullet \tau A \equiv \tau A$$

by Lemma 1 as required. By definition we have $\Theta^{-1}_{\forall X.T,0}(\Theta_{\forall X.T,0}(\rho))(R) = \mathsf{tr}(f_1 \mathsf{Eq}(A), \rho \mathsf{Eq}(B)) f_1 R$. Condition A1.3 implies that

$$(f_1 \mathsf{Eq}(A), \rho R, \rho \mathsf{Eq}(B), f_1 R) \in [\![T]\!]_2(\mathbf{C}(\mathsf{Eq}(\bar{A})), \mathsf{Eq}_{||}(R))(f_0 A, f_0 A, f_0 B, f_B)$$

and since $\mathbf{C}(\mathsf{Eq}(\bar{A})) \cong \mathsf{Eq}_{||}(\mathsf{Eq}(\bar{A}))$, by the induction hypothesis $(f_1 \mathsf{Eq}(A), \rho R, \rho \mathsf{Eq}(B), f_1 R)$ are related in $\mathsf{Eq}_{||} [\![T]\!]_2(\mathsf{Eq}(\bar{A}), R)$, i.e. $\mathsf{tr}(f_1 \mathsf{Eq}(A), \rho \mathsf{Eq}(B)) f_1 R = \rho(R)$ as required. □

The proof critically uses of the uniformity condition (A1.3) for connections. In the interpretation of ∀-types in Fig. 2, and in the proof of Theorem 22, we made some seemingly arbitrary choices: we choose to only be uniform with respect to one connection, and we used the given f_1, not the given g_1, in order to construct the isomorphism $\Theta^{-1}_{\forall X.T,0}$. The following lemma shows that these choices are actually irrelevant:

Lemma 23. *For every type judgement $\Gamma, X \vdash T$ type and $(f_0, f_1) \in [\![\forall X.T]\!]_0 \vec{A}$, $\phi \in [\![\forall X.T]\!]_1(\mathsf{Eq}\vec{A})(f, g)$, we have:*

(i) For every relation R, $\mathsf{tr}(f_1 \mathsf{Eq} R_0, \phi \mathsf{Eq} R_1) f_1 R = \phi R$.

(ii) For every relation R,

$$\mathsf{tr}(f_1 \mathsf{Eq} R_0, \phi \mathsf{Eq} R_1) f_1 R = \mathsf{tr}((\phi \mathsf{Eq} R_0)^{-1}, (g_1 \mathsf{Eq} R_1)^{-1}) g_1 R.$$

(iii) For every 2-relation Q,

$$(\phi Q_{r0}, \phi Q_{0r}, g_1 Q_{r1}, g_1 Q_{1r}) \in [\![T]\!]_2(\mathbf{C} \circ \mathsf{Eq}\vec{A}, Q)(f_0 Q_{00}, g_0 Q_{10}, g_0 Q_{01}, g_0 Q_{11}).$$

□

Here, item (i) is a technical lemma, while item (ii) says that one can equally well use g_1 as f_1 in the proof of Theorem 22. Finally item (iii) shows that in certain cases, the interpretation of terms of $\forall$-type are uniform also with respect to the other connection which is not explicitly mentioned in the logical relation for $\forall$.

5.2 Interpretation of terms

We next show how to interpret terms. A term $\Gamma; \Delta \vdash t : T$, with $|\Gamma| = n$, will be given a "standard" interpretation

$$[\![t]\!]_0 \vec{A} : [\![\Delta]\!]_0 \vec{A} \to [\![T]\!]_0 \vec{A} \ ,$$

for every $\vec{A} : \mathsf{1\text{-}Type}^n$, a relational interpretation

$$([\![t]\!]_0 \vec{R_0}, [\![t]\!]_0 \vec{R_1}, [\![t]\!]_1 \vec{R}) : [\![\Delta]\!]_1 \vec{R} \to [\![T]\!]_1 \vec{R} \ ,$$

for every $\vec{R} : \mathsf{Rel}^n$, and finally a 2-relational interpretation

$$(([\![t]\!]_0 \vec{Q}_-, [\![t]\!]_1 \vec{Q}_-), [\![t]\!]_2 \vec{Q}) : [\![\Delta]\!]_2 \vec{Q} \to [\![T]\!]_2 \vec{Q}$$

for every $\vec{Q} : \mathsf{2Rel}^n$, where we have written e.g. $[\![t]\!]_0 \vec{Q}_-$ for the map of I_0-types with components $([\![t]\!]_0 \vec{Q}_-)w = [\![t]\!]_0 \vec{Q}_w : [\![\Delta]\!]_0 \vec{Q}_w \to [\![T]\!]_0 \vec{Q}_w$ and similarly for $[\![t]\!]_1 \vec{Q}_-$. At each level, $\Delta = x_1 : T_1, \ldots, x_m : T_m$ is interpreted as the product

$$[\![x_1 : T_1, \ldots, x_m : T_m]\!]_i = [\![T_1]\!]_i \times \ldots \times [\![T_m]\!]_i \ .$$

$$[\![x_0 : T_0, \dots, x_n : T_n \vdash x_k : T_k]\!]_i \vec{X} \coloneqq \pi_k \qquad [\![\Delta, x : S \vdash t : T]\!]_i = [\![\Delta \vdash t : T]\!]_i \circ \pi_0$$

$$[\![\Delta \vdash \lambda x.\, t : S \to T]\!]_0 \vec{A}(\gamma) = \lambda s.\, [\![\Delta, x : S \vdash t : T]\!]_0 \vec{A}(\gamma, s)$$
$$[\![\Delta \vdash \lambda x.\, t : S \to T]\!]_1 \vec{R}(\overline{\gamma}) = \lambda s_0.\, \lambda s_1.\, \lambda s.\, [\![\Delta, x : S \vdash t : T]\!]_1 \vec{R}((\gamma_0, s_0), (\gamma_1, s_1), (\gamma, s))$$
$$[\![\Delta \vdash \lambda x.\, t : S \to T]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma}) = \lambda((x, p), \gamma).\, [\![\Delta, x : S \vdash t : T]\!]_2 \vec{Q}((\overline{x}, x), (\overline{p}, p))(\overline{\gamma}, \gamma)$$

$$[\![f\, t]\!]_0 \vec{A}(\gamma) = [\![f]\!]_0 \vec{A}(\gamma)\, ([\![t]\!]_0 \vec{A}(\gamma))$$
$$[\![f\, t]\!]_1 \vec{R}(\gamma_0, \gamma_1, \gamma) = [\![f]\!]_1 \vec{R}(\gamma_0, \gamma_1, \gamma, [\![t]\!]_0 \vec{R}_0(\gamma_0), [\![t]\!]_0 \vec{R}_1(\gamma_1), [\![t]\!]_1 \vec{R}(\gamma_0, \gamma_1, \gamma))$$
$$[\![f\, t]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma}) = [\![f]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma}, [\![t]\!]_0 \vec{Q}_{\overline{i}}(\overline{x}), [\![t]\!]_1 \vec{Q}_{\overline{j}}(\overline{p}), [\![t]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma}))$$

$$[\![\Lambda X.t]\!]_0 \vec{A}(\gamma) = (\lambda A.\, [\![\Delta, X; \Delta \vdash t : T]\!]_0 (\vec{A}, A)\gamma, \lambda R.\, [\![t]\!]_1 (\mathsf{Eq}(\vec{A}), R) \Theta_{\Delta,0}(\mathsf{refl}(\gamma)))$$
$$[\![\Lambda X.t]\!]_1 \vec{R}(\gamma_0, \gamma_1, \gamma) = \lambda R.\, ([\![t]\!]_1 (\vec{R}, R))(\gamma_0, \gamma_1, \gamma)$$
$$[\![\Delta \vdash \Lambda X.t : \forall X.T]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma}) = \lambda Q.\, [\![t]\!]_2 (\vec{Q}, Q)((\overline{x}, \overline{p}), \overline{\gamma})$$

$$[\![\Delta \vdash t[S] : T[S \mapsto X]]\!]_0 \vec{A}(\gamma) = \mathsf{fst}([\![t]\!]_0 \vec{A}(\gamma))([\![S]\!]_0 \vec{A})$$
$$[\![t[S]]\!]_1 \vec{R}(\gamma_0, \gamma_1, \gamma) = \mathsf{tr}(\Theta_{T,0}(\mathsf{snd}([\![t]\!]_0 \vec{R}_0 \gamma_0) \mathsf{Eq}([\![S]\!]_0 \vec{R}_0)))^{-1}([\![t]\!]_1 \vec{R}(\gamma_0, \gamma_1, \gamma)([\![S]\!]_1 \vec{R}))$$
$$[\![t[S]]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma}) = [\![t]\!]_2 \vec{Q}((\overline{x}, \overline{p}), \overline{\gamma})([\![S]\!]_2 \vec{Q})$$

Fig. 3: Interpretation of terms

The full interpretation is given in Fig. 3. Variables, term abstraction and term application are again given by projections and the exponential structure at each level. For type abstraction and type application, we use the same concepts at the meta-level, but we also have to prove that the resulting term satisfies the uniformity conditions (A0), (A1.1), (A1.2) and (A1.3). In addition, we have to put in a twist for the relational interpretation in order to validate the β- and η-rules.

Lemma 24. *The interpretation in Fig. 3 is well-defined.*

Proof. The interpretation of $\Gamma; \Delta \vdash \Lambda X.t : \forall X.T$ is type-correct, since Δ is weakened with respect to X in $\Gamma, X; \Delta \vdash t : T$. The uniformity conditions (A0), (A1.1), (A1.2) and (A1.3) can all be proven using $[\![t]\!]_2$. □

Theorem 25. *The interpretation defined in Fig. 3 is sound, i.e. if $\Gamma; \Delta \vdash s = t : T$, then there is $p_{\bar{A}} : \mathsf{Id}_{[\![T]\!]_0 \bar{A}}([\![s]\!]_0, [\![t]\!]_0)$ and $q_{\bar{R}} : \mathsf{Id}_{[\![T]\!]_1 \bar{R}}(\mathsf{tr}(p_{\bar{R}_0})([\![s]\!]_1), [\![t]\!]_1)$. (We automatically have $\mathsf{tr}(p, q)[\![s]\!]_2 \equiv [\![t]\!]_2$ by proof-irrelevance of 2-relations.)* □

This model reveals hidden uniformity not only in the "standard" interpretation of terms as functions, but also in the canonical proofs of this uniformity via the Reynolds relational interpretation of terms. In more detail: consider a term $\Gamma; \Delta \vdash t : T$ with $|\Gamma| = n$. By construction, our model shows that if $\vec{R} : \mathsf{Rel}^n$, $a : [\![\Delta]\!]_0 \vec{R}_0$, $b : [\![\Delta]\!]_0 \vec{R}_1$ and $p : [\![\Delta]\!]_1 \vec{R}(a, b)$, then $[\![t]\!]_1 \vec{R}\ p : [\![T]\!]_1 \vec{R}([\![t]\!]_0 \vec{R}_0\ a, [\![t]\!]_0 \vec{R}_1\ b)$, i.e. $[\![t]\!]_1 \vec{R}\ p$ is a proof that $[\![t]\!]_0 \vec{R}_0\ a$ and $[\![t]\!]_0 \vec{R}_1\ b$ are related at $[\![T]\!]_1 \vec{R}$. This is a proof-relevant version of Reynolds' Abstraction Theorem. Furthermore, if $\vec{Q} : \mathsf{2Rel}^n$, $(a, b, c, d) : [\![\Delta]\!]_0 \vec{Q}_{00} \times [\![\Delta]\!]_0 \vec{Q}_{10} \times [\![\Delta]\!]_0 \vec{Q}_{01} \times [\![\Delta]\!]_0 \vec{Q}_{11}$ and $(p, q, r, s) \in [\![\Delta]\!]_2 \vec{Q}(a, b, c, d)$, then

$$
\begin{aligned}
&([\![t]\!]_1 \vec{Q}_{r0}\ p, [\![t]\!]_1 \vec{Q}_{0r}\ q, [\![t]\!]_1 \vec{Q}_{r1}\ r, [\![t]\!]_1 \vec{Q}_{1r}\ s) \in \\
&\qquad\qquad\qquad [\![T]\!]_2 \vec{Q}([\![t]\!]_0 \vec{Q}_{00}\ a, [\![t]\!]_0 \vec{Q}_{10}\ b, [\![t]\!]_0 \vec{Q}_{10}\ c, [\![t]\!]_0 \vec{Q}_{11}\ d)
\end{aligned}
$$

This is the Abstraction Theorem "one level up" for the proofs $[\![t]\!]_1$, which we will put to use in the next section.

6 Theorems about Proofs for Free

In Phil Wadler's famous 'Theorems for free!' [Wadler, 1989], the fact that parametric transformations are always natural in the categorical sense is shown to have many useful and fascinating consequences. Among other things, it is shown that

$$[\![A]\!] \cong [\![\forall X.(A \to X) \to X]\!]$$

for all types A — the categorically inclined reader will recognise this as an instance of the Yoneda Lemma (see e.g. Mac Lane [1998]) for the identity functor, if only we dared to consider the right hand side of the equation to consist of *natural* transformations only. And indeed, as Wadler shows (and Reynolds already knew [1983]), all System F terms $[\![t]\!] : [\![\forall X.(A \to X) \to X]\!]$ are natural by parametricity. Hence, in proof-irrelevant parametric models of System F, indeed $[\![A]\!] \cong [\![\forall X.(A \to X) \to X]\!]$.

In a more expressive theory such as (impredicative) Martin-Löf Type Theory with proof-irrelevant identity types and function extensionality, we can go further even without a relational interpretation, as pointed out by Steve Awodey (personal communication). Taking inspiration from the Yoneda Lemma once again, we can show

$$A \cong (\Sigma t : (\Pi X : \mathsf{Set})(A \to X) \to X)\, \mathsf{isNat}(t) \qquad (1)$$

where

$$\mathsf{isNat}(t) := (\Pi X, Y : \mathsf{Set})(\Pi f : X \to Y)\mathsf{Id}_{(A \to X) \to Y}(f \circ t_X, t_Y \circ (f \circ \_))$$

expresses that t is a natural transformation (note that we need the identity type in order to state this). If we start with the interpretation of a System F term in a proof-irrelevant model of parametricity, we can automatically derive this naturality proof using Wadler's argument.

The above isomorphism (1) relied on A being a set, i.e. that A has no non-trivial higher structure. If we instead consider $A : 1\text{-}\mathsf{Type}$, the isomorphism (1) fails; instead we have

$$A \cong (\Sigma t : (\Pi X : \mathsf{Set})(A \to X) \to X)(\Sigma p : \mathsf{isNat}(t))\, \mathsf{isCoh}(p) \tag{2}$$

where

$$\mathsf{isCoh}(p) := (\Pi X, Y, Z : 1\text{-}\mathsf{Type})(\Pi f : X \to Y)(\Pi g : Y \to Z) \\ (p\, X\, Z\, (g \circ f)) \equiv (p\, Y\, Z\, g) \star (p\, X\, Y\, f)$$

expresses that the proof p is suitably coherent. Here $(p\, Y\, Z\, g) \star (p\, X\, Y\, f)$ is the operation that pastes the two proofs $p\, X\, Y\, f$ and $p\, Y\, Z\, g$ of diagrams commuting into a proof that the composite diagram commutes. Proof-irrelevant parametricity can not ensure this coherence condition, but as we will see, an extension of the usual naturality argument to proof-relevant parametricity will guarantee this extra uniformity of the proof as well.

6.1 Graph relations and graph 2-relations

Relations representing graphs of functions are key to many applications of parametricity.

Definition 26. *Let $f : A \to B$ in 1-Type. We define the* graph $\langle f \rangle$ *of f as* $\langle f \rangle := (A, B, \lambda a.\, \lambda b.\, \mathsf{Id}_B(f\, a, b)) : \mathsf{Rel}$. *This extends to an action on commuting squares: if $g : A' \to B'$, $\alpha : A \to A'$, $\beta : B \to B'$ and $p : (\Pi x : A)\, \mathsf{Id}_{B'}(g(\alpha(a)), \beta(f(a)))$, then we define $\langle \alpha, \beta \rangle = (\alpha, \beta, \lambda a.\, \lambda b.\, \lambda(r : f\, a \equiv b).\, p(a) \bullet \mathsf{ap}(\beta)(r)) : \langle f \rangle \to \langle g \rangle$.*

Abstractly, we see that $\langle f \rangle$ is obtained from $\mathsf{Eq}(B)$ by "reindexing" along (f, id) and there is a morphism $\langle f, \mathsf{id} \rangle : \langle f \rangle \to \mathsf{Eq}(B)$; in particular, we recover $\mathsf{Eq}(B)$ as $\langle \mathsf{id}_B \rangle$. Just like Eq is full and faithful, so is $\langle - \rangle : 1\text{-}\mathsf{Type}^{\to} \to \mathsf{Rel}$:

Lemma 27. *For all $f : A \to B$, $g : A' \to B'$,*

$$(\langle f \rangle \Rightarrow \langle g \rangle) \cong (\Sigma \alpha : A \to A')(\Sigma \beta : B \to B')\mathsf{Id}_{A \to B'}(g \circ \alpha, \beta \circ f) \ .$$

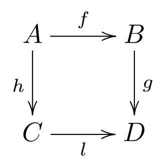
□

The main tool for deriving consequences of parametricity is the Graph Lemma, which relates the graph of the action of a functor on a morphism with its relational action on the graph of the morphism.

Theorem 28. *Let $F_0 : 1\text{-}\mathsf{Type} \to 1\text{-}\mathsf{Type}$ and $F_1 : \mathsf{Rel} \to \mathsf{Rel}$ over F_0 be functorial. If $F_1(\mathsf{Eq}(A)) \cong \mathsf{Eq}(F_0 A)$ for all A, then for any $f : A \to B$, there are morphisms $(\mathsf{id}, \mathsf{id}, \phi_{F,f}) : \langle F_0 f \rangle \to F_1 \langle f \rangle$ and $(\mathsf{id}, \mathsf{id}, \psi_{F,f}) : F_1 \langle f \rangle \to \langle F_0 f \rangle$.* □

Note that in our proof-relevant setting, this theorem does *not* construct an equivalence $\langle F_0 f\rangle \cong F_1\langle f\rangle$. Instead, we only have a logical equivalence, i.e. maps in both directions, and that seems to be enough for all known consequences of parametricity. (In a proof-irrelevant setting, the constructed logical equivalence would automatically be an equivalence.)

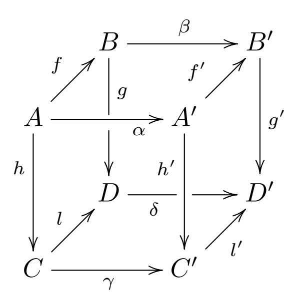
Next, we consider also graph 2-relations. Since we have multiple "equality 2-relations", one could expect also multiple graph 2-relations, but for the application we have in mind, one suffices. Given functions f, g, l and h, we write $\Box(f,g,l,h)$ for the 1-type of proofs that the square



commutes, i.e. $\Box(f,g,l,h) = (\Pi x : A)\mathsf{Id}_D(g(fx), l(hx))$. We define the 1-type of commuting squares by

$$(\text{1-}\mathsf{Type}^{\rightarrow})^{\rightarrow} :\equiv (\Sigma f : A \to B)(\Sigma g : B \to D)(\Sigma l : C \to D)(\Sigma h : A \to C)\Box(f,g,l,h)$$

A morphism $(f,g,l,h,p) \to (f',g',l',h',p')$ in $(\text{1-}\mathsf{Type}^{\rightarrow})^{\rightarrow}$ consists of four morphisms $\alpha : A \to A'$, $\beta : B \to B'$, $\gamma : C \to C'$ and $\delta : D \to D'$, and four proofs $q : \Box(\alpha, h', \gamma, h)$, $q' : \Box(\beta, f', \delta, g)$, $r : \Box(\gamma, l', \delta, l)$ and $r' : \Box(\alpha, f', \beta, f)$ such that they form a "commuting cube"



i.e. such that $p \star q \star r \equiv p' \star q' \star r'$, where $p \star q \star r$ and $p' \star q' \star r'$ are pastings of the squares that proves that both ways from one corner of the cube to the opposite one commutes. The 2-*graph* $\langle \_ \rangle_2 : (\text{1-}\mathsf{Type}^{\rightarrow})^{\rightarrow} \to \mathsf{2Rel}$ is defined by

$$\langle f,g,h,l,p\rangle_2 :\equiv (\langle f\rangle, \langle g\rangle, \langle h\rangle, \langle l\rangle, \lambda(a,b,c,d).\, \lambda(q,r,s,t).\, w(a)\bullet\mathsf{ap}(g)p\bullet q \equiv \mathsf{ap}(l)s\bullet r)$$

It also has an action on morphisms, which we omit here. The 2-relation says that the two ways to prove $h(l(a)) = d$ using p, q, r, s, t are in fact equal. Again, more abstractly, this is a "reindexing" of $\mathsf{Eq}_{||}(\langle g\rangle)$ along the morphism $(\langle f,l\rangle, \langle f, \mathsf{id}\rangle, \mathsf{id}_{\langle g\rangle}), \langle l, \mathsf{id}\rangle) : (\langle h\rangle, \langle f\rangle, \langle g\rangle, \langle l\rangle) \to (\langle g\rangle, \mathsf{Eq}(B), \langle g\rangle, \mathsf{Eq}(D))$ in Rel^4.

Lemma 29. $\langle - \rangle_2$ *is full and faithful in the sense that*

$$(\langle f,g,h,l,p\rangle_2 \to_{\mathsf{2Rel}} \langle f',g',h',l',p'\rangle_2) \cong (f,g,h,l,p) \to_{(\mathit{1}\text{-}\mathsf{Type}^{\to})^{\to}} (f',g',h',l',p')$$

□

This lemma can be used to prove a 2-relational version of the Graph Lemma:

Theorem 30 (2-relational Graph Lemma). *Let* $F_2 : \mathsf{2Rel} \to \mathsf{2Rel}$ *be functorial, and over* (F_0, F_1) *where* F_0 *and* F_1 *are as in Theorem 28. If* $F_2(\mathsf{Eq}R) \cong \mathsf{Eq}(F_1 R)$ *for all* R*, then for any* (f,g,h,l,p) *in* $(\mathit{1}\text{-}\mathsf{Type}^{\to})^{\to}$*, there are morphisms* $\phi_2 : \langle F_0 f, F_0 g, F_0 h, F_0 l, \mathsf{ap}(F_0)p\rangle_2 \to F_2\langle f,g,h,l,p\rangle_2$ *and* $\psi_2 : F_2\langle f,g,h,l,p\rangle_2 \to \langle F_0 f, F_0 g, F_0 h, F_0 l, \mathsf{ap}(F_0)p\rangle$ *in* $\mathsf{2Rel}$ *over* (ϕ,ϕ) *and* (ψ,ψ) *from Theorem 28.* □

6.2 Coherent proofs of naturality

Let us now apply the tools we have developed to the question of the coherence of the naturality proofs from parametricity. We first recall the standard theorem that holds also with proof-irrelevant parametricity:

Theorem 31 (Parametric terms are natural). *Let* $F(X)$ *and* $G(X)$ *be functorial type expressions in the free type variable* X *in some type context* Γ*. Every term* $\Gamma; - \vdash t : \forall X. F(X) \to G(X)$ *gives rise to a natural transformation* $[\![F]\!]_0 \to [\![G]\!]_0$*, i.e. if* $f : A \to B$ *then there is* $\mathsf{nat}(f) : \mathsf{Id}([\![G]\!]_0(f) \circ [\![t]\!]_0\, A, [\![t]\!]_0\, B \circ [\![F]\!]_0(f))$*.*

Proof. We construct $\mathsf{nat}(f)$ using the relational interpretation of t: By construction, $[\![t]\!]_1\, \langle f\rangle : [\![F]\!]_1(\langle f\rangle) \to [\![G]\!]_1(\langle f\rangle)$, hence using Theorem 28,

$$\psi_{G,f} \circ [\![t]\!]_1\, \langle f\rangle \circ \phi_{F,f} : (\Pi xy)\, \langle [\![F]\!]_0 f\rangle(x,y) \to \langle [\![G]\!]_0 f\rangle([\![t]\!]_0\, Ax, [\![t]\!]_0\, B\, y)$$

and since $\mathsf{refl} : \langle [\![F]\!]_0 f\rangle(a, ([\![F]\!]_0 f)a)$ for each $a : [\![F]\!]_0 A$, we can define $\mathsf{nat}(f) :\equiv \mathsf{ext}(\lambda a.\, (\psi_{G,f} \circ [\![t]\!]_1\, \langle f\rangle \circ \phi_{F,f})\, a\, (([\![F]\!]_0 f)a)\, \mathsf{refl})$. □

In order for $([\![t]\!]_0, \mathsf{nat})$ to lie in the image of the isomorphism (2), we also need the naturality proofs to be coherent. But thanks to the 2-relational interpretation, we can show that they are:

Theorem 32 (Naturality proofs are coherent). *Let* F*,* G *and* t *be as in Theorem 31. The proof* $\mathsf{nat} : \mathsf{isNat}([\![t]\!]_0)$ *is coherent, i.e. for all* $f : A \to B$ *and* $g : B \to C$*, there is a proof* $\mathsf{coh}(f,g) : \mathsf{Id}(\mathsf{nat}(g \circ f), \mathsf{nat}(g) \star \mathsf{nat}(f))$*.*

Proof. We construct $\mathsf{coh}(f,g)$ using the 2-relational interpretation of t. By construction, $[\![t]\!]_2\langle f,g,g\circ f,\mathsf{id},\mathsf{refl}\rangle_2 : [\![F]\!]_2\langle f,g,g\circ f,\mathsf{id},\mathsf{refl}\rangle_2 \to [\![G]\!]_2\langle f,g,g\circ f,\mathsf{id},\mathsf{refl}\rangle_2$, hence using Theorem 30,

$$\begin{aligned}&\phi_2 \circ [\![t]\!]_2\langle f,g,g\circ f,\mathsf{id},\mathsf{refl}\rangle_2 \circ \psi_2 :\\ &\qquad (\Pi(\bar{x},\bar{r}))\big(\bar{r} \in \langle F_0 f, F_0 g, F_0(g\circ f), \mathsf{id}, \mathsf{ap}(F_0)p\rangle_2 \bar{x}\\ &\qquad\qquad \to ([\![t]\!]_1 \bar{r}) \in \langle G_0 f, G_0 g, G_0(g\circ f), \mathsf{id}, \mathsf{ap}(G_0)p\rangle_2([\![t]\!]_0 \bar{x})\big)\end{aligned}$$

We define

$$\mathsf{coh}(f,g) := \mathsf{ext}(\lambda a.\,(\phi_2 \circ [\![t]\!]_2 \langle f, g, g\circ f, \mathsf{id}, \mathsf{refl}\rangle_2 \circ \psi_2)\,(a, (F_0 f)a, F_0(g\circ f)a, a)\,\vec{\mathsf{refl}})$$

— this works, since ϕ_2 and ψ_2 are over (ϕ,ϕ) and (ψ,ψ) respectively, since $\mathsf{nat}(h)$ is defined to be $(\phi \circ [\![t]\!]_1 \circ \psi)\mathsf{refl}$, and since the 2-relation $\langle G_0 f, G_0 g, G_0(g \circ f), \mathsf{id}, \mathsf{ap}(G_0)p\rangle_2$ exactly says that pasting the two diagrams produces the third in this case. □

7 Conclusions and future work

In this paper, we tackled the concrete problem of transporting Reynolds' theory of relational parametricity to a proof-relevant setting. This is non-trivial as one must modify Reynolds' uniformity predicate on polymorphic functions so that it itself becomes parametric. Implementing this intuition has significant mathematical ramifications: an extra layer of 2-dimensional relations is needed to formalise the idea of two proofs being related to each other. Further, there are a variety of choices to be made as to what face maps and degeneracies to use between proof-relevant relations and 2-relations. Having made these choices, we showed that the key theorems of parametricity, namely the identity extension lemma and the fundamental theorem of logical relations hold. Finally, we explored how a standard consequence of relational parametricty — namely the fact that parametricity implies naturality — also holds in the proof-relevant setting. This work complements the more proof-theoretic work on *internal* parametricity in proof-relevant frameworks [Bernardy et al., 2015, 2012; Polonsky, 2015]. Relevant is also the work on parametricity for dependent types in general [Atkey et al., 2014; Krishnaswami and Dreyer, 2013], assuming proof-irrelevance.

In terms of future work, we are extending the results of this paper to arbitrary dimensions. We have a candidate definition of higher-dimensional relations, the requisite face maps and degeneracies and we have proven the Identity Extension Lemma. What remains to do is to fully investigate the consequences. For instance, what form of higher dimensional initial algebra theorem can be proved with higher dimensional parametricity? More generally, we need to compare the methods, structures and results of higher dimensional parametricity with (where possible) Homotopy Type Theory and in particular its cubical sets model [Bezem et al., 2014], which shares many striking similarities. Finally, once the theoretical framework is settled, we will want to implement it and then use that implementation in formal proof.

Acknowledgements. We thank Bob Atkey, Peter Hancock and the anonymous reviewers for helpful discussions and comments.

References

Atkey, R.: Relational parametricity for higher kinds. In: Cégielski, P., Durand, A. (eds.) CSL 2012. LIPIcs, vol. 16, pp. 46–61. Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2012)

Atkey, R., Ghani, N., Johann, P.: A relationally parametric model of dependent type theory. In: POPL. pp. 503–515. ACM (2014)

Bernardy, J.P., Coquand, T., Moulin, G.: A presheaf model of parametric type theory. In: Ghica, D.R. (ed.) MFPS. pp. 17–33. ENTCS, Elsevier (2015)

Bernardy, J.P., Jansson, P., Paterson, R.: Proofs for free. Journal of Functional Programming 22, 107–152 (2012)

Bezem, M., Coquand, T., Huber, S.: A model of type theory in cubical sets. In: Types for Proofs and Programs (TYPES 2013). Leibniz International Proceedings in Informatics, vol. 26, pp. 107–128. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2014)

Brown, R., Higgins, P.J.: On the algebra of cubes. Journal of Pure and Applied Algebra 21(3), 233 – 260 (1981)

Brown, R., Higgins, P.J., Sivera, R.: Nonabelian Algebraic Topology: Filtered spaces, crossed complexes, cubical homotopy groupoids, EMS Tracts in Mathematics, vol. 15. European Mathematical Society Publishing House (2011)

Coquand, T., Huet, G.: The Calculus of Constructions. Information and Computation 76, 95–120 (1988)

Dunphy, B., Reddy, U.: Parametric limits. In: LICS. pp. 242–251 (2004)

Garner, R.: Two-dimensional models of type theory. Mathematical Structures in Computer Science 19(04), 687–736 (2009)

Ghani, N., Johann, P., Nordvall Forsberg, F., Orsanigo, F., Revell, T.: Bifibrational functorial semantics of parametric polymorphism. In: Ghica, D.R. (ed.) MFPS. pp. 67–83. ENTCS, Elsevier (2015a)

Ghani, N., Nordvall Forsberg, F., Orsanigo, F.: Parametric polymorphism — universally. In: de Paiva, V., de Queiroz, R., Moss, L.S., Leivant, D., de Oliveira, A.G. (eds.) WoLLIC. LNCS, vol. 9160, pp. 81–92. Springer (2015b)

Grandis, M.: The role of symmetries in cubical sets and cubical categories (on weak cubical categories, I). Cah. Topol. Gom. Diff. Catg. 50(2), 102–143 (2009)

Krishnaswami, N.R., Dreyer, D.: Internalizing relational parametricity in the extensional calculus of constructions. In: CSL. pp. 432–451 (2013)

Mac Lane, S.: Categories for the working mathematician, vol. 5. Springer (1998)

Martin-Löf, P.: An intuitionistic theory of types (1972), published in Twenty-Five Years of Constructive Type Theory

O'Hearn, P.W., Tennent, R.D.: Parametricity and local variables. Journal of the ACM 42(3), 658–709 (1995)

Polonsky, A.: Extensionality of lambda-*. In: Herbelin, H., Letouzey, P., Sozeau, M. (eds.) 20th International Conference on Types for Proofs and Programs (TYPES 2014). Leibniz International Proceedings in Informatics (LIPIcs), vol. 39, pp. 221–250. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2015)

Reynolds, J.: Types, abstraction and parametric polymorphism. In: Mason, R.E.A. (ed.) Information Processing 83. pp. 513–523 (1983)
Robinson, E., Rosolini, G.: Reflexive graphs and parametric polymorphism. In: LICS. pp. 364–371 (1994)
Strachey, C.: Fundamental concepts in programming languages. Higher Order Symbolic Computation 13(1-2), 11–49 (2000)
The Univalent Foundations Program: Homotopy Type Theory: Univalent Foundations of Mathematics. `http://homotopytypetheory.org/book` (2013)
Voevodsky, V.: The equivalence axiom and univalent models of type theory. Talk at CMU on February 4, 2010 (2010), `http://arxiv.org/abs/1402.5556`
Wadler, P.: Theorems for free! In: FPCA. pp. 347–359 (1989)
Wadler, P.: The Girard-Reynolds isomorphism (second edition). Theoretical Computer Science 375(1-3), 201–226 (2007)

A Proofs from Section 5

Proof (of Theorem 22). The proof is done by induction on type judgements. For type variables, all statements are trivial. For arrow types, this is Propositions 8 and 20.

It remains to prove (ii), (iii) and (iv) for ∀-types. In this case we only need to produce maps in both directions — they will automatically compose to the identity by proof-irrelevance of 2-relations. The structure of the proof is the same for all of the three points.

For (ii) consider $(\tau_0, \rho_0, \tau_1, \rho_1) \in \mathsf{Eq}_{||}([\![\forall X.T]\!]_1 \bar{R})(f, g, h, l)$. We want to show that $(\Psi(\tau_0), \rho_0, \Psi(\tau_1), \rho_1) \in [\![\forall X.T]\!]_2 \mathsf{Eq}_{||}(\bar{R})(f, g, h, l)$, i.e. that for every 2-relation Q,

$$(\Psi(\tau_0)Q_{r0}, \rho_0 Q_{0r}, \Psi(\tau_1)Q_{r1}, \rho_1 Q_{1r}) \in [\![T]\!]_2(\mathsf{Eq}_{||}(\bar{R}), Q)(f_0 Q_{00}, g_0 Q_{10}, h_0 Q_{01}, l_0 Q_{11}).$$

By condition A1.1, we have

$$(f_1 Q_{r0}, \rho_0 Q_{1r}, f_1 Q_{r1}, \rho_0 Q_{0r}) \in [\![T]\!]_2(\mathsf{Eq}_{||}(\bar{R}), Q)(f_0 Q_{00}, f_0 Q_{10}, h_0 Q_{01}, h_0 Q_{11})$$

and using the equalities $(f_1 \mathsf{Eq} Q_{00}, \tau_0 \mathsf{Eq} Q_{10}, h_1 \mathsf{Eq} Q_{01}, \tau_1 \mathsf{Eq} Q_{11})$, we can show

$$\begin{aligned}(f_0 Q_{00}, g_0 Q_{10}, h_0 Q_{01}, l_0 Q_{11}, \Psi(\tau_0) Q_{r0}, \rho_0 Q_{0r}, \Psi(\tau_1) Q_{r1}, \rho_1 Q_{1r})\\ \equiv (f_0 Q_{00}, f_0 Q_{10}, h_0 Q_{01}, h_0 Q_{11}, f_1 Q_{r0}, \rho_0 Q_{1r}, f_1 Q_{r1}, \rho_0 Q_{0r})\end{aligned}$$

We now transport across this equality to finish the argument.

Finally, in the other direction, if $(\rho_0, \rho_1, \rho_2, \rho_3) \in [\![\forall X.T]\!]_2 \mathsf{Eq}_{||}(\bar{R})(f, g, h, l)$, then $(\Theta(\rho_0), \rho_1, \Theta(\rho_2), \rho_3) \in \mathsf{Eq}_{||}([\![\forall X.T]\!]_1 \bar{R})(f, g, h, l)$ by straightforward calculation and the definition of $[\![\forall X.T]\!]_2$.

The case (iii) is just the same as the previous case, the only difference is that we now transport starting from condition (A1.2) and adjust the equalities along which we transport.

The last case (iv) is more complicated. Consider $(\tau_0, \tau_1, \rho_0, \rho_1) \in \mathbf{C}(\llbracket \forall X.T \rrbracket_1 \bar{R})(f, g, h, l)$. We want to show that $(\Psi(\tau_0), \Psi(\tau_1), \rho_0, \rho_1) \in \llbracket \forall X.T \rrbracket_2 \mathbf{C}(\bar{R})(f, g, h, l)$, i.e. that for every 2-relation Q,

$$(\Psi(\tau_0)Q_{r0}, \Psi(\tau_1)Q_{0r}, \rho_0 Q_{r1}, \rho_1 Q_{1r}) \in \llbracket T \rrbracket_2(\mathbf{C}(\bar{R}), Q)(f_0 Q_{00}, g_0 Q_{10}, h_0 Q_{01}, l_0 Q_{11}) \ .$$

By condition A1.3, we have

$$(h_1 Q_{r0}, h_1 Q_{1r}, \rho_0 Q_{r1}, \rho_0 Q_{0r}) \in \llbracket T \rrbracket_2(\mathbf{C}\bar{R}, Q)(h_0 Q_{00}, h_0 Q_{10}, h_0 Q_{01}, l_0 Q_{11})$$

and using the equalities

$$((\tau_1 Q_{00})^{-1}) \centerdot f_1 \mathsf{Eq} Q_{00}, (\tau_1 \mathsf{Eq} Q_{10})^{-1} \centerdot \tau_0 Q_{10}, (h_1 \mathsf{Eq} Q_{01})^{-1} \centerdot h_1 \mathsf{Eq} Q_{01}, \mathsf{refl}),$$

we can show

$$\begin{aligned}(f_0 Q_{00}, g_0 Q_{10}, h_0 Q_{01}, l_0 Q_{11}, \Psi(\tau_0)Q_{r0}, \Psi(\tau_1)Q_{0r}, \rho_0 Q_{r1}, \rho_1 Q_{1r})\\ \equiv (h_0 Q_{00}, h_0 Q_{10}, h_0 Q_{01}, l_0 Q_{11}, h_1 Q_{r0}, h_1 Q_{1r}, \rho_0 Q_{r1}, \rho_0 Q_{0r})\end{aligned}$$

We can now transport across this equality to finish the argument. This requires the use of Lemmas 23 (i) and (ii), and the fact that $(\tau_0, \tau_1, \rho_0, \rho_1) \in \mathbf{C}(\llbracket \forall X.T \rrbracket_1 \bar{R})(f, g, h, l)$.

Finally, In the other direction, if $(\rho_0, \rho_1, \rho_2, \rho_3) \in \llbracket \forall X.T \rrbracket_2 \mathbf{C}(\bar{R})(f, g, h, l)$, then $(\Theta(\rho_0), \Theta(\rho_1), \rho_2, \rho_3) \in \mathbf{C}(\llbracket \forall X.T \rrbracket_1 \bar{R})(f, g, h, l)$ by straightforward calculation and the definition of $\llbracket \forall X.T \rrbracket_2$. □

Proof (of Lemma 23).

(i) Since

$$(f_1 R, f_1 \mathsf{Eq} R_0, f_1 R, f_1 \mathsf{Eq} R_1) \in \llbracket T \rrbracket_2(\mathbf{C} \circ \mathsf{Eq}(\vec{A}), \mathsf{Eq}_=(R))(f_0 R_0, f_0 R_1, g_0 R_0, g_0 R_1)$$

and $\llbracket T \rrbracket_2(\mathbf{C} \circ \mathsf{Eq}(\vec{A}), \mathsf{Eq}_=(R)) = \llbracket T \rrbracket_2(\mathsf{Eq}_= \circ \mathsf{Eq}(\vec{A}), \mathsf{Eq}_=(R)) \cong \mathsf{Eq}_=(\llbracket T \rrbracket_1(\mathsf{Eq}\vec{A}, R))$, by Theorem 22(iii), the thesis follows.

(ii) By assumption,

$$(f_1 R, \phi \mathsf{Eq} R_0, g_1 R, \phi \mathsf{Eq} R_1) \in \llbracket T \rrbracket_2(\mathsf{Eq}_= \mathsf{Eq}\vec{A}, \mathsf{Eq}_= R)(f_0 R_0, f_1 R_1, g_0 R_0, g_0 R_1).$$

By Theorem 22(iii), $\llbracket T \rrbracket_2(\mathsf{Eq}_= \mathsf{Eq}\vec{A}, \mathsf{Eq}_= R) \cong \mathsf{Eq}_=(\llbracket T \rrbracket_1(\mathsf{Eq}\vec{A}, R))$, hence we have $\mathsf{tr}((\phi \mathsf{Eq} R_0)^{-1}, (g_1 \mathsf{Eq} R_1)^{-1}) g_1 R = \mathsf{tr}(f_1 \mathsf{Eq} R_0, \phi \mathsf{Eq} R_1) f_1 R$. If we now transport $(f_1 R, \phi \mathsf{Eq} R_0, g_1 R, \phi \mathsf{Eq} R_1)$ along the equality proof $((f_1 \mathsf{Eq} R_0)^{-1}, (f_1 \mathsf{Eq} R_1)^{-1}, (\phi \mathsf{Eq} R_0)^{-1}, (g_0 \mathsf{Eq} R_1)^{-1})$, the result follows.

(iii) By assumption,

$$(g_1 Q_{r0}, g_1 Q_{0r}, g_1 Q_{r1}, g_1 Q_{1r}) \in \llbracket T \rrbracket_2(\mathsf{Eq}_2 \vec{A}, Q)(g_0 Q_{00}, g_0 Q_{10}, g_0 Q_{01}, g_0 Q_{11})$$

We can transport $(g_1 Q_{r0}, g_1 Q_{0r}, g_1 Q_{r1}, g_1 Q_{1r})$ along the equality $((\phi \mathsf{Eq} Q_{00})^{-1}, (g_1 \mathsf{Eq} Q_{10})^{-1}, (g_1 \mathsf{Eq} Q_{01})^{-1}, (g_1 \mathsf{Eq} Q_{11})^{-1})$. By (i) and (ii), condition (A0), and $\llbracket T \rrbracket_2(\mathsf{Eq}_2 \vec{A}, Q) = \llbracket T \rrbracket_2(\mathbf{C} \circ \mathsf{Eq}\vec{A}, Q)$, the thesis follows. □

Proof (of Theorem 25). We need to check that the β- and η-rules for both term and type abstraction are respected. For term abstraction, this follows from Lemmas 13 and 16.

We next consider the η-rule for type abstraction. Let $\Gamma; \Delta \vdash t : \forall X.T$ be given. Let $[\![t]\!]_0 \vec{A}\gamma = (f_0, f_1)$. Showing $[\![\Lambda X.t[X]]\!]_0 \equiv [\![t]\!]_0$ means giving $p_0 : \mathsf{Id}(\lambda A.f_0 A, f_0)$ and $p_1 : \mathsf{Id}(\lambda R.(\mathsf{tr}(p_0 \bullet (\mathsf{snd}(([\![t]\!]_0 \bar{A}\gamma)\mathsf{Eq}(R_0))))^{-1}([\![t]\!]_1 \mathsf{Eq}(\bar{A})\Theta_{\Delta,0}(\mathsf{refl}(\gamma))\, R)), \mathsf{snd}([\![t]\!]_0 \bar{A}\gamma))$. For p_0, we choose $p_0 = \mathsf{refl}$. Note that

$$([\![t]\!]_1 \mathsf{Eq}(\bar{A})\Theta_{\Delta,0}(\mathsf{refl}(\gamma))\, R)) = \Theta_{\Delta,0}(\mathsf{refl}([\![t]\!]_0 \vec{A}\gamma))\, R =$$

$$\mathsf{tr}(f_1 \mathsf{Eq}(R_0), \mathsf{refl}) f_1 R$$

under the equivalence with respect to $\tau = \mathsf{refl}$, and

$$\mathsf{tr}(\mathsf{refl} \bullet (\mathsf{snd}(([\![t]\!]_0 \bar{A}\gamma)\mathsf{Eq}(R_0))))^{-1} = \mathsf{tr}(f_1 \mathsf{Eq}(R_0), \mathsf{refl})^{-1}.$$

In this way we can conclude

$$\begin{aligned} \mathsf{tr}(f_1 \mathsf{Eq}(R_0), \mathsf{refl})^{-1}(\Theta_{\Delta,0}(\mathsf{refl}([\![t]\!]_0 \vec{A}\gamma))\, R) &= \mathsf{tr}(f_1 \mathsf{Eq}(R_0), \mathsf{refl})^{-1} \mathsf{tr}(f_1 \mathsf{Eq}(R_0), \mathsf{refl}) f_1 R \\ &= f_1 R. \end{aligned}$$

Similarly, things are exactly lined up to make $\mathsf{tr}(\mathsf{pair}_=(p_0, p_1))([\![\Lambda X.t[X]]\!]_1) \equiv [\![t]\!]_1$ trivial.

For the β-rule, consider $\Gamma, X \vdash t \colon T$. We can use $p_{\bar{A}} = ([\![t]\!]_1 \mathsf{Eq}(\vec{A}, [\![S]\!]_0 \vec{A})\Theta_{\Delta,0}(\mathsf{refl}(\gamma)))^{-1}$ to prove $[\![(\Lambda X.t)[S]]\!]_0 \bar{A}\gamma \equiv [\![t[X \mapsto S]]\!]_0 \bar{A}\gamma$. This makes $\mathsf{tr}(p_{\bar{R}_0})([\![(\Lambda X.t)[S]]\!]_1)\bar{R}\bar{\gamma} \equiv [\![t[X \mapsto S]]\!]_1 \bar{R}\bar{\gamma}$ trivial, again using Lemma 1. $\square$