Technical Report

UCAM-CL-TR-336
ISSN 1476-2986

Number 336

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Towards a proof theory of rewriting: the simply-typed 2-λ calculus

Barnaby P. Hilken

May 1994

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

https://www.cl.cam.ac.uk/

Technical reports published by the University of Cambridge
Computer Laboratory are freely available via the Internet:

https://www.cl.cam.ac.uk/techreports/

ISSN 1476-2986

TOWARDS A PROOF THEORY OF REWRITING: THE SIMPLY-TYPED 2-λ CALCULUS

BARNABY P. HILKEN

ABSTRACT. This paper describes the simply-typed 2-λ-calculus, a language with three levels: types, terms and rewrites. The types and terms are those of the simply-typed λ-calculus, and the rewrites are expressions denoting sequences of β-reductions and η-expansions. An equational theory is imposed on the rewrites, based on 2-categorical justifications, and the word problem for this theory is solved by finding a canonical expression in each equivalence class.

The canonical form of rewrites allows us to prove several properties of the calculus, including a strong form of confluence and a classification of the long-β-η-normal forms in terms of their rewrites. Finally we use these properties as the basic definitions of a theory of categorical rewriting, and find that the expected relationships between confluence, strong normalisation and normal forms hold.

1. INTRODUCTION

In the theoretical computer science community recently there has been much interest in *proof theory*: the study of logics not in terms of their consequence relations, but in terms of their proofs. The point of interest is not just *whether* propositions are provable, but *how* they are proved, and what mathematical structure can be given to proofs. This raises the question of which proofs should be considered equivalent, and which are distinct. Traditional proof theory answers this with the notions of cut elimination (for sequent calculus) and proof normalisation (for natural deduction), which identify proofs by certain syntactic rules. Typically, the introduction followed by immediate elimination of a connective is equated with a trivial proof. This kind of syntactic rule is really justified by the fact that it works: we are left with the suspicion that there might be another way to do it.

Categorical logic provides an alternative, more mathematical, approach to the same problem, at least for intuitionistic logic. Here the propositions and proofs of a logic are taken to be the objects and arrows of a category respectively, and two proofs are equal if, and only if, the corresponding arrows are forced to be equal by the axioms of category theory. In other words, the logic is identified with a free category of a certain form, depending on the connectives of the logic. The connectives are given universal properties: conjunction as product, disjunction as coproduct and implication as exponential, for example. Since universal properties characterise objects up to isomorphism, this gives a more convincing reason for identifying proofs.

This work was carried out as part of a SERC-funded project while the author was at the University of Manchester. The paper was completed while the author was a SERC postdoctoral fellow at the University of Cambridge.

Categorical proof theory arises from the observation that the identifications justified by the category theory are the same as the traditional syntactic ones.

The aim of this work is to develop a proof theory for rewriting. Our analogy is this: the elements (terms, strings, etc.) of a rewrite system correspond to the propositions of a logic, and the rewrite relation $t \to_* s$ (t rewrites in zero or more steps to s) corresponds to the consequence relation. The analogue of a proof we call a *rewrite*, and we write $\alpha\colon t \Rightarrow s$ when α is a rewrite whose effect is to transform t into s. Just as proofs say how propositions are proved, so the rewrite α says how t is rewritten to get s. We can think of α as an algorithm—perhaps as simple as a sequence of instances of rewrite rules—which expresses the necessary computational information. The questions we wish to study are: what form do such algorithms take, what mathematical structure do they have, and when are two of them equal?

The reflexivity and transitivity of the relation $\to_*$ suggest that we push our analogy further, and try to develop categorical rewriting. We take the elements and rewrites of a rewrite system to be the objects and arrows of a category respectively. Composition of arrows is sequential composition of rewrites, corresponding to the transitivity of $\to_*$, and identity arrows are "zero-step" rewrites, corresponding to reflexivity. We can then look for categorical justification of identifications between rewrites. In particular, we would hope that Seely's description of β-reduction and η-expansion as unit and counit of an adjunction [7] would fit this framework.

In this paper we study one particular rewrite system, the simply-typed λ-calculus, in some detail. We define the types and terms in the usual way, and give a language for rewrites generated from β-reduction and η-expansion by sequential and parallel composition. We then introduce equations between rewrites which are motivated by categorical considerations similar to those of Seely. These equations lead to a simple canonical form for rewrites, which solves the word problem, and allows us to prove several results about our system.

Generalising from this example, we then define categorical rewriting by which we mean a theory of rewriting which concerns not just the relation $\to_*$ but the rewrites themselves. We give a condition on categories which ensures that they act like rewrite systems, and categorical definitions of confluence, normal forms and strong normalisation. We prove that our example has the properties of confluence and strong normalisation, and that the normal forms are precisely Huet's long $\beta\eta$-normal forms [6]. Finally we show that our definitions are linked in the expected way: strong normalisation implies existence of normal forms, and confluence implies their uniqueness (up to isomorphism).

2. The Simply-Typed 2-λ-Calculus

The simply-typed 2-λ-calculus is a language of three syntactic classes, called **types**, **terms** and **rewrites**. Each term has a **context** which gives the types of free variables which might appear in the term, and a type. Each rewrite has a source term and a target term, which share a common context and type. The well-formedness conditions are expressed by two judgements:

- $\Gamma \vdash t\colon X$ means that t is a well-formed term of type X in context Γ.
- $\Gamma \vdash \gamma\colon t \Rightarrow u\colon X$ means that γ is a well-formed rewrite with source t and target u, where t and u are well-formed terms of type X in context Γ.

The intended interpretation of the language is that the types and terms are those of the simply-typed λ-calculus, and the rewrites are algorithms which describe a sequence of β-reductions and η-expansions which can be applied to a term.

2.1. Syntax. Let $\mathbb{B}$ be a set of "basic types", with typical element B. The language is defined inductively as follows, where (in order to simplify several points) DeBruijn notation is used for variables.

Types.

$$X ::= B \mid X \to X$$

Since this is a simply-typed calculus, a type is built up from basic types using the $\to$ (function space) constructor.

A context Γ is just a list of types $X_1, \ldots, X_n$.

Terms.

$$X_1, \ldots, X_n \vdash j\colon X_j \qquad (1 \le j \le n)$$

$$\frac{X, \Gamma \vdash t\colon Y}{\Gamma \vdash \lambda t\colon X \to Y}$$

$$\frac{\Gamma \vdash t\colon X \to Y \quad \Gamma \vdash u\colon X}{\Gamma \vdash t\,u\colon Y}$$

A term t is a term of the simply-typed λ-calculus, in DeBruijn notation.

Rewrites.

$$X_1, \ldots, X_n \vdash j\colon j \Rightarrow j\colon X_j \qquad (1 \le j \le n)$$

$$\frac{X, \Gamma \vdash \gamma\colon t \Rightarrow t'\colon Y}{\Gamma \vdash \lambda\gamma\colon \lambda t \Rightarrow \lambda t'\colon X \to Y}$$

$$\frac{\Gamma \vdash \gamma\colon t \Rightarrow t'\colon X \to Y \quad \Gamma \vdash \delta\colon u \Rightarrow u'\colon X}{\Gamma \vdash \gamma\,\delta\colon t\,u \Rightarrow t'\,u'\colon Y}$$

$$\frac{\Gamma \vdash \gamma\colon t \Rightarrow t'\colon X \quad \Gamma \vdash \delta\colon t' \Rightarrow t''\colon X}{\Gamma \vdash \gamma;\delta\colon t \Rightarrow t''\colon X}$$

$$\frac{\Gamma \vdash t\colon X \to Y}{\Gamma \vdash \eta_t\colon t \Rightarrow \lambda(t^1\,1)\colon X \to Y}$$

$$\frac{X, \Gamma \vdash t\colon Y \quad \Gamma \vdash u\colon X}{\Gamma \vdash \beta_{t,u}\colon (\lambda t)\,u \Rightarrow t[u]\colon Y}$$

Rewrites are built up from β-reduction and η-expansion by sequential and parallel composition. By a simple induction, for any term $\Gamma \vdash t\colon X$ there is a rewrite $\Gamma \vdash t\colon t \Rightarrow t\colon X$, which we call an **identity** rewrite. The notations t^1 and $t[u]$ are defined below.

2.2. Substitution. For definiteness, we give our notation for substitution in some detail. The reader who is unfamiliar with DeBruijn notation should read this in some detail, noting how variable capture and other problems are dealt with.

Terms. Firstly, t^n is t with all free variables greater than or equal to n incremented by one:

$$j^n = \begin{cases} j+1 & \text{if } j \geq n, \\ j & \text{otherwise} \end{cases}$$
$$(t\,u)^n = t^n\,u^n$$
$$(\lambda t)^n = \lambda t^{n+1}$$

Next, $t[v_1, v_2, \ldots]$ is t with u_j substituted for j:

$$j[v_1, v_2, \ldots] = v_j$$
$$(t\,u)[v_1, v_2, \ldots] = t[v_1, v_2, \ldots]\,u[v_1, v_2, \ldots]$$
$$(\lambda t)[v_1, v_2, \ldots] = \lambda t[1, v_1^1, v_2^1, \ldots]$$

for brevity we write $t[u]$ for $t[u, 1, 2, \ldots]$.

Lemma 1. Some basic properties of substitution:

(1) If $X_1, \ldots, X_m \vdash t\colon Y$, $1 \leq n \leq m+1$ and X is a type, then $X_1, \ldots, X_{n-1}, X, X_n, \ldots, X_m \vdash t^n\colon Y$.
(2) If $X_1, \ldots, X_n \vdash t\colon Y$ and $\Gamma \vdash u_j\colon X_j$ for $j = 1, \ldots, n$, then $\Gamma \vdash t[u_1, \ldots, u_n]\colon Y$.
(3) $t[1, 2, \ldots] = t$.
(4) $t[u_1, u_2, \ldots][v_1, v_2, \ldots] = t[u_1[v_1, v_2, \ldots], u_2[v_1, v_2, \ldots], \ldots]$.
(5) $t^2[1] = t$.
(6) $(t[u_1, u_2, \ldots])^1 = t^1[1, u_1^1, u_2^1, \ldots]$.

Proof. These results are all straightforward structural inductions. □

Rewrites. The operation of incrementing variables extends to rewrites in a straightforward way:

$$j^n = \begin{cases} j+1 & \text{if } j \geq n, \\ j & \text{otherwise} \end{cases}$$
$$(\gamma\,\delta)^n = \gamma^n\,\delta^n$$
$$(\lambda\gamma)^n = \lambda\gamma^{n+1}$$
$$(\gamma;\delta)^n = \gamma^n;\delta^n$$
$$\eta_t^n = \eta_{t^n}$$
$$\beta_{t,u}^n = \beta_{t^{n+1},u^n}$$

There are two forms of substitution: rewrites into terms:

$$j[\gamma_1, \gamma_2, \ldots] = \gamma_j$$
$$(t\,u)[\gamma_1, \gamma_2, \ldots] = t[\gamma_1, \gamma_2, \ldots]\,u[\gamma_1, \gamma_2, \ldots]$$
$$(\lambda t)[\gamma_1, \gamma_2, \ldots] = \lambda t[1, \gamma_1^1, \gamma_2^1, \ldots]$$

and terms into rewrites:

$$
\begin{aligned}
j[v_1, v_2, \dots] &= v_j \\
(\gamma\,\delta)[v_1, v_2, \dots] &= \gamma[v_1, v_2, \dots]\,\delta[v_1, v_2, \dots] \\
(\lambda\gamma)[v_1, v_2, \dots] &= \lambda\gamma[1, v_1^1, v_2^1, \dots] \\
(\gamma; \delta)[v_1, v_2, \dots] &= \gamma[v_1, v_2, \dots]; \delta[v_1, v_2, \dots] \\
\eta_t[v_1, v_2, \dots] &= \eta_{t[v_1, v_2, \dots]} \\
\beta_{t,u}[v_1, v_2, \dots] &= \beta_{t[1, v_1^1, v_2^1, \dots], u[v_1, v_2, \dots]}
\end{aligned}
$$

Note that there are three interpretations of $t[u_1, u_2, \dots]$ as a rewrite: the identity on $t[u_1, u_2, \dots]$, the substitution of $[u_1, u_2, \dots]$ into the identity on t and substitution of identities on $[u_1, u_2, \dots]$ into t. A simple induction shows that these three are equal, so there is no ambiguity.

Lemma 2. Basic properties of substitution of rewrites:

(1) If $X_1, \dots, X_n \vdash t\colon Y$ and $\Gamma \vdash \gamma_j\colon u_j \Rightarrow u_j'\colon X_j$ for $j = 1, \dots, n$, then $\Gamma \vdash t[\gamma_1, \dots, \gamma_n]\colon t[u_1, \dots, u_n] \Rightarrow t[u_1', \dots, u_n']\colon Y$.
(2) $t[u_1, u_2, \dots][\gamma_1, \gamma_2, \dots] = t[u_1[\gamma_1, \gamma_2, \dots], u_2[\gamma_1, \gamma_2, \dots], \dots]$.
(3) $(t[\gamma_1, \gamma_2, \dots])^1 = t^1[1, \gamma_1^1, \gamma_2^1, \dots]$.

Proof. More straightforward structural inductions. □

Lemma 3. Basic properties of substitution into rewrites:

(1) If $X_1, \dots, X_m \vdash \gamma\colon t \Rightarrow u\colon Y$, $1 \le n \le m+1$ and X is a type, then $X_1, \dots, X_{n-1}, X, X_n, \dots, X_m \vdash \gamma^n\colon t^n \Rightarrow u^n\colon Y$.
(2) If $X_1, \dots, X_n \vdash \gamma\colon t \Rightarrow t'\colon Y$ and $\Gamma \vdash u_j\colon X_j$ for $j = 1, \dots, n$, then $\Gamma \vdash \gamma[u_1, \dots, u_n]\colon t[u_1, \dots, u_n] \Rightarrow t'[u_1, \dots, u_n]\colon Y$.
(3) $\gamma[1, 2, \dots] = \gamma$.
(4) $\gamma[u_1, u_2, \dots][v_1, v_2, \dots] = \gamma[u_1[v_1, v_2, \dots], u_2[v_1, v_2, \dots], \dots]$.
(5) $\gamma^2[1] = \gamma$.
(6) $(\gamma[u_1, u_2, \dots])^1 = \gamma^1[1, u_1^1, u_2^1, \dots]$.

Proof. Again, these are straightforward structural inductions. □

Lemma 4. And one property which links the two:

$$t[\gamma_1, \gamma_2, \dots][v_1, v_2, \dots] = t[\gamma_1[v_1, v_2, \dots], \gamma_2[v_1, v_2, \dots], \dots]$$

Proof. Another straightforward structural induction. □

3. The theory 2-λ

The theory 2-λ is an equational theory on the *rewrites* of the 2-λ-calculus. We write it as a judgement:

- $\Gamma \vdash \gamma = \delta\colon t \Rightarrow u\colon X$ means that γ and δ are equivalent in the theory 2-λ, where γ and δ are well-formed rewrites with source t and target u, of type X in context Γ.

The intention is to axiomatise not when two rewrites have the same effect—after all, we are only considering equations between rewrites with common source and target—but when two rewrites might be implemented identically; for example, a parallel rewrite might be implemented on a sequential machine in either order. This is an attempt to say when two rewrites represent the same algorithm.

3.1. The axiomatisation of 2-λ. The first axioms need no explanation, they merely formalise what might be called a 2-λ-theory: an equivalence which respects the syntactic structure.

(reflexivity)
$$\frac{\Gamma \vdash \gamma \colon t \Rightarrow t' \colon X}{\Gamma \vdash \gamma = \gamma \colon t \Rightarrow t' \colon X}$$

(symmetry)
$$\frac{\Gamma \vdash \gamma = \delta \colon t \Rightarrow t' \colon X}{\Gamma \vdash \delta = \gamma \colon t \Rightarrow t' \colon X}$$

(transitivity)
$$\frac{\Gamma \vdash \gamma = \delta \colon t \Rightarrow t' \colon X \quad \Gamma \vdash \delta = \epsilon \colon t \Rightarrow t' \colon X}{\Gamma \vdash \gamma = \epsilon \colon t \Rightarrow t' \colon X}$$

$$\frac{X, \Gamma \vdash \gamma = \gamma' \colon t \Rightarrow t' \colon Y}{\Gamma \vdash \lambda\gamma = \lambda\gamma' \colon \lambda t \Rightarrow \lambda t' \colon X \to Y}$$

$$\frac{\Gamma \vdash \gamma = \gamma' \colon t \Rightarrow t' \colon X \to Y \quad \Gamma \vdash \delta = \delta' \colon u \Rightarrow u' \colon X}{\Gamma \vdash \gamma\,\delta = \gamma'\,\delta' \colon t\,u \Rightarrow t'\,u' \colon Y}$$

$$\frac{\Gamma \vdash \gamma = \gamma' \colon t \Rightarrow t' \colon X \quad \Gamma \vdash \delta = \delta' \colon t' \Rightarrow t'' \colon X}{\Gamma \vdash \gamma;\delta = \gamma';\delta' \colon t \Rightarrow t'' \colon X}$$

The particular axioms which define the theory 2-λ are as follows:

(1)
$$\frac{\Gamma \vdash \gamma \colon j \Rightarrow t \colon X}{\Gamma \vdash j;\gamma = \gamma \colon j \Rightarrow t \colon X}$$

(2)
$$\frac{\Gamma \vdash \gamma \colon t \Rightarrow j \colon X}{\Gamma \vdash \gamma;j = \gamma \colon t \Rightarrow j \colon X}$$

(rewrites j act as left and right identities of composition)

(3)
$$\frac{X, \Gamma \vdash \gamma \colon t \Rightarrow t' \colon Y \quad X, \Gamma \vdash \delta \colon t' \Rightarrow t'' \colon Y}{\Gamma \vdash \lambda\gamma; \lambda\delta = \lambda(\gamma;\delta) \colon \lambda t \Rightarrow \lambda t'' \colon X \to Y}$$

(4)
$$\frac{\begin{array}{c}\Gamma \vdash \gamma \colon t \Rightarrow t' \colon X \to Y \quad \Gamma \vdash \gamma' \colon t' \Rightarrow t'' \colon X \to Y \\ \Gamma \vdash \delta \colon u \Rightarrow u' \colon X \quad \Gamma \vdash \delta' \colon u' \Rightarrow u'' \colon X\end{array}}{\Gamma \vdash (\gamma\,\delta);(\gamma'\,\delta') = (\gamma;\gamma')\,(\delta;\delta') \colon t\,u \Rightarrow t''\,u'' \colon Y}$$

(abstraction and application distribute over composition)

(5)
$$\frac{\Gamma \vdash \gamma \colon t \Rightarrow t' \colon X \quad \Gamma \vdash \delta \colon t' \Rightarrow t'' \colon X \quad \Gamma \vdash \epsilon \colon t'' \Rightarrow t''' \colon X}{\Gamma \vdash \gamma;(\delta;\epsilon) = (\gamma;\delta);\epsilon \colon t \Rightarrow t''' \colon X}$$

(composition is associative)

$$\text{(6)} \qquad \frac{\Gamma \vdash \gamma\colon t \Rightarrow t'\colon X \to Y}{\Gamma \vdash \eta_t; \lambda(\gamma^1\, 1) = \gamma; \eta_{t'}\colon t \Rightarrow \lambda(t'^1\, 1)\colon X \to Y}$$

$$\text{(7)} \qquad \frac{X, \Gamma \vdash \gamma\colon t \Rightarrow t'\colon Y \quad \Gamma \vdash \delta\colon u \Rightarrow u'\colon X}{\Gamma \vdash (\lambda\gamma)\,\delta; \beta_{t',u'} = \beta_{t,u}; \gamma[u]; t'[\delta]\colon (\lambda t)\, u \Rightarrow t'[u']\colon Y}$$

(η and β commute with rewrites of their subscripts)

$$\text{(8)} \qquad \frac{X, \Gamma \vdash t\colon Y}{\Gamma \vdash \eta_{\lambda t}; \lambda\beta_{t^2,1} = \lambda t\colon \lambda t \Rightarrow \lambda t\colon X \to Y}$$

$$\text{(9)} \qquad \frac{\Gamma \vdash t\colon X \to Y \quad \Gamma \vdash u\colon X}{\Gamma \vdash \eta_t\, u; \beta_{t^1\,1,u} = t\,u\colon t\,u \Rightarrow t\,u\colon Y}$$

(η-expansion followed by β-reduction cancels out).

Lemma 5. Basic properties of the theory, relating the equations to substitution:

(1) If $X_1, \ldots, X_n \vdash \gamma = \delta\colon t \Rightarrow t'\colon Y$ and $\Gamma \vdash u_j\colon X_j$ for $j = 1, \ldots, n$, then $\Gamma \vdash \gamma[u_1, \ldots, u_n] = \delta[u_1, \ldots, u_n]\colon t[u_1, \ldots, u_n] \Rightarrow t'[u_1, \ldots, u_n]\colon Y$.

(2) If $X_1, \ldots, X_n \vdash t\colon Y$ and $\Gamma \vdash \gamma_j = \delta_j\colon u_j \Rightarrow u'_j\colon X_j$ for $j = 1, \ldots, n$, then $\Gamma \vdash t[\gamma_1, \ldots, \gamma_n] = t[\delta_1, \ldots, \delta_n]\colon t[u_1, \ldots, u_n] \Rightarrow t[u'_1, \ldots, u'_n]\colon Y$.

(3) If $\Gamma \vdash \gamma\colon t \Rightarrow u\colon X$ then $\Gamma \vdash t; \gamma = \gamma\colon t \Rightarrow u\colon X$ and $\Gamma \vdash \gamma; u = \gamma\colon t \Rightarrow u\colon X$.

(4) If $X_1, \ldots, X_n \vdash t\colon Y$, $\Gamma \vdash \gamma_j\colon u_j \Rightarrow u'_j\colon X_j$ and $\Gamma \vdash \delta_j\colon u'_j \Rightarrow u''_j\colon X_j$ for $j = 1, \ldots, n$, then $\Gamma \vdash t[\gamma_1; \delta_1, \ldots, \gamma_n; \delta_n] = t[\gamma_1, \ldots, \gamma_n]; t[\delta_1, \ldots, \delta_n]\colon t[u_1, \ldots, u_n] \Rightarrow t[u''_1, \ldots, u''_n]\colon Y$.

(5) If $X_1, \ldots, X_n \vdash \gamma\colon t \Rightarrow t'\colon Y$ and $\Gamma \vdash \delta_j\colon u_j \Rightarrow u'_j\colon X_j$ for $j = 1, \ldots, n$, then $\Gamma \vdash \gamma[u_1, \ldots, u_n]; t'[\delta_1, \ldots, \delta_n] = t[\delta_1, \ldots, \delta_n]; \gamma[u'_1, \ldots, u'_n]\colon t[u_1, \ldots, u_n] \Rightarrow t'[u'_1, \ldots, u'_n]\colon Y$.

Proof. Yet more structural inductions. □

3.2. The categorical description of 2-λ. In this paragraph we present the author's original motivation for the theory 2-λ, which justifies the equations 1–9. It is based on Seely's description of the λ-calculus as a 2-category [7]. This motivation uses some fairly delicate notions from the theory of 2-categories. The reader who is unfamiliar with this material can safely skip the rest of this section, as neither the results nor the methods will be used in the rest of the paper.

The 2-categorical objects which occur here are either *strict* or *lax*, but not *pseudo*. We will therefore stick to the "Australian" terminology, where everything is preserved on the nose unless otherwise qualified, and the word strict is used only for emphasis.

The theory of 2-categories has several notions of adjunction (see, for example, [5]) of which we shall need the following:

Definition. A **2-natural adjunction** consists of the following data:

- two 2-categories C and D,
- two (strict) 2-functors $F, G\colon C \to D$,
- two 2-natural transformations $\sigma\colon F \Rightarrow G$ and $\tau\colon G \Rightarrow F$, and
- two modifications $\eta\colon \mathrm{id}_F \to \tau\sigma$ and $\epsilon\colon \sigma\tau \to \mathrm{id}_G$,

satisfying the triangle laws:

$$\epsilon\sigma \circ \sigma\eta = \mathrm{id}_\sigma$$
$$\tau\epsilon \circ \eta\tau = \mathrm{id}_\tau$$

In this case we say that σ is **naturally left adjoint** to τ.

Definition. Let $F\colon C \to D$ be a 2-functor. A **lax right adjoint** to F assigns to each object Y of D the following:

- an object $G(Y)$ of C and
- two 2-natural transformations $\sigma(Y)\colon C(\_, G(Y)) \Rightarrow D(F(\_), Y)$ and $\tau(Y)\colon D(F(\_), Y) \Rightarrow C(\_, G(Y))$,

such that $\sigma(Y)$ is naturally left adjoint to $\tau(Y)$.

Definition. Let C be a 2-category with finite products (in the enriched sense). We say C has **lax exponentials** if for each object X the 2-functor $X \times \_\colon C \to C$ has a lax right adjoint.

Lemma 6. Let C be a 2-category with finite products, and X, X' be objects of C. If $X \times \_$ and $X' \times \_$ have lax right adjoints, then so does $X \times X' \times \_$.

Proof. Let $X \times \_$ have lax right adjoint G^X, σ^X, τ^X etc. Then

$$G^{X\times X'}(Y) = G^{X'}(G^X(Y))$$
$$\sigma^{X\times X'}(Y)_Z = \sigma^X(Y)_{X'\times Z}\sigma^{X'}(G^X(Y))_Z$$
$$\tau^{X\times X'}(Y)_Z = \tau^{X'}(G^X(Y))_Z\tau^X(Y)_{X'\times Z}$$

defines a lax right adjoint to $X \times X' \times \_$. □

This 2-category theory is related to the theory 2-λ by a 2-categorical version of the Lambek-Lawvere correspondence. We associate a 2-category Λ with 2-λ as follows:

- The objects are contexts Γ.
- The arrows are lists of terms $[t_1, \ldots, t_n]\colon \Gamma \to X_1, \ldots, X_n$, where $\Gamma \vdash t_j\colon X_j$.
- The 2-cells are lists of equivalence classes of rewrites $[\gamma_1, \ldots, \gamma_n]\colon [t_1, \ldots, t_n] \Rightarrow [u_1, \ldots, u_n]\colon \Gamma \to X_1, \ldots, X_n$, where $\Gamma \vdash \gamma_j\colon t_j \Rightarrow u_j\colon X_j$, under the relation
- two rewrites γ and $\delta\colon t \Rightarrow u\colon \Gamma \to X$ are equivalent iff $\Gamma \vdash \gamma = \delta\colon t \Rightarrow u\colon X$.
- Horizontal composition of $[t_1, \ldots, t_n]\colon \Delta \to E$ and $[u_1, \ldots, u_m]\colon \Gamma \to \Delta$ is $[t_1[u_1, \ldots, u_m], \ldots, t_n[u_1, \ldots, u_m]]\colon \Gamma \to E$.
- Vertical composition of $[\gamma_1, \ldots, \gamma_n]\colon [t_1, \ldots, t_n] \Rightarrow [u_1, \ldots, u_n]$ and $[\delta_1, \ldots, \delta_n]\colon [u_1, \ldots, u_n] \Rightarrow [v_1, \ldots, v_n]$ is $[\gamma_1; \delta_1, \ldots, \gamma_n; \delta_n]\colon [t_1, \ldots, t_n] \Rightarrow [v_1, \ldots, v_n]$.

Proposition 7. Λ is a 2-category with finite products and lax exponentials.

Proof. That Λ is a 2-category amounts to checking various axioms, all of which are either immediate or appear in lemmas 1–5.

Products are defined by concatenation of contexts, projections are variables and universal arrows are given by concatenations of lists.

In view of lemma 6, it is enough to give a lax right adjoint to $X \times \_$. This is defined by

$$\begin{aligned} G^X(Y_1,\ldots,Y_m) &= X \to Y_1,\ldots,X \to Y_m \\ \sigma^X(Y_1,\ldots,Y_m)_\Gamma([t_1,\ldots,t_m]) &= [t_1^1\,1,\ldots,t_m^1\,1] \\ \tau^X(Y_1,\ldots,Y_m)_\Gamma([t_1,\ldots,t_m]) &= [\lambda t_1,\ldots,\lambda t_m] \\ \eta^X(Y_1,\ldots,Y_m)_\Gamma([t_1,\ldots,t_m]) &= [\eta_{t_1},\ldots,\eta_{t_m}] \\ \epsilon^X(Y_1,\ldots,Y_m)_\Gamma([t_1,\ldots,t_m]) &= [\beta_{t_1^2,1},\ldots,\beta_{t_m^2,1}] \end{aligned}$$

Again, all the work has been done in the lemmas. □

Theorem 8. Λ is the universal (free) 2-category with finite products and lax exponentials on the set $\mathbb{B}$ of basic types.

Proof. Let $\mathcal{C}$ be a 2-category with finite products and lax exponentials, the lax right adjoint to $X \times \_$ being given by G^X, σ^X, τ^X, η^X and ϵ^X. For each $B \in \mathbb{B}$ let B_C be an object of $\mathcal{C}$. We construct a 2-functor $\mathcal{F}\colon \Lambda \to \mathcal{C}$ which preserves finite products and lax exponentials as follows:

$$\begin{aligned} \mathcal{F}(X_1,\ldots,X_n) &= \mathcal{F}(X_1) \times \cdots \times \mathcal{F}(X_n) \\ \mathcal{F}(X \to Y) &= G^{\mathcal{F}(X)}(\mathcal{F}(Y)) \\ \mathcal{F}(B) &= B_C \end{aligned}$$

$$\begin{aligned} \mathcal{F}[t_1,\ldots,t_n] &= \langle \mathcal{F}(t_1),\ldots,\mathcal{F}(t_n)\rangle \\ \mathcal{F}(j) &= \pi_j \\ \mathcal{F}(\lambda t) &= \tau^{\mathcal{F}(X)}(\mathcal{F}(Y))_{\mathcal{F}(\Gamma)}(\mathcal{F}(t)) \quad \text{where } X,\Gamma \vdash t\colon Y \\ \mathcal{F}(t\,u) &= \sigma^{\mathcal{F}(X)}(\mathcal{F}(Y))_{\mathcal{F}(\Gamma)}(\mathcal{F}(t)) \circ \langle \mathcal{F}(u), \mathrm{id}\rangle \quad \text{where } \Gamma \vdash t\colon X \to Y \end{aligned}$$

$$\begin{aligned} \mathcal{F}[\gamma_1,\ldots,\gamma_n] &= \langle \mathcal{F}(\gamma_1),\ldots,\mathcal{F}(\gamma_n)\rangle \\ \mathcal{F}(j) &= 1_{\pi_j} \\ \mathcal{F}(\lambda\gamma) &= \tau^{\mathcal{F}(X)}(\mathcal{F}(Y))_{\mathcal{F}(\Gamma)}(\mathcal{F}(\gamma)) \quad \text{where } X,\Gamma \vdash \gamma\colon t \Rightarrow t'\colon Y \\ \mathcal{F}(\gamma\,\delta) &= \sigma^{\mathcal{F}(X)}(\mathcal{F}(Y))_{\mathcal{F}(\Gamma)}(\mathcal{F}(\gamma)) \circ \langle \mathcal{F}(\delta), 1\rangle \quad \text{where } \Gamma \vdash \gamma\colon t \Rightarrow t'\colon X \to Y \\ \mathcal{F}(\gamma;\delta) &= \mathcal{F}(\gamma);\mathcal{F}(\delta) \\ \mathcal{F}(\eta_t) &= \eta^{\mathcal{F}(X)}(\mathcal{F}(Y))_{\mathcal{F}(\Gamma)}(\mathcal{F}(t)) \quad \text{where } \Gamma \vdash t\colon X \to Y \\ \mathcal{F}(\beta_{t,u}) &= \epsilon^{\mathcal{F}(X)}(\mathcal{F}(Y))_{\mathcal{F}(\Gamma)}(\mathcal{F}(t)) \circ \langle \mathcal{F}(u), 1\rangle \quad \text{where } \Gamma \vdash t\colon X \to Y \end{aligned}$$

It is straightforward to check that this is well-defined, and clear that it is unique. □

4. The Canonical Form of Rewrites

In this section we solve the word problem for 2-λ, by finding a set $\mathcal{G}$ of rewrites with the property that 2-λ equates every rewrite with a unique element of $\mathcal{G}$.

Let $\mathcal{A}$, $\mathcal{E}$ and $\mathcal{G}$ be the smallest sets of rewrites closed under the following:

- Every identity rewrite is in $\mathcal{A}$.

- If $\Gamma \vdash \alpha_1 \colon t_1 \Rightarrow \lambda t \colon X \to Y$ and $\Gamma \vdash \alpha_2 \colon t[u] \Rightarrow t_2 \colon Y$ are in $\mathcal{A}$, then $\Gamma \vdash \alpha_1\, u; \beta_{t,u}; \alpha_2 \colon t_1\, u \to t_2 \colon Y$ is in $\mathcal{A}$.
- Every identity rewrite is in $\mathcal{E}$.
- If $\Gamma \vdash \epsilon_1 \colon t_1 \Rightarrow t \colon X \to Y$, $X, \Gamma \vdash \epsilon_2 \colon 1 \Rightarrow t_2 \colon X$ and $X, \Gamma \vdash \epsilon_3 \colon t^1\, t_2 \Rightarrow t_3 \colon Y$ are in $\mathcal{E}$, then $\Gamma \vdash \epsilon_1; \eta_t; \lambda(t^1\, \epsilon_2; \epsilon_3) \colon t_1 \Rightarrow \lambda t_3 \colon X \to Y$ is in $\mathcal{E}$.
- If $\Gamma \vdash \alpha \colon t \Rightarrow j \colon X$ is in $\mathcal{A}$ and $\Gamma \vdash \epsilon \colon j \Rightarrow u \colon X$ is in $\mathcal{E}$ then $\Gamma \vdash \alpha; j; \epsilon \colon t \Rightarrow u \colon X$ is in $\mathcal{G}$.
- If $\Gamma \vdash \alpha \colon t \Rightarrow \lambda t_1 \colon X \to Y$ is in $\mathcal{A}$, $\Gamma \vdash \epsilon \colon \lambda u_1 \Rightarrow u \colon X \to Y$ is in $\mathcal{E}$ and $X, \Gamma \vdash \gamma \colon t_1 \Rightarrow u_1 \colon Y$ is in $\mathcal{G}$ then $\Gamma \vdash \alpha; \lambda\gamma; \epsilon \colon t \Rightarrow u \colon X \to Y$ is in $\mathcal{G}$.
- If $\Gamma \vdash \alpha \colon t \Rightarrow t_1\, u_1 \colon Y$ is in $\mathcal{A}$, $\Gamma \vdash \epsilon \colon t_2\, u_2 \Rightarrow u \colon Y$ is in $\mathcal{E}$, and $\Gamma \vdash \gamma \colon t_1 \Rightarrow t_2 \colon X \to Y$ and $\Gamma \vdash \delta \colon u_1 \Rightarrow u_2 \colon X$ are in $\mathcal{G}$, then $\Gamma \vdash \alpha; \gamma\, \delta; \epsilon \colon t \Rightarrow u \colon Y$ is in $\mathcal{G}$.

The notation $\gamma_1; \gamma_2; \gamma_3$ is shorthand for $(\gamma_1; \gamma_2); \gamma_3$. (Of course, the choice of left rather than right bracketing is arbitrary, as long as we are consistent.)

Lemma 9. The following results are no more than observations; they are recorded here so that we can use them without further comment.

(1) Every rewrite in $\mathcal{G}$ is of the form $\alpha; \delta; \epsilon$, where $\alpha \in \mathcal{A}$, $\epsilon \in \mathcal{E}$ and δ is either j, $\lambda\gamma_1$ or $\gamma_1\, \gamma_2$.
(2) If $\Gamma \vdash \alpha \colon t \Rightarrow t' \colon X$ in $\mathcal{A}$ is not an identity rewrite, then $t = t_1\, t_2$ for some t_1, t_2.
(3) If $\Gamma \vdash \epsilon \colon t \Rightarrow t' \colon X$ in $\mathcal{E}$ is not an identity rewrite, then $t' = \lambda t_1$ for some t_1.
(4) If $\Gamma \vdash \gamma \colon 1 \Rightarrow t \colon X$ is in $\mathcal{G}$ then $\gamma = 1; 1; \epsilon$ for some $\Gamma \vdash \epsilon \colon 1 \Rightarrow t \colon X$ in $\mathcal{E}$.

In general, identity rewrites are not members of $\mathcal{G}$. However, for each term t we can define $\mathcal{I}(t)$ in $\mathcal{G}$ as follows:

$$\begin{aligned}\mathcal{I}(j) &= j; j; j\\ \mathcal{I}(\lambda t) &= \lambda t; \lambda\mathcal{I}(t); \lambda t\\ \mathcal{I}(t\, u) &= t\, u; \mathcal{I}(t)\, \mathcal{I}(u); t\, u\end{aligned}$$

Lemma 10. If $\Gamma \vdash t \colon X$ then

(1) $\Gamma \vdash \mathcal{I}(t) \colon t \Rightarrow t \colon X$ is in $\mathcal{G}$
(2) $\Gamma \vdash \mathcal{I}(t) = t \colon t \Rightarrow t \colon X$.

Proof. Structural induction. □

Substitution of rewrites in $\mathcal{G}$ is defined as follows. If $\Gamma \vdash \delta_j \colon u_j \Rightarrow u'_j \colon X_j$ for $j = 1 \ldots n$, then:

$$\begin{aligned}(\alpha; j; \epsilon)[\delta_1, \ldots, \delta_n] &= \alpha[u_1, \ldots, u_n]; ; \delta_j; ; \epsilon[u'_1, \ldots, u'_n]\\ (\alpha; \lambda\gamma; \epsilon)[\delta_1, \ldots, \delta_n] &= \alpha[u_1, \ldots, u_n]; \lambda(\gamma[1, \delta_1^1, \ldots, \delta_n^1]); \epsilon[u'_1, \ldots, u'_n]\\ (\alpha; \gamma_1\, \gamma_2; \epsilon)[\delta_1, \ldots, \delta_n] &= \alpha[u_1, \ldots, u_n]; \gamma_1[\delta_1, \ldots, \delta_n]\, \gamma_2[\delta_1, \ldots, \delta_n]; \epsilon[u'_1, \ldots, u'_n]\end{aligned}$$

Lemma 11. Let $X_1, \ldots, X_n \vdash \gamma \colon t \Rightarrow t' \colon X$ and $\Gamma \vdash \delta_j \colon u_j \Rightarrow u'_j \colon X_j$ for $j = 1 \ldots n$ be rewrites in $\mathcal{G}$. Then

(1) $\Gamma \vdash \gamma[\delta_1, \ldots, \delta_n] \colon t[u_1, \ldots, u_n] \Rightarrow t'[u'_1, \ldots, u'_n] \colon X$ is in $\mathcal{G}$
(2) $\Gamma \vdash \gamma[\delta_1, \ldots, \delta_n] = \gamma[u_1, \ldots, u_n]; t'[\delta_1, \ldots, \delta_n] \colon t[u_1, \ldots, u_n] \Rightarrow t'[u'_1, \ldots, u'_n] \colon X$.

Proof. Structural induction. □

The heart of the proof of the canonical form theorem is the definition of sequential composition of rewrites in $\mathcal{G}$. The composition of rewrites in $\mathcal{E}$ and $\mathcal{A}$ is straightforward; we use the symbol '$;;$' defined as follows, with the convention that $\alpha \in \mathcal{A}$, $\epsilon \in \mathcal{E}$ and $\alpha;\delta;\epsilon \in \mathcal{G}$.

$$t;;\alpha = \alpha \quad \text{where } t \text{ is an identity}$$
$$(\alpha_1 u;\beta_{t,u};\alpha_2);;\alpha = \alpha_1 u;\beta_{t,u};(\alpha_2;;\alpha)$$

$$\epsilon;;t = \epsilon \quad \text{where } t \text{ is an identity}$$
$$\epsilon;;(\epsilon_1;\eta_t;\lambda(t^1\,\epsilon_2;\epsilon_3)) = (\epsilon;;\epsilon_1);\eta_t;\lambda(t^1\,\epsilon_2;\epsilon_3)$$

$$\alpha;;(\alpha';\delta;\epsilon) = (\alpha;;\alpha');\delta;\epsilon$$
$$(\alpha;\delta;\epsilon');;\epsilon = \alpha;\delta;(\epsilon';;\epsilon)$$

Note that $;;$ is associative (in every possible way) and that identities in $\mathcal{A}$ and $\mathcal{E}$ are identities of $;;$.

For the sequential composition of two (composable) rewrites in $\mathcal{G}$ we use the symbol '$\dagger$' defined as follows:

$$\alpha;j;j \dagger j;j;\epsilon = \alpha;j;\epsilon$$
$$\alpha;\lambda\gamma_1;\lambda t \dagger \lambda t;\lambda\gamma_2;\epsilon = \alpha;\lambda(\gamma_1 \dagger \gamma_2);\epsilon$$
$$\alpha;\gamma_{11}\,\gamma_{12};t\,u \dagger t\,u;\gamma_{21}\,\gamma_{22};\epsilon = \alpha;(\gamma_{11} \dagger \gamma_{21})\,(\gamma_{12} \dagger \gamma_{22});\epsilon$$

$$\gamma_1;\eta_{t_1};\lambda(t_1^1\,\epsilon_{11};\epsilon_{12}) \dagger \lambda t_2;\lambda\gamma_2;\epsilon_2 =$$
$$\begin{cases} \gamma_1 \dagger \gamma_3;\eta_{t_3};\lambda(t_3^1\,\epsilon_{31};\epsilon_{32});;\epsilon_2 \\ \qquad \text{if } t_1^1\,1;\mathcal{I}(t_1^1)\,(1;1;\epsilon_{11});\epsilon_{12} \dagger \gamma_2 = t_1^1\,1;\gamma_3^1\,(1;1;\epsilon_{31});\epsilon_{32} \\ \gamma_1 \dagger \alpha_3;\lambda\gamma_3;\epsilon_2 \\ \qquad \text{if } t_1^1\,1;\mathcal{I}(t_1^1)\,(1;1;\epsilon_{11});\epsilon_{12} \dagger \gamma_2 = \alpha_3^1\,1;\beta_{t_3,1};\gamma_3 \end{cases}$$

$$\alpha_1;\gamma_{11}\,\gamma_{12};t_{21}\,t_{22} \dagger \alpha_{21}\,t_{22};\beta_{t_{23},t_{22}};\gamma_2 =$$
$$\begin{cases} \alpha_1;;\alpha_3\,t_{12};\beta_{t_3,t_{12}};\gamma_3[\gamma_{12}] \dagger \gamma_2 \\ \qquad \text{if } \gamma_{11} \dagger \alpha_{21};\lambda\mathcal{I}(t_{23});\lambda t_{23} = \alpha_3;\lambda\gamma_3;\lambda t_{23} \\ \alpha_1;\gamma_3\,(\gamma_{12};;\epsilon_{31}[t_{22}]);\epsilon_{32}[t_{22}] \dagger \gamma_2 \\ \qquad \text{if } \gamma_{11} \dagger \alpha_{21};\lambda\mathcal{I}(t_{23});\lambda t_{23} = \gamma_3;\eta_{t_3};\lambda(t_3^1\,\epsilon_{31};\epsilon_{32}) \end{cases}$$

Lemma 12. If $\Gamma \vdash \gamma_1\colon t_1 \Rightarrow t_2\colon X$ and $\Gamma \vdash \gamma_2\colon t_2 \Rightarrow t_3\colon X$ are in $\mathcal{G}$, then

(1) $\gamma_1 \dagger \gamma_2$ is well-defined and $\Gamma \vdash \gamma_1 \dagger \gamma_2\colon t_1 \Rightarrow t_3\colon X$ is in $\mathcal{G}$,
(2) $\Gamma \vdash \gamma_1 \dagger \gamma_2 = \gamma_1;\gamma_2\colon t_1 \Rightarrow t_3\colon X$.

Proof. (1) That the clauses defining $\dagger$ are exhaustive follows from lemma 9. The well-foundedness of the recursion is slightly more complicated than the simple structural inductions considered so far; we define a measure $|\gamma|_{\mathcal{G}}$ on $\mathcal{G}$ and $|\alpha|_{\mathcal{A}}$ on $\mathcal{A}$ as follows:

$$\begin{aligned} |\alpha; j; \epsilon|_{\mathcal{G}} &= |\alpha|_{\mathcal{A}} + 1 \\ |\alpha; \lambda\gamma; \epsilon|_{\mathcal{G}} &= |\alpha|_{\mathcal{A}} + |\gamma|_{\mathcal{G}} + 1 \\ |\alpha; \gamma\,\delta; \epsilon|_{\mathcal{G}} &= |\alpha|_{\mathcal{A}} + |\gamma|_{\mathcal{G}} + |\delta|_{\mathcal{G}} + 1 \end{aligned}$$

$$\begin{aligned} |t|_{\mathcal{A}} &= 0 \\ |\alpha_1\, u; \beta_{t,u}; \alpha_2|_{\mathcal{A}} &= |\alpha_1|_{\mathcal{A}} + |\mathcal{I}(t)|_{\mathcal{G}} + |\alpha_2|_{\mathcal{A}} + 1 \end{aligned}$$

and use the inductive hypothesis on n that:

- $\gamma_1 \dagger \gamma_2$ is well defined for all composable $\gamma_1, \gamma_2 \in \mathcal{G}$ such that $|\gamma_2|_{\mathcal{G}} < n$, and
- if $|\gamma_2|_{\mathcal{G}} < n$ then $|(t\,1; \mathcal{I}(t)\,(1;1;\epsilon_1); \epsilon_2) \dagger \gamma_2|_{\mathcal{G}} < n$ for all $\epsilon_1, \epsilon_2 \in \mathcal{E}$ which make the composition defined.

The proof is then straightforward.

(2) This is a fairly straightforward induction, which amounts to justifying the clauses in the definition of $\dagger$ using the rules 1–9. □

Lemma 13. Basic facts relating the various operations on $\mathcal{G}$.

(1) If $\Gamma \vdash \alpha\colon t_1 \Rightarrow t_2\colon X$ is in $\mathcal{A}$, and $\Gamma \vdash \gamma\colon t_2 \Rightarrow t_3\colon X$ and $\Gamma \vdash \delta\colon t_3 \Rightarrow t_4\colon X$ are in $\mathcal{G}$, then $\alpha;;(\gamma_1 \dagger \gamma_2) = (\alpha;;\gamma_1) \dagger \gamma_2$.

(2) If $\Gamma \vdash \gamma\colon t_1 \Rightarrow t_2\colon X$ and $\Gamma \vdash \delta\colon t_2 \Rightarrow t_3\colon X$ are in $\mathcal{G}$, and $\Gamma \vdash \epsilon\colon t_3 \Rightarrow t_4\colon X$ is in $\mathcal{E}$, then $(\gamma_1 \dagger \gamma_2);;\epsilon = \gamma_1 \dagger (\gamma_2;;\epsilon)$.

(3) If $\Gamma \vdash \gamma\colon t \Rightarrow u\colon X$ is in $\mathcal{G}$ then $\mathcal{I}(t) \dagger \gamma = \gamma = \gamma \dagger \mathcal{I}(u)$.

(4) If $X_1, \ldots, X_n \vdash \gamma\colon t \Rightarrow t'\colon X$, $X_1, \ldots, X_n \vdash \gamma'\colon t' \Rightarrow t''\colon X$, $\Gamma \vdash \delta_j\colon u_j \Rightarrow u'_j\colon X_j$ and $\Gamma \vdash \delta'_j\colon u'_j \Rightarrow u''_j\colon X_j$ are in $\mathcal{G}$ for $j = 1 \ldots n$, then $(\gamma \dagger \gamma')[\delta_1 \dagger \delta'_1, \ldots, \delta_n \dagger \delta'_n] = \gamma[\delta_1, \ldots, \delta_n] \dagger \gamma'[\delta'_1, \ldots, \delta'_n]$

Proof. Straightforward inductions, using the complexity measure $|\gamma|_{\mathcal{G}}$. □

The final result we need is the associativity of $\dagger$:

Proposition 14. If $\Gamma \vdash \gamma_1\colon t_1 \Rightarrow t_2\colon X$, $\Gamma \vdash \gamma_2\colon t_2 \Rightarrow t_3\colon X$ and $\Gamma \vdash \gamma_3\colon t_3 \Rightarrow t_4\colon X$ are in $\mathcal{G}$, then $(\gamma_1 \dagger \gamma_2) \dagger \gamma_3 = \gamma_1 \dagger (\gamma_2 \dagger \gamma_3)$.

Proof. By induction on $|\gamma_3|_{\mathcal{G}}$. There are a total of eight well-formed cases of $\gamma_1, \gamma_2, \gamma_3$, with up to four subcases each. Fortunately, six of the main cases are straightforward, and can be left to the reader. The two remaining cases are as follows:

case 1: $\gamma_1 = \alpha_1; \gamma_{11}\,\gamma_{12}; t_{21}\,t_{22}$, $\gamma_2 = t_{21}\,t_{22}; \gamma_{21}\,\gamma_{22}; t_{31}\,t_{32}$ and $\gamma_3 = \alpha_3\,t_{32}; \beta_{t_{33},t_{32}}; \gamma_{31}$. There are several subcases, corresponding to the different cases in the definition of $\dagger$:

case 1.1: $\gamma_{21} \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} = \alpha_4; \lambda\gamma_4; \lambda t_{33}$ and $\gamma_{11} \dagger \alpha_4; \lambda\mathcal{I}(t_{23}); \lambda t_{23} = \alpha_5; \lambda\gamma_5; \lambda t_{23}$.

$$\begin{aligned} \text{Then}\quad (\gamma_{11} \dagger \gamma_{21}) \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} &= \gamma_{11} \dagger \alpha_4; \lambda\gamma_4; \lambda t_{33} \\ &= \alpha_5; \lambda\gamma_5; \lambda t_{23} \dagger \lambda t_{23}; \lambda\gamma_4; \lambda t_{33} \\ &= \alpha_5; \lambda(\gamma_5 \dagger \gamma_4); \lambda t_{33} \end{aligned}$$

$$\begin{aligned} \text{so}\quad (\gamma_1 \dagger \gamma_2) \dagger \gamma_3 &= \alpha_1;;\alpha_5\,t_{12}; \beta_{t_{13},t_{12}}; (\gamma_5 \dagger \gamma_4)[\gamma_{12} \dagger \gamma_{22}] \dagger \gamma_{31} \\ &= (\alpha_1;;\alpha_5\,t_{12}; \beta_{t_{13},t_{12}}; \gamma_5[\gamma_{12}] \dagger \gamma_4[\gamma_{22}]) \dagger \gamma_{31} \\ &= \gamma_1 \dagger (\gamma_2 \dagger \gamma_3) \end{aligned}$$

case 1.2: $\gamma_{21} \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} = \alpha_4; \lambda\gamma_4; \lambda t_{33}$, $\gamma_{11} \dagger \alpha_4; \lambda\mathcal{I}(t_{23}); \lambda t_{23} = \gamma_5; \eta_{t_5}; \lambda(t_5^1\, \epsilon_{51}; \epsilon_{52})$ and $t_5^1\, 1; \mathcal{I}(t_5^1)\,(1; 1; \epsilon_{51}); \epsilon_{52} \dagger \gamma_4 = t_5^1\, 1; \gamma_6^1\,(1; 1; \epsilon_{61}); \epsilon_{62}$.

$$\begin{aligned}\text{Then } (\gamma_{11} \dagger \gamma_{21}) \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} &= \gamma_{11} \dagger \alpha_4; \lambda\gamma_4; \lambda t_{33} \\ &= \gamma_5; \eta_{t_5}; \lambda(t_5^1\, \epsilon_{51}; \epsilon_{52}) \dagger \lambda t_{23}; \lambda\gamma_4; \lambda t_{33} \\ &= \gamma_5 \dagger \gamma_6; \eta_{t_6}; \lambda(t_6^1\, \epsilon_{61}; \epsilon_{62})\end{aligned}$$

$$\begin{aligned}\text{so } (\gamma_1 \dagger \gamma_2) \dagger \gamma_3 &= \alpha_1; (\gamma_5 \dagger \gamma_6)\,(\gamma_{12} \dagger \gamma_{22};; \epsilon_{61}[t_{32}]); \epsilon_{62}[t_{32}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\, \gamma_{12}; t_5\, t_{22} \dagger (t_5^1\, 1; \gamma_6^1\,(1; 1; \epsilon_{61}); \epsilon_{62})[\gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\, \gamma_{12}; t_5\, t_{22} \dagger (t_5^1\, 1; \mathcal{I}(t_5^1)\,(1; 1; \epsilon_{51}); \epsilon_{52} \dagger \gamma_4)[\gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\,(\gamma_{12};; \epsilon_{51}[t_{22}]); \epsilon_{52}[t_{22}] \dagger \gamma_4[\gamma_{22}] \dagger \gamma_{31} \\ &= \gamma_1 \dagger (\gamma_2 \dagger \gamma_3)\end{aligned}$$

case 1.3: $\gamma_{21} \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} = \alpha_4; \lambda\gamma_4; \lambda t_{33}$, $\gamma_{11} \dagger \alpha_4; \lambda\mathcal{I}(t_{23}); \lambda t_{23} = \gamma_5; \eta_{t_5}; \lambda(t_5^1\, \epsilon_{51}; \epsilon_{52})$, $t_5^1\, 1; \mathcal{I}(t_5^1)\,(1; 1; \epsilon_{51}); \epsilon_{52} \dagger \gamma_4 = \alpha_6^1\, 1; \beta_{t_6,1}; \gamma_6$ and $\gamma_5 \dagger \alpha_6; \lambda\gamma_6; \lambda t_{33} = \alpha_7; \lambda\gamma_7; \lambda t_{33}$.

$$\begin{aligned}\text{Then } (\gamma_{11} \dagger \gamma_{21}) \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} &= \gamma_{11} \dagger \alpha_4; \lambda\gamma_4; \lambda t_{33} \\ &= \gamma_5; \eta_{t_5}; \lambda(t_5^1\, \epsilon_{51}; \epsilon_{52}) \dagger \lambda t_{23}; \lambda\gamma_4; \lambda t_{33} \\ &= \gamma_5 \dagger \alpha_6; \lambda\gamma_6; \lambda t_{33} \\ &= \alpha_7; \lambda\gamma_7; \lambda t_{33}\end{aligned}$$

$$\begin{aligned}\text{so } (\gamma_1 \dagger \gamma_2) \dagger \gamma_3 &= \alpha_1;; \alpha_7\, t_{12}; \beta_{t_{13},t_{12}}; \gamma_7[\gamma_{12} \dagger \gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; (\alpha_7; \lambda\gamma_7; \lambda t_{33})\,(\gamma_{12} \dagger \gamma_{22}); \lambda t_{33}\, t_{32} \dagger (\lambda t_{13}\, t_{12}; \beta_{t_{13},t_{12}}; \mathcal{I}(t_{13}[t_{12}])) \dagger \gamma_{31} \\ &= \alpha_1; (\gamma_5 \dagger \alpha_6; \lambda\gamma_6; \lambda t_{33})\,(\gamma_{12} \dagger \gamma_{22}); \lambda t_{33}\, t_{32} \dagger (\lambda t_{13}\, t_{12}; \beta_{t_{13},t_{12}}; \mathcal{I}(t_{13}[t_{12}])) \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\, \gamma_{12}; t_5\, t_{22} \dagger \alpha_6\, t_{22}; \beta_{t_6,t_{22}}; \gamma_6[\gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\, \gamma_{12}; t_5\, t_{22} \dagger t_5\, t_{22}; \mathcal{I}(t_5)\,(\mathcal{I}(t_{22});; \epsilon_{51}[t_{22}]); \epsilon_{52}[t_{22}] \dagger \gamma_4[\gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\,(\gamma_{12};; \epsilon_{51}[t_{22}]); \epsilon_{52}[t_{22}] \dagger \gamma_4[\gamma_{22}] \dagger \gamma_{31} \\ &= \gamma_1 \dagger (\gamma_2 \dagger \gamma_3)\end{aligned}$$

case 1.4: $\gamma_{21} \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} = \alpha_4; \lambda\gamma_4; \lambda t_{33}$, $\gamma_{11} \dagger \alpha_4; \lambda\mathcal{I}(t_{23}); \lambda t_{23} = \gamma_5; \eta_{t_5}; \lambda(t_5^1\, \epsilon_{51}; \epsilon_{52})$, $t_5^1\, 1; \mathcal{I}(t_5^1)\,(1; 1; \epsilon_{51}); \epsilon_{52} \dagger \gamma_4 = \alpha_6^1\, 1; \beta_{t_6,1}; \gamma_6$ and $\gamma_5 \dagger \alpha_6; \lambda\gamma_6; \lambda t_{33} = \gamma_7; \eta_{t_7}; \lambda(t_7^1\, \epsilon_{71}; \epsilon_{72})$.

$$\begin{aligned}\text{Then } (\gamma_{11} \dagger \gamma_{21}) \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} &= \gamma_{11} \dagger \alpha_4; \lambda\gamma_4; \lambda t_{33} \\ &= \gamma_5; \eta_{t_5}; \lambda(t_5^1\, \epsilon_{51}; \epsilon_{52}) \dagger \lambda t_{23}; \lambda\gamma_4; \lambda t_{33} \\ &= \gamma_5 \dagger \alpha_6; \lambda\gamma_6; \lambda t_{33} \\ &= \gamma_7; \eta_{t_7}; \lambda(t_7^1\, \epsilon_{71}; \epsilon_{72})\end{aligned}$$

$$\begin{aligned}\text{so}\ \ (\gamma_1 \dagger \gamma_2) \dagger \gamma_3 &= \alpha_1; \gamma_7\,(\gamma_{12} \dagger \gamma_{22};\,; \epsilon_{71}[t_{32}]); \epsilon_{72}[t_{32}] \dagger \gamma_{31} \\ &= \alpha_1; (\gamma_7; \eta_{t_7}; \lambda(t_7^1\,\epsilon_{71}; \epsilon_{72})\,(\gamma_{12} \dagger \gamma_{22}); \lambda t_{33}\,t_{32} \dagger (\lambda t_{13}\,t_{12}; \beta_{t_{13},t_{12}}; \mathcal{I}(t_{13}[t_{12}])) \dagger \gamma_{31} \\ &= \alpha_1; (\gamma_5 \dagger \alpha_6; \lambda\gamma_6; \lambda t_{33})\,(\gamma_{12} \dagger \gamma_{22}); \lambda t_{33}\,t_{32} \dagger (\lambda t_{13}\,t_{12}; \beta_{t_{13},t_{12}}; \mathcal{I}(t_{13}[t_{12}])) \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\,\gamma_{12}; t_5\,t_{22} \dagger \alpha_6\,t_{22}; \beta_{t_6,t_{22}}; \gamma_6[\gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\,\gamma_{12}; t_5\,t_{22} \dagger t_5\,t_{22}; \mathcal{I}(t_5)\,(\mathcal{I}(t_{22});\,; \epsilon_{51}[t_{22}]); \epsilon_{52}[t_{22}] \dagger \gamma_4[\gamma_{22}] \dagger \gamma_{31} \\ &= \alpha_1; \gamma_5\,(\gamma_{12};\,; \epsilon_{51}[t_{22}]); \epsilon_{52}[t_{22}] \dagger \gamma_4[\gamma_{22}] \dagger \gamma_{31} \\ &= \gamma_1 \dagger (\gamma_2 \dagger \gamma_3)\end{aligned}$$

case 1.5: $\gamma_{21} \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} = \gamma_4; \eta_{t_4}; \lambda(t_4^1\,\epsilon_{41}; \epsilon_{42})$.
Then $(\gamma_{11} \dagger \gamma_{21}) \dagger \alpha_3; \lambda\mathcal{I}(t_{33}); \lambda t_{33} = \gamma_{11} \dagger \gamma_4; \eta_{t_4}; \lambda(t_4^1\,\epsilon_{41}; \epsilon_{42})$

$$\begin{aligned}\text{so}\ \ (\gamma_1 \dagger \gamma_2) \dagger \gamma_3 &= \alpha_1; (\gamma_{11} \dagger \gamma_4)\,(\gamma_{12} \dagger \gamma_{22};\,; \epsilon_{41}[t_{32}]); \epsilon_{42}[t_{32}] \dagger \gamma_{31} \\ &= \gamma_1 \dagger (\gamma_2 \dagger \gamma_3)\end{aligned}$$

case 2: $\gamma_1 = \gamma_{11}; \eta_{t_{11}}; \lambda(t_{11}^1\,\epsilon_{11}; \epsilon_{12})$, $\gamma_2 = \lambda t_{21}; \lambda\gamma_{21}; \lambda t_{31}$ and $\gamma_3 = \lambda t_{31}; \lambda\gamma_{31}; \epsilon_3$. Again there are several subcases:

case 2.1: $t_{11}^1\,1; \mathcal{I}(t_{11}^1)\,(1; 1; \epsilon_{11}); \epsilon_{12} \dagger \gamma_{21} = t_{11}^1\,1; \gamma_4\,(1; 1; \epsilon_{41}); \epsilon_{42}$ and $t_4^1\,1; \mathcal{I}(t_4^1)\,(1; 1; \epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} = t_4^1\,1; \gamma_5\,(1; 1; \epsilon_{51}); \epsilon_{52}$.

$$\begin{aligned}\text{Then}\ \ t_{11}^1\,1; \mathcal{I}(t_{11}^1)\,(1; 1; \epsilon_{11}); \epsilon_{12} \dagger (\gamma_{21} \dagger \gamma_{31}) &= t_{11}^1\,1; \gamma_4\,(1; 1; \epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} \\ &= t_{11}^1\,1; \gamma_4\,\mathcal{I}(1); t_4^1\,1 \dagger t_4^1\,1; \gamma_5\,(1; 1; \epsilon_{51}); \epsilon_{52} \\ &= t_{11}^1\,1; (\gamma_4 \dagger \gamma_5)\,(1; 1; \epsilon_{51}); \epsilon_{52}\end{aligned}$$

$$\begin{aligned}\text{so}\ \ \gamma_1 \dagger (\gamma_2 \dagger \gamma_3) &= \gamma_{11} \dagger (\gamma_4 \dagger \gamma_5); \eta_{t_5}; \lambda(t_5^1\,\epsilon_{51}; \epsilon_{52}) \\ &= (\gamma_1 \dagger \gamma_2) \dagger \gamma_3\end{aligned}$$

case 2.2: $t_{11}^1\,1; \mathcal{I}(t_{11}^1)\,(1; 1; \epsilon_{11}); \epsilon_{12} \dagger \gamma_{21} = t_{11}^1\,1; \gamma_4\,(1; 1; \epsilon_{41}); \epsilon_{42}$, $t_4^1\,1; \mathcal{I}(t_4^1)\,(1; 1; \epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} = \alpha_5\,1; \beta_{t_5,1}; \gamma_5$ and $\gamma_4 \dagger \alpha_5; \lambda\mathcal{I}(t_5); \lambda t_5 = \alpha_6; \lambda\gamma_6; \lambda t_5$.

$$\begin{aligned}\text{Then}\ \ t_{11}^1\,1; \mathcal{I}(t_{11}^1)\,(1; 1; \epsilon_{11}); \epsilon_{12} \dagger (\gamma_{21} \dagger \gamma_{31}) &= t_{11}^1\,1; \gamma_4\,(1; 1; \epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} \\ &= t_{11}^1\,1; \gamma_4\,\mathcal{I}(1); t_4^1\,1 \dagger \alpha_5\,1; \beta_{t_5,1}; \gamma_5 \\ &= \alpha_6\,1; \beta_{t_6,1}; \gamma_6 \dagger \gamma_5\end{aligned}$$

$$\begin{aligned}\text{so}\ \ \gamma_1 \dagger (\gamma_2 \dagger \gamma_3) &= \gamma_{11} \dagger \alpha_6; \lambda(\gamma_6 \dagger \gamma_5); \epsilon_3 \\ &= \gamma_{11} \dagger \gamma_4 \dagger \alpha_5; \lambda\gamma_5; \epsilon_3 \\ &= (\gamma_1 \dagger \gamma_2) \dagger \gamma_3\end{aligned}$$

case 2.3: $t_{11}^1\,1; \mathcal{I}(t_{11}^1)\,(1; 1; \epsilon_{11}); \epsilon_{12} \dagger \gamma_{21} = t_{11}^1\,1; \gamma_4\,(1; 1; \epsilon_{41}); \epsilon_{42}$, $t_4^1\,1; \mathcal{I}(t_4^1)\,(1; 1; \epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} = \alpha_5\,1; \beta_{t_5,1}; \gamma_5$ $\gamma_4 \dagger \alpha_5; \lambda\mathcal{I}(t_5); \lambda t_5 = \gamma_6; \eta_{t_6}; \lambda(t_6^1\,\epsilon_{61}; \epsilon_{62})$ and $t_{11}^1; \gamma_6\,(1; 1; \epsilon_{61}); \epsilon_{62} \dagger \gamma_5 = t_{11}^1\,1; \gamma_7\,(1; 1; \epsilon_{71}); \epsilon_{72}$.

$$\begin{aligned}\text{Then}\ \ t_{11}^1\,1; \mathcal{I}(t_{11}^1)\,(1; 1; \epsilon_{11}); \epsilon_{12} \dagger (\gamma_{21} \dagger \gamma_{31}) &= t_{11}^1\,1; \gamma_4\,(1; 1; \epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} \\ &= t_{11}^1\,1; \gamma_4\,\mathcal{I}(1); t_4^1\,1 \dagger \alpha_5\,1; \beta_{t_5,1}; \gamma_5 \\ &= t_{11}^1; \gamma_6\,(1; 1; \epsilon_{61}); \epsilon_{62} \dagger \gamma_5 \\ &= t_{11}^1\,1; \gamma_7\,(1; 1; \epsilon_{71}); \epsilon_{72}\end{aligned}$$

$$\begin{aligned}
\text{so } \gamma_1 \dagger (\gamma_2 \dagger \gamma_3) &= \gamma_{11} \dagger \gamma_7; \eta_{t_7}; \lambda(t_7^1\, \epsilon_{71}; \epsilon_{72});\,; \epsilon_3 \\
&= \gamma_{11} \dagger \mathcal{I}(t_{11}); \eta_{t_{11}}; \lambda(t_{11}^1\, 1; t_{11}^1\, 1) \dagger \lambda(t_{11}^1\, 1); \lambda(t_{11}^1; \gamma_6\, (1;1;\epsilon_{61}); \epsilon_{62} \dagger \gamma_5); \epsilon_3 \\
&= \gamma_{11} \dagger \gamma_6; \eta_{t_6}; \lambda(t_6^1\, \epsilon_{61}; \epsilon_{62}) \dagger \lambda t_5; \lambda\gamma_5; \epsilon_3 \\
&= \gamma_{11} \dagger \gamma_4 \dagger \alpha_5; \lambda\gamma_5; \epsilon_3 \\
&= (\gamma_1 \dagger \gamma_2) \dagger \gamma_3
\end{aligned}$$

case 2.4: $t_{11}^1\, 1; \mathcal{I}(t_{11}^1)\, (1;1;\epsilon_{11}); \epsilon_{12} \dagger \gamma_{21} = t_{11}^1\, 1; \gamma_4\, (1;1;\epsilon_{41}); \epsilon_{42}$, $t_4^1\, 1; \mathcal{I}(t_4^1)\, (1;1;\epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} = \alpha_5\, 1; \beta_{t_5,1}; \gamma_5$ $\gamma_4 \dagger \alpha_5; \lambda\mathcal{I}(t_5); \lambda t_5 = \gamma_6; \eta_{t_6}; \lambda(t_6^1\, \epsilon_{61}; \epsilon_{62})$ and $t_{11}^1; \gamma_6\, (1;1;\epsilon_{61}); \epsilon_{62} \dagger \gamma_5 = \alpha_7\, 1; \beta_{t_7,1}; \gamma_7$.

$$\begin{aligned}
\text{Then } t_{11}^1\, 1; \mathcal{I}(t_{11}^1)\, (1;1;\epsilon_{11}); \epsilon_{12} \dagger (\gamma_{21} \dagger \gamma_{31}) &= t_{11}^1\, 1; \gamma_4\, (1;1;\epsilon_{41}); \epsilon_{42} \dagger \gamma_{31} \\
&= t_{11}^1\, 1; \gamma_4 \mathcal{I}(1); t_4^1\, 1 \dagger \alpha_5\, 1; \beta_{t_5,1}; \gamma_5 \\
&= t_{11}^1; \gamma_6\, (1;1;\epsilon_{61}); \epsilon_{62} \dagger \gamma_5 \\
&= \alpha_7\, 1; \beta_{t_7,1}; \gamma_7
\end{aligned}$$

$$\begin{aligned}
\text{so } \gamma_1 \dagger (\gamma_2 \dagger \gamma_3) &= \gamma_{11} \dagger \alpha_7; \lambda\gamma_7; \epsilon_3 \\
&= \gamma_{11} \dagger \mathcal{I}(t_{11}); \eta_{t_{11}}; \lambda(t_{11}^1\, 1; t_{11}^1\, 1) \dagger \lambda(t_{11}^1\, 1); \lambda(t_{11}^1; \gamma_6\, (1;1;\epsilon_{61}); \epsilon_{62} \dagger \gamma_5); \epsilon_3 \\
&= \gamma_{11} \dagger \gamma_6; \eta_{t_6}; \lambda(t_6^1\, \epsilon_{61}; \epsilon_{62}) \dagger \lambda t_5; \lambda\gamma_5; \epsilon_3 \\
&= \gamma_{11} \dagger \gamma_4 \dagger \alpha_5; \lambda\gamma_5; \epsilon_3 \\
&= (\gamma_1 \dagger \gamma_2) \dagger \gamma_3
\end{aligned}$$

case 2.5: $t_{11}^1\, 1; \mathcal{I}(t_{11}^1)\, (1;1;\epsilon_{11}); \epsilon_{12} \dagger \gamma_{21} = \alpha_4\, 1; \beta_{t_4,1}; \gamma_4$.
Then $t_{11}^1\, 1; \mathcal{I}(t_{11}^1)\, (1;1;\epsilon_{11}); \epsilon_{12} \dagger (\gamma_{21} \dagger \gamma_{31}) = \alpha_4\, 1; \beta_{t_4,1}; \gamma_4 \dagger \gamma_{31}$

$$\begin{aligned}
\text{so } \gamma_1 \dagger (\gamma_2 \dagger \gamma_3) &= \gamma_{11} \dagger \alpha_4; \lambda(\gamma_4 \dagger \gamma_{31}); \epsilon_3 \\
&= (\gamma_1 \dagger \gamma_2) \dagger \gamma_3
\end{aligned}$$

□

With each rewrite $\Gamma \vdash \gamma\colon t \Rightarrow u\colon X$ we associate a rewrite $\Gamma \vdash \mathcal{F}(\gamma)\colon t \Rightarrow u\colon X$ in $\mathcal{G}$ as follows:

$$\begin{aligned}
\mathcal{F}(j) &= j; j; j \\
\mathcal{F}(\lambda\gamma) &= \lambda t; \lambda\mathcal{F}(\gamma); \lambda u \quad \text{where } X, \Gamma \vdash \gamma\colon t \Rightarrow u\colon Y \\
\mathcal{F}(\gamma_1\, \gamma_2) &= t_1\, t_2; \mathcal{F}(\gamma_1)\, \mathcal{F}(\gamma_2); u_1\, u_2 \quad \text{where } \Gamma \vdash \gamma_j\colon t_j \Rightarrow u_j\colon X_j \\
\mathcal{F}(\gamma_1; \gamma_2) &= \mathcal{F}(\gamma_1) \dagger \mathcal{F}(\gamma_2) \\
\mathcal{F}(\eta_t) &= \mathcal{I}(t);\,; t; \eta_t; \lambda(t^1\, 1; t^1\, 1) \\
\mathcal{F}(\beta_{t,u}) &= \lambda t\, u; \beta_{t,u}; t[u];\,; \mathcal{I}(t[u])
\end{aligned}$$

Proposition 15. If $\Gamma \vdash \gamma\colon t \Rightarrow u\colon X$ then $\Gamma \vdash \mathcal{F}(\gamma) = \gamma\colon t \Rightarrow u\colon X$.

Proof. Straightforward induction. The work has already been done in lemmas 10–12. □

Proposition 16. If $\Gamma \vdash \gamma = \delta\colon t \Rightarrow u\colon X$ then $\mathcal{F}(\gamma) = \mathcal{F}(\delta)$.

Proof. By induction on the length of the derivation of $\Gamma \vdash \gamma = \delta \colon t \Rightarrow u \colon X$. The hard cases have already been done in lemmas 13–14. □

Theorem 17. The set $\mathcal{G}$ contains exactly one member of each equivalence class of the rewrites quotiented by the theory 2-λ.

Proof. This follows immediately from the last two results. □

Corollary 18. The theory 2-λ is consistent, in the sense that it does not identify everything possible.

Proof. By the theorem, it suffices to give two distinct elements of $\mathcal{G}$ with the same source and target. We give two different examples:

- If $X, \Gamma \vdash t \colon Y$ then $\Gamma \vdash \lambda(\lambda t^2\, 1); \lambda\mathcal{F}(\beta_{t^2,1}); \lambda t; \eta_{\lambda t}; \lambda(\lambda t^2\, 1; \lambda t^2\, 1) \colon \lambda(\lambda t^2\, 1) \Rightarrow \lambda(\lambda t^2\, 1) \colon X \to Y$ is in $\mathcal{G}$, but it is not equal to $\mathcal{I}(\lambda(\lambda t^2\, 1))$.
- Let $I = \lambda 1$. Then $X \vdash I\,(I\,1); \mathcal{I}(I)\,\mathcal{F}(\beta_{1,1}); I\,1 \colon I\,(I\,1) \Rightarrow I\,1 \colon X$ and $X \vdash \mathcal{F}(\beta_{1,I\,1}) \colon I\,(I\,1) \Rightarrow I\,1 \colon X$ are both in $\mathcal{G}$, but they are not equal.

□

5. The 2-λ-calculus as a rewrite system

In this section we investigate the 2-λ-calculus as a rewrite system, looking in particular at confluence, normalisation and normal forms. Of course, the underlying system is just the simply-typed λ-calculus, so the results are already known. Our aim here is to study the relationship between these properties from rewriting theory and the equations of the theory 2-λ.

From this point on, we consider rewrites up to equivalence. The formalism of the first part of the paper has done its job, and we no longer need the notion of syntactic equality. We can assume that any rewrite is in $\mathcal{G}$ even though we will use the rules of 2-λ to reason about them, and write $\beta_{t,u}$ instead of $\lambda t\, u; \beta_{t,u}; t[u]; ; \mathcal{I}(t[u])$ and $\gamma; \delta$ instead of $\gamma \dagger \delta$. The more pedantic reader can insert $\mathcal{F}$ at every appropriate point.

5.1. Confluence. The simply typed λ-calculus is well known to be confluent/ Church-Rosser/ have the diamond property: that if $\gamma_1 \colon t \Rightarrow u_1$ and $\gamma_2 \colon t \Rightarrow u_2$ are two rewrites with a common source then there exist $\delta_1 \colon u_1 \Rightarrow v$ and $\delta_2 \colon u_2 \Rightarrow v$ with common target [1]. We will prove the stronger *commuting* diamond property: that δ_1 and δ_2 can be chosen so that $\vdash \gamma_1; \delta_1 = \gamma_2; \delta_2$.

The first lemma we prove says that if two rewrites in $\mathcal{A}$ have a common source, one is a prefix of the other:

Lemma 19. If $\alpha_1 \colon t \Rightarrow u_1$ and $\alpha_2 \colon t \Rightarrow u_2$ are in $\mathcal{A}$ then either there exists $\alpha_3 \colon u_1 \Rightarrow u_2$ in $\mathcal{A}$ such that $\vdash \alpha_1; \alpha_3 = \alpha_2$ or vice versa.

Proof. By induction on the structure of α_1 and α_2.

case 1: $\alpha_1 = t$. Then $u_1 = t$, take $\alpha_3 = \alpha_2$.

case 2: $\alpha_2 = t$. Then $u_2 = t$, take $\alpha_3 = \alpha_1$.

case 3: $\alpha_j = \alpha_{j1}\, t_2; \beta_{v_j,t_2}; \alpha_{j2}$ where $t = t_1\, t_2$, $\alpha_{j1} \colon t_1 \Rightarrow \lambda v_j$.

Apply the inductive hypothesis to α_{11}, α_{21} to get (without loss of generality) $\alpha_{31} \colon \lambda v_1 \Rightarrow \lambda v_2$. But any rewrite in $\mathcal{A}$ with source a λ-term is identity, so $v_1 = v_2$ and $\alpha_{11} = \alpha_{21}$. The result then follows by applying the inductive hypothesis to $\alpha_{12} \colon v_1[t_2] \Rightarrow u_1$ and $\alpha_{22} \colon v_2[t_2] \Rightarrow u_2$. □

Next we turn our attention to rewrites in $\mathcal{E}$: they commute with any other rewrite:

Lemma 20. If $\epsilon\colon t \Rightarrow u_1$ is in $\mathcal{E}$ and $\gamma\colon t \Rightarrow u_2$ then there exist $\gamma'\colon u_1 \Rightarrow v$ and $\epsilon'\colon u_2 \Rightarrow v$ in $\mathcal{E}$ such that $\vdash \epsilon; \gamma' = \gamma; \epsilon'$.

Proof. By induction on the structure of ϵ.
case 1: $\epsilon = t$. Then $u_1 = t$, take $\gamma' = \gamma$ and $\epsilon' = u_2$.
case 2: $\epsilon = \epsilon_1; \eta_w; \lambda(w^1\, \epsilon_2; \epsilon_3)$, where $\epsilon_1\colon t \Rightarrow w$, $\epsilon_2\colon 1 \Rightarrow x$ and $\epsilon_3\colon w^1\, x \Rightarrow y$.
Apply the inductive hypothesis to ϵ_1 and γ to get ϵ_1' and γ_1, then apply it to ϵ_3 and $\gamma_1^1\, x$ to get γ_2 and ϵ_3'. We then have

$$\begin{aligned}
\epsilon; \lambda\gamma_2 &= \epsilon_1; \eta_w; \lambda(w^1\, \epsilon_2; \epsilon_3); \lambda\gamma_2 \\
&= \epsilon_1; \eta_w; \lambda(w^1\, \epsilon_2; \epsilon_3; \gamma_2) \\
&= \epsilon_1; \eta_w; \lambda(w^1\, \epsilon_2; \gamma_1^1\, x; \epsilon_3') \\
&= \epsilon_1; \gamma_1; \eta_{w'}; \lambda(w'^1\, \epsilon_2; \epsilon_3') \\
&= \gamma; \epsilon_1'; \eta_{w'}; \lambda(w'^1\, \epsilon_2; \epsilon_3')
\end{aligned}$$

so take $\gamma' = \lambda\gamma_2$ and $\epsilon' = \epsilon_1'; \eta_{w'}; \lambda(w'^1\, \epsilon_2; \epsilon_3')$. □

A final lemma to say how rewrites in $\mathcal{E}$ interact with β-reductions:

Lemma 21. If $\epsilon\colon \lambda t \Rightarrow u$ and $\epsilon'\colon u\, s \Rightarrow v$ are in $\mathcal{E}$ then there exist w, $\gamma\colon v \Rightarrow w$ and $\gamma'\colon t[s] \Rightarrow w$ such that $\vdash \epsilon\, s; \epsilon'; \gamma = \beta_{t,s}; \gamma'$.

Proof. By induction on the structure of ϵ.
case 1: $\epsilon = \lambda t$. Then $u = \lambda t$; apply lemma 20 to ϵ' and $\beta_{t,s}$.
case 2: $\epsilon = \epsilon_1; \eta_x; \lambda(x^1\, \epsilon_2; \epsilon_3)$, where $\epsilon_1\colon \lambda t \Rightarrow x$, $\epsilon_2\colon 1 \Rightarrow r$ and $\epsilon_3\colon x^1\, 1 \Rightarrow y$.
Apply lemma 20 to ϵ' and $\beta_{y,s}$ to get $\gamma_1\colon v \Rightarrow w_1$ and $\epsilon''\colon y[s] \Rightarrow w_1$, then apply the inductive hypothesis to ϵ_1 and $\epsilon_3[s]; \epsilon''$ to get $\gamma_2\colon w_1 \Rightarrow w$ and $\gamma_3\colon t[r[s]] \Rightarrow w$. We then have

$$\begin{aligned}
\epsilon\, s; \epsilon'; \gamma_1; \gamma_2 &= \epsilon_1\, s; \eta_x\, s; \lambda(x^1\, \epsilon_2; \epsilon_3)\, s; \beta_{y,s}; \epsilon''; \gamma_2 \\
&= \epsilon_1\, s; x\, \epsilon_2[s]; \epsilon_3[s]; \epsilon''; \gamma_2 \\
&= \lambda t\, \epsilon_2[s]; \epsilon_1\, r[s]; \epsilon_3[s]; \epsilon''; \gamma_2 \\
&= \lambda t\, \epsilon_2[s]; \beta_{t,r[s]}; \gamma_3 \\
&= \beta_{t,s}; t[\epsilon_2[s]]; \gamma_3
\end{aligned}$$

so take $\gamma = \gamma_1; \gamma_2$ and $\gamma' = t[\epsilon_2[s]]; \gamma_3$. □

We are now in a position to prove confluence. For this (and another) proof, some more sophisticated well-foundedness is needed: the usual proof of confluence of the λ-calculus depends upon 'finiteness of developments' [1]. Rather than set up all that machinery here, we use the fact that the simply-typed λ-calculus (with *η-reduction*) is strongly normalising. We write $||t||$ for the length of the longest β-η-reduction path starting from t.

Proposition 22. If $\gamma_1\colon t \Rightarrow u_1$ and $\gamma_2\colon t \Rightarrow u_2$ then there exist $\gamma_1'\colon u_1 \Rightarrow v$ and $\gamma_2'\colon u_2 \Rightarrow v$ such that $\vdash \gamma_1; \gamma_1' = \gamma_2; \gamma_2'$.

Proof. By induction on $||t||$.

Let $(\alpha_j; \delta_j; \epsilon_j) = \gamma_j$ for $j = 1, 2$. Then α_1 and α_2 have the same domain; by lemma 19 there exists α_3 such that (wlog) $\alpha_2 = \alpha_1; \alpha_3$.

case 1: $\delta_1 = j$. Apply lemma 20 to ϵ_1 and $(\alpha_3; \delta_2; \epsilon_2)$.

case 2: $\delta_1 = \lambda\gamma_{11}$, $\alpha_3 = \lambda t$ and $\delta_2 = \lambda\gamma_{21}$. Apply the inductive hypothesis to γ_{11} and γ_{21} to get γ_{12} and γ_{22}; two applications of lemma 20 then give the answer.

case 3: $\delta_1 = \gamma_{11}\,\gamma_{12}$ and $\alpha_3 = t_1\,t_2$. Similar to case 2.

case 4: $\delta_1 = \gamma_{11}\,\gamma_{12}$ and $\alpha_3 = \alpha_{31}\,t_2; \beta_{t_3,t_2}; \alpha_{32}$.

Apply the inductive hypothesis to $\gamma_{11}\colon t_1 \Rightarrow v_1$ and $\alpha_{31}\colon t_1 \Rightarrow \lambda t_3$ to get $\gamma_3\colon v_1 \Rightarrow w_1$ and $\gamma_4\colon \lambda t_3 \Rightarrow w_1$, and lemma 20 to ϵ_1 and $\gamma_3\,v_2$ to get γ_5 and ϵ_3. The domain of γ_4 is a lambda term, so it is of the form $(\lambda t_3; \lambda\gamma_{41}; \epsilon_4)$. Apply lemma 21 to ϵ_4 and ϵ_3 to get γ_6 and γ_7. The inductive hypothesis can now be applied to $\gamma_{41}[\gamma_{12}]; \gamma_7$ and $\alpha_{32}; \delta_2; \epsilon_2$ to get γ_8 and γ_9. We then have:

$$\begin{aligned}
\gamma_2; \gamma_9 &= \alpha_1; \alpha_{31}\,t_2; \beta_{t_3,t_2}; \alpha_{32}; \delta_2; \epsilon_2; \gamma_9 \\
&= \alpha_1; \alpha_{31}\,t_2; \lambda\gamma_{41}\,\gamma_{12}; \beta_{t_4,v_2}; \gamma_7; \gamma_8 \\
&= \alpha_1; \alpha_{31}\,t_2; \lambda\gamma_{41}\,\gamma_{12}; \epsilon_4\,v_2; \epsilon_3; \gamma_6; \gamma_8 \\
&= \alpha_1; \gamma_{11}\,\gamma_{12}; \epsilon_1; \gamma_5; \gamma_6; \gamma_8 \\
&= \gamma_1; \gamma_5; \gamma_6; \gamma_8
\end{aligned}$$

so take $\gamma_1' = \gamma_5; \gamma_6; \gamma_8$ and $\gamma_2' = \gamma_9$. □

5.2. Mellifluence. The 2-λ-calculus has another property, related to confluence, which cannot be formulated sensibly in the λ-calculus. This property is essential in relating confluence, strong normalisation and normal forms, as we show in section 6. In this section we show that every rewrite is mellifluent, where:

Definition. A rewrite $\gamma : t \Rightarrow u$ is **mellifluent** if whenever $\delta_1, \delta_2 : u \Rightarrow v$ satisfy $\gamma; \delta_1 = \gamma; \delta_2$, there exists $\gamma' : v \Rightarrow w$ such that $\delta_1; \gamma' = \delta_2; \gamma'$.

Lemma 23.

(1) Any rewrite in $\mathcal{A}$ is mellifluent.
(2) If γ_1 and γ_2 are mellifluent, then $\gamma_1; \gamma_2$ is mellifluent.
(3) If $\gamma_1; \gamma_2$ is mellifluent, then γ_2 is mellifluent.

Proof. Straightforward. □

Lemma 24. If $\gamma\colon t \Rightarrow u$ is mellifluent, then $\lambda\gamma\colon \lambda t \Rightarrow \lambda u$ is mellifluent.

Proof. Let $\gamma_1, \gamma_2\colon \lambda u \Rightarrow v$ satisfy $\lambda\gamma; \gamma_1 = \lambda\gamma; \gamma_2$. Then $\gamma_j = (\lambda u; \lambda\gamma_{j1}; \epsilon_j)$ so $\lambda\gamma; \gamma_j = (\lambda t; \lambda(\gamma; \gamma_{j1}); \epsilon_j)$ which is in canonical form, so $\epsilon_1 = \epsilon_2$ and $\gamma; \gamma_{11} = \gamma; \gamma_{21}$. But γ is mellifluent, so there exists γ_3 satisfying $\gamma_{11}; \gamma_3 = \gamma_{21}; \gamma_3$. Now by lemma 20 there exist γ_4 and ϵ_4 satisfying $\lambda\gamma_3; \epsilon_4 = \epsilon_j; \gamma_4$, and $\gamma_1; \gamma_4 = \gamma_2; \gamma_4$ so $\lambda\gamma$ is mellifluent. □

The next lemma describes a property of rewrites in $\mathcal{E}$:

Lemma 25. If $\epsilon\colon \lambda t \Rightarrow u$ in $\mathcal{E}$ then $u = \lambda u'$ and there exist $\gamma_1\colon t \Rightarrow v$ and $\gamma_2\colon u' \Rightarrow v$ such that $\epsilon; \lambda\gamma_2 = \lambda\gamma_1$.

Proof. By induction on the structure of ϵ.

case 1: $\epsilon = \lambda t$. Then $u' = t$; take $\gamma_1 = \gamma_2 = t$.

case 2: $\epsilon = \epsilon_1; \eta_{u_1}; \lambda(u_1^1\, \epsilon_2; \epsilon_3)$

Apply the inductive hypothesis to ϵ_1 to get $u_1 = \lambda u_2$, $\gamma_3 \colon t \Rightarrow v_1$ and $\gamma_4 \colon u_2 \Rightarrow v_1$. We now have $u_1^1\, \epsilon_2; \epsilon_3 \colon u_1^1\, 1 \Rightarrow u'$ and $\beta_{u_2^2,1}; \gamma_4 \colon u_1^1\, 1 \Rightarrow v_1$; apply confluence to get γ_5 and γ_6. We then have:

$$\begin{aligned}
\epsilon; \lambda\gamma_6 &= \epsilon_1; \eta_{u_1}; \lambda(u_1^1\, \epsilon_2; \epsilon_3); \lambda\gamma_6 \\
&= \epsilon_1; \eta_{u_1}; \lambda(\beta_{u_2^1,1}; \gamma_4); \lambda\gamma_5 \\
&= \epsilon_1; \lambda\gamma_4; \lambda\gamma_5 \\
&= \lambda(\gamma_3; \gamma_5)
\end{aligned}$$

so take $\gamma_1 = \gamma_3; \gamma_5$ and $\gamma_2 = \gamma_6$. □

From this follows that η-expansion is mellifluent:

Lemma 26. The rewrite $\eta_t \colon t \Rightarrow \lambda(t^1\, 1)$ is mellifluent.

Proof. Let $\gamma_1, \gamma_2 \colon \lambda(t^1\, 1) \Rightarrow v$ satisfy $\eta_t; \gamma_1 = \eta_t; \gamma_2$. Then $\gamma_j = (\lambda(t^1\, 1); \lambda\gamma_{j1}; \epsilon_j)$, where $\epsilon_j \colon \lambda u_j \Rightarrow v$. By lemma 25, there exist γ_{j2}, γ_{j3} such that $\epsilon_j; \lambda\gamma_{j3} = \lambda\gamma_{j2}$; apply confluence to γ_{13} and γ_{23} to get γ_{14} and γ_{24}.

We will take $\gamma_3 = \lambda(\gamma_{13}; \gamma_{14}) = \lambda(\gamma_{23}; \gamma_{24})$, and prove that $\gamma_1; \gamma_3 = \gamma_2; \gamma_3$.

If $\gamma_{j5} = \gamma_{j1}; \gamma_{j2}; \gamma_{j4}$, then $\eta_t; \lambda\gamma_{15} = \eta_t; \lambda\gamma_{25}$ and $\gamma_j; \gamma_3 = \gamma_{j5}$. Let $\gamma_{j5} = (\alpha_j; \delta_j; \epsilon_{j1})$

case 1: $\alpha_1 = \alpha_2 = t^1\, 1$. Then $\delta_j = \gamma_{j6}^1\, (1; 1; \epsilon_{j2})$, and $\eta_t; \lambda\gamma_{j5} = \gamma_{j6}; \eta_w; \lambda(w^1\, \epsilon_{j2}; \epsilon_{j1})$ which is in canonical form. Therefore, $\gamma_{16} = \gamma_{26}$, $\epsilon_{12} = \epsilon_{22}$ and $\epsilon_{11} = \epsilon_{21}$, so $\gamma_{15} = \gamma_{25}$.

case 2: $\alpha_j = \alpha_{j1}^1\, 1; \beta_{w,1}; \alpha_{j2}$. Then $\eta_t; \lambda\gamma_{j5} = (\alpha_{j1}; \lambda(\alpha_{j2}; \delta_j; \epsilon_{j1}); \lambda z)$ which is in canonical form; matching up as before gives $\gamma_{15} = \gamma_{25}$.

case 3: One of each. This case is impossible, since the two canonical forms cannot match. □

Finally:

Proposition 27. All the rewrites of the 2-λ-calculus are mellifluent.

Proof. By induction on $||u||$, where $\gamma : t \Rightarrow u$.

Let $\gamma = (\alpha; \delta; \epsilon)$. By lemma 23, it is sufficient to prove that $\delta \colon t' \Rightarrow u'$ and $\epsilon \colon u' \Rightarrow u$ are mellifluent; note that $||u'|| \leq ||u||$. First consider ϵ:

case 1: $\epsilon = t$. This is identity, therefore mellifluent.

case 2: $\epsilon = \epsilon_1; \eta_v; \lambda(v^1\, \epsilon_2; \epsilon_3)$. Then $v^1\, \epsilon_2; \epsilon_3 \colon v^1\, 1 \Rightarrow u_2$ where $\lambda u_2 = u$ so $||u_2|| < ||u||$ so by the inductive hypothesis, $v^1\, \epsilon_2; \epsilon_3 \colon v^1\, 1 \Rightarrow u_2$ is mellifluent, and by lemma 24, $\lambda(v^1\, \epsilon_2; \epsilon_3)$ is mellifluent. Also, $\epsilon_1 \colon u' \Rightarrow v$, and $||v|| < ||\lambda(v^1\, 0)|| \leq ||u||$ so by the inductive hypothesis, ϵ_1 is mellifluent. Finally, η_v is mellifluent by lemma 26, so ϵ is mellifluent.

Next we consider $\delta \colon t' \Rightarrow u'$:

case 1: $\delta = j$. This is identity, therefore mellifluent.

case 2: $\delta = \lambda\gamma_1$. This is mellifluent by the inductive hypothesis and lemma 24.

case 3: $\delta = \gamma_1\, \gamma_2$. Then $t' = t_1\, t_2$ and $u' = u_1\, u_2$. Let $\gamma_3, \gamma_4 \colon u' \Rightarrow v$ satisfy $\delta; \gamma_3 = \delta; \gamma_4$ and proceed by cases of γ_3 and γ_4:

case 3.1: $\gamma_j = (u_1\, u_2; \gamma_{j1}\, \gamma_{j2}; \epsilon_j)$ for $j = 3, 4$. Then $\delta; \gamma_j = (t'; (\gamma_1; \gamma_{j1})\, (\gamma_2; \gamma_{j2}); \epsilon_j)$ which is in canonical form, so $\epsilon_3 = \epsilon_4$ and $\gamma_k; \gamma_{3k} = \gamma_k; \gamma_{4k}$ for $k = 1, 2$. By the inductive hypothesis, there exist γ_{5k} satisfying $\gamma_{3k}; \gamma_{5k} = \gamma_{4k}; \gamma_{5k}$, and by confluence there exist γ_6, ϵ_6 satisfying $\gamma_{51}\,\gamma_{52}; \epsilon_6 = \epsilon_j; \gamma_6$. Then $\gamma_3; \gamma_6 = \gamma_4; \gamma_6$, so δ is mellifluent.

case 3.2: $\gamma_j = (\alpha_j\, u_2; \beta_{v_j,u_2}; \gamma_{j1})$ for $j = 3, 4$. Then $\alpha_3 = \alpha_4$ and $v_3 = v_4$ by lemma 19, so $\delta; \gamma_j = (\delta; \alpha_j\, u_2; \beta_{v_j,u_2}); \gamma_{j1}$. But $\delta; \alpha_j\, u_2; \beta_{v_j,u_2} \colon t_1 \Rightarrow v_j[u_2]$ and $||v_j[u_2]|| < ||\lambda v_j\, u_2|| \leq ||u'||$, so by the inductive hypothesis it is mellifluent, and there exists γ_5 satisfying $\gamma_{31}; \gamma_5 = \gamma_{41}; \gamma_5$. Then $\gamma_3; \gamma_5 = \gamma_4; \gamma_5$, so δ is mellifluent.

case 3.3: $\gamma_3 = (u'; \gamma_{31}\,\gamma_{32}; \epsilon_3)$ and $\gamma_4 = (\alpha_4\, u_2; \beta_{v_4,u_2}; \gamma_{41})$, or vice versa. This case cannot arise since $\delta; \gamma_4$ and $\delta; \gamma_3$ then have different canonical forms. □

5.3. Normal Forms. Since the 2-λ-calculus has η-*expansion*, the relevant normal forms are Huet's long-β-η-normal forms [6]. In this section we characterise them entirely in terms of how they can be rewritten.

In standard rewriting theory, normal forms cannot be rewritten at all. This is not the case in the 2-λ-calculus, as there are usually further irrelevant η-expansions, and always identity rewrites. However, every rewrite whose source is a long-β-η-normal form is reversible, where

Definition.

- A rewrite $\gamma \colon t \Rightarrow u$ is **reversible** if there exists $\delta \colon u \Rightarrow t$ such that $\gamma; \delta = t$.
- A term t is **normal** if every rewrite $\gamma \colon t \Rightarrow u$ is reversible.

and conversely, every normal term is a long-β-η-normal form.

Recall that the definition of long-$\beta\eta$-normal form is, in our notation,

- $t \colon X \to Y$ is in long-$\beta\eta$-normal form iff $t = \lambda t'$ where $t' \colon Y$ is in long-$\beta\eta$-normal form.
- $t \colon B$ is in long-$\beta\eta$-normal form iff t is in *reduced* form.
- j is in reduced form.
- $t_1\, t_2$ is in reduced form iff t_1 is in reduced form and t_2 is in long-$\beta\eta$-normal form.
- λt is not in reduced form.

We will prove that the arrows from long-$\beta\eta$-normal forms are reversible. First we must define the corresponding notion for reduced forms. Let $\mathcal{P}$ be the smallest set of rewrites such that:

- All identity rewrites are in $\mathcal{P}$.
- If $\gamma \in \mathcal{P}$ then $\gamma; \epsilon \in \mathcal{P}$, for all ϵ in $\mathcal{E}$.
- If $\gamma_1; \gamma_2 \in \mathcal{P}$ then $\gamma_1 \in \mathcal{P}$.

Note that we have defined $\mathcal{P}$ for each type independently. In particular, for base types B, the second clause does not apply, and $\gamma \in \mathcal{P}$ iff γ is reversible.

The following lemma describes a closure property of $\mathcal{P}$ which relates rewrites of different type:

Lemma 28. If $\gamma_1 \colon t_1 \Rightarrow u_1 \colon X \to Y \in \mathcal{P}$ and $t_2 \colon X$ is normal, then for any $\gamma_2 \colon t_2 \Rightarrow u_2$, $\gamma_1\,\gamma_2 \in \mathcal{P}$.

Proof. We prove that this property is preserved by the three clauses defining $\mathcal{P}$.

- If γ_1 is identity then $\gamma_1\,\gamma_2$ is reversible, so a member of $\mathcal{P}$.

- If γ_1 has this property then we prove that $\gamma_1; \epsilon$ does by structural induction on ϵ.
  case 1: $\epsilon = u_1$. Then $\gamma_1; \epsilon = \gamma_1$.
  case 2: $\epsilon = \epsilon_1; \eta_u; \lambda(u^1\, \epsilon_2; \epsilon_3)$. Then $(\gamma_1; \epsilon\, \gamma_2); \beta_{v,u_2} = (t_1; (\gamma_1; \epsilon_1)\, (\gamma_2; \epsilon_2[v]); \epsilon_3[v])$ which, by inductive hypothesis, is a member of $\mathcal{P}$. Therefore $\gamma_1; \epsilon \in \mathcal{P}$ as required.
- If $\gamma_1; \gamma_3 \in P$ has this property, then $(\gamma_1\, \gamma_2); (\gamma_3\, u_2) = \gamma_1; \gamma_3\, \gamma_2$ is a member of $\mathcal{P}$, so $\gamma_1\, \gamma_2 \in \mathcal{P}$ as required.

□

We are now ready to prove half our theorem:

Proposition 29.
- If $\gamma\colon t \Rightarrow u$ and t is in long-$\beta\eta$-normal form, then γ has a left inverse.
- If $\gamma\colon t \Rightarrow u$ and t is in reduced form, then $\gamma \in \mathcal{P}$.

Proof. By structural induction on t. We proceed by cases:

case 1: $t\colon X \to Y$ is in long-$\beta\eta$-normal form. Then $t = \lambda t_1$ where t_1 is in l$\beta\eta$nf, and $\gamma = (t; \lambda\gamma_1; \epsilon)$. By lemma 25, there exist $\gamma_2\colon u_1 \Rightarrow v$ and $\gamma_3\colon u_2 \Rightarrow v$ s.t. $\epsilon; \lambda\gamma_3 = \lambda\gamma_2$. Then $\gamma_1; \gamma_2\colon t_1 \Rightarrow v$ and by inductive hypothesis, has a left inverse γ_4. Now $\lambda\gamma_3; \gamma_4$ is a left inverse for γ.

case 2: $t\colon B$ is in long-$\beta\eta$-normal form. Then t is in reduced form, and by the second inductive hypothesis, $\gamma \in \mathcal{P}$. As remarked above, this means γ is reversible.

case 3: $t = j$ is in reduced form. Then $\gamma = (j; j; \epsilon)$ which is certainly in $\mathcal{P}$.

case 4: $t = t_1\, t_2$ is in reduced form. Then $\gamma = (t; \gamma_1\, \gamma_2; \epsilon)$ by a simple induction. By inductive hypothesis $\gamma_1 \in \mathcal{P}$, and t_2 is normal, so $\gamma_1\, \gamma_2 \in \mathcal{P}$ by lemma 28. Therefore $\gamma \in \mathcal{P}$ as required. □

The next lemma tells us more about the rewrites in $\mathcal{P}$:

Lemma 30. Every $\gamma \in \mathcal{P}$ is of the form $(t; \delta; \epsilon)$ where δ satisfies one of the following:
- $\delta = j$
- $\delta = \lambda\gamma_1$ and $\gamma_1 \in \mathcal{P}$
- $\delta = \gamma_1\, \gamma_2$ and $\gamma_1 \in \mathcal{P}$, γ_2 is reversible.

Proof. We prove that this property is preserved by the three clauses defining $\mathcal{P}$. It is clear that all identities are of this form, and that it is preserved by composition with rewrites of the form ϵ. It remains to prove that if $\gamma_1; \gamma_2 \in \mathcal{P}$ is of one of the three forms above, then so is γ_1. The proof is by induction on $|\gamma_2|_{\mathcal{G}}$.

Let $\gamma_j = (\alpha_j; \delta_j; \epsilon_j)$; it is clear from the definition of † that $\alpha_1 = t$. We proceed by cases of ϵ_1 and α_2:

case 1: $\epsilon_1 = \alpha_2 = u$. There are three subcases, depending on the form of δ_j:

case 1.1: $\delta_j = i$. Then γ_1 is of the required form.

case 1.2: $\delta_j = \lambda\gamma_{j1}$. Then $\gamma_1; \gamma_2 = (t; \lambda(\gamma_{11}; \gamma_{21}); \epsilon_2)$ and $\gamma_{11}; \gamma_{21} \in \mathcal{P}$, so $\gamma_{11} \in \mathcal{P}$ and γ_1 is of the required form.

case 1.3: $\delta_j = \gamma_{j1}\, \gamma_{j2}$. Then $\gamma_1; \gamma_2 = (t; (\gamma_{11}; \gamma_{21})\, (\gamma_{12}; \gamma_{22}); \epsilon_2)$ and $\gamma_{11}; \gamma_{21} \in \mathcal{P}$, $\gamma_{12}; \gamma_{22}$ is reversible. Then $\gamma_{11} \in \mathcal{P}$ and γ_{12} is reversible, so γ_1 is of the required form.

case 2: $\alpha_2 = u$, $\epsilon_1 \neq u$. Then $\gamma_1; \gamma_2 = (t; \delta_1; v); (\epsilon_1; \gamma_2)$ and by inductive hypothesis, $(t; \delta_1; v)$ is of the required form. Therefore γ_1 is also.

case 3: $\epsilon_1 = u$, $\alpha_2 = \alpha_{21}\, t_2; \beta_{t_1,t_2}; \alpha_{22}$. Then $\delta_1 = \gamma_{11}\, \gamma_{12}$, and we proceed by cases of $\gamma_{11} \dagger \alpha_{21}$:

case 3.1: $\gamma_{11}; \alpha_{21} = (\alpha_3; \lambda\gamma_3; \lambda t_1)$. Then $\gamma_1; \gamma_2 = (\alpha_3\, v; \beta_{w,v}; \ldots)$, contradicting the hypothesis that it is of the given form.

case 3.2: $\gamma_{11}; \alpha_{21} = (\gamma_3; \eta_w; \lambda(w^1\, \epsilon_{31}; \epsilon_{32}))$. Then

$$\gamma_1; \gamma_2 = (t; \gamma_3\, (\gamma_{12}; \epsilon_{31}[u]); \epsilon_{32}[u]); (\alpha_{22}; \delta_2; \epsilon_2)$$

By inductive hypothesis, $\gamma_3 \in \mathcal{P}$ and $\gamma_{12}; \epsilon_{31}[u]$ is reversible, so $\gamma_{11} \in P$ and γ_{12} is reversible, as required. □

We can now prove the other half of the theorem:

Proposition 31. Let t be a term. Then

- If every $\gamma\colon t \Rightarrow u$ is reversible, then t is in long-$\beta\eta$-normal form.
- If every $\gamma\colon t \Rightarrow u$ is in $\mathcal{P}$, and t is not of the form λt_1, then t is in reduced form.

Proof. By structural induction on t. We proceed by cases:

case 1: $t\colon X \to Y$ and every $\gamma\colon t \Rightarrow u$ is reversible. Then in particular, $\eta_t\colon t \Rightarrow \lambda(t^1\, 1)$ is reversible, and its inverse γ has the form $(\lambda(T^1\, 1); \lambda\gamma_1; \epsilon)$. Then $t = \lambda t_1$. Let $\gamma_1\colon t_1 \Rightarrow u_1$. Then $\lambda\gamma_1\colon t \Rightarrow \lambda u_1$ is reversible, with inverse $(\lambda u_1; \lambda\gamma_2; \epsilon_2)$, say. Then $\lambda\gamma_1; (\lambda u_1; \lambda\gamma_2; \epsilon_2) = (\lambda t_1; \lambda(\gamma_1; \gamma_2); \epsilon_2)$ and γ_1 is reversible. By inductive hypothesis, therefore, t_1 is in long-$\beta\eta$-normal form, and so is t.

case 2: $t\colon B$ and every $\gamma\colon t \Rightarrow u$ is reversible. Then every such γ is in $\mathcal{P}$, and since t cannot be a lambda term, t is in reduced form by the second inductive hypothesis. Therefore t is in long-$\beta\eta$-normal form.

case 3: $t = j$. Then t is in reduced form.

case 4: $t = t_1\, t_2$ and every $\gamma\colon t \Rightarrow u$ is in $\mathcal{P}$. If $t_1 = \lambda t_{11}$, then $\beta\colon t \Rightarrow t_{11}[t_2]$, contradicting lemma 30.

Let $\gamma_1\colon t_1 \Rightarrow u_1$. Then $\gamma_1\, t_2\colon t \Rightarrow u_1\, t_2$ is in $\mathcal{P}$, and by lemma 30 γ_1 is in $\mathcal{P}$. By inductive hypothesis, therefore, t_1 is in reduced form.

Let $\gamma_2\colon t_2 \Rightarrow u_2$. Then $t_1\, \gamma_2\colon t \Rightarrow t_1\, u_2$ is in $\mathcal{P}$, and by lemma 30 γ_2 is reversible. By the first inductive hypothesis, therefore, t_2 is in long-$\beta\eta$-normal form, so t is in reduced form.

□

Putting these two together, we have proved

Corollary 32. A term is in long-$\beta\eta$-normal form iff it is normal.

We need one more property of reversible arrows, related to mellifluence:

Lemma 33. If $\gamma\colon t \Rightarrow u$ and $\zeta_1, \zeta_2\colon u \Rightarrow v$ satisfy $\gamma; \zeta_1 = \gamma; \zeta_2$ where ζ_1, ζ_2 are reversible, then $\zeta_1 = \zeta_2$.

Proof. It is convenient to define a set $\mathcal{Q}$ of rewrites whose normal forms are built up entirely from rewrites in $\mathcal{E}$:

- If $\epsilon : j \Rightarrow t$ is in $\mathcal{E}$ then $j; j; \epsilon \in \mathcal{Q}$
- If $\epsilon : \lambda t \Rightarrow u$ is in $\mathcal{E}$ and $\zeta : s \Rightarrow t$ is in $\mathcal{Q}$, then $\lambda s; \lambda\zeta; \epsilon$ is in $\mathcal{Q}$.
- If $\epsilon : t_1\, t_2 \Rightarrow u$ is in $\mathcal{E}$ and $\zeta_1 : s_1 \Rightarrow t_1$ and $\zeta_2 : s_2 \Rightarrow t_2$ are in $\mathcal{Q}$, then $s_1\, s_2; \zeta_1\, \zeta_2; \epsilon$ is in $\mathcal{Q}$.

By lemma 30, $\mathcal{P} \subseteq \mathcal{Q}$, so every reversible rewrite is in $\mathcal{Q}$. A straightforward induction shows that the composition of two rewrites in $\mathcal{Q}$ is in $\mathcal{Q}$.

Th proof is by induction on $|\zeta_j|_{\mathcal{G}}$. Let $(\alpha; \delta; \epsilon)$ be the canonical form of γ, and $(; \theta_j; \epsilon_j)$ that of ζ_j. There are three cases of θ_1 and θ_2:

case 1: $\theta_1 = \theta_2 = i$. Then $\epsilon = i$ and $\delta = i$ so $\gamma; \zeta_j = (\alpha; i; \epsilon_j)$. Matching canonical forms gives $\epsilon_1 = \epsilon_2$, so $\zeta_1 = \zeta_2$ as required.

case 2: $\theta_j = \lambda\zeta_{j1}$. Proceed by cases of ϵ:

case 2.1: $\epsilon = u$. Then $\delta = \lambda\gamma_1$ and $\gamma; \zeta_j = (\alpha; \lambda(\gamma_1; \zeta_{j1}); \epsilon_j)$, so $\epsilon_1 = \epsilon_2$ and $\gamma_1; \zeta_{11} = \gamma_1; \zeta_{21}$. By inductive hypothesis, $\zeta_{11} = \zeta_{21}$ so $\zeta_1 = \zeta_2$ as required.

case 2.2: $\epsilon = \epsilon_3; \eta_x; \lambda(x^1\, \epsilon_4; \epsilon_5)$. Then $(x^1\, 1; x^1\, \epsilon_4; \epsilon_5); \zeta_{j1}$ is in $\mathcal{Q}$, so equals $(x^1\, 1; \zeta_{j2}^1\, \epsilon_{j1}; \epsilon_{j2})$ for some ζ_{j2} in $\mathcal{Q}$. Therefore $\gamma; \zeta_j = (\alpha; \delta; \epsilon_3); (\zeta_{j2}; \eta_y; \lambda(y^1\, \epsilon_{j1}; \epsilon_{j2}); \epsilon_j)$ and by inductive hypothesis, $\zeta_{12} = \zeta_{22}$, $\epsilon_{11} = \epsilon_{21}$, $\epsilon_{12} = \epsilon_{22}$ and $\epsilon_1 = \epsilon_2$. Therefore $(x^1\, 1; x^1\, \epsilon_4; \epsilon_5); \zeta_{11} = (x^1\, 1; x^1\, \epsilon_4; \epsilon_5); \zeta_{21}$, and by the inductive hypothesis $\zeta_{11} = \zeta_{21}$. So $\zeta_1 = \zeta_2$ as required.

case 3: $\theta_j = \zeta_{j1}\, \zeta_{j2}$. Then $\epsilon = u_1\, u_2$ and $\delta = \gamma_1\, \gamma_2$ so $\gamma; \zeta_j = (\alpha; (\gamma_1; \zeta_{j1})\, (\gamma_2; \zeta_{j2}); \epsilon_j)$. Therefore $\gamma_1; \zeta_{11} = \gamma_1; \zeta_{21}$, $\gamma_2; \zeta_{12} = \gamma_2; \zeta_{22}$ and $\epsilon_1 = \epsilon_2$. By inductive hypothesis, $\zeta_{11} = \zeta_{21}$ and $\zeta_{12} = \zeta_{22}$, so $\zeta_1 = \zeta_2$ as required.

□

6. General results

In this section we use the results we have proved about 2-λ as the basic definitions of a general theory of rewriting. We relate these definitions in such a way as to suggest strongly that they are the correct generalisations of the notions of rewrite system, confluence, normal form and strong normalisation. In this way we provide a framework for the proof theory of rewriting which could be applied to many other systems.

This theory is based on *category theory*: we see the elements (terms, strings, or whatever) of the rewrite system as the objects, and the rewrites as the arrows of a category. The category theory we use is all elementary (unlike that in section 3.2) so the non-categorical reader need only look up a few definitions before continuing. We will write the identity on x as 1_x, and composition in diagramatic order: if $f\colon x \to y$ and $g\colon y \to z$ then $f; g\colon x \to z$, as this seems much more natural than applicative order when talking about rewrite systems.

Definition. A **rewriting category** is a category satisfying the following axiom:

- If $f\colon x \to y$ and $g, h\colon y \to z$ are such that $f; g = f; h$ then there exists $k\colon z \to w$ such that $g; k = h; k$.

This is precisely the property which we called mellifluence in section 5.2.

Our main example of a rewriting category is that constructed from the 2-λ calculus. The objects are the terms (in context) and the arrows are the rewrites in $\mathcal{G}$ (or equivalently, equivalence classes of rewrites). Identities are given by $\mathcal{I}$ and composition by †; the category axioms were proved in section 4 and the rewriting category axiom is proposition 27. The importance of this property is simply that it is needed in most of the proofs which follow.

Definition.

- An object x of a rewriting category is **normal** if every arrow $f\colon x \to y$ is split monic, i.e. there exists $g\colon y \to x$ such that $f; g = 1_x$.
- An object y is **weakly normalising** if there exists $f\colon y \to x$ for some normal x. In this case we call x a **normal form** of y.
- A rewriting category is **weakly normalising** if every object is weakly normalising.

Corollary 32 states that the normal forms of 2-λ are precisely the long-$\beta\eta$-normal forms. Since every term of the simply typed λ-calculus has a long-β-η-normal form, 2-λ is weakly normalising.

Lemma 34.

(1) Any arrow between normal objects is an isomorphism.
(2) If $f\colon x \to y$ then any normal form of y is a normal form of x.
(3) Let x be a normal object in a rewriting category, and $f\colon x \to y$. Then the map $g\colon y \to x$ satisfying $f; g = 1_x$ is unique.

Proof. (1) Let x and y be normal objects, and $f\colon x \to y$. Then because x is normal, there exists $g\colon y \to x$ such that $f; g = 1_x$. Similarly, because y is normal, there exists $h\colon x \to y$ such that $g; h = 1_y$. Now, $f = f; (g; h) = (f; g); h = h$, so it is iso.

(2) If $g\colon y \to z$ with z normal then $f; g\colon x \to z$.

(3) Let $g_1, g_2\colon y \to x$ both satisfy $f; g_j = 1_x$. Then by the rewriting category axiom, there exists $h\colon x \to z$ such that $g_1; h = g_2; h$. But h must be monic because x is normal, so $g_1 = g_2$. □

Definition.

- An object x of a rewriting category is **confluent** if for all pairs $f_1\colon x \to y_1$ and $f_2\colon x \to y_2$ there exist z, $g_1\colon y_1 \to z$ and $g_2\colon y_2 \to z$ such that $f_1; g_1 = f_2; g_2$.
- A rewriting category is **confluent** if every object is confluent.

Proposition 22 states that 2-λ is confluent. Note that confluence and the rewriting category axiom are precisely the conditions for a calculus of fractions [3]. This means that we can calculate the free groupoid on a confluent rewriting category in a particularly simple way. This groupoid can be interpreted as the equational theory generated by the rewrite system.

Lemma 35. Let x be an object in a rewriting category. Then

(1) if x is confluent and $f\colon x \to y$ then y is confluent
(2) if x is normal then x is confluent
(3) if x is confluent and $f\colon x \to y$ then any normal form of x is a normal form of y
(4) if x is confluent then all its normal forms are isomorphic.

Proof. (1) Let $g_1\colon y \to z_1$ and $g_2\colon y \to z_2$. Then $f; g_1\colon x \to z_1$ and $f; g_2\colon x \to z_2$ so by confluence of x there exist $h_1\colon z_1 \to w$ and $h_2\colon z_2 \to w$ such that $f; g_1; h_1 = f; g_2; h_2$. Now by the rewriting category property there exists $k\colon w \to v$ such that $g_1; h_1; k = g_2; h_2; k$ so two arrows which complete the commuting diamond are $h_1; k$ and $h_2; k$.

(2) Let $f_1\colon x \to y_1$ and $f_2\colon x \to y_2$. Since x is normal there exist $g_1\colon y_1 \to x$ and $g_2\colon y_2 \to x$ such that $f_1; g_1 = 1_x = f_2; g_2$. But this shows that x is confluent.

(3) Let $g\colon x \to z$ where z is normal. Since x is confluent there exist $h_1\colon y \to w$ and $h_2\colon z \to w$ such that $f; h_1 = g; h_2$. But z is normal, so there exists $k\colon w \to z$ such that $h_2; k = 1_z$. Now $h_1; k\colon y \to z$ (and $f; h_1; k = g$).

(4) By part (3), if x has two normal forms, then there is an arrow between them. But by lemma 34, this arrow is iso. □

Definition.

- Let D be a diagram in a category. We call a cocone $\mu\colon D \to x$ over D **separating** if for any other cocone $\nu\colon D \to y$ there is *at most one* arrow $f\colon x \to y$ such that $\mu; f = \nu$.
- An object x of a rewriting category is **strongly normalising** if every filtered diagram containing x has a separating cocone.
- A rewriting category is **strongly normalising** if every object is strongly normalising.

Note that if we replace 'at most one' with 'exactly one' in the definition of separating cocone, it becomes the definition of colimiting cocone. The following lemma, together with lemma 33 shows that 2-λ is strongly normalising:

Lemma 36. Let x be an object of a rewriting category.

(1) If x is confluent and weakly normalising, then any filtered diagram containing it has a cocone whose apex is normal.
(2) If whenever $f\colon x \to y$, $g_1, g_2 : y \to z$ are such that $f; g_1 = f; g_2$ then $g_1 = g_2$, then any cocone over a diagram containing x with vertex y is separating.

Proof. (1) Let D be a filtered diagram containing x, and $e\colon x \to v$ for v normal. Define $\mu\colon D \to v$ as follows:

- For each object $y \in D$ there exist z_y, $f_y\colon y \to z_y$ and $g_y\colon x \to z_y$ in D (since D is filtered). By lemma 35, there exists $h_y\colon z_y \to v$ s.t. $g_y; h_y = e$. Then

$$\mu_y = f_y; h_y\colon y \to v \tag{10}$$

- For each arrow $k\colon y \to y'$ in D there exist w, $l\colon z_y \to w$, $l'\colon z_{y'} \to w$ s.t. $g_y; l = g_{y'}; l'$ and $f_y; l = k; f_{y'}; l'$, since D is filtered. Then there exists $m\colon w \to v$ s.t. $g_y; l; m = e$, and by the rewriting category axiom, $h_y = l; m$ and $h_{y'} = l'; m$. Now

$$k; \mu_{y'} = k; f_{y'}; h_{y'} = k; f_{y'}; l'; m = f_y; l; m = f_y; h_y = \mu_y$$

so μ is a cocone.

(2) Straightforward. □

Lemma 37. Let x be an object in a rewriting category. Then

(1) if x is strongly normalising and $f\colon x \to y$ then y is strongly normalising
(2) if x is normal then x is strongly normalising

Proof. (1) Let D be a filtered diagram containing y. Then there is a diagram D' formed by adjoining one new object x and one new arrow $f\colon x \to y$ to D. D' is filtered and contains x, so has a separating cocone, but a separating cocone over D' restricts to one over D.

(2) Immediate from lemma 36 □

Proposition 38. If x is strongly normalising then it is weakly normalising, i.e. it has a normal form.

Proof. Let $\preccurlyeq$ be the partial order on arrows $f\colon x \to y$ induced by $(x \downarrow C)$: so $[f] \preccurlyeq [f']$ iff there exists $g\colon y \to y'$ such that $f; g = f'$. We will prove that every chain in this poset has an upper bound.

Let $[f_j] \preccurlyeq [f_{j+1}]$ be such a chain, and choose $g_j\colon f_j \to f_{j+1}$ in $(x \downarrow C)$. The resulting diagram in C is linear, so filtered, so has a separating cocone. The image of this cocone in the partial order is an upper bound for the chain.

So every chain is bounded and we can apply Zorn's lemma to find a maximal element $[h]$, where $h\colon x \to z$. Now consider the full subcategory of $(x \downarrow C)$ of arrows in the equivalence class $[h]$. This category is filtered because of the rewriting category property and maximality, so its image in C has a separating cocone $\mu\colon [h] \to v$. We will show that v is normal.

Let $f = h; \mu_h\colon x \Rightarrow v$. Now if $g\colon v \Rightarrow u$ then by maximality $f; g \in [h]$ so there exists $g'\colon u \Rightarrow v$ st. $f; g; g' = f$, and by separation, $g; g' = 1$. □

The combination of lemma 35 and proposition 38 means that if x is confluent and strongly normalising then it has a normal form, unique up to isomorphism. However, the proof is unnecessarily complicated and non-constructive, using the axiom of choice. The next result gives a simple construction of the normal form in the confluent case.

Lemma 39. Let x be an object in a rewriting category C, and let $P\colon (x \downarrow C) \to C$ be the usual projection functor. Then

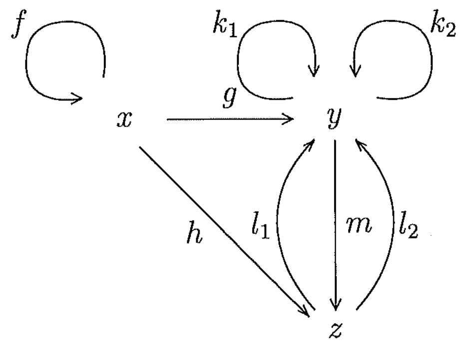
(1) x is confluent iff $(x \downarrow C)$ is filtered.
(2) If $\mu\colon P \to y$ is separating then y is normal.

Proof. (1) The two conditions for filteredness of the slice category are precisely the rewriting category property and confluence of x.

(2) The map $\mu_{1_x}\colon x \to y$ is an object of $(x \downarrow C)$, so $\mu_{\mu_1}\colon y \to y$, and by separation $\mu_{\mu_1} = 1_y$. If $f\colon y \to z$, then $\mu_1; f\colon x \to z$ is an object of $(x \downarrow C)$, so $\mu_{\mu_1;f}\colon z \to y$. Then $f; \mu_{\mu_1;f} = \mu_{\mu_1} = 1$ □

We have proved all the expected relationships between confluence, weak and strong normalisation, and even found a simple condition for confluence + weak normalisation to imply strong normalisation. We now give an example to show that some such condition is necessary.

Let C be the category with three objects x y and z, and eight non-identity arrows:



with composition defined by

$$
\begin{array}{ccc}
f;f=f & f;g=g & f;h=h \\
g;k_1=g & g;k_2=g & g;m=h \\
h;l_1=g & h;l_2=g & \\
k_1;k_1=k_1 & k_1;k_2=k_2 & k_1;m=m \\
k_2;k_1=k_1 & k_2;k_2=k_2 & k_2;m=m \\
l_1;k_1=l_1 & l_1;k_2=l_2 & l_1;m=1_z \\
l_2;k_1=l_1 & l_2;k_2=l_2 & l_2;m=1_z \\
m;l_1=k_1 & m;l_2=k_2 &
\end{array}
$$

Then C is a confluent rewriting category and z is normal, but x is not strongly normalising because none of the three cocones over

$$x \xrightarrow{f} x \xrightarrow{f} x \xrightarrow{f} \cdots$$

is separating.

7. Conclusions

This treatment of the λ-calculus shows that rewriting does, in some cases at least, have a natural proof theory. We have given a syntactic formulation in terms of equations between expressions denoting sequences of rewrites, and shown that this theory has a categorical semantics, characterising β-reduction and η-expansion as the unit and counit of an adjunction, and a tractable word problem. As a corollary to this we have shed light on a well-known problem: the relationship between η-expansion, strong normalisation and long β-η-normal forms.

This technique could usefully be applied to many other systems. The language of rewrites used here would generalise almost immediately to other λ-calculi such as 'system F' [4] and the 'calculus of constructions' [2], and since the proofs given here do not depend on the type structure (except in the definition of long β-η-normal forms) we expect the same results to hold. Other important calculi arise by adding other type constructors, such as those of product, coproduct and Σ types. Here the need for η-expansion is even greater, as η-contraction leads to the lack of confluence. We conjecture that with the definitions given here, these systems can be made confluent and strongly normalising.

Although categorical term rewriting is not as general as conditional term rewriting, its algebraic character makes it much more tractable for mathematical study. Many systems which require a conditional approach, from commutativity to fair nondeterminism, might be described categorically by chosing appropriate equations between rewrites. This is a fertile field for further work.

References

1. H. P. Barendregt. *The Lambda Calculus*, volume 103 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, 1984.
2. T. Coquand and G. Huet. The calculus of constructions. *Information and Computation*, 76(2/3):95–120, 1988.

3. P. Gabriel and M. Zisman. *Calculus of Fractions and Homotopy Theory*, volume 35 of *Ergebnisse der Mathematik*. Springer-Verlag, 1967.
4. Jean-Yves Girard, Yves Lafont, and Paul Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.
5. J. W. Gray. *Formal Category Theory I: Adjointness for 2-categories*, volume 391 of *Springer Lecture Notes in Mathematics*. Springer-Verlag, 1974.
6. G. Huet. *Résolution d'équations dans des langages d'ordre* $1, 2, \ldots, \omega$. PhD thesis, Université de Paris VII, 1976.
7. R. Seely. Modeling computations: A 2-categorical framework. In *Proceedings of the Second Symposium on Logic in Computer Science*, pages 65–71. IEEE Computer Science Press, June 1987.

Computer Laboratory, University of Cambridge, New Museums Site, Pembroke Street, Cambridge CB2 3QG

E-mail address: bph1000@cl.cam.ac.uk