

The Combination of Dynamic and Static Typing from a Categorical Perspective

Harley Eades III and Michael Townsend

Augusta University

{heades,mitownsend}@augusta.edu

Categories and Subject Descriptors D.1.1 [Programming Technique]: [Applicative (Functional) Programming]; D.3.1 [Formal Definitions and Theory]: [Semantics]; F.3.3 [Studies of Program Constructs]: [Type structure]; F.4.1 [Mathematical Logic and Formal Languages]: [Lambda calculus and related systems]

General Terms typed lambda-calculus, untyped lambda-calculus, gradual typing, static typing, dynamic typing, categorical model, functional programming

Keywords static typing, dynamic typing, gradual typing, categorical semantics, retract

Abstract

Gradual typing was first proposed by Siek and Taha in 2006 as a way for a programming language to combine the strengths of both static and dynamic typing. However one question we must ask is, what is gradual typing? This paper contributes to answering this question by providing the first categorical model of gradual typing using the seminal work of Scott and Lambek on the categorical models of the untyped and typed λ -calculus. We then extract a functional programming language, called Grady, from the categorical model using the Curry-Howard-Lambek correspondence that combines both static and dynamic typing, but Grady is an annotated language and not a gradual type system. Finally, we show that Siek and Taha’s gradual type system can be translated into Grady, and that their original annotated language is equivalent in expressive power to Grady.

1. Introduction

(Scott 1980) showed how to model the untyped λ -calculus within a cartesian closed category, \mathcal{C} , with a distinguished object we will call $?$ – read as the type of untyped terms – such that the object $?$ is a retract of $?$. That is, there are morphisms $\text{squash} : (? \rightarrow ?) \rightarrow ?$ and $\text{split} : ? \rightarrow (? \rightarrow ?)$ where $\text{squash}; \text{split} = \text{id} : (? \rightarrow ?) \rightarrow (? \rightarrow ?)$. For example, taking these morphisms as

¹ We will use the terms “object” and “type” interchangeably.

² We denote composition of morphisms by $f; g : A \rightarrow C$ given morphisms $f : A \rightarrow B$ and $g : B \rightarrow C$.

terms in the typed λ -calculus we can define the prototypical looping term $(\lambda x.x x)(\lambda x.x x)$ by $(\lambda x : ?.(split\ x)\ x)$ (squash $(\lambda x : ?.(split\ x)\ x)$).

In the same volume as Scott (Lambek 1980) showed that cartesian closed categories also model the typed λ -calculus. Suppose we want to model the typed λ -calculus with pairs and natural numbers. That is, given two types A_1 and A_2 there is a type $A_1 \times A_2$, and there is a type Nat . Furthermore, we have first and second projections, and zero and successor functions. This situation can easily be modeled by a cartesian closed category \mathcal{C} – see Section 2 for the details – but also add to \mathcal{C} the type of untyped terms $?$, squash, and split. At this point \mathcal{C} is a model of both the typed and the untyped λ -calculus. However, the two theories are really just sitting side by side in \mathcal{C} and cannot really interact much.

Suppose \mathcal{T} is a discrete category with the objects Nat and 1 (the terminal object or empty product) and $T : \mathcal{T} \rightarrow \mathcal{C}$ is a full and faithful functor. This implies that \mathcal{T} is a subcategory of \mathcal{C} , and that \mathcal{T} is the category of atomic types. Then for any type A of \mathcal{T} we add to \mathcal{C} the morphisms $\text{box} : TA \rightarrow ?$ and $\text{unbox} : ? \rightarrow TA$ such that $\text{box}; \text{unbox} = \text{id} : TA \rightarrow TA$ making TA a retract of $?$. This is the bridge allowing the typed world to interact with the untyped one. We can think of box as injecting typed data into the untyped world, and unbox as taking it back. Notice that the only time we can actually get the typed data back out is if it were injected into the untyped world initially. In the model this is enforced through composition, but in the language this will be enforced at runtime, and hence, requires the language to contain dynamic typing. Thus, what we have just built up is a categorical model that offers a new perspective of how to combine static and dynamic typing.

(Siek and Taha 2006) define gradual typing to be the combination of both static and dynamic typing that allows for the programmer to program in dynamic style, and thus, annotations should be suppressed. This means that a gradually typed program can utilize both static types which will be enforced during compile time, but may also utilize dynamic typing that will be enforced during runtime. Therefore, gradual typing is the best of both worlds.

Siek and Taha’s gradually typed functional language is the typed λ -calculus with the type of untyped terms $?$ and the following rules:

$$\frac{\Gamma \vdash t_1 : ? \quad \Gamma \vdash t_2 : A}{\Gamma \vdash t_1 t_2 : ?} \quad \frac{\Gamma \vdash t_1 : A_1 \rightarrow B \quad \Gamma \vdash t_2 : A_2 \quad A_1 \sim A_2}{\Gamma \vdash t_1 t_2 : B}$$

The premise $A \sim B$ is read, the type A is consistent with the type B , and is defined by the following rules:

$$\frac{\overline{A \sim A} \quad \overline{A \sim ?} \quad \overline{? \sim A}}{\frac{A_1 \sim A_2 \quad B_1 \sim B_2}{A_1 \rightarrow B_1 \sim A_2 \rightarrow B_2} \quad \frac{A_1 \sim A_2 \quad B_1 \sim B_2}{A_1 \times B_1 \sim A_2 \times B_2}}$$

This is a reflexive and symmetric relation, but is a non-transitive relation. If we squint we can see split, squash, box, and unbox hiding in the definition of the previous rules, but they have been suppressed. We will show that when one uses either of the two typing rules then one is really implicitly using a casting morphism built from split, squash, box, and unbox. In fact, the consistency relation $A \sim B$ can be interpreted as such a morphism. Then the typing above can be read semantically as a saying if a casting morphism exists, then the type really can be converted into the necessary type.

Contributions. This paper offers the following contributions:

- A new categorical model for gradual typing for functional languages. We show how to interpret (Siek and Taha 2006)’s gradual type system in the categorical model outlined above. As far as the authors are aware this is the first categorical model for gradual typing.
- We then extract a functional programming language called Grady from the categorical model via the Curry-Howard-Lambek correspondence. This is not a gradual type system, but can be seen as an alternative annotated language in which Siek and Taha’s gradual type system can be translated to.
- A proof that Grady is as expressive as (Siek and Taha 2006)’s annotated language and vice versa. We give a type directed translation of Siek and Taha’s annotated language to Grady and vice versa, then we show that these translations preserve evaluation.
- Having the untyped λ -calculus along side the typed λ -calculus can be a lot of fun. We show how to Church encode typed data, utilize the Y-combinator, and even obtain terminating recursion on natural numbers by combining the Y-combinator with a natural number eliminator. Thus, obtaining the expressive power of Gödel’s system T (Girard et al. 1989).

Related work. TODO

2. Categorical Model

Definition 1. Suppose \mathcal{C} is a category. Then an object A is a **retract** of an object B if there are morphisms $i : A \rightarrow B$ and $r : B \rightarrow A$ such that the following diagram commutes:

$$\begin{array}{ccc} A & \xrightarrow{i} & B \\ & \searrow & \downarrow r \\ & & A \end{array}$$

Definition 2. An **untyped λ -model**, $(\mathcal{C}, ?, \text{split}, \text{squash})$, is a cartesian closed category \mathcal{C} with a distinguished object $?$ and two morphisms $\text{squash} : (? \rightarrow ?) \rightarrow ?$ and $\text{split} : ? \rightarrow (? \rightarrow ?)$ making the object $?$ a retract of $?$.

Theorem 3 (Scott (1980)). An untyped λ -model is a sound and complete model of the untyped λ -calculus.

Definition 4. An object \mathbb{N} of a category \mathcal{C} with a terminal object 1 is a **natural number object (NNO)** if and only if there are morphisms $z : 1 \rightarrow \mathbb{N}$ and $s : \mathbb{N} \rightarrow \mathbb{N}$ such that for any other object A and morphisms $z' : 1 \rightarrow A$ and $s' : A \rightarrow A$ there is a unique morphism $u : \mathbb{N} \rightarrow A$ making the following diagram

commute:

$$\begin{array}{ccccc} 1 & \xrightarrow{z} & \mathbb{N} & \xrightarrow{s} & \mathbb{N} \\ & \searrow z' & \downarrow u & & \downarrow u \\ & & A & \xrightarrow{s'} & A \end{array}$$

Definition 5. A **gradual λ -model**, $(\mathcal{T}, \mathcal{C}, ?, \text{split}, \text{squash}, \text{box}, \text{unbox})$, where \mathcal{T} and \mathcal{C} are cartesian closed categories with NNOS, $(\mathcal{C}, ?, \text{split}, \text{squash})$ is an untyped λ -model, $T : \mathcal{T} \rightarrow \mathcal{C}$ is a cartesian closed embedding – a full and faithful cartesian closed functor that is injective on objects and preserves the NNO – and for every object, A , of \mathcal{T} there are morphisms $\text{box}_A : TA \rightarrow ?$ and $\text{unbox}_A : ? \rightarrow TA$ making TA a retract of $?$.

3. Grady

References

- Jean-Yves Girard, Yves Lafont, and Paul Taylor. *Proofs and Types (Cambridge Tracts in Theoretical Computer Science)*. Cambridge University Press, April 1989.
- Joachim Lambek. From lambda calculus to cartesian closed categories. *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 376–402, 1980.
- Dana Scott. Relating theories of the lambda-calculus. In *To H.B. Curry: Essays on Combinatory Logic, Lambda-Calculus and Formalism* (eds. Hindley and Seldin), pages 403–450. Academic Press, 1980.
- Jeremy G Siek and Walid Taha. Gradual typing for functional languages. In *Scheme and Functional Programming Workshop*, volume 6, pages 81–92, 2006.

A. The Complete Spec of Grady

term var, x, y, z

index, k

$t ::=$

x	variable
triv	unit
squash_S	injection of the retract
split_S	surjection of the retract
box_T	generalize to the untyped universe
unbox_T	specialize the untyped universe to a specific type
$\langle A \rangle t$	type cast
$\lambda x : A. t$	λ -abstraction
$t_1 t_2$	function application
(t_1, t_2)	pair constructor
$\text{fst } t$	first projection
$\text{snd } t$	second projection
$\text{succ } t$	successor function
0	zero
(t)	S

$s ::=$

x	simple values
triv	unit
0	zero

$v ::=$

x	values
triv	unit
0	zero

		$\langle ? \rangle_S$	untyped cast
h	::=	triv split_S squash_S box_T unbox_T $\lambda x : A. t$ (t_1, t_2) $\text{fst } t$ $\text{snd } t$ $\text{succ } t$ 0	head-normal forms
T	::=	1 Nat $T_1 \rightarrow T_2$ $T_1 \times T_2$ (T)	terminating types unit type natural number type function type cartesian product type S
S	::=	$? \rightarrow ?$ $? \times ?$	
C	::=	1 Nat	atomic type unit type natural number type
A, B	::=	1 Nat $?$ $A_1 \rightarrow A_2$ $A_1 \times A_2$ (A)	type unit type natural number type untyped universe function type cartesian product type S
Γ	::=	\cdot $\Gamma, x : A$	typing context empty context cons
vd	::=	\vdash \nmid	

$\boxed{A \sim B}$ A is consistent with B

$$\begin{array}{c}
 \overline{A \sim A} \\
 \overline{A \sim ?} \\
 \overline{? \sim A} \\
 \frac{A_1 \sim A_2 \quad B_1 \sim B_2}{A_1 \rightarrow B_1 \sim A_2 \rightarrow B_2} \\
 \frac{A_1 \sim A_2 \quad B_1 \sim B_2}{A_1 \times B_1 \sim A_2 \times B_2}
 \end{array}$$

$\boxed{\Gamma vdt : A}$ t has type A in context Γ

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A} \quad \text{VAR}$$

$\overline{\Gamma \vdash \text{box}_T : T \rightarrow ?}$	BOX
$\overline{\Gamma \vdash \text{unbox}_T : ? \rightarrow T}$	UNBOX
$\overline{\Gamma \vdash \text{squash}_S : S \rightarrow ?}$	SQUASH
$\overline{\Gamma \vdash \text{split}_S : ? \rightarrow S}$	SPLIT
$\overline{\Gamma \vdash \text{triv} : 1}$	UNIT
$\overline{\Gamma \vdash 0 : \text{Nat}}$	ZERO
$\overline{\Gamma \vdash t : \text{Nat}}$	SUCC
$\overline{\Gamma \vdash \text{succ } t : \text{Nat}}$	
$\frac{\Gamma \vdash t_1 : A_1 \quad \Gamma \vdash t_2 : A_2}{\Gamma \vdash (t_1, t_2) : A_1 \times A_2}$	PAIR
$\frac{\Gamma \vdash t : A_1 \times A_2}{\Gamma \vdash \text{fst } t : A_1}$	FST
$\frac{\Gamma \vdash t : A_1 \times A_2}{\Gamma \vdash \text{snd } t : A_2}$	SND
$\frac{\Gamma, x : A_1 \vdash t : A_2}{\Gamma \vdash \lambda x : A_1. t : A_1 \rightarrow A_2}$	LAM
$\frac{\Gamma \vdash t_1 : A_1 \rightarrow A_2 \quad \Gamma \vdash t_2 : A_1}{\Gamma \vdash t_1 t_2 : A_2}$	APP
$\frac{\Gamma \vdash t_1 : ? \quad \Gamma \vdash t_2 : A}{\Gamma \vdash t_1 t_1 : ?}$	
$\frac{\Gamma \vdash t_1 : A_1 \rightarrow B \quad \Gamma \vdash t_2 : A_2 \quad A_1 \sim A_2}{\Gamma \vdash t_1 t_2 : B}$	
$\frac{\Gamma \vdash t : A \quad A \sim B}{\Gamma \vdash \langle B \rangle t : B}$	
$\boxed{\Gamma \vdash t_1 \rightsquigarrow t_2 : A}$	t_1 reduces to t_2 with type A in context Γ
$\frac{\Gamma \vdash s : A}{\Gamma \vdash s \rightsquigarrow s : A}$	RD-VALUES
$\frac{\Gamma \vdash t : T}{\Gamma \vdash \text{unbox}_T (\text{box}_T t) \rightsquigarrow t : T}$	RD-RETRACT
$\frac{\Gamma \nmid t : T}{\Gamma \vdash \text{unbox}_T (\text{box}_{T'} t) \rightsquigarrow \text{wrong} : \text{TypeError}}$	RD-TWRONG
$\frac{h \neq \text{box}_T t}{\Gamma \vdash \text{unbox}_T h \rightsquigarrow \text{wrong} : \text{TypeError}}$	RD-HWRONG
$\frac{\Gamma \vdash t : S}{\Gamma \vdash \text{split}_S (\text{squash}_S t) \rightsquigarrow t : S}$	RD-RETRACTU
$\frac{\Gamma, x : A_1 \vdash t_2 : A_2 \quad \Gamma \vdash t_1 : A_1}{\Gamma \vdash (\lambda x : A_1. t_2) t_1 \rightsquigarrow [t_1/x] t_2 : A_2}$	RD-BETA
$\frac{\Gamma \vdash t : A_1 \rightarrow A_2 \quad x \notin \text{FV}(t)}{\Gamma \vdash \lambda x : A_1. t x \rightsquigarrow t : A_1 \rightarrow A_2}$	RD-ETA
$\frac{\Gamma \vdash t_1 : A_1 \quad \Gamma \vdash t_2 : A_2}{\Gamma \vdash \text{fst } (t_1, t_2) \rightsquigarrow t_1 : A_1}$	RD-PROJ1
$\frac{\Gamma \vdash t_1 : A_1 \quad \Gamma \vdash t_2 : A_2}{\Gamma \vdash \text{snd } (t_1, t_2) \rightsquigarrow t_2 : A_2}$	RD-PROJ2

$$\begin{array}{c}
\frac{\Gamma \vdash t : A_1 \times A_2}{\Gamma \vdash (\mathbf{fst} \, t, \mathbf{snd} \, t) \rightsquigarrow t : A_1 \times A_2} \text{RD_ETAP} \\
\\
\frac{\Gamma, x : A_1 \vdash t \rightsquigarrow t' : A_2}{\Gamma \vdash \lambda x : A_1. t \rightsquigarrow \lambda x : A_1. t' : A_1 \rightarrow A_2} \text{RD_LAM} \\
\\
\frac{\Gamma \vdash t_1 \rightsquigarrow t'_1 : A_1 \rightarrow A_2 \quad \Gamma \vdash t_2 : A_1}{\Gamma \vdash t_1 \, t_2 \rightsquigarrow t'_1 \, t_2 : A_2} \text{RD_APP1} \\
\\
\frac{\Gamma \vdash t_1 : A_1 \rightarrow A_2 \quad \Gamma \vdash t_2 \rightsquigarrow t'_2 : A_1}{\Gamma \vdash t_1 \, t_2 \rightsquigarrow t_1 \, t'_2 : A_2} \text{RD_APP2} \\
\\
\frac{\Gamma \vdash t \rightsquigarrow t' : A_1 \times A_2}{\Gamma \vdash \mathbf{fst} \, t \rightsquigarrow \mathbf{fst} \, t' : A_1} \text{RD_FST} \\
\\
\frac{\Gamma \vdash t \rightsquigarrow t' : A_1 \times A_2}{\Gamma \vdash \mathbf{snd} \, t \rightsquigarrow \mathbf{snd} \, t' : A_2} \text{RD_SND} \\
\\
\frac{\Gamma \vdash t_1 \rightsquigarrow t'_1 : A_1 \quad \Gamma \vdash t_2 : A_2}{\Gamma \vdash (t_1, t_2) \rightsquigarrow (t'_1, t_2) : A_1 \times A_2} \text{RD_PAIR1} \\
\\
\frac{\Gamma \vdash t_1 : A_1 \quad \Gamma \vdash t_2 \rightsquigarrow t'_2 : A_2}{\Gamma \vdash (t_1, t_2) \rightsquigarrow (t_1, t'_2) : A_1 \times A_2} \text{RD_PAIR2}
\end{array}$$

$\boxed{\Gamma \vdash t_1 \rightsquigarrow t_2 : A}$ Reduction for annotated Siek16

$$\begin{array}{c}
\frac{\Gamma \vdash v : A}{\Gamma \vdash v \rightsquigarrow v : A} \text{RDA_VALUES} \\
\\
\frac{\Gamma \vdash \mathbf{drop-cast} \, v : C}{\Gamma \vdash \langle C \rangle v \rightsquigarrow \mathbf{drop-cast} \, v : C} \text{RDA_CASTA} \\
\\
\frac{\Gamma \vdash t : A_1 \rightarrow B_1 \quad (A_1 \rightarrow B_1) \sim (A_2 \rightarrow B_2)}{\Gamma \vdash \langle A_2 \rightarrow B_2 \rangle t \rightsquigarrow \lambda y : A_2. \langle B_2 \rangle (t \, \langle A_1 \rangle y) : A_2 \rightarrow B_2} \text{RDA_CASTLAM} \\
\\
\frac{\Gamma \vdash t : A_1 \times B_1 \quad (A_1 \times B_1) \sim (A_2 \times B_2)}{\Gamma \vdash \langle A_2 \times B_2 \rangle t \rightsquigarrow (\langle A_2 \rangle (\mathbf{fst} \, t), \langle B_2 \rangle (\mathbf{snd} \, t)) : A_2 \times B_2} \text{RDA_CASTPAIR} \\
\\
\frac{\Gamma \vdash t_1 \rightsquigarrow t_2 : A \quad A \sim B}{\Gamma \vdash \langle B \rangle t_1 \rightsquigarrow \langle B \rangle t_2 : B} \text{RDA_CAST} \\
\\
\frac{\Gamma, x : A_1 \vdash t_2 : A_2 \quad \Gamma \vdash t_1 : A_1}{\Gamma \vdash (\lambda x : A_1. t_2) \, t_1 \rightsquigarrow [t_1/x] t_2 : A_2} \text{RDA_BETA} \\
\\
\frac{\Gamma \vdash t : A_1 \rightarrow A_2 \quad x \notin \mathbf{FV}(t)}{\Gamma \vdash \lambda x : A_1. t \rightsquigarrow t : A_1 \rightarrow A_2} \text{RDA_ETA} \\
\\
\frac{\Gamma, x : A_1 \vdash t \rightsquigarrow t' : A_2}{\Gamma \vdash \lambda x : A_1. t \rightsquigarrow \lambda x : A_1. t' : A_1 \rightarrow A_2} \text{RDA_LAM} \\
\\
\frac{\Gamma \vdash t_1 \rightsquigarrow t'_1 : A_1 \rightarrow A_2 \quad \Gamma \vdash t_2 : A_1}{\Gamma \vdash t_1 \, t_2 \rightsquigarrow t'_1 \, t_2 : A_2} \text{RDA_APP1} \\
\\
\frac{\Gamma \vdash t_1 : A_1 \rightarrow A_2 \quad \Gamma \vdash t_2 \rightsquigarrow t'_2 : A_1}{\Gamma \vdash t_1 \, t_2 \rightsquigarrow t_1 \, t'_2 : A_2} \text{RDA_APP2} \\
\\
\frac{\Gamma \vdash t \rightsquigarrow t' : A_1 \times A_2}{\Gamma \vdash \mathbf{fst} \, t \rightsquigarrow \mathbf{fst} \, t' : A_1} \text{RDA_FST} \\
\\
\frac{\Gamma \vdash t \rightsquigarrow t' : A_1 \times A_2}{\Gamma \vdash \mathbf{snd} \, t \rightsquigarrow \mathbf{snd} \, t' : A_2} \text{RDA_SND} \\
\\
\frac{\Gamma \vdash t : A_1 \times A_2}{\Gamma \vdash (\mathbf{fst} \, t, \mathbf{snd} \, t) \rightsquigarrow t : A_1 \times A_2} \text{RDA_ETAP} \\
\\
\frac{\Gamma \vdash t_1 \rightsquigarrow t'_1 : A_1 \quad \Gamma \vdash t_2 : A_2}{\Gamma \vdash (t_1, t_2) \rightsquigarrow (t'_1, t_2) : A_1 \times A_2} \text{RDA_PAIR1} \\
\\
\frac{\Gamma \vdash t_1 : A_1 \quad \Gamma \vdash t_2 \rightsquigarrow t'_2 : A_2}{\Gamma \vdash (t_1, t_2) \rightsquigarrow (t_1, t'_2) : A_1 \times A_2} \text{RDA_PAIR2}
\end{array}$$