A General Framework for Relational Parametricity

Kristina Sojakova and Patricia Johann

January 2018

Abstract

Reynolds' original theory of *relational parametricity* intended to capture the idea that polymorphically typed System F programs preserve all relations between inputs. But as Reynolds himself later showed, his theory can only be formalized in a meta-theory with an impredicative universe, such as Martin-Löf Type Theory. A number of more abstract treatments of relational parametricity have since appeared; however, as we show, none of these truly generalize Reynolds' original theory, in the sense of having it as a direct instance. Indeed, they all require certain strictness conditions that Reynolds' theory does not satisfy. To correct this, we develop an abstract framework for relational parametricity that does deliver Reynolds' theory as a direct instance in a natural way. This framework is parametric and uniform with respect to a choice of meta-theory, which allows us to obtain the well-known PER model of Longo and Moggi as a direct instance in a natural way as well. Moreover, we demonstrate on a concrete example that our notion of parametricity also encompasses *proof-relevant* parametric models, which does not seem to be the case for the well-known definitions. Our framework is thus both *descriptive*, in that it accounts for well-known models, and *prescriptive*, in that it identifies properties that good models of relational parametricity should satisfy. It is constructed using the new notion of a *split λ2-fibration with isomorphisms*, introduced in this paper, which relaxes certain strictness requirements on split λ2-fibrations. Our main theorem is a generalization of Seely's classical construction of sound models for System F from split λ2-fibrations: we prove that the canonical model of System F induced by every split λ2-fibration with isomorphisms validates System F's entire equational theory on the nose, independently of the parameterizing meta-theory.

1 Introduction

Reynolds [11] introduced the notion of *relational parametricity* to model the extensional behavior of programs in System F [4], the formal calculus at the core of all polymorphic functional languages. His goal was to give a type $\alpha \vdash T(\alpha)$ an *object interpretation* T_0 and a *relational interpretation* T_1, where T_0 takes sets to sets and T_1 takes relations $R \subseteq A \times B$ to relations $T_1(R) \subseteq T_0(A) \times T_0(B)$. A term $\alpha; x : S(\alpha) \vdash t(\alpha, x) : T(\alpha)$ was to be interpreted as a map t_0 associating to each set

A a function $t_0(A) : S_0(A) \to T_0(A)$. The interpretations were to be given inductively on the structure of T and t in such a way that they implied two key theorems: the *Identity Extension Lemma*, stating that if R is the equality relation on A then $T_1(R)$ is the equality relation on $T_0(A)$; and the *Abstraction Theorem*, stating that, for any relation $R \subseteq A \times B$, $t_0(A)$ and $t_0(B)$ map arguments related by $S_1(R)$ to results related by $T_1(R)$. A similar result holds for types and terms with any number of free variables.

In Reynolds' treatment of relational parametricity, if U is the type $\alpha \vdash S(\alpha) \to T(\alpha)$, for example, then $U_0(A)$ is the set of functions $f : S_0(A) \to T_0(A)$ and, for $R \subseteq A \times B$, $U_1(R)$ relates $f : S_0(A) \to T_0(A)$ to $g : S_0(B) \to T_0(B)$ iff f and g map arguments related by $S_r(R)$ to results related by $T_1(R)$. Similarly, if V is the type $\cdot \vdash \forall\alpha.S(\alpha)$, then V_0 consists of those polymorphic functions taking any set A and returning an element of $S_0(A)$, and V_1 relates two such functions f and g iff, for an arbitrary relation $R \subseteq A \times B$, $f(A)$ and $f(B)$ are related by $S_1(R)$. The Abstraction Theorem then guarantees that, for any term $t : \forall\alpha.\alpha \to \alpha$, the interpretation t_0 of t is related to itself by the relational interpretation of $\forall\alpha.\alpha \to \alpha$. So if we fix a set A, fix $x \in A$, and define a relation on A by $R := \{(x, x)\}$, then $t_0(A)$ must be related to itself by the relational interpretation of $\alpha \vdash \alpha \to \alpha$ applied to R. This means that $t_0(A)$ must carry arguments related by R to results related by R. Since x is related to itself by R, $t_0(A)\, x$ must be related to itself by R, so that $t_0(A)\, x$ must be x. That is, t_0 must be the polymorphic identity function. Significantly, Reynolds' approach to relational parametricity allows the derivation of this conclusion just from the type of t. Such applications of relational parametricity are useful in many different scenarios, *e.g.*, when proving invariance of polymorphic functions under changes of data representation, equivalences of programs, and "free theorems" [15].

The well-known problem ([12]) with Reynolds' treatment of relational parametricity is that the universe of sets is not impredicative, and hence the aforementioned "set" V_0 cannot be formed. This issue can be resolved if we instead work in a meta-theory which *does* have an impredicative universe; a natural choice is the impredicative version of extensional Martin-Löf Type Theory (MLTT). We thus have two canonical relationally parametric models of System F: *i)* the PER model of Longo and Moggi [6], internal to the theory of ω-sets and realizable functions, and *ii)* Reynolds' original model[1], internal to MLTT.

Since Reynolds' original paper, a number of more abstract treatments of relational parametricity inspired by Reynolds have appeared; see, *e.g.*, [1, 2, 3, 5, 7, 13]. These works all aim to distill the essence of parametric polymorphism into a notion of a *relationally parametric model* of System F — *i.e.*, a model for which appropriate versions of Reynolds' Abstraction Theorem and Identity Extension Lemma can be formulated and verified. Here we aim to do more, namely to transform this essence into a *semantic framework* (rather than just a reusable blueprint) for constructing such models. But:

What constitutes a good framework for relational parametricity?

Our answer is that such a framework should:

[1] Since there are no set-theoretic models of System F, by the phrase "Reynolds' original model" we will always mean the version of his model that is internal to extensional MLTT with an impredicative universe. The need for impredicativity is inherited from Reynolds' original construction, and is not a new requirement.

1. *Deliver a relationally parametric model for each instantiation of its parameters, from which it uniformly produces such models. In particular, it should allow a choice of a suitable meta-theory (MLTT, the theory of ω-sets, etc.)*

2. *Admit the two canonical relationally parametric models mentioned above as direct instances in a natural, uniform way.*

3. *Identify properties that good models of System F parametricity should be expected to satisfy.*

The existing approaches for constructing frameworks for parametricity can be divided into external (*e.g.*, [2, 5, 7, 13]) and internal (*e.g.*, [1, 3]). In essence, both approaches use a split $\lambda 2$-fibration to construct a sound model of System F in the standard way à la Seely [14]. The external approach starts with an arbitrary split $\lambda 2$-fibration and endows it with enough additional structure that the model it induces can reasonably be considered relationally parametric. The internal approach, on the other hand, constructs the split $\lambda 2$-fibration in a particular way that ensures the resulting model is relationally parametric "by construction". Surprisingly, no existing framework for relational parametricity, whether produced by an internal or external approach, has Reynolds' original model as a direct instance. Our second criterion for a good framework for relational parametricity thus remains unsatisfied.

To see why, we first note that existing internally produced frameworks require certain strictness conditions that the syntactic model does not satisfy. For example, let $\alpha \vdash S(\alpha)$ and $\alpha \vdash T(\alpha)$ be two types, with object interpretations S_0 and T_0 and relational interpretations S_1 and T_1. The interpretation of the product $\alpha \vdash S(\alpha) \times T(\alpha)$ should be an appropriate product of interpretations; that is, the object interpretation should map a set A to $S_0(A) \times T_0(A)$ and the relational interpretation should map a relation R to $S_1(R) \times T_1(R)$, where the product of two relations is defined in the obvious way. For the Identity Extension Lemma to hold, we need $S_1(\mathsf{Eq}(A)) \times T_1(\mathsf{Eq}(A))$ to be the same as $\mathsf{Eq}(S_0(A) \times T_0(A))$. Here, the equivalence relation $\mathsf{Eq}(A)$ on a set A maps $(a, b) : A \times A$ to the type $\mathsf{Id}(a, b)$ of proofs of equality between a and b, so that a and b are related iff $\mathsf{Id}(a, b)$ is inhabited, *i.e.*, iff a is equal to b. By the induction hypothesis, $S_1(\mathsf{Eq}(A))$ is $\mathsf{Eq}(S_0(A))$, and similarly for T, so we need to show that $\mathsf{Eq}(S_0(A)) \times \mathsf{Eq}(T_0(A))$ is $\mathsf{Eq}(S_0(A) \times T_0(A))$. But this is not the case since the identity type on a product is in general not *identical* to the product of identity types, but rather just suitably *isomorphic*.[2]

To see why existing externally produced frameworks also fail to directly subsume Reynolds' original model, we note that the need to replace identity by isomorphism in the statement of the Identity Extension Lemma has some major implications. First, it takes us out of the canonical framework of [1], which requires strict commutativity with Eq. Secondly, since we now work up to isomorphism, we want to ensure that the interpretations of types preserve isomorphisms, *i.e.*, if $A \cong B$ then $T_0(A) \cong T_0(B)$, and similarly for T_1. Without this requirement it can happen that, *e.g.*, the interpretations of $\alpha \vdash S(\alpha)$ and $\cdot \vdash T$ commute with Eq up to isomorphism, but the interpretation of the

[2] Even working in a version of MLTT with the univalence principle would not help us here: it would make the two types *propositionally equal* but still not identical.

type $\cdot \vdash S[\alpha := T]$ does not. This means that types should no longer be interpreted by *discrete* functors $T_0 : |\mathsf{Set}| \to \mathsf{Set}$ and $T_1 : |\mathcal{R}| \to \mathcal{R}$ as in [1], since in this case there is no telling what T_0 and T_1 do to isomorphic sets and relations. Finally, working up to isomorphism also takes us out of the scope of direct application of Seely's standard construction: since some of the structure (namely, the adjoints that interpret $\forall$-types) commutes with substitution only up to isomorphism, the $\lambda 2$-fibration we obtain will not be *split*. Altogether this means that Reynolds' original model is not an instance of any of these externally produced frameworks, at least not directly, since they only show how to construct models of System F from split $\lambda 2$-fibrations.

At this point we could appeal to some of the known strictification techniques for other type theories (see, *e.g.*, [10]) to turn our $\lambda 2$-fibration into a split one, and attempt to make the resulting fibration an instance of one of the external frameworks. However, this is not ideal for our purposes since the strictification process generally changes the basic interpretation (of, *e.g.*, the $\forall$-types). Although the new interpretation is equivalent to the original one, in a precise sense that depends on the technique used, the user of the framework usually has a specific interpretation in mind that is appropriate to the intended application of parametricity. For example, in our reasoning about polymorphic identity above, the proof relied on the fact that the relational interpretation of the type $\forall\alpha.\alpha \to \alpha$ was the suitably chosen one. Having a strictification procedure baked into a framework would make it less usable for practical purposes.

In this paper we take a new approach, potentially of independent interest. We first generalize the notion of a split $\lambda 2$-fibration to the new notion of a *split $\lambda 2$-fibration with isomorphisms*, and then show that every split $\lambda 2$-fibration with isomorphisms induces a sound model of System F *directly*, *i.e.*, without the need for strictification. To say when such a model is relationally parametric, we axiomatize the notions of sets and relations in the form of a *reflexive graph category*, here a generalization of the eponymous notion from [1] that accommodates meta-theories suitable not only for semantic models, but for syntactic models as well. A reflexive graph category provides two *face maps* (called ∂_0 and ∂_1 in [1]), which represent the domain and codomain projections, and a *degeneracy* (called I in [1]), which represents the equality functor. The models constructed by our framework interpret each type as a face map- and degeneracy-preserving *reflexive graph functor*, and each term as a face map- and degeneracy-preserving *reflexive graph natural transformation*.

The main contributions of this paper are as follows:

- *We demonstrate that existing frameworks for relational parametricity for System F fail to directly subsume both canonical models of relational parametricity for System F.*

- *We solve this problem by developing a* good *abstract framework for relational parametricity that is parameterized on a choice of meta-theory, delivers both canonical relationally parametric models of System F as direct instances in a uniform way, and identifies properties that good models of System F parametricity should be expected to satisfy,* e.g.*, the guarantee that the interpretations of terms, not just types, suitably commute with the degeneracy.*

- *We construct our interpretation by first introducing the novel notion of a* split

λ2-fibration with isomorphisms*, which allows type formers to commute with substitution only up to isomorphism, and then proving a generalization of Seely's result ensuring that the canonical model induced by any such fibration validates System F's equational theory on the nose.*

- *We give a novel definition of a* parametric model of System F*, which is a hybrid of the external and internal approaches, and show that it subsumes both canonical models (expressed as instances of our framework), as well as the proof-relevant model adapted from [8] and described in Example 75.*

Fibrational Preliminaries We give a brief introduction to fibrations, mainly to settle notation. More details can be found in, *e.g.*, [5].

Definition 1. *Let $U : \mathcal{E} \to \mathcal{B}$ be a functor. A morphism $g : Q \to P$ in $\mathcal{E}$ is* cartesian *over $f : X \to Y$ in $\mathcal{B}$ if $Ug = f$ and, for every $g' : Q' \to P$ in $\mathcal{E}$ with $Ug' = f \circ v$ for some $v : UQ' \to X$, there is a unique $h : Q' \to Q$ with $Uh = v$ and $g' = g \circ h$. A functor $U : \mathcal{E} \to \mathcal{B}$ is a* fibration *if, for every object P of $\mathcal{E}$ and morphism $f : X \to UP$ of $\mathcal{B}$, there is a cartesian morphism in $\mathcal{E}$ with codomain P over f.*

If $U : \mathcal{E} \to \mathcal{B}$ is a fibration then $\mathcal{E}$ is its *total category* and $\mathcal{B}$ is its *base category*. An object P in $\mathcal{E}$ is *over* its image UP, and similarly for morphisms. A morphism is *vertical* if it is over id. We write $\mathcal{E}(X)$ for the *fiber over* an object X in $\mathcal{B}$, *i.e.*, the subcategory of $\mathcal{E}$ of objects over X and morphisms over id_X.

If $U : \mathcal{E} \to \mathcal{B}$ is a fibration, we call a cartesian morphism over f with codomain P a *cartesian lifting* of f with codomain P with respect to U. A cartesian lifting of f with codomain P with respect to U need not be unique, but it is always unique up to vertical isomorphism. We are interested in fibrations in which representative cartesian liftings are specified, or chosen.

Definition 2. *A fibration $U : \mathcal{E} \to \mathcal{B}$* is cloven *if it comes with a* choice *of cartesian liftings,* i.e.*, with one cartesian lifting of f with codomain P with respect to U regarded as primary amongst all such cartesian liftings for each morphism f in $\mathcal{B}$ and object P in $\mathcal{E}$.*

We emphasize that the choice of cartesian liftings is part of the structure that is given when a fibration is cloven. In this case, When a fibration is cloven we will use the phrase "the cartesian lifting" of f with codomain P to refer to the chosen such lifting, which we will denote by $f^{\S}_P$. Any time we consider categorical objects (*e.g.*, categories, functors, etc.) with particular structures (*e.g.*, products, adjoints, etc.) in this paper, we intend that those structures are chosen in this sense. of Definition 2.

When $U : \mathcal{E} \to \mathcal{B}$ is cloven we write $f^{\S}_P$ for the specified cartesian lifting of f with codomain P and we write f^*P for the domain of this lifting.

The function mapping each object P of $\mathcal{E}$ to the domain f^*P of $f^{\S}_P$ extends to a functor $f^* : \mathcal{E}_Y \to \mathcal{E}_X$ mapping each morphism $k : P \to P'$ in $\mathcal{E}_Y$ to the unique morphism f^*k such that $k \circ f^{\S}_P = f^{\S}_{P'} \circ f^*k$. The universal property of $f^{\S}_{P'}$ ensures the existence and uniqueness of f^*k. We call f^* the *substitution functor along f*. We will be especially interested in cloven fibrations for which the substitution functors induced by these liftings are well-behaved.

Definition 3. *A cloven fibration $U : \mathcal{E} \to \mathcal{B}$ is* split *if its substitution functors are such that* $\mathsf{id}^* = \mathsf{id}$ *and* $(g \circ f)^* = f^* \circ g^*$.

We will later require even more structure of our split fibrations:

Definition 4. *A split fibration $U : \mathcal{E} \to \mathcal{B}$ has a* split generic object *if there is an object Ω in $\mathcal{B}$, together with a collection of isomorphisms θ_X mapping each morphism from X to Ω in $\mathcal{B}$ to an object of the fiber $\mathcal{E}_X$ that is natural in X,* i.e.*, is such that $\theta_Y(fg) = g^*(\theta_X(f))$ for every $f : X \to \Omega$ and $g : Y \to X$.*

2 Reflexive Graph Categories

Although Reynolds himself showed that his original approach to relational parametricity does not work in set theory, we can still use it as a guide for designing an abstract framework for parametricity. Instead of sets and relations, we consider abstract notions of "0-relations" and "1-relations", and require them to be related as follows: *i)* for any 1-relation R, there are two canonical ways of projecting a 0-relation out of R, corresponding to the domain and codomain operations, *ii)* for any 0-relation A, there is a canonical way of turning it into a 1-relation, corresponding to the equality relation on A, and *iii)* if we start with a 0-relation A, turn it into a 1-relation according to *ii)*, and then project out a 0-relation according to *i)*, we get A back. This suggests that our abstract relations and the canonical operations on them can be organized into a reflexive graph structure: categories $\mathcal{X}_0$, $\mathcal{X}_1$ and functors $\mathbf{f}_\top, \mathbf{f}_\bot : \mathcal{X}_1 \to \mathcal{X}_0$, $\mathbf{d} : \mathcal{X}_0 \to \mathcal{X}_1$ such that $\mathbf{f}_\top \circ \mathbf{d} = \mathsf{id} = \mathbf{f}_\bot \circ \mathbf{d}$, as is done in [1].

Since there are no set-theoretic models of System F ([12]), the reflexive graph structure identified above must to be internal to some ambient category $\mathcal{C}$. For Reynolds' original model, the ambient category has types $A : \mathbb{U}_1$ as objects and terms $f : \Sigma_{A,B:\mathbb{U}_1} A \to B$ as morphisms.[3] Here, $\mathbb{U}_1$ is the universe one level above the impredicative universe $\mathbb{U}_0$; we will denote $\mathbb{U}_0$ simply by $\mathbb{U}$ below. This ensures that $\mathbb{U}$ is an object in $\mathcal{C}$. To model relations, we introduce:

$$\mathsf{isProp}(A) :\equiv \Pi_{a,b:A}\, \mathsf{Id}(a,b)$$
$$\mathsf{Prop} :\equiv \Sigma_{A:\mathbb{U}}\, \mathsf{isProp}(A)$$

The type Prop of *propositions* singles out those types in $\mathbb{U}$ with the property that any two inhabitants, if they exist, are equal. Propositions can be used to model relations as follows: in Reynolds' original model, $a : A$ is related to $b : B$ in at most one way under any relation R (either $(a,b) \in R$ or not), so the type of proofs that $(a,b) \in R$ is a proposition. Conversely, given $R : A \times B \to \mathsf{Prop}$, we consider a and b as related by R iff $R(a,b)$ is inhabited.

To see the universe $\mathbb{U}$ as a category Set internal to $\mathcal{C}$ we take its object of objects Set_0 to be $\mathbb{U}$ and define its object of morphisms by $\mathsf{Set}_1 :\equiv \Sigma_{A,B:\mathbb{U}} A \to B$. We define the category R of relations by giving its objects R_0 and R_1 of objects and morphisms,

[3] As is standard, by "types" and "terms" we mean equivalence classes of types and terms under definitional equality.

respectively:

$$\begin{aligned}
\mathsf{R}_0 &\coloneqq \Sigma_{A,B:\mathbb{U}}\, A \times B \to \mathsf{Prop} \\
\mathsf{R}_1 &\coloneqq \Sigma_{((A_1,A_2),R_A),((B_1,B_2),R_B):\mathsf{R}_0} \Sigma_{(f,g):(A_1 \to B_1)\times(A_2 \to B_2)} \\
&\qquad \Pi_{(a_1,a_2):A_1\times A_2} R_A(a_1,a_2) \to R_B(f(a_1),g(a_2))
\end{aligned}$$

We clearly have two internal functors from R to Set corresponding to the first and second projections. We also have an internal functor Eq from Set to R that constructs an equality relation, defined by

$$\begin{aligned}
\mathsf{Eq}\, A &\coloneqq ((A,A),\mathsf{Id}_A) \\
\mathsf{Eq}\,((A,B),f) &\coloneqq \big((\mathsf{Eq}\, A, \mathsf{Eq}\, B),(f,f),\mathsf{ap}_f\big)
\end{aligned}$$

Here, the term $\mathsf{ap}_f : \mathsf{Id}_A(a_1,a_2) \to \mathsf{Id}_B(f(a_1),f(a_2))$, which is defined as usual by Id-induction, witnesses the fact that f respects equality.

These observations motivate the next two definitions, in which we denote the category of categories and functors internal to $\mathcal{C}$ by $\mathsf{Cat}(\mathcal{C})$, and assume $\mathcal{C}$ is locally small and has all finite products.

Definition 5. *A* reflexive graph structure $\mathcal{X}$ *on a category C consists of:*

- *objects $\mathcal{X}(0)$ and $\mathcal{X}(1)$ of C*
- *distinct arrows $\mathcal{X}(\mathbf{f}_\star) : \mathcal{X}(1) \to \mathcal{X}(0)$ for $\star : \mathbf{Bool}$*
- *arrow $\mathcal{X}(\mathbf{d}) : \mathcal{X}(0) \to \mathcal{X}(1)$*

such that $\mathcal{X}(\mathbf{f}_\star) \circ \mathcal{X}(\mathbf{d}) = \mathsf{id}$.

The requirement that $\mathcal{X}(\mathbf{f}_\top)$ and $\mathcal{X}(\mathbf{f}_\bot)$ are distinct serves a purpose similar to that served by the requirement in Definition 8.6.2 of [5] that the fibre category $\mathbb{F}_1$ over the terminal object in $\mathbb{C}$ is the category of relations in the preorder fibration $\mathbb{D} \to \mathbb{E}$ on the fibre category $\mathbb{E}_1$ over the terminal object in $\mathbb{B}$. Both conditions ensure that the relations considered are not just the trivial (*i.e.*, equality) ones but can, in particular, also be heterogeneous. Moreover, since the two face maps are distinct, any morphism generated by face maps and degeneracies must be one of the seven distinct maps $\mathsf{id}_{\mathcal{X}(0)}, \mathsf{id}_{\mathcal{X}(1)}, \mathcal{X}(\mathbf{f}_\star), \mathcal{X}(\mathbf{d})$, and $\mathcal{X}(\mathbf{d}) \circ \mathcal{X}(\mathbf{f}_\star)$ for $\star : \mathbf{Bool}$. Unlike in [5], where relations are obtained in a standard way as predicates (given by a preorder fibration) over a product, we do not assume that relations are constructed in any specific way, but rather only that the abstract operations on relations suitably interact.

Definition 6. *A* reflexive graph category (on $\mathcal{C}$) *is a reflexive graph structure on* $\mathsf{Cat}(\mathcal{C})$.

Example 7 (PER model)**.** *The ambient category $\mathcal{C}$ is taken to be the category of ω-sets, given in Definition 6.3 of [6]. We construct a reflexive graph functor, which we call $\mathcal{R}$, as follows. The internal category $\mathcal{R}(0)$ of 0-relations is the category $\mathbf{M}'$ given in Definition 8.4 of [6]. Informally, the objects of $\mathbf{M}'$ are partial equivalence relations on $\mathbb{N}$, and the morphisms are realizable functions that respect such relations. To define the internal category $\mathcal{R}(1)$ of relations, we first construct its object of objects. The carrier* *of this ω-set is the set of pairs of the form $R := ((A_0, A_1), R_A)$, where A_0 and A_1 are partial equivalence relations and R_A is a saturated predicate on the product PER $A_1 \times A_2$. A saturated predicate on a PER A is a predicate on $\mathbb{N}$ such that $R(a)$ implies $a \sim_A a$ and, moreover, $a_1 \sim_A a_2$ and $R(a_1)$ imply $R(a_2)$. To finish the construction of our object of objects for $\mathcal{R}(1)$ we take any pair $((A, B), R)$ as above to be realized by any natural number.*

The carrier of the object of morphisms for $\mathcal{R}(1)$ comprises all pairs of the form

$$\big(\big(((A_1, B_1), R_1), ((A_2, B_2), R_2)\big), \big(\{m_1\}_{A_1 \to A_2}, \{m_2\}_{B_1 \to B_2}\big)\big)$$

satisfying the condition that, for any k, l such that $R_1(\langle k, l\rangle)$ holds, $R_2\big(\langle m_1 \cdot k, m_2 \cdot l\rangle\big)$ holds as well. The first component records the domain and codomain of the morphism and the second component is a pair of equivalence classes under the specified exponential PERs. As in [6], we denote the application of the n^{th} partial recursive function to a natural number a in its domain by $n \cdot a$. To finish the construction of the object of morphisms for $\mathcal{R}(1)$, we take a pair of pairs as above to be realized by a natural number k iff $\mathsf{fst}(k) \sim_{A_1 \to A_2} m_1$ and $\mathsf{snd}(k) \sim_{B_1 \to B_2} m_2$.

We clearly have two internal functors from $\mathcal{R}(1)$ to $\mathcal{R}(0)$, corresponding to the first and second projections. We also have an equality functor Eq from $\mathcal{R}(0)$ to $\mathcal{R}(1)$ defined by

$$\mathsf{Eq}\, A := ((A, A), R_A)$$
$$\mathsf{Eq}\, ((A, B), \{m\}_{A \to B}) := \big((\mathsf{Eq}\, A, \mathsf{Eq}\, B), (\{m\}_{A \to B}, \{m\}_{A \to B})\big)$$

where $R_A(k)$ iff $\mathsf{fst}(k) \sim_A \mathsf{snd}(k)$.

Example 8 (Reynolds' model)**.** *We obtain a reflexive graph category $\mathcal{R}$ by taking $\mathcal{R}(0) := \mathsf{Set}$, $\mathcal{R}(1) := \mathsf{R}$, and $\mathcal{R}(\mathbf{d}) := \mathsf{Eq}$, and letting $\mathcal{R}(\mathbf{f}_\top)$ and $\mathcal{R}(\mathbf{f}_\perp)$ be morphisms corresponding to the first and second projections, respectively.*

If $\mathcal{X}$ is a reflexive graph category, then the discrete graph category $|\mathcal{X}|$ and the product reflexive graph category $\mathcal{X}^n$ for $n \in \mathbb{N}$ are defined in the obvious ways: $|\mathcal{X}(l)|$ has the same objects as $\mathcal{X}(l)$ but only the identity morphisms, and $(\mathcal{X} \times \mathcal{Y})(l) = \mathcal{X}(l) \times \mathcal{Y}(l)$ for $l \in \{0, 1\}$. For the latter, the product on the right-hand side is a product of internal categories, which exists because $\mathcal{C}$ has finite products by assumption.

If C is an internal category, we denote by C_0 and C_1 the objects of $\mathcal{C}$ representing the objects and morphisms of C, respectively. If $F : C \to D$ is an internal functor, we denote by $F_0 : C_0 \to D_0$ and $F_1 : C_1 \to D_1$ the arrows of $\mathcal{C}$ representing the object and morphisms parts of F, respectively. Also:

Notation 9. *We will use the following notation with respect to an internal category C in $\mathcal{C}$:*

- *Given a "generalized object" $a : J \to C_0$ (with J arbitrary), we denote by $\mathsf{id}_C[a]$ the arrow $\mathsf{id}_C \circ a$, where $\mathsf{id}_C : C_0 \to C_1$ is the arrow representing identity morphisms in C.*
- *For a "generalized morphism" $f : J \to C_1$ (with J arbitrary), we denote by $\mathsf{s}_C[f]$ and $\mathsf{t}_C[f]$ the arrows $\mathsf{s}_C \circ f$ and $\mathsf{t}_C \circ f$ respectively, where $\mathsf{s}_C, \mathsf{t}_C : C_1 \to C_0$ are the arrows representing the source and target operations in C.*

- *For generalized morphisms $f, g : J \to C_1$ such that $\mathsf{t}_C[f] = \mathsf{s}_C[g]$, we denote by $g \circ_C f$ the arrow $\mathsf{comp}_C \circ \langle f, g \rangle$, where comp_C : pullback$(\mathsf{t}_C, \mathsf{s}_C) \to C_1$ is the arrow representing composition in C, its domain pullback$(\mathsf{t}_C, \mathsf{s}_C)$ is the pullback of the two arrows t_C and s_C, and $\langle f, g \rangle$ is the canonical morphism into this pullback.*

- *We say that $f : J \to C_1$ is an isomorphism if there exists a $g : J \to C_1$ such that* $\mathsf{s}_C[f] = \mathsf{t}_C[g]$, $\mathsf{s}_C[g] = \mathsf{t}_C[f]$ *and* $f \circ_C g = \mathsf{id}_C[\mathsf{s}_C[g]]$, $g \circ_C f = \mathsf{id}_C[\mathsf{s}_C[f]]$. *If such a g exists, it is necessarily unique and hence will be denoted by f^{-1}.*

Given a reflexive graph category $\mathcal{X}$ axiomatizing the 0-and 1-relations, an obvious first attempt at pushing Reynolds' original idea through is to take the interpretation $[\![T]\!]$ of a type $\overline{\alpha} \vdash T$ with n free type variables to be a pair $([\![T]\!](0), [\![T]\!](1))$, where $[\![T]\!](0) : |\mathcal{X}(0)|^n \to \mathcal{X}(0)$ and $[\![T]\!](1) : |\mathcal{X}(1)|^n \to \mathcal{X}(1)$ are functions giving the "set" and "relation" interpretations of the type T. Although this approach will need some tweaking — we will need to endow $[\![T]\!](0)$ and $[\![T]\!](1)$ with actions on *some* morphisms — it suggests:

Definition 10. *Let $\mathcal{X}$ and $\mathcal{Y}$ be reflexive graph categories. A* reflexive graph functor *$\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ is a pair $(\mathcal{F}(0), \mathcal{F}(1))$ of functors such that $\mathcal{F}(0) : \mathcal{X}(0) \to \mathcal{Y}(0)$ and $\mathcal{F}(1) : \mathcal{X}(1) \to \mathcal{Y}(1)$.*

Writing T_0 for $[\![T]\!](0)$ and T_1 for $[\![T]\!](1)$, we recall from the introduction that T_0 and T_1 should be appropriately related via the domain and codomain projections and the equality functor. Since the two face maps $\mathcal{X}(\mathbf{f}_\star)$ now model the projections, and the degeneracy $\mathcal{X}(\mathbf{d})$ models the equality functor, we end up with the following conditions:

i) for each object $\overline{R}$ in $\mathcal{X}(1)^n$, we have $\mathcal{X}(\mathbf{f}_\star)\, T_1(\overline{R}) = T_0(\mathcal{X}(\mathbf{f}_\star)^n\, \overline{R})$

ii) for each object $\overline{A}$ in $\mathcal{X}(0)^n$, we have $\mathcal{X}(\mathbf{d})\, T_0(\overline{A}) = T_1(\mathcal{X}(\mathbf{d})^n\, \overline{A})$

We examine what these conditions imply for Reynolds' model by considering the product $\alpha \vdash S(\alpha) \times T(\alpha)$ of two types $\alpha \vdash S(\alpha)$ and $\alpha \vdash T(\alpha)$. By the induction hypothesis, S and T are interpreted as pairs (S_0, S_1) and (T_0, T_1), where $S_0, T_0 : \mathsf{Set}_0 \to \mathsf{Set}_0$ and $S_1, T_1 : \mathsf{R}_0 \to \mathsf{R}_0$ satisfy *i)* and *ii)*. The interpretation of a product should be a product of interpretations, *i.e.*, $(S \times T)_0\, A \coloneqq S_0(A) \times T_0(A)$ and $(S \times T)_1\, R \coloneqq S_1(R) \times T_1(R)$. It remains to be seen that this interpretation satisfies *i)* and *ii)*. Fix a relation R on A and B. Condition *i)* entails that $S_1(R) \coloneqq ((S_0(A), S_0(B)), R_S)$ and $T_1(R) \coloneqq ((T_0(A), T_0(B)), R_T)$ for some R_S and R_T. Thus $S_1(R) \times T_1(R)$ has the form $\big((S_0(A) \times T_0(A), S_0(B) \times T_0(B)), R_{S \times T}\big)$, where $R_{S \times T}$ maps a pair of pairs $((a, b), (c, d))$ to $R_S(a, c) \times R_T(b, d)$. We can see that *i)* is satisfied simply by construction, which leads us to define:

Definition 11. *A reflexive graph functor $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ is* face map-preserving *if the following diagram in* $\mathsf{Cat}(\mathcal{C})$ *commutes for all $\star \in \mathbf{Bool}$:*

$$\begin{array}{ccc}
\mathcal{X}(1) & \xrightarrow{\mathcal{F}(1)} & \mathcal{Y}(1) \\
{\scriptstyle \mathcal{X}(\mathbf{f}_\star)}\downarrow & & \downarrow{\scriptstyle \mathcal{Y}(\mathbf{f}_\star)} \\
\mathcal{X}(0) & \xrightarrow[\mathcal{F}(0)]{} & \mathcal{Y}(0)
\end{array}$$

To verify condition *ii)* for the interpretation of the product of $S \times T$, fix a set A. Condition *ii)* for S gives that $S_1(\mathsf{Eq}(A))$ is $\mathsf{Eq}(S_0(A))$ and similarly for T. We thus need to show that $\mathsf{Eq}(S_0(A)) \times \mathsf{Eq}(T_0(A))$ is $\mathsf{Eq}(S_0(A) \times T_0(A))$. But while the domains and codomains of these two relations agree – all are $S_0(A) \times T_0(A)$ – the former relation maps $((a, b), (c, d))$ to $\mathsf{Id}(a, c) \times \mathsf{Id}(b, d)$, while the latter maps it to $\mathsf{Id}((a, b), (c, d))$. These two types are not identical, but they *are* isomorphic, (*i.e.*, there are functions back and forth that compose to identity on both sides). We could attempt to remedy the situation by working in intensional MLTT with the univalence principle and quotienting arrows in the ambient category $\mathcal{C}$ by propositional equality. The isomorphism between $\mathsf{Id}((a, b), (c, d))$ and $\mathsf{Id}(a, c) \times \mathsf{Id}(b, d)$ would become an equivalence and, via univalence, a propositional equality. This would make the interpretation of $S \times T$ degeneracy-preserving as desired. However, this approach fails elsewhere: if we quotient arrows in $\mathcal{C}$ by propositional equality, then two morphisms $((A, B), f)$ and $((B', C), g)$ in Set such that $\mathsf{Id}(B, B')$ is inhabited become composable. But there is no natural way to compose them since B and B' can be propositionally equal in many different ways.

Instead, we relax condition *ii)* to allow an isomorphism $\varepsilon_T(\overline{A}) : \mathcal{X}(\mathbf{d})\, T_0(\overline{A}) \cong T_1(\mathcal{X}(\mathbf{d})^n\, \overline{A})$. In fact, we can require more: the respective domains and codomains of $\mathcal{X}(\mathbf{d})\, T_0(\overline{A})$ and $T_1(\mathcal{X}(\mathbf{d})^n\, \overline{A})$ coincide by condition *i)*, so we can insist that both projections map the isomorphism $\varepsilon_T(\overline{A})$ to the identity morphism on $T_0(\overline{A})$. This coherence condition is a natural counterpart to the equation $\mathcal{X}(\mathbf{f}_\star) \circ \mathcal{X}(\mathbf{d}) = \mathsf{id}$ and turns out to be not just a design choice but a necessary requirement: in Reynolds' model, for instance, the proof that the interpretations of $\forall$-types (as defined later) suitably commute with the functor Eq depends precisely on the fact that the morphisms underlying the maps $\varepsilon_T(\overline{A})$ are identities. We thus define:

Definition 12. *A reflexive graph functor $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ is* degeneracy-preserving *if the following diagram in $\mathsf{Cat}(\mathcal{C})$ commutes up to a given natural isomorphism $\varepsilon_{\mathcal{F}}$ satisfying the coherence condition $\mathcal{Y}(\mathbf{f}_\star)_1 \circ \varepsilon_{\mathcal{F}} = \mathsf{id}_{\mathcal{Y}(0)}[\mathcal{F}(0)_0]$ for $\star \in \mathbf{Bool}$:*

$$\begin{array}{ccc}
\mathcal{X}(0) & \xrightarrow{\mathcal{F}(0)} & \mathcal{Y}(0) \\
{\scriptstyle \mathcal{X}(\mathbf{d})}\downarrow & & \downarrow{\scriptstyle \mathcal{Y}(\mathbf{d})} \\
\mathcal{X}(1) & \xrightarrow[\mathcal{F}(1)]{} & \mathcal{Y}(1)
\end{array}$$

Thus, as a first approximation, we can try to interpret a type $\overline{\alpha} \vdash T$ with n free type variables as a face map- and degeneracy-preserving reflexive graph functor $(T_0, T_1) : |\mathcal{X}|^n \to \mathcal{X}$. Reynolds' original idea for interpreting terms suggests that the interpretation of a term $\overline{\alpha}; x : S \vdash t : T$ should be a (vacuously) natural transformation $t_0 : S_0 \to T_0$. As observed in [3], the Abstraction Theorem can then be formulated as follows: there is a (vacuously) natural transformation $t_1 : S_1 \to T_1$ such that, for any object $\overline{R}$ in $\mathcal{X}(1)^n$, we have $\mathcal{X}(\mathbf{f}_\star)\, t_1(\overline{R}) = t_0(\mathcal{X}(\mathbf{f}_\star)^n\, \overline{R})$. To see that this does indeed give what we want, we revisit Reynolds' model. There, the face maps are the domain and codomain projections and an object $\overline{R}$ in $\mathcal{X}(1)^n$ is an n-tuple of relations. Denote $\mathcal{X}(\mathbf{f}_\top)^n\ \overline{R}$ by $\overline{A}$ and $\mathcal{X}(\mathbf{f}_\bot)^n\ \overline{R}$ by $\overline{B}$. Then $t_1(\overline{R})$ is a morphism of relations from $S_1(\overline{R})$ to $T_1(\overline{R})$ and, since S_1 and T_1 are face map-preserving, $S_1(\overline{R}) \coloneqq \big((S_0(\overline{A}), S_0(\overline{B})), R_S\big)$ and $T_1(\overline{R}) \coloneqq \big((T_0(\overline{A}), T_0(\overline{B})), R_T\big)$ for some R_S and R_T. By definition, $t_1(\overline{R})$ gives a pair of maps $(f, g) : \big(S_0(\overline{A}) \to T_0(\overline{A}), S_0(\overline{B}) \to T_0(\overline{B})\big)$, together with a map $h : \Pi_{(a_1,a_2):S_0(\overline{A})\times S_0(\overline{B})} R_S(a_1, a_2) \to R_B(f(a_1), g(a_2))$ stating precisely that f and g map related inputs to related outputs. Since $\mathcal{X}(\mathbf{f}_\top)\, t_1(\overline{R})$ is $\big((S_0(\overline{A}), T_0(\overline{A})), f\big)$ and $\mathcal{X}(\mathbf{f}_\bot)\, t_1(\overline{R})$ is $\big((S_0(\overline{B}), T_0(\overline{B})), g\big)$, the requirement that $\mathcal{X}(\mathbf{f}_\star)\, t_1(\overline{R})$ is $t_0(\mathcal{X}(\mathbf{f}_\star)^n\, \overline{R})$ implies that the maps underlying $t_0(\overline{A})$ and $t_0(\overline{B})$ must be f and g, respectively, and so must indeed map related inputs to related outputs. Pairing the natural transformations t_0 and t_1 motivates:

Definition 13. *Let $\mathcal{F}, \mathcal{G} : \mathcal{X} \to \mathcal{Y}$ be reflexive graph functors. A* reflexive graph natural transformation *$\eta : \mathcal{F} \to \mathcal{G}$ is a pair $(\eta(0), \eta(1))$ of natural transformations $\eta(0) : \mathcal{F}(0) \to \mathcal{G}(0)$ and $\eta(1) : \mathcal{F}(1) \to \mathcal{G}(1)$.*

The Abstraction Theorem then further suggests defining:

Definition 14. *A reflexive graph natural transformation $\eta : \mathcal{F} \to \mathcal{G}$ between two face map-preserving reflexive graph functors is* face map-preserving *if for any $\star \in \mathbf{Bool}$ we have*

$$\mathcal{Y}(\mathbf{f}_\star)_1 \circ \eta(1) = \eta(0) \circ \mathcal{X}(\mathbf{f}_\star)_0$$

The interpretation of a term $\overline{\alpha}; x : S \vdash t : T$ should then be a face map-preserving natural transformation from (S_0, S_1) to (T_0, T_1). We also have the dual notion:

Definition 15. *A reflexive graph natural transformation $\eta : \mathcal{F} \to \mathcal{G}$ between two degeneracy-preserving reflexive graph functors $(\mathcal{F}, \varepsilon_\mathcal{F})$ and $(\mathcal{G}, \varepsilon_\mathcal{G})$ is* degeneracy-preserving *if for any $\star \in \mathbf{Bool}$, we have*

$$(\eta(1) \circ \mathcal{X}(\mathbf{d})_0) \circ_{\mathcal{Y}(1)} \varepsilon_\mathcal{F} = \varepsilon_\mathcal{G} \circ_{\mathcal{Y}(1)} (\mathcal{Y}(\mathbf{d})_1 \circ \eta(0))$$

Intuitively, the above equation represents the commutativity of the following diagram in the internal category $\mathcal{Y}(1)$:

$$\begin{array}{ccc}
\mathcal{Y}(\mathbf{d})_0 \circ \mathcal{F}(0)_0 & \xrightarrow{\ \varepsilon_\mathcal{F}\ } & \mathcal{F}(1)_0 \circ \mathcal{X}(\mathbf{d})_0 \\
{\scriptstyle \mathcal{Y}(\mathbf{d})_1 \circ \eta(0)} \big\downarrow & & \big\downarrow {\scriptstyle \eta(1) \circ \mathcal{X}(\mathbf{d})_0} \\
\mathcal{Y}(\mathbf{d})_0 \circ \mathcal{G}(0)_0 & \xrightarrow[\ \varepsilon_\mathcal{G}\ ]{} & \mathcal{G}(1)_0 \circ \mathcal{X}(\mathbf{d})_0
\end{array}$$

There is no explicit analogue of Definition 15 in Reynolds' model. This is because Reynolds' model (as well as the PER model) is proof-irrelevant, in the precise sense that the functor $\langle \mathcal{X}(\mathbf{f}_\perp), \mathcal{X}(\mathbf{f}_\top) \rangle$ is faithful. The proof-irrelevance implies that any face map-preserving natural transformation is automatically degeneracy-preserving as well. Since we aim to subsume both proof-irrelevant and *proof-relevant* models (Example 75), we explicitly require the reflexive graph natural transformations interpreting terms to be degeneracy-preserving (as is also done in [1]). From now on we will only consider reflexive graph functors and natural transformations that are both face map- and degeneracy-preserving and will omit specific mentions of these properties.

We have the usual laws of identity and composition of reflexive graph functors and natural transformations:

Definition 16. *Given a reflexive graph category $\mathcal{X}$, the identity reflexive graph functor $1_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$ is defined as follows:*

- $1_{\mathcal{X}}(l)$ *is the identity functor on* $\mathcal{X}(l)$
- $\varepsilon_{1_{\mathcal{X}}} \coloneqq \mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0]$

Definition 17. *Given two reflexive graph functors $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ and $\mathcal{G} : \mathcal{Y} \to \mathcal{Z}$, let $\mathcal{G} \circ \mathcal{F} : \mathcal{X} \to \mathcal{Z}$ be the reflexive graph functor defined as follows:*

- $(\mathcal{G} \circ \mathcal{F})(l) \coloneqq \mathcal{G}(l) \circ \mathcal{F}(l)$
- $\varepsilon_{\mathcal{G}\circ\mathcal{F}} \coloneqq (\mathcal{G}(1)_1 \circ \varepsilon_{\mathcal{F}}) \circ_{\mathcal{Z}(1)} (\varepsilon_{\mathcal{G}} \circ \mathcal{F}(0)_0)$

Definition 18. *Given a reflexive graph functor $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$, the identity reflexive graph natural transformation $1_{\mathcal{F}} : \mathcal{F} \to \mathcal{F}$ is defined by $1_{\mathcal{F}}(l) = \mathsf{id}_{\mathcal{Y}(l)}[\mathcal{F}(l)_0]$.*

Definition 19. *Given reflexive graph functors $\mathcal{F}, \mathcal{G}, \mathcal{H} : \mathcal{X} \to \mathcal{Y}$ and reflexive graph natural transformations $\eta_1 : \mathcal{F} \to \mathcal{G}$ and $\eta_2 : \mathcal{G} \to \mathcal{H}$, let $\eta_2 \circ \eta_1 : \mathcal{F} \to \mathcal{H}$ be the reflexive graph natural transformation defined by $(\eta_2 \circ \eta_1)(l) \coloneqq \eta_2(l) \circ_{\mathcal{Y}(l)} \eta_1(l)$.*

Definition 20. *Given reflexive graph functors $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ and $\mathcal{G}_1, \mathcal{G}_2 : \mathcal{Y} \to \mathcal{Z}$, and a reflexive graph natural transformation $\eta : \mathcal{G}_1 \to \mathcal{G}_2$, let $\eta \circ \mathcal{F} : \mathcal{G}_1 \circ \mathcal{F} \to G_2 \circ \mathcal{F}$ be the reflexive graph natural transformation defined by $(\eta \circ \mathcal{F})(l) \coloneqq \eta(l) \circ \mathcal{F}(l)_0$.*

Definition 21. *Given reflexive graph functors $\mathcal{F}_1, \mathcal{F}_2 : \mathcal{X} \to \mathcal{Y}$ and $\mathcal{G} : \mathcal{Y} \to \mathcal{Z}$, and a reflexive graph natural transformation $\eta : \mathcal{F}_1 \to \mathcal{F}_2$, let $\mathcal{G} \circ \eta : \mathcal{G} \circ \mathcal{F}_1 \to \mathcal{G} \circ \mathcal{F}_2$ be the reflexive graph natural transformation defined by $(\mathcal{G} \circ \eta)(l) \coloneqq \mathcal{G}(l)_1 \circ \eta(l)$.*

One basic example of a reflexive graph functor which will be used often and will end up interpreting type variables is the projection:

Definition 22. *Given a reflexive graph category $\mathcal{X}$ and $0 \leq i < n$, the "i-th projection" reflexive graph functor $\mathsf{pr}^n_i : \mathcal{X}^n \to \mathcal{X}$ is defined as follows:*

- $\mathsf{pr}^n_i(l)$ *is the internal functor projecting out the i-th component*
- $\varepsilon_{\mathsf{pr}^n_i} \coloneqq \mathsf{id}_{\mathcal{X}(1)}\left[\mathcal{X}(\mathbf{d})_0 \circ \mathsf{pr}^n_i(0)_0\right]$

Dually, we have the following:

Definition 23. *Given reflexive graph functors* $\mathcal{F}_0, \ldots, \mathcal{F}_{m-1} : \mathcal{X} \to \mathcal{Y}$*, let* $\langle \mathcal{F}_0, \ldots, \mathcal{F}_{m-1} \rangle$ *be the reflexive graph functor from* $\mathcal{X}$ *to* $\mathcal{Y}^m$ *defined as follows:*

- $\langle \mathcal{F}_0, \ldots, \mathcal{F}_{m-1} \rangle(l) \coloneqq \langle \mathcal{F}_0(l), \ldots, \mathcal{F}_{m-1}(l) \rangle$
- $\varepsilon_{\langle \mathcal{F}_0, \ldots, \mathcal{F}_{m-1} \rangle} \coloneqq \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}} \rangle$

Definition 24. *Given reflexive graph functors* $\mathcal{F}_0, \ldots, \mathcal{F}_{m-1}, \mathcal{G}_0, \ldots, \mathcal{G}_{m-1} : \mathcal{X} \to \mathcal{Y}$*, and reflexive graph natural transformations* $\eta_0 : \mathcal{F}_0 \to \mathcal{G}_0, \ldots, \eta_{m-1} : \mathcal{F}_{m-1} \to \mathcal{G}_{m-1}$*, let* $\langle \eta_0, \ldots, \eta_{m-1} \rangle : \langle \mathcal{F}_0, \ldots, \mathcal{F}_{m-1} \rangle \to \langle \mathcal{G}_0, \ldots, \mathcal{G}_{m-1} \rangle$ *be the reflexive graph natural transformation defined by* $\langle \eta_0, \ldots, \eta_{m-1} \rangle(l) \coloneqq \langle \eta_0(l), \ldots, \eta_{m-1}(l) \rangle$*.*

Lemma 25. *We have the following properties:*

1. *The identity reflexive graph functor serves as the identity for the composition of reflexive graph functors.*
2. *The composition of reflexive graph functors is associative.*
3. *The identity reflexive graph natural transformation serves as the identity for the composition of reflexive graph natural transformations.*
4. *The composition of reflexive graph natural transformations is associative.*
5. *The composition* $(-) \circ \mathcal{F}$ *of a reflexive graph functor and a reflexive graph natural transformation is functorial.*
6. *The composition* $\mathcal{G} \circ (-)$ *of a reflexive graph natural transformation and a reflexive graph functor is functorial.*

3 Reflexive Graph Categories with Isomorphisms

As noted above, if we try to interpret a type $\overline{\alpha} \vdash T$ as a reflexive graph functor $[\![T]\!] : \mathcal{X}^n \to \mathcal{X}$ we encounter a problem with contravariance. Specifically, if $\alpha \vdash S$ and $\alpha \vdash T$ are types, then to interpret the function type $\alpha \vdash S \to T$ as the exponential of $[\![S]\!]$ and $[\![T]\!]$, $[\![S \to T]\!](0)$ must map each object X to the exponential $S_0(X) \Rightarrow T_0(X)$ and each morphism $f : X \to Y$ to a morphism from $S_0(X) \Rightarrow T_0(X)$ to $S_0(Y) \Rightarrow T_0(Y)$. The only natural way to construct such a morphism is to furnish morphisms $f_A : S_0(Y) \to S_0(X)$ and $f_B : T_0(X) \to T_0(Y)$. An obvious instantiation is $f_B \coloneqq T_0(f)$, but the same does not work for f_A because $S_0(f)$ goes in the wrong direction. This is a well-known problem that is unrelated to parametricity and, as we can see, already occurs on the "set" level.

The usual solution is to require the domains of the functors interpreting types to be discrete, so that $[\![T]\!] : |\mathcal{X}|^n \to \mathcal{X}$. However, as noted in the introduction, this will not work in our setting. Consider types $\alpha \vdash S(\alpha)$ and $\cdot \vdash T$. By the induction hypothesis, $[\![S]\!] : |\mathcal{X}| \to \mathcal{X}$ and $[\![T]\!] : \mathbf{1} \to \mathcal{X}$ are face map- and degeneracy-preserving reflexive graph functors. The interpretation of the type $\cdot \vdash S[\alpha \coloneqq T]$ should be given by the composition $[\![S]\!] \circ [\![T]\!] : \mathbf{1} \to \mathcal{X}$, and this should be a face map- and degeneracy-preserving functor. While preservation of face maps is easy to prove, preservation of degeneracies poses a problem: we need $S_1(T_1)$ to be isomorphic to the degeneracy $\mathbf{d}(S_0(T_0))$. By assumption, T_1 is isomorphic to the degeneracy $\mathbf{d}(T_0)$, and $S_1(\mathbf{d}(T_0))$ is isomorphic to $\mathbf{d}(S_0(T_0))$, so if we knew that S_1 mapped isomorphic relations to isomorphic relations we would be done. But since the domain of S_1 is $|\mathcal{X}(1)|$, there is no reason that it should preserve non-identity isomorphisms of $\mathcal{X}(1)$.

In this paper we solve the contravariance problem in a different way. We first note that the issue does not arise if $S_0(f)$ is an isomorphism, even a non-identity one. This leads us to require, for each $l \in \{0, 1\}$, a wide subcategory $\mathcal{M}(l) \subseteq \mathcal{X}(l)$ such that every morphism in $\mathcal{M}(l)$ is in fact an isomorphism. Formally:

Definition 26. *Given a reflexive graph category $\mathcal{X}$, a* reflexive graph subcategory *of $\mathcal{X}$ is a reflexive graph category $\mathcal{M}$ together with a reflexive graph "inclusion" functor $\mathcal{I} : \mathcal{M} \to \mathcal{X}$ such that*

- $\mathcal{I}(l)_0$ *and* $\mathcal{I}(l)_1$ *are monic for* $l \in \{0, 1\}$
- $\mathcal{I}(0) \circ \mathcal{M}(\mathbf{f}_\star) = \mathcal{X}(\mathbf{f}_\star) \circ \mathcal{I}(1)$ *for* $\star \in \mathbf{Bool}$
- $\mathcal{I}(1) \circ \mathcal{M}(\mathbf{d}) = \mathcal{X}(\mathbf{d}) \circ \mathcal{I}(1)$

The subcategory $(\mathcal{M}, \mathcal{I})$ is wide *if $\mathcal{I}(l)_0$ is an isomorphism for $l \in \{0, 1\}$.*

The last two conditions in Definition 26 guarantee that $\mathcal{I}$ preserves face maps and degeneracies on the nose. To simplify the presentation, we treat $\mathcal{M}(l)$ as a subcategory of $\mathcal{X}(l)$ and avoid explicit mentions of $\mathcal{I}$ unless otherwise indicated.

Definition 27. *A* reflexive graph category with isomorphisms *is a reflexive graph category $\mathcal{X}$ together with a wide reflexive graph subcategory $(\mathcal{M}, \mathcal{I})$ such that every morphism in $\mathcal{M}(l)$, $l \in \{0, 1\}$, is an isomorphism.*

We view $\mathcal{M}(l)$ as selecting the "good" isomorphisms of $\mathcal{X}(l)$, in the sense that a morphism $f : J \to \mathcal{X}(l)$ is good iff $f^\dagger$ exists. Given a reflexive graph category with isomorphisms $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ we can now interpret a type $\overline{\alpha} \vdash T$ with n free type variables as a reflexive graph functor $[\![T]\!] : \mathcal{M}^n \to \mathcal{M}$. It is important that $[\![T]\!]$ carries (tuples of) good isomorphisms to good isomorphisms: if $[\![T]\!]$ were instead a functor from $\mathcal{M}^n$ to $\mathcal{X}$, then it would not be possible to define substitution (see Definition 31).

One obvious choice is to take $\mathcal{M} := |\mathcal{X}|$. Then $[\![T]\!] : |\mathcal{X}|^n \to |\mathcal{X}|$ and $\varepsilon_{[\![T]\!]}$ is necessarily the identity natural transformation, so $[\![T]\!]$ preserves degeneracies on the nose. This instantiation shows that, despite being motivated by Reynolds' model, for which the Identity Extension Lemma holds only up to isomorphism, our framework can also uniformly subsume strict models of parametricity, for which the Identity Extension Lemma holds on the nose.

Example 28 (PER model, continued)**.** *We take $\mathcal{M} := |\mathcal{R}|$.*

Example 29 (Reynolds' model, continued)**.** *For each l, we take the objects of $\mathcal{M}(l)$ to be the objects of $\mathcal{R}(l)$, and the morphisms of $\mathcal{M}(l)$ to be* all *isomorphisms of $\mathcal{R}(l)$. For example, when $l = 0$,*

$$\begin{aligned}\mathcal{M}(0)_1 := \{(i, j) : \mathsf{Set}_1 \times \mathsf{Set}_1 \;\&\; \\ i_{\mathsf{d}} = j_{\mathsf{c}} \times i_{\mathsf{c}} = j_{\mathsf{d}} \times j \circ i = \mathsf{id} \times i \circ j = \mathsf{id}\}\end{aligned}$$

Here and at several places below we write $a = b$ for $\mathsf{Id}(a, b)$ and $\{x : A \;\&\; B(x)\}$ for $\Sigma_{x:A} B(x)$ to enhance readability. Moreover, $\circ$ and id are composition and identity in the category Set, and we use the subscripts $(\cdot)_{\mathsf{d}}$ and $(\cdot)_{\mathsf{c}}$ to denote the domain and codomain, respectively, of a morphism in Set. The first (or second) projection gives the required mono from $\mathcal{M}(0)_1$ to Set_1.

With this infrastructure in place we can now interpret a term $\overline{\alpha}; x : S \vdash t : T$ as a natural transformation from $\mathcal{I} \circ [\![S]\!]$ to $\mathcal{I} \circ [\![T]\!]$. Importantly, the components of such a natural transformation are drawn from $\mathcal{X}(l)$, rather than just $\mathcal{M}(l)$, as would be the case if we interpreted t as a natural transformation from $[\![S]\!]$ to $[\![T]\!]$. In fact, this latter interpretation would not even be sensible, since not every term gives rise to an isomorphism (most do not).

We want to interpret a type context of length n as the natural number n, types with n free type variables as reflexive graph functors from $\mathcal{M}^n$ to $\mathcal{M}$, and terms with n free type variables as natural transformations between reflexive graph functors with codomain $\mathcal{X}$. Following the standard procedure, we first define, for each n, a category $\mathcal{M}^n \to \mathcal{M}$ to interpret expressions with n free type variables, and then combine these categories using the usual Grothendieck construction. This gives a fibration whose fiber over n is $\mathcal{M}^n \to \mathcal{M}$.

Definition 30. *The category $\mathcal{M}^n \to \mathcal{M}$ is defined as follows:*

- *the objects are face map- and degeneracy-preserving reflexive graph functors from $\mathcal{M}^n$ to $\mathcal{M}$*
- *the morphisms from $\mathcal{F}$ to $\mathcal{G}$ are the face map- and degeneracy-preserving reflexive graph natural transformations from $\mathcal{I} \circ \mathcal{F}$ to $\mathcal{I} \circ \mathcal{G}$*

To move between the fibers we need a notion of substitution:

Definition 31. *For any m-tuple $\mathbf{F} := (\mathcal{F}_0, \ldots, \mathcal{F}_{m-1})$ of objects in $\mathcal{M}^n \to \mathcal{M}$, the functor $\mathbf{F}^*$ from $\mathcal{M}^m \to \mathcal{M}$ to $\mathcal{M}^n \to \mathcal{M}$ is defined by $\mathbf{F}^*(\mathcal{G}) := \mathcal{G} \circ \langle \mathcal{F}_0, \ldots, \mathcal{F}_{m-1} \rangle$ for objects and $\mathbf{F}^*(\eta) := \eta \circ \langle \mathcal{F}_0, \ldots, \mathcal{F}_{m-1} \rangle$ for morphisms.*

When giving a categorical interpretation of System F, a category for interpreting type contexts is also required:

Definition 32. *The* category of contexts $\mathsf{Ctx}(\mathcal{X})$ *is given by:*

- *objects are natural numbers*
- *morphisms from n to m are m-tuples of objects in $\mathcal{M}^n \to \mathcal{M}$*
- *the identity $\mathsf{id}_n : n \to n$ has as its i^{th} component the i^{th} projection functor pr_i^n*
- *given morphisms $\mathbf{F} : n \to m$ and $\mathbf{G} = (\mathcal{G}_1, \ldots, \mathcal{G}_k) : m \to k$, the i^{th} component of the composition $\mathbf{G} \circ \mathbf{F} : n \to k$ is $\mathbf{F}^*(G_i)$*

That this is indeed a category follows from the lemma below:

Lemma 33. *We have the following:*

i) For any morphism $\mathbf{F} = (\mathcal{F}_0, \dots, \mathcal{F}_{m-1}) : n \to m$ *in* $\mathsf{Ctx}(\mathcal{R})$ *and* $0 \leq i < m$, *we have* $\mathbf{F}^\star(\mathsf{pr}_i^m) = \mathcal{F}_i$.

ii) For any natural number n, $(\mathbf{1}_n)^\star$ *is the identity functor on* $|\mathcal{R}|^n \to \mathcal{R}$.

iii) For morphisms $\mathbf{F} : n \to m$, $\mathbf{G} : m \to k$ *in* $\mathsf{Ctx}(\mathcal{R})$, *we have* $(\mathbf{G} \circ \mathbf{F})^\star = \mathbf{F}^\star \circ \mathbf{G}^\star$.

Proof. Parts $i)$, $ii)$ are easy to show. For part $iii)$ let $\mathbf{F} = (\mathcal{F}_0, \dots, \mathcal{F}_{m-1})$ and $\mathbf{G} = (\mathcal{G}_0, \dots, \mathcal{G}_{k-1})$. Fix an object $\mathcal{H}$ in $\mathcal{M}^k \to \mathcal{M}$. The first component of $(\mathbf{G} \circ \mathbf{F})^\star(\mathcal{H})$ is the reflexive graph functor whose component at level l is

$$\mathcal{H}(l) \circ \Big\langle \mathcal{G}_0(l) \circ \langle \mathcal{F}_0(l), \dots, \mathcal{F}_{m-1}(l) \rangle, \dots, \mathcal{G}_{k-1}(l) \circ \langle \mathcal{F}_0(l), \dots, \mathcal{F}_{m-1}(l) \rangle \Big\rangle$$

On the other hand, the first component of $\mathbf{F}^\star(\mathbf{G}^\star(\mathcal{H}))$ is a reflexive graph functor whose component at level l is

$$\mathcal{H}(l) \circ \langle \mathcal{G}_0(l), \dots, \mathcal{G}_{k-1}(l) \rangle \circ \langle \mathcal{F}_0(l), \dots, \mathcal{F}_{m-1}(l) \rangle$$

which is clearly equal to the above. The second component of $(\mathbf{G} \circ \mathbf{F})^\star(\mathcal{H})$ is the morphism

$$\begin{aligned}
&\Big(\mathcal{H}(1)_1 \circ \Big\langle \big(\mathcal{G}_0(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \dots, \varepsilon_{\mathcal{F}_{m-1}} \rangle\big) \circ_{\mathcal{M}(1)} \big(\varepsilon_{\mathcal{G}_0} \circ \langle \mathcal{F}_0(0)_0, \dots, \mathcal{F}_{m-1}(0)_0 \rangle\big), \dots, \\
&\qquad \big(\mathcal{G}_{k-1}(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \dots, \varepsilon_{\mathcal{F}_{m-1}} \rangle\big) \circ_{\mathcal{M}(1)} \big(\varepsilon_{\mathcal{G}_{k-1}} \circ \langle \mathcal{F}_0(0)_0, \dots, \mathcal{F}_{m-1}(0)_0 \rangle\big) \Big\rangle\Big) \circ_{\mathcal{M}(1)} \\
&\Big(\varepsilon_{\mathcal{H}} \circ \Big\langle \mathcal{G}_0(0)_0 \circ \langle \mathcal{F}_0(0)_0, \dots, \mathcal{F}_{m-1}(0)_0 \rangle, \dots, \\
&\qquad \mathcal{G}_{k-1}(0)_0 \circ \langle \mathcal{F}_0(0)_0, \dots, \mathcal{F}_{m-1}(0)_0 \rangle \Big\rangle\Big)
\end{aligned}$$

On the other hand, the second component of $\mathbf{F}^\star(\mathbf{G}^\star(\mathcal{H}))$ is the morphism

$$\begin{aligned}
&\mathcal{H}(1)_1\Big(\Big\langle \mathcal{G}_0(1)_1(\langle \varepsilon_{\mathcal{F}_0}, \dots, \varepsilon_{\mathcal{F}_{m-1}} \rangle), \dots, \mathcal{G}_{k-1}(1)_1(\langle \varepsilon_{\mathcal{F}_0}, \dots, \varepsilon_{\mathcal{F}_{m-1}} \rangle) \Big\rangle\Big) \circ_{\mathcal{M}(1)} \\
&\quad \Big(\Big(\mathcal{H}(1)_1(\langle \varepsilon_{\mathcal{G}_0}, \dots, \varepsilon_{\mathcal{G}_{k-1}} \rangle) \circ_{\mathcal{M}(1)} \big(\varepsilon_{\mathcal{H}} \circ \langle \mathcal{G}_0(0)_0, \dots, \mathcal{G}_{k-1}(0)_0 \rangle\big)\Big) \circ \\
&\qquad \langle \mathcal{F}_0(0)_0, \dots, \mathcal{F}_{m-1}(0)_0 \rangle \Big)
\end{aligned}$$

We now have the chain of equalities in Figure 1, where the first equality follows by definition of $\circ_{\mathcal{M}(1)^m}$; the second one follows by functoriality of $\mathcal{H}(1)$; and the third one follows since $\circ_{\mathcal{M}(1)}$ commutes with precomposition in the ambient category. We use the same color to denote rewriting of equal subexpressions. This finishes the proof that $(\mathbf{G} \circ \mathbf{F})^\star$ and $\mathbf{F}^\star \circ \mathbf{G}^\star$ agree on objects. The proof that they agree on morphisms is easy. □

Defining the product $n \times 1$ in $\mathsf{Ctx}(\mathcal{R})$ to be the natural number sum $n + 1$, we see that $\mathsf{Ctx}(\mathcal{R})$ enjoys sufficient structure to model the construction of System F type contexts:

$$
\begin{aligned}
&\Big(\mathcal{H}(1)_1 \circ \Big\langle \big(\mathcal{G}_0(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle\big) \circ_{\mathcal{M}(1)} \big(\varepsilon_{\mathcal{G}_0} \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\big), \ldots, \\
&\qquad \big(\mathcal{G}_{k-1}(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle\big) \circ_{\mathcal{M}(1)} \big(\varepsilon_{\mathcal{G}_{k-1}} \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\big)\Big\rangle\Big) \circ_{\mathcal{M}(1)} \\
&\quad \Big(\varepsilon_{\mathcal{H}} \circ \Big\langle \mathcal{G}_0(0)_0 \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle, \ldots, \\
&\qquad\qquad \mathcal{G}_{k-1}(0)_0 \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\Big\rangle\Big) \\
\overset{(1)}{=}\; &\Big(\mathcal{H}(1)_1 \circ \Big(\Big\langle \mathcal{G}_0(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle, \ldots, \mathcal{G}_{k-1}(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle\Big\rangle \circ_{\mathcal{M}(1)^n} \\
&\qquad \big(\langle \varepsilon_{\mathcal{G}_0}, \ldots, \varepsilon_{\mathcal{G}_{k-1}}\rangle \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\big)\Big)\Big) \circ_{\mathcal{M}(1)} \\
&\quad \Big(\varepsilon_{\mathcal{H}} \circ \langle \mathcal{G}_0(0)_0, \ldots, \mathcal{G}_{k-1}(0)_0\rangle \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\Big) \\
\overset{(2)}{=}\; &\Big(\mathcal{H}(1)_1 \circ \Big\langle \mathcal{G}_0(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle, \ldots, \mathcal{G}_{k-1}(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle\Big\rangle\Big) \circ_{\mathcal{M}(1)} \\
&\quad \Big(\mathcal{H}(1)_1 \circ \langle \varepsilon_{\mathcal{G}_0}, \ldots, \varepsilon_{\mathcal{G}_{k-1}}\rangle \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\Big) \circ_{\mathcal{M}(1)} \\
&\quad \Big(\varepsilon_{\mathcal{H}} \circ \langle \mathcal{G}_0(0)_0, \ldots, \mathcal{G}_{k-1}(0)_0\rangle \circ \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\Big) \\
\overset{(3)}{=}\; &\Big(\mathcal{H}(1)_1 \circ \Big\langle \mathcal{G}_0(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle, \ldots, \mathcal{G}_{k-1}(1)_1 \circ \langle \varepsilon_{\mathcal{F}_0}, \ldots, \varepsilon_{\mathcal{F}_{m-1}}\rangle\Big\rangle\Big) \circ_{\mathcal{M}(1)} \\
&\quad \Big(\Big(\mathcal{H}(1)_1 \circ \langle \varepsilon_{\mathcal{G}_0}, \ldots, \varepsilon_{\mathcal{G}_{k-1}}\rangle \circ_{\mathcal{M}(1)} \big(\varepsilon_{\mathcal{H}} \circ \langle \mathcal{G}_0(0)_0, \ldots, \mathcal{G}_{k-1}(0)_0\rangle\big)\Big) \circ \\
&\qquad \langle \mathcal{F}_0(0)_0, \ldots, \mathcal{F}_{m-1}(0)_0\rangle\Big)
\end{aligned}
$$

Figure 1: Equalities for The Proof of Lemma 33

Lemma 34. *The category* $\mathsf{Ctx}(\mathcal{R})$ *has a terminal object 0 and products* $(-) \times 1$.

Proof. The product of n and 1 is $n + 1$; the first projection has as its i-th component the "i-th projection functor" pr_i^{n+1} and the second projection has as its sole component the "n-th projection functor" pr_n^{n+1}. □

We can now combine the categories $\mathsf{Ctx}(\mathcal{R})$ and $\mathcal{M}^n \to \mathcal{M}$ analogously to the Grothendieck construction to obtain the total category of our $\lambda 2$-fibration; the category $\mathcal{M}^n \to \mathcal{M}$ emerges as the fiber over the object n of the base category $\mathsf{Ctx}(\mathcal{R})$.

Definition 35. *The category* $\int_n \mathcal{M}^n \to \mathcal{M}$ *is defined as follows:*

- *objects are pairs* $(n, \mathcal{F})$*, where* $\mathcal{F}$ *is an object in* $\mathcal{M}^n \to \mathcal{M}$
- *morphisms from* $(n, \mathcal{F})$ *to* $(m, \mathcal{G})$ *are pairs* $(\mathbf{F}, \eta)$*, where* $\mathbf{F} : n \to m$ *is a morphism in* $\mathsf{Ctx}(\mathcal{X})$ *and* $\eta : \mathcal{F} \to \mathbf{F}^*(\mathcal{G})$ *is a morphism in* $\mathcal{M}^n \to \mathcal{M}$
- *the identity on* $(n, \mathcal{F})$ *is the pair* $(\mathsf{id}_n, \mathsf{id}_{\mathcal{F}})$*, where* $\mathsf{id}_n : n \to n$ *is the identity in* $\mathsf{Ctx}(\mathcal{X})$ *and* $\mathsf{id}_{\mathcal{F}} : \mathcal{F} \to \mathcal{F}$ *is the identity in* $\mathcal{M}^n \to \mathcal{M}$
- *the composition of two morphisms* $(\mathbf{F}, \eta_1) : (n, \mathcal{F}) \to (m, \mathcal{G})$ *and* $(\mathbf{G}, \eta_2) : (m, \mathcal{G}) \to (k, \mathcal{H})$ *is the pair* $(\mathbf{G} \circ \mathbf{F}, \mathbf{F}^*(\eta_2) \circ \eta_1)$*, where the first composition is in* $\mathsf{Ctx}(\mathcal{X})$ *and the second composition is in* $\mathcal{M}^n \to \mathcal{M}$

Since the set of objects of $\mathcal{M}^n \to \mathcal{M}$ is, by definition, (isomorphic to) the set of morphisms from n to 1 in $\mathsf{Ctx}(\mathcal{R})$, we have not only that $\int_n \mathcal{M}^n \to \mathcal{M}$ is the total category of a fibration over $\mathsf{Ctx}(\mathcal{R})$, but that this fibration is actually a split fibration with a split generic object:

Lemma 36. *The forgetful functor from $\int_n \mathcal{M}^n \to \mathcal{M}$ to $\mathsf{Ctx}(\mathcal{R})$ is a split fibration with split generic object 1.*

Proof. Given any morphism $\mathbf{F} : n \to m$ in $\mathsf{Ctx}(\mathcal{R})$ and an object $\mathcal{G}$ in $\mathcal{M}^m \to \mathcal{M}$, the cartesian lifting of $\mathbf{F}$ with respect to $\mathcal{G}$ is defined to be the morphism $(\mathbf{F}, \mathsf{id}_{\mathbf{F}^*(\mathcal{G})}) : (n, \mathbf{F}^*(\mathcal{G})) \to (m, \mathcal{G})$ in the total category $\int_n \mathcal{M}^n \to \mathcal{M}$. The induced reindexing functor is precisely $\mathbf{F}^*$. $\square$

4 Cartesian Closed Reflexive Graph Categories With Isomorphisms

To appropriately interpret arrow types we need the category $\mathcal{M}^n \to \mathcal{M}$ to be cartesian closed. For this we require more structure on the underlying reflexive graph category with isomorphisms. We define:

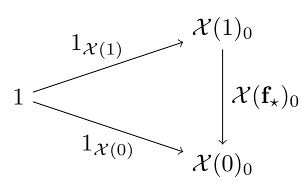
Definition 37. *An internal category C in $\mathcal{C}$* has a terminal object *if it comes equipped with an arrow $1_C : 1 \to C_0$ with the following universal property:*

- *for any object $a : J \to C_0$ (with J arbitrary), there is a unique morphism $!_C(a) : J \to C_1$ such that*

$$\mathsf{s}_C[!_C(a)] = a$$
$$\mathsf{t}_C[!_C(a)] = 1_C \circ !J$$

It is possible to show that the above definition is equivalent to the standard one given *e.g.*, in Section 7.2 of [5]. However, the explicit version will be more useful for us.

Definition 38. *A reflexive graph category $\mathcal{X}$* has terminal objects *if for each $l \in \{0,1\}$ the category $\mathcal{X}(l)$ has a terminal object. The terminal objects are* stable under face maps *if for any $\star \in \mathbf{Bool}$, the canonical morphism witnessing the commutativity of the diagram below is the identity:*



The terminal objects are stable under degeneracies *if the canonical morphism $\eta^1_{\mathcal{X}}$ witnessing the commutativity of the diagram below is an isomorphism:*

$$
\begin{array}{ccc}
 & & \mathcal{X}(0)_0 \\
 & \nearrow^{1_{\mathcal{X}(0)}} & \Big\downarrow \mathcal{X}(\mathbf{d})_0 \\
1 & & \\
 & \searrow_{1_{\mathcal{X}(1)}} & \\
 & & \mathcal{X}(1)_0
\end{array}
$$

Definition 39. *A reflexive graph category* $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ *with isomorphisms* has terminal objects *if* $\mathcal{X}$ *has terminal objects. The terminal objects are* stable under face maps *if the terminal objects in* $\mathcal{X}$ *are stable under face maps. The terminal objects are* stable under degeneracies *if the terminal objects in* $\mathcal{X}$ *are stable under degeneracies and the (iso)morphism* $\eta^1_{\mathcal{X}}$ *is in the image of* $\mathcal{I}(1)$*.*

Lemma 40. *If a reflexive graph category* $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ *with isomorphisms has terminal objects stable under face maps and degeneracies, then for each* n*, the category* $\mathcal{M}^n \to \mathcal{M}$ *has a terminal object.*

Proof. We define the terminal object in $\mathcal{M}^n \to \mathcal{M}$ to be 1_n, where

- $1_n(l)_0 \coloneqq 1_{\mathcal{X}(l)} \circ !(\mathcal{M}(l)^n_0)$
- $1_n(l)_1 \coloneqq \mathsf{id}_{\mathcal{M}(l)}[1_{\mathcal{X}(l)}] \circ !(\mathcal{M}(l)^n_1)$
- $\varepsilon_{1_n} \coloneqq \eta^1_{\mathcal{X}} \circ !(\mathcal{M}(0)^n_0)$

To show that 1_n is indeed a terminal object, take another object $\mathcal{F}$. The universal morphism from $\mathcal{F}$ into our candidate terminal object is the reflexive graph natural transformation whose component at level l is $!_{\mathcal{X}(l)}(\mathcal{F}(l)_0)$. To prove naturality, we need to show that

$$\left(!_{\mathcal{X}(l)}(\mathcal{F}(l)_0) \circ \mathsf{t}_{\mathcal{M}(l)^n}\right) \circ_{\mathcal{X}(l)} \mathcal{F}(l)_1 = 1_n(l)_1 \circ_{\mathcal{X}(l)} \left(!_{\mathcal{X}(l)}(\mathcal{F}(l)_0) \circ \mathsf{s}_{\mathcal{M}(l)^n}\right)$$

The target of both sides is $1_{\mathcal{X}(l)} \circ !(\mathcal{M}(l)^n_1)$ so the equality follows from the universal property of $1_{\mathcal{X}(l)}$. To prove that the candidate universal morphism is degeneracy-preserving, we need to show that

$$\left(!_{\mathcal{X}(1)}(\mathcal{F}(1)_0) \circ \mathcal{M}(\mathbf{d})^n_0\right) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}} = \varepsilon_{1_n} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ !_{\mathcal{X}(0)}(\mathcal{F}(0)_0)\right)$$

The target of both sides is $1_{\mathcal{X}(1)} \circ !(\mathcal{M}(0)^n_0)$ so the equality again follows from the universal property of $1_{\mathcal{X}(1)}$. The preservation of face maps follows by the exact same argument. This shows that our candidate universal morphism is indeed a proper morphism. Its uniqueness is obvious, once again by the universal property of $1_{\mathcal{X}(l)}$. □

Definition 41. *An internal category* C *in* $\mathcal{C}$ has products *if it comes equipped with arrows* $\times_C : C_0 \times C_0 \to C_0$ *and* $\mathsf{fst}_C, \mathsf{snd}_C : C_0 \times C_0 \to C_1$ *such that*

$$\mathsf{s}_C[\mathsf{fst}_C] = \times_C \quad \textit{and} \quad \mathsf{t}_C[\mathsf{fst}_C] = \mathsf{fst}[C_0, C_0]$$
$$\mathsf{s}_C[\mathsf{snd}_C] = \times_C \quad \textit{and} \quad \mathsf{t}_C[\mathsf{snd}_C] = \mathsf{snd}[C_0, C_0]$$

with the following universal property:

- *for any objects $a, b, c : J \to C_0$ and morphisms $f, g : J \to C_1$ (with J arbitrary) such that*

$$\mathsf{s}_C[f] = c \quad \textit{and} \quad \mathsf{t}_C[f] = a$$
$$\mathsf{s}_C[g] = c \quad \textit{and} \quad \mathsf{t}_C[g] = b$$

there is a unique morphism $\langle f, g\rangle_C : J \to C_1$ such that

$$\begin{aligned}
&\mathsf{s}_C[\langle f, g\rangle_C] = c \\
&\mathsf{t}_C[\langle f, g\rangle_C] = a \times_C b \\
&\mathsf{fst}_C[a, b] \circ_C \langle f, g\rangle_C = f \\
&\mathsf{snd}_C[a, b] \circ_C \langle f, g\rangle_C = g
\end{aligned}$$

where we write $a \times_C b$, $\mathsf{fst}_C[a, b]$, $\mathsf{snd}_C[a, b]$ for the arrows $\times_C \circ \langle a, b\rangle$, $\mathsf{fst}_C \circ \langle a, b\rangle$, $\mathsf{snd}_C \circ \langle a, b\rangle$.

If C has products, then we have the following:

- for any objects $a, b, c, d : J \to C_0$ and morphisms $f, g : J \to C_1$ such that

$$\mathsf{s}_C[f] = a \quad \textit{and} \quad \mathsf{t}_C[f] = c$$
$$\mathsf{s}_C[g] = b \quad \textit{and} \quad \mathsf{t}_C[g] = d$$

there exists a unique morphism $f \times_C g : J \to C_1$ such that

$$\begin{aligned}
&\mathsf{s}_C[f \times_C g] = a \times_C b \\
&\mathsf{t}_C[f \times_C g] = c \times_C d \\
&\mathsf{fst}_C[c, d] \circ_C (f \times_C g) = f \circ_C \mathsf{fst}_C[a, b] \\
&\mathsf{snd}_C[c, d] \circ_C (f \times_C g) = g \circ_C \mathsf{snd}_C[a, b]
\end{aligned}$$

Using this observation, it is possible to show that above definition is equivalent to the standard one given *e.g.*, in Section 7.2 of [5].

Definition 42. *A reflexive graph category $\mathcal{X}$* has products *if for each $l \in \{0, 1\}$ the category $\mathcal{X}(l)$ has products. The products are* stable under face maps *if for any $\star \in$ **Bool**, the canonical morphism witnessing the commutativity of the diagram below is the identity:*

$$\begin{array}{ccc}
\mathcal{X}(1)_0 \times \mathcal{X}(1)_0 & \xrightarrow{\times_{\mathcal{X}(1)}} & \mathcal{X}(1)_0 \\
{\scriptstyle \mathcal{X}(\mathbf{f}_\star)_0 \times \mathcal{X}(\mathbf{f}_\star)_0} \downarrow & & \downarrow {\scriptstyle \mathcal{X}(\mathbf{f}_\star)_0} \\
\mathcal{X}(0)_0 \times \mathcal{X}(0)_0 & \xrightarrow[\times_{\mathcal{X}(0)}]{} & \mathcal{X}(0)_0
\end{array}$$

The products are stable under degeneracies *if the canonical morphism $\eta^{\times}_{\mathcal{X}}$ witnessing the commutativity of the diagram below is an isomorphism:*

$$\begin{array}{ccc}
\mathcal{X}(0)_0 \times \mathcal{X}(0)_0 & \xrightarrow{\times_{\mathcal{X}(0)}} & \mathcal{X}(0)_0 \\
{\scriptstyle \mathcal{X}(\mathbf{d})_0 \times \mathcal{X}(\mathbf{d})_0}\downarrow & & \downarrow{\scriptstyle \mathcal{X}(\mathbf{d})_0} \\
\mathcal{X}(1)_0 \times \mathcal{X}(1)_0 & \xrightarrow[\times_{\mathcal{X}(1)}]{} & \mathcal{X}(1)_0
\end{array}$$

Notation 43. *If $\mathcal{X}$ has products stable under degeneracies, we write $\eta^{\times}_{\mathcal{X}}[a, b]$ for the composition $\eta^{\times}_{\mathcal{X}} \circ \langle a, b\rangle$ whenever $a, b : J \to \mathcal{X}(0)_0$ are two objects.*

If $\mathcal{X}$ has products stable under degeneracies, we have:

- for any objects $a, b : J \to \mathcal{X}(0)_0$, the following diagrams commute:

$$\begin{array}{ccc}
\mathcal{X}(\mathbf{d})_0 \circ (a \times_{\mathcal{X}(1)} b) & \xrightarrow{\eta^{\times}_{\mathcal{X}}[a,b]} & (\mathcal{X}(\mathbf{d})_0 \circ a) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ a) \\
& {\scriptstyle \mathcal{X}(\mathbf{d})_1 \circ \mathsf{fst}_{\mathcal{X}(0)}[a,b]} \searrow & \downarrow{\scriptstyle \mathsf{fst}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ a, \mathcal{X}(\mathbf{d})_0 \circ b]} \\
& & \mathcal{X}(\mathbf{d})_0 \circ a
\end{array}$$

$$\begin{array}{ccc}
\mathcal{X}(\mathbf{d})_0 \circ (a \times_{\mathcal{X}(1)} b) & \xrightarrow{\eta^{\times}_{\mathcal{X}}[a,b]} & (\mathcal{X}(\mathbf{d})_0 \circ a) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ a) \\
& {\scriptstyle \mathcal{X}(\mathbf{d})_1 \circ \mathsf{snd}_{\mathcal{X}(0)}[a,b]} \searrow & \downarrow{\scriptstyle \mathsf{snd}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ a, \mathcal{X}(\mathbf{d})_0 \circ b]} \\
& & \mathcal{X}(\mathbf{d})_0 \circ a
\end{array}$$

- The isomorphism $\eta^{\times}_{\mathcal{X}}$ is coherent, *i.e.*, for any objects $a, b : J \to \mathcal{X}(0)_0$:

$$\mathcal{X}(\mathbf{f}_\star)_1 \circ \eta^{\times}_{\mathcal{X}}[a, b] = \mathsf{id}$$

- The isomorphism $\eta^{\times}_{\mathcal{X}}$ is natural, *i.e.*, for any objects $a, b, c, d : J \to \mathcal{X}(0)_0$ and morphisms $f, g : J \to \mathcal{X}(0)_1$ such that

$$\mathsf{s}_{\mathcal{X}(0)}[f] = a \;\; \textit{and} \;\; \mathsf{t}_{\mathcal{X}(0)}[f] = c$$
$$\mathsf{s}_{\mathcal{X}(0)}[g] = b \;\; \textit{and} \;\; \mathsf{t}_{\mathcal{X}(0)}[g] = d$$

  the following diagram commutes:

$$\begin{array}{ccc}
\mathcal{X}(\mathbf{d})_0 \circ (a \times_{\mathcal{X}(0)} b) & \xrightarrow{\mathcal{X}(\mathbf{d})_1 \circ (f \times_{\mathcal{X}(0)} g)} & \mathcal{X}(\mathbf{d})_0 \circ (c \times_{\mathcal{X}(0)} d) \\
{\scriptstyle \eta^{\times}_{\mathcal{X}}[a,b]} \downarrow & & \downarrow {\scriptstyle \eta^{\times}_{\mathcal{X}}[c,d]} \\
(\mathcal{X}(\mathbf{d})_0 \circ a) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ b) & \xrightarrow[(\mathcal{X}(\mathbf{d})_1 \circ f) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_1 \circ g)]{} & (\mathcal{X}(\mathbf{d})_0 \circ c) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ d)
\end{array}$$

Definition 44. *A reflexive graph category $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ with isomorphisms* has products *if $\mathcal{X}$ has products and for any $f, g : J \to \mathcal{X}(l)_1$, $f \times_{\mathcal{X}(l)} g$ is in the image of $\mathcal{I}(l)$ whenever f and g are. The products are* stable under face maps *if the products in $\mathcal{X}$ are stable under face maps. The products are* stable under degeneracies *if the products in $\mathcal{X}$ are stable under degeneracies and the (iso)morphism $\eta^{\times}_{\mathcal{X}}$ is in the image of $\mathcal{I}(1)$.*

Lemma 45. *If a reflexive graph category $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ with isomorphisms has products stable under face maps and degeneracies, then for each n, the category $\mathcal{M}^n \to \mathcal{M}$ has products.*

Proof. Fix $\mathcal{F}$ and $\mathcal{G}$ in $\mathcal{M}^n \to \mathcal{M}$. We define $\mathcal{F} \times \mathcal{G}$ by:

- $(\mathcal{F} \times \mathcal{G})(l)_0 \coloneqq \mathcal{F}(l)_0 \times_{\mathcal{X}(l)} \mathcal{G}(l)_0$
- $(\mathcal{F} \times \mathcal{G})(l)_1 \coloneqq \mathcal{F}(l)_1 \times_{\mathcal{X}(l)} \mathcal{G}(l)_1$
- $\varepsilon_{\mathcal{F} \times \mathcal{G}} \coloneqq (\varepsilon_{\mathcal{F}} \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}) \circ_{\mathcal{M}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]$

The first projection out of $\mathcal{F} \times \mathcal{G}$ is defined as the reflexive graph natural transformation whose component at level l is $\mathsf{fst}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0]$. To prove naturality – with respect to $\mathcal{F} \times \mathcal{G}$ and $\mathcal{G}$ – we observe the following chain of equalities:

$$\begin{aligned}
&\big(\mathsf{fst}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0] \circ \mathsf{t}_{\mathcal{M}(l)^n}\big) \circ_{\mathcal{X}(l)} (\mathcal{F}(l)_1 \times_{\mathcal{X}(l)} \mathcal{G}(l)_1) \\
&= \mathsf{fst}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\big] \circ_{\mathcal{X}(l)} (\mathcal{F}(l)_1 \times_{\mathcal{X}(l)} \mathcal{G}(l)_1) \\
&= \mathcal{F}(l)_1 \circ_{\mathcal{X}(l)} \mathsf{fst}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}\big] \\
&= \mathcal{F}(l)_1 \circ_{\mathcal{X}(l)} \big(\mathsf{fst}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big)
\end{aligned}$$

The first and third equalities are clear and the second follows by the definition of $\times_{\mathcal{X}(l)}$ on morphisms. To prove degeneracy-preservation – with respect to $\varepsilon_{\mathcal{F} \times \mathcal{G}}$ and $\varepsilon_{\mathcal{G}}$ – we observe the following chain of equalities:

$$\begin{aligned}
&\big(\mathsf{fst}_{\mathcal{X}(1)}[\mathcal{F}(1)_0, \mathcal{G}(1)_0] \circ \mathcal{M}(\mathbf{d})_0^n\big) \circ_{\mathcal{X}(1)} \big(\varepsilon_{\mathcal{F}} \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\big) \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \\
&= \mathsf{fst}_{\mathcal{X}(1)}\big[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(l)_0 \circ \mathcal{M}(\mathbf{d})_0^n\big] \circ_{\mathcal{X}(1)} (\varepsilon_{\mathcal{F}} \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}) \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \\
&= \varepsilon_{\mathcal{F}} \circ_{\mathcal{X}(1)} \mathsf{fst}_{\mathcal{X}(1)}\big[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0, \mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0\big] \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \\
&= \varepsilon_{\mathcal{F}} \circ_{\mathcal{X}(1)} \big(\mathcal{X}(\mathbf{d})_1 \circ \mathsf{fst}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]\big)
\end{aligned}$$

The first equality is clear, the second follows by definition of $\times_{\mathcal{X}(1)}$ on morphisms, and the third follows by definition of $\eta^{\times}_{\mathcal{X}}$. The preservation of face maps follows by the exact same argument. This shows that the first projection is indeed a proper morphism. The second projection is defined analogously.

To show that $\mathcal{F} \times \mathcal{G}$ with the aforementioned projections is indeed a product, fix $\mathcal{H}$ and $\eta_{\mathcal{F}} : \mathcal{H} \to \mathcal{F}$, $\eta_{\mathcal{G}} : \mathcal{H} \to \mathcal{G}$. The universal morphism from $\mathcal{H}$ into $\mathcal{F} \times \mathcal{G}$ is the reflexive graph natural transformation whose component at level l is $\langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)}$. To show naturality – with respect to $\mathcal{H}$ and $\mathcal{F} \times \mathcal{G}$ – we need to establish the equality

$$\big(\langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)} \circ \mathsf{t}_{\mathcal{M}(l)^n}\big) \circ_{\mathcal{X}(l)} \mathcal{H}(l)_1 = \\ (\mathcal{F}(l)_1 \times_{\mathcal{X}(l)} \mathcal{G}(l)_1) \circ_{\mathcal{X}(l)} \big(\langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)} \circ \mathsf{s}_{\mathcal{M}(l)^n}\big)$$

The target of the two morphisms is a product, so it suffices to check that their compositions with the first and second projections coincide. The chain of equalities below establishes this for the first projection. Equalities (1) and (5) are clear; equalities (2) and (4) follow by the definition of $\langle \cdot, \cdot \rangle_{\mathcal{X}(l)}$; equality (3) follows from the naturality of $\eta_{\mathcal{F}}$; and equality (6) follows by the definition of $\times_{\mathcal{X}(l)}$ on morphisms. The case of the second projection is entirely analogous.

$$\begin{aligned}
&\mathsf{fst}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\big] \circ_{\mathcal{X}(l)} \big(\langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)} \circ \mathsf{t}_{\mathcal{M}(l)^n}\big) \circ_{\mathcal{X}(l)} \mathcal{H}(l)_1 \\
\overset{(1)}{=}\ & \Big(\big(\mathsf{fst}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0] \circ_{\mathcal{X}(l)} \langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)}\big) \circ \mathsf{t}_{\mathcal{M}(l)^n}\Big) \circ_{\mathcal{X}(l)} \mathcal{H}(l)_1 \\
\overset{(2)}{=}\ & (\eta_{\mathcal{F}}(l) \circ \mathsf{t}_{\mathcal{M}(l)^n}) \circ_{\mathcal{X}(l)} \mathcal{H}(l)_1 \\
\overset{(3)}{=}\ & \mathcal{F}(l)_1 \circ_{\mathcal{X}(l)} (\eta_{\mathcal{F}}(l) \circ \mathsf{s}_{\mathcal{M}(l)^n}) \\
\overset{(4)}{=}\ & \mathcal{F}(l)_1 \circ_{\mathcal{X}(l)} \Big(\big(\mathsf{fst}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0] \circ_{\mathcal{X}(l)} \langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)}\big) \circ \mathsf{s}_{\mathcal{M}(l)^n}\Big) \\
\overset{(5)}{=}\ & \mathcal{F}(l)_1 \circ_{\mathcal{X}(l)} \mathsf{fst}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}\big] \circ_{\mathcal{X}(l)} \big(\langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)} \circ \mathsf{s}_{\mathcal{M}(l)^n}\big) \\
\overset{(6)}{=}\ & \mathsf{fst}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\big] \circ_{\mathcal{X}(l)} (\mathcal{F}(l)_1 \times_{\mathcal{X}(l)} \mathcal{G}(l)_1) \circ_{\mathcal{X}(l)} \\
& \quad \big(\langle \eta_{\mathcal{F}}(l), \eta_{\mathcal{G}}(l) \rangle_{\mathcal{X}(l)} \circ \mathsf{s}_{\mathcal{M}(l)^n}\big)
\end{aligned}$$

To prove that our candidate universal morphism is degeneracy-preserving – with respect to $\varepsilon_{\mathcal{H}}$ and $\varepsilon_{\mathcal{F}\times\mathcal{G}}$ – we need to establish the equality

$$\big(\langle \eta_{\mathcal{F}}(1), \eta_{\mathcal{G}}(1) \rangle_{\mathcal{X}(1)} \circ \mathcal{M}(\mathbf{d})_0^n\big) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{H}} = \\ (\varepsilon_{\mathcal{F}} \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}) \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(1)} \big(\mathcal{X}(\mathbf{d})_1 \circ \langle \eta_{\mathcal{F}}(0), \eta_{\mathcal{G}}(0) \rangle_{\mathcal{X}(0)}\big)$$

Again the target of the two morphisms is a product so it suffices to check that their compositions with the first and second projections coincide. The chain of equalities below establishes this for the first projection. Equality (1) is clear; equalities (2) and (4) follow by the definition of $\langle \cdot, \cdot \rangle_{\mathcal{X}(l)}$; equality (3) follows by the degeneracy-preservation of $\eta_{\mathcal{F}}$; equality (5) follows by the functoriality of $\mathcal{X}(\mathbf{d})$; equality (6) follows by the definition of $\eta^{\times}_{\mathcal{X}}$; and equality (7) follows by the definition of $\times_{\mathcal{X}(1)}$ on morphisms. The case of the second projection is entirely analogous, which shows that our candidate universal morphism is degeneracy-preserving. The preservation of face maps is shown by the exact same argument. Thus our candidate universal morphism is indeed a proper morphism. Its universality and uniqueness are obvious, again by the universal property of $\times_{\mathcal{X}(l)}$.

$$\begin{aligned}
&\mathsf{fst}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n] \circ_{\mathcal{X}(1)} \left(\langle \eta_{\mathcal{F}}(1), \eta_{\mathcal{G}}(1)\rangle_{\mathcal{X}(1)} \circ \mathcal{M}(\mathbf{d})_0^n\right) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{H}} \\
\overset{(1)}{=}\ & \left(\left(\mathsf{fst}_{\mathcal{X}(1)}[\mathcal{F}(1)_0, \mathcal{G}(1)_0] \circ_{\mathcal{X}(1)} \langle \eta_{\mathcal{F}}(1), \eta_{\mathcal{G}}(1)\rangle_{\mathcal{X}(1)}\right) \circ \mathcal{M}(\mathbf{d})_0^n\right) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{H}} \\
\overset{(2)}{=}\ & (\eta_{\mathcal{F}}(1) \circ \mathcal{M}(\mathbf{d})_0^n) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{H}} \\
\overset{(3)}{=}\ & \varepsilon_{\mathcal{F}} \circ_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_1 \circ \eta_{\mathcal{F}}(0)) \\
\overset{(4)}{=}\ & \varepsilon_{\mathcal{F}} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \left(\mathsf{fst}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(0)} \langle \eta_{\mathcal{F}}(0), \eta_{\mathcal{G}}(0)\rangle_{\mathcal{X}(0)}\right)\right) \\
\overset{(5)}{=}\ & \varepsilon_{\mathcal{F}} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \mathsf{fst}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]\right) \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \langle \eta_{\mathcal{F}}(0), \eta_{\mathcal{G}}(0)\rangle_{\mathcal{X}(0)}\right) \\
\overset{(6)}{=}\ & \varepsilon_{\mathcal{F}} \circ_{\mathcal{X}(1)} \mathsf{fst}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0, \mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0] \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(1)} \\
& \quad \left(\mathcal{X}(\mathbf{d})_1 \circ \langle \eta_{\mathcal{F}}(0), \eta_{\mathcal{G}}(0)\rangle_{\mathcal{X}(0)}\right) \\
\overset{(7)}{=}\ & \mathsf{fst}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n] \circ_{\mathcal{X}(1)} (\varepsilon_{\mathcal{F}} \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}) \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(1)} \\
& \quad \left(\mathcal{X}(\mathbf{d})_1 \circ \langle \eta_{\mathcal{F}}(0), \eta_{\mathcal{G}}(0)\rangle_{\mathcal{X}(0)}\right)
\end{aligned}$$

□

Definition 46. *An internal category C in $\mathcal{C}$ with products* has exponentials *if it comes equipped with arrows $\Rightarrow_C : C_0 \times C_0 \to C_0$ and $\mathsf{eval}_C : C_0 \times C_0 \to C_1$ such that*

$$\mathsf{s}_C[\mathsf{eval}_C] = \left((\Rightarrow_C) \times_C \mathsf{fst}[C_0, C_0]\right) \quad \textit{and} \quad \mathsf{t}_C[\mathsf{eval}_C] = \mathsf{snd}[C_0, C_0]$$

with the following universal property:

- *for any objects $a, b, c : J \to C_0$ and morphism $f : J \to C_1$ (with J arbitrary) such that*

$$\mathsf{s}_C[f] = c \times_C a \quad \textit{and} \quad \mathsf{t}_C[f] = b$$

  there is a unique morphism $\lambda_C[a, b, c, f] : J \to C_1$ such that

$$\begin{aligned}
&\mathsf{s}_C[\lambda_C[a, b, c, f]] = c \\
&\mathsf{t}_C[\lambda_C[a, b, c, f]] = (a \Rightarrow_C b) \\
&\mathsf{eval}_C[a, b] \circ_C \left(\lambda_C[a, b, c, f] \times_C \mathsf{id}_C[a]\right) = f
\end{aligned}$$

where we write $a \Rightarrow_C b$, $\mathsf{eval}_C[a, b]$ for the arrows $(\Rightarrow_C \circ \langle a, b\rangle)$, $\mathsf{eval}_C \circ \langle a, b\rangle$.

If C has exponentials, then we have the following:

- for any objects $a, b, c, d : J \to C_0$ and morphisms $f, g : J \to C_1$ such that

$$\begin{aligned}
&\mathsf{s}_C[f] = c \quad \textit{and} \quad \mathsf{t}_C[f] = a \\
&\mathsf{s}_C[g] = b \quad \textit{and} \quad \mathsf{t}_C[g] = d
\end{aligned}$$

  there exists a unique morphism $f \Rightarrow_C g : J \to C_1$ such that

$$\begin{aligned}
&\mathsf{s}_C[f \Rightarrow_C g] = (a \Rightarrow_C b) \\
&\mathsf{t}_C[f \Rightarrow_C g] = (c \Rightarrow_C d) \\
&\mathsf{eval}_C[c, d] \circ_C \left((f \Rightarrow_C g) \times_C \mathsf{id}_C[c]\right) = \\
&\quad g \circ_C \mathsf{eval}_C[a, b] \circ_C (\mathsf{id}_C[a \Rightarrow_C b] \times_C f)
\end{aligned}$$

Using this observation, it is possible to show that above definition is equivalent to the standard one given *e.g.*, in Section 7.2 of [5].

Definition 47. *A reflexive graph category $\mathcal{X}$ with products* has exponentials *if for each $l \in \{0,1\}$, the category $\mathcal{X}(l)$ has exponentials. Assuming the products are stable under face maps, we say the exponentials are* stable under face maps *if for any $\star \in \mathbf{Bool}$, the canonical morphism witnessing the commutativity of the diagram below is the identity:*

$$\begin{array}{ccc}
\mathcal{X}(1)_0 \times \mathcal{X}(1)_0 & \xrightarrow{\Rightarrow_{\mathcal{X}(1)}} & \mathcal{X}(1)_0 \\
{\scriptstyle \mathcal{X}(\mathbf{f}_\star)_0 \times \mathcal{X}(\mathbf{f}_\star)_0}\downarrow & & \downarrow{\scriptstyle \mathcal{X}(\mathbf{f}_\star)_0} \\
\mathcal{X}(0)_0 \times \mathcal{X}(0)_0 & \xrightarrow[\Rightarrow_{\mathcal{X}(0)}]{} & \mathcal{X}(0)_0
\end{array}$$

Assuming the products are stable under degeneracies, we say the exponentials are stable under degeneracies *if the canonical morphism $\eta^{\Rightarrow}_{\mathcal{X}}$ witnessing the commutativity of the diagram below is an isomorphism:*

$$\begin{array}{ccc}
\mathcal{X}(0)_0 \times \mathcal{X}(0)_0 & \xrightarrow{\Rightarrow_{\mathcal{X}(0)}} & \mathcal{X}(0)_0 \\
{\scriptstyle \mathcal{X}(\mathbf{d})_0 \times \mathcal{X}(\mathbf{d})_0}\downarrow & & \downarrow{\scriptstyle \mathcal{X}(\mathbf{d})_0} \\
\mathcal{X}(1)_0 \times \mathcal{X}(1)_0 & \xrightarrow[\Rightarrow_{\mathcal{X}(1)}]{} & \mathcal{X}(1)_0
\end{array}$$

Notation 48. *If $\mathcal{X}$ has exponentials stable under degeneracies, we write $\eta^{\Rightarrow}_{\mathcal{X}}[a,b]$ for the composition $\eta^{\Rightarrow}_{\mathcal{X}} \circ \langle a, b\rangle$ whenever $a, b : J \to \mathcal{X}(l)_0$ are two objects.*

If $\mathcal{X}$ has products and exponentials stable under degeneracies, we have:

- for any objects $a, b : J \to \mathcal{X}(0)_0$, the following diagram commutes:

$$\begin{array}{ccc}
\mathcal{X}(\mathbf{d})_0 \circ \big((a \Rightarrow_{\mathcal{X}(0)} b) \times_{\mathcal{X}(0)} a\big) & \xrightarrow{\eta^{\times}_{\mathcal{X}}[a \Rightarrow_{\mathcal{X}(0)} b,\, a]} & \big(\mathcal{X}(\mathbf{d})_0 \circ (a \Rightarrow_{\mathcal{X}(1)} b)\big) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ a) \\
 & & \downarrow{\scriptstyle \eta^{\Rightarrow}_{\mathcal{X}}[a,b] \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ a]} \\
{\scriptstyle \mathcal{X}(\mathbf{d})_1 \circ \mathsf{eval}_{\mathcal{X}(0)}[a,b]}\downarrow & & \big((\mathcal{X}(\mathbf{d}) \circ a) \Rightarrow_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d}) \circ b)\big) \times_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ a) \\
 & \swarrow{\scriptstyle \mathsf{eval}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ a,\, \mathcal{X}(\mathbf{d})_0 \circ b]} & \\
\mathcal{X}(\mathbf{d})_0 \circ b & &
\end{array}$$

- The isomorphism $\eta^{\Rightarrow}_{\mathcal{X}}$ is coherent, *i.e.*, for any objects $a, b : J \to \mathcal{X}(0)_0$, the following holds:
$$\mathcal{X}(\mathbf{f}_\star)_1 \circ \eta^{\Rightarrow}_{\mathcal{X}}[a, b] = \mathsf{id}$$

- The isomorphism $\eta^{\Rightarrow}_{\mathcal{X}}$ is natural, *i.e.*, for any objects $a, b, c, d : J \to \mathcal{X}(0)_0$ and morphisms $f, g : J \to \mathcal{X}(0)_1$ such that
$$\mathsf{s}_{\mathcal{X}(0)}[f] = a \;\; and \;\; \mathsf{t}_{\mathcal{X}(0)}[f] = c$$
$$\mathsf{s}_{\mathcal{X}(0)}[g] = b \;\; and \;\; \mathsf{t}_{\mathcal{X}(0)}[g] = d$$

  the following diagram commutes:

$$\begin{array}{ccc}
\mathcal{X}(\mathbf{d})_0 \circ (a \Rightarrow_{\mathcal{X}(0)} b) & \xrightarrow{\mathcal{X}(\mathbf{d})_1 \circ (f \Rightarrow_{\mathcal{X}(0)} g)} & \mathcal{X}(\mathbf{d})_0 \circ (c \Rightarrow_{\mathcal{X}(0)} d) \\
\Big\downarrow {\scriptstyle \eta^{\Rightarrow}_{\mathcal{X}}[a,b]} & & \Big\downarrow {\scriptstyle \eta^{\Rightarrow}_{\mathcal{X}}[c,d]} \\
(\mathcal{X}(\mathbf{d})_0 \circ a) \Rightarrow_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ b) & \xrightarrow[(\mathcal{X}(\mathbf{d})_1 \circ f) \Rightarrow_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_1 \circ g)]{} & (\mathcal{X}(\mathbf{d})_0 \circ c) \Rightarrow_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ d)
\end{array}$$

Definition 49. *A reflexive graph category $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ with isomorphisms and products* has exponentials *if $\mathcal{X}$ has exponentials and for any $f, g : J \to \mathcal{X}(l)_1$, $f \Rightarrow_{\mathcal{X}(l)} g$ is in the image of $\mathcal{I}(l)$ whenever f and g are. Assuming the products are stable under face maps, we say the exponentials are* stable under face maps *if the exponentials in $\mathcal{X}$ are stable under face maps. Assuming the products are stable under degeneracies, we say the exponentials are* stable under degeneracies *if the exponentials in $\mathcal{X}$ are stable under degeneracies and the (iso)morphism $\eta^{\Rightarrow}_{\mathcal{X}}$ is in the image of $\mathcal{I}(1)$.*

Lemma 50. *If a reflexive graph category $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$ with isomorphisms has products and exponentials stable under face maps and degeneracies, then for each n, the category $\mathcal{M}^n \to \mathcal{M}$ has exponentials.*

Proof. Fix $\mathcal{F}$ and $\mathcal{G}$ in $\mathcal{M}^n \to \mathcal{M}$. We define $\mathcal{F} \Rightarrow \mathcal{G}$ by:

- $(\mathcal{F} \Rightarrow \mathcal{G})(l)_0 \coloneqq \mathcal{F}(l)_0 \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_0$
- $(\mathcal{F} \Rightarrow \mathcal{G})(l)_1 \coloneqq \mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1$
- $\varepsilon_{\mathcal{F} \Rightarrow \mathcal{G}} \coloneqq (\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}) \circ_{\mathcal{M}(1)} \eta^{\Rightarrow}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]$

The evaluation morphism for $\mathcal{F} \Rightarrow \mathcal{G}$ is defined as the reflexive graph natural transformation whose component at level l is $\mathsf{eval}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0]$. To prove naturality – with respect to $(\mathcal{F} \Rightarrow \mathcal{G}) \times \mathcal{F}$ and $\mathcal{G}$ – we observe the chain of equalities below. The first and third equalities follow since $\times_{\mathcal{X}(1)}$ suitably commutes with $\circ_{\mathcal{X}(1)}$; the second equality follows by definition of $\Rightarrow_{\mathcal{X}(l)}$ on morphisms; and the fourth equality follows since the product of identities is again an identity.

$$
\begin{aligned}
&\left(\mathsf{eval}_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0] \circ \mathsf{t}_{\mathcal{M}(l)^n}\right) \circ_{\mathcal{X}(l)} \Big(\big(\mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1\big) \times_{\mathcal{X}(l)} \mathcal{F}(l)_1 \Big) \\
\overset{(1)}{=}\ & \mathsf{eval}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\big] \circ_{\mathcal{X}(l)} \\
& \Big(\big(\mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1\big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}] \Big) \circ_{\mathcal{X}(l)} \\
& \Big(\mathsf{id}_{\mathcal{X}(l)}\big[(\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}) \Rightarrow_{\mathcal{X}(l)} (\mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n})\big] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1 \Big) \\
\overset{(2)}{=}\ & \mathcal{G}(l)_1 \circ_{\mathcal{X}(l)} \mathsf{eval}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}\big] \circ_{\mathcal{X}(l)} \\
& \Big(\mathsf{id}_{\mathcal{X}(l)}\big[(\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}) \Rightarrow_{\mathcal{X}(l)} (\mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n})\big] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1^{-1} \Big) \circ_{\mathcal{X}(l)} \\
& \Big(\mathsf{id}_{\mathcal{X}(l)}\big[(\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}) \Rightarrow_{\mathcal{X}(l)} (\mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n})\big] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1 \Big) \\
\overset{(3)}{=}\ & \mathcal{G}(l)_1 \circ_{\mathcal{X}(l)} \left(\mathsf{eval}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0, \mathcal{G}(l)_0\big] \circ \mathsf{s}_{\mathcal{M}(l)^n}\right) \circ_{\mathcal{X}(l)} \\
& \Big(\mathsf{id}_{\mathcal{X}(l)}\big[(\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}) \Rightarrow_{\mathcal{X}(l)} (\mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n})\big] \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}] \Big) \\
\overset{(4)}{=}\ & \mathcal{G}(l)_1 \circ_{\mathcal{X}(l)} \left(\mathsf{eval}_{\mathcal{X}(l)}\big[\mathcal{F}(l)_0, \mathcal{G}(l)_0\big] \circ \mathsf{s}_{\mathcal{M}(l)^n}\right)
\end{aligned}
$$

To prove the degeneracy-preservation of the evaluation morphism – with respect to $\varepsilon_{(\mathcal{F} \Rightarrow \mathcal{G}) \times \mathcal{F}}$ and $\varepsilon_{\mathcal{G}}$ – we observe the chain of equalities below.

$$
\begin{aligned}
&\left(\mathsf{eval}_{\mathcal{X}(1)}[\mathcal{F}(1)_0, \mathcal{G}(1)_0] \circ \mathcal{M}(\mathbf{d})_0^n\right) \circ_{\mathcal{X}(1)} \\
& \bigg(\Big(\big(\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\big) \circ_{\mathcal{X}(1)} \eta_{\mathcal{X}}^{\Rightarrow}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \Big) \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}} \bigg) \circ_{\mathcal{X}(1)} \\
& \eta_{\mathcal{X}}^{\times}\big[\mathcal{F}(0)_0 \Rightarrow_{\mathcal{X}(0)} \mathcal{G}(0)_0, \mathcal{F}(0)_0\big] \\
\overset{(1)}{=}\ & \mathsf{eval}_{\mathcal{X}(1)}\big[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n\big] \circ_{\mathcal{X}(1)} \\
& \Big(\big(\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\big) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n] \Big) \circ_{\mathcal{X}(1)} \\
& \left(\eta_{\mathcal{X}}^{\Rightarrow}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \\
& \eta_{\mathcal{X}}^{\times}\big[\mathcal{F}(0)_0, \mathcal{F}(0) \Rightarrow_{\mathcal{X}(0)} \mathcal{G}(0)_0\big] \\
\overset{(2)}{=}\ & \varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \mathsf{eval}_{\mathcal{X}(1)}\big[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0, \mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0\big] \circ_{\mathcal{X}(1)} \\
& \Big(\mathsf{id}_{\mathcal{X}(1)}\big[(\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0) \Rightarrow_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0)\big] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}^{-1} \Big) \circ_{l_2} \\
& \left(\eta_{\mathcal{X}}^{\Rightarrow}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \\
& \eta_{\mathcal{X}}^{\times}\big[\mathcal{F}(0)_0, \mathcal{F}(0) \Rightarrow_{\mathcal{X}(0)} \mathcal{G}(0)_0\big] \\
\overset{(3)}{=}\ & \varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \mathsf{eval}_{\mathcal{X}(1)}\big[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0, \mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0\big] \circ_{\mathcal{X}(1)} \\
& \Big(\eta_{\mathcal{X}}^{\Rightarrow}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0] \Big) \circ_{\mathcal{X}(1)} \\
& \eta_{\mathcal{X}}^{\times}\big[\mathcal{F}(0)_0, \mathcal{F}(0) \Rightarrow_{\mathcal{X}(0)} \mathcal{G}(0)_0\big] \\
\overset{(4)}{=}\ & \varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \mathsf{eval}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]\right)
\end{aligned}
$$

The first and third equalities follow since $\times_{\mathcal{X}(1)}$ suitably commutes with $\circ_{\mathcal{X}(1)}$; the second equality follows by definition of $\Rightarrow_{\mathcal{X}(1)}$ on morphisms; and the fourth equality follows by definition of $\eta^{\rightrightarrows}_{\mathcal{X}}$. The preservation of face maps follows by the exact same argument. This shows that the evaluation morphism is indeed a proper morphism.

To show that $\mathcal{F} \Rightarrow \mathcal{G}$ with the aforementioned evaluation is an exponential, fix $\mathcal{H}$ and $\eta : \mathcal{F} \times \mathcal{H} \to \mathcal{G}$. The universal morphism from $\mathcal{H}$ into $\mathcal{F} \Rightarrow \mathcal{G}$ is the reflexive graph natural transformation whose component at level l is $\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)]$. To show naturality – with respect to $\mathcal{H}$ and $\mathcal{F} \Rightarrow \mathcal{G}$ – we need to establish the equality

$$\begin{aligned}&\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{t}_{\mathcal{M}(l)^n}\big) \circ_{\mathcal{X}(l)} \mathcal{H}(l)_1 = \\ &\quad \big(\mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1\big) \circ_{\mathcal{X}(l)} \big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big)\end{aligned}$$

The target of the two morphisms is an exponential, so it suffices to check that taking a product of each morphism with the identity and postcomposing with the evaluation morphism yields the same result. Moreover, since $\mathsf{id}_{\mathcal{X}(l)}[\mathcal{H}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1$ is an isomorphism, it suffices to show that a further precomposition with this isomorphism yields the same result. To this end we observe the chain of equalities below. Equalities (1), (5), (7), (8), and the green part of (2) follow since $\times_{\mathcal{X}(l)}$ suitably commutes with $\circ_{\mathcal{X}(l)}$; equality (4) and the red part of (2) follow by the definition of $\lambda_{\mathcal{X}(l)}$; equality (3) follows by the degeneracy-preservation of η; and equality (6) follows by the definition of $\Rightarrow_{\mathcal{X}(l)}$.

$$
\begin{aligned}
&\mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \\
&\quad \Big(\Big(\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{t}_{\mathcal{M}(l)^n}\big) \circ_{\mathcal{X}(l)} \mathcal{H}(l)_1\Big) \times_{\mathcal{X}(l)} \\
&\quad \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\Big) \circ_{\mathcal{X}(l)} \big(\mathsf{id}_{\mathcal{X}(l)}[\mathcal{H}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1\big) \\
\overset{(1)}{=}\ &\mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{t}_{\mathcal{M}(l)^n}\big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\Big) \circ_{\mathcal{X}(l)} \\
&\quad \big(\mathcal{H}(l)_1 \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\big) \circ_{\mathcal{X}(l)} \big(\mathsf{id}_{\mathcal{X}(l)}[\mathcal{H}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1\big) \\
\overset{(2)}{=}\ &(\eta(l) \circ \mathsf{t}_{\mathcal{M}(l)^n}) \circ_{\mathcal{X}(l)} (\mathcal{H}(l)_1 \times_{\mathcal{X}(l)} \mathcal{F}(l)_1) \\
\overset{(3)}{=}\ &\mathcal{G}(l)_1 \circ_{\mathcal{X}(l)} (\eta(l) \circ \mathsf{s}_{\mathcal{M}(l)^n}) \\
\overset{(4)}{=}\ &\mathcal{G}(l)_1 \circ_{\mathcal{X}(l)} \mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}]\Big) \\
\overset{(5)}{=}\ &\mathcal{G}(l)_1 \circ_{\mathcal{X}(l)} \mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \\
&\quad \Big(\mathsf{id}_{\mathcal{X}(l)}\big[(\mathcal{F}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}) \Rightarrow_{\mathcal{X}(l)} (\mathcal{G}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n})\big] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1^{-1}\Big) \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big) \times_{\mathcal{X}(l)} \mathcal{F}(l)_1\Big) \\
\overset{(6)}{=}\ &\mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1\big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\Big) \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big) \times_{\mathcal{X}(l)} \mathcal{F}(l)_1\Big) \\
\overset{(7)}{=}\ &\mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1\big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\Big) \circ_{\mathcal{X}(l)} \\
&\quad \Big(\big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\Big) \circ_{\mathcal{X}(l)} \\
&\quad \big(\mathsf{id}_{\mathcal{X}(l)}[\mathcal{H}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1\big) \\
\overset{(8)}{=}\ &\mathsf{eval}_{\mathcal{X}(l)}\left[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}, \mathcal{G}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}\right] \circ_{\mathcal{X}(l)} \Big(\Big(\big(\mathcal{F}(l)_1^{-1} \Rightarrow_{\mathcal{X}(l)} \mathcal{G}(l)_1\big) \circ_{\mathcal{X}(l)} \\
&\quad \big(\lambda_{\mathcal{X}(l)}[\mathcal{F}(l)_0, \mathcal{G}(l)_0, \mathcal{H}(l)_0, \eta(l)] \circ \mathsf{s}_{\mathcal{M}(l)^n}\big)\Big) \times_{\mathcal{X}(l)} \mathsf{id}_{\mathcal{X}(l)}[\mathcal{F}(l)_0 \circ \mathsf{t}_{\mathcal{M}(l)^n}]\Big) \circ_{\mathcal{X}(l)} \\
&\quad \big(\mathsf{id}_{\mathcal{X}(l)}[\mathcal{H}(l)_0 \circ \mathsf{s}_{\mathcal{M}(l)^n}] \times_{\mathcal{X}(l)} \mathcal{F}(l)_1\big)
\end{aligned}
$$

To prove that our candidate universal morphism is degeneracy-preserving – with respect to $\varepsilon_{\mathcal{H}}$ and $\varepsilon_{\mathcal{F} \Rightarrow \mathcal{G}}$ – we need to establish the following equality:

$$\begin{aligned}&\left(\lambda_{\mathcal{X}(1)}[\mathcal{F}(1)_0, \mathcal{G}(1)_0, \mathcal{H}(1)_0, \eta(1)] \circ \mathcal{M}(\mathbf{d})_0^n\right) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{H}} = \\ &\quad \left(\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\right) \circ_{\mathcal{X}(1)} \eta_{\mathcal{X}}^{\Rightarrow}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(1)} \\ &\quad \left(\mathcal{X}(\mathbf{d})_1 \circ \lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)]\right)\end{aligned}$$

The target of the two morphisms is an exponential, so it suffices to check that taking a product of each morphism with the identity and postcomposing with the evaluation morphism yields the same result. Moreover, since

$$\left(\mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{H}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \eta_{\mathcal{X}}^{\times}[\mathcal{F}(0)_0, \mathcal{H}(0)_0]$$

is an isomorphism, it suffices to show that a further precomposition with this isomorphism yields the same result. To this end we observe the two chains of equalities below. Equalities (1), (7), (9), (10), and the green part of (2) follow since $\times_{\mathcal{X}(1)}$ suitably commutes with $\circ_{\mathcal{X}(1)}$; equality (4) and the red part of (2) follow by the definition of $\lambda_{\mathcal{X}(l)}$; equality (3) follows by the degeneracy-preservation of η; equality (5) follows by the functoriality of $\mathcal{X}(\mathbf{d})$; equality (6) follows by the definition of $\eta_{\mathcal{X}}^{\Rightarrow}$; the orange part of equality (8) follows by the definition of $\Rightarrow_{\mathcal{X}(1)}$ on morphisms; and the purple part of equality (8) follows by the naturality of $\eta_{\mathcal{X}}^{\times}$. The preservation of face maps is shown by the exact same argument.

This shows that our candidate universal morphism is a proper morphism. Its universality and uniqueness are obvious by the universal property of $\Rightarrow_{\mathcal{X}(l)}$.

$$\begin{aligned}&\mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n\right] \circ_{\mathcal{X}(1)} \\ &\quad \Bigg(\Big(\left(\lambda_{\mathcal{X}(1)}[\mathcal{F}(1)_0, \mathcal{G}(1)_0, \mathcal{H}(1)_0, \eta(1)] \circ \mathcal{M}(\mathbf{d})_0^n\right) \circ_{\mathcal{X}(1)} \varepsilon_{\mathcal{H}}\Big) \\ &\quad \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Bigg) \circ_{\mathcal{X}(1)} \left(\mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{H}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \\ &\quad \eta_{\mathcal{X}}^{\times}[\mathcal{F}(0)_0, \mathcal{H}(0)_0] \\ \overset{(1)}{=}\ &\mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n\right] \circ_{\mathcal{X}(1)} \\ &\quad \Big(\left(\lambda_{\mathcal{X}(1)}[\mathcal{F}(1)_0, \mathcal{G}(1)_0, \mathcal{H}(1)_0, \eta(1)] \circ \mathcal{M}(\mathbf{d})_0^n\right) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Big) \circ_{\mathcal{X}(1)} \\ &\quad \left(\varepsilon_{\mathcal{H}} \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\right) \circ_{\mathcal{X}(1)} \left(\mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{H}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \\ &\quad \eta_{\mathcal{X}}^{\times}[\mathcal{F}(0)_0, \mathcal{H}(0)_0] \\ \overset{(2)}{=}\ &\left(\eta(1) \circ \mathcal{M}(\mathbf{d})_0^n\right) \circ_{\mathcal{X}(1)} \left(\varepsilon_{\mathcal{H}} \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \eta_{\mathcal{X}}^{\times}[\mathcal{F}(0)_0, \mathcal{H}(0)_0] \\ \overset{(3)}{=}\ &\varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \eta(0)\right) \\ \overset{(4)}{=}\ &\varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \Bigg(\mathcal{X}(\mathbf{d})_1 \circ \Big(\mathsf{eval}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(0)} \\ &\quad \left(\lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)] \times_{\mathcal{X}(0)} \mathsf{id}_{\mathcal{X}(0)}[\mathcal{F}(0)_0]\right)\Big)\Bigg) \\ \overset{(5)}{=}\ &\varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \mathsf{eval}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]\right) \circ_{\mathcal{X}(1)} \\ &\quad \Big(\mathcal{X}(\mathbf{d})_1 \circ \left(\lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)] \times_{\mathcal{X}(0)} \mathsf{id}_{\mathcal{X}(0)}[\mathcal{F}(0)_0]\right)\Big)\end{aligned}$$

$$
\begin{aligned}
&\varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \left(\mathcal{X}(\mathbf{d})_1 \circ \mathsf{eval}_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0]\right) \circ_{\mathcal{X}(1)} \\
&\quad \Big(\mathcal{X}(\mathbf{d})_1 \circ \left(\lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)] \times_{\mathcal{X}(0)} \mathsf{id}_{\mathcal{X}(0)}[\mathcal{F}(0)_0]\right)\Big) \\
\overset{(6)}{=}\ &\varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0, \mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0\right] \circ_{\mathcal{X}(1)} \\
&\quad \left(\eta^{\rightrightarrows}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0]\right) \circ_{\mathcal{X}(1)} \\
&\quad \eta^{\times}_{\mathcal{X}}\left[\mathcal{F}(0)_0 \Rightarrow_{\mathcal{X}(1)} \mathcal{G}(0)_0, \mathcal{F}(0)_0\right] \circ_{\mathcal{X}(1)} \\
&\quad \Big(\mathcal{X}(\mathbf{d})_1 \circ \left(\lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)] \times_{\mathcal{X}(0)} \mathsf{id}_{\mathcal{X}(0)}[\mathcal{F}(0)_0]\right)\Big) \\
\overset{(7)}{=}\ &\varepsilon_{\mathcal{G}} \circ_{\mathcal{X}(1)} \mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0, \mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0\right] \circ_{\mathcal{X}(1)} \\
&\quad \left(\mathsf{id}_{\mathcal{X}(1)}\left[(\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0) \Rightarrow_{\mathcal{X}(1)} (\mathcal{X}(\mathbf{d})_0 \circ \mathcal{G}(0)_0)\right] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}^{-1}\right) \circ_{\mathcal{X}(1)} \\
&\quad \left(\eta^{\rightrightarrows}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \\
&\quad \eta^{\times}_{\mathcal{X}}\left[\mathcal{F}(0)_0 \Rightarrow_{\mathcal{X}(1)} \mathcal{G}(0)_0, \mathcal{F}(0)_0\right] \circ_{\mathcal{X}(1)} \\
&\quad \Big(\mathcal{X}(\mathbf{d})_1 \circ \left(\lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)] \times_{\mathcal{X}(0)} \mathsf{id}_{\mathcal{X}(0)}[\mathcal{F}(0)_0]\right)\Big) \\
\overset{(8)}{=}\ &\mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n\right] \circ_{\mathcal{X}(1)} \\
&\quad \Big(\left(\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\right) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Big) \circ_{\mathcal{X}(1)} \\
&\quad \left(\eta^{\rightrightarrows}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \\
&\quad \Big(\left(\mathcal{X}(\mathbf{d})_1 \circ \lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)]\right) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{F}(0)_0]\Big) \circ_{\mathcal{X}(1)} \\
&\quad \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{H}(0)_0] \\
\overset{(9)}{=}\ &\mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n\right] \circ_{\mathcal{X}(1)} \\
&\quad \Big(\left(\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\right) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Big) \circ_{\mathcal{X}(1)} \\
&\quad \Big(\eta^{\rightrightarrows}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Big) \circ_{\mathcal{X}(1)} \\
&\quad \Big(\left(\mathcal{X}(\mathbf{d})_1 \circ \lambda_{\mathcal{X}(0)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)]\right) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Big) \circ_{\mathcal{X}(1)} \\
&\quad \left(\mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{H}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{H}(0)_0] \\
\overset{(10)}{=}\ &\mathsf{eval}_{\mathcal{X}(1)}\left[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n, \mathcal{G}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n\right] \circ_{\mathcal{X}(1)} \\
&\quad \Bigg(\Big(\left(\varepsilon_{\mathcal{F}}^{-1} \Rightarrow_{\mathcal{X}(1)} \varepsilon_{\mathcal{G}}\right) \circ_{\mathcal{X}(1)} \eta^{\rightrightarrows}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{G}(0)_0] \circ_{\mathcal{X}(1)} \\
&\quad \left(\mathcal{X}(\mathbf{d})_1 \circ \lambda_{\mathcal{X}(1)}[\mathcal{F}(0)_0, \mathcal{G}(0)_0, \mathcal{H}(0)_0, \eta(0)]\right)\Big) \times_{\mathcal{X}(1)} \mathsf{id}_{\mathcal{X}(1)}[\mathcal{F}(1)_0 \circ \mathcal{M}(\mathbf{d})_0^n]\Bigg) \circ_{\mathcal{X}(1)} \\
&\quad \left(\mathsf{id}_{\mathcal{X}(1)}[\mathcal{X}(\mathbf{d})_0 \circ \mathcal{H}(0)_0] \times_{\mathcal{X}(1)} \varepsilon_{\mathcal{F}}\right) \circ_{\mathcal{X}(1)} \eta^{\times}_{\mathcal{X}}[\mathcal{F}(0)_0, \mathcal{H}(0)_0]
\end{aligned}
$$

□

Definition 51. *A reflexive graph category with isomorphisms is* cartesian closed *if it has terminal objects, products, and exponentials, all stable under face maps and degeneracies.*

Example 52. *[PER model, continued] Terminal objects, products, and exponentials are defined for $\mathcal{R}$ in the obvious ways, inheriting from the corresponding constructs on PERs. It is not hard to check that all of these constructs are preserved on the nose by the two face maps (projections) and the degeneracy (equality functor), and thus, in our terminology, are stable under face maps and degeneracies.*

Example 53. *[Reynolds' model, continued] Here, too, terminal objects, products, and exponentials are defined for $\mathcal{R}$ in the obvious ways, relating two pairs iff their first and second components are related, and two functions iff they map related arguments to related results. It is easy to see that all of these constructs are preserved on the nose (*i.e.*, up to* definitional *equality) by the projections, and thus are stable under face maps. Unlike in the PER model though, they are only preserved by the equality functor* Eq *up to (the canonical) isomorphism. For example, as discussed just after Definition 11, the two types $\mathsf{Id}((a,b),(c,d))$ and $\mathsf{Id}(a,c) \times \mathsf{Id}(b,d)$ for $(a,b),(c,d) : A \times B$ are not identical, although they are isomorphic under the canonical (iso)morphism $\eta^{\times}[A,B] : \mathsf{Eq}(A \times B) \to \mathsf{Eq}(A) \times \mathsf{Eq}(B)$. A similar situation arises for function types $A \to B$: by function extensionality, $\mathsf{Id}(f,g)$ and $\Pi_{a,a':A}\mathsf{Id}(f(a),g(a'))$ are isomorphic, but not necessarily identical, via $\eta^{\Rightarrow}[A,B]$. Nevertheless, we still get stability under degeneracies since we explicitly allowed for this possibility in Definition 49.*

As Examples 52 and 53 show, cartesian closed reflexive graph categories with isomorphisms suitably generalize the structure of sets and relations, and, as we saw above, allow us to interpret unit, product, and function types in a straightforward way.

Lemma 54. *If a reflexive graph category with isomorphisms $(\mathcal{X},(\mathcal{M},\mathcal{I}))$ has terminal objects, products, and exponentials, all stable under face maps and degeneracies, then the forgetful functor from the category $\int_n \mathcal{M}^n \to \mathcal{M}$ to $\mathsf{Ctx}(\mathcal{X})$ is a cartesian closed split fibration with a split generic object $\Omega := 1$.*

5 Reflexive Graph Models of Parametricity

It remains to interpret $\forall$-types. For this, the following notation will be useful.

Notation 55. *Let $\mathcal{B}$ be a category with a distinguished object Ω and products $(-) \times \Omega$.*

- *For $n \in \mathbb{N}$, let $\mathsf{q}_n^{\Omega} : \Omega^n \times \Omega \to \Omega$ be the second projection.*
- *For $n,k \in \mathbb{N}$, let $\mathsf{p}_n^{\Omega}(k) : \Omega^{n+k+1} \to \Omega^{n+k}$ be the "weakening morphism" that drops the k-th variable in the context (counting from the right), defined inductively by*

$$\mathsf{p}_n^{\Omega}(0) := \mathsf{fst}[\Omega^n, \Omega]$$
$$\mathsf{p}_n^{\Omega}(k+1) := \mathsf{p}_n^{\Omega}(k) \times \mathsf{id}$$

- *For $n, k \in \mathbb{N}$, and a morphism $A : \Omega^n \to \Omega$, let $\mathsf{s}^\Omega_n(k, A) : \Omega^{n+k} \to \Omega^{n+k+1}$ be the "substitution morphism" that substitutes A for the k-th variable in the context (counting from the right), defined inductively by*

$$\mathsf{s}^\Omega_n(0, A) \coloneqq \langle \mathsf{id}, A\rangle$$
$$\mathsf{s}^\Omega_n(k+1, A) \coloneqq \mathsf{s}^\Omega_n(k, A) \times \mathsf{id}$$

To interpret $\forall$-types we need to know that, in the forgetful fibration from Lemma 54, each weakening functor $\mathsf{p}^\Omega_n(0)^*$ for $n \in \mathbb{N}$ has a right adjoint $\forall_n$.

Remark 56. *In [1], only $\forall_0$ is required, with the intention that $\forall_n$ can be derived from $\forall_0$ using partial application, but this does not seem to work since partial application of an indexed functor might not be an indexed functor.*

For our running examples we have:

Example 57 (PER model, continued)**.** *Define the adjoint $\forall_n$ by*

$$\forall_n \mathcal{F}(0)\ \overline{A} \coloneqq \big\{(m, k) \mid \textit{for all } A : \mathsf{Ob}(\mathcal{R}(0)), m \sim_{\mathcal{F}(0)\ (\overline{A}, A)} k,$$
$$\textit{and for all } R : \mathsf{Ob}(\mathcal{R}(1)), \mathcal{F}(1)(\overline{\mathsf{Eq}\, A}, R)\ \langle m, k\rangle\big\}$$
$$\forall_n \mathcal{F}(1)\ \overline{R} \coloneqq \Big(\big(\forall_n \mathcal{F}(0)\ \overline{R_0}, \forall_n \mathcal{F}(0)\ \overline{R_1}\big),$$
$$\big\{m \mid \textit{for all } R : \mathsf{Ob}(\mathcal{R}(1)), \mathcal{F}(1)\, (\overline{R}, R)\, m\big\}\Big)$$

where for a relation $R \coloneqq ((A_0, A_1), R_A)$ we write R_0 for A_0, R_1 for A_1, and, par abus de notation*, write R for both the entire relation and just its predicate part R_A. We will employ a similar convention for Reynolds' model. To define $\forall_n$ on a morphism $\eta : \mathcal{F} \to \mathcal{G}$, we define*

$$\forall_n \eta(0)\ \overline{A} \coloneqq \Big(\big(\forall_n \mathcal{F}(0)\ \overline{A}, \forall_n \mathcal{G}(0)\ \overline{A}\big), \{m \cdot 0\}_{(\forall_n \mathcal{F}(0)\ \overline{A}) \to (\forall_n \mathcal{G}(0)\ \overline{A})}\Big)$$

Here m is any natural number realizing $\eta(0)\ \overline{A}$. Crucial observations are that all natural transformations are "uniformly realized" in the sense that there is a natural number realizing each such transformation, and since all PERs are defined to be realized by all natural numbers, each is suitably uniform. In particular, if η were not uniformly realized in the above sense then $\forall_n$ would not be well-defined on morphisms. These observations can be used to show that, in the category-theoretic setting (rather than the setting of ω-sets), the family of adjoints $\forall$ cannot exist precisely because ad hoc *natural transformations —* i.e.*, natural transformations that are not uniformly realizable, even though each of their components may indeed be realizable — are not excluded.*

Example 58 (Reynolds' model, continued)**.** *On the set level, the adjoint $\forall_n$ is defined as follows:*

$$\forall_n \mathcal{F}(0)\,\overline{A} \coloneqq \Big\{ f_0 : \Pi_{A:\mathbb{U}} \mathcal{F}(0)(\overline{A}, A) \;\&$$
$$f_1 : \Pi_{R:\mathsf{R}_0} \mathcal{F}(1)\,(\overline{\mathsf{Eq}\,A}, R)\,(f_0(R_0), f_0(R_1)) \;\&$$
$$\Pi_{i:\mathcal{M}(0)_1} \mathcal{F}(0)\big(\overline{\mathsf{id}_{\mathcal{M}(0)}(A)}, i\big)\; f_0(i_{\mathsf{d}}) = f_0(i_{\mathsf{c}}) \Big\}$$

The last condition says that f_0 is functorial in its argument, in the sense that if i is an isomorphism, then $f_0(i_{\mathsf{d}})$ and $f_0(i_{\mathsf{c}})$ are suitably related via the obvious isomorphism between $\mathcal{F}(0)\,(\overline{A}, i_{\mathsf{d}})$ and $\mathcal{F}(0)\,(\overline{A}, i_{\mathsf{c}})$. This condition, which does not have an analogue in the set-theoretic presentation of Reynolds' model, is needed because we do not work with discrete domains as is common in other presentations of parametricity. A very similar condition does appear, e.g.*, in the definition of parametric limits for the category of sets in [1]. An analogous condition "one level up" asserting the functoriality of f_1 is automatically satisfied since the codomain of f_1 is a proposition. Finally, on the relation level, we define $\forall_n \mathcal{F}(1)\,\overline{R}$ to be the relation with domain $\forall_n \mathcal{F}(0)\,\overline{R_0}$ and codomain $\forall_n \mathcal{F}(0)\,\overline{R_1}$ mapping $((f_0, f_1), (g_0, g_1))$ to $\Pi_{R:\mathsf{R}_0} \mathcal{F}(1)\,(\overline{R}, R)\,(f_0(R_0), g_0(R_1))$.*

To see that the above definition indeed gives a degeneracy-preserving reflexive graph functor, fix $\overline{A}$. We want to show that the two relations $\mathsf{Eq}\,(\forall_n \mathcal{F}(0)\,\overline{A})$ and $\forall_n \mathcal{F}(1)\,\overline{\mathsf{Eq}(A)}$ are isomorphic. The domains and codomains of these relations are all the same – $\forall_n \mathcal{F}(0)\,\overline{A}$ – so we let both of the underlying maps of the isomorphism be identities (as also required by the coherence condition on the isomorphism). Fix $((f_0, f_1), (g_0, g_1)) : (\forall_n \mathcal{F}(0)\,\overline{A}) \times (\forall_n \mathcal{F}(0)\,\overline{A})$. We need functions going back and forth between $\mathsf{Id}((f_0, f_1), (g_0, g_1))$ and $\Pi_{R:\mathsf{R}_0} \mathcal{F}(1)\,(\overline{\mathsf{Eq}(A)}, R)\,(f_0(R_0), g_0(R_1))$. Such functions will automatically be mutually inverse since the types in question are propositions.

Going from left to right is easy using Id-induction and f_1. To go from right to left, fix $\phi : \Pi_{R:\mathsf{R}_0} \mathcal{F}(1)\,(\overline{\mathsf{Eq}(A)}, R)\,(f_0(R_0), g_0(R_1))$. To show $\mathsf{Id}((f_0, f_1), (g_0, g_1))$ it suffices to show $\mathsf{Id}(f_0, g_0)$ since the type of g_1 (or f_1) is a proposition. By function extensionality, it suffices to show pointwise equality between f_0 and g_0. So fix B. The only thing we can do with ϕ is to apply it to $\mathsf{Eq}(B)$, which gives us $\phi(\mathsf{Eq}(B)) : \mathcal{F}(1)\,(\overline{\mathsf{Eq}(A)}, \mathsf{Eq}(B))\,(f_0(B), g_0(B))$. The relation $\mathcal{F}(1)\,(\overline{\mathsf{Eq}(A)}, \mathsf{Eq}(B))$ is isomorphic to $\mathsf{Eq}\,\mathcal{F}(0)\,(\overline{A}, B)$ via $\varepsilon_{\mathcal{F}}(\overline{A}, B)^{-1}$. Applying $\varepsilon_{\mathcal{F}}(\overline{A}, B)^{-1}$ to $(f_0(B), g_0(B))$ and $\phi(\mathsf{Eq}(B))$ thus gives us $\mathsf{Id}\big(\mathbf{f}_\top(\varepsilon_{\mathcal{F}}(\overline{A}, B))^{-1}\; f_0(A), \mathbf{f}_\bot(\varepsilon_{\mathcal{F}}(\overline{A}, B))^{-1}\; g_0(B)\big)$. The coherence condition on $\varepsilon_{\mathcal{F}}$ tells us that $\mathbf{f}_\star(\varepsilon_{\mathcal{F}}(\overline{A}, B))$ is the identity on $\mathcal{F}(0)(\overline{A}, B)$, hence so is $\mathbf{f}_\star(\varepsilon_{\mathcal{F}}(\overline{A}, B))^{-1}$. This gives us $\mathsf{Id}(f_0(A), g_0(B))$ as desired.

With a family of adjoints $\forall_n$ in hand, we are almost in a position to use Seely's result [14], which constructs a model of System F from a split $\lambda 2$-fibration. The only missing piece is showing that the adjoints are natural, *i.e.*, that they satisfy the following Beck-Chevalley condition: for any morphism $f : n \to m$ in $\mathsf{Ctx}(\mathcal{X})$ and object $X : \mathcal{M}^{m+1} \to \mathcal{M}$, the canonical morphism $\theta_\forall(f, X)$ from $f^*(\forall_m(X))$ to $\forall_n((f \times \mathsf{id})^*(X))$ is the identity. But here we hit a snag. In Reynolds' model, the type

$\mathcal{F}^*(\forall_1(\mathcal{G}))$ for $\mathcal{F} : \mathcal{M}$ and $\mathcal{G} : \mathcal{M}^2 \to \mathcal{M}$ has the form

$$\big\{f_0 : \Pi_{A:\mathsf{Set}}\mathcal{G}(0)\,(\mathcal{F}(0), A)\ \&\ f_1 : \Pi_{R:\mathsf{R}_0}\mathcal{G}(1)\,(\mathsf{Eq}\,\mathcal{F}(0), R)\,(f_0(R_0), f_0(R_1))\ \&\ \ldots\big\}$$

whereas the type $\forall_0((\mathcal{F} \times \mathsf{id})^*(\mathcal{G}))$ has the form

$$\big\{f : \Pi_{A:\mathsf{Set}}\mathcal{G}(0)\,(\mathcal{F}(0), A)\ \&\ f_1 : \Pi_{R:\mathsf{R}_0}\mathcal{G}(1)\,(\mathcal{F}(1), R)\,(f_0(R_0), f_0(R_1))\ \&\ \ldots\big\}$$

Since $\mathcal{F}$ is degeneracy-preserving, $\mathcal{F}(1)$ is *isomorphic* to $\mathsf{Eq}\,\mathcal{F}(0)$. But these are not necessarily identical, so $\theta_\forall(f, X)$ is not necessarily an identity, and we therefore cannot directly invoke Seely's result.

However, we can still show that $\theta_\forall(f, X)$ is an *isomorphism*, so what we have is essentially a "pseudo-natural" hyperdoctrine [9]. At this point, we could try to adapt one of the known strictification techniques for other theories (see, *e.g.*, [10]), but this has the disadvantage mentioned in the introduction of changing the underlying interpretation. Instead, we generalize the notion of a split $\lambda2$-fibration to allow the type constructors to commute with substitution only up to canonical isomorphisms.

The main obstacle to verifying the equational theory in the more general setting when type formers are allowed to commute with substitution only up to canonical isomorphisms is that the substitution of isomorphic types may yield non-isomorphic results. Consider, for instance, types $\cdot \vdash S$ and $\alpha, \beta \vdash T$, and let S' denote the weakened type $\alpha \vdash S$. By assumption, weakening is modeled by the weakening functor $\mathsf{p}_n^\Omega(k)^*$, so $[\![S']\!]$ is isomorphic to $\mathsf{p}_0^\Omega(0)^*\,[\![S]\!]$. Since substitution (of X) is modeled by the substitution functor $\mathsf{s}_n^\Omega(k, X)^*$, both $\mathsf{s}_1^\Omega(0, [\![S']\!])^*\,[\![T]\!]$ and $\mathsf{s}_1^\Omega(0, \mathsf{p}_0^\Omega(0)^*\,[\![S]\!])^*\,[\![T]\!]$ should model the substitution $T[\beta := S']$, up to isomorphism. But there is no *a priori* reason that these two types *should* be isomorphic: the $\lambda2$-structure ensures that any object A in the fiber over n can be identified with a morphism $A : n \to 1$ but it gives no guarantee that if A and B are isomorphic then A^* and B^* will be naturally isomorphic, or even related in any way whatsoever (the PER model furnishes a counterexample). If X arises as an interpretation of a System F type, then we can construct an isomorphism between $\mathsf{s}_n^\Omega(k, A)^*\,X$ and $\mathsf{s}_n^\Omega(k, B)^*\,X$ by induction on the structure of X but, of course, this property does not extend to arbitrary objects.

We solve this problem by *requiring* that, for "good" isomorphisms from A to B, the functors $\mathsf{s}_n^\Omega(k, A)^*$ and $\mathsf{s}_n^\Omega(k, B)^*$ are naturally isomorphic. To select the "good" isomorphisms, we introduce:

Definition 59. *A* wide split subfibration *of a split fibration $U : \mathcal{E} \to \mathcal{B}$ is a restriction $U' : \mathcal{E}' \to \mathcal{B}$ of U, where $\mathcal{E}'$ is a wide subcategory of $\mathcal{E}$ with the property that, for any object X of $\mathcal{E}$ and $f : Y \to UX$ in $\mathcal{B}$, the cartesian lifting of f with respect to U is cartesian with respect to U'.*

There is a similar definition in [5], but it omits the requirement that the lifting of f be *cartesian* with respect to U', as opposed to merely belonging to $\mathcal{E}'$. Under [5]'s definition, a subfibration is not necessarily a fibration.

Definition 60. A split fibration with isomorphisms *is a split fibration $U : \mathcal{E} \to \mathcal{B}$ (the* underlying fibration*), together with a wide split subfibration $U' : \mathcal{E}' \to \mathcal{B}$ of U (the* fibration of isomorphisms*), satisfying the following properties:*

1. *U has a split generic object Ω in $\mathcal{B}$.*
2. *$\mathcal{B}$ has a terminal object and products $(-) \times \Omega$.*
3. *For $n \in \mathbb{N}$, every morphism in $\mathcal{E}'(\Omega^n)$ is an isomorphism.*
4. *For any $n, k \in \mathbb{N}$ and morphism $i : A \to B$ in $\mathcal{E}'(\Omega^n)$, there is a natural isomorphism $\phi_n(k, i)$ between $\mathsf{s}^{\Omega}_n(k, A)^*$, $\mathsf{s}^{\Omega}_n(k, B)^* : \mathcal{E}(\Omega^{n+k+1}) \to \mathcal{E}(\Omega^{n+k})$ such that:*
   (a) $\phi_n(0, i)\, \mathsf{q}^{\Omega}_n = i$
   (b) $\phi_n(k + 1, i)\, \mathsf{q}^{\Omega}_{n+k+1} = \mathsf{id}_{\mathsf{q}^{\Omega}_{n+k}}$
   (c) $\phi_n(0, i)\, (\mathsf{p}^{\Omega}_n(0)^\star\, X) = \mathsf{id}_X$ *for every object X of $\mathcal{E}(\Omega^n)$*
   (d) $\phi_n(k+1, i)\, (\mathsf{p}^{\Omega}_{n+k+1}(0)^*\, X) = \mathsf{p}^{\Omega}_{n+k}(0)^*\, (\phi_n(k, i)\, X)$ *for every object X of $\mathcal{E}(\Omega^{n+k+1})$*

Condition 3 justifies our choice of terminology in Definition 60. We could instead require that *every* vertical morphism in $\mathcal{E}'$ be an isomorphism, but this is not necessary because, when constructing models of System F, we only need to consider fibers over powers of the generic object Ω. Conditions *4 a)* through *4 d)* ensure that when X is the interpretation of a System F type, $\phi_n(k, i)\, X$ is precisely the isomorphism determined by induction on the structure of X. Similar conditions are needed any time we impose more structure on the fibration. For example:

Definition 61. *A split fibration with isomorphisms* has fibered terminal objects *if the following conditions hold:*

1. *For $n \in \mathbb{N}$, the fiber $\mathcal{E}(\Omega^n)$ has a terminal object 1_n.*
2. *Beck-Chevalley: for any morphism $f : \Omega^n \to \Omega^m$ in $\mathcal{B}$, the canonical morphism $\theta_1(f) : f^*(1_m) \to 1_n$ is in $\mathcal{E}'$.*
3. *For any $n, k \in \mathbb{N}$ and morphism $i : A \to B$ in $\mathcal{E}'(\Omega^n)$, we have*

$$\theta_1(\mathsf{s}^{\Omega}_n(k, B)) \circ \phi_n(k, i)(1_{n+k+1}) = \theta_1(\mathsf{s}^{\Omega}_n(k, A))$$

Definition 62. *A split fibration with isomorphisms* has fibered products *if the following conditions hold:*

1. *For $n \in \mathbb{N}$, the fiber $\mathcal{E}(\Omega^n)$ has products $\times_n$ that preserve membership in $\mathcal{E}'$.*
2. *Beck-Chevalley: for any morphism $f : \Omega^n \to \Omega^m$ in $\mathcal{B}$ and objects X, Y in $\mathcal{E}(\Omega^m)$, the canonical morphism below is in $\mathcal{E}'$:*

$$\theta_\times(f, X, Y) : f^*(X \times_m Y) \to \big(f^*(X) \times_n f^*(Y)\big)$$

3. *For any $n, k \in \mathbb{N}$, morphism $i : A \to B$ in $\mathcal{E}'(\Omega^n)$, and objects X, Y in $\mathcal{E}(\Omega^{n+k+1})$, we have*

$$\theta_\times\big(\mathsf{s}_n^\Omega(k, B), X, Y\big) \circ \phi_n(k, i)\,(X \times_{n+k+1} Y) = \\ \big((\phi_n(k, i)\,X) \times_{n+k} (\phi_n(k, i)\,Y)\big) \circ \theta_\times\big(\mathsf{s}_n^\Omega(k, A), X, Y\big)$$

Definition 63. *A split fibration with isomorphisms and fibered products* has fibered exponentials *if the following conditions hold:*

1. *For $n \in \mathbb{N}$, the fiber $\mathcal{E}(\Omega^n)$ has exponentials $\Rightarrow_n$ that preserve membership in $\mathcal{E}'$.*

2. *Beck-Chevalley: for any morphism $f : \Omega^n \to \Omega^m$ in $\mathcal{B}$ and objects X, Y in $\mathcal{E}(\Omega^m)$, the canonical morphism below is in $\mathcal{E}'$:*

$$\theta_\Rightarrow(f, X, Y) : f^*(X \Rightarrow_m Y) \to \big(f^*(X) \Rightarrow_n f^*(Y)\big)$$

3. *For any $n, k \in \mathbb{N}$, morphism $i : A \to B$ in $\mathcal{E}'(\Omega^n)$, and objects X, Y in $\mathcal{E}(\Omega^{n+k+1})$, we have*

$$\theta_\Rightarrow\big(\mathsf{s}_n^\Omega(k, B), X, Y\big) \circ \phi_n(k, i)\,(X \Rightarrow_{n+k+1} Y) = \\ \big((\phi_n(k, i)\,X)^{-1} \Rightarrow_{n+k} (\phi_n(k, i)\,Y)\big) \circ \theta_\Rightarrow\big(\mathsf{s}_n^\Omega(k, A), X, Y\big)$$

Definition 64. *A* split cartesian closed fibration with isomorphisms *is a split fibration with isomorphisms that has fibered terminal objects, products, and exponentials.*

Definition 65. *A split fibration with isomorphisms* has simple Ω-products *if the following conditions hold:*

1. *For $n \in \mathbb{N}$, the weakening functor $\mathsf{p}_n^\Omega(0)^* : \mathcal{E}(\Omega^n) \to \mathcal{E}(\Omega^{n+1})$ has a right adjoint $\forall_n$, and these adjoints preserve membership in $\mathcal{E}'$.*

2. *Beck-Chevalley: for any morphism $f : \Omega^n \to \Omega^m$ in $\mathcal{B}$ and object X in $\mathcal{E}(\Omega^{m+1})$, the canonical morphism below is in $\mathcal{E}'$:*

$$\theta_\forall(f, X) : f^*(\forall_m(X)) \to \forall_n((f \times \mathsf{id})^*(X))$$

3. *For any $n, k \in \mathbb{N}$, morphism $i : A \to B$ in $\mathcal{E}'(\Omega^n)$, and object X in $\mathcal{E}(\Omega^{n+k+2})$, we have*

$$\theta_\forall\big(\mathsf{s}_n^\Omega(k, B), X\big) \circ \phi_n(k, i)\,(\forall_{n+k+1}(X)) = \\ \forall_{n+k}(\phi_n(k+1, i)\,X) \circ \theta_\forall\big(\mathsf{s}_n^\Omega(k, A), X\big)$$

We can now state the key definition of this section:

Definition 66. *A* split λ2-fibration with isomorphisms *is a split cartesian closed fibration with isomorphisms that has simple Ω-products as specified in Definition 65.*

Any split $\lambda 2$-fibration trivially gives rise to a split $\lambda 2$-fibration with isomorphisms by taking the total category $\mathcal{E}'$ of the fibration of isomorphisms to consist of only the chosen cartesian morphisms in $\mathcal{E}$, which forces every vertical morphism in $\mathcal{E}'$ to be an identity. Our first main result generalizes Seely's [14] well-known one:

Theorem 67. *Every split $\lambda 2$-fibration $U : \mathcal{E} \to \mathcal{B}$ with isomorphisms gives a sound model of System F in which:*

- *every type context Γ is interpreted as an object $[\![\Gamma]\!]$ in $\mathcal{B}$*
- *every type $\Gamma \vdash T$ is interpreted as an object $[\![\Gamma \vdash T]\!]$ in the fiber $\mathcal{E}([\![\Gamma]\!])$*
- *every term context $\Gamma; \Delta$ is interpreted as an object $[\![\Gamma \vdash \Delta]\!]$ in the fiber $\mathcal{E}([\![\Gamma]\!])$*
- *every term $\Gamma; \Delta \vdash t : T$ is interpreted as a morphism $[\![\Gamma; \Delta \vdash t : T]\!]$ from $[\![\Gamma; \Delta]\!]$ to $[\![\Gamma \vdash T]\!]$ in the fiber $\mathcal{E}([\![\Gamma]\!])$*

Proof sketch. We outline the part of the proof which utilitizes the extra structure of a split λ-2 fibration with isomorphisms (as opposed to a split λ-2 fibration). The interpretation of System F types proceeds by induction of the structure of the type. Let $[T]_k$ stand for the weakening of the type expression T by inserting an extra free variable before the k-th free variable and $[T]_k^A$ stand for the substitution of the type expression A for the k-th free variable in the type expression T. We have the following:

Lemma 68. *If $\Gamma_1, \Gamma_2 \vdash T$ type with $|\Gamma_1| = n$, $|\Gamma_2| = k$, then there is an isomorphism $i_n(k, T) : \mathsf{p}_n^{\Omega}(k)^* \, [\![T]\!] \to [\![[T]_k]\!]$ that belongs to $\mathcal{E}'$.*

Proof. By induction on the structure of T. □

Lemma 69. *If $\Gamma_1, X, \Gamma_2 \vdash T$ type and $\Gamma_1 \vdash A$ type with $|\Gamma_1| = n$, $|\Gamma_2| = k$, then there is an isomorphism $j_n(k, A, T) : \mathsf{s}_n^{\Omega}(k, [\![A]\!])^* \, [\![T]\!] \to [\![[T]_k^A]\!]$ that belongs to $\mathcal{E}'$.*

Proof. By induction on the structure of T. □

We define i and j on term contexts Δ instead of just a single type T in the obvious way. The interpretation of System F terms proceeds by induction of the structure of the term. If $\Gamma; \Delta \vdash \Lambda X.t : \forall X.T$ with $|\Gamma| = n$, and $\Gamma; [\Delta]_0 \vdash t : T$, we define $[\![\Lambda X.t]\!]$ to be the composition

$$
\begin{array}{c}
[\![\Delta]\!] \\
\Big\downarrow \eta_n([\![\Delta]\!]) \\
\forall_n(\mathsf{p}_n^{\Omega}(0)^* \, [\![\Delta]\!]) \\
\Big\downarrow \textit{via } i_n(0,\Delta) \\
\forall_n \, [\![[\Delta]_0]\!] \\
\Big\downarrow \textit{via } [\![t]\!] \\
\forall_n \, [\![T]\!]
\end{array}
$$

If $\Gamma; \Delta \vdash t : \forall X.T$, $\Gamma \vdash A$ type, $\Gamma; \Delta \vdash t \; A : [T]_0^A$ with $|\Gamma| = n$, we define $[\![t \, A]\!]$ to be the composition

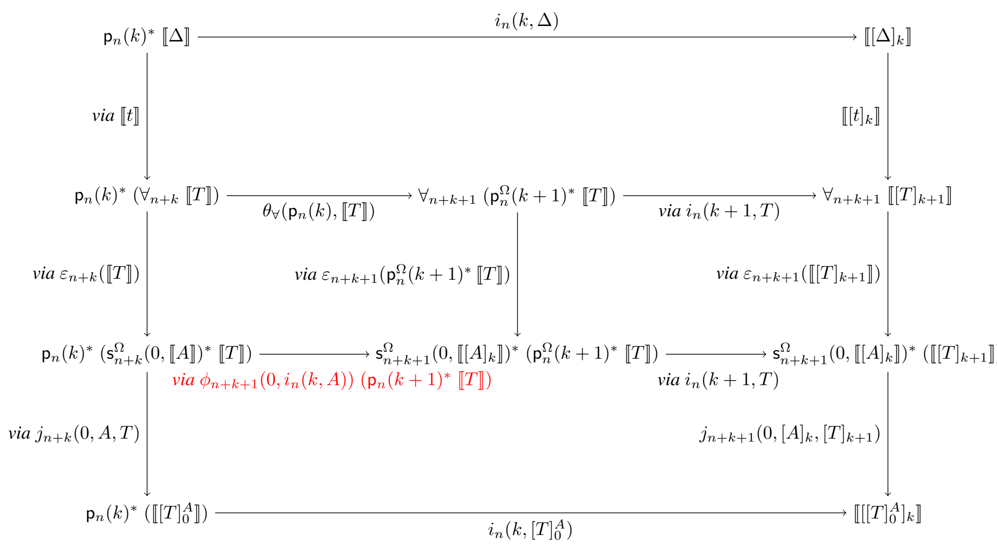
$$
\begin{array}{c}
[\![\Delta]\!] \\
\Big\downarrow [\![t]\!] \\
\forall_n \, [\![T]\!] \\
\Big\downarrow \mathsf{s}_n^{\Omega}(0, [\![A]\!])^* \; (\varepsilon_n([\![T]\!])) \\
\mathsf{s}_n^{\Omega}(0, [\![A]\!])^* \, [\![T]\!] \\
\Big\downarrow j_n(0, A, T) \\
[\![[T]_0^A]\!]
\end{array}
$$

Let $[t]_k$ stand for the weakening of the term expression t by inserting an extra free type variable before the k-th free type variable. We now need the following lemma:

Lemma 70. *If $\Gamma_1, \Gamma_2; \Delta \vdash t : T$ with $|\Gamma_1| = n$, $|\Gamma_2| = k$, then the following diagram commutes:*

$$
\begin{array}{ccc}
\mathsf{p}_n(k)^* \, [\![\Delta]\!] & \xrightarrow{\; i_n(k,\Delta) \;} & [\![[\Delta]_k]\!] \\
\text{via } [\![t]\!] \Big\downarrow & & \Big\downarrow [\![[t]_k]\!] \\
\mathsf{p}_n(k)^* \, [\![T]\!] & \xrightarrow[\; i_n(k,T) \;]{} & [\![[T]_k]\!]
\end{array}
$$

Proof. By induction on the structure of T. The interesting case is when $T \coloneqq t\,A$. This requires us to show that the outer square in the diagram below commutes:



The top rectangle commutes by induction hypothesis; middle left by the definition of $\theta_\forall$, the naturality of $\phi_{n+k+1}(0, i_n(k, A))$, and condition *4 c)* in Definition 60; middle right by the naturality of ε_{n+k+1}; and the bottom rectangle by the naturality of $\phi_{n+k+1}(0, i_n(k, A))$ and induction on the structure of T, using properties *4 a) - d)* of Definition 60 and *3* of Definitions 61, 62, 63. □

We have a similar lemma for the substitution of a type for a variable in a term, as well as similar lemmas for the weakening of a term by a term variable and the substitution of a term for a variable in a term. The syntactic equalities can now be verified by induction of the derivation. □

We now want to specify when a model of System F given by a split $\lambda 2$-fibration with isomorphisms according to Theorem 67 is relationally parametric. For this we will use our second main result, which shows that every cartesian closed reflexive graph category with isomorphisms naturally gives rise to a split cartesian closed fibration with isomorphisms.

Theorem 71. *Given a cartesian closed reflexive graph category with isomorphisms $(\mathcal{X}, (\mathcal{M}, \mathcal{I}))$, there is a canonical split cartesian closed fibration with isomorphisms whose underlying fibration is the forgetful functor from $\int_n \mathcal{M}^n \to \mathcal{M}$ to $\mathsf{Ctx}(\mathcal{X})$.*

Proof sketch. To define the fibration of isomorphisms, we specify a wide subcategory of $\int_n \mathcal{M}^n \to \mathcal{M}$ as follows: a morphism $(\mathbf{F}, \eta)$ from (n, F) to (m, G) belongs to this subcategory iff there is a reflexive graph natural transformation $\eta' : F \to \mathbf{F}^*(G)$ with $\mathcal{I} \circ \eta' = \eta$. (We note that such an η' is necessarily unique.) In other words, the "good" morphisms are precisely those reflexive graph natural transformations whose components at each level l are "good", in the sense of belonging to the image of $\mathcal{I}(l)$. The forgetful functor from $\int_n \mathcal{M}^n \to \mathcal{M}$ to $\mathsf{Ctx}(\mathcal{X})$ was already recognized as a split cartesian closed fibration with a split generic object $\Omega := 1$ in Lemma 54, which takes care of condition *1* in Definition 60. Lemma 34 ensures that the base category has the required terminal object and products with Ω, which takes care of condition *2*. Since each morphism in $\mathcal{M}(l)$ is required to be an isomorphism, condition *3* is satisfied, and since the products and exponentials of morphisms are required to preserve membership in the image of $\mathcal{I}(l)$, so is condition *1* of Definitions 61, 62, 63. The canonical morphisms in condition *2* of Definitions 61, 62, 63 are all identities since reindexing preserves terminal objects, products, and exponentials on the nose. Finally, we define $\phi_n(k, i, H) := \mathcal{I} \circ H \circ \langle \upsilon_n(k, i, 0), \ldots, \upsilon_n(k, i, n + k)\rangle$ for an object H in $\mathcal{M}^{n+k+1} \to \mathcal{M}$. Here, $\upsilon_n(k, i, j)$ for $j \leq n + k$ is the reflexive graph natural transformation between the j^{th} components of $\mathsf{s}^\Omega_n(k, A)$ and $\mathsf{s}^\Omega_n(k, B)$, respectively, given by induction on k by:

$$\upsilon_n(0, i, j) := \begin{cases} \mathsf{id}_{\pi^n_j} & \text{for } j < n \\ i & \text{for } j = n \end{cases}$$

and

$$\upsilon_n(k + 1, i, j) := \begin{cases} \mathsf{p}^\Omega_{n+k}(0)^*\, \upsilon_n(k, i, j) & \text{for } j < n + k + 1 \\ \mathsf{id}_{\pi^{n+k+1}_{n+k}} & \text{for } j = n + k + 1 \end{cases}$$

Conditions *4 a) - d)* of Definition 60 and *3* of Definitions 61, 62, 63 can now be verified by routine calculation. □

If $\mathcal{R}$ is a cartesian closed reflexive graph category with isomorphisms, we denote by $\mathsf{Fwl}(\mathcal{R})$ the canonical split cartesian closed fibration with isomorphisms whose existence is guaranteed by Theorem 71. To formulate an abstract definition of a parametric model, we will appropriately relate a split $\lambda 2$-fibration with isomorphisms U to $\mathsf{Fwl}(\mathcal{R})$. To see how, we revisit the simplest model, namely the term model. In the associated split $\lambda 2$-fibration U_{term}, the fiber over $n \in \mathbb{N}$ consists of types and terms with n free type variables. If $\mathcal{U}$ is the category consisting of closed types and terms between them, then $\mathcal{U}$ induces a split cartesian closed fibration, U_{set}, whose fiber over n consists of functors from $|\mathcal{U}|^n \to \mathcal{U}$ and natural transformations between them. Here, a type $\overline{\alpha} \vdash T$ with n free variables can be seen as functor from $|\mathcal{U}|^n \to \mathcal{U}$, and a term $\overline{\alpha}; x : S \vdash t : T$ as a natural transformation between S and T. We thus have a morphism of split cartesian closed fibrations $\mu : U_{term} \to U_{set}$. However, unlike U_{term}, U_{set} does not admit the family of adjoints required to make it a $\lambda 2$-fibration. We can view U_{term} as a version of U_{set} that "enriches" the functors and natural transformations with enough extra information to ensure that the desired adjoints exist: in this example, the information that the maps involved are not *ad hoc*, but come from syntax. Since these adjunctions are only applicable to non-empty contexts, no such "enrichment" should be necessary for objects and morphisms over the *terminal* object. And indeed, the restriction of μ to the fibers over the respective terminal objects is clearly an equivalence. These observations motivate our main definition:

Definition 72. *Let $\mathcal{R}$ be a cartesian closed reflexive graph category with isomorphisms. A* parametric model of System F over $\mathcal{R}$ *is a split $\lambda 2$-fibration with isomorphisms U, together with a morphism $\mu : U \to \mathsf{Fwl}(\mathcal{R})$ of split cartesian closed fibrations with isomorphisms whose restriction to the fibers of U and $\mathsf{Fwl}(\mathcal{R})$ over the terminal objects is full, faithful, and essentially surjective.*

Theorem 73 (PER model)**.** *Let $\mathcal{R}_{PER}$ be the cartesian closed reflexive graph category with isomorphisms defined in Examples 7, 28, and 52. The family of adjoints defined in Example 57 makes $\mathsf{Fwl}(\mathcal{R}_{PER})$ into a split $\lambda 2$-fibration with isomorphisms, and hence into a parametric model of System F over $\mathcal{R}_{PER}$.*

Theorem 74 (Reynolds' model)**.** *Let $\mathcal{R}_{REY}$ be the reflexive graph category with isomorphisms defined in Examples 8, 29, and 53. The family of adjoints defined in Example 58 makes $\mathsf{Fwl}(\mathcal{R}_{REY})$ into a split $\lambda 2$-fibration with isomorphisms, and hence into a parametric model of System F over $\mathcal{R}_{REY}$.*

6 A Proof-Relevant Model of Parametricity

We now describe a proof-relevant version of Reynolds' model, in which witnesses of relatedness are not necessarily unique but are instead themselves related via a yet higher notion of a relation – a 2-relation. The construction of such a model is the subject of [8], but the development there seems to contain a major technical gap: it is unclear how to prove the $\forall$-case in Lemma 9.4 in [8], whose proof is not given and we cannot reconstruct. Since this lemma is crucial to the soundness of the interpretation, it is unknown whether the result of [8] can be salvaged as-is. Here we reuse the main ideas from [8] but put them on a solid technical footing.

Example 75. *We use the same ambient category as in Example 8 and reuse the (internal) category Set of types. The category R of relations is almost the same as in Example 8, except that relations are now proof-relevant,* i.e., $\mathsf{R}_0 := \Sigma_{A,B:\mathsf{Set}} A \times B \to \mathbb{U}$. *As before, we have two face maps $\mathbf{f}_\top, \mathbf{f}_\bot : \mathcal{R} \to \mathsf{Set}$ projecting out the domain and codomain of a relation and a degeneracy $\mathsf{Eq} : \mathsf{Set} \to \mathcal{R}$ constructing the equality relation. Given relations R on A and B and S on C and D, to relate two witnesses $p : R(a,b)$ and $q : S(c,d)$ we should know* a priori *how a relates to c and b to d. This motivates defining the category $\mathsf{2R}$ of* 2-relations*, whose objects Q are tuples $(Q^{0\top}, Q^{1\top}, Q^{0\bot}, Q^{1\bot})$ of relations forming a square*

$$\begin{array}{ccc} A & \xrightarrow{Q^{0\top}} & B \\ {\scriptstyle Q^{1\top}}\downarrow & & \downarrow{\scriptstyle Q^{1\bot}} \\ C & \xrightarrow[Q^{0\bot}]{} & D \end{array}$$

together with a Prop-valued predicate (also denoted Q) on the type of tuples of the form $((a,b,c,d),(p,q,r,s))$, where $p : Q^{0\top}(a,b)$, $q : Q^{1\top}(a,c)$, $r : Q^{0\bot}(c,d)$, and $s : Q^{1\bot}(b,d)$. This gives four face maps $\mathbf{f}_{0\top}, \mathbf{f}_{0\bot}, \mathbf{f}_{1\top}, \mathbf{f}_{1\bot} : \mathsf{2R} \to \mathsf{R}$, one for each edge.

We have two degeneracies from R *to* $\mathsf{2R}$*, one replicating a relation* R *horizontally and one vertically. More precisely, given* R*, we obtain the 2-relation* $\mathsf{Eq}_{=}(R)$ *by placing* R *on top and bottom, with equality relations* $\mathsf{Eq}(R_0)$ *and* $\mathsf{Eq}(R_1)$ *as vertical edges, and mapping* $\big((a,b,a,b),(p,-,r,-)\big)$ *to* $\mathsf{Id}(p,r)$*. The symmetric version* $\mathsf{Eq}_{||}(R)$ *places* R *on left and right and assumes equality relations as horizontal edges. But we also have two other ways of turning a relation* R *into a 2-relation: the functor* $\mathbf{C}_\top$ *places* R *on top and left, and* $\mathbf{C}_\bot(R)$ *places* R *on bottom and right, filling the remaining edges with equalities. The functors* $\mathbf{C}_\top$ *and* $\mathbf{C}_\bot$ *are called* connections. *We define terminal objects, products, exponentials, and isomorphisms in the obvious way.*

Just like in Reynolds' model, we have $\mathbf{f}_\star \circ \mathsf{Eq} = \mathsf{id}$*. We also have further equalities:*

- $\mathbf{f}_{0\star} \circ \mathsf{Eq}_{=} = \mathsf{id}$
- $\mathbf{f}_{1\star} \circ \mathsf{Eq}_{=} = \mathsf{Eq} \circ \mathbf{f}_\star$ *for a fixed* $\star \in \mathbf{Bool}$
- $\mathbf{f}_{1\star} \circ \mathsf{Eq}_{||} = \mathsf{id}$
- $\mathbf{f}_{0\star} \circ \mathsf{Eq}_{||} = \mathsf{Eq} \circ \mathbf{f}_\star$ *for a fixed* $\star \in \mathbf{Bool}$
- $\mathbf{f}_{l\star} \circ \mathbf{C}_\star = \mathsf{id}$ *for* $l \in \{0,1\}$ *and a fixed* $\star \in \mathbf{Bool}$
- $\mathbf{f}_{l\overline{\star}} \circ \mathbf{C}_\star = \mathsf{Eq} \circ \mathbf{f}_{\overline{\star}}$ *for* $l \in \{0,1\}$ *and a fixed* $\star \in \mathbf{Bool}$

Moreover, the compositions $\mathsf{Eq}_{=} \circ \mathsf{Eq}$, $\mathsf{Eq}_{||} \circ \mathsf{Eq}$, $\mathbf{C}_\top \circ \mathsf{Eq}$, $\mathbf{C}_\bot \circ \mathsf{Eq}$ *are all naturally isomorphic.*

The structure described above induces two split cartesian closed fibrations of interest: the first one is induced by combining the first two levels, the categories Set *and* R*, into a cartesian closed reflexive graph category with isomorphisms* R_{PREL}*; this is the fibration* $\mathsf{Fwl}(\mathsf{R}_{PREL})$*. We recall that the objects of* $\mathsf{Fwl}(\mathsf{R}_{PREL})$ *over* n *are pairs* $\{\mathcal{F}(l) : \mathcal{M}(l)^n \to \mathcal{M}(l)\}_{l\in\{0,1\}}$ *of functors that commute with the two face maps from* R *to* Set *on the nose, as well as with the degeneracy* Eq *up to a suitably coherent natural isomorphism*$\varepsilon_\mathcal{F}$*. The morphisms are pairs* $\{\eta(l) : \mathcal{F}(l) \to \mathcal{G}(l)\}_{l\in\{0,1\}}$ *of natural transformations that respect both face maps from* R *to* Set *and the degeneracy* Eq.

The second fibration, which we call Fwl_{2D}*, is induced in much the same way, but taking into account all three levels. This means that the objects over* n *are triples* $\{\mathcal{F}(l) : \mathcal{M}(l)^n \to \mathcal{M}(l)\}_{l\in\{0,1,2\}}$ *of functors that commute with* all *face maps — the two from from* R *to* Set *as well as the four from* $\mathsf{2R}$ *to* R *— on the nose and* all *degeneracies* $\mathsf{Eq}, \mathsf{Eq}_{=}, \mathsf{Eq}_{||}$ *and connections* $\mathbf{C}_\top$*,* $\mathbf{C}_\bot$ *up to suitably coherent natural isomorphisms. Here "suitably coherent" means taking into account not only the equality* $\mathbf{f}_\star \circ \mathsf{Eq} = \mathsf{id}$ *but the additional equalities involving* $\mathsf{Eq}_{=}, \mathsf{Eq}_{||}, \mathbf{C}_\top, \mathbf{C}_\bot$ *as well. This means that, for example, the image of the isomorphism witnessing the commutativity of* $\mathcal{F}$ *with* $\mathsf{Eq}_{=}$ *under the face map* $\mathbf{f}_{1\star}$ *is precisely* $\varepsilon_\mathcal{F} \circ \mathbf{f}_\star$*.*

Analogously, the morphisms are triples $\{\eta(l) : \mathcal{F}(l) \to \mathcal{G}(l)\}_{l\in\{0,1,2\}}$ *of natural transformations that respect all face maps, degeneracies, and connections. We have the obvious forgetful morphism of split cartesian closed fibrations with isomorphisms from* Fwl_{2D} *to* $\mathsf{Fwl}(\mathsf{R}_{PREL})$ *that only retains the structure pertaining to levels 0 and 1.*

The fibration FwI_{2D} admits a family of adjoints to weakening functors as follows. The adjoint $\forall_n\, \mathcal{F}(0)\, \overline{A}$ is the type

$$\begin{aligned}\{&f_0 : \Pi_{A:\mathbb{U}}\mathcal{F}(0)(\overline{A}, A)\,\&\\ &f_1 : \Pi_{R:\mathsf{R}_0}\mathcal{F}(1)(\overline{\mathsf{Eq}\,A}, R)\,(f_0(R_0), f_0(R_1))\,\&\\ &f_2 : \Pi_{Q:2\mathsf{R}_0}\mathcal{F}(2)(\overline{\mathsf{Eq}_=(\mathsf{Eq}_2(A))}, Q)\,\big((f_0\,Q_0^{0\top}, f_0\,Q_1^{0\top}, f_0\,Q_1^{1\top}, f_0\,Q_1^{0\perp}),\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad(f_1\,Q^{0\top}, f_1\,Q^{1\top}, f_1\,Q^{0\perp}, f_1\,Q^{1\perp})\big)\,\&\\ &\Pi_{i:M(0)}\mathcal{F}(0)(\overline{\mathsf{id}_{\mathcal{M}(0)}(A)}, i)\,f_0(i_{\mathsf{c}}) = f_0(i_{\mathsf{d}})\,\&\\ &\Pi_{i:M(1)}\mathcal{F}(1)\big(\overline{\mathsf{id}_{\mathcal{M}(1)}(\mathsf{Eq}\,A)}, i\big)\,\big(f_0\,(i_{\mathsf{d}})_0, f_0\,(i_{\mathsf{d}})_1\big)\,f_1(i_{\mathsf{d}}) = f_1(i_{\mathsf{c}})\}\end{aligned}$$

In the type of f_2, we could have just as well used any of the other functors $\mathsf{Eq}_{\|}$, $\mathbf{C}_\top$, $\mathbf{C}_\perp$ instead of $\mathsf{Eq}_=$ since as observed above, their compositions with Eq are all naturally isomorphic. We next define $\forall_n\, \mathcal{F}(1)\, \overline{R}$ to be the relation with domain $\forall_n\, \mathcal{F}(0)\, \overline{R_0}$ and codomain $\forall_n\, \mathcal{F}(0)\, \overline{R_1}$ mapping $((f_0, f_1, f_2), (g_0, g_1, g_2))$ to

$$\begin{aligned}\{&\phi \quad\;\; : \Pi_{R:\mathsf{R}_0}\mathcal{F}(1)\,(\overline{R}, R)\,(f_0(R_0), g_0(R_1))\,\&\\ &\phi_{\mathsf{Eq}_=} : \Pi_{Q:2\mathsf{R}_0}\mathcal{F}(2)\,(\overline{\mathsf{Eq}_=\,R}, Q)\,\big((f_0\,Q_0^{0\top}, f_0\,Q_1^{0\top}, g_0\,Q_1^{1\top}, g_0\,Q_1^{0\perp}),\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad(f_1\,Q^{0\top}, \phi\,Q^{1\top}, g_1\,Q^{0\perp}, \phi\,Q^{1\perp})\big)\,\&\\ &\phi_{\mathsf{Eq}_{\|}} : \Pi_{Q:2\mathsf{R}_0}\mathcal{F}(2)\,(\overline{\mathsf{Eq}_{\|}\,R}, Q)\,\big((f_0\,Q_0^{0\top}, g_0\,Q_1^{0\top}, f_0\,Q_1^{1\top}, g_0\,Q_1^{0\perp}),\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad(\phi\,Q^{0\top}, f_1\,Q^{1\top}, \phi\,Q^{0\perp}, g_1\,Q^{1\perp})\big)\,\&\\ &\phi_{\mathbf{C}_\top} : \Pi_{Q:2\mathsf{R}_0}\mathcal{F}(2)\,(\overline{\mathbf{C}_\top\,R}, Q)\,\big((f_0\,Q_0^{0\top}, g_0\,Q_1^{0\top}, g_0\,Q_1^{1\top}, g_0\,Q_1^{0\perp}),\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad(\phi\,Q^{0\top}, \phi\,Q^{1\top}, g_1\,Q^{0\perp}, g_1\,Q^{1\perp})\big)\,\&\\ &\phi_{\mathbf{C}_\perp} : \Pi_{Q:2\mathsf{R}_0}\mathcal{F}(2)\,(\overline{\mathbf{C}_\perp\,R}, Q)\,\big((f_0\,Q_0^{0\top}, f_0\,Q_1^{0\top}, f_0\,Q_1^{1\top}, g_0\,Q_1^{0\perp}),\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad(f_1\,Q^{0\top}, f_1\,Q^{1\top}, \phi\,Q^{0\perp}, \phi\,Q^{1\perp})\big)\,\&\\ &\Pi_{i:M(1)}\mathcal{F}(1)(\overline{\mathsf{id}_{\mathcal{M}(1)}(R)}, i)\,\big(f_0\,(i_{\mathsf{d}})_0, g_0\,(i_{\mathsf{d}})_1\big)\,\phi_1(i_{\mathsf{d}}) = \phi_1(i_{\mathsf{c}})\,\}\end{aligned}$$

The component $\phi_{\mathsf{Eq}_=}$ asserts that ϕ appropriately interacts with the degeneracy $\mathsf{Eq}_=$ and similarly for the analogous components $\phi_{\mathsf{Eq}_{\|}}$, $\phi_{\mathbf{C}_\top}$, $\phi_{\mathbf{C}_\perp}$. We define $\forall_n\, \mathcal{F}(2)\, \overline{Q}$ to be the 2-relation with underlying tuple of relations

$$\big(\forall_n\mathcal{F}(1)\,\overline{Q^{0\top}}, \forall_n\,\mathcal{F}(1)\,\overline{Q^{1\top}}, \forall_n\,\mathcal{F}(1)\,\overline{Q^{0\perp}}, \forall_n\,\mathcal{F}(1)\,\overline{Q^{1\perp}}\big)$$

mapping $\big(((f_0, f_1, f_2), (g_0, g_1, g_2), (h_0, h_1, h_2), (l_0, l_1, l_2)), ((\phi_0, \ldots), (\phi_1, \ldots), (\phi_2, \ldots), (\phi_3, \ldots))\big)$ to the proposition

$$\begin{aligned}\Pi_{Q:2\mathsf{R}_0}\mathcal{F}(2)\,(\overline{Q}, Q)\,\big(&(f_0\,Q_0^{0\top}, g_0\,Q_1^{0\top}, h_0\,Q_0^{0\perp}, l_0\,Q_1^{0\perp}),\\ &(\phi_0\,Q^{0\top}, \phi_1\,Q^{1\top}, \phi_2\,Q^{0\perp}, \phi_3\,Q^{1\perp})\big)\end{aligned}$$

Finally, unlike the frameworks [1, 2, 3, 5, 7, 13], our definition of a parametric model recognizes the above proof-relevant model:

Theorem 76 (Proof-relevant model)**.** *The family of adjoints defined in Example 75 makes* Fwl_{2D} *into a split* $\lambda 2$*-fibration with isomorphisms, and hence into a parametric model of System F over* R_{PREL}*.*

Proof sketch. Faithfulness follows because having $\eta(0), \eta(1)$ fixed, there is a unique way to define $\eta(2)$: since η has to respect the degeneracy $\mathsf{Eq}_=$, we must have

$$\eta(2) \circ_{\mathcal{R}(2)} \varepsilon^{\overline{\overline{=}}}_{\mathcal{F}} = \varepsilon^{\overline{\overline{=}}}_{\mathcal{G}} \circ_{\mathcal{R}(2)} \mathsf{Eq}_=(\eta(1))$$

where $\varepsilon^{\overline{\overline{=}}}_{\mathcal{F}}$, $\varepsilon^{\overline{\overline{=}}}_{\mathcal{G}}$ are the natural isomorphisms witnessing the fact that $\mathcal{F}$, $\mathcal{G}$ by assumption preserve $\mathsf{Eq}_=$ (we could have used any of the other functors $\mathsf{Eq}_{\|}, \mathbf{C}_\top, \mathbf{C}_\bot$ as well). This gives at most one possible value for $\eta(2)$. Fullness follows since the triple $\{\eta(l)\}_{l\in\{0,1,2\}}$ with $\eta(2)$ as given above indeed respects all face maps, degeneracies, and connections (in fact it is only necessary to check the respecting of face maps since the predicates at level 2 are proof-irrelevant). Finally, essential surjectivity follows from the fact that the reflexive graph functor $(\mathcal{F}(0), \mathcal{F}(1))$ is isomorphic to the reflexive graph functor $(\mathcal{F}(0), \mathsf{Eq}(\mathcal{F}(0)))$ via the reflexive graph natural transformation $(\mathsf{id}, \varepsilon_{\mathcal{F}})$ (the fact that this transformation is face-map preserving again uses the coherence $\mathcal{R}(\mathbf{f}_\star) \circ \varepsilon_{\mathcal{F}} = \mathsf{id}$). But $(\mathcal{F}(0), \mathsf{Eq}(\mathcal{F}(0)))$ clearly belongs to the image since it can be extended *e.g.*, to the triple $\big(\mathcal{F}(0), \mathsf{Eq}(\mathcal{F}(0)), \mathsf{Eq}_=(\mathsf{Eq}(\mathcal{F}(0)))\big)$. □

7 Discussion

We can now be more specific about how our approach compares to the external approaches in [2, 5, 7, 13], all of which are based on a reflexive graph of split $\lambda 2$-fibrations. We already noted that the fibration corresponding to Reynolds' model is not a split $\lambda 2$-fibration, but only a split $\lambda 2$-fibration *with isomorphisms*. It thus is not a direct instance of the definitions found in [2, 5, 7, 13]. The same is true for the fibration corresponding to the proof-relevant model, but there we have an even bigger problem: it is unclear how to define the family of adjoints for the second fibration (called r in [5]) of "heterogeneous" reflexive graph functors in a way that is compatible with the adjoint structure on the original fibration. This is because unlike in the proof-irrelevant case, the definition of $\forall_n \mathcal{F}(1)$ now has conditions such as $\phi_=$ which are only meaningful for "homogeneous" reflexive graph functors, *i.e.*, those where the domain and codomain of $\mathcal{F}(1)(\overline{R})$ are given by the *same* functor $\mathcal{F}(1)$, albeit applied to different arguments ($\overline{R_0}$ vs. $\overline{R}_1$).

We indicate three directions for future work. Readers interested with *applications of parametricity* will notice that we do not require conditions such as fullness or (op)cartesianness of certain maps or well-pointedness of certain categories. This follows the spirit of [5], where the notion of *parametricity* pertains to the suitable interaction with (what we call) face maps and degeneracies. Specific *applications* such as establishing the Graph Lemma and the existence of initial algebras are left for another occasion. Readers fond of *type theory* might wonder about possible models expressed in the *intensional* version of MLTT. Although currently there are no well-known models for which the latter would be the right choice of meta-theory, that might change with more research into higher notions of parametricity. Finally, readers familiar with *cubical sets* no doubt recognized the structure of sets, relations, and 2-relations with face maps, degeneracies, and connections from the last section as the first few levels of the cubical hierarchy, and wonder whether one can formulate the analogous notion of $2, 3, \ldots$ or even ∞-parametricity using this hierarchy. We conjecture the answer to be a *YES!* and plan to pursue this question in future work.

Acknowledgments This research is supported by NSF awards 1420175 and 1545197. We thank Steve Awodey for helpful discussions.

References

[1] B. Dunphy and U. Reddy. Parametric limits. In *Logic in Computer Science*, pages 242–251, 2004.

[2] N. Ghani, F. Nordvall Forsberg, and A. Simpson. Comprehensive parametric polymorphism: Categorical models and type theory. In *Foundations of Software Science and Computation Structures*, pages 3–19, 2016.

[3] N. Ghani, P. Johann, F. Nordvall Forsberg, F. Orsanigo, and T. Revell. Bifibrational Functorial Semantics for Parametric Polymorphism. In *Mathematical Foundations of Program Semantics*, pages 165–181, 2015.

[4] J.-Y. Girard, P. Taylor, and Y. Lafont. *Proofs and Types*. Cambridge University Press, 1989.

[5] B. Jacobs. *Categorical Logic and Type Theory*. Elsevier, 1999.

[6] G. Longo and E. Moggi. Constructive natural deduction and its ω-set interpretation. *Mathematical Structures in Computer Science*, 1(2):215–254, 2009.

[7] Q. Ma and J. Reynolds. Types, abstraction, and parametric polymorphism. In *Mathematical Foundations of Program Semantics*, pages 1–40, 1992.

[8] F. Orsanigo. *Bifibrational Parametricity: From Zero to Two Dimensions*. PhD thesis, University of Strathclyde, 2017.

[9] A. Pitts. Polymorphism is set theoretic, constructively. *Category Theory and Computer Science*, pages 12–39, 1987.

[10] A. M. Pitts. Categorical logic. In *Handbook of Logic in Computer Science, Volume 5. Algebraic and Logical Structures*, chapter 2. Oxford University Press, 2000.

[11] J. Reynolds. Types, Abstraction, and Parametric Polymorphism. In *International Federation for Information Processing*, pages 513–523, 1983.

[12] J. C. Reynolds. Polymorphism is not set-theoretic. *Semantics of Data Types*, pages 145–156, 1984.

[13] E. Robinson and G. Rosolini. Reflexive graphs and parametric polymorphism. In *Logic in Computer Science*, pages 364–371, 1994.

[14] R. A. Seely. Categorical semantics for higher order polymorphic lambda calculus. *Journal of Symbolic Logic*, pages 969–989, 1987.

[15] P. Wadler. Theorems for free! In *Functional Programming and Computer Architecture*, pages 347–359, 1987.