Semantics for a Higher Order Functional Programming Language for Quantum Computation
Benoît Valiron

To cite this version:

HAL Id: tel-00483944
https://tel.archives-ouvertes.fr/tel-00483944
Submitted on 17 May 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Benoît Valiron

Thesis Submitted to the Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

© Benoît Valiron, Ottawa, Canada, 2008

1The Ph.D. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics
Pour mon père.
Abstract

The objective of this thesis is to develop a semantics for higher-order quantum information.

Following the work done in the author’s M.Sc. thesis, we study a lambda calculus for quantum computation with classical control. The language features two important properties. The first one, arising from the so-called no-cloning theorem of quantum computation, is the need for a distinction between duplicable and non-duplicable elements. For keeping track of duplicability at higher-order, we use a type system inspired by the resource-sensitive linear logic. The second important aspect is the probability inherent to measurement, the only operation for retrieving classical data from quantum data. This forces us into choosing a reduction strategy for being able to define an operational semantics.

We address the question of a denotational semantics in two respects. First, we restrict the study to the strictly linear aspect of the language. Doing so, we suppress the need for distinguishing between duplicable and non-duplicable elements and we can focus on the description of quantum features at higher-order. Using the category of completely positive maps as a framework, we build a fully abstract denotational model of the strictly linear fragment of the language.

The study of the full language is more demanding. For dealing with the probabilistic aspect of the language, we use a method inspired by Moggi and build a computational model with a distinction between values and computations. For the distinction between duplicability and non-duplicability in the calculus, we adapt Bierman’s linear category, where the duplication is considered as a comonad with specific properties. The resulting model is what we call a linear category for duplication. Finally, we only focus on the fragment of the language that contains the aforementioned elements, and remove the classical Boolean and quantum Boolean features to get a generic computational linear lambda-calculus. In this idealized setting, we show that such a language have a full and complete interpretation in a linear category for duplication.
Acknowledgements

Numerous people allowed me to write this thesis. In particular, I would like to thank:

- Peter, who supported me for yet another degree;
- Caroline, who came all the way to Halifax;
- my parents, for being supportive all that time;
- Rob, for feeding me with coffee;
- the administrative staff of both Dalhousie and Ottawa University, for being so helpful;
- and finally Gilles, for printing more that 1000 pages of this opus of content.
# Contents

Abstract iii  
Acknowledgements iv  
List of Figures ix  
List of Tables x  
1 Introduction 1  
2 Notions of Category Theory 6  
  2.1 Categories and Functors 6  
  2.2 Natural Transformations 7  
  2.3 Adjoint Functors 8  
  2.4 Products and Coproducts 9  
  2.5 Monads 10  
  2.6 Comonads 11  
  2.7 Symmetric Monoidal Categories and Comonoids 13  
    2.7.1 Monoidal Categories 13  
    2.7.2 Commutative Comonoids 13  
  2.8 Monoidal Comonads and Coalgebras 14  
3 Quantum Computation 17  
  3.1 Mathematical Formalism 17  
    3.1.1 Generalities on Finite Dimensional Hilbert Spaces 17  
    3.1.2 Tensor Products of Hilbert Spaces 19  
    3.1.3 Completely Positive Maps 19  
    3.1.4 Superoperators 20  
    3.1.5 The 2-Dimensional Hilbert Space 21  
  3.2 Quantum Foundations 23  
    3.2.1 Quantum Bits 23  
    3.2.2 Operations 23  
    3.2.3 Mixed States 25  
  3.3 Quantum Effects 26  
    3.3.1 No Cloning 26  
    3.3.2 Entanglement 26  
    3.3.3 Bell’s Inequalities 27  
  3.4 Some Algorithms and Uses of Quantum Effects 28  
    3.4.1 Teleportation 28
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.3.1 Monoidal Structure</td>
<td>129</td>
</tr>
<tr>
<td>9.3.2 Monadic Structure</td>
<td>139</td>
</tr>
<tr>
<td>9.3.3 Comonadic Structure</td>
<td>145</td>
</tr>
<tr>
<td>9.3.4 The Category $\mathcal{C}_\lambda$ is a Linear Category for Duplication</td>
<td>147</td>
</tr>
<tr>
<td>10 Proof of Theorem $\mathcal{O}_2$</td>
<td>148</td>
</tr>
<tr>
<td>10.1 A Handy Tool: Neutral Terms</td>
<td>148</td>
</tr>
<tr>
<td>10.2 Term Rewriting System Number One</td>
<td>152</td>
</tr>
<tr>
<td>10.3 Term Rewriting System Number Two</td>
<td>158</td>
</tr>
<tr>
<td>10.3.1 The Rewriting System</td>
<td>158</td>
</tr>
<tr>
<td>10.3.2 Height of Terms</td>
<td>160</td>
</tr>
<tr>
<td>10.3.3 Degree of Terms</td>
<td>161</td>
</tr>
<tr>
<td>10.3.4 Last Notion of Measure: Number of a Term</td>
<td>163</td>
</tr>
<tr>
<td>10.3.5 Putting Everything Together</td>
<td>164</td>
</tr>
<tr>
<td>10.4 Proof of Theorem $\mathcal{O}_2$</td>
<td>172</td>
</tr>
<tr>
<td>11 Categorical Semantics</td>
<td>174</td>
</tr>
<tr>
<td>11.1 Denotational Semantics</td>
<td>174</td>
</tr>
<tr>
<td>11.1.1 Interpretation of the Type System</td>
<td>174</td>
</tr>
<tr>
<td>11.1.2 Interpretation of the Language</td>
<td>174</td>
</tr>
<tr>
<td>11.2 Soundness of the Denotation</td>
<td>182</td>
</tr>
<tr>
<td>11.3 Completeness</td>
<td>198</td>
</tr>
<tr>
<td>11.4 Towards a Denotational Model</td>
<td>200</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>201</td>
</tr>
<tr>
<td><strong>Index</strong></td>
<td>207</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>The Bloch sphere</td>
<td>22</td>
</tr>
<tr>
<td>3.2</td>
<td>Quantum teleportation protocol</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>Dense coding protocol</td>
<td>30</td>
</tr>
<tr>
<td>4.1</td>
<td>The ORAM model</td>
<td>35</td>
</tr>
</tbody>
</table>
# List of Tables

4.1 Rules for constructing quantum flow-charts ........................................ 38
5.1 Intuitionistic logic: natural deduction rules ........................................ 46
5.2 Interpretation of the simply typed lambda calculus ................................ 48
6.1 Reductions rules of the quantum lambda calculus ................................... 60
6.2 Typing rules ....................................................................................... 62
7.1 Typing rules for the linear quantum lambda calculus ................................ 71
7.2 Reduction rules for the linear quantum lambda calculus ............................ 76
7.3 Denotational semantics for typing judgements ................................. 84
7.4 Axiomatic equivalence ......................................................................... 94
9.1 Typing rules for the linear-non-linear quantum lambda calculus .......... 111
9.2 Axiomatic equivalence axioms: beta-eta-rules ...................................... 124
9.3 Axiomatic equivalence axioms: computation rules ............................... 124
9.4 Axiomatic equivalence: derived rules ................................................. 125
9.5 Definitions of maps and operations on maps in $\mathcal{C}$ ......................... 147
11.1 Interpretation of core values .............................................................. 177
11.2 Interpretation of extended values ...................................................... 178
11.3 Interpretation of computations ......................................................... 179
Chapter 1

Introduction

Quantum computation is a paradigm of computation where information is potentially non-local, can be non-duplicable, and for which one of the core features is inherently probabilistic. However, using quantum computation, numerous hard problems, such as factoring, can be performed efficiently. Higher-order computation brings new light on these algorithms, but raises several challenging problems such as dealing with non-duplicable functions, or marrying duplicability and probabilistic behavior. By taking the problem both at the concrete level of completely positive maps and at the abstract level of categorical structures, this thesis addresses the issue of the semantics of higher-order quantum computation.

Background

Quantum computation. While classical computation uses data encoded on objects governed by the laws of classical physics, quantum computation uses data encoded on the state of particles governed by the laws of quantum physics. The basic unit of data in quantum computation is a quantum bit, which can be modeled by a normalized vector in a 2-dimensional Hilbert space $\mathcal{H}$. Choosing an orthonormal basis $\{|0\rangle, |1\rangle\}$, one can consider $|0\rangle$ and $|1\rangle$ as vectors of information, leading to potential superposition of information. The state of a system of two or more quantum bits is a normalized vector in $\mathcal{H} \otimes \mathcal{H} \otimes \ldots \otimes \mathcal{H}$. The permissible operations on quantum bits are determined by the laws of quantum physics. The only possible operations are initializations, unitary operations and measurements (and combinations thereof). Here, measurement is an operation that sends a quantum bit in state $\alpha|0\rangle + \beta|1\rangle$ to $|0\rangle$ with probability $|\alpha|^2$ and $|1\rangle$ with probability $|\beta|^2$, while simultaneously returning a Boolean 0 or 1, respectively.

Shor [1994] has shown that quantum computers can factor an integer in a time that is polynomial in its number of digits. It is not known whether any classical algorithm can solve the problem in polynomial time. The factoring problem has numerous implications in cryptography. This discovery has focused attention on quantum computing. Quantum computing is also able to bring change in other domains, such as database manipulation, with algorithms to query elements in databases, and such as numerical methods, with the ability to perform Fourier transform efficiently [Nielsen and Chuang, 2002].

From the point of view of quantum programming languages, there are several features of quantum computation that deserve special attention. First, unlike classical data, quantum data cannot in general be duplicated: there is no map $(\alpha|0\rangle + \beta|1\rangle) \mapsto (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle)$. This is known as the "no-cloning theorem" [Wootters and Zurek, 1982]. Second, quantum data can be entangled: Consider the two-quantum bit state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. Although one can speak of “quantum
bit number 1” and “quantum bit number 2”, there is no way of writing the system as $|\phi\rangle \otimes |\psi\rangle$. Finally, measurement, which is the only operation for converting quantum data to classical data, is inherently a probabilistic operation.

**Semantics of programming languages.** When dealing with a programming language, an important question is how to describe the behavior of programs, and to characterize the set of the possible behaviors. A semantics (Pierce, 2002) for a programming language serves as a description of properties that valid programs satisfy. This permits to get insights into the capabilities of the language and the logic underlying it. This can help for example in understanding the power and the limitations of the computational paradigm considered, or more pragmatically to build automated tools to check for the validity of a piece of code, along a set of specifications.

An operational semantics is a description of how valid programs behave by sequences of computational steps. Depending on the programming language considered, this can take the form of a rewriting system of the code itself (in the case, for example, of the simply-typed lambda calculus), or it can be more involved and require the description of an abstract machine (as, for example, in the case of the Turing machine).

An axiomatic semantics is a description of local syntactic equivalences of code. The goal is to provide an equational logic of programs, in order to exhibit general syntactic properties of the language. It is helpful to describe a system of normal forms in which every program can be rewritten. An important property of an axiomatic semantics is soundness. We say that an axiomatic semantics is sound if two axiomatically equivalent programs are operationally equivalent.

A denotational semantics is built by drawing a parallel between the programs in the language and a mathematically structured set. The properties of the language are understood as mathematical and logical properties, giving a handle to answer questions about the language. A very strong property for a denotational semantics is full abstraction: the semantics is fully abstract when any two programs have the same denotation precisely when they are operationally equivalent in every context. When the interpretation is also onto, the denotation is fully complete.

Sometimes, at first it is only possible to describe the axiomatic structure of the language. In this situation, by studying the relations between pieces of code it can be possible to describe the categorical structure underlying the logic and infer a categorical semantics.

**Higher-order computation and lambda calculus.** An essential paradigm in computation is higher-order computation. A higher-order function is a function that inputs or outputs a “black box”, which is itself a function. The canonical formalism for expressing higher-order functions is the lambda calculus developed by Church (1936) and Kleene (1935). This endeavour was shown (Turing, 1936) to have the same computational power as Turing machines.

The lambda calculus not only permits the design of all computable functions, but it also allows one to draw a correspondence between proofs in logic and programs. This correspondence, known as the Curry-Howard isomorphism (Girard et al., 1990), is the one of the main tools used for passing from an axiomatic semantics to a categorical semantics (Lambek and Scott, 1986). It is thus a powerful handle to grasp the structure behind potentially intricate computational schemes.

**Pure and impure lambda calculus.** In its most general formulation, the lambda calculus is called pure, or purely functional. This means that programs are completely determined by their output result on each possible input. It is fairly restrictive and unfortunately rules out many interesting side effects, such as non-determinism, probabilistic outcomes, or input-outputs. To give a semantics to languages with side-effects, Moggi (1989, 1991) defines a lambda calculus, the computational lambda calculus, where two classes of terms are considered: the values and the computations. Using
an axiomatic equivalence and some categorical tools, he describes a general model of computations with side effects.

Semantics of Quantum Programming Languages

Besides a few existing algorithms, quantum computation is a relatively fresh area of research with no single best way of thinking about the processes in play. Understanding the semantics of quantum programming languages is therefore a useful tool in understanding the capabilities of quantum computation and for deciding on the most appropriate form to use as a computational scheme. Being in the intersection of physics, computer science and mathematics, several paradigms to process quantum computation exist, and a wide range of approaches are studied (Gay, 2006).

Some studies focus primarily on the unitary evolution of the quantum states, and view measurement as a meta-operation outside of the formalism. Benioff (1982) and Deutsch (1985) built a quantum Turing machine, where the tape, the head and the states are encoded as quantum information, Van Tonder (2003, 2004) described a lambda calculus encoded on strings of quantum bits, Altenkirch and Grattage (2005a,b) studied a first order functional language with a notion of quantum test compiling into a quantum circuit.

Other approaches consider measurements as the driving force in the computation. This is the one-way model of Raussendorf and Briegel (2001), further extended with an axiomatic and a denotational semantics in Danos et al. (2007).

It is also possible to see measurements as an active part in the computation and to allow interleaved unitary operations and measurements. One example is the QRAM model of Knill (1996) and Bettelli et al. (2003). Here, a quantum computer consists of a classical computer with a quantum device attached to it. In this configuration, called “classical control” (Selinger, 2004a), the operation of the machine is controlled by a classical program which emits a sequence of instructions to the quantum device for performing measurements and unitary operations. Several programming languages have been proposed to deal with such a model, see for example the works of Sanders and Zuliani (2000), Bettelli et al. (2003, 2004), Selinger (2004a).

The denotational aspect of quantum computation has been studied in various ways. A first trend of research, started by Abramsky and Coecke (2004), focuses on the basic structures required for performing quantum computation, using a categorical framework. This work gave birth to a range of publications, such as Coecke (2004), Coecke and Pavlovic (2007), Selinger (2004a,b) and Coecke and Paquette (2006). Other approaches focus on the description of quantum computation in terms of superoperators. Selinger (2004a) defines a flow-chart language and provides a fully complete semantics for the language in terms of superoperators. Selinger (2004a) is an attempt to generalize this semantics at higher-order, by looking for a suitable category of normed vector spaces. However, none of the studied descriptions capture the required structure. Another related work is the one of Girard (2004), describing a model of linear logic in terms of normed vector spaces. However, although being a *-autonomous category, it does not yield the correct answer at base types, as pointed out by Selinger (2004b).

A Lambda Calculus for Quantum Computation

Selinger and Valiron (2006a, 2005) and Valiron (2004a,b) developed a higher-order language for quantum computation. The language is a typed lambda calculus together with classical and quantum data, featuring creations of quantum bits, measurements and unitary operations. For example, in such a language one can build a function that inputs a Boolean and returns a function from quantum bits to quantum bits based on the input bit.
The first important feature built into the language is the distinction between duplicable and non-duplicable elements. This arises from the no-cloning property of quantum computation. So if \( x : qbit \) is a variable representing a quantum bit, and \( y : bit \) is a variable representing a classical bit, then it is legal to write \( f(y, y) \), but not \( g(x, x) \). In order to keep track of duplicability at higher-order types we use a type system based on linear logic (Girard, 1987). We use the duplicability operator “!” to mark classical types.

The second feature of quantum computation included in the language is the probabilistic nature of the measurement. This was taken care of by defining an abstract machine together with an operational semantics in the form of a call-by-value reduction strategy.

The most significant results of (Valiron, 2004a) were the proof that the type system prevents a typed program from ending up in an error state, and the description a type inference algorithm for deciding whether a given program is typeable or not.

Semantics for Higher-Order Quantum Computation

We now turn to the question of a semantics for the quantum lambda calculus of (Valiron, 2004a).

As explained before, we use the duplicability operator “!” to mark classical types. In the categorical semantics, this operator gives rise to a comonad as in the work of Seely (1989) and Benton et al. (1992). Another account of mixing copyable and non-copyable data is Coecke and Pavlovic (2007), where the copyability is internal to objects. However, it is not clear whether this approach scales well to higher order.

To model the probabilistic effect of the measurement operator in our call-by-value setting, our semantics requires a computational monad in the sense of Moggi (1991). The coexistence of the computational monad and the duplicability comonad in the same category is what makes our semantics interesting and novel. It differs from the work of Benton and Wadler (1996), who considered a monad and a comonad one two different categories, arising from a single adjunction.

The computational aspects of linear logic have been extensively explored by many authors, including Abramsky (1993), Benton et al. (1992, 1993), Bierman (1993), and Wadler (1992). However, these works contain explicit lambda terms to witness the structural rules of linear logic, for example, \( x : !A \triangleright \text{derelict}(x) : A \). By contrast, in our language, structural rules are implicit at the term level, so that \( !A \) is regarded as a subtype of \( A \) and one writes \( x : !A \triangleright x : A \). As it was shown in Selinger and Valiron (2006a), linearity information can automatically be inferred by the type checker. This allows the programmer to program as in a regular non-linear language.

This use of subtyping is the main technical complication in our proof of well-definedness of the semantics. This is because one has to show that the denotation is independent of the choice of a potentially large number of possible derivations of a given typing judgement. We are forced to introduce a Church-style typing system, and to prove that the semantics finally does not depend on the additional type annotations.

Another technical choice we made in our language concerns the relation between the exponential “!” and the pairing operation. Linear logic only requires \( !A \otimes !B \triangleright !(A \otimes B) \) and not the opposite implication. However, in our programming language setting, we find it natural to identify a classical pair of values with a pair of classical values, and therefore we will have an isomorphism \( !A \otimes !B \cong !(A \otimes B) \).

Plan of the Thesis

The content of the thesis is summarized as follows.

In Chapter 2, we state the notions of category theory used in the thesis. In Chapter 3, we give a brief overview of quantum computation, and we describe in Chapter 4 the works in the
literature concerned with the semantics of quantum programming languages. In particular, we detail in Section 4.3 the flow-chart language of Selinger (2004b) and its semantics in terms of superoperators, used in Chapter 7.

The last chapter on general background is Chapter 5, giving a brief introduction on lambda calculus and its categorical interpretation. We review the simply-typed lambda calculus, its computational extension for dealing with side-effects, and the computational interpretation of intuitionistic linear logic.

Chapter 6 summarizes the results contained in Valiron (2004a). The quantum lambda calculus that will be referred to all along the remainder of the thesis is defined here, together with its type system and its operational semantics.

The next chapters are the main contribution of this thesis.

Chapter 7 contains a precise study of the linear fragment of the quantum lambda calculus. This fragment is typed, its operational semantics is described in the same way as in Chapter 6, and a denotational semantics in terms of completely positive maps is provided. Using the fact that the semantics is full at first order, the semantics is proved to be fully abstract.

Chapter 8 raises the issue of the structure of the full language. Using Moggi’s computational model and Bierman’s linear categories, a categorical structure candidate for being a model of the language, called linear category for duplication, is provided.

Chapter 9 describes a version of the quantum lambda calculus stripped from the classical and the quantum Boolean structure. The language is reduced to its core features, namely the distinction between linearity and non-linearity, the higher-order, the pairing and the computational aspect. An axiomatic semantics is described, yielding a syntactic category that is proved to be a linear category for duplication. The result requires Theorem 9.2.7, whose proof requires some technical machinery. Chapter 10 explains this machinery: We introduce two rewriting systems. They converge to a special class of terms, called the neutral terms, for which the theorem is trivial. The main difficulty is to prove a weak-normalization result for these rewriting systems. We use techniques from Girard et al. (1990, Ch. 4) by defining a well-founded measure of convergence, decreasing with each reduction of the rewriting system.

Chapter 11 develops the interpretation of the computational lambda calculus for duplication into a general linear category for duplication, and the proof that the interpretation is sound and complete.
Chapter 2

Notions of Category Theory

In this chapter we give a brief exposition of the definitions and results of category theory used in the remainder of the thesis. For a complete introduction to category theory, [Mac Lane (1998)] is a solid reference. An approach to category theory oriented towards semantic aspects of computation can be found in [Lambek and Scott (1989)].

2.1 Categories and Functors

The content of this section follows [Lambek and Scott (1989)].

Definition 2.1.1. A category \(C\) consists of the following three elements.

- A class of objects, denoted by \(|C|\).
- A class of morphisms. To each morphism \(f\) is associated a domain \(A\) and a codomain \(B\), where \(A, B\) are objects. We say \(f\) is a morphism (or an arrow, or a map) from \(A\) to \(B\) and we write \(f : A \to B\). For any given \(A\) and \(B\), we define the homset of \(A\) to \(B\) as the class of morphisms from \(A\) to \(B\), and we denote it \(C(A, B)\), \(\text{hom}_C(A, B)\), or \(\text{hom}(A, B)\) if there is no ambiguity. For each \(A\) there is a special morphism \(\text{id}_A : A \to A\), called the identity.
- A binary operator \(\circ\) called composition such that if \(f : A \to B\) and \(g : B \to C\) are morphisms, then \(g \circ f : A \to C\) is a morphism. We will also write \(g \circ f\) as \(f;g\).

The composition satisfies two properties: if \(f : A \to B\), \(g : B \to C\) and \(h : C \to D\) are morphisms, then
\[
(f;g);h = f;(g;h) \quad \text{and} \quad f; id_B = f = id_A; f.
\]
We say that the category is locally small if for all objects \(A, B\) the class \(\text{hom}(A, B)\) is a set. We say it is small if moreover the class of objects is a set.

Definition 2.1.2. We say that a morphism \(f : A \to B\) in a category \(C\) is an isomorphism if there exists a morphism \(f^{-1} : B \to A\) such that \(f ; g = \text{id}_A\) and \(g ; f = \text{id}_B\).

Example 2.1.3. The trivial category \(1\). Its class of objects is the singleton \(\{\ast\}\) and its unique arrow is the identity on \(\ast\). Composition is defined trivially.

Example 2.1.4. The category \(\text{Set}\) of sets. Its objects are arbitrary sets and its morphisms arbitrary mappings between sets.
Example 2.1.5. Given a category $\mathcal{C}$, the opposite category $\mathcal{C}^{\text{op}}$ consisting of the objects of $\mathcal{C}$ and the arrows of $\mathcal{C}$ in “opposite” direction: $f$ is an arrow from $B$ to $A$ in $\mathcal{C}^{\text{op}}$ if and only if $f : A \to B$ is an arrow of $\mathcal{C}$. We write $f^{\text{op}}$ for the arrow in $\mathcal{C}^{\text{op}}$ corresponding to $f$ in $\mathcal{C}$.

Example 2.1.6. Given two categories $\mathcal{C}$ and $\mathcal{D}$ one defines a category $\mathcal{C} \times \mathcal{D}$ with the following data.

- Objects are pairs $(A, B)$ where $A \in |\mathcal{C}|$ and $B \in |\mathcal{D}|$.
- An arrow $(A, B) \to (C, D)$ is a pair of arrows $(f : A \to C, g : B \to D)$, where $f \in \mathcal{C}$ and $g \in \mathcal{D}$.
- The identity arrow $id_{A,B}$ is the pair $(id_A, id_B)$.
- The composition of two maps $(f, f') : (A, A') \to (B, B')$ and $(g, g') : (B, B') \to (C, C')$ is the map $(f; g, f'; g') : (A, A') \to (C, C')$.

Definition 2.2.1. Given two categories $\mathcal{C}$ and $\mathcal{D}$, a functor $F$ from $\mathcal{C}$ to $\mathcal{D}$ is a map from $|\mathcal{C}|$ to $|\mathcal{D}|$ and from $\text{hom}_\mathcal{C}(A, B)$ to $\text{hom}_\mathcal{D}(FA, FB)$ for all $A, B$ in $\mathcal{C}$ preserving identities and composition:

$$F(id_A) = id_{FA} \quad \text{and} \quad F(f; g) = Ff; Fg.$$ 

A functor $F : \mathcal{C} \to \mathcal{D}$ is full if for all objects $A$ and $B$ of $\mathcal{C}$ the map

$$F_{A,B} : \text{hom}_\mathcal{C}(A, B) \to \text{hom}_\mathcal{D}(FA, FB)$$

is surjective. The functor $F$ is faithful if the maps $F_{A,B}$ are injective, and fully faithful if they are bijective. A full embedding is a fully faithful functor that is also injective on objects.

Example 2.1.7. The diagonal functor on $\mathcal{C}$ is the functor $\triangle_\mathcal{C} : \mathcal{C} \to \mathcal{C} \times \mathcal{C}$ sending $A$ to $(A, A)$ and $f$ to $(f, f)$. It is faithful but not in general full.

Example 2.1.8. The terminal functor on $\mathcal{C}$ is the functor $\bigcirc_\mathcal{C} : \mathcal{C} \to \mathbf{1}$ sending every object on $\star$ and every morphism on $id_\star$. It is full but not in general faithful.

Example 2.1.9. Consider three categories $\mathcal{C}, \mathcal{D}$ and $\mathcal{E}$.

1. Given a functor $F : \mathcal{C} \to \mathcal{D}$ and a functor $G : \mathcal{D} \to \mathcal{E}$, the map $GF : \mathcal{C} \to \mathcal{E}$ defined as the composition of $F$ and $G$ is also a functor.

2. The map $id_\mathcal{C} : \mathcal{C} \to \mathcal{C}$ sending an object $A$ to itself and a map $f : A \to B$ to itself is the identity functor. \hfill \square

Definition 2.1.10. One defines the category $\text{Cat}$ of small categories as the category where objects are small categories and morphisms are functors of small categories. The identity morphism on $\mathcal{C}$ is the identity functor $id_\mathcal{C}$ and the composition of arrows is the composition of functors.

2.2 Natural Transformations

The content of this section follows (Mac Lane 1998).

Definition 2.2.1. Given two functors $F, G : \mathcal{C} \to \mathcal{D}$, a natural transformation $n$ from $F$ to $G$, denoted as $n : F \to G$, is a collection of maps $n_A : FA \to GA$ in $\mathcal{D}$ for each object $A$ of $\mathcal{C}$ such that for all morphism $f : A \to B$ in $\mathcal{C}$ the following diagram commutes:

$$\begin{array}{ccc}
FA & \xrightarrow{n_A} & GA \\
Ff & \downarrow & \Downarrow & \downarrow & \Downarrow & \downarrow \\
FB & \xrightarrow{n_B} & GB.
\end{array}$$
2.3. Adjoint Functors

We say that a diagram commutes when for every two vertices $X, Y$ in the diagram, all the paths from $X$ to $Y$ (following arrows) yield equal morphisms. In this particular case, this means $n_A; Gf = Ff; n_B$. The diagrammatic notation for a natural transformation $n : F \rightarrow G$ is

$$
\begin{array}{c}
\text{C} \\
\downarrow \psi \\
\text{D}
\end{array}
\xymatrix{
F \ar[rr]^n & & G
}
$$

(2.2.1)

We say that a natural transformation $n$ is a natural isomorphism if for all objects $A$ the map $n_A$ is an isomorphism.

**Definition 2.2.2.** Given three functors $F, G, H : C \rightarrow D$ and two natural transformations $m : F \rightarrow G$ and $n : G \rightarrow H$, one defines the composition $m; n : F \rightarrow H$ as the collection of maps $m_A; n_A : FA \rightarrow HA$. One defines the identity on $F$ as the collection of arrows $id_{FA} : FA \rightarrow FA$.

**Definition 2.2.3.** Consider two categories $C$ and $D$. One defines the functor category $\text{Func}(C, D)$, or $D^C$, as the category whose objects are functors $C \rightarrow D$ and arrows natural transformations $F \rightarrow G$, where $F, G : C \rightarrow D$ are functors.

**Definition 2.2.4.** Consider four categories, four functors and one natural transformation in the following situation:

$$
\begin{array}{c}
\text{A} \\
\downarrow F \\
\text{B} \\
\downarrow G \\
\text{C} \\
\downarrow H \\
\text{D}
\end{array}
\xymatrix{
A \\
\downarrow F \\
B \\
\downarrow G \\
C \\
\downarrow H \\
D
}
$$

One defines the two natural transformations $nF : GF \rightarrow HF$, $Kn : KG \rightarrow KH$ as $(nF)_A = n_{FA}$ and $(Kn)_A = K(n_A)$.

**Lemma 2.2.5.** Consider the situation

$$
\begin{array}{c}
\text{A} \\
\downarrow F \\
\text{B} \\
\downarrow G \\
\text{C} \\
\downarrow H \\
\text{D}
\end{array}
\xymatrix{
A \\
\downarrow F \\
B \\
\downarrow G \\
C \\
\downarrow H \\
D
}
$$

Then $K(nF) = (Kn)F$, $(Kn; m) = (Kn; Km)$ and $(n; m)L = (nL; mL)$.

**Definition 2.2.6** (Mac Lane, 1998, p. 93). Let $C$ and $D$ be two categories. We say that $C$ and $D$ are equivalent if there exists a pair of functors $F : C \rightarrow D$ and $G : D \rightarrow C$ together with natural isomorphisms $id_C \cong G \circ F$ and $id_D \cong F \circ G$. We say that $F$ and $G$ are equivalences of category.

2.3 Adjoint Functors

The following definitions and lemmas are taken directly from [Lambek and Scott, 1989, pp. 13–15]

**Definition 2.3.1.** An adjunction between categories $C$ and $D$ is given by a quadruple $(F, U, \eta, \epsilon)$, where $F : C \rightarrow D$ and $U : D \rightarrow C$ are functors and $\eta : id_C \rightarrow UF$ and $\epsilon : FU \rightarrow id_D$ are natural transformations such that $(\eta U); (U \epsilon) = id_U$ and $(F \eta); (\epsilon F) = id_F$. One says that $U$ is right adjoint to $F$, or that $F$ is left adjoint to $U$, and one calls $\eta$ the unit and $\epsilon$ the counit of the adjunction.
2.4. Products and Coproducts

Lemma 2.3.2. An adjunction \((F, U, \eta, \epsilon)\) between locally small categories \(C\) and \(D\) gives rise to and is determined by a natural isomorphism \(\tau\)

\[
\tau_{A,B} : D(FA, B) \cong C(A, UB).
\]

between the two functors \(D(F-, -)\) and \(C(-, U-)\) of type \(C^{op} \times D \to \text{Set}\). □

Lemma 2.3.3 (Uniqueness). Adjoint functors determine each other uniquely up to natural isomorphism. □

Lemma 2.3.4 (Composition). Given the two following adjunctions:

\[ (F : C \to D, \ U : D \to C, \ \eta : id_C \to UF, \ \epsilon : FU \to id_D), \]

\[ (F' : D \to E, \ U' : E \to D, \ \eta' : id_D \to U'F', \ \epsilon' : F'U' \to id_E), \]

the quadruple

\[ (F'F : C \to E, \ UU' : E \to C, \ \eta;U\eta'F : id_C \to UU'F'F, \ F'\epsilon U';\epsilon' : F'FUU' \to id_E) \]

is an adjunction between \(C\) and \(E\). □

Definition 2.3.5. We say that the adjunction \((F', UU')\) of Lemma 2.3.4 is the composition of \((F, U)\) and \((F', U')\).

Remark 2.3.6. The new structure is an adjunction because the two equations of Definition 2.3.1, namely

\[
((\eta;U\eta'F)(UU'));((UU')(F'\epsilon U';\epsilon')) = id_{UU'}, \quad (2.3.1)
\]

\[
((F'F)(\eta;U\eta'F));((F'\epsilon U';\epsilon')(F'F)) = id_{F'F}, \quad (2.3.2)
\]

are satisfied.

2.4 Products and Coproducts

Definition 2.4.1 [Mac Lane, 1998]. An object \(T\) in a category \(C\) is called a terminal object if for each object \(A\) there exists a unique map \(\square_A : A \to T\).

Definition 2.4.2. An object \(\bot\) in a category \(C\) is called an initial object if for each object \(A\) there exists a unique map \(\Box_A : \bot \to A\).

Lemma 2.4.3 [Lambek and Scott, 1989]. The category \(C\) has a terminal (respectively initial) object if and only if the functor \(\square_C : C \to 1\) has a right (respectively left) adjoint.

Definition 2.4.4 [Lambek and Scott, 1989]. Given a category \(C\) and two objects \(A\) and \(B\), the product of \(A\) and \(B\) is, if it exists, the data consisting of an object \(A \times B\) and two maps \(\pi^1_{A,B} : A \times B \to A\) and \(\pi^2_{A,B} : A \times B \to B\), such that for all maps \(f : C \to A\) and \(g : C \to B\) there exists a unique map \((f, g) : C \to A \times B\) with the following equations holding for all \(h : C \to A \times B\):

\[
\langle f, g \rangle; \pi^1_{A,B} = f \quad (2.4.1)
\]

\[
\langle f, g \rangle; \pi^2_{A,B} = g \quad (2.4.2)
\]

\[
\langle h; \pi^1_{A,B}, h; \pi^2_{A,B} \rangle = h \quad (2.4.3)
\]

We say that the category \(C\) has binary products if there is a product for all \(A\) and \(B\).
2.5 Monads

Definition 2.4.5 [Mac Lane, 1998]. Given a category $\mathcal{C}$ and two objects $A$ and $B$ in $\mathcal{C}$, the coproduct of $A$ and $B$ is, if it exists, the data consisting of an object $A + B$ and two maps $i^\gamma_{A,B} : A \to A + B$ and $i^\delta_{A,B} : B \to A + B$ such that for all maps $f : A \to C$ and $g : B \to C$ there exists a unique map $[f,g] : A + B \to C$ with the following equations holding for all $h : A + B \to C$:

\begin{align}
  i^\gamma_{A,B} [f,g] &= f \\
  i^\delta_{A,B} [f,g] &= g \\
  [i^\gamma_{A,B}; h, i^\delta_{A,B}; h] &= h
\end{align}

We say that the category $\mathcal{C}$ has binary coproducts if there is a coproduct for all $A$ and $B$.

Lemma 2.4.6. Binary products (respectively binary coproducts) in a category $\mathcal{C}$ induce a functor $\times$ (respectively $+$): $\mathcal{C} \times \mathcal{C} \to \mathcal{C}$.

Lemma 2.4.7 [Lambek and Scott, 1989]. The category $\mathcal{C}$ has binary products (respectively binary coproducts) if and only if the functor $\Delta_C : \mathcal{C} \to \mathcal{C} \times \mathcal{C}$ has a right (respectively left) adjoint.

Definition 2.4.8 [Lambek and Scott, 1989]. We say that a category $\mathcal{C}$ is cartesian if it has a terminal object $T$ and binary products $\times$. The category is cartesian closed if for all $A$ the functor $A \times (-)$ has a right adjoint $A \Rightarrow (-)$.

2.5 Monads

Definition 2.5.1 [Mac Lane, 1998]. A monad over a category $\mathcal{C}$ is a triple $(T, \eta, \mu)$ where $T : \mathcal{C} \to \mathcal{C}$ is a functor, $\eta : id \Rightarrow T$ and $\mu : T^2 \Rightarrow T$ are natural transformations and the following diagrams commute:

\begin{center}
\begin{tikzcd}
T^3 A \arrow{r}{T^2 \mu_A} \arrow{d}{\mu_T}
\arrow{rd}[swap]{\mu_A}
\end{tikzcd}
\quad
\begin{tikzcd}
T^2 A \arrow{r}{T \mu_A} \arrow{d}{\mu_T}
\arrow{rd}[swap]{\mu_A}
\end{tikzcd}
\quad
\begin{tikzcd}
TA \arrow{r}{\eta_T A} \arrow{d}{\mu_T}
\arrow{rd}[swap]{\mu_A}
\end{tikzcd}
\quad
\begin{tikzcd}
TA \arrow{r}{\eta_T A} \arrow{d}{\mu_T}
\arrow{rd}[swap]{\mu_A}
\end{tikzcd}
\end{center}

The natural transformation $\mu$ is called the multiplication of the monad and $\eta$ the unit of the monad.

Definition 2.5.2 [Manes, 1976]. A Kleisli triple over a category $\mathcal{C}$ is a triple $(T, \eta, -^*)$ where $T : |\mathcal{C}| \to |\mathcal{C}|$, $\eta_A : A \to TA$ for all $A$ and $f^* : TA \to TB$ for all $f : A \to TB$, and the following equations hold:

\begin{align}
  \eta_A^* &= id_{TA}, \\
  \eta_A^* f^* &= f \text{ for } f : A \to TB, \\
  f^* g^* &= (f; g^*)^* \text{ for } f : A \to TB \text{ and } g : B \to TC.
\end{align}

Lemma 2.5.3 [Manes, 1976]. There is a one-to-one correspondence between Kleisli triples and monads.

Proof. Given a Kleisli triple $(T, \eta, -^*)$, we construct a monad $(T, \eta, \mu)$ by extending $T$ to arrows: $T(f) = (f; \eta_B)^*$ for $f : A \to B$, and by defining $\mu_A = id_{TA}^*$. Conversely, given a monad $(T, \eta, \mu)$, $-^*$ is defined by $f^* = (Tf); \mu_B$ for $f : A \to TB$. $\square$

Definition 2.5.4 [Manes, 1976]. Given a Kleisli triple $(T, \eta, -^*)$ over $\mathcal{C}$, the Kleisli category $\mathcal{C}_T$ is defined as follows:

- the objects of $\mathcal{C}_T$ are the objects of $\mathcal{C}$,
2.6. Comonads

- the set $C_T(A, B)$ of morphisms from $A$ to $B$ in $C_T$ is $C(A, TB)$,
- the identity on $A$ is $\eta_A : A \rightarrow TA$,
- The composition of $f \in C_T(A, B)$ and $g \in C_T(B, C)$ in $C_T$ is $f ; g^* : A \rightarrow TC$.

**Lemma 2.5.5.** Consider a monad $(T, \mu, \eta)$ over a category $C$. There is an adjunction $C F \rightarrow C T: U \downarrow \downarrow \perp$ arising from the monad, where $F(A) = A$, $U(A) = TA$, and $F$ and $U$ perform the following operations on morphisms:

\[
F : (A \xrightarrow{f} B) \mapsto (A \xrightarrow{f} B \xrightarrow{\eta_B} TB), \\
U : (A \xrightarrow{g} TB) \mapsto (TA \xrightarrow{g^*} TB).
\]

The unit $\eta$ of the adjunction is the unit of the monad, and the counit $\epsilon$ is the map $TA \xrightarrow{id_T A} TA$ in the category $C$.

**Lemma 2.5.6.** Conversely, given an adjunction $(F, U, \eta, \epsilon)$, there is a canonical monad arising from it, by setting $T = UF$. The unit of the monad is the unit of the adjunction, and the multiplication $\mu$ of the monad is $U\epsilon F : UFUF \rightarrow UF$.

### 2.6 Comonads

A dual notion to monad is the notion of comonad. We are going to use it extensively in the rest of the thesis.

**Definition 2.6.1** (Mac Lane, 1998). A comonad in a category $C$ is a tuple $(L, \epsilon, \delta)$ where $L : C \rightarrow C$ is a functor and $\epsilon : L \rightarrow I$ and $\delta : L \rightarrow L^2$ are natural transformations which render commutative the diagrams

\[
\begin{align*}
LA & \xrightarrow{\delta_A} L^2A \\
L^2A & \xrightarrow{\delta_{LA}} L^3A,
\end{align*}
\]

\[
\begin{align*}
LA & \xrightarrow{\epsilon_{LA}} L^2A \\
L^2A & \xrightarrow{\epsilon_{LA}} LA.
\end{align*}
\]

By analogy with monads, we define the map $(-)_*$:

\[
f : LA \rightarrow B \\
f_* : LA \xrightarrow{\delta_A} L^2A \xrightarrow{Lf} LB
\]

**Definition 2.6.2.** Given a comonad $(L, \epsilon, \delta)$ over a category $C$, the co-Kleisli category $C_L$ is defined as follows:

- the objects of $C_L$ are the objects of $C$,
- the set $C_L(A, B)$ of morphisms from $A$ to $B$ in $C_L$ is $C(LA, B)$,
2.6. Comonads

- the identity on \( A \) is \( \epsilon_A : LA \to A \),
- The composition of \( f \in C_L(A, B) \) and \( g \in C_L(B, C) \) in \( C_L \) is \( f \circ g : LA \to C \).

**Definition 2.6.3.** Given a comonad \((L, \delta, \epsilon)\) on a category \( C \), an \( L \)-coalgebra is a pair \((A, h_A : A \to LA)\) where \( A \) is an object and \( h_A \) is an arrow of \( C \), such that both

\[
\begin{align*}
A & \xrightarrow{h_A} LA, \\
LA & \xrightarrow{Lh_A} L^2A,
\end{align*}
\]

commute. A morphism \( f : (A, h_A) \to (B, h_B) \) of \( L \)-coalgebras is an arrow \( f : A \to B \) of \( C \) which renders commutative the diagram

\[
\begin{aligned}
A & \xrightarrow{f} B \\
LA & \xrightarrow{Lf} LB.
\end{aligned}
\]

The category \( C^L \) of \( L \)-coalgebras of \( C \) is the category whose objects are \( L \)-coalgebras and whose maps are morphisms of \( L \)-coalgebras. It is called the co-Eilenberg-Moore category of the comonad \( L \).

**Lemma 2.6.4.** Given a comonad \((L, \delta, \epsilon)\) on a category \( C \), for each object \( A \) in \( C \) the pair \((LA, \delta_A : LA \to L^2A)\) forms a coalgebra. Given any map \( f : A \to B \), the map \( Lf : (LA, \delta_A) \to (LB, \delta_B) \) is a coalgebra map. \( \square \)

**Lemma 2.6.5.** Given a category \( C \) and a comonad \((L, \delta, \epsilon)\) on it, the coalgebra \((LA, \delta_A)\) has the following universal property: If \( (X, h_X) \) is any \( L \)-coalgebra, and if \( f : X \to A \) is any map in \( C \), there exists a unique coalgebra map \( f^\#: (X, h_X) \to (LA, \delta_A) \) such that

\[
\begin{aligned}
LA & \xrightarrow{\epsilon_A} A, \\
X & \xrightarrow{f} A.
\end{aligned}
\]

Moreover, \( f^\# = h_X \circ Lf \). \( \square \)

**Corollary 2.6.6.** Given a category \( C \) and a comonad \((L, \delta, \epsilon)\), there is a one-to-one correspondence between coalgebra morphisms \( g : (LA, \delta_A) \to (LB, \delta_B) \) and maps of \( C \) of the form \( f : LA \to B \). \( \square \)

**Definition 2.6.7.** The coalgebra \((LA, \delta_A)\) in Lemma 2.6.4 is called a (co)free coalgebra.

**Corollary 2.6.8.** There is a full embedding of \( C_L \) into \( C^L \) sending the object \( A \) of \( C_L \) to \((LA, \delta_A)\) and sending the morphism \( f : LA \to B \) into \( f_* : LA \to LB \). \( \square \)

**Lemma 2.6.9.** Given an adjunction \((F, U, \eta, \epsilon)\) there is a canonical comonad arising from it, by setting \( T = FU \), the counit of the comonad is the counit of the adjunction, and the comultiplication \( \delta \) of the comonad is \( F\eta U : FUFU \to FU \). \( \square \)
2.7 Symmetric Monoidal Categories and Comonoids

2.7.1 Monoidal Categories

Definition 2.7.1 [Mac Lane, 1998]. A monoidal category \( C \) is a tuple \((C, \otimes, \top, \alpha, \lambda, \rho)\) where \( C \) is a category, \( \otimes: C \times C \to C \) is a functor, \( \top \) is an object of \( C \) and \( \alpha, \lambda, \rho \) are three natural isomorphisms

\[
\alpha_{A,B,C}: A \otimes (B \otimes C) \to (A \otimes B) \otimes C,
\]

\[
\lambda_A: \top \otimes A \to A,
\]

\[
\rho_A: A \otimes \top \to A
\]
such that the diagrams

\[
\begin{array}{c}
A \otimes (B \otimes (C \otimes D)) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]

\[
\begin{array}{c}
A \otimes ((B \otimes C) \otimes D) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]

\[
\begin{array}{c}
A \otimes (B \otimes (C \otimes D)) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]

\[
\begin{array}{c}
A \otimes (B \otimes C) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]

commute. We call the category strict monoidal when \( \alpha_{A,B,C}, \lambda_A \) and \( \rho_A \) are identity morphisms.

Definition 2.7.2 [Mac Lane, 1998]. A monoidal category is said to be symmetric when it is equipped with a natural isomorphism \( \sigma_{A,B}: A \otimes B \to B \otimes A \) such that the diagrams

\[
\begin{array}{c}
A \otimes (B \otimes C) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]

\[
\begin{array}{c}
A \otimes (C \otimes B) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]

commute.

2.7.2 Commutative Comonoids

Definition 2.7.3. In a symmetric monoidal category \( C \), a commutative comonoid object is an object \( A \) of \( C \) equipped with two arrows \( \hat{\Delta}_A: A \to \top \) and \( \Delta_A: A \to A \otimes A \) such that the following diagrams commute

\[
\begin{array}{c}
A \otimes (A \otimes A) \\
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\end{array}
\]
A morphism \( f : (A, \triangle_A, \Delta_A) \to (B, \triangle_B, \Delta_B) \) of commutative comonoids is an arrow \( f : A \to B \) in \( C \) such that

\[
\begin{align*}
A \otimes A &\xrightarrow{\rho_A} A & A \otimes A &\xrightarrow{\Delta_A} A \otimes A, \\
A &\xrightarrow{\sigma_A} A & A &\xrightarrow{\lambda_A} A
\end{align*}
\]

Definition 2.7.4 [Selinger 2001]. We call a diagonal structure on a symmetric monoidal category \((C, \otimes, \top)\) a family \(\{(A, \triangle_A, \Delta_A)\}_{A \in |C|}\) of commutative comonoids that respects the symmetric monoidal structure of \(C\), that is, such that the following diagrams commute:

\[
\begin{array}{ccc}
\delta_A \otimes B &\xrightarrow{\alpha} A \otimes B & B \\
\lambda_T \otimes = \rho_T & & \lambda_T \\
T \otimes T &\xrightarrow{\lambda_T \otimes = \rho_T} T, & T,
\end{array}
\]

\[
\begin{array}{ccc}
\Delta_A \otimes B &\xrightarrow{\triangle_A \otimes B} (A \otimes B) \otimes (A \otimes B) & (A \otimes A) \otimes (B \otimes B) \\
\delta_{\otimes A,B} & & \delta_{\otimes A,B}
\end{array}
\]

where \(sw\) is defined as the natural transformation constructed from \(\alpha\)'s and \(\sigma\)'s swapping the two middle objects, that is \(\alpha; (\alpha^{-1} \otimes \text{id}); ((\text{id} \otimes \sigma) \otimes \text{id}); (\alpha \otimes \text{id}); \alpha^{-1}\).

Theorem 2.7.5 [Selinger 2001]. Let \((C, \otimes, \top, \alpha, \lambda, \rho, \sigma)\) be a symmetric monoidal category. It is a cartesian category if and only if there exist two natural transformations \(\Delta_A : A \to A \otimes A\) and \(\triangle_A : A \to \top\) such that \(\{(A, \Delta_A, \Delta_A)\}_{A \in |C|}\) forms a diagonal structure. In this case the cartesian structure is given by

\[
\begin{align*}
\circ_A : A &\xrightarrow{\circ_A} \top, \\
\pi_{A,B}^1 : A \otimes B &\xrightarrow{A \otimes \top \circ_A} A \otimes A & \rho_A, A, \\
\pi_{A,B}^2 : A \otimes B &\xrightarrow{\top \otimes B \circ_A} A \otimes B & \lambda_B, B,
\end{align*}
\]

where \(f : C \to A\) and \(g : C \to B\). Given a cartesian structure, we construct the diagonal structure with \(\Delta_A = (\text{id}_A, \text{id}_A)\) and \(\triangle_A = \circ_A\).

\[\square\]

2.8 Monoidal Comonads and Coalgebras

This section is taken from [Eilenberg and Kelly 1965].

Definition 2.8.1. A (lax) monoidal functor \(F\) between two monoidal categories \((C, \otimes, \top)\) and \((D, \otimes', \top')\) consists of a functor \(F : C \to D\) together with two natural transformations

\[
\begin{align*}
d^F_{A,B} : FA \otimes FB &\to F(A \otimes B), \\
d^F_{\top} : \top' &\to F\top,
\end{align*}
\]
called coherence maps, with the following coherence equations:

\[
\begin{array}{ccc}
(F A \otimes' F B) \otimes' F C & \xrightarrow{\sigma} & F A \otimes' (F B \otimes' F C) \\
\end{array}
\]

\[
\begin{array}{ccc}
d^F_{A,B} & & \phi^F_{A,B} \\
F(A \otimes B) \otimes F C & \xrightarrow{\phi_{A,B} \otimes F} & F A \otimes (B \otimes F C) \\
d^F_{A,B} & & \\
F((A \otimes B) \otimes C) & \xrightarrow{F\phi} & F(A \otimes (B \otimes C)). \\
\end{array}
\]

If the categories are symmetric, the functor is said to be lax symmetric monoidal if the following diagram also commutes:

\[
\begin{array}{ccc}
F A \otimes' F B & \xrightarrow{\sigma} & F B \otimes' F A \\
d^F_{A,B} & & \phi^F_{A,B} \\
F(A \otimes B) & \xrightarrow{F\phi} & F(B \otimes A). \\
\end{array}
\]

We call the functor strong monoidal if the natural transformations \(d^F_{A,B}\) and \(\phi^F\) are isomorphisms.

**Theorem 2.8.2** (Mac Lane, 1998, p. 257). Any monoidal category \(C\) is categorically equivalent, via a strong monoidal functor \(F : C \rightarrow D\) and a strong monoidal functor \(G : D \rightarrow C\), to a strict monoidal category \(D\).

**Lemma 2.8.3.** Let \((C, \otimes_C, \tau_C)\), \((D, \otimes_D, \tau_D)\) and \((E, \otimes_E, \tau_E)\) be symmetric monoidal categories.

1. The identity functor \(1_C : C \rightarrow C\) together with \(d^C_\otimes = id: \tau_C : \tau_C \rightarrow \tau_C\) and \(d^{1_C}_{A \otimes_C B} = id_{A \otimes_C B} : A \otimes_C B \rightarrow A \otimes_C B\) is a monoidal functor.

2. The functor \(\triangle_C\) from \(C\) to \(C\) sending \(A\) to \(\tau_C\) and \(f \mapsto id_{\tau_C}\) is monoidal, with

\[
d^\triangle_C = id_{\tau_C} : \tau_C \rightarrow \tau_C,
\]

\[
d^\triangle_{A,B} = \lambda_{\tau_C} = \rho_{\tau_C} : \triangle_C A \otimes C B = \tau_C \otimes \tau_C \rightarrow \tau_C = \triangle_C (A \otimes C B).
\]

3. The diagonal functor \(\Delta_C\) from \(C\) to \(C\) sending \(A\) to \(A \otimes A\) and \(f \mapsto f \otimes f\) is monoidal with

\[
d^\Delta_C : \tau \xrightarrow{\lambda^\tau_1 = \rho_\tau^{-1}} \tau \otimes \tau,
\]

\[
d^\Delta_{A,B} : \Delta_C A \otimes \Delta_C B = (A \otimes A) \otimes (B \otimes B),
\]

\[
\xrightarrow{\text{sw}_{A,B}} (A \otimes B) \otimes (A \otimes B) = \Delta_C (A \otimes B).
\]

4. If \((F, d^F, d^F_{A,B}) : C \rightarrow D\) and \((G, d^G, d^G_{A,B}) : D \rightarrow E\) are any monoidal functor, so is \(G \circ F : C \rightarrow E\) together with

\[
\begin{array}{ccc}
d^G_{A,B} & & \phi^G_{A,B} \\
G(F A \otimes E G B) & \xrightarrow{G(\phi^F_{A,B})} & GF(A \otimes E B) \\
d^G_{A,B} & & \\
G(F A \otimes E G B) & \xrightarrow{G\phi^F} & GF(A \otimes E B). \\
\end{array}
\]
Definition 2.8.4. In the setting of Definition 2.8.1 if \((F, d)\) and \((F', d')\) are two symmetric monoidal functors, we say that a natural transformation \(n : C \rightarrow D\) is monoidal if the following diagrams commute:

\[
\begin{array}{ccc}
FA \otimes FB & \xrightarrow{n_{A,B}} & F' A \otimes F'B \\
\downarrow d_{A,B} & & \downarrow d'_{A,B} \\
F(A \otimes B) & \xrightarrow{n_{A\otimes B}} & F'(A \otimes B),
\end{array}
\]

Remark 2.8.5. The natural transformation \(n\) in Definition 2.8.4 could be called symmetric since the category is symmetric. However, since \(\sigma\) is not a structure but a property of monoidal functors, it does not yield any equation. In that sense the adjective “symmetric” does not apply to \(n\).

Definition 2.8.6 (Benton et al., 1992; Bierman, 1993). Let \((C, \otimes, \top)\) be a (symmetric) monoidal category. A (symmetric) monoidal comonad on \(C\) is a comonad \((L, \delta, \epsilon)\):

- equipped with two natural transformations \(d_{A,B} : LA \otimes LB \rightarrow L(A \otimes B)\) and \(d_{\top} : L\top \rightarrow \top\) making \((L, d)\) a lax (symmetric) monoidal functor,

- such that \(\delta\) and \(\epsilon\) are monoidal natural transformations, i.e. such that the following diagrams commute:

\[
\begin{array}{ccc}
LA \otimes LB & \xrightarrow{d_{A,B}} & L(A \otimes B), \\
\downarrow \epsilon_{A\otimes B} & & \downarrow \epsilon_{A\otimes B} \\
A \otimes B & \xrightarrow{\epsilon_A \otimes \epsilon_B} & A \otimes B,
\end{array}
\]

\[
\begin{array}{ccc}
L^2 A \otimes L^2 B & \xrightarrow{d_{L^2 A, L^2 B}} & L(LA \otimes LB) \\
\downarrow \delta_{A\otimes B} & & \downarrow \delta_{A\otimes B} \\
L^2(A \otimes B) & \xrightarrow{d_{\top, \top}} & L^2(\top, \top)
\end{array}
\]

Theorem 2.8.7. Let \((C, \otimes, \top)\) be a symmetric monoidal category with structure \(\alpha, \lambda, \rho\) and \(\sigma\). Let \((L, \delta, \epsilon, d)\) be a symmetric monoidal comonad on \(C\). Then the co-Eilenberg-Moore category \(CL\) is also symmetric monoidal when equipped with

- the functor \(\otimes' : CL \times CL \rightarrow CL\), defined as \((A, h_A) \otimes' (B, h_B) = (A \otimes B, h_A \circ h_B); d_{A,B}\) on objects and \(f \otimes' g = f \otimes g\) on arrows,

- the unit object \(\top' = (\top, d_{\top})\),

- the structure maps

\[
\begin{align*}
\alpha'_{(A,h_A),(B,h_B),(C,h_C)} &= \alpha_{A,B,C}, \\
\lambda'_{(A,h_A)} &= \lambda_A, \\
\sigma'_{(A,h_A),(B,h_B)} &= \sigma_{A,B}, \\
\rho'_{(A,h_A)} &= \rho_A.
\end{align*}
\]

Definition 2.8.8. In Theorem 2.8.7 we call \(\otimes'\) the induced tensor and \((\otimes', \top', \alpha, \lambda, \rho, \sigma)\) the induced symmetric monoidal structure. We identify \(\otimes'\) with \(\otimes\) and \(\top'\) with \(\top\).
Chapter 3

Quantum Computation

Whether the notion of data is thought of concretely or abstractly, it is usually supposed to behave classically: a piece of data is supposed to be clonable, erasable, readable as many time as needed, and is not supposed to change when left untouched.

Quantum computation is a paradigm where data can be encoded with a medium governed by the law of quantum physics. Although only rudimentary quantum computers have been built so far, the laws of quantum physics are mathematically well described. It is therefore possible to try to understand the capabilities of quantum computation, and quantum information turns out to behave substantially differently from usual (classical) information.

In this chapter we will describe the mathematics needed for quantum computation, and we will have a brief overview of the theory of quantum computation. A more complete study of the subject can be found in (Nielsen and Chuang, 2002) and in (Preskill, 1999).

3.1 Mathematical Formalism

This section will cover the common wisdom we need about Hilbert spaces and their operators. All considered vector spaces are assumed to be finite-dimensional. For a more complete mathematical analysis of the subject, see e.g. (Lang, 1993, Ch. V).

3.1.1 Generalities on Finite Dimensional Hilbert Spaces

Definition 3.1.1 (Lang, 1993, p. 95). Let $\mathcal{E}$ be a finite-dimensional vector space over the complex numbers $\mathbb{C}$. A sesquilinear form on $\mathcal{E}$ is a map $\phi: \mathcal{E} \times \mathcal{E} \to \mathbb{C}$, which is linear in $y$: $\phi(x, \alpha y + \beta z) = \alpha \phi(x, y) + \beta \phi(x, z)$, and semi-linear in $x$: $\phi(\alpha x + \beta t, y) = \overline{\alpha} \phi(x, y) + \beta \phi(t, y)$, where $\overline{\alpha}$ is the complex conjugation of $\alpha$: $a + bi = a - bi$, for $a, b \in \mathbb{R}$. We say the form is hermitian if moreover $\phi(x, y) = \overline{\phi(y, x)}$. It is positive definite if $\phi(x, x) \geq 0$ with equality if and only if $x = 0$. If a form is hermitian and positive definite, we will call it a scalar product. In this case we often denote it as $\langle .|.angle$.

Definition 3.1.2. Let $\mathcal{E}$ be a complex vector space. A norm on $\mathcal{E}$ is a map $\| \cdot \|: \mathcal{E} \to \mathbb{R}$ such that

1. $|v| \geq 0$ for all $v$, with equality if and only if $v = 0$;
2. $|\alpha v| = |\alpha||v|$;
3. $|v + w| \leq |v| + |w|$.
Lemma 3.1.3. Given a scalar product $\langle \cdot, \cdot \rangle$ on a complex vector space $\mathcal{E}$, the map $\|\cdot\| : \mathcal{E} \to \mathbb{C}$ mapping $v$ to $\sqrt{\langle v|v \rangle}$ is a norm.

Definition 3.1.4. Given a scalar product on a complex vector space $\mathcal{E}$. The norm in Lemma 3.1.3 is called the induced norm.

Definition 3.1.5. A vector space together with a norm is called a normed vector space.

Definition 3.1.6. Let $\mathcal{H}$ be a finite dimensional vector space over the field $\mathbb{C}$ of complex numbers. The space $\mathcal{H}$ is a Hilbert space if it comes equipped with a scalar product. When speaking about a given Hilbert space, we will always consider it to be normed by the induced norm.

Definition 3.1.7. An operator $A$ on a Hilbert space $\mathcal{H}$ is a linear map $A : \mathcal{H} \to \mathcal{H}$. A functional on $\mathcal{H}$ is a linear map $\lambda : \mathcal{H} \to \mathbb{C}$. The space of functionals on $\mathcal{H}$ is denoted by $\mathcal{H}^*$. 

Lemma 3.1.8. Let $\lambda$ be a functional on $\mathcal{H}$. Then there exists $x_0 \in \mathcal{H}$ such that for all $y \in \mathcal{H}$, $\lambda(y) = \langle x_0|y \rangle$. 

Lemma 3.1.9. Let $\phi$ be a sesquilinear form on a Hilbert space $\mathcal{H}$. Then the map $\lambda_x : y \mapsto \phi(x,y)$ is a functional.

Let $A$ be an operator on $\mathcal{H}$. Define $\phi_A$ by $\phi_A(x,y) = \langle x|Ay \rangle$. Then $\phi_A$ is a sesquilinear form on $\mathcal{H}$. For each $x$, $\lambda_x(y) = \phi_A(x,y)$ is a functional. Therefore there exists some $x_0$ such that $\lambda_x(y) = \langle x_0|y \rangle$. Define $A^*$ as $A^*x = x_0$: we have $\langle x|Ay \rangle = \langle A^*x|y \rangle$.

Definition 3.1.10. One calls $A^*$ the adjoint of $A$. If $A = A^*$, we say $A$ is self-adjoint, or hermitian.

Lemma 3.1.11 (Lang, 1993, p. 106). The map $A \mapsto A^*$ satisfies the following properties:

$$(A + B)^* = A^* + B^*, \quad (\alpha A)^* = \overline{\alpha} A^*, \quad (AB)^* = B^* A^*.$$

Lemma 3.1.12 (Lang, 1993, p. 107). If for all $x$, $\langle x|Ax \rangle = 0$ then $A = 0$.

Theorem 3.1.13 (Spectral Theorem, Lang, 2002). Let $A$ be a hermitian operator on a finite dimensional Hilbert space $\mathcal{H}$. Then $\mathcal{H}$ has an orthogonal basis consisting of eigenvectors of $A$.

Definition 3.1.14. We say that an operator $A$ on $\mathcal{H}$ is positive if it is hermitian and if all of its eigenvalues are positive. Equivalently, $A$ is positive if the form $\phi_A$ is positive definite. We say $A$ is a density matrix if moreover, the sum of the eigenvalues is less or equal to 1.

Lemma 3.1.15. Let $\mathcal{H}$ be a finite dimensional Hilbert space, and $A$ an invertible operator. Then the following conditions are equivalent:

1. $A^* = A^{-1}$;
2. $\|Ax\| = |x|$ for all $x \in \mathcal{H}$;
3. $\langle Ax|Ay \rangle = \langle x|y \rangle$ for all $x, y \in \mathcal{H}$;
4. $\|Ax\| = 1$ for every unit vector $x \in \mathcal{H}$.

Definition 3.1.16. An invertible operator satisfying the conditions of Lemma 3.1.15 is said to be unitary.
3.1. Mathematical Formalism

3.1.2 Tensor Products of Hilbert Spaces

Definition 3.1.17. Consider two finite Hilbert spaces $V$ and $W$ with respective bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_k)$. We define the tensor product of $V$ and $W$, denoted $V \otimes W$, as the space generated by the basis of syntactic symbols:

$$(e_1 \otimes f_1, \ldots, e_1 \otimes f_k, \ldots, e_n \otimes f_1, \ldots, e_n \otimes f_k).$$

The tensor product of two elements $v = \sum \lambda_i \cdot e_i$ and $w = \sum \mu_j \cdot f_j$ is

$$v \otimes w = \sum_{i,j} \lambda_i \mu_j \cdot e_i \otimes f_j.$$ 

We define the scalar product in $V \otimes W$ as follows:

$$\langle e_i \otimes f_j | e_k \otimes f_l \rangle = \langle e_i | e_k \rangle \langle f_j | f_l \rangle.$$ 

Lemma 3.1.18. If $U$, $V$ and $W$ are finite-dimensional Hilbert spaces, there exists a one-to-one correspondence $\Phi$ between linear maps $U \otimes V \to W$ and linear maps $U \to V \otimes W$.

Proof. Choose a basis $(e_i)_i$ for $V$ and a basis $(f_j)_j$ for $W$. Then if $f : U \otimes V \to W$, we define $g = \Phi(f) : U \to V \otimes W$ as the map

$$g(u) = \sum_i e_i \otimes f(u \otimes e_i).$$

If $g : U \to V \otimes W$, and if

$$g(u) = \sum_{i,j} \rho_{i,j} \cdot e_i \otimes f_j,$$

we define $f = \Phi^{-1}(g) : U \otimes V \to W$ as the map

$$f(u \otimes e_i) = \sum_j \rho_{i,j} \cdot e_i \otimes f_j.$$ 

Convention 3.1.19. If $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{k \times k}$ are two square matrices of respective size $n \times n$ and $k \times k$ seen as operators respectively on $\mathbb{C}^n$ and on $\mathbb{C}^k$, the matrix $A \otimes B$ is an operator on $\mathbb{C}^{nk}$ and we represent it as the matrix of blocks

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,n}B \end{pmatrix}.$$ 

3.1.3 Completely Positive Maps

In this section, we focus primarily on Hilbert spaces of the form $\mathbb{C}^{n \times n}$, spaces of of $n \times n$-matrices. We choose as basis for $\mathbb{C}^{n \times n}$ the canonical basis $(E_{i,j})_{i,j \in \{1, \ldots, n\}}$, where $(E_{i,i})_{k,l} = 0$ for all $k$ and $l$ except when $k = i$ and $l = j$. We use the notations and definitions of Selinger [2004].

Definition 3.1.20. A linear map $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ is said to be a positive map if for all positive matrices $A$, $F(A)$ is positive.\footnote{Note that we identified $V$ with $V^*$. Although the result is correct since we are in finite dimension, we lose the basis independence of the correspondence.}
Definition 3.1.21. A linear map $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ is said to be completely positive if for all $i$, if $id_i$ is the identity on $i \times i$-matrices, $id_i \otimes F$ is a positive map $\mathbb{C}^{n_i \times n_i} \to \mathbb{C}^{k_i \times k_i}$.

Definition 3.1.22. The characteristic matrix of a linear map $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ is defined as the $nk \times nk$-matrix of blocks

$$\chi_F = \begin{pmatrix}
F(E_{1,1}) & \cdots & F(E_{1,n}) \\
\vdots & \ddots & \vdots \\
F(E_{n,1}) & \cdots & F(E_{n,n})
\end{pmatrix}.$$

Lemma 3.1.23. Let $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ be a linear operator, and let $\chi_F$ be its characteristic matrix. Then $F(A) = UAU^*$ for some $U \in \mathbb{C}^{k \times n}$ if and only if there exists a vector $v \in \mathbb{C}^{nk}$ such that $vv^* = \chi_F$.

Theorem 3.1.24 (Choi, 1975). Let $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ be a linear operator, and let $\chi_F$ be its characteristic matrix. Then the following are equivalent:

1. $F$ is completely positive
2. $\chi_F$ is positive
3. $F$ is of the form $F(A) = \sum_i U_i A U_i^*$ for some finite sequence of matrices $U_1, \ldots, U_l$ in $\mathbb{C}^{k \times n}$. $$\square$$

3.1.4 Superoperators

The following definitions and results are from [Selinger, 2004b].

Definition 3.1.25 (Löwner order). We define a relation $\sqsubseteq$ on $\mathbb{C}^{n \times n}$ as follows. Let $A, B$ be matrices in $\mathbb{C}^{n \times n}$. Then $A \sqsubseteq B$ if and only if $B - A$ is positive.

Definition 3.1.26. A completely positive map $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ is called a superoperator if for all positive matrices $A \in \mathbb{C}^{n \times n}$, we have $tr(F(A)) \leq tr(A)$.

Theorem 3.1.27 (Kraus Representation Theorem). Let $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ be a completely positive operator. The following are equivalent:

1. $F$ is a superoperator;
2. $(id_n \otimes tr_k)(\chi_F) \sqsubseteq I_n$, where $I_n$ is the identity matrix and $tr_k : \mathbb{C}^{k \times k} \to \mathbb{C}$ is the trace operation;
3. $F$ is of the form

$$F(A) = \sum_i U_i A U_i^*$$

for matrices $U_1, \ldots, U_n$ with $\sum_i U_i^* U_i \sqsubseteq I_n$. $$\square$$

Definition 3.1.28. A matrix $U \in \mathbb{C}^{k \times n}$ is called contraction if $U$ is a sub-matrix of some unitary $U'$, that is, if there exist matrices $U_1$, $U_2$, and $U_3$ such that

$$U' = \begin{pmatrix}
U \\
U_1 \\
U_2
\end{pmatrix}$$

is unitary. A linear function $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{k \times k}$ is called a contraction if it is of the form $F(A) = UAU^*$ for some contraction $U$. $\square$
Lemma 3.1.29. Let $U \in \mathbb{C}^{n \times m}$ be a contraction. Then the matrix
\[
\begin{pmatrix}
U \\
0
\end{pmatrix} \in \mathbb{C}^{(n+1) \times m}
\]
is also a contraction.

Proof. With the notation of the definition, consider the unitary matrix
\[
\begin{pmatrix}
U & 0 & U_1 \\
0 & 1 & 0 \\
U_2 & 0 & U_3
\end{pmatrix}.
\]

Lemma 3.1.30. A matrix $U \in \mathbb{C}^{k \times n}$ is a contraction if and only if $UU^* \subseteq I_k$, and if and only if $U^*U \subseteq I_n$.

Definition 3.1.31. If $\text{tr}_n$ is the trace on matrices in $\mathbb{C}^{n \times n}$, we define the partial trace operator $\mathbb{C}^{kn \times kn} \to \mathbb{C}^{k \times k}$ as
\[
\text{tr}_n^k = \text{tr}_n \otimes \text{id}_{\mathbb{C}^k \times k}.
\]

Theorem 3.1.32 (Naimark Theorem). Every superoperator $F : \mathbb{C}^{n \times n} \to \mathbb{C}^{m \times m}$ can be factored as $F = E \circ G$, where $G$ is a contraction and $E$ is a partial trace operator.

3.1.5 The 2-Dimensional Hilbert Space

We will now study in more detail what we saw in the previous section, applied to the 2-dimensional Hilbert space. We identify it with $\mathbb{C}^2$ and we call the canonical basis $e_1 = (1,0)$ and $e_2 = (0,1)$. The results of this section can be found in [Preskill, 1999].

Lemma 3.1.33. The set of hermitian matrices on $\mathcal{H}$ is spanned as a real vector space by the following basis:
\[
I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

Definition 3.1.34. The matrices $\sigma_x, \sigma_y, \sigma_z$ are called the Pauli matrices.

Remark 3.1.35. The eigenvectors and eigenvalues of the Pauli matrices are as follows:
\[
\sigma_x (\frac{1}{2}) = \frac{1}{2}, \quad \sigma_y (\frac{1}{2}) = \frac{1}{2}, \quad \sigma_z (\frac{1}{2}) = (\frac{1}{2}), \quad \sigma_x (\frac{1}{2}) = (\frac{1}{2}), \quad \sigma_y (\frac{1}{2}) = (\frac{1}{2}), \quad \sigma_z (\frac{1}{2}) = (\frac{1}{2}).
\]

Definition 3.1.36. Let $\mathcal{H}$ be a Hilbert space. A ray $[v]$ in $\mathcal{H}$ is an equivalence class of vectors in $\mathcal{H}$ that differ by multiplication by non-zero complex scalars: so $[v]$ contains all $\alpha v$, $\alpha \in \mathbb{C} \setminus \{0\}$.

Lemma 3.1.37. There is a group action from the group of linear operators of $\mathcal{H}$ on the set of rays of $\mathcal{H}$.

Lemma 3.1.38. If $v$ and $w$ are orthogonal, so are any two elements of their respective rays.

Lemma 3.1.39. The rays of $\mathbb{C}^2$ are in one-to-one correspondence with the points on the unit sphere of $\mathbb{R}^3$. In spherical coordinates, a point $(\theta, \varphi)$ with $\theta \in [0, \pi], \psi \in [-\pi, \pi]$ is in correspondence with the ray of representative
\[
r(\theta, \varphi) = \left( \begin{array}{c} \cos \frac{\theta}{2} e^{i\varphi} \\ e^{i\varphi} \sin \frac{\theta}{2} \end{array} \right).
\]

The correspondence is pictured in Figure 3.1.
3.1. Mathematical Formalism

One calls this unit sphere the **Bloch sphere**. In this representation, the Pauli matrices acts on the rays as follows:

- $\sigma_x$ sends $r(\theta, \varphi)$ onto
  
  \[
  \begin{pmatrix}
  e^{i\varphi} \sin \frac{\theta}{2} \\
  \cos \frac{\theta}{2}
  \end{pmatrix}
  \approx
  \begin{pmatrix}
  \cos \frac{\pi - \theta}{2} \\
  -e^{i\varphi} \sin \frac{\pi - \theta}{2}
  \end{pmatrix}
  \]

  thus sends the point $(\theta, \varphi)$ of the unit sphere of $\mathbb{R}^3$ (in spherical coordinate) to $(\pi - \theta, -\varphi)$, that is, in cartesian coordinates,

  \[
  \begin{pmatrix}
  \sin \theta \cos \varphi \\
  \sin \theta \sin \varphi \\
  \cos \theta
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  \sin \theta \cos \varphi \\
  -\sin \theta \sin \varphi \\
  -\cos \theta
  \end{pmatrix}
  \quad \text{or} \quad 
  \begin{pmatrix}
  x \\
  y \\
  z
  \end{pmatrix}
  \mapsto
  \begin{pmatrix}
  x \\
  -y \\
  -z
  \end{pmatrix}.
  \]

  The operator $\sigma_x$ acts on the sphere as an axial symmetry around $Ox$. The two eigenvectors $(\frac{1}{2})$ and $(\frac{1}{2})$ respectively correspond to the point $(1,0,0)$ and to the point $(-1,0,0)$ on the sphere.

- $\sigma_y$ sends $r(\theta, \varphi)$ onto
  
  \[
  \begin{pmatrix}
  e^{i\varphi} \sin \frac{\theta}{2} \\
  \cos \frac{\theta}{2}
  \end{pmatrix}
  \approx
  \begin{pmatrix}
  \cos \frac{\pi - \theta}{2} \\
  e^{i(\pi - \varphi)} \sin \frac{\pi - \theta}{2}
  \end{pmatrix}
  \]

  i.e. maps $(\theta, \varphi)$ to $(\pi - \theta, \pi - \varphi)$, i.e. computes an axial symmetry around $Oy$. The two eigenvectors $(\frac{1}{2})$ and $(\frac{1}{2})$ respectively correspond to the point $(0,1,0)$ and to the point $(0,-1,0)$ on the sphere.

- Finally, $\sigma_z$ sends $r(\theta, \varphi)$ onto
  
  \[
  \begin{pmatrix}
  e^{i\varphi} \sin \frac{\theta}{2} \\
  \cos \frac{\theta}{2}
  \end{pmatrix}
  \approx
  \begin{pmatrix}
  \cos \frac{\theta}{2} \\
  e^{i(\pi - \varphi)} \sin \frac{\theta}{2}
  \end{pmatrix}
  \]

  i.e. maps $(\theta, \varphi)$ to $(\theta, \pi - \varphi)$, i.e. computes an axial symmetry around $Oz$. The two eigenvectors $(\frac{1}{2})$ and $(\frac{1}{2})$ respectively correspond to the point $(0,0,1)$ and to the point $(0,0,1)$ on the sphere.
Lemma 3.1.40. Two rays in $\mathbb{C}^2$ are orthogonal if and only if their representations on the Bloch sphere are antipodal.

3.2 Quantum Foundations

Quantum computation is somehow the art of making the theory of finite Hilbert spaces into a theory of computation: In quantum computation, information is encoded on states of quantum particles, states whose mathematical representation are rays in some Hilbert space.

Convention 3.2.1. In the following, we use the Dirac convention for writing vectors:

- vectors are written as kets: $|\phi\rangle$;
- functionals are written as bras: $\langle \psi |$;
- scalar products, which are by Lemma 3.1.9 the application of some functional to some vectors, are written as usual: $\langle \psi | \phi \rangle$.

3.2.1 Quantum Bits

One Quantum Bit

In classical computation, the smallest unit of data is the bit, element of the two-elements set $\{0, 1\}$. In quantum computation, the smallest unit of data is a quantum bit, or qubit, defined as a ray in a 2-dimensional Hilbert space. We write the canonical basis of this space as $\{|0\rangle, |1\rangle\}$.

Using the results of Section 3.1.5, one says that

$$
(0, |1\rangle) = \left(0, \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|1\rangle - |1\rangle)\right)
$$

is the basis along $z$,

$$
(0, |1\rangle) = \left(0, \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \frac{1}{\sqrt{2}}(|1\rangle - i|1\rangle)\right)
$$

is the basis along $y$.

Several Quantum Bits

When considering a system of $n$ quantum bits, the state of the system is a ray in $\mathcal{H}^\otimes n = \mathcal{H}^\otimes n$, where $\mathcal{H}$ is the 2-dimensional Hilbert space. As stated in Definition 3.1.17, we take the canonical basis for $\mathcal{H}^\otimes n$ to be $\{|i_1\ldots i_n\rangle \mid i_1, \ldots, i_n \in \{0, 1\}\}$, where we write $|i_1\ldots i_n\rangle$ for $|i_1\rangle \otimes \cdots \otimes |i_n\rangle$.

3.2.2 Operations

Measurements

Observables. In quantum physics, a measurement in a Hilbert space $\mathcal{H}$ is associated with a self-adjoint operator $A$ of $\mathcal{H}$. Physicists call $A$ an observable.

As we saw in Theorem 3.1.18 to a self-adjoint operator one can associate a set of orthogonal eigenspaces with their corresponding eigenvalues. If $|\phi\rangle$ is an eigenvector of $A$ with eigenvalue $a$, the numerical outcome of the measurement of $|\phi\rangle$ with the observable $A$ is the value $a$. 

What does happen if $\phi$ is not an eigenvector? For simplicity assume that $\mathcal{H}$ is finite and that the basis of eigenvectors of $A$ is $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$ such that the eigenvalue corresponding to $|\phi_i\rangle$ is $a_i$. Then $|\phi\rangle$ can be re-written as $|\phi\rangle = \sum_{i=1}^{n} \alpha_i |\phi_i\rangle$, and measuring with $A$ makes the system randomly collapse to a new state, namely a basis vector $|\phi_i\rangle$. This new state is $|\phi_i\rangle$ with probability $|\alpha_i|^2 = |\langle \phi | \phi_i \rangle|^2$. In this situation, the measurement outputs the numerical value $a_i$.

For the purpose of measurement the numerical eigenvalues have little importance. What is really important is that two eigenvalues corresponding to two distinct eigenspaces differ. When measuring a state of $n$ quantum bits, we say that the outcome of the measurement of the state is $x_1 \ldots x_n$ if the state collapsed to $|x_1 \ldots x_n\rangle$. For example, we say:

"The measurement of the state $|\phi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$ in the standard basis results in the outcome 01 with probability $|\langle 01 | \phi \rangle|^2 = |\beta|^2$."

**Pauli matrices.** Given a quantum bit state $|\phi\rangle$, for $i = x, y, z$ the Pauli matrix $\sigma_i$ provides an observable, whose measurement projects the state $|\phi\rangle$ in the basis along axis $i$.

From Section 3.1.5, the bases along $x, y$ and $z$ have special properties: if a quantum bit is in one of the basis element of the base along $i$, for some $i = x, y$ or $z$, then measuring the observable $\sigma_j$, $j \neq i$ will yield 0 with probability $\frac{1}{2}$ and 1 with probability $\frac{1}{2}$. Thus somehow the information stored in the basis along $i$ is not accessible from the two other ones.

**Convention 3.2.2.** In the remainder of the thesis, when speaking of “a measurement” without further precision, we always consider the measurement to be along the basis $z$.

**Unitary gates**

A quantum mechanical system, if not measured (by the operator or by the environment), evolves along a unitary path. That is, loosely speaking, all we can do is to apply rotations on the system. We call these actions unitary gates.

The ability to perform one gate or another depends on the physical implementation of the quantum bits. In general, only a few gates are needed to approximate any given unitary matrix, a result known as the Solovay-Kitaev Theorem. We already met the Pauli matrices in Lemma 3.2.3.

The matrix $\sigma_2$ is also called the not gate:

$$
N = \begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}.
$$

Two other usual gates are the Hadamard gate and the phase-shift gate:

$$
H = \frac{1}{\sqrt{2}} \begin{pmatrix}
1 & 1 \\
1 & -1
\end{pmatrix},
\quad V_\theta = \begin{pmatrix}
1 & 0 \\
0 & e^{i\theta}
\end{pmatrix}.
$$

Finally, written in the canonical basis ($|00\rangle, |01\rangle, |10\rangle, |11\rangle$), we have the exchange gate and the controlled-not gate:

$$
X = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\quad N_C = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{pmatrix}.
$$

The $N_C$-gate comes from a generic process: If $U$ is a one-qubit gate, one constructs $U_C$, called controlled $U$, to be the two-qubits gate:

$$
U_C (|0\rangle \otimes |\phi\rangle) = |0\rangle \otimes |\phi\rangle,
\quad U_C (|1\rangle \otimes |\phi\rangle) = |1\rangle \otimes U|\phi\rangle.
$$

The gate $U$ is applied on the second qubit depending on the value of the first qubit.
3.2. Quantum Foundations

Quantum Circuit

It is possible to write tensors and compositions of gates as wires and boxes on wires. For example, the map \( H \otimes V \theta \) can be written as

![Diagram of a quantum circuit with wires and boxes.]

Some gates have specific representations:

- The gate \( N \) is written as \( \oplus \).
- A controlled-\( U \) gate is written as \( \cdot U \).
- The gate \( N_C \) is written as \( \cdot \).
- The gate \( X \) is simply \( \cdot \).
- Measurements will be represented by dashed boxes followed by a dashed wire, to indicate that we now have a classical bit, as follows:

![Diagram of a measurement.]

3.2.3 Mixed States

The representation of a state of a system of quantum bits as a ray in some Hilbert space is a valid description of an isolated quantum system. It does not work when only part of the system is considered. Indeed, suppose we are given a two-qubit system \( |\phi_{AB}\rangle \) in the state

\[
\alpha |0_A\rangle \otimes |0_B\rangle + \beta |0_A\rangle \otimes |1_B\rangle + \gamma |1_A\rangle \otimes |0_B\rangle + \delta |1_A\rangle \otimes |1_B\rangle,
\]

where \( (|0_A\rangle, |1_A\rangle) \) and \( (|0_B\rangle, |1_B\rangle) \) are some basis of \( \mathbb{C}^2 \) and suppose we only have access to qubit \( A \).

Consider an observable \( M \) on \( A \). Using the formalism developed with isolated systems, if we perform a measurement with this observable, the total expectation of the measurement is

\[
\langle \phi_{AB} | (M \otimes I_B) | \phi_{AB} \rangle = \left( \frac{\alpha^* \langle 0_A | + \gamma^* \langle 1_A |}{\sqrt{\alpha^* \langle 0_A | + \gamma^* \langle 1_A |}}M \left( \frac{\alpha |0_A\rangle + \gamma |1_A\rangle}{\sqrt{\alpha^* \langle 0_A | + \gamma^* \langle 1_A |}} \right) + \frac{\beta^* \langle 0_A | + \delta^* \langle 1_A |}{\sqrt{\beta^* \langle 0_A | + \delta^* \langle 1_A |}}M \left( \frac{\beta |0_A\rangle + \delta |1_A\rangle}{\sqrt{\beta^* \langle 0_A | + \delta^* \langle 1_A |}} \right) \right) = \text{tr}(M \rho_A),
\]

where \( \rho_A \) is the density matrix (as in Definition 3.1.14) equal to

\[
\left( \beta |0_A\rangle + \delta |1_A\rangle \right) \left( \beta^* \langle 0_A | + \delta^* \langle 1_A | \right) + \left( \alpha |0_A\rangle + \gamma |1_A\rangle \right) \left( \alpha^* \langle 0_A | + \gamma^* \langle 1_A | \right),
\]

that is

\[
\left( \frac{\alpha^2}{\beta^2} \right) + \left( \frac{\beta^2}{\gamma^2} \right)
\]

in the basis \( (|0_A\rangle, |1_A\rangle) \). Note that if \( M \) is the projection onto \( |0_A\rangle \), as expected we get \( |\alpha|^2 + |\beta|^2 \).

And if it is the projection onto \( |1_A\rangle \), we get \( |\delta|^2 + |\gamma|^2 \).
3.3. Quantum Effects

Thus, if we are only considering the quantum bit $A$, one can safely work with the density matrix $\rho_A$ in place of the whole system. This representation is called a mixed state: Since $|\alpha|^2 + |\gamma|^2 + |\beta|^2 + |\delta|^2 = 1$, if

$$
\alpha' = \frac{\alpha}{|\alpha|^2 + |\gamma|^2}, \quad \gamma' = \frac{\gamma}{|\alpha|^2 + |\gamma|^2}, \quad \beta' = \frac{\beta}{|\beta|^2 + |\delta|^2}, \quad \delta' = \frac{\delta}{|\beta|^2 + |\delta|^2},
$$

then the matrix $\rho_A$ represents a probability distribution over pure states

$$
(|\alpha|^2 + |\gamma|^2)\left\{\alpha'|0_A\rangle + \gamma'|1_A\rangle\right\} + (|\beta|^2 + |\delta|^2)\left\{\beta'|0_A\rangle + \delta'|1_A\rangle\right\}.
$$

We can re-encode the state of an $n$-quantum bit system $|\phi\rangle$ as the $2^n \times 2^n$ density matrix $\rho = |\phi\rangle\langle\phi|$. A unitary operation $U$ on this state yields the density matrix $U\rho U^*$, and the measurement of an observable $M$ yields the expected value $\text{tr}(M\rho)$.

3.3 Quantum Effects

We list in this section a few specifics of quantum computation, following [Preskill 1999].

3.3.1 No Cloning

Consider a quantum bit in some unknown state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$. Then it is not possible to “clone” the state and get the two-qubit state $|\phi\rangle \otimes |\phi\rangle$. The reason is simple: the only operations we are allowed to perform are unitaries and measurements, which are linear maps. The cloning operation being non-linear, it is therefore not implementable.

It is however possible to build a “copying” map $G$

$$
\alpha|0\rangle + \beta|1\rangle \mapsto \alpha|00\rangle + \beta|11\rangle.
$$

But it does not satisfy the usual commutativity rule for duplication: if $f: \mathcal{H} \rightarrow \mathcal{H}$ is any linear map, then $G \circ f \neq (f \otimes f) \circ G$.

3.3.2 Entanglement

Another special property of quantum information is the superposition: the 2-qubit state

$$
\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)
$$

is a valid state, but cannot be written as $|\phi\rangle \otimes |\psi\rangle$. We say the two quantum bits are entangled.

The interesting aspect of entanglement is that the measurements of the qubits will be correlated. For example, measuring the first qubit in the state (3.3.1) with respect to the standard basis yields 0 with probability $\frac{1}{2}$ and 1 with probability $\frac{1}{2}$. In the former case, the state of the 2-qubit system is $|00\rangle$, thus measuring the second qubit yields 0 with probability 1. In the latter, the state of the system is $|11\rangle$, and measuring the second qubit yields 1 with probability 1.

A useful basis for 2-qubit systems is the following set of entangled states:

$$
\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).
$$
3.3. Quantum Effects

3.3.3 Bell’s Inequalities

This section is taken from (Preskill, 1999).

From a classical perspective, a measurement is a special form of coin-toss. A hidden variable theory is a theory where a probability $\lambda$ is associated to each state and to each axis such that the measurement along this axis outputs $\lambda$. The question is then whether it is possible to build a hidden variable theory that would agree with the prediction of quantum theory while still satisfying the strong locality criterion (Einstein et al., 1935), stating:

"Suppose that $A$ and $B$ are space-like separated systems. Then in a complete description of physical reality an action performed on system $A$ must not modify the description of system $B"."

Bell (1964) shows that it is not possible. The argument goes as follows. Consider a quantum machine that maximally entangles two quantum bits $A$ and $B$

$$|\phi_{AB}\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)$$

and sends qubit $A$ to Alice and qubit $B$ to Bob. Suppose that they can independently choose one of the following axes (in the Bloch sphere, see Figure 3.1) to measure:

$$a = (0, 0, 1), \quad b = \left(\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right), \quad c = \left(\frac{-\sqrt{3}}{2}, 0, -\frac{1}{2}\right).$$

They live in the $xz$-plane of the Bloch sphere and correspond respectively to the bases in $\mathbb{C}^2$:

$$|0_a\rangle = |0\rangle, \quad |1_a\rangle = |1\rangle$$

$$|0_b\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle, \quad |1_b\rangle = \frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle,$$

$$|0_c\rangle = \frac{1}{2}|0\rangle - \frac{1}{2}|1\rangle, \quad |1_c\rangle = \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle.$$

What is the probability of obtaining the same output when measuring $A$ and $B$ with respect to two different bases?

In a local hidden variable theory, the result of measuring $A$ and $B$ along each of the possible axis is predetermined. Let $P_{x,y}$ be the probability of obtaining the same output while measuring $A$ along $x$ and $B$ along $y$. For any local hidden variable theory,

$$P_{a,b} + P_{b,c} + P_{c,a} \geq 1, \quad (3.3.2)$$

since for any possible distribution of measurement values, there will always be two values that will be equal (see (Preskill, 1999) for a complete discussion). However, the quantum theory provides the value

$$P_{x,y} = \langle \phi_{AB}|(|0_x\rangle \langle 0_x| \otimes |0_y\rangle \langle 0_y|)|\phi_{AB}\rangle + \langle \phi_{AB}|(|1_x\rangle \langle 1_x| \otimes |1_y\rangle \langle 1_y|)|\phi_{AB}\rangle$$

The computation shows that $P_{x,y} = \frac{1}{4}$ for $x \neq y$, $x, y = a, b, c$. In particular,

$$P_{a,b} + P_{b,c} + P_{c,a} = \frac{3}{4} < 1,$$

which violates Equation (3.3.2).
3.4 Some Algorithms and Uses of Quantum Effects

The main power of quantum computation comes from entanglement. We describe in the following three algorithms that we will consider in the remainder of the thesis. For more algorithms, such as the quantum fast Fourier transform, or Shor’s factorization algorithm [Shor, 1994], a good reference is [Aharonov, 1999].

3.4.1 Teleportation

One use of entanglement is the “teleportation” of a quantum bit state via a classical channel. The idea is the following: Alice and Bob share an entangled pair of qubits

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Alice wants to send a qubit in an unknown state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob. The teleportation procedure, summarized in Figure 3.2, can be described in three steps:

1. At location A, create the initial entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ with the two last qubits using the circuit

   $|0\rangle$  
   $\uparrow$  
   $H$  
   $|0\rangle$

   Alice keeps qubit 2 and stays at location A, while Bob takes qubit 3 and goes to location B.

2. Alice, to send qubit 1 in state $|\phi\rangle$ to Bob, applies a rotation on her two qubits 1 and 2, using

   $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ with the two last qubits using

   $|0\rangle$  
   $\uparrow$  
   $H$  
   $|0\rangle$

   Alice keeps qubit 2 and stays at location A, while Bob takes qubit 3 and goes to location B.
3.4. Some Algorithms and Uses of Quantum Effects

She then measures the 2-qubit resulting state, gets two classical bits \((x, y)\) and sends them to Bob.

3. Bob, to transform qubit 3 to state \(|\phi\rangle\) applies on it the transformation \(U_{xy}\) defined as follows:

\[
\text{if } M \text{ outputs } 00, \quad U_{00} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\
\text{if } M \text{ outputs } 01, \quad U_{01} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\
\text{if } M \text{ outputs } 10, \quad U_{10} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\
\text{if } M \text{ outputs } 11, \quad U_{11} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\]

Note that the entanglement of qubits 2 and 3 can be done ahead of time (so long as they stay entangled).

Proof of the correctness of the protocol. The rotation of step (2) performs the following computation

\[
\begin{align*}
\text{CNOT} & \quad H \otimes \text{id} \\
|00\rangle & \quad \mapsto |00\rangle \quad \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle), \\
|01\rangle & \quad \mapsto |01\rangle \quad \mapsto \frac{1}{\sqrt{2}}(|01\rangle + |11\rangle), \\
|10\rangle & \quad \mapsto |11\rangle \quad \mapsto \frac{1}{\sqrt{2}}(|01\rangle - |11\rangle), \\
|11\rangle & \quad \mapsto |10\rangle \quad \mapsto \frac{1}{\sqrt{2}}(|00\rangle - |10\rangle).
\end{align*}
\]

If we apply it to the two first qubits of

\[
(\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle)
\]

we get

\[
\frac{1}{2} \left( \alpha(|000\rangle + |100\rangle) + \alpha(|011\rangle + |111\rangle) + \beta(|010\rangle - |110\rangle) + \beta(|001\rangle - |101\rangle) \right)
\]

\[
= \frac{1}{2} \begin{pmatrix} |00\rangle \otimes (\alpha|0\rangle + \beta|1\rangle) \\ + |01\rangle \otimes (\beta|0\rangle + \alpha|1\rangle) \\ + |10\rangle \otimes (\alpha|0\rangle - \beta|1\rangle) \\ + |11\rangle \otimes (\alpha|1\rangle - \beta|0\rangle) \end{pmatrix}
\]

If we measure the two first qubits, the third qubit becomes

\[
\begin{align*}
\alpha|0\rangle + \beta|1\rangle & \quad \text{if } 00 \text{ was measured}, \\
\beta|0\rangle + \alpha|1\rangle & \quad \text{if } 01 \text{ was measured}, \\
\alpha|0\rangle - \beta|1\rangle & \quad \text{if } 10 \text{ was measured}, \\
\alpha|1\rangle - \beta|0\rangle & \quad \text{if } 11 \text{ was measured}.
\end{align*}
\]

Finally note that if \(U_{xy}\) is applied in the case where \(x, y\) was measured, then the state of the last qubit is \(\alpha|0\rangle + \beta|1\rangle\).
3.4. Some Algorithms and Uses of Quantum Effects

3.4.2 Dense Coding

The teleportation algorithm allows to “send” one qubit using two classical bits. It is possible to perform the reverse operation, namely to send two classical bits using an entangled pair and a quantum channel.

The situation is similar. Alice and Bob share an entangled state of 2 qubits, Alice has qubit 1, Bob has qubit 2. This time, Bob wants to send the two classical bits \((x, y)\) to Alice. He can use the protocol summarized in Figure 3.3 called dense coding. Like the teleportation algorithm, the procedure is again in three steps:

1. At location B, Alice and Bob create an entangled state of two quantum bits 1 and 2, using the circuit

   \[ |0\rangle \rightarrow H |0\rangle \oplus |0\rangle. \]

   Then, Alice goes to location A with qubit 1, while Bob stays at location B with qubit 2.

2. If Bob wants to send two classical bits \((x, y)\) to Alice, he needs to apply the gate \(U_{x,y}\) as defined in the teleportation protocol (Equation 3.4.1), and send the modified qubit 2 to Alice.

3. Alice, to retrieve the two classical bits, applies a rotation on qubits 1 and 2 using the gate

   \[ \text{and a measurement, retrieving the two bits } (x, y). \]
3.4. Some Algorithms and Uses of Quantum Effects

3.4.3 The Deutsch Algorithm

The Deutsch algorithm is for finding out whether a boolean function is balanced or constant. In classical computation, two calls to the function are needed. In quantum computation, one can find it out in only one call. The algorithm takes as input an unknown circuit $U_f$ computing

$$U_f(|x \rangle \otimes |y \rangle) = |x \rangle \otimes |y \oplus f(x) \rangle$$

for some boolean function $f$, where $\oplus$ stands for the boolean addition. The quantum circuit for the algorithm is the following:

$$|0\rangle \xrightarrow{H} |0\rangle \xrightarrow{H} |x\rangle \xrightarrow{U_f} |0 \oplus f1\rangle$$

$$|1\rangle \xrightarrow{H} |y \oplus f0\rangle \xrightarrow{U_f} |1 \oplus f1\rangle,$$

where we write $fx$ in place of $f(x)$. To retrieve the answer, we have to measure the first qubit: it is 0 when the function is balanced and 1 when it is not. Note that the input of this algorithm is a “black-box”, in other terms a function from two qubits to two qubits.

Proof that the procedure is correct. This will compute the following thing:

$$(H \otimes id)(U_f)(H \otimes H)(|0 \rangle \otimes |1 \rangle)$$

$$= (H \otimes id)(U_f) \frac{1}{\sqrt{2}}(|0 \rangle + |1 \rangle) \otimes \frac{1}{\sqrt{2}}(|0 \rangle - |1 \rangle)$$

$$= \frac{1}{2}(H \otimes id)(U_f)(|00 \rangle + |10 \rangle - |01 \rangle - |11 \rangle)$$

$$= \frac{1}{2}(H \otimes id) \left( |0 \rangle \otimes |0 \oplus f0 \rangle + |1 \rangle \otimes |0 \oplus f1 \rangle - |0 \rangle \otimes |1 \oplus f0 \rangle - |1 \rangle \otimes |1 \oplus f1 \rangle \right)$$

$$= \frac{1}{2 \sqrt{2}} \left( |0 \rangle \otimes (|f0 \rangle + |f1 \rangle) + (|0 \rangle - |1 \rangle) \otimes (|f1 \rangle - |f0 \rangle) \right)$$

$$= \frac{1}{2 \sqrt{2}} \left( |0 \rangle \otimes (|f0 \rangle + |f1 \rangle) - (|0 \rangle - |1 \rangle) \otimes (|f1 \rangle - |f0 \rangle) \right).$$

If $f0 = f1$, then $|f0 \rangle - |f1 \rangle - |1 \oplus f0 \rangle + |1 \oplus f1 \rangle = 0$ and the result is

$$\frac{1}{\sqrt{2}} |0 \rangle \otimes (|f1 \rangle - |1 \oplus f1 \rangle).$$

If $f0 = 1 \oplus f1$, then $|f0 \rangle + |f1 \rangle - |1 \oplus f0 \rangle - |1 \oplus f1 \rangle = 0$ and the result is

$$\frac{1}{\sqrt{2}} |1 \rangle \otimes (|1 \oplus f1 \rangle - |f1 \rangle).$$

So the value of the measurement of the first qubit is 0 if the function $f$ is balanced, and 1 in the other case. \qed
Chapter 4

A Tour in Existing Models of Quantum Computation

Chapter 3 developed the formalism and the techniques needed to perform quantum computation. This chapter surveys a few selected models for manipulating quantum programs and protocols. A more complete set can be found in e.g., (Gay, 2006).

The goal is twofold. First, we want to be able to write programs and codes manipulating quantum data. Numerous attempts have been done in the past years to provide a more easy way to write (and read) quantum algorithms. Second, we aim at a deeper understanding of the structure of quantum information by developing a semantics for quantum computation.

4.1 The Various Paradigms

For building languages and understanding algorithms, several paradigms of computation have been studied.

4.1.1 Unitary Gates as Computation.

Families of Quantum Circuits

One formal system for writing quantum algorithms is the quantum circuit formalism described by Deutsch [1989] and sketched in Section 3.2.2. In this formalism, the driving forces in the computation are the unitary operations. An algorithm is composed of three steps:

1. allocation of \( n \) quantum bits;
2. application of a list of unitary gates;
3. measurement of all or a part of the \( n \) quantum bits.

The description of an algorithm is a family of quantum circuits, one for each possible number of input qubits. The family needs to be uniform, i.e., efficiently computable by a classical computer. As for classical boolean circuits, it is then possible to speak of the complexity of an algorithm by taking the limit of the ratio of the number of gates used in the circuit versus the number of input quantum bits.
The Quantum Turing Machine

A Turing machine is an automaton consisting of the following elements:

1. an infinite tape, divided into cells, each cell containing a letter from a given alphabet;
2. a head, able to read and write into cells, and to move to the left or to the right along the tape;
3. a finite set of states and a state register, recording the state of the machine. One of the states is called initial and a subset of the states are called final;
4. a transition table, mapping each possible pair (letter, state) to a triple consisting of a new state, a letter and a direction (left or right).

At the beginning the state register is in the initial state. At each time-step, the automaton reads the cell of tape below the head. Based on the letter read and its state, using the transition table, it writes on the tape the corresponding letter and moves in the given direction. It iterates this course of action until a final state is reached. A program consists of a given alphabet, a set of states and a transition table.

Benioff (1980) and Deutsch (1985) describe a Turing machine where the tape, the head, and the states of the automaton are encoded as a quantum state. The transition table is described in a unitary operation. This model provides a tool to reason about expressiveness and complexity of quantum algorithms: As was shown by Nishimura and Ozawa (2002), the resulting notion of computation is equivalent to that of the quantum circuit model.

QML or the Quantum Test

Altenkirch and Grattage (2005a,b) describe a first-order language for manipulating quantum bits. This calculus understands duplication as sharing, or copying:

$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha|00\rangle + \beta|11\rangle.$$  

The language is written on a classical device and compiles into a quantum circuit. Due to the copying operation, contraction can be done blindly but weakening (that is, measurement) has to be kept track of.

The language features superposition of quantum states, leading to a “quantum” test. For example, the gate $\mathbf{N}_C$ is defined as:

$$\mathbf{N}_C \langle c, x \rangle = \begin{cases} 
\text{if}^q x \text{ then } \langle c, \mathbf{N} x \rangle \text{ else } \langle c, x \rangle. 
\end{cases}$$

The main drawback of this formalism is that the quantum test

$$\text{if}^q x \text{ then } \mathbf{M} \text{ else } \mathbf{N}$$

is well defined if and only if $\mathbf{M}$ and $\mathbf{N}$ have orthogonal representations. Therefore, in general one cannot say before running the computation whether a program is valid or not.

Van Tonder’s Quantum Lambda Calculus

Van Tonder (2004) describes a lambda calculus where the terms are encoded on strings of quantum bits. The language is built using the terms

$$\text{Term } M, N ::= x \mid c \mid MN \mid \lambda x. M,$$

$$\text{Constant } c ::= 0 \mid 1 \mid H \mid \cdots,$$
where the constants $c$ contains booleans 0 and 1, as representatives of the quantum bits $|0\rangle$ and $|1\rangle$, and representatives for a chosen set of unitary gates. For example, the identity is $|\lambda x.x\rangle$. The unitaries apply on 0 and 1 in the naive way: $|H0\rangle$ should reduce to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. This language is purely quantum, and the reduction is encoded as a unitary map.

A trick needs to be used to overcome the fact that the reduction is not in general a reversible operation (Consider for example $(\lambda x.1)0 \rightarrow 1$). Given any function $\beta : X \rightarrow Y$, it is possible to produce a reversible map $X \rightarrow X \times Y$ invertible on its range by mapping $x \mapsto (x, \beta(x))$. If $\beta$ is a reduction of terms, we can make it reversible by appending to the successive terms in the reduction its history:

$$M \rightarrow (M, \beta(M)) \rightarrow (M, \beta(M), \beta^2(M)) \rightarrow \cdots$$

However, a naive implementation of the lambda calculus into quantum bits does not provide a consistent equational theory. Indeed the term $\frac{1}{\sqrt{2}}(|(\lambda x.0)0\rangle + |(\lambda x.0)1\rangle) = |\lambda x.0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ should be $\beta$-equivalent to $\frac{1}{\sqrt{2}}(|0\rangle + |0\rangle) = \sqrt{2}|0\rangle$, which is not valid since its norm is greater than 1.

As it turns out, one needs to distinguish between terms that are in superposition (such as $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and the ones that are not. For the system to be consistent, terms in superposition should not be duplicated nor erased, and should be the same, modulo the 0’s and the 1’s.

### 4.1.2 Concurrent Quantum Computation

Quantum computation has important implications in cryptography. The design of quantum protocols to securely communicate data is a field of ongoing research.

From a theoretic side, a question one might ask is whether the tools that were developed to analyse the safety of classical distributed systems can be used in the context of quantum systems. The simultaneous works of Lalire and Jorrand (2004) and Gay and Nagarajan (2005) develop a process algebraic approach to concurrent quantum computation. Starting from the $\pi$-calculus, quantum features, such as quantum bits, unitaries and measurements, are added to the syntax, and an operational semantics is developed. Gay and Nagarajan (2005) gives a typed language and prove some safety properties for well-typed protocols; for example the fact that a quantum bit cannot be shared among several processes, and the soundness of the type system.

### 4.1.3 Measurement-Based Quantum Computation

The unitary operations are one side of the quantum coin. The other side is the measurement. It turns out that it is possible to perform quantum computation with “only” measurements, as it was described first by Briegel and Raussendorf (2002); Raussendorf and Briegel (2001) and then computationally by Danos et al. (2007).

The idea is the following: A state of highly entangled quantum bits is prepared and then a series of measurements are performed. The computation is “run” by the projections occurring during the measurements. In its most general formulation, the calculus happens to be as powerful as the regular quantum-circuit computational model, and is slightly better in terms of time complexity.

Formally, one works with a pattern, composed of a finite set $V$, a set of inputs $I \subset V$, a set of outputs $O \subset V$, and a sequence of operations of the following kinds (where $i, j \in V$):

- 1-quantum measurements $M^i_j$: measurement of qubit $i$ in the $xy$-plane, along the axes

$$\frac{1}{\sqrt{2}}(|0\rangle + e^{i\alpha}|1\rangle), \quad \frac{1}{\sqrt{2}}(|0\rangle - e^{i\alpha}|1\rangle).$$
4.2. Formalism of Hilbert Spaces

A recent research thread initiated by Abramsky and Coecke (2004) aims at providing an abstract framework in which one could interpret quantum computation. The goal is to extract from the
4.2. Formalism of Hilbert Spaces

category of Hilbert spaces the required structures for performing quantum computation.

4.2.1 Dagger Compact-Closed Categories

Abstract notions of adjointness, self-adjointness, unitaries, bases and scalar products arise in the context of dagger compact-closed categories with biproducts:

Definition 4.2.1 (Selinger, 2005). A dagger category is a category $C$ together a functor $\dagger : C^{\text{op}} \to C$ such that for every $f : A \to B$, $f^{\dagger\dagger} = f$ and for every object $A$, $A^{\dagger} = A$.

Definition 4.2.2 (Selinger, 2005). A dagger compact-closed category is a dagger symmetric monoidal category that is also compact closed (see e.g. Kelly and Laplaza, 1980), and such that the following diagram commutes for all $A$:

\[
\begin{array}{c}
\top \\
\downarrow \eta_A \\
A \otimes A^* \\
\downarrow \sigma_{A,A^*} \\
A^* \otimes A.
\end{array}
\]

Various notions of Hilbert spaces can be defined in the context of dagger compact-closed categories with biproducts:

- A scalar is a morphism $s : \top \to \top$.
- A vector is a morphism $v : \top \to A$.
- The adjoint of a morphism $f : A \to B$ is the morphism $f^{\dagger} : B \to A$.
- The scalar product of $v, w : \top \to A$ is $v; w^{\dagger} : \top \to \top$.
- A unitary morphism is a morphism $f : A \to B$ such that $f^{\dagger}$ is its inverse.
- A basis for an object $A$ is a unitary morphism $b_A : n\cdot\top \to A$,

where $n\cdot\top$ is a sum of $n$ $\top$’s. In this case, the dimension of $A$ is said to be $n$.

- We can define the matrix of a morphism $f : A \to B$ where $A$ and $B$ are finite dimensional as the map $n_{A} : A_{\top} \to A$.

These definitions satisfy various equations related to Hilbert spaces and hermitian positive operators. For example, (Selinger, 2005) shows that completely positive maps can be defined in this context. The paper provides a categorical construction for building a category of completely positive maps out of any dagger compact-closed category.

Although no formal syntax comes with the semantics, a colourful graphical language is provided, making proofs of properties relatively easy to follow. However, this graphical language does not mix well with the need for biproducts: in the desired quantum computational applications, the notion of bases is heavily used.
4.2.2 Classical Objects

In an attempt to remove biproducts, Coecke and Pavlovic (2007) describe a new notion called classical objects. A classical object is an object $X$ in the category together with a map $\delta_X : X \to X \otimes X$, a map $\epsilon_X : X \to \top$, and a few equations. In particular, these equations make $(X, \epsilon_X, \delta_X)$ a commutative comonoid object, as in Definition 2.7. The notion of classical object captures the notion of basis without the needs for biproducts. The idea is the following. A generic object $X$ in the category is like a generic finite dimensional Hilbert space, with no associated basis. Associating a structure $(\delta_X, \epsilon_X)$ to the object $X$ is the same thing as associating a basis to $X$. In Hilbert spaces, the maps $\delta_X$ and $\epsilon_X$ are

$$
\delta_X : e_i \mapsto e_i \otimes e_i, \quad \epsilon_X : e_i \mapsto 1,
$$

for some basis $(e_1, \ldots, e_n)$ of $X$. With enough structure, it is possible to simulate what was done with biproducts. In particular, as was shown by Coecke and Paquet (2006), an abstract version of the Naimark theorem can be proved using the graphical language.

4.3 A Flow-Chart Language

Selinger (2004b) describes a semantics for QRAM-based first-order quantum computation. Instead of starting from the category of Hilbert spaces, the paper describes computations as maps in the category of superoperators.

4.3.1 The language

The language is described by a flow-chart notation: it is a super-set of a classical flow-chart language. A program is a graph together with a cursor that follows the wires, with data attached to it. A graph is constructed from the rules in Table 4.1. The box new qbit $b = i$ creates a new quantum bit of name $b$ with value $\lvert i \rangle$. The box new bit $b = i$ creates a new boolean variable named $b$ with value $i$. The gate $q_1, \ldots, q_n \ast U_n$ applies the unitary gate $U_n$ on the corresponding variables. The gate permute $\phi$ permutes the context $\Delta$ according to the permutation $\phi$. The gate branch $b$ performs a test on the boolean $b$ and meas $q$ measures the quantum bit $q$ and performs a test on the output.

4.3.2 The Category of Superoperators

The language has a full and complete representation in the category $Q$ of superoperators. This category is defined as the refinement of the category $CPM$.

**Definition 4.3.1.** The category $CPM$ is defined as follows:

- The objects are signatures $\sigma = n_1, \ldots, n_k$, i.e., finite tuples of positive integers,
- the morphisms $\sigma \to \tau$ are completely positive maps $V_\sigma \to V_\tau$, where $V_{n_1,\ldots,n_k}$ is the Hilbert space $\mathbb{C}^{n_1 \times n_1} \times \cdots \times \mathbb{C}^{n_k \times n_k}$.

We use the notation $B_\sigma$ for the canonical basis of $V_\sigma$.

The objects in this category can be seen as spaces of block matrices. Following the construction in Lemma 3.1.13, one can define a tensor and a coproduct structure, as follows:

**Definition 4.3.2.** Given two signatures $\sigma = (n_1, \ldots, n_k)$ and $\tau = (m_1, \ldots, m_l)$, we define $\sigma \otimes \tau$ and $\sigma \oplus \tau$ to be respectively

$$
\sigma \otimes \tau = (n_1 m_1, \ldots, n_1 m_l, n_2 m_1, \ldots, n_k m_l), \quad \sigma \oplus \tau = (n_1, \ldots, n_k, m_1, \ldots, m_l).
$$
Lemma 4.3.3. Let \( \sigma \) and \( \tau \) be two signatures. Then we have \( V_{\sigma \otimes \tau} \cong V_\sigma \otimes V_\tau \) and \( V_{\sigma \oplus \tau} \cong V_\sigma \oplus V_\tau \).

Definition 4.3.4. We define a map between arrows
\[
\Phi : \text{CPM}(\sigma \otimes \tau, \sigma') \xrightarrow{\cong} \text{CPM}(\sigma, \tau \otimes \sigma')
\]
as in Lemma 4.3.3 using the canonical basis.

Lemma 4.3.5. The category \( (\text{CPM}, \otimes, (1)) \) is a symmetric monoidal category. The map \( \Phi \) makes \( \text{CPM} \) symmetric monoidal closed. The structure \( \oplus \) and the signature \( (0) \) provide a structure of biproducts.

Definition 4.3.6. Given a signature \( \sigma = (n_1, \ldots, n_k) \), a tuple \( A = (A_1, \ldots, A_k) \) in \( V_\sigma \) is hermitian (respectively positive) if each matrix \( A_i \) is hermitian (respectively positive). We extend the Löwner order of Definition 3.1.25 to tuples of matrices by letting \( A \sqsubseteq B \) if \( B - A \) is positive. We define the trace of the matrix tuple \( A \) to be the sum of the traces of its components: \( \text{tr}(A) = \sum \text{tr}(A_i) \).

Definition 4.3.7. Let \( \sigma = (n_1, \ldots, n_k) \) and \( \tau = (m_1, \ldots, m_l) \) be two signatures. A linear function \( F : V_\sigma \to V_\tau \) is a tuple \( (F_{1,1}, \ldots, F_{1,k}, \ldots, F_{n,1}, \ldots, F_{n,k}) \), where \( F_{i,j} \) is the component \( \mathbb{C}^{n_i} \to \mathbb{C}^{m_j} \). The characteristic matrix tuple \( \chi_F \) of the function \( F \) is
\[
(\chi_{F_{1,1}}, \ldots, \chi_{F_{1,k}}, \ldots, \chi_{F_{n,1}}, \ldots, \chi_{F_{n,k}}).
\]
Note that since \( F : V_\sigma \to V_\tau, \chi_F \in V_{\sigma \otimes \tau} \).

Definition 4.3.8. The category \( Q \) has the same objects as \( \text{CPM} \) and as arrows superoperators. A superoperator is, as in Definition 4.3.8, a completely positive map \( f : \sigma \to \tau \) such that for all \( A \in V_\sigma, \text{tr}(f(A)) \leq \text{tr}(A) \).
4.3. A Flow-Chart Language

Lemma 4.3.9. The category $\mathcal{Q}$ is symmetric monoidal and has coproducts, with the same construct as for CPM. It is not, however, closed.

Definition 4.3.10. The trace characteristic matrix tuple of a linear function $F : V_\sigma \to V_\tau$ is defined to be $\chi^{(\text{tr})}_F = \chi_{\text{tr} \circ F} \in V_\sigma$, that is, the characteristic matrix tuple of $\text{tr}_\tau \otimes F$. Equivalently, $\chi^{(\text{tr})}_F = (\text{id}_\sigma \otimes \text{tr}_\tau)(\chi_F)$.

Completely positive maps and superoperators in this generalized context still satisfy a version of Theorem 3.1.27.

Theorem 4.3.11 (Generalized Kraus Representation Theorem). Consider the signatures \(\sigma\) of the form \((n_1, \ldots, n_s)\) and \(\tau\) of the form \((m_1, \ldots, m_t)\). Let \(F : V_\sigma \to V_\tau\) be a linear function. Then the following holds:

1. \(F\) is completely positive if and only if \(\chi_F\) is positive.
2. \(F\) is a superoperator if and only if \(\chi_F^{(\text{tr})} \subseteq I_\sigma\), where \(I_\sigma \in V_\sigma\) is the tuple consisting of identity matrices.
3. \(F\) is a superoperator if and only if it can be written in the form

\[
F(A_1, \ldots, A_s) = \left( \sum U_{i1} A_i U_{i1}^*, \ldots, \sum U_{i1} A_i U_{i1}^* \right),
\]

for matrices $U_{ij} \in \mathbb{C}^{m_j \times n_j}$, where $\sum_{j,l} U_{ij}^* U_{ijl} \subseteq I_n$ for all $i$. Here, $l$ ranges over a finite set. $\blacksquare$

4.3.3 Interpretation of the Flow-Chart Language

It is possible to interpret a flow chart diagram as a superoperator. We state in this section the few results of [Selinger 2004] used in Chapter 7.

Definition 4.3.12. Let \(\sigma = (n_1, \ldots, n_s)\) be a signature, and let \(\sigma' = n_1 + \ldots + n_s\) be an integer regarded as a simple signature. The measurement operator \(\mu_\sigma : V_\sigma \to V_\sigma\) is defined as

\[
\mu_\sigma \left( \begin{array}{cccc}
A_{1,1} & \cdots & A_{1,s} \\
\vdots & \ddots & \vdots \\
A_{s,1} & \cdots & A_{s,s}
\end{array} \right) = (A_{1,1}, A_{2,2}, \ldots, A_{s,s}),
\]

where $A_{ij} \in \mathbb{C}^{n_i \times n_j}$.

Definition 4.3.13. We extend the definition of partial trace of Definition 3.1.31 to any map of the form $(\text{tr}_\tau \otimes \text{id}_\sigma) : V_\sigma \otimes V_\tau \to V_\tau$.

Definition 4.3.14. We extend the definition of contractions of Definition 5.1.28 to linear maps. We say that $f : V_\sigma \to V_\tau$ is a contraction if it is of the form

\[
f(A_1, \ldots, A_s) = (U_1 A_1 U_1^*, \ldots, U_s A_s U_s^*)
\]

for contractions $U_i \in \mathbb{C}^{n_i \times n_i}$, where $\sigma = (n_1, \ldots, n_s)$ and $\tau = (m_1, \ldots, m_s)$.

Theorem 4.3.15. Every superoperator $F : V_\sigma \to V_\tau$ can be factored as $F = M \circ E \circ G$, where $G$ is a contraction, $E$ is a partial trace operator, and $M$ is a measurement operator. $\blacksquare$

Theorem 4.3.16. The interpretation of any flow-chart in CPM yields a superoperator. Conversely, any superoperator is the image of a flow chart diagram under the interpretation. $\blacksquare$
4.4 Extension to Higher-Order

A natural avenue of research for finding a model of higher-order quantum computation is the generalization of the category of superoperators. Selinger (2004c) investigated possible categories of normed vector spaces as model of higher order quantum computation. A challenging problem turns out to be the matching between the norm for the tensor of two spaces and the required norm on the space of functions. This study did not find any satisfactory model.

Another work in this direction is the paper of Girard (2004). In this paper, the author describes a models of linear logic based on normed vector spaces, following the same constructions as for coherent spaces (Girard, 1987). The resulting category, although based on the right sort of spaces, does not yield the correct answer at base types, as pointed out by Selinger (2004c).
Chapter 5

Lambda Calculus and Semantics of Higher-Order Computation

In this chapter I will briefly describe the semantics for classical higher-order computation. The canonical tool for this purpose is a semantics, a mathematical model that will allow us to exhibit underlying structures and invariants.

For an complete introduction on lambda calculus, see (Barendregt, 1984). An in-depth survey of categorical logic can be found in (Lambek and Scott, 1989). Finally, references on linear logic and its semantics can be found in (Bierman, 1995; Girard, 1987; Troelstra, 1992).

5.1 Lambda Calculus

The origin of higher-order computation lies in the 1930’s when Kleene (1935a,b) described a formal system called lambda calculus in order to capture the notion of computable function. Lambda calculus was shown (Church, 1936) universal in the sense that any computable function can be interpreted as an expression in the language.

In this section we briefly describe the core lambda calculus. For a thoughtful development of the theory see for example (Barendregt, 1984).

5.1.1 The Language

In the lambda calculus, functions are described explicitly as formulae. The function \( f : x \mapsto f(x) \) is written as the lambda-abstraction \( \lambda x.f(x) \). The application of a function to an argument is simply the juxtaposition of the two expressions. Thus

“ let \( f : x \mapsto x^2 \) in \( f(2) \) ”

becomes \( (\lambda x.x^2)(2) \) in lambda calculus.

Formally, a lambda-term is an expression made of the following grammar, written in BNF form:

\[
Term \ M, N ::= \ x \mid (\lambda x.M) \mid (MN) \mid \langle M, N \rangle \mid \text{fst}(M) \mid \text{snd}(M) \mid *,
\]

where \( x \) ranges over an infinite set of variables. The term \( \lambda x.M \) stands for the function that inputs \( x \) and outputs \( M \), the term \( MN \) represents the application of a function \( M \) to an argument \( N \), the term \( \langle M, N \rangle \) represents the pair composed of \( M \) and \( N \), \( * \) is the 0-tuple, and \( \text{fst}(M) \) and \( \text{snd}(M) \)

\footnote{The BNF formulation was first introduced by Naur et al. (1956) for a formal definition of the language ALGOL’60.}
the left and right projections of the pairing. We speak of term combinator or term operator when referring to each token in the grammar definition.

The formalism extends transparently to higher-order functions: In lambda calculus, we can simply write \( \lambda f.\lambda g.\lambda x.f(g(x)) \) for the composition operator \( f, g \mapsto f \circ g \). And the expression

\[
\text{"let } H : (f, g) \mapsto (x \mapsto f(g(x))) \text{ in }
\text{let } a : x \mapsto x^2 \text{ in }
\text{let } b : x \mapsto x + 3 \text{ in } (H(a,b))(z) \quad \]

for some variable \( z \) simply becomes

\[
(\lambda f.\lambda g.\lambda x.f(g(x)))(\lambda x.x^2)(\lambda x.x + 3)(z). \tag{5.1.1}
\]

5.1.2 Free and Bound Variables

In (5.1.1), in the term \( (\lambda f.\lambda g.\lambda x.f(g(x))) \) it should be understood that \( x \) is local, and that the \( x \) in \( (\lambda x.x^2) \) does not interfere with the one of the previous formula. We say that the variable \( x \) in \( \lambda x.x^2 \) is bound by the lambda-abstraction. On the other hand, the variable \( z \) is not bound by anything: we say it is free. We define the set of free variables of a given term by induction as follows:

\[
\begin{align*}
FV(x) & = \{x\}, \\
FV(\lambda x.M) & = FV(M) \setminus \{x\}, \\
FV(MN) & = FV(M) \cup FV(N), \\
FV(\text{fst}(M)) & = FV(\text{snd}(M)) = FV(M), \\
FV(*) & = \{\}. 
\end{align*}
\]

A variable that is not free is said to be bound. A term is called closed if it does not have any free variable, and open if it has at least one free variable.

5.1.3 \( \alpha \)-Equivalence

It should be understood that \( \lambda x.x^2 \) is the same term as, say, \( \lambda t.t^2 \). We say that the two terms are \( \alpha \)-equivalent. However, \( \lambda f.\lambda g.\lambda x.f(g(x)) \) and \( \lambda g.\lambda g.\lambda x.g(g(x)) \) are not \( \alpha \)-equivalent: the \( g \) in the expression refers to the innermost \( \lambda g \).

**Definition 5.1.1** [Barendregt, 1984]. For any term \( M \), we define the syntactic variable substitution \( M[x := y] \) as the term \( M \) where all syntactic occurrences of \( x \) (whether they are free or not) are replaced by \( y \). We define a change of bound variables in \( M \) to be the replacement of a subterm \( \lambda x.N \) by \( \lambda t.N[t/x] \), provided that \( t \) does not occur anywhere in \( N \).

We say that \( M \) is \( \alpha \)-equivalent to \( N \), written \( M \equiv_{\alpha} N \), if \( N \) results from \( M \) by a series of changes of bound variables.

**Convention 5.1.2.** Terms are considered up to \( \alpha \)-equivalence from now on.

5.1.4 Operational Meaning of Lambda Calculus

We now discuss the operational meaning of the calculus by defining a rewriting system.

**Substitution**

A term of the form \( (\lambda x.M)N \) is understood as “\( M \) with \( N \) in place of \( x \)”. The process of replacing \( N \) by \( x \) is called substitution, and we write the resulting term by \( M[N/x] \). Of course, we are only allowed to modify the free occurrences of \( x \), and one should check that the process does not bind any free variables. For example, the term \( (\lambda x.yx)((fx)/y) \) is not the same as \( \lambda x.(fx)x \): indeed, the variable \( x \) in \( fx \) is free. Instead, using the \( \alpha \)-equivalent formulation \( (\lambda y.t)((fx)/y) \) one sees that the term is in fact \( \lambda y.(tx)t \).
Definition 5.1.3. Given a term $M$ and a term $P$, we define the substitution of $x$ in $M$ by $P$, written $M[P/x]$, by the following:

\[
x[P/x] = P,
\]

\[
\ast[P/x] = \ast,
\]

\[
(MN)[P/x] = (M[P/x])(N[P/x]),
\]

\[
((M,N))[P/x] = (M[P/x],N[P/x]),
\]

\[
(fst(M))[P/x] = fst(M[P/x]),
\]

\[
(snd(M))[P/x] = snd(M(P/x)),
\]

\[
(\lambda y.M)[P/x] = \lambda y.(M[P/x]) \quad \text{(if } y \neq x \text{ and } y \notin FV(P)).
\]

We write $M[M_1/x_1, \ldots, M_n/x_n]$ for $M[M_1/x_1] \ldots [M_n/x_n]$.

Reduction of Terms

The simply-typed lambda calculus is a language describing operations: application of a function to an argument, projection of a pair. To express these operations we develop a set of rules, called reduction rules. They come in two flavors: the $\beta$-reductions, which perform operations and enforce the implicit meaning of the term:

\[
(\lambda x.M)N \rightarrow_\beta M[N/x], \quad (\beta_\lambda)
\]

\[
fst(M,N) \rightarrow_\beta M, \quad (\beta_1^\top)
\]

\[
snd(M,N) \rightarrow_\beta N, \quad (\beta_2^\top)
\]

and $\eta$-reductions simplifying terms:

\[
\lambda x.(Mx) \rightarrow_\eta M, \quad (\eta_\lambda)
\]

\[
(fst(M), snd(M)) \rightarrow_\eta M. \quad (\eta_0)
\]

Each term on the left side of these rule is called a redex.

A context is a “term with a hole”, defined formally by the syntax

\[
C[-] ::= [-] \mid (\lambda x.C[-]) \mid (C[-]M) \mid (MC[-]) \mid (M,C[-]) \mid (C[-], M) \mid fst(C[-]) \mid snd(C[-]),
\]

where $M$ stands for any term. Given a term $N$ and a context $C[-]$, we write $C[N]$ for the context $C[-]$ where $[-]$ is syntactically replaced with $N$. In this setting, we naturally extend the notion of reduction, as follows: Given any context $C[-]$, if $N$ is a redex reducing to $N'$, then the reduction $\rightarrow_\iota$ extends to $C[N] \rightarrow_\iota C[N']$, for $i = \beta, \eta$. A term with no subterm in redex form is said to be in normal form.

We denote by $\rightarrow^*$ the reflexive transitive closure of the relation $\rightarrow$. The relation $\rightarrow^*_\beta$ satisfies a property called the Church-Rosser theorem that we only mention here. For a proof of this result consult e.g. [Barendregt, 1984].

Theorem 5.1.4 (Church-Rosser theorem). Suppose that $M$, $N$ and $P$ are lambda-terms such that $M \rightarrow_\beta^* N$ and such that $N \rightarrow_\beta^* P$. Then there exists a lambda-term $Z$ such that $N \rightarrow_\beta^* Z$ and $P \rightarrow_\beta^* Z$.

5.1.5 Typed Lambda Calculus

In mathematics, to a function one usually associates a domain and a codomain. So for example, when we mention the identity function $id$, it is always associated to a given space. Although one may lose some generality, we gain in legibility.
Type System

Similarly, in order to be able to manipulate lambda-terms more easily, one can associate a type to each lambda-term. A type is an expression providing some information about the term. One can associate a type system to the untyped lambda calculus defined in Section 6.1 as follows:

\[
\text{Type } A, B \ ::= \ \iota | \top | A \times B | A \Rightarrow B.
\]

We choose an arbitrary constant type \(\iota\), and types are constructed by induction. If \(A\) and \(B\) are types, then \(A \times B\) is the type for binary products, and \(A \Rightarrow B\) is the type for functions from \(A\) to \(B\). Finally, \(\top\) stands for the type of the 0-tuple.

A typed term is the data consisting of the term and its type. The type of an open term is a function of the type of its free variables. For example, the type of \(\lambda x.(yx)\) should be a function type \(A \Rightarrow B\), depending on the type of \(y\). We define the notion of a typing judgement to formalize the concept. A typing judgement is an expression \(\Delta, x_1 : A_1, \ldots, x_n : A_n \triangleright M : B\) where the free variables of \(M\) are contained in \(\{x_1, \ldots, x_n\}\), and \(B\) is the type of \(M\) when each \(x_i\) has type \(A_i\). The set \(\{x_1 : A_1, \ldots, x_n : A_n\}\) is called the typing context of the judgement, and it might be empty.

Typing Rules

The typed version of the lambda calculus presented in Section 6.1 is called the simply-typed lambda calculus. A typing judgement is said to be valid if it is derived from a typing derivation, or typing tree. A valid typing derivation is constructed using the following inductive rules:

\[
\begin{align*}
\Delta, x : A \triangleright x & \quad \text{(ax)}, \\
\Delta, x : A \triangleright M & \quad \Delta \triangleright \lambda x.M & \quad \text{(\(\lambda\))}, \\
\Delta \triangleright M & \quad \Delta \triangleright N \quad \Delta \triangleright N : A & \quad \text{(app)}, \\
\Delta \triangleright (M, N) & \quad \Delta \triangleright A & \quad \text{(\(\times\))}.1, \\
x : A \triangleright \top & \quad \text{(\(\top\)),I}, \\
\Delta \triangleright M & \quad \Delta \triangleright A \times B & \quad \text{(\(\times\))E1}, \\
\Delta \triangleright \text{fst}(M) & \quad \Delta \triangleright A & \quad \text{(\(\times\))E2}.
\end{align*}
\]

Properties of Typed Terms

On one hand, we have a description of some intended properties of the language through a type system: for example, a term \(M\) of type \(A \Rightarrow B\) can only by applied to an argument of type \(A\). On the other hand, we have a rewriting system by means of reduction rules, accounting for the operational behavior of the language. Are these two descriptions compatible?

The main results that should be satisfied (and are verified in the case of the simply-typed lambda calculus) are the following.

**Lemma 5.1.5** (Substitution). If \(\Delta \triangleright N : A\) and \(\Delta, x : A \triangleright M : B\) are valid typing judgements, then \(\Delta \triangleright M[\text{\(N/x\)}] : B\) is valid.

**Theorem 5.1.6** (Subject reduction). If \(\Delta \triangleright M : A\) and if \(M \rightarrow_{\beta\eta} M'\) then \(\Delta \triangleright M' : A\) is valid.

They state that the type of a term remains constant along the reduction process. A third important result states the consistency of well-typed closed terms, in the sense that a well-typed closed term will reduce until reaching a normal form. The result is stated as follows:
5.2. Proofs as Computations

Theorem 5.1.7 (Progress). If $M$ is a closed, well-typed term, then either $M$ is a normal form or $M \rightarrow_{\beta} M'$.

Axiomatic Equivalence

We say that two valid typing judgements $\Delta \vdash M : A$ and $\Delta \vdash N : A$ are axiomatically equivalent, written $\Delta \vdash M \approx_{ax} N : A$, if the relation can be derived from

- $\Delta \vdash (\lambda x. M) N \approx_{ax} M[N/x] : A$,
- $\Delta \vdash \lambda x. (M x) \approx_{ax} M : A \rightarrow B$,
- $\Delta \vdash \text{fst}(M, N) \approx_{ax} M : A$,
- $\Delta \vdash \text{snd}(M, N) \approx_{ax} N : B$,
- $\Delta \vdash (\text{fst}(M), \text{snd}(M)) \approx_{ax} M : A \times B$,
- $\Delta \vdash x \approx_{ax} \ast : \top$,

and from one congruence rule per term construct, such as

$$\frac{\Delta \vdash M \approx_{ax} M' : A \Rightarrow B \quad \Delta \vdash N \approx_{ax} N' : A}{\Delta \vdash MN \approx_{ax} M'N' : B}$$

Curry Style Versus Church Style

While defining the type system, we wrote “A typed term is the data consisting of the term and its type.”

There are actually two distinct ways of typing a term. One can adopt the Church-style typing convention, and consider that all of its subterms are typed. Or one can adopt the Curry-style typing convention, and see only the outside term as typed. From the point of view of the programmer, it is simpler to type a lambda-term in a Curry-style fashion. However, to be able to check whether a typed term is valid or not it is easier to have it Church-style typed.

These two ways of typing terms might or might not be equivalent: it is a question of understanding the meaning of the term. For example, consider the Curry-style typing judgement

$$x : A \vdash (\lambda f.x)(\lambda y.y) : A.$$ 

The type of the variable $x$ is given. However, the type of the variable $y$ is not mentioned anywhere and might be anything. For each possible type $B$, the typing derivation

$$\frac{x : A, f : B \Rightarrow B \Rightarrow x : A}{x : A \vdash \lambda f.x : (B \Rightarrow B) \Rightarrow A} \quad \frac{x : A \vdash \lambda y.y : B \Rightarrow B}{x : A \vdash (\lambda f.x)(\lambda y.y) : A}$$

is valid. Therefore the data consisting of a set of typed variables, a term and a type does not uniquely determine a typing derivation. In a Curry-style type system, the axiomatic equivalence is really defined on typing derivations, and not typing judgements.

Obviously, in this simple example the term reduces to $x$, and thus all possible typing derivations live in the same axiomatic equivalence class. But in some more complicated language, this might not be the case, and the correspondence between typing derivations and typing judgements might not be one-to-one.

5.2 Proofs as Computations

Let us consider a typing derivation. We can concentrate on the meaning of terms, and understand the typing tree as a computation. But it is also possible to focus on the types, and see the typing tree as a proof in some logical system.
5.2. Proofs as Computations

A sequent in intuitionistic logic is a pair \( (\Gamma, B) \), written as \( \Gamma \vdash B \), where \( \Gamma \) is a multiset (possibly empty) of formulae \( \{ A_1, \ldots, A_n \} \) and \( B \) is a formula. We read the sequent as “\( A_1 \) and \( A_2 \) and... and \( A_n \) implies \( B \)”, and we call \( \Gamma \) the context of the sequent.

A sequent is called valid if it can be derived from a proof. Proofs are trees of formulae defined inductively. In intuitionistic logic a proof is said to be constructive: the proof of \( A \times B \) should include the proof of \( A \) and the proof of \( B \). The proof of \( A \Rightarrow B \) should be a map from the proofs of \( A \) to the proofs of \( B \).

We will concentrate on natural deduction style proof rules, first introduced by Gentzen (1934), where we mainly deal with the right side of the sequents. The proofs associated with the formulae presented above are constructed using the rules in Table 5.1. The rules come in two flavors:

1. the axiom \( \text{Ax} \) and two structural rules: the weakening rule \( W \) and the contraction rule \( C \),

2. introduction and elimination for each connective (except for the null-ary connective \( \top \) which has only an introduction rule).

Table 5.1: Intuitionistic logic: natural deduction rules

\[
\begin{align*}
\Gamma \vdash A & \quad \Gamma \vdash B \\
\Gamma \vdash A \times B & \quad \Gamma \vdash A \\
\Gamma, A \vdash B & \quad \Gamma, B \vdash A \\
\Gamma \vdash B & \quad \Gamma, A \vdash B \quad \Gamma, A, A \vdash B \\
\Gamma \vdash A & \quad \Gamma \vdash B \quad \Gamma, A \vdash B \\
\end{align*}
\]

5.2.1 Intuitionistic Logic

The logic behind the simply typed lambda calculus described in Section 5.1.5 is called intuitionistic logic. Formulae are constructed as expressions coming from the grammar

\[
\text{Formula } A, B ::= \iota | \top | A \times B \mid A \Rightarrow B,
\]

where \( \iota \) ranges over a set of constants, \( \times \) is the conjunction, and \( \Rightarrow \) is the implication.

A sequent in intuitionistic logic is a pair \( (\Gamma, B) \), written as \( \Gamma \vdash B \), where \( \Gamma \) is a multiset (possibly empty) of formulae \( \{ A_1, \ldots, A_n \} \) and \( B \) is a formula. We read the sequent as “\( A_1 \) and \( A_2 \) and... and \( A_n \) implies \( B \)”, and we call \( \Gamma \) the context of the sequent.

A sequent is called valid if it can be derived from a proof. Proofs are trees of formulae defined inductively. In intuitionistic logic a proof is said to be constructive: the proof of \( A \times B \) should include the proof of \( A \) and the proof of \( B \). The proof of \( A \Rightarrow B \) should be a map from the proofs of \( A \) to the proofs of \( B \).

We will concentrate on natural deduction style proof rules, first introduced by Gentzen (1934), where we mainly deal with the right side of the sequents. The proofs associated with the formulae presented above are constructed using the rules in Table 5.1. The rules come in two flavors:

1. the axiom \( \text{Ax} \) and two structural rules: the weakening rule \( W \) and the contraction rule \( C \),

2. introduction and elimination for each connective (except for the null-ary connective \( \top \) which has only an introduction rule).

5.2.2 Curry-Howard Isomorphism

Being able to write proofs, we now turn to the question of an equivalence relation of proofs. Given two proofs of the same sequent, when can we say that they are equal? For example, it is reasonable to expect that

\[
\begin{align*}
\frac{\Gamma \vdash A \quad \text{Ax}}{\Gamma \vdash A} & \quad \begin{array}{c}
\frac{\Gamma \vdash B \quad \text{Ax}}{\Gamma \vdash B} \\
\frac{\Gamma \vdash A \times B \quad \times_1}{\Gamma \vdash A} & \quad \text{and} & \quad \frac{\text{Ax}}{\Gamma \vdash A}
\end{array}
\end{align*}
\]

are the same proof. Indeed, in the first case we do not use \( A, B \vdash B \): the proof only “forgets” it in the last step.

By comparing the rules for building proofs and the typing rules of the simply-typed lambda calculus, one can give a computational interpretation of proofs. It is known as the Curry-Howard isomorphism. Under this scheme, one can say that two proofs are equivalent if they represent two axiomatically equivalent lambda-terms.
If we re-interpret the reduction rules for the lambda calculus in the context of proof of intuitionistic linear logic, one finds a formal definition of the “similar look” of the two proofs of Equation 5.2.1. The first proof corresponds to the typing judgement \( x : A, y : B \vdash \pi_1((x, y)) : A \) and the second to the typing judgement \( x : A, y : B \vdash x : A \).

Interpreted in the context of proofs, the \( \beta \)-reduction rules correspond to the removal of an introduction followed by an elimination. We can construct the following rewrite system of proofs:

\[
\begin{array}{c}
\pi_1 \quad \pi_2 \\
\vdots \quad \vdots \\
\Gamma \vdash A \\
\Gamma \vdash B \\
\Gamma \vdash A \times B \quad \times_{I} \\
\Gamma \vdash A \\
\Gamma \vdash A \times B \quad \times_{E_1} \\
\Gamma \vdash A,
\end{array}
\]

\[
\begin{array}{c}
\pi_1 \quad \pi_2 \\
\vdots \quad \vdots \\
\Gamma \vdash A \\
\Gamma \vdash B \\
\Gamma \vdash A \times B \quad \times_{I} \\
\Gamma \vdash A, \\
\end{array}
\]

\[
\begin{array}{c}
\pi_1 \quad \pi_2 \\
\vdots \quad \vdots \\
\Gamma, A \vdash B \\
\Gamma \vdash A \\
\Gamma \vdash A 	imes B \quad \times_{I} \\
\Gamma \vdash A, \\
\end{array}
\]

\[
\begin{array}{c}
\pi_1 \quad \pi_2 \\
\vdots \quad \vdots \\
\Gamma, A \vdash B \\
\Gamma \vdash B \\
\Gamma \vdash A \\
\Gamma \vdash A \Rightarrow B \Rightarrow_{I} \\
\Gamma \vdash A, \\
\end{array}
\]

\[
\begin{array}{c}
\pi_1 \quad \pi_2 \\
\vdots \quad \vdots \\
\Gamma, A \vdash B \\
\Gamma \vdash B \\
\Gamma \vdash A \Rightarrow B \Rightarrow_{E_2} \\
\Gamma \vdash B, \\
\end{array}
\]

where \( \pi_1^\ast \) is \( \pi_1 \) without the \( A \) in the context and whenever an axiom of \( \pi_1 \) is of the form \( \Gamma, A, \Delta \vdash A \) it is replaced by the proof

\[
\begin{array}{c}
\pi_2 \\
\vdots \\
\Gamma, \Delta \vdash A.
\end{array}
\]

5.3 Categorical Logic

A proof can be set in relation with a term in a suitable language. However, while this describes the equivalence classes of proofs, it does not give out the internal structure of the theory.

The structure of the proof system can be extracted using a categorical description of the situation. In categorical logic, we consider a category whose objects are formulae and whose arrows \( A \to B \) are equivalence classes of proofs of the sequent \( A \vdash B \). Equivalently, objects are types and arrows are equivalence classes of typing judgements \( x : A \vdash M : B \).

In this section we describe the category generated by the language of Section 5.2.1 (or, equivalently, of the proof theory defined in Section 5.2.4).

**Definition 5.3.1.** Given a cartesian closed category \( (C, \times, \Rightarrow, T) \), let

\[
\Phi : \text{hom}(A \times B, C) \sim \text{hom}(A, B \Rightarrow C)
\]

be the canonical isomorphism and \( \varepsilon_{A,B} : (A \Rightarrow B) \times A \Rightarrow B \) the counit of the adjunction.

Given a map \( \psi \) from type variables to objects of \( C \), one defines an interpretation of any type by induction:

\[
[A \times B]_{\psi} = [A]_{\psi} \times [B]_{\psi}, \quad [A \Rightarrow B]_{\psi} = [A]_{\psi} \Rightarrow [B]_{\psi},
\]

and one can interpret valid typing judgements of the simply-typed lambda calculus as morphisms in \( C \):

\[
[x_1 : A_1, \ldots, x_n : A_n \vdash M : B]_{\psi} : [A_1]_{\psi} \times \cdots \times [A_n]_{\psi} \Rightarrow [B]_{\psi}.
\]
5.4 Lambda Calculus and Side Effects

Axiom:
\[ \Delta, x : A \triangleright x : A \]
\[ \pi^2 : [\Delta] \times [A] \rightarrow [A] \]

Unit:
\[ \Delta \triangleright \ast : \top \]
\[ \bigcirc_{\Delta} : [\Delta] \rightarrow T \]

Product:
\[ \Delta \triangleright M : A \quad \Delta \triangleright N : A \]
\[ \Delta \triangleright \langle M, N \rangle : A \times B \]
\[ f : [\Delta] \rightarrow [A] \quad g : [\Delta] \rightarrow [B] \]
\[ (f, g) : [\Delta] \rightarrow [A] \times [B] \]

Application:
\[ \Delta \triangleright M : A \Rightarrow B \quad \Delta \triangleright N : A \]
\[ \Delta \triangleright MN : B \]
\[ f : [\Delta] \rightarrow [A] \Rightarrow [B] \quad g : [\Delta] \rightarrow [A] \]
\[ \Phi(f) : [\Delta] \rightarrow [A] \Rightarrow [B] \]

Abstraction:
\[ \Delta, x : A \triangleright M : B \]
\[ \Delta \triangleright \lambda x. M : A \Rightarrow B \]
\[ f : [\Delta] \times [A] \rightarrow [B] \]
\[ \Phi(f) : [\Delta] \rightarrow [A] \Rightarrow [B] \]

Table 5.2: Interpretation of the simply typed lambda calculus

They are defined by induction as in Table 5.2. The denotation of a context \( \Delta = x_1 : A_1, \ldots, x_n : A_n \) is defined by
\[ [\Delta]_{\psi} = [A_1]_{\psi} \times \cdots \times [A_n]_{\psi}. \]
If the context \( \Delta \) is empty, then \( [\Delta] = T \). For legibility one drops the \( \psi \) on the symbol \([\[\_\]\]_\psi].

Lemma 5.3.2. The interpretation is sound: If \( \Delta \triangleright M \approx_{ax} M' : A \) then
\[ [\Delta \triangleright M : A] = [\Delta \triangleright M' : A]. \]

Definition 5.3.3. Let \( \mathcal{C}_{\lambda}^{st} \) be the category defined as follows:
\begin{itemize}
  \item Objects are types of the simply typed lambda calculus.
  \item Arrows from \( A \) to \( B \) are \( \beta \)-equivalence classes of typing judgements \( x : A \triangleright M : B \).
  \item The identity map on \( A \) is \( x : A \triangleright x : A \), and the composition of \( x : A \triangleright M : B \) and \( y : B \triangleright N : C \) is \( x : A \triangleright (\lambda y. N)M : C \).
\end{itemize}

Lemma 5.3.4. The category \( \mathcal{C}_{\lambda}^{st} \) is cartesian closed.

Theorem 5.3.5. The simply-typed lambda calculus is an internal language for cartesian categories: In \( \mathcal{C}_{\lambda}^{st} \), we have
\[ [x : A \triangleright M : B]_{\mathcal{C}_{\lambda}^{st}} \approx_{ax} (x : A \triangleright M : B). \]

5.4 Lambda Calculus and Side Effects

5.4.1 Pure Versus Impure Calculus.

The simply typed lambda calculus as model of computation is slightly restrictive: It allow only pure computations, that is, computations with no side effects. A program is said to produce some side effect if the returned value does not entirely describe the behaviour of the program. Here are several examples of side effects:
5.4. Lambda Calculus and Side Effects

1. Probabilistic choice. A program that tosses a coin along the computation will have side effects: two runs of the program might not give the same output value.

2. Non-termination. A program that can diverge (i.e. not return anything) will have side effects, since no value can describe the behaviour of the program.

3. External state. A program that read and/or write some data on some external state is subject to side-effects: the value the program output might depend on the external state.

5.4.2 Reduction Strategies

In the event of side effects, the order of reduction of the redexes might influence the output, or even the behaviour of the program (in the case of divergence, for example). The description of a program allowing side effects has to come with a description of values and a choice of reduction strategy. The values are special terms that do not reduce anymore under the chosen strategy. For a simply-typed lambda calculus with side effects, the values are defined as

\[
Value \quad V, W \quad ::= \quad x \mid \lambda x.M \mid \langle V, W \rangle \mid *.
\]

Two main strategies are the following:

1. Call-by-name. In this strategy, one first applies functions to arguments before evaluating them. The rules for the application is

\[
\frac{M \rightarrow M'}{MN \rightarrow M'N} \quad , \quad \frac{(\lambda x.M)N \rightarrow M[N/x]}{\langle V, W \rangle \rightarrow *}.
\]

2. Call-by-value. In this strategy, one reduces arguments to values before applying a function to them. Rules for application can be written as follows:

\[
\frac{N \rightarrow N'}{MN \rightarrow MN'} \quad , \quad \frac{M \rightarrow M'}{MV \rightarrow MV'} \quad , \quad \frac{(\lambda x.M)V \rightarrow M[V/x]}{\langle V, W \rangle \rightarrow *}.
\]

In the following we will concentrate on a model for call-by-value based computation.

5.4.3 Towards a Semantics

The categorical analysis developed in Section 5.3 does not apply to a lambda calculus with side effects. In order to make categorical sense of the calculus, one has to separate the analysis of the values from the analysis of the computations (that is, the arbitrary terms).

The semantics is due to Moggi (1988, 1989, 1991). A categorical model for a lambda calculus with side effects is a cartesian category \((C, \times, T)\) together with a strong monad \((T, \mu, \eta, t)\).

Definition 5.4.1 Moggi (1991). A strong monad over a monoidal category \(C\) is a monad \((T, \eta, \mu)\) together with a natural transformation \(t_{A,B} : A \otimes TB \rightarrow T(A \otimes B)\), call the tensorial strength, such
that the diagrams

\[
\begin{array}{cccc}
T \otimes TA & \xrightarrow{\lambda} & TA & (A \otimes B) \otimes TC \xrightarrow{\alpha} A \otimes (B \otimes TC) \\
\downarrow{\tau} & & \downarrow{t} & \\
A \otimes B & & T((A \otimes B) \otimes C) & A \otimes T(B \otimes C)
\end{array}
\]

commute.

**Remark 5.4.2** [Moggi, 1991]. If the category \( C \) is symmetric, the tensorial strength \( t \) induces two natural transformations \( TA \otimes TB \to T(A \otimes B) \), namely

\[
\begin{align*}
\Psi_1 : TA \otimes TB & \xrightarrow{\sigma_{T,A,TB}} TB \otimes TA \xrightarrow{t_{TB,A}} T(TB \otimes A) \xrightarrow{[\sigma_{TB,A,T}]} T(A \otimes B), \\
\Psi_2 : TA \otimes TB & \xrightarrow{t_{TA,B}} T(TA \otimes B) \xrightarrow{[\sigma_{T,B,A}]} T(A \otimes B).
\end{align*}
\]

Note that since \( \sigma_{T,A,B} ; t_{B,A} : TA \otimes B \to T(B \otimes A) \), \( (\sigma_{T,A,B} ; t_{B,A})^* \) is a morphism \( T(TA \otimes B) \to T(B \otimes A) \). Note also that \( \Psi_1 \) and \( \Psi_2 \) might not be equal: the map \( \Psi_1 \) “evaluates” the second variable and then the first one. The map \( \Psi_2 \) does the opposite. The strong monad is called *commutative* if \( \Psi_1 = \Psi_2 \).

**Lemma 5.4.3.** If \( T \) is a commutative strong monad on a symmetric monoidal category \( C \), then the two natural transformations \( \Psi_1 \) and \( \Psi_2 \) make both \( T \) monoidal.

The category \( C \) is not quite cartesian closed. Indeed, from a computation \( x : A \triangleright M : B \) one retrieve a value \( \lambda x.M : A \Rightarrow B \). We ask the category to have exponentials in the following sense.

**Definition 5.4.4.** A symmetric monoidal category \((C, \otimes, \top)\) together with a strong monad \((T, \eta, \mu)\) is said to have \( T \)-exponentials [Moggi, 1991], or *Kleisli exponentials*, if it is equipped with a bifunctor \( \Rightarrow : C^{op} \times C \to C \), and a natural isomorphism

\[
\Phi : C(A, B \Rightarrow C) \xrightarrow{\cong} C(A \otimes B, TC).
\]

**Lemma 5.4.5.** The map \( \Phi \) induces a natural transformation \( \varepsilon_{A,B} : (A \Rightarrow B) \otimes A \to TB \) defined by \( \Phi(id_{A \Rightarrow B}) \).

### 5.4.4 Computational Model for Call-By-Value

The typed lambda calculus developed by [Moggi, 1991] for dealing with side-effects is called a *computational lambda calculus*. It is interpreted in a *computational model*, that is:

- a cartesian category \((C, \times, T)\),
- with a strong monad \((T, \mu, \eta, t)\) satisfying the mono requirement,
- and Kleisli exponentials \( A \Rightarrow B \) for every \( A, B \) objects in \( C \).

The values are interpreted as maps in \( C \), and the computations in \( C_T \). The category \( C \) is called the category of *values* and the category \( C_T \) the category of *computations*. 
5.5 Intuitionistic Linear Logic

Linear logic is a resource-sensitive logic developed by Girard (1987). It tries to address the fact that in classical logic, the introduction rule for the conjunction $\times$

$$\frac{\Delta \triangleright A}{\Delta \triangleright A \times B} \times_I$$

can be derived from two distinct rules, respectively an additive and a multiplicative one:

$$\frac{\Delta \triangleright A}{\Delta \triangleright A \times B} \times_I, \quad \frac{\Delta \triangleright A}{\Delta, \Gamma \triangleright A \times B} \times^m_I,$$

where $\Delta$ and $\Gamma$ are disjoint. If these two rules are “the same” in classical logic, it is because of the existence of the weakening and contraction rules

$$\frac{\Delta \triangleright A}{\Delta, A \triangleright B} W, \quad \frac{\Delta, A, A \triangleright B}{\Delta, A \triangleright B} C.$$

Using ($W$) and ($\times^m_I$) one can recover ($\times^a_I$):

$$\frac{\Delta, \Gamma \triangleright A}{\Delta, \Gamma, \Delta, \Gamma \triangleright A \times B} \times^m_I W,$$

and using ($C$) and ($\times^a_I$) one can recover ($\times^m_I$):

$$\frac{\Delta \triangleright A}{\Delta, \Gamma \triangleright A} W, \quad \frac{\Gamma \triangleright A}{\Delta, \Gamma \triangleright A \times B} \times^a_I W.$$

In linear logic, the contraction and weakening rules are not valid in general, so for example $A$ and $A \times A$ are not isomorphic, and the product is split into its additive and multiplicative components $\times$ and $\otimes$.

In the intuitionistic multiplicative fragment, formulae are defined by

$$A, B ::= \iota | \top | A \otimes B | A \multimap B | !A,$$

where $\iota$ ranges over a set of constants, $\otimes$ is the multiplicative conjunction, called the tensor, $\top$ stands for the unit of the tensor, and $A \multimap B$ is the implication corresponding to the tensor. $!(\!)$ is a unary constructor for allowing weakening and contraction:

$$\frac{!A, !A \triangleright B}{!A \triangleright B} (C), \quad \frac{!A \triangleright B}{!A \triangleright B} (W).$$

5.6 Linear Calculi and their Interpretations

Numerous interpretations of intuitionistic linear logic have been proposed, both at an abstract level and at a computational level.

Very good references for this section are the Ph.D. theses of Bierman (1993) and Maneglia (2004) and the review of Melliès (2002).
5.6. Linear Calculi and their Interpretations

5.6.1 Earlier Models

One of the first well-known categorical descriptions of the system of proofs of linear logic was (Seely, 1989). The paper discusses the required structures for a model of linear logic, where the structure of the operator “!” is captured by a comonad arising in the context of an adjunction between a cartesian category and a symmetric monoidal category.

In his thesis, Lafont (1988a,b) also describes a model of intuitionistic linear logic where the modality ! is this time defined in terms of a comonoid structure. The logic is thought of as a type system. In that spirit, the type !A is defined via its constructors and destructors, as would be the types of integers or booleans.

Abramsky (1993) and Wadler (1992) describe a computational interpretation of intuitionistic linear logic, with a typed lambda calculus whose typing rules follows closely the inference rules of the logic. The calculus is a lambda calculus with special term constructs to account for duplicable terms.

5.6.2 Bierman’s Linear Category

The formulation generalizing all of the above descriptions is the one proposed by Bierman (1993) in his Ph.D. thesis, simplified afterward by Benton et al. (1992, 1993). The categorical description comes with a complete lambda calculus that we skip for the purpose of this thesis.

For the definition of the category, we prefer here the terminology given in Schalk (2004), and use the concept of linear exponential comonad.

**Definition 5.6.1.** Let \((C, \otimes, \top, \alpha, \lambda, \rho, \sigma)\) be a symmetric monoidal category. Let \((L, \delta, \epsilon, d^L, d^L)\) be a monoidal comonad. We say that \(L\) is a linear exponential comonad provided that

1. each object in \(C\) of the form \(LA\) is equipped with a commutative comonoid \((LA, \triangle_A, \diamond_A)\), where \(\triangle_A : LA \rightarrow LA \otimes LA\) and \(\diamond_A : LA \rightarrow \top\);
2. \(\triangle_A\) and \(\diamond_A\) are monoidal natural transformations, i.e. the following diagram

\[
\begin{array}{ccc}
LA \otimes LB & \xrightarrow{\triangle_A \otimes \triangle_B} & (LA \otimes LA) \otimes (LB \otimes LB) \\
\downarrow d^L_{A,B} & & \downarrow \swarrow \\
L(A \otimes B) & \xrightarrow{\triangle_{(A \otimes B)}} & L(A \otimes B) \otimes L(A \otimes B)
\end{array}
\]  

(5.6.1)

and the diagrams

\[
\begin{align}
L \top & \xrightarrow{d^L_\top} L \top \otimes L \top \\
\otimes & \xrightarrow{\lambda_T} \top \otimes T \\
\circ & \xrightarrow{d^L_\top \otimes id} L \top \otimes L \top \\
\circ & \xrightarrow{id \otimes \lambda_T} L \top \otimes L \top
\end{align}
\]  

(5.6.2)

\[
\begin{align}
LA \otimes LB & \xrightarrow{\diamond_A \otimes \diamond_B} \top \otimes \top \\
\downarrow d^L_{A,B} & & \downarrow \lambda_T \\
L(A \otimes B) & \xrightarrow{\diamond_{A \otimes B}} L(A \otimes B)
\end{align}
\]  

(5.6.3)

\[
\begin{align}
T & \xrightarrow{id} T \\
\otimes & \xrightarrow{d^L_\top} L \top \\
\circ & \xrightarrow{\diamond_T} T
\end{align}
\]  

(5.6.5)

commutes.
3. The maps

\[ \triangle_A : (LA, \delta_A) \to (LA \otimes LA, (\delta_A \otimes \delta_A); d_A), \]
\[ \diamond_A : (LA, \delta_A) \to (T, d^L_T) \]

are \( L \)-coalgebra morphisms, i.e.

\[
\begin{array}{ccc}
LA & \xrightarrow{\triangle_A} & LA \otimes LA \\
& \downarrow^{\delta_A} & \downarrow^{\delta_A \otimes \delta_A} \\
L^2 A & \xrightarrow{\delta_A} & L^2 A \otimes L^2 A \\
& \downarrow^{d^L_A} & \downarrow^{d^L_A} \\
LA & \xrightarrow{\diamond_A} & \top \\
& \downarrow^{\delta_A} & \downarrow^{\delta_A} \\
L^2 A & \xrightarrow{\diamond_A} & L^2 A \otimes L^2 A;
\end{array}
\]

4. Every map \( \delta_A \) is a comonoid morphism \( (LA, \diamond_A, \triangle_A) \to (L^2 A, \diamond_L A, \triangle_L A) \), i.e.

\[
\begin{array}{ccc}
LA & \xrightarrow{\triangle_A} & L^2 A, \\
& \downarrow^{\delta_A \otimes \delta_A} & \downarrow^{\delta_A} \\
LA \otimes LA & \xrightarrow{\delta_A \otimes \delta_A} & L^2 A \otimes L^2 A \\
& \downarrow^{d^L_A} & \downarrow^{d^L_A} \\
LA & \xrightarrow{\diamond_A \otimes \diamond_A} & \top \\
& \downarrow^{\delta_A} & \downarrow^{\delta_A} \\
L^2 A & \xrightarrow{\diamond_A \otimes \diamond_A} & L^2 A \otimes L^2 A.
\end{array}
\]

**Definition 5.6.2** [Bierman, 1993]. A linear category is a symmetric monoidal category \((C, \otimes, T)\) with finite products \((\times, 1)\), together with a linear exponential comonad \(L\).

**Benton** [1994] gave a clean formulation of the above category in terms of adjunctions. We follow **Barber** [1997] by not mentioning the cartesian closedness, and use the exposition of **Mellies** [2002] as reference.

**Definition 5.6.3.** A linear-non-linear category consists of

- A symmetric monoidal closed category \((C, \otimes, T)\);
- a category \((M, \times, 1)\) with finite products;
- a symmetric monoidal adjunction

\[
\begin{array}{ccc}
M & \xrightarrow{U} & C.
\end{array}
\]

**Lemma 5.6.4.** Every linear category defines a linear-non-linear category, where \((M, \times, 1)\) is the category of coalgebras of the comonad.

**Theorem 5.6.5.** Every linear-non-linear category is a linear category.
Chapter 6

A Lambda Calculus for Quantum Computation

The question this thesis tries to address is the question of the meaning and the description of higher-order quantum information, in the paradigm of the QRAM model. That is, programs are written on a classical device but are given access to quantum information.

In this chapter we will describe a typed lambda calculus for quantum computation, together with an operational semantics. The requirements we have in mind for the language are the following:

1. The language should have classical control. That is, the structural aspect of the programs (pairing, tests, etc.) should be classical. Such a language will be described in Section 6.1.

2. We want the program to manipulate quantum data. So quantum information should be thought of as part of the code as much as possible. We give in Section 6.2 an abstract machine and an operational semantics for giving sense to the language.

3. We want the language to be as natural as possible for the programmer. In particular, the compiler should be the one to make sure that valid programs do not duplicate quantum variables. We do not want to have special term constructs for this. In Section 6.3 we define a type system to rule out invalid programs.

The work presented in this chapter appeared in my M.Sc. thesis [Valiron, 2004a]. It was also published in [Selinger and Valiron, 2006a, 2007].

6.1 The Language

We adapt the lambda calculus described in Section 5.1 to manipulate classical data and quantum data. We focus primarily on the classical booleans 0 and 1, on the measurement, the creation of quantum bits and the unitary operations.

Definition 6.1.1. A lambda-term is an expression made of the following core grammar, written in BNF form:

\[
\text{Term} \quad M, N, P \quad ::= \quad x \mid MN \mid \lambda x.M \mid \text{if } M \text{ then } N \text{ else } P \mid 0 \mid 1 \mid \text{meas} \mid \text{new} \mid U \mid \ast \mid \langle M, N \rangle \mid \text{let } \langle x, y \rangle = M \text{ in } N,
\]

where \( x \) ranges over an infinite set of term variables and \( U \) a set of unitary gates; the term if \( M \) then \( N \) else \( P \) stands for a test on \( P \) with output \( M \) if true, \( N \) else; the constant terms 0 and 1 stand for the
Definition 6.1.4. Given a term $H \langle M, \ldots \rangle$ for example, a fair coin can be implemented as follows:

Example 6.1.5. Using this calculus, one can produce terms that compute quantum algorithms.

We identify terms up to $\alpha$-equivalence, defined as in Section 5.1.3. We give the full definition for further reference.

Convention 6.1.2. We sometimes use the shorthand notations

$$\langle M_1, \ldots, M_n \rangle = \langle M_1, (M_2, \ldots) \rangle,$$

$$M_1 M_2 \ldots M_n = \ldots ((M_1 M_2) M_3) \ldots M_n,$$

$$\text{let } x = M \text{ in } N = \langle \lambda x. N \rangle M,$$

$$\lambda (x, y). M = \lambda z. (\text{let } (x, y) = z \text{ in } N),$$

$$\lambda x_1 \ldots x_n. M = \lambda x_1 \lambda x_2 \ldots \lambda x_n. M,$$

$$\text{let } x, y_1 \ldots y_n = M \text{ in } N \Rightarrow \text{let } x = (\lambda y_1 \ldots y_n. M) \text{ in } N.$$

The notions of free and bound variables and of substitution are defined as in Chapter 5. We give the full definition for further reference.

Definition 6.1.3. The set $FV(M)$ of free variables of a term $M$ is defined as follows:

$$FV(*) = \emptyset,$$

$$FV(U) = \emptyset,$$

$$FV(M \text{ new}) = \emptyset,$$

$$FV(M \text{ meas}) = \emptyset,$$

$$FV(0) = \emptyset,$$

$$FV(1) = \emptyset,$$

$$FV(\text{ let } (x, y) = M \text{ in } N) = FV(M) \cup (FV(N) \setminus \{x, y\}).$$

Definition 6.1.4. Given a term $M$ and a term $P$, we define the substitution of $x$ in $M$ by $P$, written $M[P/x]$, by the following:

$$x[P/x] = P,$$

$$c[P/x] = c,$$

$$*P/x = *,$$

$$(M N)[P/x] = (M[P/x])(N[P/x]),$$

$$(\langle M, N \rangle)[P/x] = (M[P/x], N[P/x]),$$

$$(\text{if } M \text{ then } N_1 \text{ else } N_2)[P/x] = \text{if } M[P/x] \text{ then } N_1[P/x] \text{ else } N_2[P/x],$$

$$\langle \lambda x. M \rangle[P/x] = \lambda x. M,$$

$$\langle \lambda y. M \rangle[P/x] = \lambda y. (M[P/x]) \quad (\text{if } y \neq x \text{ and } y \not\in FV(P))$$

If $y = x$ or $z = x$:

$$(\text{let } (y, z) = M \text{ in } N)[P/x] = (\text{let } (y, z) = (M[P/x]) \text{ in } N)$$

If $y \neq x$ and $z \neq x$ (and $y, z \not\in FV(P)$):

$$(\text{let } (y, z) = M \text{ in } N)[P/x] = (\text{let } (y, z) = (M[P/x]) \text{ in } (N[P/x])).$$

Example 6.1.5. Using this calculus, one can produce terms that compute quantum algorithms. For example, a fair coin can be implemented as follows:

$$\text{meas}(H(\text{new } 0)),$$

where $H$ is the Hadamard gate.
6.2 Operational Semantics

Although we want the language to perform quantum computation, no quantum bit was added into the terms of the language. For example, we might want to represent the constant function outputting a quantum bit, as follows:

\[ \lambda x. (\alpha |0\rangle + \beta |1\rangle). \]

The reason why we do not allow such a notation is the problem of entanglement. As noted in Section 3.3.2, it is not always possible to write a two quantum bit system in the form \(|\phi\rangle \otimes |\psi\rangle\). Therefore, if \(x\) is the variable corresponding to the first quantum bit and \(y\) the one corresponding to the second in an entangled state, it is not possible to write classically the term

\[ (\lambda f. fx)(\lambda t. (gy)t), \]

as there is no means of writing \(x\) and \(y\).

### 6.2.1 Abstract Machine

This non-local nature of quantum information forces us to introduce a level of indirection for representing the state of a quantum program. Following the scheme of the QRAM model, we describe an abstract machine consisting of a lambda-term \(M\), a array of quantum bits \(Q\) (the QRAM), and a map linking variables in \(M\) to quantum bits in \(Q\):

**Definition 6.2.1.** A quantum closure is represented by a triple \([Q, L, M]\), where

- \(Q\) is a normalized vector of \(\otimes_{i=1}^{n} \mathbb{C}^2\), for some \(n \geq 0\), called the quantum array.
- \(M\) is a lambda term,
- \(L\) is an injective function from a set \(|L|\) of term variables to \(\{1, \ldots, n\}\), where \(FV(M) \subseteq |L|\).

\(L\) is called the linking function.

We write \(|Q| = n\). If \(L(x_i) = i\), we will sometimes write \(L\) as the ordered list \(|x_1 \cdots x_n|\). The idea is that the variable \(x_i\) is bound in the term \(M\) to qubit number \(L(x_i)\) of the state \(Q\). We also call the pair \((Q, L)\) a quantum context. If there are no quantum bits, i.e., if \(Q \in \mathbb{C}\), we write \(Q = |\rangle\). Similarly, if \(L\) is empty we write \(L = |\rangle\).

The purpose of the linking function is to assign specific free variables of \(M\) to specific quantum bits in \(Q\). The notion of \(\alpha\)-equivalence extends naturally to quantum closure, for instance, the states \([|1\rangle, |x\rangle, \lambda y.x]\) and \([|1\rangle, |z\rangle, \lambda y.z]\) are equivalent. We therefore extend the notion of \(\alpha\)-equivalence to quantum closures:

\[ [Q, |x \cdots y \cdots z\rangle, M] =_\alpha [Q, |x \cdots y' \cdots z\rangle, M[y'/y]] \]

if \(y' \not\in FV(M) \cup \{x, \ldots, y, \ldots, z\}\).

### 6.2.2 Evaluation Strategy

Although we solved the problem of entanglement, there is another issue that prevents us from blindly using substitution, namely the probabilistic nature of the measurement.

Consider the term **plus** to be the boolean addition function, defined as

\[ \text{plus} = \lambda xy. \text{if} \ x \ \text{then} \ (\text{if} \ y \ \text{then} \ 0 \ \text{else} \ 1) \ \text{else} \ (\text{if} \ y \ \text{then} \ 1 \ \text{else} \ 0). \]

Now, consider the term \(M = (\lambda x. \text{plus} \ x \ x)(\text{meas}(H(\text{new} \ 0)))\). Depending of the choice made in reducing the term, we obtain a different answer, as shown below, using an intuitive reduction system.
6.2. Operational Semantics

Call-By-Value.

Reducing this in the empty environment, using the call-by-value reduction strategy (as defined in Section 5.4.2), we obtain the following reductions:

\[
\begin{align*}
\langle |\rangle, |\rangle, M \rightarrow & \ [0], |p\rangle, (\lambda x. \textbf{plus} \ x \ x)(\text{meas}(H \ p)) \\
& \rightarrow \ \left\{ \begin{array}{ll}
\langle |0\rangle, |p\rangle, (\lambda x. \textbf{plus} \ x \ x)(0) \\
\langle |1\rangle, |p\rangle, (\lambda x. \textbf{plus} \ x \ x)(1) \\
\end{array} \right. \\
& \rightarrow \ \left\{ \begin{array}{ll}
\langle |0\rangle, |p\rangle, 0 \\
\langle |1\rangle, |p\rangle, 0 \\
\end{array} \right. \\
\end{align*}
\]

where the two branches are taken with probability \( \frac{1}{2} \) each. Thus, under call-by-value reduction, this program produces the boolean value 0 with probability 1.

Call-By-Name

Reducing the same term under the call-by-name strategy, we obtain in one step

\[
\begin{align*}
\langle |\rangle, |\rangle, M \rightarrow & \ [\lambda y. \text{meas}(H \ \text{new} \ 0)], \textbf{plus}(\text{meas}(H \ (\text{new} \ 0))), \textbf{meas}(H \ \text{new} \ 0) \\
\end{align*}
\]

and then with probability \( \frac{1}{4} \),

\[
\begin{align*}
[01], |pq\rangle, 1], \ [10], |pq\rangle, 1], \ [00], |pq\rangle, 0], \ [11], |pq\rangle, 0].
\end{align*}
\]

Therefore, the boolean output of this function is 0 or 1 with equal probability.

Mixed Strategy

Finally, if we mix the two reduction strategies, the program can even reduce to an ill-formed term. Namely, reducing by call-by-value until we reach the term

\[
\begin{align*}
\langle |\rangle, |\rangle, M \rightarrow & \ [\lambda y. \text{meas}(H \ \text{new} \ 0)], \textbf{meas}(H \ (\text{new} \ 0)), \textbf{meas}(H \ \text{new} \ 0)] \\
\end{align*}
\]

and then changing to call-by-name, we obtain in one step the term

\[
\begin{align*}
\langle |\rangle, |\rangle, M \rightarrow & \ [\lambda y. \text{meas}(H \ \text{new} \ 0)], \textbf{meas}(H \ (\text{new} \ 0)), \textbf{meas}(H \ \text{new} \ 0)] \\
\end{align*}
\]

which is not a valid program since there are two occurrences of the quantum bit \( p \).

Remark 6.2.2. In call-by-name, a measurement of the form \( \text{meas} \ M \) is carried over along \( \beta \)-reductions. There is no possible way to “force” the measurement to happen. Therefore, one cannot obtain the behavior of the reductions in (6.2.1) using a call-by-name procedure. However, it is possible to simulate a call-by-name reduction in call-by-value by encapsulating terms in lambda-abstractions. For example, to carry over the term \( \text{meas} \ M \) using a call-by-value reduction strategy, one can write the term as \( \lambda y. \text{meas} \ M \), with \( y \) a fresh variable.

Convention 6.2.3. Following Remark 6.2.2, we consider the call-by-value reduction strategy in the remainder of this chapter.
6.2.3 Probabilistic Reduction Systems

In order to formalize the operational semantics of the quantum lambda calculus, we introduce the notion of a probabilistic reduction system. Note that several notions of probabilistic rewrite systems have been studied in the literature (see e.g. Bournez and Hovrup, 2003; Sen et al., 2003). However, these works study rewrite system in a very general setting. For the purpose of this work, where our probabilistic reduction is quite simple, we define an abstract notion of probabilistic reduction only to capture the notion of value and error state and not to study the system itself.

**Definition 6.2.4.** A probabilistic reduction system is a tuple \((X, U, R, \text{prob})\) where \(X\) is a set of states, \(U \subseteq X\) is a subset of value states, \(R \subseteq (X \setminus U) \times X\) is a set of reductions, and \(\text{prob} : R \to [0, 1]\) is a probability function, where \([0, 1]\) is the real unit interval. Moreover, we impose the following conditions:

- For any \(x \in X\), \(R_x = \{ x' \mid (x, x') \in R \}\) is finite.
- \(\sum_{x' \in R_x} \text{prob}(x, x') \leq 1\)

We call \(\text{prob}\) the one-step reduction, and denote \(x \xrightarrow{p} y\) to be \(\text{prob}(x, y) = p\). Let us extend \(\text{prob}\) to the \(n\)-step reduction

\[
\text{prob}^0(x, y) = \begin{cases} 
0 & \text{if } x \neq y \\
1 & \text{if } x = y
\end{cases}
\]

\[
\text{prob}^1(x, y) = \begin{cases} 
\text{prob}(x, y) & \text{if } (x, y) \in R \\
0 & \text{else}
\end{cases}
\]

\[
\text{prob}^{n+1}(x, y) = \sum_{z \in R_x} \text{prob}(x, z) \text{prob}^n(z, y),
\]

and the notation is extended to \(x \xrightarrow{n} p y\) to mean \(\text{prob}^n(x, y) = p\).

**Definition 6.2.5.** We say that \(y\) is reachable in one step with non-zero probability from \(x\), denoted \(x \xrightarrow{>0} y\) when \(x \xrightarrow{p} y\) with \(p > 0\). We say that \(y\) is reachable with non-zero probability from \(x\), denoted \(x \xrightarrow{>0} y\) when there exists \(n > 0\) such that \(x \xrightarrow{n} y\) with \(p > 0\).

**Definition 6.2.6.** We can then compute the probability to reach \(u \in U\) from \(x\): It is a function from \(X \times U\) to \(\mathbb{R}\) defined by \(\text{prob}_U(x, u) = \sum_{n=0}^{\infty} \text{prob}^n(x, u)\). The total probability for reaching \(U\) from \(x\) is

\[
\text{prob}_U(x) = \sum_{n=0}^{\infty} \sum_{u \in U} \text{prob}^n(x, u).
\]

On the other hand, there is also the probability to diverge from \(x\), or never reaching anything. This value is

\[
\text{prob}_\infty(x) = \lim_{n \to \infty} \sum_{y \in X} \text{prob}^n(x, y).
\]

We define the error probability of \(x\) to be the number

\[
\text{prob}_{err}(x) = 1 - \text{prob}_U(x) - \text{prob}_\infty(x).
\]

**Lemma 6.2.7.** For all \(x \in X\), \(\text{prob}_U(x) + \text{prob}_\infty(x) \leq 1\). That is, \(0 \leq \text{prob}_{err}(x) \leq 1\).

**Definition 6.2.8.** In addition to the notion of reachability with non-zero probability, there is also a weaker notion of reachability, given by \(R\): We will say that \(y\) is reachable in one step from \(x\), written \(x \xrightarrow{\sim} y\), if \(xRy\). By the properties of \(R\), \(x \xrightarrow{>0} y\) implies \(x \xrightarrow{\sim} y\). As usual, \(\sim^*\) denotes the transitive reflexive closure of \(\sim\), and we say that \(y\) is reachable from \(x\) if \(x \xrightarrow{\sim^*} y\).
Definition 6.2.9. In a probabilistic reduction system, a state  \( x \) is called an error-state if  \( x \not\in U \) and  \( \sum_{x' \in X} \text{prob}(x, x') < 1 \). An element  \( x \in X \) is consistent if there is no error-state  \( e \) such that  \( x \leadsto^* e \).

Lemma 6.2.10. If  \( x \) is consistent, then  \( \text{prob}_{\text{err}}(x) = 0 \).

Remark 6.2.11. We need the weaker notion of reachability  \( x \leadsto y \), in addition to reachability with non-zero probability  \( x \leadsto_{\geq 0} y \), because a null probability of getting a certain result is not an absolute warranty of its impossibility. In the QRAM, suppose we have a qubit in state  \( |0\rangle \). Measuring it cannot theoretically yield the value 1, but in practice, this might happen with small probability, due to imprecision of the physical operations. Therefore, when we prove subject reduction (see Theorem 6.4.5), we will use the stronger notion. In short: a type-safe program should not crash, even in the event of random QRAM errors.

Remark 6.2.12. The converse of Lemma 6.2.10 is false. For instance, if  \( X = \{a, b\}, U = \emptyset, a \leadsto_1 a, \) and  \( a \leadsto_0 b, \) then  \( b \) is an error state, and  \( b \) is reachable from  \( a \), but only with probability zero. Hence  \( \text{prob}_{\text{err}}(a) = 0 \), although  \( a \) is inconsistent.

6.2.4 Reductive System

We define a probabilistic call-by-value reduction procedure for the quantum lambda calculus. Note that, although the reduction itself is probabilistic, the choice of which redex to reduce at each step is deterministic.

Definition 6.2.13. A value is a term of the following form:

\[
\text{Value} \quad V, W ::=} x \mid \lambda x. M \mid 0 \mid 1 \mid \text{meas} \mid \text{new} \mid U \mid \ast \mid \langle V, W \rangle.
\]

A quantum closure of the form  \( [Q, L, V] \) where  \( V \) is a value is called a quantum value-state.

The reduction rules are shown in Table 6.1. We write  \( [Q, L, M] \leadsto_p [Q', L', M'] \) for a single-step reduction of states which takes place with probability  \( p \). In the rule for reducing the term  \( U(x_j, \ldots, x_n) \),  \( U \) is an  \( n \)-ary built-in unitary gate,  \( j_1, \ldots, j_n \) are pairwise distinct, and  \( Q' \) is the quantum state obtained from  \( Q \) by applying this gate to qubits  \( j_1, \ldots, j_n \). In the rule for measurement,  \( |Q_0\rangle \) and  \( |Q_1\rangle \) are normalized states of the form

\[
|Q_0\rangle = \sum_j \alpha_j|\phi^0_j\rangle \otimes |0\rangle \otimes |\psi^1_j\rangle, \quad |Q_1\rangle = \sum_j \beta_j|\phi^0_j\rangle \otimes |1\rangle \otimes |\psi^1_j\rangle,
\]

where  \( |\phi^0_j\rangle \) and  \( |\psi^1_j\rangle \) are  \( i \)-qubit states (so that the measured qubit is the one pointed to by  \( x_i \)). In the rule for  \( \text{new} \),  \( Q \) is an  \( n \)-qubit state, so that  \( Q \otimes |i\rangle \) is an  \( (n + 1) \)-qubit state.

Definition 6.2.14. We define a weaker relation  \( \leadsto \). This relation models the transformations that can happen in the presence of decoherence and imprecision of physical operations. We define  \( [Q, M] \leadsto [Q', M'] \) to be  \( [Q, M] \leadsto_p [Q', M'] \), even when  \( p = 0 \), plus the additional rule, if  \( Q \) and  \( Q' \) are vectors of equal dimensions:  \( [Q, M] \leadsto [Q', M] \).

Lemma 6.2.15. Let  \( X \) be the state of all quantum closures and  \( U \) the set of quantum value-states. Let  \( \text{prob} \) be the function such that for  \( x, y \in X \),  \( \text{prob}(x, y) = p \) if  \( x \leadsto_p y \) and  \( 0 \) else. Then  \( (X, U, \leadsto, \to) \) is a probabilistic reduction system.

Remark 6.2.16. This probabilistic reduction system has error states, for example, the states  \( [Q, L, H(x.x)] \) or  \( [Q, [xyz], U(x, x)] \). Such error states correspond to run-time errors. In the next section, we introduce a type system designed to rule out such error states.
We will now define a type system designed to eliminate all run-time errors arising from the reduction system of the previous section, such as the ones shown in Remark \[6.3.4(1)\].

6.3.1 Types

In the lambda calculus we just defined, we need to be able to account for higher-order, products (and unit), classical booleans and quantum booleans. Since we do not want to have specific term construct for dealing with duplication and non-duplication, the information has to be encoded into the type system. We use a type system inspired from linear logic. By default, a term of type \(\mathbf{A}\) is assumed to be non-duplicable, and duplicable terms are given the type \(!\mathbf{A}\) instead.

**Definition 6.3.1.** Formally, we define the following type system for the lambda calculus of Section \[6.1\] as follows:

\[
\text{qType } \mathbf{A}, \mathbf{B} ::= \text{bit } | \text{qbit } | !\mathbf{A} | (\mathbf{A} \rightarrow \mathbf{B}) | \top | (\mathbf{A} \otimes \mathbf{B}).
\]

The constant types \text{bit} and \text{qbit} stand respectively for the classical and the quantum booleans; the type \(!\mathbf{A}\) is for duplicable elements of type \(\mathbf{A}\); the type \(\mathbf{A} \rightarrow \mathbf{B}\) for functions from \(\mathbf{A}\) to \(\mathbf{B}\); the type \(\mathbf{A} \otimes \mathbf{B}\) for pairs of elements of types \(\mathbf{A}\) and \(\mathbf{B}\); finally \(\top\) stands for the type of the term \(*\).

We call the operator 
\(\text{"!"}\) the exponential.

**Convention 6.3.2.** We write \(!!\mathbf{A}\) for \(!!!\ldots!!\mathbf{A}\), with \(n\) repetitions of \(!\). We also use the notation \(\mathbf{A} \otimes^n\) for the \(n\)-fold tensor product \(\mathbf{A} \otimes \cdots \otimes \mathbf{A} = (\ldots (\mathbf{A} \otimes \mathbf{A}) \cdots \otimes \mathbf{A})\).

The typing rules will ensure that any value of type \(!\mathbf{A}\) is duplicable. However, there is no harm in using it only once; thus, such a value should also have type \(\mathbf{A}\). We use a subtyping relation \[Breazu-Tannen et al., 1988; Pierce, 2002\] to capture this notion.
6.3. Type System

Definition 6.3.3. We define a subtyping relation $\prec$ on types as follows, using the overall condition on $n$ and $m$ that $(m = 0) \lor (n \geq 1)$:

\[
\begin{align*}
\n^n \text{bit} & \prec \n^n \text{bit} \quad (\text{bit}), \\
\n^n \text{qbit} & \prec \n^n \text{qbit} \quad (\text{qbit}), \\
\n^n \top & \prec \n^n \top \quad (\top), \\
A_1 \prec B_1 & \quad A_2 \prec B_2 \quad (\otimes), \\
\n^n (A_1 \otimes B_1) & \prec \n^n (B_1 \otimes B_2) \quad (\otimes), \\
A \prec A' & \quad B \prec B' \quad (\neg \circ).
\end{align*}
\]

Lemma 6.3.4. $(\text{qType}, \prec)$ is reflexive and transitive. If we define an equivalence relation $\equiv$ by $A \equiv B$ iff $A \prec B$ and $B \prec A$, $(\text{qType}/\equiv, \prec)$ is a poset.

Proof. The proof uses the fact that $(m = 0) \lor (n \geq 1)$ can be rewritten as $(n = 0) \Rightarrow (m = 0)$ (where $\Rightarrow$ means “implies”). Then the result follows from the reflexivity and the transitivity of the implication.

Lemma 6.3.5. If $A \prec \! B$, then $A = \! A'$ for some type $A'$. Dually, if $A$ is not of the form $\! A'$ and if $A \prec \! B$, then $B$ is not of the form $\! B'$.

Proof. By inspection of the cases.

6.3.2 Typing Rules

In their definition in Section 6.3, we implicitly considered the terms to be typed in a Church-style fashion. Considering terms this way is closer to the programmer approach: the types of the internal subterms have to be figured out by the compiler.

Definition 6.3.6. A typing judgement is the given of a set $\Delta = \{x_1 : A_1, \ldots, x_n : A_n\}$ of typed variables, of a term $M$ and of a type $B$, written as $\Delta \vdash M : B$. We call the set $\Delta$ the typing context, or simply the context.

Given two contexts $\Delta_1$ and $\Delta_2$, we write $(\Delta_1, \Delta_2)$ for the context consisting of the variables in $\Delta_1$ and in $\Delta_2$. When described in this form, it is assumed that $|\Delta_1| \cap |\Delta_2| = \emptyset$.

A typing derivation is called valid if it can be inferred from the rules of Table 6.2. We write $\vdash \Delta$ for a context of the form $\{x_1 : \! A_1, \ldots, x_n : \! A_n\}$. We use the notation $|\Delta|$ to represent the set of (untyped) variables $\{x_1, \ldots, x_n\}$ contained in $\Delta$. In the table, the symbol $c$ spans the set of term constants $\{\text{meas}, \text{new}, U, 0, 1\}$. To each constant $c$ we associate a term $A_c$, as follows:

\[
A_0, A_1 = \text{bit}, \quad A_{\text{new}} = \text{bit} \prec \text{qbit}, \quad A_U = \text{qbit} \otimes \text{qbit} \otimes \text{qbit}, \quad A_{\text{meas}} = \text{qbit} \prec \! \text{bit}.
\]

Remark 6.3.7. The type $\! A_c$ is understood as being the “most generic” one for $c$, as enforced by the typing rule $(a2)$. For example, we defined $A_{\text{new}}$ to be $\text{bit} \prec \! \text{qbit}$. Since new can take any type $B$ such that $\! A_c \prec : B$, the term $\text{new}$ can be typed with all the types in the poset:

\[
\vdash \text{bit} \prec \! \text{qbit} \prec \! \text{bit} \prec \! \text{qbit} \prec \! \text{bit} \prec \! \text{qbit}.
\]

This implies, as expected, that no created quantum bit can have the type $\! \text{qbit}$.

Remark 6.3.8. Note that the type system enforces the requirement that variables holding quantum data cannot be freely duplicated; thus $\lambda x.(x, x)$ is not a valid term of type $\text{qbit} \prec \! \text{qbit} \otimes \text{qbit}$. On the other hand, we allow variables to be discarded freely.

Note also that due to rule $(a2)$ the term $\lambda x.M$ is duplicable only if all the free variables of $M$ (other than $x$) are duplicable. This follows the call-by-value approach: since $\lambda x.M$ is a value, if it is duplicated it will be duplicated as a syntactic string of symbols, thus duplicating every single free variable of it (other than $x$).
### 6.4 Properties of the Type System

We briefly state the properties of the type system that are similar to the properties of the simply-typed lambda calculus of Section 6.1.5. For a full development, we refer the reader to [Valiron, 2004a].

#### 6.4.1 Safety Properties

**Lemma 6.4.1.** If $\Gamma \vdash \Delta$ and $A : B$, and if $\Delta \vdash M : A$ is a valid typing judgement, then $\Gamma \vdash M$.

**Lemma 6.4.2.** If $V$ is a value and if $\Delta \vdash V : A$ is a valid typing judgement, where for all $x \in FV(V)$ then $x$ is duplicable (i.e. there exists a type $B$ such that $\{x : !B\} \subseteq \Delta$), then $\Delta \vdash V : !A$ is valid.

**Lemma 6.4.3 (Substitution).** If $V$ is a value, and if $\Delta, \Gamma_1, x : A \vdash M : B$ and $\Delta, \Gamma_2 \vdash V : A$ are two valid typing judgements, then $\Delta, \Gamma_1, \Gamma_2 \vdash M[V/x] : B$ is valid.
Remark 6.4.4. Note that all the rules of affine intuitionistic logic are derived rules of our type system except for a general promotion rule: Indeed, although the typing judgement $\vdash $ meas 0 : qbit is valid, the judgement $\vdash $ meas 0 : !qbit is not valid. However, as stated in Lemma 6.4.2, the promotion rule is true for values: $\Delta \vdash V : A$ $\lnot \Delta \vdash V' : !A$.

This point will be further developed in Chapter 8 when we will look for a categorical model of the language.

Theorem 6.4.5 (Subject reduction). Given a program $[Q, L, M]$ of type $B$ such that there exists $[Q', L', M']$ with $[Q, L, M] \rightarrow^* [Q', L', M']$, then $[Q', L', M']$ is a program of type $B$.

Theorem 6.4.6 (Progress). Let $[Q, L, M]$ be a program of type $B$. Then $[Q, L, M]$ is not in an error state in the sense of Definition 6.2.9. In particular, either it is a value, or there exists some state $[Q', L', M']$ such that $[Q, L, M] \rightarrow p [Q', L', M']$. Moreover, the total probability of all possible single-step reductions from $[Q, L, M]$ is 1.

6.4.2 Type Inference Algorithm

The fact that a well-typed term behaves well is a good property. However, it still remains to provide a type for the term. For a large class of classical lambda calculus, such a task can be automated using a type inference algorithm.

In these cases, the algorithm is built on the fact that each term admits a principal type, that is, a most general type from which all the other types can be derived. For example, in simply typed lambda calculus, the term $M = \lambda fx.fx$ admits a principal type of the form $(X \Rightarrow Y) \Rightarrow (X \Rightarrow Y)$. Here, $X$ and $Y$ are type variables, and any valid type can be written in this form, for some value of $X$ and $Y$.

The quantum lambda calculus we defined does not satisfy the principal type property. Indeed, consider again the term $M = \lambda fx.fx$. Naively, a principal type would be (1) the most general and (2) the smallest with respect to the subtyping relation. The most general type would have a structure of the form $(X \rightarrow Y) \rightarrow (X \rightarrow Y)$. Since $(X \rightarrow Y) \rightarrow (X \rightarrow Y)$ and $(X \rightarrow Y) \rightarrow (X \rightarrow Y)$ are not valid types for $M$, the valid types for the term $M$, ordered by the subtyping relation (from left to right) are as follows:

$!((X \rightarrow Y) \rightarrow (X \rightarrow Y))$ $\rightarrow (X \rightarrow Y) \rightarrow (X \rightarrow Y)$

$!((X \rightarrow Y) \rightarrow (X \rightarrow Y))$ $\rightarrow (X \rightarrow Y) \rightarrow (X \rightarrow Y)$

$!((X \rightarrow Y) \rightarrow (X \rightarrow Y))$ $\rightarrow (X \rightarrow Y) \rightarrow (X \rightarrow Y)$

Note that there is no smallest element.

It is however possible to find a type inference algorithm for the quantum lambda calculus. Indeed, well-typed terms in the quantum lambda calculus are well-typed in a simply-typed version of the language. That is to say, a valid typing derivation in the quantum lambda calculus is an intuitionistic typing derivation indexed with “!”.

Following this idea it is possible to extract a type inference algorithm for the quantum lambda calculus. The algorithm proceeds by first finding the most general intuitionistic type for the given term, and then placing “!” whenever it is necessary. We refer the reader to [Valiron, 2004a] for the full discussion.
6.5 Examples of Algorithms

It turns out that some quantum algorithms can be efficiently understood as elements of higher-order types. We give in this section the implementation of two non-trivial quantum algorithm, the Deutsch algorithm and the teleportation procedure.

The latter one is of particular interest, since the usual formulation (see Section 3.4.1) is completely sequential.

6.5.1 The Deutsch Algorithm

The simplest example to interpret in a higher-order setting is the Deutsch algorithm. We refer the reader to Section 3.4.3 for its definition. It can be implemented in the quantum lambda calculus in the following way:

\[
\begin{align*}
\text{let } & \text{Deutsch } U_f = \\
& \text{let } \text{tens } f g = \lambda \langle x, y \rangle. \langle fx, gy \rangle \\
& \text{let } \langle x, y \rangle = (\text{tens } H (\lambda x.x))(U_f(H(\text{new } 0), H(\text{new } 1))) \text{ in meas } x,
\end{align*}
\]

using the notations of Convention 6.1.2. Note that \(U_f\) is a variable that stands for a function from a two-qubit state to a two-qubit state. And indeed the function \text{Deutsch} is a higher-order function:

\[\triangleright \text{Deutsch} : !((\text{qbit} \otimes \text{qbit} \rightarrow \text{qbit} \otimes \text{qbit}) \rightarrow \text{bit})\]

is a well-typed typing judgment. Note that \text{Deutsch} is duplicable, and that \(U_f\) does not need to be duplicable, since it is used only once.

Also note that ideally, instead of \(U_f\) the algorithm takes the Boolean function \(f\). Although one can write the type of the function \(U : f \mapsto U_f\) as

\[\text{(bit } \rightarrow \text{ bit}) \rightarrow (\text{qbit} \otimes \text{qbit} \rightarrow \text{qbit} \otimes \text{qbit}),\]

this function is not representable in the quantum lambda calculus.

6.5.2 The Teleportation Procedure

Consider the teleportation algorithm as it is described in Section 3.4.1. We can embed each quantum circuit part of the procedure in a function. There is a function \text{EPR} : !((\text{qbit} \otimes \text{qbit}) \rightarrow \text{bit}) that creates an entangled state, as in the step (1):

\[\text{EPR} = \lambda x. \text{CNOT}(H(\text{new } 0), \text{new } 0).\]

There is a function \text{BellMeasure} : !(\text{qbit} \rightarrow ((\text{qbit} \rightarrow \text{bit} \otimes \text{bit}))) that takes two qubits, rotates and measures them, as in step (2):

\[\text{BellMeasure} = \lambda q_2. \lambda q_1. (\text{let } \langle x, y \rangle = \text{CNOT}(q_1, q_2) \text{ in } \langle \text{meas}(Hx), \text{meas } y \rangle)\]

We also can define a function \text{U} : !((\text{qbit} \rightarrow (\text{bit} \otimes \text{bit} \rightarrow \text{qbit}))) that takes a qubit \(q\) and two bits \(x, y\) and returns \(U_{xy}q\), as in step (3):

\[\text{U} = \lambda q. \lambda x, y. \text{if } x \text{ then } (\text{if } y \text{ then } U_{11}q \text{ else } U_{10}q) \text{ else } (\text{if } y \text{ then } U_{01}q \text{ else } U_{00}q),\]
where the operators $U_{xy}$ are defined as in Section 3.4. The teleportation procedure can be seen as the creation of two non-duplicable functions $f$ and $g$

$$f : \text{qbit} \rightarrow \text{bit} \otimes \text{bit},$$
$$g : \text{bit} \otimes \text{bit} \rightarrow \text{qbit},$$

such that $(g \circ f)(x) = x$ for an arbitrary qubit $x$. We can construct such a pair of functions by the following code:

```plaintext
let ⟨x, y⟩ = EPR * in
let f = BellMeasure x in
let g = U y in ⟨f, g⟩.
```

Note that, since $f$ and $g$ depend on the state of the qubits $x$ and $y$, respectively, these functions cannot be duplicated, which is reflected in the fact that the types of $f$ and $g$ do not contain a top-level "!".

The Dense Coding Procedure

Note that the two functions $f$ and $g$ defined in the previous section also satisfy a dual property, namely that $(f \circ g)(x, y) = (x, y)$ for an arbitrary pair of classical bits. It turns out that this is precisely the higher-order description of the dense coding protocol of Section 3.4.2. The pair of functions $⟨f, g⟩$ creates an isomorphism between $\text{qbit}$ and $\text{bit} \otimes \text{bit}$. This is not in contradiction with the non-duplicability of quantum information since the functions are non-duplicable: one can use the function $g$ only once to “retrieve” the quantum bit encoded into two classical booleans.

Tensor and Space-Like Separation

These two protocols have in common the fact that two parts of the algorithm take place in two different places. This space-like separation is modeled in the type system by the tensor product. It is interesting to note that an interpretation of these two complex protocols can be efficiently interpreted as one single higher-order term.

### 6.5.3 Type Derivation of the Teleportation Protocol

To illustrate the linear type system from Section 6.3.2, we give a complete derivation of the type of the quantum teleportation term. The notation $⟨L.x.y.z⟩$ means that Lemma $x,y,z$ is used. Computing some subtypes:

1. $α_2 !^n α <: \alpha$
2. $α_2 !^m β <: β$
3. $ω, 1, 2 !^k (α → !^m β) <: (!^n α → β)$
4. $⟨L.6.3.4⟩ A <: A$
5. $D, 4 !A <: A$

Computing the type of $\text{EPR}$:

6. $\text{const}, 3 \triangleright \text{new} : \text{bit} \rightarrow \text{qbit}$
7. $\text{const}, 5 \triangleright 0 : \text{bit}$
Computing the type of \textit{BellMeasure}:

\begin{verbatim}
8  app, 6, 7  \triangleright new 0:qbit
9  const, 3  \triangleright H:qbit \rightarrow qbit
10 app, 9, 8  \triangleright H(new 0):qbit
11 \otimes .I, 10, 9  \triangleright \langle H(new 0), new 0 \rangle : qbit \otimes qbit
12 const, 3  x: \top \triangleright \text{CNOT}(qbit \otimes qbit) \rightarrow (qbit \otimes qbit)
13 app, 12, 11 x: \top \triangleright \text{CNOT}(H(new 0), new 0):qbit \otimes qbit
14 \lambda_2, 13  \triangleright \lambda x. \text{CNOT}(H(new 0), new 0)!(\top \rightarrow (qbit \otimes qbit))
\end{verbatim}

\begin{verbatim}
15 var, 1  y:qbit \triangleright y:qbit
16 const, 3  \triangleright \text{meas}:qbit \rightarrow \text{bit}
17 app, 16, 15 y:qbit \triangleright \text{meas}:y:bit
18 var, 1  x:qbit \triangleright x:qbit
19 app, 9, 18 x:qbit \triangleright Hx:qbit
20 app, 16, 19 x:qbit \triangleright \text{meas}(Hx):\text{bit}
21 var, 1  q_1:qbit \triangleright q_1:qbit
22 var, 1  q_2:qbit \triangleright q_2:qbit
23 \otimes .I, 21, 22 q_2:qbit, q_1:qbit \triangleright \langle q_1, q_2 \rangle : qbit \otimes qbit
24 const, 3  \triangleright \text{CNOT}(qbit \otimes qbit) \rightarrow (qbit \otimes qbit)
25 app, 24, 23 q_2:qbit, q_1:qbit \triangleright \text{CNOT}(q_1, q_2) : qbit \otimes qbit
26 \otimes .E, 20, 17 x:qbit, y:qbit \triangleright \langle \text{meas}(Hx), \text{meas} y \rangle : \text{bit} \otimes \text{bit}
27 \otimes .E, 25, 26 q_2:qbit, q_1:qbit \triangleright \text{let}: \langle x, y \rangle = \text{CNOT}(q_1, q_2)
\text{in} \langle \text{meas}(Hx), \text{meas} y \rangle : \text{bit} \otimes \text{bit}
28 \lambda_1, 27 q_2:qbit \triangleright \lambda q_1.(\text{let}: \langle x, y \rangle = \text{CNOT}(q_1, q_2)
\text{in} \langle \text{meas}(Hx), \text{meas} y \rangle) : qbit \rightarrow \text{bit} \otimes \text{bit}
29 \lambda_2, 28 \triangleright \lambda q_2.\lambda q_1.(\text{let}: \langle x, y \rangle = \text{CNOT}(q_1, q_2)
\text{in} \langle \text{meas}(Hx), \text{meas} y \rangle)!:(qbit \rightarrow \text{bit} \otimes \text{bit})
\end{verbatim}

Computing the type of U:

\begin{verbatim}
30 var, 1  q:qbit \triangleright q:qbit
31 const, 3  \triangleright U_{ij}:qbit \rightarrow qbit
32 app, 30, 31 q:qbit \triangleright U_{ij}:qbit
33 var, 1  y:bit \triangleright y:!bit
34 var, 1  x:bit \triangleright x:!bit
35 if, 33, 32, 32 q:qbit, y:bit \triangleright \text{if} y \text{ then } U_{11}q \text{ else } U_{10}q:qbit
36 if, 34, 35, 35 q:qbit, x:bit, y:bit \triangleright \text{if} x \text{ then } (\text{if} y \text{ then } U_{11}q \text{ else } U_{10}q)\text{ else} (\text{if} y \text{ then } U_{01}q \text{ else } U_{00}q):qbit
37 \lambda_1, 36 q:qbit \triangleright \lambda (x, y).(\text{if} x \text{ then } (\text{if} y \text{ then } U_{11}q \text{ else } U_{10}q))
\text{ else} (\text{if} y \text{ then } U_{01}q \text{ else } U_{00}q):qbit \otimes \text{bit} \rightarrow \text{qbit}
38 \lambda_2, 37 \triangleright \lambda q.\lambda (x, y).(\text{if} x \text{ then } (\text{if} y \text{ then } U_{11}q \text{ else } U_{10}q)
\end{verbatim}
\[ \text{6.5. Examples of Algorithms} \]

\[ \text{else (if y then } U_{01q} \text{ else } U_{00q}) :! (qbit \to (\text{bit} \otimes \text{bit} \to \text{qbit}) \]

Finally, computing the type of the pair \( (f, g) \):

\[
\begin{align*}
39 & \quad \top \quad \triangleright \ast : \top \\
40 & (L \ \text{6.4.2}) \quad \triangleright \ EPR : \top \to qbit \otimes qbit \\
41 & \quad \text{app}, 40, 39 \quad \triangleright \ EPR*: qbit \otimes qbit \\
42 & (L \ \text{6.4.2}) \quad \triangleright \ BellMeasure : qbit \to (qbit \to \text{bit} \otimes \text{bit}) \\
43 & \quad \text{var}, 1 \quad x : qbit \triangleright x : qbit \\
44 & \quad \text{app}, 42, 43 \quad x : qbit \triangleright \ BellMeasure \quad x : qbit \to \text{bit} \otimes \text{bit} \\
45 & \quad \text{var}, 1 \quad y : qbit \triangleright y : qbit \\
46 & (L \ \text{6.4.2}) \quad \triangleright \ U : qbit \to (\text{bit} \otimes \text{bit} \to qbit) \\
47 & \quad \text{app}, 46, 45 \quad y : qbit \triangleright \ U \ y : \text{bit} \otimes \text{bit} \to qbit \\
48 & \quad \text{var}, 1 \quad f : qbit \to \text{bit} \otimes \text{bit} \triangleright f : qbit \to \text{bit} \otimes \text{bit} \\
49 & \quad \text{var}, 1 \quad g : \text{bit} \otimes \text{bit} \to qbit \triangleright g : \text{bit} \otimes \text{bit} \to qbit \\
50 & \quad \otimes, 48, 49 \quad g : \text{bit} \otimes \text{bit} \to qbit, f : qbit \to \text{bit} \otimes \text{bit} \triangleright (f, g) : (qbit \to (\text{bit} \otimes \text{bit}) \otimes (\text{bit} \otimes \text{bit} \to qbit)) \\
51 & \quad \text{let}, 47, 50 \quad f : qbit \to \text{bit} \otimes \text{bit}, y : qbit \triangleright \text{letg} = U \ y \ \text{in} \ (f, g) : (qbit \to (\text{bit} \otimes \text{bit}) \otimes (\text{bit} \otimes \text{bit} \to qbit)) \\
52 & \quad \text{let}, 44, 51 \quad x : qbit, y : qbit \triangleright \text{letf} = \text{BellMeasure} \ x \ \text{in} \ \text{letg} = U \ y \ \text{in} \ (f, g) : (qbit \to (\text{bit} \otimes \text{bit}) \otimes (\text{bit} \otimes \text{bit} \to qbit)) \\
53 & \quad \text{let}, 41, 52 \quad \triangleright \text{let}(x, y) = EPR* \ \text{in} \ \text{letf} = \text{BellMeasure} \ x \ \text{in} \ \text{letg} = U \ y \ \text{in} \ (f, g)) : (qbit \to (\text{bit} \otimes \text{bit}) \otimes (\text{bit} \otimes \text{bit} \to qbit))
\]

\[ \text{6.5.4 Reduction of the Teleportation Term} \]

As an illustration of the reduction rules of the quantum lambda calculus we show the detailed reduction of the term from the teleportation example from Section 6.4.2. The reduction of the teleportation term corresponds to the equality \( g \circ f = \text{id} \). We use the following abbreviations:

\[
\begin{align*}
M_{p, p'} & := \text{let } f = \text{BellMeasure } p \ \text{in} \ \text{let } g = U \ p' \ \text{in} \ g(f \ p_0) \\
B_{p_1} & := \lambda q_1. (\text{let } \langle p, p' \rangle = \text{CNOT}(q_1, p_1) \ \text{in} \ \langle \text{meas}(H \ p), \text{meas} \ p' \rangle) \\
U_{p_2} & := \lambda (x, y). (\text{if } x \ \text{then } (\text{if } y \ \text{then } U_{11} \ p_2 \ \text{else } U_{10} \ p_2) \ \text{else } (\text{if } y \ \text{then } U_{01} \ p_2 \ \text{else } U_{00} \ p_2))
\end{align*}
\]

The reduction of the term is then as follows:

\[
\begin{align*}
& \quad \left[ \begin{array}{c}
\text{let } \langle p, p' \rangle = \text{EPR*} \\
\text{let } f = \text{BellMeasure } p \\
\text{in } \text{let } g = U \ p' \\
\text{in } g(f \ p_0)
\end{array} \right] \\
\quad \rightarrow_1 \left[ \langle \alpha(0) + \beta(1) \rangle \otimes (0), \text{let } \langle p, p' \rangle = \text{CNOT}(H(\text{new } 0), \text{new } 0) \ \text{in} \ M_{p, p'} \right] \\
\quad \rightarrow_1 \left[ (\langle \alpha(0) + \beta(1) \rangle \otimes |0 \rangle, \text{let } \langle p, p' \rangle = \text{CNOT}(H \ p_1, \text{new } 0) \ \text{in} \ M_{p, p'} \right]
\end{align*}
\]
6.5. Examples of Algorithms

\[ \rightarrow_1 \left[ (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \text{let } \langle p, p' \rangle = \text{CNOT}\langle p_1, \text{new } 0 \rangle \text{ in } M_{p,p'} \right] \]

\[ \rightarrow_1 \left[ (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle, \text{let } \langle p, p' \rangle = \text{CNOT}\langle p_1, p_2 \rangle \text{ in } M_{p,p'} \right] \]

\[ \rightarrow_1 \left[ (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \text{let } \langle p, p' \rangle = \langle p_1, p_2 \rangle \text{ in } M_{p_1,p_2} \right] \]

\[ \rightarrow_1 \left[ (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \text{let } f = \text{BellMeasure } p \right. \]

\[ \left. \text{in let } g = U_{p_2} \text{ in } g(f \circ p_0) \right] \]

\[ \rightarrow_1^* \left[ (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), U_{p_2}(B_{p_0}, p_0) \right] \]

\[ \rightarrow_1 \left[ (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), U_{p_2} \left( \text{let } \langle p, p' \rangle = \text{CNOT}\langle p_0, p_1 \rangle \text{ in } \langle \text{meas}(H_p), \text{meas } p' \rangle \right) \right] \]

\[ \rightarrow_1 \left[ \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \alpha(011) \\
+ \beta(110) + \beta(101) \end{array} \right), U_{p_2} \left( \text{let } \langle p, p' \rangle = \langle p_0, p_1 \rangle \text{ in } \langle \text{meas}(H_p), \text{meas } p' \rangle \right) \right] \]

\[ \rightarrow_1 \left[ \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \alpha(011) \\
+ \beta(110) + \beta(101) \end{array} \right), U_{p_2} \langle \text{meas } p_0, \text{meas } p_1 \rangle \right] \]

\[ \rightarrow_1 \left[ \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \alpha(011) \\
+ \beta(110) + \beta(101) \end{array} \right), U_{p_2} \langle \text{meas } p_0, \text{meas } p_1 \rangle \right] \]

\[ \left\{ \begin{array}{c} \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \alpha(011) \\
+ \beta(110) + \beta(101) \end{array} \right), U_{p_2} \langle 0, \text{meas } p_1 \rangle \right\} \]

\[ \left\{ \begin{array}{c} \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \beta(001) \\
+ \beta(010) + \beta(011) \end{array} \right), U_{p_2} \langle 1, \text{meas } p_1 \rangle \right\} \]

\[ \left\{ \begin{array}{c} \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \beta(001) \\
+ \beta(010) + \beta(011) \end{array} \right), U_{p_2} \langle 0, 0 \rangle \right\} \rightarrow_1^* \left\{ \begin{array}{c} \alpha(000) + \beta(001) \\
+ \beta(010) + \beta(011) \end{array} \right), U_{p_2} \langle 0, 0 \rangle \right\} \]

\[ \left\{ \begin{array}{c} \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(100) - \beta(101) \\
+ \beta(110) - \beta(111) \end{array} \right), U_{p_2} \langle 1, 0 \rangle \right\} \rightarrow_1^* \left\{ \begin{array}{c} \alpha(100) - \beta(101) \\
+ \beta(110) - \beta(111) \end{array} \right), U_{p_2} \langle 1, 0 \rangle \right\} \]

\[ \left\{ \begin{array}{c} \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(000) + \beta(001) \\
+ \beta(010) + \beta(011) \end{array} \right), U_{p_2} \langle 0, 1 \rangle \right\} \rightarrow_1^* \left\{ \begin{array}{c} \alpha(000) + \beta(001) \\
+ \beta(010) + \beta(011) \end{array} \right), U_{p_2} \langle 0, 1 \rangle \right\} \]

\[ \left\{ \begin{array}{c} \frac{1}{\sqrt{2}} \left( \begin{array}{c} \alpha(100) - \beta(101) \\
+ \beta(110) - \beta(111) \end{array} \right), U_{p_2} \langle 1, 1 \rangle \right\} \rightarrow_1^* \left\{ \begin{array}{c} \alpha(100) - \beta(101) \\
+ \beta(110) - \beta(111) \end{array} \right), U_{p_2} \langle 1, 1 \rangle \right\} \]

6.5.5 Reduction of the Superdense Coding Term

As another example of the reduction rules, we give the reduction of the superdense coding example from Section 6.5.3. This reduction shows the equality \( f \circ g = \text{id} \). Of the four possible cases, we only give one case, namely \( (f \circ g)(0,1) = (0,1) \); the remaining cases are similar. We use the same abbreviations as above.

\[ \left[ \begin{array}{c} \text{let } \langle p, p' \rangle = \text{EPR}\ast \\
\text{let } f = \text{BellMeasure } p \right. \]

\[ \left. \text{in let } g = U_{p'} \text{ in } f(g(0,1)) \right] \]
6.6 Towards a Denotational Semantics

In this chapter, we described a lambda calculus for dealing with quantum information. Given a typing judgement
\[ x : A \vdash M : B, \]
we can now give an operational description of the behavior of the term \( M \).

Another approach for understanding the language is to understand this expression as the syntactic representation of a function from some space \( A \) to some space \( B \). A denotational semantics is the description of a theory of mathematically structured spaces, where types are interpreted as spaces and typing judgements as structure-preserving maps between these spaces.

The difficulty in the case of the quantum lambda calculus is two-fold:

1. We need to account for probabilistic and quantum effects,

2. while keeping track of duplicable and non-duplicable data.

The remainder of this thesis is devoted to the study of these problems.
Chapter 7

The Linear Fragment

In this chapter, we restrict the quantum lambda calculus to the purely linear case, namely the fragment of the language where each value, classical and quantum, must be used exactly once. The linear quantum lambda calculus differs from its nonlinear cousin in that it is less sensitive to the evaluation order of terms. We give a denotational semantics for this language in a category of completely positive maps, and we show that it is fully abstract with respect to the operational semantics.

The plan of the chapter is as follows. First we briefly describe the language and the type system. Then we develop for it an operational semantics similar to the one developed in Section 6.2.4 and we define a notion of axiomatic equivalence of terms. Finally we build a denotational semantics for the language, and we show the soundness of the axiomatic equivalence with respect to the operational semantics, as well as the full abstraction of the denotational semantics. An extended abstract of the results presented in this chapter has been published in Selinger and Valiron (2006b).

7.1 A Linear Lambda Calculus for Quantum Computation

In Chapter 6, we defined a lambda calculus for quantum computation with classical control. This language was defined in a Curry-style manner. Although this makes sense from the point of view of the programmer, this typing paradigm is harder to tackle for the semantics. Indeed, we shall define the interpretation of a valid typing judgement by induction on its typing derivation. A one-to-one correspondence between typing judgements and typing derivation is therefore required.

We begin by re-adapting the definitions and results from Chapter 6 to the requirements of this study.

Definition 7.1.1. The linear quantum lambda calculus has the following (explicitly typed) terms:

Types $A, B ::= \text{bit} | \text{qbit} | A \rightarrow B | \top | A \otimes B$,

Term $M, N, P ::= x^A | MN | \lambda x^A.M | (M, N) | \text{let } (x^A, y^B) = M \text{ in } N |$

if $M$ then $N$ else $P | 0 | 1 | \text{new} | \text{meas} | U |

let $x = M \text{ in } N | * | \Omega_{x_{a_1}, \ldots, x_{a_n}}.$

$\Omega$ is a non-terminating term and let $x = M \text{ in } N$ is used to match 0-tuples. The remaining terms are as in Definition 6.1.1. We identify terms up to $\alpha$-equivalence. The set of free variables of a term $M$ is written $FV(M)$. In the following, we often write $c$ for an arbitrary constant of the language, i.e., $0, 1, \text{meas}, \text{new}, U,$ or $*$. 
7.1. A Linear Lambda Calculus for Quantum Computation

\[
\begin{align*}
x : A &\triangleright x^A : A \\
\Gamma_1 \triangleright P : \text{bit} &\quad \Gamma_2 \triangleright M : A \quad \Gamma_2 \triangleright N : A \\
\Gamma_1, \Gamma_2 \triangleright \text{if } P \text{ then } M \text{ else } N : A &
\end{align*}
\]

\[
\begin{align*}
\Gamma_1 \triangleright M : A \rightarrow B &\quad \Gamma_2 \triangleright N : A \\
\Gamma_1, \Gamma_2 \triangleright MN : B &
\end{align*}
\]

\[
\begin{align*}
\Gamma_1 \triangleright M_1 : A_1 &\quad \Gamma_2 \triangleright M_2 : A_2 \\
\Gamma_1, \Gamma_2 \triangleright \langle M_1, M_2 \rangle : A_1 \otimes A_2 &
\end{align*}
\]

\[
\begin{align*}
\Gamma_1 \triangleright M : A_1 \otimes A_2 &\quad \Gamma_2, x_1 : A_1, x_2 : A_2 \triangleright N : A \\
\Gamma_1, \Gamma_2 \triangleright \text{let } \langle x_1, x_2 \rangle = M \text{ in } N : A &
\end{align*}
\]

\[
\begin{align*}
\Gamma_1 \triangleright M : \top &\quad \Gamma_2 \triangleright N : A \\
\Gamma_1, \Gamma_2 \triangleright \text{let } * = M \text{ in } N : A &
\end{align*}
\]

Table 7.1: Typing rules for the linear quantum lambda calculus

We use the same shortcuts as in Conventions 6.1.2 and 6.3.2. Also, as we did in the previous chapter, to each term constant \(c\), we associate a fixed type \(A_c\), namely

\[
\begin{align*}
\text{meas} &\triangleright \text{qbit} \rightarrow \text{bit}, \\
0 &\triangleright \text{bit}, \\
\text{new} &\triangleright \text{bit} \rightarrow \text{qbit}, \\
U &\triangleright \text{qbit}^n \rightarrow \text{qbit}^n.
\end{align*}
\]

The only difference is the suppression of the type constructor “!” and, therefore, of subtyping.

As before, a typing context \(\Delta\) is a set of typed variables, written as \(\Delta = \{x_1 : A_1, \ldots, x_n : A_n\}\). A typing judgement is an expression \(\Delta \triangleright M : A\), where \(\Delta\) is a list of distinct typed variables called a typing context, \(M\) is a term, and \(A\) is a type. We say that a typing judgement is valid if it follows from the typing rules given in Table 7.1. In this table, if \(\Delta = \{x_1 : A_1, \ldots, x_n : A_n\}\) we write \(\Omega^\Delta\) in place of \(\Omega^A_{x_1, \ldots, x_n}\).

**Remark 7.1.2.** Note that, unlike the language of the previous chapter, there is no type constructor \(!A\). Also note that we added the term construct \(\text{let } * = M \text{ in } N\). Operationally, if we understand a “command” as a function that outputs the unit term \(*\), this term is used for being able to “use” a command in a program. Semantically, it is simply the destructor associated to \(\top\).

**Definition 7.1.3.** Given a term \(M\), we define the set of its free variables using the same definition as in Definition 6.1.3, modulo the indexing on the term. We extend the definition of \(FV\) to the terms \(\Omega\) and \(\text{let } * = M \text{ in } N\) as follows:

\[
FV(\Omega^A_{x_1, \ldots, x_n}) = \{x_1, \ldots, x_n\}, \quad FV(\text{let } * = M \text{ in } N) = FV(M) \cup FV(N).
\]

**Lemma 7.1.4.** Suppose that \(\Delta \triangleright M : A\) is a valid typing judgement. Then the set of free variables \(FV(M)\) is equal to \(|\Delta|\).

**Proof.** Consider a typing derivation \(\pi\) of \(\Delta \triangleright M : A\). We prove the result by induction on \(\pi\).

\((ax_1)\). The derivation is of the form \(x : A \triangleright x : A\), and \(FV(x) = \{x\}\).

\((ax_2)\). The derivation is of the form \(\triangleright c : A\), and \(FV(c) = \emptyset\).

\((\otimes I)\). The derivation is of the form \(\triangleright \ast : \top\), and \(FV(\ast) = \emptyset\).
(Ω). The derivation is of the form \( \Delta \triangleright \Omega_\Delta : A \), and \( FV(\Omega) = |\Delta| \).

\[ FV(if \; P \; then \; N_1 \; else \; N_2 : A, \; where \; \Delta \triangleright P : bit, \; \Delta \triangleright N_1 : A, \; \Delta \triangleright N_2 : A. \]

By induction hypothesis, \( FV(P) = |\Delta| \), \( FV(N_1) = |\Gamma| \) and \( FV(N_2) = |\Gamma| \). Therefore,

\[ FV(if \; P \; then \; N_1 \; else \; N_2) = FV(P) \cup FV(N_1) \cup FV(N_2) = |\Delta| \cup |\Gamma| = |\Delta, \Gamma|. \]

(\( \lambda \)). The derivation is of the form \( \Delta \triangleright \lambda x.N : A \rightarrow B \), where \( \Delta, x : A \rightarrow N : B \). By induction, \( FV(N) = |\Delta, x : A| \), which is \( |\Delta| \cup \{x\} \). Thus \( FV(\lambda x.N) = |\Delta| \).

(app). The derivation is of the form \( \Delta \triangleright N_1 N_2 : A \), where \( \Delta_1 \triangleright N_1 : C \) and \( \Delta_2 \triangleright N_2 : A \). By induction hypothesis, \( FV(N_1) = |\Gamma_1| \) and \( FV(N_2) = |\Gamma_2| \). Therefore, \( FV(N_1 N_2) = FV(N_1) \cup FV(N_2) = |\Gamma_1, \Gamma_2| \).

(\( \odot \)). The derivation is of the form \( \Delta \triangleright N_1 N_2 : B \odot C \), where \( \Delta \triangleright N_1 : C \) and \( \Delta_2 \triangleright N_2 : A \). By induction hypothesis, \( FV(N_1) = |\Gamma_1| \) and \( FV(N_2) = |\Gamma_2| \). Therefore, \( FV(N_1 N_2) = FV(N_1) \cup FV(N_2) = |\Gamma_1, \Gamma_2| \).

(\( \triangleright \)). The derivation is of the form \( \Delta \triangleright \langle N_1, N_2 \rangle : B \triangleright C \), where \( \Delta \triangleright N_1 : N_2 \). By induction hypothesis, \( FV(N_1) = |\Gamma_1| \) and \( FV(N_2) = |\Gamma_2| \). Therefore, \( FV(N_1 N_2) = FV(N_1) \cup FV(N_2) = |\Gamma_1, \Gamma_2| \).

(\( \triangleright \)). The derivation is of the form \( \Delta \triangleright \langle \langle x^B, y^C \rangle \rangle = \langle N_1, N_2 \rangle : A \), where \( \Delta \triangleright N_1 : A \) and \( \Delta_2 \triangleright N_2 : A \). By induction hypothesis, \( FV(N_1) = |\Gamma_1| \) and \( FV(N_2) = |\Gamma_2| \). Therefore, \( FV(\langle \langle x^B, y^C \rangle \rangle = \langle N_1, N_2 \rangle) = FV(N_1) \cup (FV(N_2) \setminus \{x, y\}) = |\Gamma_1, \Gamma_2| \).

Lemma 7.1.5. Consider a valid typing judgement \( \Delta \triangleright M : A \). Then the type \( A \) is completely determined by \( \Delta \) and \( M \), and there exists only one derivation for the typing judgement.

Proof. The proof is done by induction on the size of the term \( M \), using Lemma 7.1.4 and noticing that for every term construct only one possible typing rule does apply.

Definition 7.1.6. Let \( \Gamma \triangleright N : A \) be a valid typing judgement, let \( M \) be a term such that \( FV(M) \cap |\Gamma| = \emptyset \), and let \( x \) be a term variable. We define the substitution of \( x \) in \( M \) by \( N \), written \( M[N/x] \), as we did in Definition 7.1.4 (modulo the type indexation). We extend this definition to the two term constructs \( \Omega_\Delta^\text{c} \) and let \( \ast = M \) in \( N \) as follows. If \( \Delta = (\Delta', x : B) \) for some type \( B \), we define \( \Omega_\Delta^\text{c} N/x \) as the term \( \Omega_{\Delta'}^\text{c} \). If \( x \not\in |\Delta| \), the substitution \( \Omega_\Delta^\text{c} N/x \) is simply equal to \( \Omega_\Delta^\text{c} \). Finally, the rule for the last construct is \( (let \; \ast = M_1 \; in \; M_2)[N/x] = \langle let \; \ast = M_1 \; in \; M_2[N/x] \rangle \). Note that we only make use of the typing context of \( N \) in the rule concerning \( \Omega \).

Lemma 7.1.7. In the context of Definition 7.1.6, suppose that the variable \( x \) does not belong to \( FV(M) \). Then \( M[N/x] = M \).

Proof. The proof is done by induction on \( M \).

\( M \equiv y \). Then \( y \not\equiv x \), and for all \( N \), \( y[N/x] = y \).

\( M \equiv c \) and \( M \equiv \ast \). In both cases, \( M[N/x] = N \) by definition.

\( M \equiv \Omega_\Delta^\text{c} \). By definition, \( FV(\Omega) = |\Delta| \). Since \( x \) does not belong to \( FV(M) \), it does neither belong to \( |\Delta| \). Therefore, from the definition of the substitution, \( \Omega_{\Delta'}^\text{c} N/x = \Omega_{\Delta'}^\text{c} \Delta \).
7.1. A Linear Lambda Calculus for Quantum Computation

Let $M = \lambda y.P$. If $y = x$, then $x \notin \text{FV}(\lambda y.P)$ and by definition $(\lambda y.P)[N/x] = \lambda y.P$. If $y \neq x$, for $x$ not being a free variable of $M$ one needs $x$ not to be a free variable of $P$. In which case induction hypothesis applies: $P[N/x] = P$. Therefore, $M[N/x] = N$.

If $P = P_1P_2$, since $\text{FV}(P_1P_2) = \text{FV}(P_1) \cup \text{FV}(P_2)$, if $x \notin \text{FV}(M)$, $x$ is neither in $\text{FV}(P_1)$ nor in $\text{FV}(P_2)$. Since $(P_1P_2)[N/x] = (P_1[N/x])(P_2[N/x])$, by induction hypothesis this means that $(P_1P_2)[N/x] = P_1P_2$.

If $P = \text{let } (\Delta, x): \lambda y.P \text{ in } \text{P}_1 \text{ if } x \notin \text{FV}(M), x$ is neither in $\text{FV}(P_1)$ nor in $\text{FV}(P_2)$. Since $\text{FV}((P_1, P_2)) = \text{FV}(P_1) \cup \text{FV}(P_2)$, if $x \notin \text{FV}(M)$, $x$ is neither in $\text{FV}(P_1)$ nor in $\text{FV}(P_2)$. Since $(\text{let } (\Delta, x): \lambda y.P \text{ in } \text{P}_1 \text{ if } x \notin \text{FV}(M)) [N/x] = (P_1[N/x], x) \text{ else } P_3[N/x])$, by induction hypothesis this means that $(\text{let } (\Delta, x): \lambda y.P \text{ in } \text{P}_1 \text{ if } x \notin \text{FV}(M)) [N/x] = (P_1P_2)[N/x] = (P_1P_2)[N/x] = (P_1P_2)[N/x]$.

If $P = (\lambda y.P)$, then $x$ is not a free variable of $P$. Therefore, $M[N/x] = \lambda y.P$. For induction hypothesis applies:

- Case $P = P_1P_2$. Then $\text{FV}(P_1P_2) = \text{FV}(P_1) \cup \text{FV}(P_2)$, if $x \notin \text{FV}(M)$, $x$ is neither in $\text{FV}(P_1)$ nor in $\text{FV}(P_2)$. Since $(P_1P_2)[N/x] = (P_1[N/x])(P_2[N/x])$, by induction hypothesis this means that $(P_1P_2)[N/x] = (P_1P_2)[N/x]$.

- Case $P = (\lambda y.P)$. Then $x$ is not a free variable of $P$. Therefore, $M[N/x] = \lambda y.P$. For induction hypothesis applies:

  - If $P = P_1P_2$. Then $\text{FV}(P_1P_2) = \text{FV}(P_1) \cup \text{FV}(P_2)$, if $x \notin \text{FV}(M)$, $x$ is neither in $\text{FV}(P_1)$ nor in $\text{FV}(P_2)$. Since $(P_1P_2)[N/x] = (P_1[N/x](P_2[N/x]))$, by induction hypothesis this means that $(P_1P_2)[N/x] = (P_1P_2)[N/x]$.

- Case $P = (\lambda y.P)$. Then $x$ is not a free variable of $P$. Therefore, $M[N/x] = \lambda y.P$. For induction hypothesis applies:

  - If $P = P_1P_2$. Then $\text{FV}(P_1P_2) = \text{FV}(P_1) \cup \text{FV}(P_2)$, if $x \notin \text{FV}(M)$, $x$ is neither in $\text{FV}(P_1)$ nor in $\text{FV}(P_2)$. Since $(P_1P_2)[N/x] = (P_1[N/x](P_2[N/x]))$, by induction hypothesis this means that $(P_1P_2)[N/x] = (P_1P_2)[N/x]$.

Lemma 7.1.8 (Substitution). Suppose that $\Delta, x: A \Rightarrow M: B$ and $\Gamma \Rightarrow N: A$ are two valid typing judgements such that $|\Delta| \cap |\Gamma| = \emptyset$. Then $\Delta, \Gamma \Rightarrow M[N/x]: B$ is a valid typing judgement.

Proof. We prove the result by induction on the size of $M$.

Case $M = y$. The typing rule for obtaining $\Delta, x: A \Rightarrow M: B$ is $(ax_1)$. That is $A = B$, $x = y$ and $\Delta$ is empty. Thus $\Delta, \Gamma \Rightarrow M[N/x]: B$ becomes $\Gamma \Rightarrow N: A$, which is valid by hypothesis.

Case $M = c$. The typing rule for obtaining $\Delta, x: A \Rightarrow M: B$ is $(ax_2)$. Thus $A, x: A \Rightarrow c: B$ is valid if $(\Delta, x: A)$ is empty, and it is not. This case is therefore not to be considered.

Case $M = \Omega^B_{\Delta, x: A[N/x]}$. The typing rule for obtaining $\Delta, x: A \Rightarrow M: B$ is $(\Omega)$. Thus $\Delta' = (\Delta, x: A)$. Since $\Omega^B_{\Delta, x: A[N/x]} = \Omega^B_{\Delta', \Gamma}$, the result is obtained by applying back the typing rule $(\Omega)$.

Case $M = \lambda y^C.P$. The typing rule for obtaining $\Delta, x: A \Rightarrow M: B$ is $(\lambda)$. Thus $B = C \Rightarrow D$ and $\Delta, x: A, y: C \Rightarrow P: D$. A corollary is that $y \neq x$, thus that $M[N/x] = \lambda y^C.(P[N/x])$.

By induction hypothesis, $\Delta, \Gamma, y: C \Rightarrow P[N/x]: D$ is valid. Applying rule $(\lambda)$, we get $\Delta, \Gamma \Rightarrow \lambda y^C.(P[N/x]): C \Rightarrow D$, which is the desired result.
Case $M \equiv \text{if } P \text{ then } N_1 \text{ else } N_2$. The only possible typing rule for obtaining $\Delta, x : A \vdash M : B$ is (if). Then $(\Delta, x : A)$ splits into $(\Delta_1, \Delta_2)$, where

$$\Delta_1 \vdash P : \text{bit}, \quad \Delta_2 \vdash N_1 : B, \quad \Delta_2 \vdash N_2 : B.$$ 

Note that the contexts $\Delta_1$ and $\Delta_2$ do not intersect. There are two possible cases:

$x \in |\Delta_1|$. In this case $\Delta_0 = (\Delta'_1, x : A)$. We are in the setting of the lemma: by induction hypothesis, $\Delta'_1, \Gamma \vdash P[V/x] : \text{bit}$.

From Lemma 7.1.2 $FV(N_1) = |\Delta_2|$ and $FV(N_2) = |\Delta_2|$. Therefore $x$ is neither in $FV(N_1)$ nor in $FV(N_2)$. From Lemma 7.1.4 $M[N/x] = (if P[V/x] \text{ then } N_1 \text{ else } N_2)$.

Thus, applying (if), we get that $\Delta, \Gamma \vdash if P[V/x] \text{ then } N_1 \text{ else } N_2 : B$ is valid.

$x \in |\Delta_2|$. In this case $\Delta_1 = (\Delta'_2, x : A)$. We are in the setting of the lemma: by induction hypothesis, $\Delta'_2, \Gamma \vdash N_1[V/x] : B$ and $\Delta'_2, \Gamma \vdash N_2[V/x] : B$

From Lemma 7.1.2 $FV(P) = |\Delta_1|$. Therefore $x$ is not in $FV(P)$. From Lemma 7.1.4 $M[N/x] = (if P \text{ then } N_1[N/x] \text{ else } N_2[N/x])$.

Thus, applying (if), we get that $\Delta, \Gamma \vdash if P \text{ then } N_1[N/x] \text{ else } N_2[N/x] : B$ is valid.

Cases $M \equiv N_1 N_2$, $(N_1, N_2)$ and (let $* = N_1 \text{ in } N_2$). In each case the term $M$ can be written in the form $f(N_1, N_2)$, and there is only one possible typing rule for obtaining $\Delta, x : A \vdash M : B$, namely respectively $(app)$, $(\otimes. I)$ and $(\top. E)$. Then $(\Delta, x : A)$ splits into $(\Delta_1, \Delta_2)$, where $\Delta_1 \vdash N_1 : C_1$ and $\Delta_2 \vdash N_2 : C_2$, for some types $C_1$ and $C_2$ depending on the rule used:

<table>
<thead>
<tr>
<th>If the rule used is...</th>
<th>$f(N_1, N_2)$ is...</th>
<th>$C_1$ is...</th>
<th>$C_2$ is...</th>
<th>and $B$ is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>(app)</td>
<td>$N_1 N_2$</td>
<td>$C \rightarrow D$</td>
<td>$C$</td>
<td>$D$</td>
</tr>
<tr>
<td>(\otimes. I)</td>
<td>$(N_1, N_2)$</td>
<td>$C$</td>
<td>$D$</td>
<td>$C \otimes D$</td>
</tr>
<tr>
<td>(\top. E)</td>
<td>let $* = N_1 \text{ in } N_2$</td>
<td>$\top$</td>
<td>$C$</td>
<td>$C$</td>
</tr>
</tbody>
</table>

Note that for $i = 1, 2$, the contexts $\Delta_i$ and $\Gamma$ do not intersect. There are two possible cases: $x \in |\Delta_1|$, or $x \in |\Delta_2|$. Since they are symmetric by exchanging the role of the indices 1 and 2, we only prove one of them.

Suppose that $x \in |\Delta_1|$. In this case $\Delta_1 = (\Delta'_1, x : A)$. We are in the setting of the lemma: by induction hypothesis, $\Delta'_1, \Gamma \vdash N_1[V/x] : C_1$. From Lemma 7.1.2 $FV(N_2) = |\Delta_2|$. Therefore $x$ is not in $FV(N_2)$. From Lemma 7.1.4 $M[N/x] = f(N_1[N/x], N_2)$.

Applying back the typing rule used, the judgement $\Delta, \Gamma \vdash f(N_1[N/x], N_2) : B$ is valid.

Case $M \equiv (\text{let } (y^C, z^D) = N_1 \text{ in } N_2)$. This case is similar to the previous ones: There is only one possible typing rule for obtaining $\Delta, x : A \vdash M : B$, namely $(\otimes. E)$. Then $(\Delta, x : A)$ splits into $(\Delta_1, \Delta_2)$, where $\Delta_1 \vdash N_1 : C \otimes D$ and $\Delta_2, y : C, z : D \vdash N_2 : B$.

Note that for $i = 1, 2$, the contexts $\Delta_i$ and $\Gamma$ do not intersect. There are two possible cases: $x \in |\Delta_1|$, or $x \in |\Delta_2|$. 

$x \in |\Delta_1|$. In this case $\Delta_1 = (\Delta'_1, x : A)$. We are in the setting of the lemma: by induction hypothesis, $\Delta'_1, \Gamma \vdash N_1[V/x] : C \otimes D$. From Lemma 7.1.2 $FV(N_2) = |\Delta_2, y : C, z : D|$. Therefore $x$ is not in $FV(N_2)$. From Lemma 7.1.4 $M[N/x] = (\text{let } (y^C, z^D) = N_1[N/x] \text{ in } N_2)$.

Applying back the typing rule $(\otimes. E)$, the judgement $\Delta, \Gamma \vdash \text{let } (y^C, z^D) = N_1[N/x] \text{ in } N_2 : B$ is valid.
Definition 7.2.1. A *quantum closure* is a triple \([Q, L, M]\) where \(Q\) is a normalized vector in \(\otimes_{i=1}^{n} \mathbb{C}^2\), for some \(n \geq 0\), \(L\) is a bijective function from a set \(|L|\) of term variables to \(\{0, \ldots, n - 1\}\), and \(M\) is a term. As before, \(Q\) is called a *quantum array*, and \(L\) is called a *linking function*. Note that the only difference from Definition 6.2.1 lies in the fact that \(L\) is a bijective function. We use the notations of Definition 6.2.1 and consider quantum closure up to \(\alpha\)-equivalence. Finally, we use the notion of *well-typed quantum closure* and *program* of Definition 6.3.9.

Definition 7.2.2. We define the call-by-value reduction strategy on program by structural induction. For this purpose we need the notion of a value. As in Definition 6.2.13 we define a *value term* to be of the form

\[ V, W ::= x | c | * | \lambda x. M | (V, W) \]

A *value program* is a program of the form \([Q, L, V]\), where \(V\) is a value term. We set the rules to be the “classical” ones found in Table 6.2.2 plus the following “quantum” rules. In the first two rules, decompose \(Q\) into

\[ Q = \sum_j \alpha_j \psi_j^0 \otimes |0\rangle \otimes \tilde{\psi}_j^0 + \sum_j \beta_j \psi_j^1 \otimes |1\rangle \otimes \tilde{\psi}_j^1, \]

where \(\psi_j^0, \psi_j^1 \in \mathbb{C}^{2^{n-1}}\) and \(\tilde{\psi}_j^0, \tilde{\psi}_j^1 \in \mathbb{C}^{2^{n-1}}\) are basis vectors. Then, if

\[ \alpha = \sum_j |\alpha_j|^2, \quad \beta = \sum_j |\beta_j|^2, \]

\[ \begin{align*}
  [Q, |x_1 \cdots x_n\rangle, meas x_1] &\rightarrow_0 \sum_j [\psi_j^0 \otimes \tilde{\psi}_j^0, |x_1 \cdots x_{i-1} x_{i+1} \cdots x_n\rangle, 0], \\
  [Q, |x_1 \cdots x_n\rangle, meas x_1] &\rightarrow_0 \sum_j [\psi_j^1 \otimes \tilde{\psi}_j^1, |x_1 \cdots x_{i-1} x_{i+1} \cdots x_n\rangle, 1].
\end{align*} \]

In the second set of rules, if \(w\) is a fresh term variable not yet in use:

\[ \begin{align*}
  [Q, |x_1 \cdots x_n\rangle, new 0] &\rightarrow_0 \sum_j [Q \otimes |0\rangle, |x_1 \cdots x_n w\rangle, w].
\end{align*} \]
7.2. Operational Semantics

\[
\begin{align*}
[Q,L,\Omega] & \rightarrow^\gamma [Q,L,\Omega] & [Q,L,(\lambda x^B.M)V] & \rightarrow^\beta_1 [Q,L,M[V/x]] \\
[Q,L,let\ (x^B,y^C) = (V,W)\ in\ N] & \rightarrow^\omega [Q,L,N[V/x,W/y]] \\
[Q,L,let\ * = *\ in\ N] & \rightarrow^\gamma_1 [Q,L,N] \\
[Q,L,if\ 0\ then\ M\ else\ N] & \rightarrow^\downarrow_1 [Q,L,N] \\
[Q,L,if\ 1\ then\ M\ else\ N] & \rightarrow^\downarrow_1 [Q,L,M] \\
[Q,L,N] & \rightarrow_p^s [Q',L',N'] & [Q,L,M] & \rightarrow_p^s [Q',L',M'] \\
[Q,L,MN] & \rightarrow_p^s [Q',L',MN'] & [Q,L,MV] & \rightarrow_p^s [Q',L',MV] \\
[Q,L,\langle M_1,M_2 \rangle] & \rightarrow_p^s [Q',L',\langle M'_1,M'_2 \rangle] & [Q,L,\langle V_1,M_2 \rangle] & \rightarrow_p^s [Q',L',\langle V'_1,M'_2 \rangle] \\
[Q,L,P] & \rightarrow_p^s [Q',L',P'] \\
[Q,L,let\ P\ then\ M\ else\ N] & \rightarrow_p^s [Q',L',if\ P'\ then\ M\ else\ N] \\
[Q,L,M] & \rightarrow_p^s [Q',L',M'] \\
[Q,L,let\ (x_1,x_2) = M\ in\ N] & \rightarrow_p^s [Q',L',let\ \langle x_1,x_2 \rangle = M'\ in\ N] \\
[Q,L,M] & \rightarrow_p^s [Q',L',M'] \\
[Q,L,let\ * = M\ in\ N] & \rightarrow_p^s [Q',L',let\ * = M'\ in\ N]
\end{align*}
\]

Table 7.2: Reduction rules for the linear quantum lambda calculus

\[
[Q,|x_1 \cdots x_n\rangle, new\ 1] \rightarrow_n^{o_1} [Q \otimes |1\rangle, |x_1 \cdots x_n w\rangle, w].
\]

Finally, if \(Q'\) is the result of applying \(U\) to the quantum bits \(L(x_1), \ldots, L(x_n)\) in \(Q\):

\[
[Q,L,U\langle x_1, \ldots, x_n \rangle] \rightarrow^U [Q',L,\langle x_1, \ldots, x_n \rangle].
\]

Note that each reduction rule is named, and has a probability associated to it.

**Remark 7.2.3.** Note that since we want a linear language, we modified the measurement rule from the rule in Definition 6.2.4.3 by deleting the quantum bit measured from the quantum array. The two other differences from Table 6.1 are the (classical) rules concerning \(\Omega\) and let \(* = M\ in\ N.\) The former was defined to be a non-terminating term: it remains unchanged along the reduction. The latter matches 0-tuples. The rule follows the fact that the only 0-tuple is \(*.

**Definition 7.2.4.** A program \([Q,L,M]\) reducing with the \(\omega\)-rule is called a fixed point.

**Lemma 7.2.5.** Suppose that \(P\) is a value program. Then it does not reduce. Now suppose that \(P\) is a fixed point. Then the only possible programs \(P'\) such that \(P \rightarrow_p^s P'\) are fixed points.

**Proof.** For the former, proof by case distinction. For the latter, proof by induction on the reduction \(P \rightarrow_p^s P'.\) □

**Lemma 7.2.6.** Suppose that \(V\) is a value such that \(x_1 : qbit, \ldots, x_n : qbit \triangleright V : A \rightarrow B.\) Then \(V\) is either of the form \(\lambda x^A.N\) or it is one of the three term constants new, meas and \(U.\) If \(V\) is of type \(A \otimes B,\) then it is of the form \(\langle V_1, V_2 \rangle.\) If it is of type \(bit,\) it is either 0 or 1. Finally, if it is of type \(qbit,\) it is a term variable.

**Proof.** The proof for the type \(A \rightarrow B\) is done by case distinction: \(V\) cannot be a variable \(x_i\) since \(qbit\) is not of the form \(A \rightarrow B;\) it cannot be \(*\) and it cannot be of the form \(\langle V_1, V_2 \rangle.\) Therefore it is
either a lambda-abstraction or a term constant. The only term constants with the right types are the one offered in the lemma. The proof for the type $A \otimes B$ is done similarly. The proofs for the last two cases are done by case distinction.

### 7.2.2 Safety Properties

**Theorem 7.2.7** (Subject reduction). If $P$ is a program of type $A$ such that $P \rightarrow_{\rho} P'$, then $P'$ is also a program of type $A$.

**Proof.** We prove the result by induction on the derivation of the relation $P \rightarrow_{\rho} P'$. If $P \equiv [Q, L, M]$, we suppose that $\text{FV}(M) = \{x_1, \ldots, x_n\} \subseteq |L|$.

**Case ($\rightarrow_{\rho}^\rho$).** The only possible typing rule yielding $x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright (x^B, y^C)$ is $\langle V, W \rangle$ in $M : A$ is $(\otimes, E)$. It means that there exists $\Delta$ and $\Gamma$ such that $(\Delta, \Gamma) = (x_1 : \text{qbit}, \ldots, x_n : \text{qbit})$ and such that $\Delta \triangleright (x^B, y^C)$ is of $\text{qbit}$ and $\Gamma \triangleright [V, W] : M$ of $\text{qbit}$ is valid. The former can only be originated from typing rule $(\otimes)$. Thus $\Delta = (\Gamma_1, \Gamma_2)$ with $\Gamma_1 \triangleright V : B$ and $\Gamma_2 \triangleright W : C$.

From Lemma 7.2.4, $\Delta_1 \triangleright M[W/y] : A$ is valid. From Lemma 7.2.1, $\Delta_1, \Gamma_2 \triangleright M[V/x][W/y] : A$ is valid. Therefore the typing judgement $\Delta, \Gamma \triangleright M[V/x, W/y] : A$ is valid and the quantum context $[Q, [x_1 \ldots x_n], M[V/x, W/y]]$ is a program of type $A$.

**Case ($\rightarrow_{\Gamma}^\Gamma$).** The only possible typing rule for getting $x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright \ast = \ast$ in $N : A$ is $(\top, E)$. It means that there exists $\Delta$ and $\Gamma$ such that $(\Delta, \Gamma) = (x_1 : \text{qbit}, \ldots, x_n : \text{qbit})$ and such that $\Delta \triangleright \ast : \top$ and $\Gamma \triangleright M : A$ are valid.

The typing judgement $\Delta \triangleright \ast : \top$ being originated from $(\top, E)$, one has $|\Delta| = \emptyset$. Therefore $\Gamma = (x_1 : \text{qbit}, \ldots, x_n : \text{qbit})$.

This makes $[Q, [x_1 \ldots x_n], N]$ a program of type $A$.

**Cases ($\rightarrow_{\Gamma_0}^\Gamma$) and ($\rightarrow_{\Gamma_1}^\Gamma$).** The only possible typing rule yielding $x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright \text{if } 0 \text{ then } M \text{ else } N : A$ is (if). It means that there exists $\Delta$ and $\Gamma$ such that $(\Delta, \Gamma) = (x_1 : \text{qbit}, \ldots, x_n : \text{qbit})$ and such that $\Delta \triangleright 1 : \text{bit}$, $\Gamma \triangleright M : A$ and $\Gamma \triangleright N : A$ are valid.

The typing judgement $\Delta \triangleright 1 : \text{bit}$ being originated from $(\text{ax}_2)$, one has $|\Delta| = \emptyset$. Therefore $\Gamma = (x_1 : \text{qbit}, \ldots, x_n : \text{qbit})$.

Therefore, $[Q, [x_1 \ldots x_n], M]$ and $[Q, [x_1 \ldots x_n], N]$ are programs of type $A$.

**Case ($\rightarrow_{\gamma}^\gamma$).** In this case, the term $P$ and the term $P'$ are the same: therefore, if $P$ is a program of type $A$, so is $P'$.

**Cases ($\rightarrow_{m_0}^\rho$) and ($\rightarrow_{m_1}^\rho$).** In these cases, $\text{FV}(\text{meas } x_i) = \{x_i\}$. The typing judgement $x_i : \text{qbit} \triangleright \text{meas } x_i : A$ being valid, since $\text{meas} : \text{qbit} \rightarrow \text{bit}$, by typing rule $(\text{app})$ we have $A = \text{bit}$. Therefore $\triangleright 0 : A$ and $\triangleright 1 : A$ are valid: in both cases, since one removes $x_i$ from $|L|$, the program $P'$ is of type $A$.

**Cases ($\rightarrow_{m_0}^\rho$) and ($\rightarrow_{m_1}^\rho$).** In these cases, $\text{FV}(\text{new } 0) = \text{FV}(\text{new } 1) = \emptyset$. The judgements $\triangleright \text{new } 0 : A$ and $\triangleright \text{new } 1 : A$ are valid, thus since $A_{\text{new}} = \text{bit} \rightarrow \text{qbit}$, the type $A$ is equal to $\text{qbit}$. Thus $P'$ is a program of type $A$. 

\[ \square \]
Case \((-1^0)\). Since \(A_U = qbit^\otimes n \rightarrow qbit^\otimes n\) for some \(n\), one has
\[
x_1 : qbit, \ldots, x_n : qbit \triangleright U(x_1, \ldots, x_n) : qbit^\otimes n.
\]

Thus \(A = qbit^\otimes n\). But we also have \(x_1 : qbit, \ldots, x_n : qbit \triangleright (x_1, \ldots, x_n) : qbit^\otimes n\). Therefore, \(P'\) is a program of type \(A\).

The cases corresponding to the induction steps are directly coming from Lemma 7.1.4.

**Theorem 7.2.8 (Progress).** If \(P\) is a program of type \(A\), then either \(P\) is a value program, or \(P\) reduces using the reduction system of Definition 7.2.2.

**Proof.** Consider a valid program \([Q, x_1 \ldots x_n], M]\) of type \(A\). Then \(x_1 : qbit, \ldots, x_n : qbit \triangleright M : A\) is valid. Let \(\pi\) be a typing derivation of this typing judgement. We prove the result by induction on \(\pi\).

**Cases** \((ax_1), (ax_2), (\top. I), (\lambda)\). In all of these cases, the term \(M\) is a value. Thus \(P\) is a value program.

**Case** \((if)\). The term \(M\) is of the form \(if M' then N_1 else N_2\). There exist contexts \(\Delta\) and \(\Gamma\) such that \((\Delta, \Gamma) = (x_1 : qbit, \ldots, x_n : qbit)\) and such that \(\Delta \triangleright M' : bit\) and \(\Gamma \triangleright N_1, N_2 : A\). From Lemma 7.1.4, \(\Delta = FV(M')\). Thus the quantum context \([Q, x_1 \ldots x_n], M'\] is a program of type \(bit\). By induction hypothesis, either it is a value program or it reduces to some closure \([Q', L', M''\].

The only value of type \(bit\) being 0 and 1, if it is a value, then \(M'\) is respectively the term constant 0 and 1, and \(P\) reduces using respectively rules \((-\varepsilon_0)\) and \((-\varepsilon_1)\). If it is not a value, then by congruence \(P\) reduces to \([Q', L', if M''N_1N_2]\).

Thus, in any case the program \(P\) reduces.

**Case** \((app)\). The term \(M\) is of the form \(N_1N_2\). There exist contexts \(\Delta\) and \(\Gamma\) such that \((\Delta, \Gamma) = (x_1 : qbit, \ldots, x_n : qbit)\) and such that \(\Delta \triangleright N_1 : B \rightarrow A\) and \(\Gamma \triangleright N_2 : B\) for some type \(B\).

If \(N_2\) is not a value: then the program \([Q, x_1 \ldots x_n], N_2]\) reduces by induction hypothesis, and by congruence we are done.

If \(N_2\) is a value, and if \(N_1\) is not a value, then this time \([Q, x_1 \ldots x_n], N_1]\) reduces by induction hypothesis, and by congruence we are done.

If \(N_2\) and \(N_1\) are values, then from Lemma 7.2.4, \(N_1\) can only be one of the following:

- \(N_1 \equiv \lambda x.B.N\). Then \(P\) reduces using rule \((-\beta)\)
- \(N_1 \equiv new\). Then \(B = bit\). The only values of type \(bit\) being 0 and 1, the program \(P\) reduces using either rule \((-\varepsilon_0)\) or \((-\varepsilon_1)\).
- \(N_1 \equiv meas\). Then \(B = qbit\). The only values of type \(qbit\) being variables, \(N_2\) must be a term variable. Since \(|\Gamma| \subseteq \{x_1 \ldots x_n\}\), \(N_2 = x_j\) for some \(j\). Then \(P\) reduces using either rule \((-\varepsilon_0)\) or \((-\varepsilon_1)\).
- \(N_1 \equiv U^k\). Then \(B = qbit^\otimes k\). The only values of type \(qbit^\otimes k\) being of the form \(\langle y_1, \ldots y_k\rangle\), and since \(|\Gamma| \subseteq \{x_1 \ldots x_n\}\), \(N_2 = \langle x_{\sigma(1)}, \ldots, x_{\sigma(k)}\rangle\) for some injective map \(\sigma : \{1, \ldots, k\} \rightarrow \{1, \ldots, n\}\). Then \(P\) reduces using rule \((-\rightarrow U)\).

**Case** \((\otimes. I)\). The term \(M\) is of the form \(\langle N_1, N_2 \rangle\) and \(A = B_1 \otimes B_2\). There exist contexts \(\Delta\) and \(\Gamma\) such that \((\Delta, \Gamma) = (x_1 : qbit, \ldots, x_n : qbit)\) and such that \(\Delta \triangleright N_1 : B_1\) and \(\Gamma \triangleright N_2 : B_2\).

If \(N_1\) is not a value: then the program \([Q, x_1 \ldots x_n], N_1]\) reduces by induction hypothesis, and by congruence we are done.
If \( N_1 \) is a value but not \( N_2 \), then this time \( [Q, x_1 \ldots x_n], N_2 \) reduces by induction hypothesis, and by congruence we are done.

If they are both value, so is \( M \) and by definition \( P \) is a value program.

**Case** \((\otimes E)\). The term \( M \) is of the form let \( \la x^B, y^C \ra = N_1 \) in \( N_2 \). There exist contexts \( \Delta \) and \( \Gamma \) such that \((\Delta, \Gamma) = (x_1 : \text{qbit}, \ldots , x_n : \text{qbit})\) and such that \( \Delta \triangleright N_1 : B \otimes C \) and \( \Gamma, x : B, y : C \triangleright N_2 : A \) are valid.

If \( N_1 \) is not a value: then since \( |\Delta| \subseteq \{x_1 \ldots x_n\} \), the program \([Q, x_1 \ldots x_n], N_1\] reduces by induction hypothesis and \( P \) reduces using congruence.

If \( N_1 \) is a value, then it is of the form \( \{V, W\} \), and then \( P \) reduces using rule \((\to_\otimes)\).

**Case** \((\top E)\). The term \( M \) is of the form let \( * = N_1 \) in \( N_2 \). There exist contexts \( \Delta \) and \( \Gamma \) such that \((\Delta, \Gamma) = (x_1 : \text{qbit}, \ldots , x_n : \text{qbit})\) and such that \( \Delta \triangleright N_1 : \top \) and \( \Gamma \triangleright N_2 : A \) are valid.

If \( N_1 \) is not a value: then since \( |\Delta| \subseteq \{x_1 \ldots x_n\} \), the program \([Q, x_1 \ldots x_n], N_1\] reduces by induction hypothesis and \( P \) reduces using congruence.

If \( N_1 \) is a value, then it is equal to \( * \), and then \( P \) reduces using rule \((\to_\top)\).

**Case** \((\Omega)\). The term \( M \) is of the form \( \Omega_{x_1 \ldots x_n}^A \). Therefore \( P \) reduces using the reduction rule \((\to_\omega)\). \(\square\)

### 7.2.3 Normalization

The language being linear, the reduction has a strong relation on the length of the terms.

**Definition 7.2.9.** Let \( l(M) \) be the length of a term \( M \), defined recursively as an integer greater or equal to 1 as follows:

\[
l(x^A) = l(c) = l(*) = l(\Omega) = 1,
\]

\[
l(\lambda x^A. M) = 1 + l(M),
\]

\[
l(\text{let } \la x^B, y^C \ra = M \text{ in } N) = 1 + l(M) + l(N),
\]

\[
l(\text{let } * = M \text{ in } N) = 1 + l(M) + l(N),
\]

\[
l(\text{if } P \text{ then } M \text{ else } N) = 1 + l(P) + \max(l(M), l(N)),
\]

\[
l(MN) = 1 + l(M) + l(N),
\]

\[
l(MN) = 1 + l(M) + l(N).
\]

**Lemma 7.2.10.** Suppose that \( \Delta, x : A \triangleright M : B \) and \( \Gamma \triangleright N : A \) are valid. Then \( l(M[N/x]) < l(M) + l(N) \).

**Proof.** The proof is done by structural induction on \( M \).

\( M \equiv x^C \). The computation goes like this: \( l(x[N/x]) = l(N) < l(x) + l(N) \).

\( M \equiv c, \ast, \Omega, \) and \( y^C \) (with \( y \neq x \)). In all of these cases, \( M[N/x] = M \). Therefore \( l(M[N/x]) = l(M) \), which by definition is strictly smaller than \( l(M) + l(N) \).

\( M \equiv \lambda y^C. M' \). If \( y = x \) then \( M[N/x] = M \). As in the previous case, we obtain the result using the definition.

If \( y \neq x \), then \( M[N/x] = \lambda y^C. (M'[N/x]) \). By induction hypothesis, \( l(M'[N/x]) < l(M') + l(N) \). Therefore, since \( l(M) = 1 + l(M') \), \( l(M[N/x]) = 1 + l(M'[N/x]) < 1 + l(M') + l(N) = l(M) + l(N) \).
7.2. Operational Semantics

\[ M \equiv \text{if } P \text{ then } N_1 \text{ else } N_2. \] We have \( l(M) = 1 + l(P) + \max(l(N_1), l(N_2)) \) and
\[ M[N/x] = \text{if } P[N/x] \text{ then } N_1[N/x] \text{ else } N_2[N/x]. \]

The valid judgement \( \Delta, x : A \triangleright M : B \) can only come from the typing rule \((i)\). That is, \((\Delta, x : A)\) splits into \((\Gamma_1, \Gamma_2)\), where \( \Gamma_1 \triangleright P : \text{bit} \) and \( \Gamma_2 \triangleright N_1, N_2 : A \).

If \( x \in |\Gamma_1| \), then \( x \notin |\Gamma_2| \), and from Lemma 7.1.7, \( x \) is neither in \( FV(N_1) \) nor in \( FV(N_2) \), therefore from Lemma 7.1.4, \( N_2[N/x] = N_2 \) and \( N_1[N/x] = N_1 \). The length of \( M[N/x] \) is then \( 1 + l(P[N/x]) + \max(l(N_1), l(N_2)) \). By induction hypothesis, \( l(P[N/x]) < l(P) + l(N) \), thus \( l(M[N/x]) < l(M) + l(N) \).

If \( x \in |\Gamma_2| \), then \( x \notin |\Gamma_1| \), and from Lemma 7.1.7, \( x \notin FV(P) \), therefore from Lemma 7.1.4, \( P[N/x] = N_1 \). The length of \( M[N/x] \) is then
\[ 1 + l(P) + \max(l(N_2[N/x]), l(N_1[N/x])). \]

By induction hypothesis, \( l(N_1[N/x]) < l(N_1) + l(N) \) and \( l(N_2[N/x]) < l(N_2) + l(N) \), thus as expected \( l(M[N/x]) < l(M) + l(N) \).

In the four remaining cases. The term \( M \) is respectively of the form
\[
\begin{align*}
&\text{let } ⟨y^C, z^D⟩ = N_1 \text{ in } N_2, \\
&\text{let } s = N_1 \text{ in } N_2, \\
&\text{let } s = N_1[N/x] \text{ in } N_2[N/x], \\
&\text{let } s = N_1[N/x] \text{ in } N_2[N/x],
\end{align*}
\]

In all of these cases, we have \( l(M) = l(N_1) + l(N_2) + 1 \). In the first cases, by \( \alpha \)-equivalence one can assume that \( x \neq y \). Therefore, \( M[N/x] \) becomes respectively:
\[
\begin{align*}
&\text{let } ⟨y^C, z^D⟩ = N_1[N/x] \text{ in } N_2[N/x], \\
&\text{let } s = N_1[N/x] \text{ in } N_2[N/x], \\
&\text{let } s = N_1[N/x] \text{ in } N_2[N/x], \\
&\text{let } s = N_1[N/x] \text{ in } N_2[N/x],
\end{align*}
\]

The valid judgement \( \Delta, x : A \triangleright M : B \) can only come from respectively the typing rule \((\otimes.E)\), \((\text{app})\), \((\vee.E)\), \((\otimes.I)\). That is, \((\Delta, x : A)\) splits into \((\Gamma_1, \Gamma_2)\), where
\[ \Gamma_1 \triangleright N_1 : C_1, \quad \Gamma_2, \Lambda \triangleright N_2 : C_2, \]

for some types \( C_1, C_2 \) and some context \( \Lambda \), with \( x \notin |\Lambda| \).

If \( x \in |\Gamma_1| \), then \( x \notin |\Gamma_2| \), and from Lemma 7.1.4, \( x \notin FV(N_2) \), therefore from Lemma 7.1.4, \( N_2[N/x] = N_2 \). The length of \( M[N/x] \) is then \( 1 + l(N_1[N/x]) + l(N_2) \). By induction hypothesis, \( l(N_1[N/x]) < l(N_1) + l(N) \), thus \( l(M[N/x]) < l(M) + l(N) \).

If \( x \in |\Gamma_2| \), then \( x \notin |\Gamma_1| \), and from Lemma 7.1.4, \( x \notin FV(N_1) \), therefore from Lemma 7.1.4, \( N_1[N/x] = N_1 \). The length of \( M[N/x] \) is then \( 1 + l(N_2[N/x]) + l(N_1) \). By induction hypothesis, \( l(N_2[N/x]) < l(N_2) + l(N) \), thus as expected \( l(M[N/x]) < l(M) + l(N) \).

\[ \square \]

**Lemma 7.2.11.** If \([Q, L, M]\) is a program such that \([Q, L, M] \rightarrow^\omega [Q', L', M']\), then either \( x \neq \omega \) and \( l(M') < l(M) \) or \( \omega = \omega \) and \( l(M') = l(M) \). In the latter case, \([Q, L, M] = [Q', L', M']\).

**Proof.** The proof is done by induction on the derivation of the reduction \([Q, L, M] \rightarrow^\omega [Q', L', M']\).

**Case \( \rightarrow^\omega \).** The term \( M \) and the term \( M' \) are respectively of the form \((Ax^A.N)N'\) and \( N'[N/x] \).

Since \( P \) is a program, \( M \) is well-typed: By Lemma 7.2.10, since \( l(M) = l(N) + l(N) + 2 \), the length \( l(M') \) verifies \( l(M') < l(N') + l(N) < l(M) \).
Lemma 7.2.5) or a value (in which case it will remain so for all larger such that
\[ l_n(N) < l(N) + l(V) + l(W) < l(M). \]

Case \((-\tau)\). The term \(M\) and the term \(M'\) are respectively of the form (let \(* = *\ in N) and \(N\).
Since \(l(M) = 2 + l(N)\), as requested \(l(M') < l(M)\).

Cases \((-\mu_0^L)\) and \((-\mu_1^L)\). The term \(M\) and the term \(M'\) are respectively of the form if \(0\) then \(N\ else N'\, or \(1\) then \(N\ else N'\ and \(N\). In both cases \(l(M) = 2 + \max(l(N), l(N'))\), this strictly larger than both \(l(N)\) and \(l(N')\: the inequality \(l(M') < l(M)\) is verified.

Case \((-\omega)\). In this case, \(\kappa = \omega\ and \([Q, L, M] = [Q', L', M']\).
Cases \((-\omega_0)\) and \((-\omega_1)\). The term \(M\) has the form \(\text{meas} \: x\): \(l(M) = 3\). The term \(M'\) is the term constant \(0\) or the term constant \(1\): \(l(M') = 1 < l(M)\).
Cases \((-\omega_0)\) and \((-\omega_1)\). The term \(M\) has the form \(\text{new} \: 0\ or \text{new} \: 1\: l(M) = 3\). The term \(M'\) is a term variable: \(l(M') = 1 < l(M)\).
Case \((-\mu)\). The term \(M\) is of the form \(U(x_1, \ldots, x_n)\) and the term \(M'\) the form \(x_1, \ldots, x_n\).

Therefore
\[ l(M') = l(x_1, \ldots, x_n) < 2 + l(x_1, \ldots, x_n) = 1 + l(U) + l(x_1, \ldots, x_n) = l(M). \]

The cases corresponding to the congruence are straightforward using the induction hypothesis.

**Theorem 7.2.12** (Normalization). Let \(P = [Q, L, M] \ be a program, and let \((P_n)_n\) be a sequence
of programs \(P_n = [Q_n, L_n, M_n]\) such that \(P_0 = P\ and such that for all \(n, either P_n \to_{P_n} P_{n+1}\ or P_n = P_{n+1} if P_n\ does not reduce. Then \(P_{l(M)}\ is either a value or a fixed point, and for all \(n > l(M)\), \(P_n = P_{l(M)}\).

**Proof.** If \((P_n)_n\) is the sequence described in the hypothesis, we show by induction on \(n\) that either
\(P_n\ is a value, or it is a fixed point, or \(l(M_n) \leq l(M) - n\: if \(n = 0\: since \(P_0 = P\, we are done. Suppose the result is true for \(n\). Then either \(P_n\ is a value, or from Theorem 7.2.8 there exists \(P'\ such that \(P_n \to P'\, and thus \(P_{n+1}\ is such that \(P_n \to_{\kappa_n} P_{n+1}\). From Lemma 7.2.11 either \(\kappa_n = \omega\ or \(l(M_{n+1}) < l(M_n)\). Since by induction hypothesis \(l(M_n) \leq l(M) - n\, by definition we have \(l(M_{n+1}) \leq l(M) - (n + 1)\).

We have \(l(M_n) \geq 1\ for all \(n\). Therefore the \(n\) for which \(P_n\ reduce are smaller than \(l(M)\).
And for \(n = l(M)\), \(P_n\ is either a fixed point (in which case it will remain so for all larger \(n\ by Lemma 7.2.12\) or a value (in which case it will remain so for all larger \(n\ since values do not reduce, also by Lemma 7.2.5). \)

**7.2.4 Quantum Context and Reduction**

When describing the reduction rules, we carefully separated the quantum context from the lambda-term.
In this subsection we show that the precise order of quantum bits in the quantum array does not matter.

**Definition 7.2.13.** If \(\sigma\ is a permutation of \(\{1, \ldots, n\}\), we extend \(\sigma\ to \(\mathbb{N}\ with \(\sigma(j) = j for j > n\),
and we define \(\tilde{\sigma}\ to be the corresponding permutation of quantum bits
\[ \tilde{\sigma}\{x_1 \cdots x_n \cdots x_{n+k}\} = \{x_{\sigma(1)} \cdots x_{\sigma(n)} x_{n+1} \cdots x_{n+k}\}. \]
We say that \((Q_1, L_1)\) is \(\sigma\)-equivalent to \((Q_2, L_2)\) if \(Q_1\) and \(Q_2\) have the same size, \(Q_2 = \sigma(Q_1)\) and \(L_2 = \sigma^{-1}\circ L_1\). We write \((Q_1, L_1) =_\sigma (Q_2, L_2)\). We define an equivalence relation called \(\alpha\)-equivalence on quantum contexts by \((Q_1, L_1) =_\alpha (Q_2, L_2)\) if there exists a \(\sigma\) such that \((Q_1, L_1) =_\sigma^* (Q_2, L_2)\).

The \(\alpha\)-equivalence is sound with respect to the semantics:

**Lemma 7.2.14.** Suppose that \([Q, x_1 \ldots x_n], M \xrightarrow{p} [Q', y_1 \ldots y_m], M'\]. Then for all permutations \(\sigma\) on \(\{1, \ldots, n\}\), there exists a permutation \(\tau\) on \(\{1, \ldots, m\}\) such that \([\sigma Q, \sigma[x_1 \ldots x_n], M \xrightarrow{p} [\sigma Q', \sigma[y_1 \ldots y_m], M']\].

**Proof.** The proof is done by induction on the derivation of the reduction \([Q, x_1 \ldots x_n], M \xrightarrow{p} [Q', y_1 \ldots y_m], M'\]. In the classical cases of Table 8.2, since the quantum context is not touched, choose \(\tau = \sigma\) and apply the same rule that was used at first. For the rules \((\rightarrow_0^\alpha)\) and \((\rightarrow_{-1}^\alpha)\), again, since \(\sigma\) only act on the \(n\) first quantum bits of the context, applying the permutation before or after the addition of a new quantum bit to the quantum array does not change the output. Thus, choosing \(\tau = \sigma \cup \{n + 1 \mapsto n + 1\}\) solves the problem.

Now, suppose that we are in the case of the rule \((\rightarrow_U^0)\). In this situation we apply a unitary map on some quantum bits in the array, referred to by \(L\). Since we perform \(\sigma\) both on \(L\) and on \(Q\), it does not modify the actual unitary map performed. Therefore, again choosing \(\tau = \sigma\) gives us the answer.

In the cases \((\rightarrow_0^\alpha)\) and \((\rightarrow_{-1}^\alpha)\), the sizes of the quantum array and of the linking function change along the reduction by the removal of the entity number \(i\). However, in the program \([\sigma Q, \sigma[x_1 \ldots x_n], \text{meas} x_{\sigma(i)}]\) the reduction operation will remove entity number \(\sigma(i)\). Therefore, if \(f\) and \(g\) are the maps

\[
\begin{align*}
  f(j) &= \begin{cases} 
    j & \text{if } j < i, \\
    j + 1 & \text{if } j \geq i,
  \end{cases} \\
  g(j) &= \begin{cases} 
    j & \text{if } j < \sigma(i), \\
    j - 1 & \text{if } j \geq \sigma(i),
  \end{cases}
\end{align*}
\]

define \(\tau\) as \(g \circ \sigma \circ f\).

Finally, the induction cases are easy application of the induction hypothesis. \(\square\)

### 7.2.5 Reduction to Values

In the reduction process, what we are really seeking is the final result of the computation. In this section we explicate the relation of programs to values.

**Definition 7.2.15.** We use the notion defined in Section 6.2.3. If \(X\) is the set of closed valid programs and \(U\) the set of values, let \(\text{prob}_U : X \times U \rightarrow [0, 1]\) be the map \(\text{prob}_U(P, V)\) that returns the total probability for a program \(P\) to end up on the value \(V\) in zero or more steps. This function is called the big-step reduction.

We also define the small-step reduction operation \(\text{prob} : X \times X \rightarrow [0, 1]\): for closed programs \(P, P'\), we define \(\text{prob}(P, P') = p\) if there is a single-step probabilistic reduction \(P \xrightarrow{p} P', \text{prob}(V, V) = 1\) if \(V\) is a value program, and \(\text{prob}(P, P') = 0\) in all other cases. Note that for all well-typed \(P\), \(\sum_{P' \in X} \text{prob}(P, P') = 1\).

**Lemma 7.2.16.** The reduction system defined in Definition 7.2.3 is a probabilistic reduction system.

**Proof.** By case distinction. \(\square\)

**Definition 7.2.17.** Given a set \(Z\), let \(\mathcal{C}Z\) be the set of probability distributions over it. We curry the small-step reduction \(\text{prob} : X \times X \rightarrow [0, 1]\) to the map \(\text{prob}' : \mathcal{C}X \rightarrow \mathcal{C}X\): \(\text{prob}'(\sum_i \alpha_i P_i) = \sum_i \alpha_i \sum_{P' \in X} \text{prob}(P_i, P'_i)\). Similarly, we curry the function \(\text{prob}_U : X \times U \rightarrow [0, 1]\) to the map \(\text{prob}'_U : \mathcal{C}X \rightarrow \mathcal{C}U\): \(\text{prob}'_U(\sum_i \alpha_i P_i) = \sum_i \alpha_i \text{prob}_U(P_i, V)\).
Definition 7.2.18. If $P$ is a closed well-typed program of type $\text{bit}$, and $b \in \{0, 1\}$, we define $(P \downarrow b) = \sum_{V \in U_b} \text{prob}'(P, V)$, where $U_b$ is the set of valid programs with term the value $b$. We say that $P$ evaluates to $b$ with probability $P \downarrow b$. The definition of $P \downarrow b$ can be extended in the same way to probability distributions of programs.

Lemma 7.2.19. Let $P : A$ be a program such that $P \rightarrow^\kappa_\rho P'$. Then either $\text{prob}'(P) = P'$, $\rho = 1$ and $\kappa \neq m_0, m_1$, or $\text{prob}'(P) = \alpha P_0 + \beta P_1$, with $P \rightarrow^{m_0} P_0$ and $P \rightarrow^{m_1} P_1$.

Proof. By induction on the reduction $P \rightarrow^\kappa_\rho P'$.

Definition 7.2.20. We define a formal probability distribution of quantum closures to be $\Gamma \models \sum_i \rho_i [Q_i, L_i, M_i] : A$, where each $\Gamma \models [Q_i, L_i, M_i] : A$ is valid and $\sum \rho_i \leq 1$. The distribution is said to be closed if $|\Gamma| = \emptyset$.

Lemma 7.2.21. If $\models P : A$ is valid, so is $\models \text{prob}'P : A$ and $\models \text{prob}'U : A$.

Proof. The lemma is a corollary of Theorem 7.2.7.

Due to the strong normalization theorem, applying the map $\text{prob}'U$ is applying the map $\text{prob}'$ finitely many times.

Theorem 7.2.22. If $[Q, L, M]$ is a program, then the reduction satisfies

$$\text{prob}^{d(M)}[Q, L, M] = \sum_{i=1}^n \rho_i [Q_i, L_i, V_i] + \sum_{j=1}^m \rho_j' P_j,$$

with $\text{prob}'U[Q, L, M] = \sum_{i=1}^n \rho_i [Q_i, L_i, V_i]$, the $P_j$ being fixed points, $\sum \rho_i + \sum \rho_j' = 1$, and $n + m < 2^{d(M)}$.

Proof. Corollary of Theorem 7.2.7.

7.3 Denotational Semantics

In Section 4.3, the notion of completely positive map is used to model the notion of quantum computation. We aim to show that the linear subset of the quantum lambda calculus has the category $\text{CPM}$ as a fully abstract model. Note that the interpretation will not be “onto” all completely positive maps, but will only be “onto up to scalar multiplies”.

7.3.1 Modeling the Linear Quantum Lambda Calculus

We interpret the language in the category defined in Section 4.3.2.

Lemma 7.3.1. The following statements are valid.

1. Let $\sigma, \tau$ be any signatures. Then $0 : \sigma \rightarrow \tau$ is completely positive.

2. The maps

$$
\begin{align*}
t &: (a, b) &\rightarrow& 2, \\
p &: (a \ b) &\rightarrow& (a, d)
\end{align*}
$$

are completely positive.
Proof. The maps defined in the lemma are completely positive since they have positive characteristic matrices.

\[\begin{align*}
[x : A \triangleright x : A](v) &= v & [\triangleright 0 : \text{bit}](x) &= (x, 0) \\
[\triangleright * : \top](x) &= x & [\triangleright 1 : \text{bit}](x) &= (0, x) \\
[\Delta \triangleright \Omega : A] &= 0 : [\Delta] \rightarrow [A] \\
[\triangleright \text{new} : \text{bit} \rightarrow \text{qbit}] &= \Phi(t) : 1 \rightarrow [\text{bit}] \otimes [\text{qbit}] \\
[\triangleright \text{meas} : \text{qbit} \rightarrow \text{bit}] &= \Phi(p) : 1 \rightarrow [\text{qbit}] \otimes [\text{bit}] \\
[\triangleright U : \text{qbit}^2 \rightarrow \text{qbit}] &= \Phi(U) : 1 \rightarrow 2^{2n} \\
[\Delta, x : A \triangleright M : B] &= f : [\Delta] \otimes [A] \rightarrow [B] \\
[\Delta \triangleright \lambda x. M : A \rightarrow B] &= \Phi(f) : [\Delta] \rightarrow [A] \otimes [B] \\
[\Delta \triangleright M : A \rightarrow B] &= \Phi(g) : [\Delta] \rightarrow [A] \otimes [B] \\
[\Gamma \triangleright N : A] &= f : [\Gamma] \rightarrow [A] \\
[\Delta, \Gamma \triangleright MN : B] &= x \otimes y \mapsto g(x \otimes (fy)) : [\Delta] \otimes [\Gamma] \rightarrow [B] \\
[\Gamma \triangleright P : \text{bit}] &= x \mapsto (px, qx) : [\Delta] \rightarrow [\text{bit}] \\
[\Delta \triangleright M : A \otimes B] &= f : [\Delta] \rightarrow [A] \otimes [B] \\
[\Gamma \triangleright N : A] &= g : [\Gamma] \rightarrow [A] \\
[\Delta, \Gamma \triangleright \text{let} \langle x, y \rangle = M \text{ in } N : C] &= u \otimes v \mapsto g(v \otimes (fu)) : [\Delta] \otimes [\Gamma] \rightarrow [C] \\
[\Delta \triangleright \text{let} * = M \text{ in } N : C] &= u \otimes v \mapsto (fu)(gv) : [\Delta] \otimes [\Gamma] \rightarrow [C] \\
\end{align*}\]

Table 7.3: Denotational semantics for typing judgements.

3. The maps

\[
\begin{align*}
1 &\mapsto 1, 1 & 1 &\mapsto 1, 1 \\
a &\mapsto (a, 0) & a &\mapsto (0, a).
\end{align*}
\]

are completely positive.

Proof. The maps defined in the lemma are completely positive since they have positive characteristic matrices.

Definition 7.3.2. We set the denotation of types to be

\[
[\text{bit}] = (1, 1), \quad [A \otimes B] = [A] \otimes [B], \\
[qbit] = (2), \quad [A \rightarrow B] = [A] \otimes [B],
\]

and the denotation of contexts to be

\[
[x_1 : A_1, \ldots, x_n : A_n] = [A_1] \otimes \cdots \otimes [A_n], \quad [\emptyset] = 1.
\]

The denotation of a typing judgement of the form $\Delta \triangleright M : A$ is a linear map

\[
[\Delta \triangleright M : A] : [\Delta] \rightarrow [A],
\]

\[
[\Delta \triangleright \lambda x. M : A \rightarrow B] = \Phi(f) : [\Delta] \rightarrow [A] \otimes [B],
\]

\[
[\Delta \triangleright M : A \rightarrow B] = \Phi(g) : [\Delta] \rightarrow [A] \otimes [B],
\]

\[
[\Gamma \triangleright N : A] = f : [\Gamma] \rightarrow [A] \\
[\Delta, \Gamma \triangleright MN : B] = x \otimes y \mapsto g(x \otimes (fy)) : [\Delta] \otimes [\Gamma] \rightarrow [B] \\
[\Gamma \triangleright P : \text{bit}] = x \mapsto (px, qx) : [\Delta] \rightarrow [\text{bit}] \\
[\Delta \triangleright M : A \otimes B] = f : [\Delta] \rightarrow [A] \otimes [B] \\
[\Gamma \triangleright N : A] = g : [\Gamma] \rightarrow [A] \\
[\Delta, \Gamma \triangleright \text{let} \langle x, y \rangle = M \text{ in } N : C] = u \otimes v \mapsto g(v \otimes (fu)) : [\Delta] \otimes [\Gamma] \rightarrow [C] \\
[\Delta \triangleright \text{let} * = M \text{ in } N : C] = u \otimes v \mapsto (fu)(gv) : [\Delta] \otimes [\Gamma] \rightarrow [C]
\]

\[
[\Delta \triangleright M : A] : [\Delta] \rightarrow [A],
\]
7.3. Denotational Semantics

defined inductively on the typing derivation of \( \Delta \vdash M : A \) as in Table 7.3. Note that from Lemma 7.3.1, this typing derivation is unique, so the definition makes sense. Here, \( \Phi : \text{hom}(\sigma \otimes \sigma', \tau) \to \text{hom}(\sigma, \sigma' \otimes \tau) \) is the bijection of Definition 4.3.4 and \( \iota \) and \( p \) are respectively the quantum bits creation and the measurement operation of Lemma 7.3.1.

Definition 7.3.3. We also define the denotation of a valid quantum closure. Consider a valid quantum closure \( \Delta \vDash [Q, L, M] : A \) where \( L = |x_1 \cdots x_n y_1 \cdots y_m \rangle \) and \( |Q| = n + m \). The quantum context \( (Q, L) \) can be seen as a map \( g : 1 \to 2^{\otimes n} \otimes 2^{\otimes m} \) such that \( g(1) = Q \). Then if \( f \) is the morphism

\[
[\Delta, x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \vdash M : A] : [\Delta] \otimes 2^{\otimes n} \to [A],
\]

one defines \( [\Delta \vDash [Q, L, M] : A] \) as the composition

\[
[\Delta] \xrightarrow{id[\Delta] \otimes g} [\Delta] \otimes 2^{\otimes n} \otimes 2^{\otimes m} \xrightarrow{f \otimes id_{2^{\otimes m}}} [A] \otimes 2^{\otimes m} \xrightarrow{id[\Delta] \otimes \text{tr}} [A].
\]

One extends this definition to probabilistic distributions of valid quantum closures using linearity:

\[
[\Delta \vDash \sum_i \rho_i P_i : A] = \sum_i \rho_i [\Delta \vDash P_i : A].
\]

Lemma 7.3.4. Let \( \Delta \vdash M : A \) be a valid typing judgement. Then \( [\Delta \vdash M : A] \) is a completely positive map. Let \( \Gamma \vDash [Q, L, M] : A \) be a valid quantum closure. Then \( [\Gamma \vDash [Q, L, M] : A] \) is a completely positive map.

Proof. We prove that \( [\Delta \vdash M : A] \) is a completely positive map by induction on the typing derivation of \( \Delta \vdash M : A \).

(\(ax_1\)) and (\(\top.I\)). The identity map is completely positive since \( \text{CPM} \) is a category.

(\(ax_2\)). The denotations of the term constants \text{meas}, \text{new}, \text{U}, 0 and 1 are completely positive map from Lemma 7.3.1.

(\(\Omega\)). The map 0 is completely positive from Lemma 7.3.1.

(\(\lambda\)). From the monoidal closedness of \( \text{CPM} \), the image of \( f \) by \( \Phi \) is completely positive.

(app). Since \( \Phi^{-1} \) sends completely positive matrices to completely positive matrices. Thus \( g \) is completely positive. The final denotation is completely positive since composition and tensors of completely positive maps are completely positive.

(if). The denotation can be reformulated as

\[
[\Delta] \otimes [\Gamma] \xrightarrow{[P] \otimes id} [\Gamma] \otimes (1 \oplus 1) \otimes [\Gamma] = [\Gamma] \otimes [\Gamma] \xrightarrow{f \otimes g} [A] \otimes [A] \xrightarrow{\text{tr} \otimes id_{[A]}} [A].
\]

(\(\otimes.I\)), (\(\otimes.E\)) and (\(\top.E\)). In these three cases, the denotation is the tensor of two completely positive maps: it is therefore a completely positive map.

This proves that the denotation of a valid term is a completely positive map. The fact that the denotation of \( \Gamma \vDash [Q, L, M] : A \) is also completely positive comes from the fact that completely positive maps are closed under composition and tensor, and that the trace operator and the identity are themselves completely positive maps.
Lemma 7.3.5 (Substitution). Suppose that $|\Gamma| \cap |\Delta| = \emptyset$ and
\[
[\Delta, x : A \triangleright M : B] = G : [\Delta] \otimes [A] \rightarrow [B],
\]
\[
[\Gamma \triangleright N : A] = F : [\Gamma] \rightarrow [A],
\]
\[
[\Delta, \Gamma \triangleright M[N/x] : B] = H : [\Delta] \otimes [\Gamma] \rightarrow [B].
\]
Then for all $d \in V_{[\Delta]}$ and all $g \in V_{[\Gamma]}$, we have $H(d \otimes g) = G(d \otimes (Fg))$.

Proof. We show that $H = (id_{[\Delta]} \otimes F); G$ by induction on the typing derivation of $\Delta, x : A \triangleright M : B$.

(\(ax_1\)). We are considering the judgement $x : A \triangleright x : A$. We have $|\Delta| = \emptyset$, $A = B$ and $F = id_{[A]}$.
Thus $x[N/x] = N$, we have as required $H = G = F ; G$.

(\(ax_2\)) and (\(\top J\)). These cases are not to be considered: the typing contexts must contain $(x : A)$.

(\(\Omega\)). Since $\Omega[N/x] = \Omega$, and since $(id_{[\Delta]} \otimes 0); G = 0$, the result is true.

(\(\lambda\)). In this case, $M$ is of the form $\lambda y^C.M'$ and $A$ is of the form $C \rightarrow D$. Also, the typing judgement $\Delta, y : C, x : A \triangleright M' : D$ is valid. Note that $y \neq x$. By induction hypothesis, if $G'$ is the map $[\Delta, y : C, x : A \triangleright M' : D]$, the map $[\Delta, y : C, \Gamma \triangleright M'[N/x] : D]$ is $H'$ verifying for all $d \in V_{[\Delta]}$, for all $g \in V_{[\Gamma]}$.
Let $d \in V_{[\Delta]}$.
\[
\Phi(H')(d \otimes g) = \sum_{c \in B_{[C]}} c \otimes H'(d \otimes c \otimes g) = \sum_{c \in B_{[C]}} c \otimes G'(d \otimes c \otimes F(g)) = \Phi(G'(d \otimes F(g))),
\]
that is, $[\Delta, \Gamma \triangleright (\lambda y^C.M')][N/x] : C \rightarrow D]$.

(app). The term $M$ is of the form $M_1M_2$ and the context $(\Delta, x : A)$ splits into $(\Delta_1, \Delta_2)$, such that $\Delta_1 \triangleright M_1 : C \rightarrow B$ and $\Delta_2 \triangleright M_2 : C$ for some type $C$.
Let $G_1 = [\Delta_1 \triangleright M_1 : C \rightarrow B]$, $G_2 = [\Delta_2 \triangleright M_2 : C]$.

There are two cases

$x \in |\Delta_1|$. The context $\Delta_1$ splits into $(\Delta'_1, x : A)$. From Lemma 7.3.4, the variable $x$ belongs to $FV(M_1)$ but not to $FV(M_2)$.

Since $G(d \otimes a \otimes d') = \Phi^{-1}(G_1)(d \otimes G_2(d') \otimes F(g))$, we have as required
\[
H(d \otimes d' \otimes g) = G(d \otimes d' \otimes F(g)).
\]

$x \in |\Delta_2|$. The context $\Delta_1.frm~e$ splits into $(\Delta'_2, x : A)$. From Lemma 7.3.4, $x$ belongs to $FV(M_2)$ but not to $FV(M_1)$.

Since $G(d \otimes a \otimes d') = \Phi^{-1}(G_1)(d' \otimes G_2(d \otimes a))$, we have as required
\[
H(d' \otimes d \otimes g) = G(d' \otimes d \otimes F(g)).
\]
(if). The term $M$ is of the form if $P$ then $M_1$ else $M_2$ and the context $(\Delta, x : A)$ splits into $(\Delta_1, \Delta_2)$, such that $\Delta_1 \vdash P : \text{bit}$ and $\Delta_2 \vdash M_1, M_2 : A$. Let

$$(p, q) = [\Delta_1 \vdash P : \text{bit}], \quad G_1 = [\Delta_2 \vdash M_1 : A], \quad G_2 = [\Delta_2 \vdash M_2 : A].$$

There are two cases.

$x \in |\Delta_1|$. The context $\Delta_1$ splits into $(\Delta_1', x : A)$. From Lemma 7.2.7, $x \in \text{FV}(P)$ and $x$ is either in $\text{FV}(M_1)$ nor in $\text{FV}(M_2)$. From Lemma 7.2.7, $M[N/x] = \text{if } P[N/x] \text{ then } M_1 \text{ else } M_2$. Suppose that $g \in V[\Gamma_1], d' \in V[\Delta_1], d \in V[\Delta_2]$. Let $(p', q') = [\Delta_1', \Gamma \vdash P[N/x] : \text{bit}]$. By induction hypothesis, this is

$$(p'(d' \odot g), q'(d' \odot g)) = (p', q')(d' \odot g) = (p, q)(d' \odot F(g)), q(d' \odot F(g))),$$

thus $[\Delta_1', \Gamma, \Delta_2 \vdash P[N/x] \text{ then } M_1 \text{ else } M_2 : A]$ is

$$H(d' \odot g \odot d) = p(d' \odot F(g))G_1(d) + q(d' \odot F(g))G_2(d) = G(d' \odot d \odot F(g)),$$

that is the requested form.

$x \in |\Delta_2|$. The context $\Delta_2$ splits into $(\Delta_2', x : A)$. From Lemma 7.2.7, $x \in \text{FV}(P)$ and $x$ is either in $\text{FV}(M_1)$ nor in $\text{FV}(M_2)$. From Lemma 7.2.7, $M[N/x] = \text{if } P[N/x] \text{ then } M_1 \text{ else } M_2$. Suppose that $g \in V[\Gamma_1], d' \in V[\Delta_2], d \in V[\Delta_2]$. Let $G_1' = [\Delta_2', \Gamma \vdash M_1[N/x] : A], \quad G_2' = [\Delta_2', \Gamma \vdash M_2[N/x] : A].$

By induction hypothesis, this is

$$G_1' = G_1(d' \odot F(g)), \quad G_2' = G_2(d' \odot F(g)).$$

thus $[\Delta_1, \Delta_2', \Gamma \vdash P[N/x] \text{ else } M_2[N/x] : A]$ is

$$H(d \odot d' \odot g) = p(d)G_1(d' \odot F(g)) + q(d)G_2(d' \odot F(g)) = G(d \odot d' \odot F(g)),$$

which is in the requested form.

The three last cases are proved similarly.

\[ \Box \]

**Lemma 7.3.6.** Suppose that $P : A$ is a program of type $A$ such that $\text{prob}'P = \sum_i \rho_i P_i$, then

$$[P : A] = \sum_i \rho_i [P_i : A]$$

**Proof.** Suppose that $P = [Q, L, M]$, where $L = [x_1 \ldots x_n]$. If $P$ does not reduce, then $\text{prob}'P = P$ and we are done. If $P$ reduce to some $P'$ via some rule $\kappa$, we use Lemma 7.2.19 and proceed by case distinction:

First, suppose that $\kappa \neq m_0, m_1$. We prove the result by induction on a derivation of the reduction.

$(-\beta)$. In this case, $M \equiv (\lambda x^B.N)N'$, and then $P' = [Q, L, N[N'/x]]$. By definition and Theorem 7.2.20 since $P$ is of type $A$ we have

$$x_1 : \text{qbit, \ldots, } x_n : \text{qbit} \triangleright (\lambda x^B.N)N' : A, \quad x_1 : \text{qbit, \ldots, } x_n : \text{qbit} \triangleright N[N'/x] : A.$$

The former being valid, its context splits as $(\Delta, \Gamma)$ where $\Delta, x : B \triangleright N : A$ and $\Gamma \triangleright N' : B$. Let $F = [\Delta, x : B \triangleright N : A]$ and $G = [\Gamma \triangleright N' : B]$. By definition, $[\Delta, \Gamma \triangleright (\lambda x^B.N)N' : A]$ is equal to $d \odot g \mapsto F(d \odot G(g))$, if $d \in V[\Delta]$ and $g \in V[\Gamma]$. By Lemma 7.3.6 this is precisely the denotation of the latter. Thus, $[P' : A] = 1 \cdot [P : A]$. 


7.3. Denotational Semantics

$\rightarrow_1^\circ$. In this case, $M \equiv \langle \text{let } (x^B, y^C) = (V, W) \text{ in } N \rangle$, and then $P' = [Q, L, N[V/x, W/y]]$. By definition and Theorem 7.3.5, since $P$ is of type $A$ we have

\[
x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright M : A, \quad x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright N[V/x][W/y] : A.
\]

The former being valid, its context splits as $(\Delta, \Gamma_1, \Gamma_2)$ where $\Delta, x : B, y : C \triangleright N : A, \Gamma_1 \triangleright V : B$ and $\Gamma_2 \triangleright W : C$. Let

\[
F = [\Delta, x : B, y : C \triangleright N : A], \quad G_1 = [\Gamma_1 \triangleright V : B], \quad G_2 = [\Gamma_2 \triangleright W : C].
\]

By definition, $[\Gamma_1, \Gamma_2 \triangleright (V, W) : B \otimes C]$ is the map $G_1 \otimes G_2$. By definition still, the map $[\Delta, \Gamma_1, \Gamma_2 \triangleright M : A]$ is $((id \otimes G_1 \otimes G_2):F)$. By Lemma 7.3.5, this is the denotation of the latter. Thus, $[\models P' : A] = 1 \cdot [\models P : A].$

$\rightarrow^1_t$. In this case, $M \equiv \langle \text{let } \ast = \ast \text{ in } N \rangle$, and then $P' = [Q, L, N]$. Since $P$ is of type $A$ we have

\[
x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright M : A, \quad x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright N : A.
\]

Since $[\triangleright \ast : T] = id_1$, the denotation of the former is equal to the denotation of the latter: we have as requested $[\models P' : A] = 1 \cdot [\models P : A].$

$\rightarrow_{1q}^0$ and $\rightarrow_{1f}^0$. In this case, $M \equiv \langle (\text{if } 0 \text{ then } N \text{ else } N') \rangle$ (respectively $M \equiv \langle (\text{if } 1 \text{ then } N \text{ else } N') \rangle$), and then $P' = [Q, L, N']$. Since $P$ is of type $A$ we have

\[
x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright \text{if } 0 \text{ then } N \text{ else } N' : A, \\
x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright \text{if } 1 \text{ then } N \text{ else } N' : A, \\
x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \triangleright N, N' : A,
\]

and $\triangleright 0, 1 : \text{bit}$. Since $[\triangleright 0 : \text{bit}] = (a, 0)$ and $[\triangleright 1 : \text{bit}] = (a, 0)$, the denotation of $M$ is then equal to the denotation of $N'$ (respectively, of $N$). Therefore $[\models P' : A] = 1 \cdot [\models P : A].$

$\rightarrow^0_1$. Since $P = P'$, they have the same denotation.

$\rightarrow_{1q}^0$ and $\rightarrow_{1f}^1$. We prove the case $\rightarrow_{1q}^0$, the other one is similar. The term $M$ is of the form $\text{new } 0$, and $A = \text{qbit}$. The denotation of $\triangleright M : \text{qbit}$ is the map $f$ such that $f(1) = |0\rangle\langle 0|$. The denotation of $P$ is

\[
G = 1 \xrightarrow{g} 2^{\otimes n} \xrightarrow{f \otimes id} 2 \otimes 2^{\otimes n} \xrightarrow{id \otimes tr} 2,
\]

where $g$ is the map $g(1) = QQ$. Since $Q$ is a normalized vector, $G(1) = |0\rangle\langle 0|$. Now, $P' = [Q \otimes |0\rangle, |x_1 \ldots x_n w\rangle, w] : \text{qbit}$ and its denotation is

\[
G' = 1 \xrightarrow{g'} 2 \otimes 2^{\otimes n} \xrightarrow{id \otimes id} 2 \otimes 2^{\otimes n} \xrightarrow{id \otimes tr} 2,
\]

where $g'(1) = |0\rangle\langle 0| \otimes QQ$. The map $G'$ is therefore equal to $G$.

$\rightarrow_{1f}^1$. This case is similar: the denotation of $P$ is of the form

\[
G = 1 \xrightarrow{g} 2^{\otimes i} \otimes 2^{\otimes (n-i)} \xrightarrow{f \otimes id} 2^{\otimes i} \otimes 2^{\otimes (n-i)} \xrightarrow{id \otimes tr} 2^{\otimes i},
\]

where $F = \Phi^{-1}(U)$, and $g$ is $QQ$. The denotation of $P'$ is

\[
G' = 1 \xrightarrow{g'} 2^{\otimes i} \otimes 2^{\otimes (n-i)} \xrightarrow{id \otimes id} 2^{\otimes i} \otimes 2^{\otimes (n-i)} \xrightarrow{id \otimes tr} 2^{\otimes i},
\]

where $g'$ is $g; (\Phi^{-1}(U) \otimes id)$. Therefore they are the same map.
The induction cases are trivial.

Finally, suppose that \( k = m_0 \) or \( m_1 \). Then \( P \rightarrow^{m_0} P_0 \) and \( P \rightarrow^{m_1} P_0 \) for some \( P_0 \) and some \( P_1 \), we prove the result again by induction on the derivation of one of the reduction.

**Base case.** We use the notation in Definition \[\text{7.2.2}\]. In both cases, \( M \) is of the form \( \text{meas} \ x_1 \), and \( A = \text{bit} \). By \( \alpha \)-equivalence one can assume that \( i = 1 \): in this case \( Q \) is of the form

\[
\sum_j \alpha_j |0\rangle \otimes |\tilde{\psi}_j^0\rangle + \sum_j \beta_j |1\rangle \otimes |\tilde{\psi}_j^1\rangle = |0\rangle \otimes \sum_j \alpha_j |\tilde{\psi}_j^0\rangle + |1\rangle \otimes \sum_j \beta_j |\tilde{\psi}_j^1\rangle.
\]

Since the denotation \([x_1 : \text{qbit} \triangleright \text{meas} \ x_1 : \text{bit}]\) is the map \( p \) of Definition \[\text{7.3.9}\] the denotation \([= [Q, L, M] : \text{bit}]\) is the map

\[
G = 1 \xrightarrow{g} 2 \otimes 2^{\otimes (n-1)} \xrightarrow{prob \otimes tr} 1 \otimes 1.
\]

The value \( g(1) \) is \( QQ^\ast \), that is

\[
|0\rangle\langle 0| \otimes \sum_{j,k} \alpha_j \alpha_k^* |\tilde{\psi}_j^0\rangle \langle \tilde{\psi}_j^0| + |1\rangle\langle 1| \otimes \sum_{j,k} \beta_j \beta_k^* |\tilde{\psi}_j^1\rangle \langle \tilde{\psi}_j^1|
\]

\[
+ |0\rangle\langle 1| \otimes \sum_{j,k} \alpha_j \beta_k^* |\tilde{\psi}_j^0\rangle \langle \tilde{\psi}_j^1| + |1\rangle\langle 0| \otimes \sum_{j,k} \beta_j \alpha_k^* |\tilde{\psi}_j^1\rangle \langle \tilde{\psi}_j^0|.
\]

Since \( |\tilde{\psi}_j^0\rangle, |\tilde{\psi}_j^1\rangle, |\tilde{\psi}_j^0\rangle,\) and \( |\tilde{\psi}_j^1\rangle \) are basis vectors, applying \( id \otimes \text{tr} \) on \( g(1) \) yields

\[
\alpha |0\rangle\langle 0| + \beta |1\rangle\langle 1| + \delta |0\rangle\langle 1| + \gamma |1\rangle\langle 0|,
\]

for some values \( \delta, \gamma \), and where \( \alpha \) and \( \beta \) are as in Definition \[\text{7.2.2}\]. Thus, \( G(1) \) is precisely \( (\alpha, \beta) = [\text{prob}^\prime P : A] \).

The induction cases are trivial.

**Lemma 7.3.7.** Suppose that \( P : A \) is a fixed point. Then \([P]\) is the zero map.

**Proof.** Write \( P \) as \([Q, L, M]\). The proof is done by structural induction on \( M \).

**Theorem 7.3.8** (Soundness). Given a program \( P : A \),

\[
[P : A] = [\text{prob}^\prime L : P : A].
\]

**Proof.** By induction on \( n \), using Lemma \[7.3.7\] one can prove that \([\text{prob}^n P]\) = \([\text{prob} P]\). From Theorem \[7.2.2\] there exists an integer \( n \) such that \( \text{prob}^n P = \text{prob}^\prime P + \sum_i P_i \) with \( P_i \) being fixed points. Therefore, from Lemma \[7.3.7\] \([P : A]\) = \([\text{prob}^n P]\) = \([\text{prob}^\prime L : P : A]\).

### 7.3.2 Fullness of the First-Order Fragment

In this section we will show a result analogous to Theorem \[1.3.16\] namely that any superoperator is the image of a term. For proving this result, we use the same techniques as in (Selinger, 2004).

**Lemma 7.3.9.** Suppose that \( n = 2^s \) and \( m = 2^t \), and let \( F : C^{n \times n} \rightarrow C^{m \times m} \) be a contraction. Then there exists a valid typing judgement \( x_1 : \text{qbit}, \ldots, x_s : \text{qbit} \triangleright M : \text{qbit}^{\otimes t} \) whose denotation is the map \( F \).
Proof. By definition, the map $F$ can be written in the form $F(A) = UAU^*$, where $U \in \mathbb{C}^{m \times n}$ is a contraction. There exists matrices $U_1$, $U_2$ and $U_3$ such that
\[
U' = \begin{pmatrix} U_1 & U_2 \\ U_3 \\ \end{pmatrix}
\]
is unitary of size $\mathbb{C}^{t \times l}$. Without loss of generality, one can assume that $l = 2^p$: one can always extend $U'$ to
\[
\begin{pmatrix} U' & 0 \\ 0 & I_{2^p-l} \\ \end{pmatrix}
\]
where $I_{2^p-l}$ is the identity matrix. The requested program performs the following:

- Allocation of the variables $y_1 : \text{qbit}, \ldots, y_{l-s} : \text{qbit}$, and let the list of variables $(z_1, \ldots, z_t)$ be $(x_1, \ldots, x_s, y_1, \ldots, y_{l-s})$;
- application of the unitary $U'$ to $(z_1, \ldots, z_t)$;
- measurement of the variables $z_1, \ldots, z_t$: if the outcome is $(0, \ldots, 0)$, the program returns the term $(z_1, \ldots, z_t)$. Otherwise it outputs the term $\Omega_{x_1,y_1,\ldots,x_s,y_{l-s}}^{\text{bit} \otimes \text{qbit}}$.

Lemma 7.3.10. Let $F : \mathbb{C}^{2^s \times 2^l} \rightarrow \mathbb{C}^{2^s \times 2^l}$ be a superoperator and $E \circ G$ is decomposition from Theorem 7.3.3. One can construct another decomposition $F' \circ G'$, where $G' : \mathbb{C}^{2^s \times 2^l} \rightarrow \mathbb{C}^{2^{t+k} \times 2^{t+k}}$ is a contraction and $E' : \mathbb{C}^{2^{t+k} \times 2^{t+k}} \rightarrow \mathbb{C}^{2^s \times 2^l}$ is a partial trace operator, for some integer $k$.

Proof. The proof uses Lemma 7.3.12 to adjust the size of the output space of the sub-unitary $E'$ to a power of 2.

Lemma 7.3.11. Let $E : \mathbb{C}^{2^{t+k} \times 2^{t+k}} \rightarrow \mathbb{C}^{2^s \times 2^l}$ be a partial trace operator. Then there exists a valid typing judgement $x_1 : \text{bit}, \ldots, x_{t+k} : \text{bit} \triangleright M : \text{bit}^{\otimes t}$ whose denotation is $E$.

Proof. The requested term $M$ measures the variables $x_{t+1}, \ldots, x_{t+k}$, perform a test on the result and in all cases returns $(x_1, \ldots, x_t)$.

Lemma 7.3.12. Let $\sigma = (1,1)^{\otimes m} \otimes 2^{\otimes m}$, and $\sigma' = 2^{\otimes (m+n)}$. Then the measurement operator $\mu_\sigma : V_\sigma \rightarrow V_{\sigma'}$ is the image of some typing judgement $x : \text{bit}^{\otimes (m+n)} \triangleright M : \text{bit}^{\otimes m} \otimes \text{qbit}^{\otimes m}$.

Proof. The requested term $M$ measures the variables $x_{t+1}, \ldots, x_{t+k}$ and stores the results in $(z_1 : \text{bit}, \ldots, z_k : \text{bit})$. It outputs the product $(z_1, \ldots, z_k, x_1, \ldots, x_t)$.

Theorem 7.3.13. For any types $A = U_1 \otimes \cdots \otimes U_n$ and $B = V_1 \otimes \cdots \otimes V_m$, where each $U_i$ and $V_j$ is either bit or qbit, if $F : [A] \rightarrow [B]$ is any superoperator, then there exists a valid typing judgement $x : A \triangleright M : B$ of denotation $F$.

Proof. Without loss of generality, one can assume that $A = \text{bit}^{\otimes k} \otimes \text{qbit}^{\otimes l}$ and that $B = \text{bit}^{\otimes s} \otimes \text{qbit}^{\otimes t}$. The map $F : [A] \otimes [B]$ is of the form $(F_1, \ldots, F_{2^k})$, where each $F_i : V_{(1,1)^{\otimes i} \otimes \text{bit}^{\otimes 2^k}}$ is a superoperator. From Theorem 7.3.13, $F_i$ splits into $M_i \circ E_i \circ G_i$, where $G_i$ is a contraction, $E_i$ is a partial trace operator, and $M_i$ is a measurement operator. Using Lemma 7.3.12, one can assume that all of them goes from images of tensors of bit’s and qbit’s. Therefore, using Lemma 7.3.12 and Lemma 7.3.14, the $F_i$’s are representable by some terms $M_i$. The function $F$ is represented by a term that does a test on its input bit and perform the corresponding $M_i$.

Corollary 7.3.14. For every type $A = U_1 \otimes \cdots \otimes U_n$, where each $U_i$ is either bit or qbit, and for every hermitian positive element $v \in V_{[A]}$ of trace at most 1, $v = \triangleright M : A(1)$ for some closed valid term $M$. 

Proof. Consider the function \( F : \mathbb{C} \to V_{\{A\}} \) defined by \( F(1) = v \). Since \( \chi_F = v \) is positive, and since \( tr(v) \leq 1 \), \( F \) is a superoperator. From Theorem 7.3.13, it is the image of some typing judgement \( \triangleright M : A \). \hfill \Box

7.3.3 Fullness up to Scalar Multiple

Theorem 7.3.13 does not hold for generic types. However, a weaker statement does hold. We explicate it in this section.

Definition 7.3.15. Given a type \( A \), we define the canonical first-order representation \( \ll A \) of \( A \) by the following: \( \llbit = \text{bit} \), \( \llqbit = \text{qbit} \), \( \ll\top = \top \), \( \ll(A \otimes B) = \ll A \otimes \ll B \), \( \ll(A \triangleright B) = \ll A \triangleright \ll B \).

Lemma 7.3.16. For all types \( A \), \( [A] = [\ll A] \).

Proof. Proof by structural induction on \( A \). \hfill \Box

Definition 7.3.17. We define the \( \ll \)-size \( s(A) \) of a type \( A \) as follows. First, \( s(\text{bit}) = s(\text{qbit}) = s(\top) = 1 \). Then, \( s(A \triangleright B) = 2^{s(A)} + s(B) \). Finally, \( s(A \otimes B) = s(A) + s(B) \).

Lemma 7.3.18. For all types \( A \), \( s(A) \geq 1 \). Also, \( s(\ll A) \leq s(A) \), with equality if and only if \( A = \ll A \).

Proof. Proof by structural induction on \( A \). \hfill \Box

Next, we want to write a correspondence between types \( A \) and \( \ll A \), in the form of typing judgements \( x : A \triangleright \ll x \ll \ll A \) and \( x : \ll A \triangleright \ll x \ll \ll A \) whose denotations are multiples of the identity function. The correspondences are built by induction on the \( \ll \)-size of \( A \):

- knowing the result for \( A \) we deduce it for \( \text{bit} \triangleright A \), \( \text{qbit} \triangleright A \) and \( \top \triangleright A \);
- knowing the result for \( A \) and \( B \) we deduce it for \( A \otimes B \);
- knowing the result for \( C \triangleright (D \triangleright B) \) we deduce it for \( (C \otimes D) \triangleright B \);
- knowing the result for \( C \triangleright D \) and \( (\ll C \otimes \ll D) \triangleright B \) we deduce it for \( (C \triangleright D) \triangleright B \).

The detailed definition goes as follows:

Definition 7.3.19. We define the terms \( \ll x \ll A \) and \( \ll x \ll r \) by induction on the \( \ll \)-size of \( A \). First,

\[
\begin{align*}
\ll x \ll \text{bit} &= \ll x \ll \text{bit} = x, \\
\ll x \ll qbit &= \ll x \ll qbit = x, \\
\ll x \ll \top &= \ll x \ll \top = x.
\end{align*}
\]

Then,

\[
\begin{align*}
\ll x \ll A \otimes B &= \text{let } (y, z) = x \in \langle \ll y \ll A, \ll z \ll B \rangle, \\
\ll x \ll A \triangleright B &= \text{let } (y, z) = x \in \langle \ll y \ll r, \ll z \ll r \rangle.
\end{align*}
\]

For types of the form \( A \triangleright B \), we first consider the case where \( A = \top \), \( A = \text{bit} \) and \( A = \text{qbit} \):

\[
\begin{align*}
\ll x \ll \text{bit} \triangleright B &= \langle *, \text{let } t^B = x^* \text{ in } \ll t^B \rangle, \\
\ll x \ll \text{qbit} \triangleright B &= \text{let } (y^T, z^B) = x \text{ in } \lambda t^B. \text{let } * = y \in \ll t^B \rangle, \\
\ll x \ll \text{bit} \triangleright B &= \text{if } \text{meas}(H(\text{new } 0)) \text{ then } \langle 1, \text{let } t^B = x^1 \text{ in } \ll t^B \rangle, \\
&\quad \text{else } \langle 0, \text{let } t^B = x^0 \text{ in } \ll t^B \rangle, \\
\ll x \ll \text{qbit} \triangleright B &= \langle y^\text{bit}, z^B \rangle = x \text{ in } \ll y^\text{bit}, z^B \rangle.
\end{align*}
\]
7.3. Denotational Semantics

\[ \lambda_{\text{bit}} \text{ if } x \text{ then } \text{ if } \text{ then } \mathcal{Y}_{f_0}^B \text{ else } \Omega \]

\[ \text{else if } \text{ then } \Omega \text{ else } \mathcal{Y}_{f_0}^B \]

\[ \mathcal{Y}_{x^B}^{q_{\text{bit}} \rightarrow B} = \text{let } (u,v) = \mathcal{N}_C(H(\text{new } 0), \text{new } 0) \text{ in } (u, \text{let } t^B = xv \text{ in } \mathcal{Y}_{f_0}^B) \]

\[ \mathcal{Y}_{x^D}^{q_{\text{bit}} \rightarrow B} = \text{let } (y^D,z^B) = x \text{ in } \lambda^D \text{ if } \text{meas } y^D \text{ then } \text{else } \Omega \text{ else } \mathcal{Y}_{f_0}^B \]

Then, when \( A \) is of the form \( C \odot D \), we proceed the following way:

\[ \mathcal{Y}_{x^B}^{(C \odot D) \rightarrow B} = \text{let } f^{C \rightarrow (D \rightarrow B)} = \lambda^C, u^D(x(t,u)) \text{ in } \]

\[ \text{let } (u^C,v^D) = \mathcal{Y}_{f_0}^{C \rightarrow (D \rightarrow B)} \text{ in } \]

\[ \text{let } (y^D,z^B) = v \text{ in } (u,y,z) \]

\[ \mathcal{Y}_{x^D}^{(C \odot D) \rightarrow B} = \lambda^D \text{, let } (y^D,z^B) = t \text{ in } \text{let } (u^C,v^B) = x \text{ in } \]

\[ \text{let } (u^C,v^D) = x \text{ in } \]

\[ \text{let } a = (w,(s,v)) \text{ in } (\mathcal{Y}_{x^B}^{C \rightarrow (D \rightarrow B)} y) z \]

Finally, when \( A \) is of the form \( C \rightarrow D \), we reduce \( C \rightarrow D \) to \( \$C \odot \$D \) to fall back on the previous case. The details are as follows:

\[ \mathcal{Y}_{x^B}^{(C \rightarrow D) \rightarrow B} = \text{let } f^{(\$C \odot \$D) \rightarrow B} = \lambda^C \odot D.x(\mathcal{Y}_{f_0}^{C \rightarrow (D \rightarrow B)}) \text{ in } \]

\[ \mathcal{Y}_{x^D}^{(C \rightarrow D) \rightarrow B} = \lambda^D \text{, $C \rightarrow D$} \mathcal{Y}_{x^D}^{(C \odot D) \rightarrow B} \mathcal{Y}_{x^D}^{C \rightarrow D} \]

**Lemma 7.3.20.** For every types \( A \), the typing judgements \( x : A \triangleright \mathcal{Y}_{x^B}^A : \$A \) and \( x : \$A \triangleright \mathcal{Y}_{x^B}^A : A \) are valid. Moreover, there exist constants \( \lambda, \lambda' > 0 \), such that \( [\mathcal{Y}_{x^B}^A] = \lambda \text{id}_{[A]} \) and \( [\mathcal{Y}_{x^B}^A] = \lambda' \text{id}_{[A]} \).

**Proof.** The proof that the typing judgements are valid is an easy induction on the \( \$\)-size of \( A \). We provide the computation for the difficult cases, namely when the type is \( \text{bit} \rightarrow B \) and \( \text{qbit} \rightarrow B \).

**Case \( x^B_\text{\text{bit}} \rightarrow B \).** We are considering

\[ f = [x : \text{bit} \odot \$B \triangleright \mathcal{Y}_{x^B}^{\text{bit} \rightarrow B} : \text{bit} \rightarrow B] : [B] \oplus [B] \rightarrow [B] \oplus [B]. \]

An element \( (a,b) \in [\$B] \oplus [\$B] \) corresponds to a probability \( a \) to get the boolean 0 and a probability \( b \) to get the boolean 1 for the bit. This means that \( f(a,b) = ([\mathcal{Y}_{x^B}^\text{bit}](a), [\mathcal{Y}_{x^B}^\text{bit}](b)) \).

By induction hypothesis, \( [\mathcal{Y}_{x^B}^\text{bit}] = \lambda \text{id} \) for some \( \lambda > 0 \). Thus we have then \( f = \lambda \text{id} \).

**Case \( x^B_\text{\text{qbit}} \rightarrow B \).** We are considering

\[ f = [x : \text{bit} \rightarrow B \triangleright \mathcal{Y}_{x^B}^{\text{bit} \rightarrow B} : \text{qbit} \odot \$B] : [B] \oplus [B] \rightarrow [B] \oplus [B]. \]

A function \( g \in V_{\text{bit} \rightarrow B} \) is sent to \( \frac{1}{2}([\mathcal{Y}_{x^B}^\text{bit}](g(1,0)), [\mathcal{Y}_{x^B}^\text{bit}](g(0,1))) \), where we identify the function \( g \) and its representation. By induction hypothesis, \( [\mathcal{Y}_{x^B}^\text{bit}] = \lambda \text{id} \) for some \( \lambda > 0 \). Thus \( g \in V_{\text{bit} \rightarrow B} \), represented by \( (g(1,0), g(0,1)) \), is sent to \( \frac{1}{2} \lambda g \). This means that \( f = \frac{1}{2} \lambda \text{id} \).
Case $\uparrow_{q}^{\text{bit} \rightarrow B}$. We are considering

\[ f = [x : \text{bit} \otimes [B] \triangleright \uparrow_{q}^{\text{bit} \rightarrow B} : \text{bit} \rightarrow [B] : 2 \otimes [B] \rightarrow 2 \otimes [B]]. \]

We interpret the term as in the definition. Suppose that $x = (y, z)$ and $t$ are respectively the matrix of blocks and the quantum bits

\[ \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (aA, bA, cB, dD), \quad (cC, dC, cD, dD), \]

where the blocks $A$, $B$, $C$ and $D$ belongs to $V_{[B]}$. Then the products $(y, t, z)$ and $(y', t', z)$ are respectively

\[ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (aA + bB + cC + dD), \]

Applying $H$ on $y'$ yields a matrix with first entry

\[ \frac{1}{2} \lambda(aA + bB + cC + dD), \]

if $[\uparrow_{q}^{B}] = \lambda \text{id}$, corresponding to the probability for $y''$ to be $|0\rangle$. Thus,

\[ f \left( \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right) \left( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = \frac{1}{2} \lambda(aA + bB + cC + dD), \]

meaning that $x$ maps to a multiple of the function $\text{bit} \rightarrow [B]$ of same denotation. That is, $f = \frac{1}{2} \lambda \text{id}$.

Case $\uparrow_{\text{bit} \rightarrow B}$. We are considering

\[ f = [x : \text{bit} \rightarrow B \triangleright \uparrow_{\text{bit} \rightarrow B} : \text{bit} \otimes [B] : 2 \otimes [B] \rightarrow 2 \otimes [B]]. \]

In this case, the function $f(x)$ creates a pair of two quantum bits in state

\[ \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \]

and applies (a multiple of) the function $f$ on each $2 \times 2$ block: it is precisely (a multiple of) the characteristic function of $x$: the function $f$ is indeed a multiple of the identity.

**Theorem 7.3.21** (Fullness up to scalar multiple). For any types $A$ and $B$, there exists a constant $\lambda > 0$ such that if $F : [A] \rightarrow [B]$ is any superoperator, then there is a valid typing judgement $x : A \triangleright M : B$ of denotation $\lambda F$.

**Proof.** For the given types $A$ and $B$, consider the terms $\uparrow_{A}$ and $\uparrow_{B}$. From Lemma they have for denotation respectively $\lambda \text{id}$ and $\lambda' \text{id}$, for some $\lambda, \lambda'$. We claim that any superoperator $[A] \rightarrow [B]$ can be interpreted modulo a factor of $\lambda \lambda'$.

Indeed, consider a superoperator $F : [A] \rightarrow [B]$. From Theorem there exists a valid typing judgement $y : \uparrow_{A} \triangleright N : \uparrow_{B}$ of denotation $F$. Let $a : A \triangleright M : B$ be the term

\[ a : A \triangleright \text{let } y = \uparrow_{a} \text{ in let } b = N \text{ in } \uparrow_{b}, \]

It has denotation $\lambda \lambda' F$. 

\[ \square \]
7.4. Equivalence Classes of Terms

We call it the axiomatic equivalence relation way: if \( \Delta \vDash \) when \( \Gamma \vdash \).

Definition 7.4.1. A first notion of equality of terms can be defined by a set of syntactic rules. This is known as the axiomatic equivalence.

A first notion of equality of terms can be defined by a set of syntactic rules. This is known as the axiomatic equivalence.

\begin{align*}
(\beta) & \quad \Gamma \vdash (\lambda x. M)N \quad \approx_{ax} \quad M[N/x] : A \\
(\eta) & \quad \Gamma \vdash \lambda x. Mx \quad \approx_{ax} \quad M : A \to A \\
(\beta_{\odot}) & \quad \Gamma \vdash \text{let } (x, y) = (N, P) \text{ in } M \quad \approx_{ax} \quad M[N/x, P/y] : A \\
(\eta_{\odot}) & \quad \Gamma \vdash \text{let } (x, y) = M \text{ in } (x, y) \quad \approx_{ax} \quad M : A \otimes B \\
(\beta) & \quad \Gamma \vdash x = x \text{ in } M \quad \approx_{ax} \quad M : A \\
(\eta_{x}) & \quad \Gamma \vdash \text{let } * = M \text{ in } * \quad \approx_{ax} \quad M : \top \\
(\beta_{\odot}) & \quad \Gamma \vdash \text{if } 1 \text{ then } M \text{ else } N \quad \approx_{ax} \quad M : A \\
(\beta^{0}_{\odot}) & \quad \Gamma \vdash \text{if } 0 \text{ then } M \text{ else } N \quad \approx_{ax} \quad N : A \\
(\Omega) & \quad \Gamma \vdash M[\Omega/x] \quad \approx_{ax} \quad \Omega : A \\
(id) & \quad \Gamma \vdash \text{meas}(\text{new } M) \quad \approx_{ax} \quad M : \text{bit}
\end{align*}

Table 7.4: Axiomatic equivalence

7.4 Equivalence Classes of Terms

Being able to build terms, we need some tools to compare them. One can compare them through syntactic manipulations, or one can have a finer approach using the two semantics we have built: in the case of the operational semantics, the behavior of the terms is what defines the equivalence, and in the case of the denotational semantics, the equivalence is expressed by the denotation of the terms.

7.4.1 Axiomatic Equivalence

A first notion of equality of terms can be defined by a set of syntactic rules. This is known as the axiomatic equivalence.

Definition 7.4.1. We define an equivalence relation \( \approx_{ax} \) on typing judgements. We write the relation as \( \Gamma \vdash M \approx_{ax} N : A \), and we define it to be the smallest relation satisfying the rules in Table 7.4 the alpha-equivalence and one congruence rule (\( \xi \)) per term constructor. This means for example:

\[
\begin{align*}
\Gamma \vdash M & \approx_{ax} M' : A \to B \\
\Delta \vdash N & \approx_{ax} N' : A
\end{align*}
\]

\( \Gamma, \Delta \vdash MN \approx_{ax} M'N' : B \quad (\xi_{\text{app}}) \)

\[
\begin{align*}
\Gamma & \vdash \lambda x. M \approx_{ax} \lambda x. M' : A \to B \\
\Delta, x : A & \vdash M \approx_{ax} M' : B
\end{align*}
\]

\( \Gamma \vdash \lambda x. M \approx_{ax} \lambda x. M' : A \to B \quad (\xi_{\lambda}) \)

We call it the axiomatic equivalence relation. We extend it to quantum closures in the following way: if \( \Delta \models [Q, L, M] : A \) and \( \Delta \models [Q, L, M'] : A \) are two well-typed quantum closures where \( \text{FV}(M) \setminus |\Delta| = \text{FV}(M') \setminus |\Delta| = \{x_1, \ldots, x_n\} \), then we write \( \Delta \models [Q, L, M] \approx_{ax} [Q, L, M'] : A \) when

\[
\Delta, x_1 : \text{qbit}, \ldots, x_n : \text{qbit} \vdash M \approx_{ax} M' : A.
\]

Lemma 7.4.2. The order of the arguments in an application does not matter. Similarly, one can apply the arguments as a pairing or sequentially. More precisely:

\[
\begin{align*}
[Q, L, (\lambda xy. M)N] & \approx_{ax} [Q, L, ((\lambda y. M)P)N] \\
[Q, L, \text{let } (x, y) = (N, P) \text{ in } M] & \approx_{ax} [Q, L, ((\lambda xy. M)N)P]
\end{align*}
\]

Proof. By application of rule (\( \beta \)) and rule (\( \beta_{\odot} \)).
7.4.2 Operational Context

To say that two arbitrary terms have the same behavior, we need a way to observe them. The only observable types at our disposal are the types bit and ⊤. So the fact that two terms $M$ and $M'$ have the same behavior can be understood as the fact that in whichever context $C[-]$ we “use” them, if $C[-] : \text{bit}$, then $C[M]$ reduces to 0, respectively, 1, with the same probability as $C[M']$. Such a term $C[-]$ is called an operational context.

**Definition 7.4.3.** We define a formal operational context to be a formula defined by the following BNF:

$$C[-] ::= [-] | (C[-]M) | (MC[-]) | \lambda x.C[-] | (C[-], M) | (M, C[-]) | \text{let } (x, y) = C[-] \text{ in } M | \text{let } (x, y) = M \text{ in } C[-] | \text{let } * = C[-] \text{ in } M | \text{let } * = M \text{ in } C[-] | \text{if } C[-] \text{ then } M \text{ else } N | \text{if } M \text{ then } C[-] \text{ else } C'[−].$$

We call $[-]$ the hole of the context.

The notions of well-typed contexts and free variables in contexts are defined the same ways as for terms. Note that there exists a new notion: the notion of captured variables, which are the variables whose scope includes the hole. We can make this more precise by speaking of typed contexts:

**Definition 7.4.4.** A typed operational context is a typing tree with root $\Gamma' \triangleright C[-] : B$, considering the additional axiom $\Gamma \triangleright [-] : A$, i.e., a typing tree of the form

$$\begin{align*}
\Gamma & \triangleright [-] : A \\
\Gamma' & \triangleright C[-] : B.
\end{align*}$$

We say that this context is of type $B$, with free variables $\Gamma'$, a hole of type $A$, and captured variables $\Gamma$. We also use the notation $\Gamma' \triangleright C[\Gamma \triangleright - : A] : B$ for a typed operational context.

**Lemma 7.4.5.** If

$$\begin{align*}
\Gamma & \triangleright [-] : A \\
\Gamma' & \triangleright C[-] : B.
\end{align*}$$

is a valid typing derivation, then so is

$$\begin{align*}
\Delta, \Gamma & \triangleright [-] : A \\
\Delta, \Gamma' & \triangleright C[-] : B,
\end{align*}$$

provided the variables that occur in $\Delta$ are fresh.

**Proof.** By easy induction on the typing derivation of $\Gamma' \triangleright C[-] : B$. 

7.4.3 Operational Equivalence

We define a notion of operational equivalence, based on the reduction rules and observations of type bit, as in [Danos and Harmer (2002)](Danos_and_Harmer_2002). (Equivalently, it would suffice to consider observations of type ⊤).
7.5. Soundness and Full Abstraction

Definition 7.4.6. Let $\triangleright C[\Gamma \triangleright - : A] : \text{bit}$ be a closed typed operational context of type $\text{bit}$, and let $R = [Q, L, M]$ be a well-typed quantum closure with typing judgement $\Gamma \models [Q, L, C[M]] : A$. In this case we define the substitution $C[R]$ by $[Q, L, C[M]]$, where $M$ is syntactically replacing $[-]$ in $C[-]$. We linearly extend this definition to probabilistic distributions of quantum closures of the form $\Gamma \models \sum_i \rho_i R_i : A$ by setting $C[\sum_i \rho_i R_i] = \sum_i \rho_i C[R_i]$.

Lemma 7.4.7. In Definition 7.4.6, the substitution is well typed and a valid typing judgement for it is $\Gamma \models [Q, L, C[M]] : \text{bit}$.

Proof. This lemma is a direct consequence of Lemma 7.4.5. □

Definition 7.4.8. Given two well-typed quantum closures $\Gamma \models R, R' : A$, we say that $R$ is operationally equivalent to $R'$ with respect to $\Gamma$ if for all closed typed operational contexts $\triangleright C[\Gamma \triangleright - : A] : \text{bit}$, $C[R] \downarrow_0 = C[R'] \downarrow_0$ and $C[R] \downarrow_1 = C[R'] \downarrow_1$. In this case, we write $\Gamma \models R \approx_{\text{op}} R' : A$. If $M, M'$ are terms, we say that $\Gamma \triangleright M \approx_{\text{op}} M' : A$ if $\Gamma \models [Q, L, C[M]] : \text{bit}$.

7.4.4 Denotational Equivalence

The last equivalence we can define is the denotational equivalence. This equivalence relation is simply stated:

Definition 7.4.9. We say that two typing judgements $\Gamma \triangleright M, M' : A$ are denotationally equivalent if $[\Gamma \triangleright M : A]$ and $[\Gamma \triangleright M' : A]$ are the same map in $\text{CPM}$. In that case we write $\Gamma \models R \approx_{\text{den}} R' : A$. We extend this definition to quantum closures: $\Gamma \models R \approx_{\text{den}} R' : A$ is true if $[\Gamma \models R : A] = [\Gamma \models R' : A]$.

7.5 Soundness and Full Abstraction

The three defined equivalence relations we have built have the expected behavior: The axiomatic equivalence is sound with respect to the operational equivalence and the denotational semantic is fully abstract with respect to the operational semantics:

Theorem 7.5.1 (Soundness). If $\Gamma \triangleright M \approx_{\text{ax}} M' : A$ then $\Gamma \triangleright M \approx_{\text{op}} M' : A$.

Remark 7.5.2. An immediate consequence of soundness and Lemma 7.4.5 is that the order of evaluation does not affect the outcome:

$$\Gamma \triangleright ((\lambda xy. R)M)N \approx_{\text{op}} ((\lambda yx. R)N)M.$$ 

Theorem 7.5.3 (Full abstraction). The denotational semantics is fully abstract with respect to the operational equivalence of typing judgements, i.e.

$$[\Gamma \triangleright M : A] = [\Gamma \triangleright M' : A] \text{ if and only if } \Gamma \triangleright M \approx_{\text{op}} M' : A.$$ 

Remark 7.5.4. The presence of the non-terminating term $\Omega$ is necessary for full abstraction to hold. Without it, every program terminates with probability 1, and there is only one definable map $\text{bit} \to \top$. Thus, although $\lambda f.(f0)$ and $\lambda f.(f1)$ of type $(\text{bit} \to \top) \to \top$ have different denotations, no context will distinguish them.
7.5. Soundness and Full Abstraction

7.5.1 Proof of the Soundness Theorem

Assume Theorem 7.5.8. It suffices to show that if $\Gamma \vdash M \approx_{ax} M' : A$ then $\Gamma \vdash M \approx_{den} M' : A$. We show this by structural induction on the proof of $\Gamma \vdash M \approx_{ax} M' : A$.

The cases $(\beta)$, $(\beta_0)$, $(\beta_1)$ and $(\beta_2)$ are similar as in the proof of Lemma 7.5.6.

We detail the remaining cases

$(\eta)$. If $[\Gamma \vdash M : A \rightarrow B] = \Phi(g) : [[\Delta]] \rightarrow [[A]] \otimes [[B]]$, then the denotation $[[\Gamma, x : A \rightarrow Mx : B]]$ is a map $f : [[\Gamma]] \otimes [[A]] \rightarrow [[B]]$, with

$$f(x \otimes y) = g(x \otimes \text{id}_A(y)) = g(x \otimes y).$$

$(\eta_\circ)$. If $f = [\Gamma \vdash M : A \otimes B]$, if $g = [\Gamma \vdash \text{let } (x, y) = M \text{ in } (x, y) : A \otimes B]$ and if $a \in [[\Gamma]]$, we have $g(a) = (\text{id}_A \otimes \text{id}_B)(f(a)) = f(a)$.

$(\eta_\star)$. If $f = [\Gamma \vdash M : \top]$, if $g = [\Gamma \vdash \text{let } * = M \text{ in } * : \top]$, and if $a \in [[\Gamma]]$, then $g(a)$ is equal to $\text{id}_\top(1) \cdot f(a)$, that is, $f(a)$.

$(\eta_f)$. In this case, the context $\Gamma$ splits into $(\Gamma_1, \Gamma_2)$, where $\Gamma_1 \triangleright M : \text{bit}$ and $\Gamma_2, x : \text{bit} \triangleright M : A$. If $(p, q)$ is the denotation of the former and $f$ the denotation of the latter, if $a \in [[\Gamma_1]]$ and $b \in [[\Gamma_2]]$, $p(a)f(0, b) + q(a)f(b, 0) = f(0, b \cdot p(a)) + f(b \cdot q(a), 0) = f((p(a), q(a)) \otimes b)$.

$(id)$. This last case is done by noticing that $[[x : \text{bit} \triangleright \text{meas}(\text{new } x) : \text{bit}]]$ is the identity.

7.5.2 Full Abstraction: Preliminary Lemmas

Lemma 7.5.5. For any two programs $P, P'$ of type bit, they have the same denotation if and only if for all $b \in \{0, 1\}$, $P \Downarrow b = P' \Downarrow b$.

Proof. Consider two well-typed programs $P, P'$ : bit. Suppose they have the same denotation $f$. Then from Lemma 7.5.8 $f$ is also the denotation of $\text{prob}_f^P$ and of $\text{prob}_f^{P'}$. But by definition, $[\text{prob}_f^P] = (p, q)$, where $p = (\text{prob}_f^P) \Downarrow 1 = P \Downarrow 0$ and $q = (\text{prob}_f^P) \Downarrow 1 = P \Downarrow 1$, and similarly for $P'$. Thus $P \Downarrow 1 = P' \Downarrow 1$ and $P \Downarrow 0 = P' \Downarrow 0$. The argument being reversible, we get the other implication.

Lemma 7.5.6. If $[\Gamma \vdash M : A] = [\Gamma \vdash M' : A]$ and $C[\Gamma \vdash \top : A] : \text{bit}$ is a valid context, then $[[C[M] : \text{bit}]] = [[C[M'] : \text{bit}]]$.

Proof. The proof uses Lemma 7.5.8.

Lemma 7.5.7. Let $A$ be a type and let $v \neq v'$ be hermitian positive elements in $V[A]$. Then there exists a valid typing judgement $x : A \triangleright M : \text{bit}$ such that $[[M](v) \neq [M](v')]$.

Proof. By Lemma 7.5.20 it suffices, without loss of generality, to consider the case where $A = \text{bit}^{\otimes n} \otimes \text{qbit}^{\otimes m}$. Then $[A] = (2^m, \ldots, 2^m), v = (v_1, \ldots, v_{2^m})$ and $v' = (v'_1, \ldots, v'_{2^m})$. If $v$ and $v'$ differ, there exists some $i$ for which $v_i \neq v'_i$. By Theorem 5.1.13 there exists orthonormal bases of eigenvectors corresponding to $v_i$ and $v'_i$. Since they differ, there exists a vector $w$ such that $w^*v_iw \neq w^*v'_iw$ (otherwise, $v_i$ and $v'_i$ would have the same eigenvectors and eigenvalues, making them equal).

Consider the map

$$f : (2^m, \ldots, 2^m) \rightarrow 1,1 \quad \begin{cases} (w_1, \ldots, w_n) \rightarrow (w^*w_i, 0) \end{cases}$$

From Theorem 7.5.1 this is a superoperator. By definition, $f(v) \neq f(v')$ and by Theorem 7.5.13 it is representable by some valid typing judgement $x : A \triangleright M : \text{bit}$.
Lemma 7.5.8. Given any type \( A \) and any hermitian positive \( v \in V_A \), there exists a closed term \( M : A \) and \( \lambda > 0 \) such that \([M](1) = \lambda v\).

Proof. If \( v = 0 \), let \( M = \Omega \). Else, from Corollary 7.3.14, there exists a valid typing judgement \( \triangleright N : \& A \) such that \( v / \text{Tr}(v) = [\triangleright N : A](1) \). Then let \( M = \text{let } x = N \text{ in } \Upsilon_A x \uparrow : A \). From Lemma 7.3.20, \([\triangleright M : A](1) = v\) and \( v \) are collinear. \( \square \)

7.5.3 Proof of the Full Abstraction Theorem

If \([\Gamma \triangleright M : A] = [\Gamma \triangleright M' : A]\), take any valid context \( C[\Gamma \triangleright - : A] : \text{bit} \) for those two terms. Then from Lemma 7.5.6, \([\triangleright C[M] : \text{bit}](Fv) \neq [\triangleright C[M'] : \text{bit}](Gv)\). Now consider \( C[\Gamma \triangleright - : A] : \text{bit} \) defined by

\[
\text{let } (x_1, \ldots, x_n) = R \text{ in let } x = [-] \text{ in } S.
\]

Then \([C[M]](1) \neq [C[M']])(1)\), hence by Lemma 7.5.8, \( C[M] \downarrow b \neq C[M'] \downarrow b\), for some \( b \in \{0, 1\}\). It follows that \( M \not\approx_{op} M'\), which completes the proof of full abstraction. \( \square \)
Chapter 8

Structure of the Linear-Non-Linear Fragment

In the previous chapter, we described a semantics for the linear fragment of the quantum lambda calculus of Chapter 6. Doing so, we omitted the issue of duplicability. This chapter and the following ones are concerned with the study of the linear-non-linear fragment of the lambda calculus described in Chapter 6. An extended abstract of the work presented here has been published in [Selinger and Valiron, 2008].

8.1 Computations and Values

In Chapter 7, we developed an interpretation of computations as completely positive linear maps between specific vector spaces. Due to the strict linearity, there was no need to distinguish between values and computations within the denotation. In the case of the general quantum lambda calculus, the computations are not the whole story. Due to the probabilistic behavior of the measurement, in order to give a categorical description of the quantum lambda calculus, the analysis of the values has to be separated from the analysis of the computations. For this purpose we will use the semantics described in Section 5.4.3 and describe a computational model to handle the probabilistic side effect.

8.2 Duplicability Versus Non-Duplicability

While the strong monad handles the probabilistic nature of the computation, it does not deal with the question of the non-duplicability. [Coecke and Pavlovic, 2007] offer a solution for duplication at first order using the notion of classical objects. The idea is that the boolean structure (1, 1) admits a “copy” operator such as

\[
\delta_{\text{bit}} : \begin{align*}
1,1 & \quad \rightarrow \quad 1,1,1 \\
(a,b) & \quad \rightarrow \quad (a,0,0,b),
\end{align*}
\]

sending the Boolean 0 to the pair (0, 0) and the Boolean 1 to the pair (1, 1). This approach does not generalize well at higher-order. Indeed, consider the set of functions (\(\top \rightarrow \text{bit}\)) and the set of pairs of functions (\(\top \rightarrow \text{bit} \otimes \top \rightarrow \text{bit}\)). Using the denotation in Chapter 7, a function \(f : \top \rightarrow \text{bit}\) is represented by a pair \((a,b)\) where \(a\) is the probability of outputting the Boolean 0 and \(b\) the probability of getting the Boolean 1. Thus, in the denotation, it is the same entity as a boolean. If we define \(\delta_{\top \rightarrow \text{bit}}\) as the same “copy” operator, \(\delta_{\top \rightarrow \text{bit}}(f)\) becomes the pair of functions that outputs
\[ \langle 0,0 \rangle \text{ with probability } a \text{ and } \langle 1,1 \rangle \text{ with probability } b \]. This is not how we expect the duplication of \( f \) to behave.

### 8.2.1 Computations as Proofs

Instead of working at the level of the denotational semantics, we will study the language at a syntactic level and work out the required structure from there, using the powerful Curry-Howard isomorphism of Section 5.2.2.

Our type system has a particular type construct, namely "\(!\)", attached to the subtyping relation defined in Definition 6.3.3. As was noted in (Valiron, 2004a), it is possible to rewrite this subtyping relation in the following way:

\[
\alpha \lessdot \alpha \quad (ax) \quad A \lessdot B \quad (D) \quad !A \lessdot !B \quad (!)
\]

This formulation of the subtyping relation guides us to the theory of linear logic.

### 8.3 Structure of the Exponential "\(!\)"

The model of linear logic offered by Bierman (1993) and described in Section 5.6.2 is almost a perfect fit. However, it is slightly too general for our purpose. In this section we describe the modifications that are needed.

#### 8.3.1 Idempotency

Indeed, the "\(!\)" operator of the type system of the quantum lambda calculus is idempotent, in the sense that \( x : !A \vdash x : !!A \) and \( x : !!A \vdash x : !A \) are inverse one of each other. This is not reflected in the definition of a generic comonad, where the map

\[
LLA \xrightarrow{\epsilon_A} LA \xrightarrow{\delta_A} LLA
\]

is not equal to the identity.

**Definition 8.3.1** (Taylor (1999)). A comonad \((L, \epsilon, \delta)\) on some category is said to be **idempotent** if \( \delta : L \rightarrow LL \) is an isomorphism.

**Example 8.3.2.** Consider any category \( C \). The functor \( id_C \) is an idempotent comonad, with the comultiplication and the counit being the identity natural transformation.

Based on the idea of classical objects of Coecke and Pavlovic (2007), a non-trivial example can be constructed. The setting is the category of finite Hilbert spaces. There is no canonical “copy” operation \( E \rightarrow E \otimes E \) on a given space \( E \) unless a basis has been arbitrarily chosen. We define the category \( ClasObj \) based on the category \( fdHilb \) of finite Hilbert spaces.

**Example 8.3.3.** Let \( ClasObj \) be the category with

- Objects: pairs \((E, B_E)\) where \( E \) is a finite dimensional Hilbert space and \( B_E \) is a set of orthonormal vectors of \( E \). We write \( \langle B_E \rangle \) for the subspace of \( E \) generated by \( B_E \), and \( E^\perp_{B_E} \) for the subspace of \( E \) orthogonal to \( \langle B_E \rangle \).
8.3. Structure of the Exponential

Let us start with the notions of idempotent comonads. Let $(E, B_E)$ be an object of the category $fdHilb$, and let $(F, B_F)$ be another such object. We want to define a map $\Delta_{(F, B_F)} : F \rightarrow F \otimes F$, which we will call the diagonal map. This map is defined as follows:

$$\Delta_{(F, B_F)}(f) = f \otimes f$$

for any morphism $f : E \rightarrow F$. The diagonal map should satisfy the following properties:

1. **Idempotence:**
   $$\Delta^2_{(F, B_F)} = \Delta_{(F, B_F)}$$

2. **Monoidality:**
   $$\Delta_{(F, B_F)}(g \circ f) = \Delta_{(F, B_F)}(g) \otimes \Delta_{(F, B_F)}(f)$$

We want to be able to define a map $(\epsilon, \delta) : (E, B_E) \rightarrow (F, B_F)$ such that $(\epsilon, \delta)$ is an idempotent comonad. Then we can define a functor $L : fdHilb \rightarrow fdHilb$ as follows:

$$L((E, B_E)) = (\epsilon \circ f, \delta \circ f)$$

for any morphism $f : E \rightarrow F$. The functor $L$ should satisfy the following properties:

1. **Functoriality:**
   $$L(f \circ g) = L(f) \circ L(g)$$

2. **Naturality:**
   $$L(\epsilon) = \epsilon$$

3. **Monoidal property:**
   $$L(\otimes) = L \otimes L$$

4. **Idempotence:**
   $$L(\Delta^2_{(F, B_F)}) = \Delta^2_{L(F, B_F)}$$

5. **Comonad property:**
   $$L(\epsilon \circ f, \delta \circ f = \epsilon \circ L(f), \delta \circ L(f))$$

We define a functor $L$ from $fdHilb$ to $fdHilb$ as follows: For any object $(E, B_E)$, $L(E, B_E)$ is the object $((B_E), B_E)$ and for any morphism $f : (E, B_E) \rightarrow (F, B_F)$, $Lf$ is the restriction of the linear map $f : E \rightarrow F$ to the subspace generated by $B_E$.

If $\epsilon_{(E, B_E)} : L(E, B_E) \rightarrow (E, B_E)$ is the canonical linear injection and if $\delta_{(E, B_E)} : L(E, B_E) \rightarrow L(L(E, B_E))$ is the linear identity, $(L, \epsilon, \delta)$ is an idempotent comonad.

**Lemma 8.3.4.** If $(L, \epsilon, \delta)$ is an idempotent comonad, then $\delta L = L \delta$, $L \epsilon = \epsilon L$ and $\delta^{-1} = L \epsilon$.

**Proof.** Equation (2.6.2) states that $\delta; L \delta = \delta; L \delta$. Multiplying on the left by $\delta^{-1}$, we get $L \delta = \delta L$. Equation (2.6.3) says $\delta; \epsilon L = \delta; L \epsilon$. Multiplying on the left by $\delta^{-1}$, we get $L \epsilon = \epsilon L$. Finally, since $\delta; L \epsilon = id$, again multiplying on the left by $\delta^{-1}$ we get the last required equation: $L \epsilon = \delta^{-1}$.

**Lemma 8.3.5.** Suppose that $(L, \epsilon, \delta, d^L, d^L)$ is an idempotent monoidal comonad. Then $\epsilon_{\top}; d^L_\top = id_{L \top}$.

**Proof.** Consider the following diagram:

The equality is obtained from Lemma 8.3.4, the lower triangle by naturality of $\epsilon$ and the upper triangle by functoriality of $L$ on the monoidality of $\epsilon$ (Equation (2.8.8)).

**Example 8.3.6.** The category $ClasObj$ from Example 8.3.3 inherits a symmetric monoidal structure from $fdHilb$ in the following manner:

$$(E, B_E) \otimes (F, B_F) = (E \otimes F, \{ \epsilon \otimes f : \epsilon \in B_E, f \in B_F \}), \quad \top = (\mathbb{C}, \{1\})$$

and the tensor product of two morphisms $f$ and $g$ is the morphism of Hilbert spaces $f \otimes g$. If we define the maps

$$d^L_\top : \top \rightarrow L \top, \quad d^L_{(E, B_E), (F, B_F)} : L(E, B_E) \otimes L(F, B_F) \rightarrow L((E, B_E) \otimes (F, B_F)),$$

as the identities of the respective Hilbert spaces $\mathbb{C}$ and $E \otimes F$, then the comonad $(L, \epsilon, \delta, d^L, d^L)$ is monoidal.

### 8.3.2 Coherence Property for Idempotent Comonads

We want to be able to define a map $[A] \rightarrow [B]$, image of the relation $A <: B$. Since by Lemma 8.3.4 $(::)$ is an ordering relation, we would like to have a similar notion on the images. The category $\mathcal{C}$ has an idempotent comonad. In the following we will show a coherence property for the comonad, and use it to define the image of $A <: B$. 

8.3. Structure of the Exponential

Definition 8.3.7. Consider a category $\mathcal{C}$ with an idempotent comonad $(L, \delta, \epsilon)$. Let $(F_1, \ldots, F_n)$ be a list of functors $F_i : \mathcal{C}^{m_i} \times (\mathcal{C}^{\text{op}})^l \to \mathcal{C}$. Suppose $A$ is the set of elements of the form

$$A, B ::= a \mid F_i(A_1, \ldots, A_{m_i}, B_1, \ldots, B_{l_i}) \mid LA,$$

where $a$ spans over a given alphabet $A$. Given a map $G : A \to |\mathcal{C}|$, we build the $A$-$G$-category $\mathcal{C}^G$ as follows:

- The set of objects is $A$.
- Arrows are arrows of $\mathcal{C}$, defined by induction as follows:
  
  - for all objects $A$ in $\mathcal{C}^G$, the arrows $id_{G(A)}$, $\epsilon_{G(A)}$ and $\delta_{G(A)}$ are arrows of $\mathcal{C}^G$.
  
  - If $f : A \to B$ is an arrow in $\mathcal{C}^G$ then so are $L f : LA \xrightarrow{\mathcal{L}(f)} LB$,

$$F_i(X_1, \ldots, f, \ldots, X_l, Y_1, \ldots, Y_{m_i}) = F_i(G(X_1), \ldots, f, \ldots, G(X_l), G(Y_1), \ldots, G(Y_{m_i})),$$

$$F_l(X_1, \ldots, X_l, Y_1, \ldots, f, \ldots, Y_{m_l}) = F_l(G(X_1), \ldots, G(X_l), G(Y_1), \ldots, f, \ldots, G(Y_{m_l})).$$

Finally, if $f : A \to B$ and $g : B \to C$ are arrows of $\mathcal{C}^G$, $(f; g) : A \to C$ is also an arrow of $\mathcal{C}^G$.

Theorem 8.3.8. Consider a category $\mathcal{C}$, an alphabet $A$ and a list of functors $(F_1, \ldots, F_n)$ as in definition 8.3.7. If $f : A \to B$ and $g : A \to B$ are two maps of $\mathcal{C}^G$, they are equal.

Following Kelly’s methodology [Kelly and Laplaza, 1980], we define the free category verifying the conditions, and show that this free category is in fact a poset. Without loss of generality, we consider only one $F_i$, that we name $\kappa : \mathcal{C}^{\text{op}} \times \mathcal{C} \to \mathcal{C}$.

Definition 8.3.9. Given an alphabet $A$, we define the graph $G$ as follows: the vertices of $G$ are

$$\text{Vertices } A, B ::= a \mid (A \times B) \mid LA,$$

where $a$ ranges over $A$, and the arrows are two-fold. First, for each vertex $A$ there exists arrows $\delta_A : LA \to LLA$ and $\epsilon_A : LA \to A$. Then for each arrow $f : A \to B$ all expansions of $f$ with respect to $(L, \kappa)$ are arrows in the graph.

Then we define $K$ to be the free category generated by $G$. We write $id_A$ for the identity on $A$. We call the arrows $id_A$, $\epsilon_A$ and $\delta_A$ elementary arrows.

We define a set of relations $\mathcal{R}$. This set contains first the $\mathcal{L}(A)$ and $\mathcal{L}(B)$, the equations of functoriality of $L$ and of naturality of $\epsilon$ and $\delta$, so that $(L, \delta, \epsilon)$ is a comonad. It contains the equations for the bifunctoriality of $\kappa$. We also request equations $L \epsilon_A ; \delta_A = \mathcal{L}^2 \epsilon_A$ and $\epsilon_{LA} = L \epsilon_A$. Furthermore, we request all expansions of the above relations with respect to $(L, \kappa)$ to belong to $\mathcal{R}$.

Finally, we define the category $\mathcal{D}$ to be $\mathcal{K}/\mathcal{R}$.

Lemma 8.3.10. In $\mathcal{D}$, $\kappa$ is a bifunctor and $(L, \delta, \epsilon)$ is an idempotent comonad.

Proof. All the requested equations are there. □

Definition 8.3.11. We define a map $K : |\mathcal{D}| \to |\mathcal{D}|$ by induction

$$K(a) = a \quad K(L^2 A) = K(LA) \quad K(La) = La \quad K(A \times B) = (KA) \times (KB)$$
8.3. Structure of the Exponential

\[ K(L(A \times B)) = L((KA) \times (KB)) \]

and a set of maps \( \zeta_A : A \to KA \), one for each object \( A \):

\[
\begin{align*}
\zeta_a &= \text{id}_a : a \to a, & \zeta_{La} &= \text{id}_{La} : La \to La, \\
\zeta_{L^2A} &= L^2A \xrightarrow{\delta^{-1}} LA \xrightarrow{\zeta_{LA}} K(LA), & \zeta_{A \times B} &= (KA) \ltimes B \xrightarrow{\zeta_A \ltimes \zeta_B} A \times (KB).
\end{align*}
\]

(8.3.1)

We define subsets \( L\text{SubVert} \) and \( \text{SubVert} \) of vertices:

\[
\begin{align*}
L\text{SubVert} &::= a \mid A \ltimes B, \\
\text{SubVert} &::= U \mid LU.
\end{align*}
\]

An element in \( \text{SubVert} \) is called a subvertex, and an element of \( L\text{SubVert} \) a linear subvertex. We define a partial ordering relation \( \preceq \) on subvertices (where \( A, B \in \text{SubVert} \) and \( U, V \in L\text{SubVert} \)):

\[
\begin{align*}
a &\preceq a & U &\preceq V & U &\preceq V & A &\preceq A' & B &\preceq B' \\
LU &\preceq V & LU &\preceq LV & A' \ltimes B &\preceq A \ltimes B'.
\end{align*}
\]

and a map \( m_{A,B} \) for each \( A \preceq B \):

\[
\begin{align*}
U &\xrightarrow{m_{U,V}} V \\
LU &\xrightarrow{L\mu_{U,V}} LV
\end{align*}
\]

\[
\begin{align*}
A &\xrightarrow{m_{A,A'}} A' \\
B &\xrightarrow{m_{B,B'}} B'
\end{align*}
\]

\[
\begin{align*}
m_a &= a \xrightarrow{\text{id}_a} a \\
m_{A' \ltimes B, A \ltimes B'} &= A' \ltimes B \xrightarrow{m_{A',A} \times m_{B,B'}} A \times B'
\end{align*}
\]

Lemma 8.3.12. For all \( A \), \( \zeta_A \) is an isomorphism.

\[ \text{Proof.} \text{ Proof by induction on } A, \text{ knowing that } \delta_A \text{ is an isomorphism.} \]

Lemma 8.3.13. For any vertex \( A \), \( KA \) is a subvertex.

\[ \text{Proof.} \text{ Proof by induction on } A. \]

Lemma 8.3.14. For any vertex \( A \), there exists a unique linear subvertex \( U \) and a unique \( n \geq 0 \) such that \( A = L^nU \). In that case, \( KA = U \) if \( n = 0 \), \( KA = LU \) otherwise.

\[ \text{Proof.} \text{ Proof by induction on } A. \]

Lemma 8.3.15. If \( U \) is a subvertex, then \( K(LU) = L(KU) \).

\[ \text{Proof.} \text{ Proof by case distinction.} \]

Lemma 8.3.16. If \( U \) is a subvertex, then

\[
LU \xrightarrow{\zeta_{LU}} K(LU) = LU \xrightarrow{L\zeta_U} L(KU)
\]

\[ \text{Proof.} \text{ Proof using Lemma 8.3.12 and by case distinction on the possible } \zeta_{LU}. \]

Lemma 8.3.17. \( \preceq \) is an partial ordering relation on subvertices.

\[ \text{Proof.} \text{ For } A, B \text{ and } C \text{ subvertices.} \]
8.3. Structure of the Exponential

- Reflexivity: $A \leq A$: proved by structural induction on $A$.
- Antisymmetry: $A \leq B$ and $B \leq A$ implies $A = B$: proved by structural induction on $A$.
- Transitivity: $A \leq B$ and $B \leq C$ implies $A \leq C$: proved by structural induction on $A$.

Lemma 8.3.18. If $A \leq B$ and $B \leq C$ are in $\text{SubVert}$,

$$A \xrightarrow{m_{A,B}} B \xrightarrow{m_{B,C}} C = A \xrightarrow{m_{A,C}} C. \quad (8.3.3)$$

Proof. We prove by induction on the derivation of $A \leq C$ that for all $B$ such that $A \leq B$ and $B \leq C$, Equation (8.3.3) is valid.

- If it is an axiom, then $A = C$ are $a$ or $\top$ and the only for $B$ is to be the same as $A$ and $B$: we are done.
- Suppose it is true for sub-derivations of $LU \leq LW$, $U$ and $W$ being linear subvertices. By case distinction the only possible $B$ such that $LU \leq B$ and $B \leq LW$ is $B = LV$, where $V$ is a linear subvertex. Then by functoriality of $L$ and and by induction hypothesis, the result is true.
- Suppose it is true for sub-derivations of $LU \leq W$, $U$ and $W$ being linear subvertices. By case distinction the only possible $B$ such that $LU \leq B$ and $B \leq W$ are $B = V$ or $LV$, $V$ a linear subvertex.

- If $B = V$: then $m_{LU,V} = LU \xrightarrow{\epsilon_U} U \xrightarrow{m_{U,V}} V$, and $m_{LU,V}; m_{V,W}$ is

\[
\begin{tikzpicture}
  \node (LU) at (0,0) {$LU$};
  \node (U) at (2,0) {$U$};
  \node (V) at (4,0) {$V$};
  \node (W) at (4,-2) {$W$};
  \node (mLUW) at (0,-2) {$m_{LU,W}$};
  \node (mUV) at (2,-2) {$m_{U,V}$};
  \node (muV) at (4,-2) {$m_{V,W}$};
  \draw[->] (LU) to node[above]{$m_{LU,V}$} (V);
  \draw[->] (U) to node[above]{$m_{U,V}$} (V);
  \draw[->] (LU) to node[below]{$m_{LU,W}$} (W);
  \draw[->] (U) to node[below]{$m_{U,W}$} (W);
  \draw[->] (V) to node[right]{$m_{V,W}$} (W);
\end{tikzpicture}
\]

where the (a) and (c) commute by definition, and (b) commutes by induction hypothesis.

- If $B = LV$: then $m_{LU,LV} = LU \xrightarrow{Lm_{U,V}} LV$, $m_{LV,W} = LV \xrightarrow{\epsilon_V} V \xrightarrow{m_{V,W}} W$, and $m_{LU,LV}; m_{LV,W}$ is

\[
\begin{tikzpicture}
  \node (LU) at (0,0) {$LU$};
  \node (U) at (2,0) {$U$};
  \node (LV) at (4,0) {$LV$};
  \node (V) at (4,-2) {$V$};
  \node (W) at (4,-4) {$W$};
  \node (mLUW) at (0,-4) {$m_{LU,W}$};
  \node (muV) at (2,-2) {$m_{U,V}$};
  \node (muV) at (4,-2) {$m_{V,W}$};
  \draw[->] (LU) to node[above]{$Lm_{U,V}$} (LV);
  \draw[->] (U) to node[above]{$m_{U,V}$} (V);
  \draw[->] (LV) to node[below]{$m_{LV,W}$} (W);
  \draw[->] (U) to node[below]{$m_{U,W}$} (W);
  \draw[->] (LV) to node[right]{$m_{LV,W}$} (W);
\end{tikzpicture}
\]

where (a) commutes by naturality of $\epsilon$, (b) and (d) by definition and (c) by induction hypothesis.
• If it is true for sub-derivations of \((D'' \times E') \leq (D \times E'')\), then by case distinction the only possible \(B\) such that \((D'' \times E') \leq B\) and \(B \leq (D \times E'')\) is \(B = (D' \times E')\) for some vertices \(D'\) and \(E'\). Then \(D \leq D' \leq D''\) and \(E \leq E' \leq E''\), and by induction hypothesis we get the result.

\[\square\]

**Lemma 8.3.19.** For any subvertex \(A\), \(m_{A,A} = id_A\).

*Proof.* Proof by induction on \(A\).

**Lemma 8.3.20.** Given any arrow \(f : A \rightarrow B\), \(K_A \leq K_B\).

*Proof.* We prove it by induction on the construction of \(f\): first the result is true for any elementary arrow and for the identity on objects. Then it is true by induction on any expansion of such arrows, using the definition. Finally, from Lemma 8.3.17 \(\leq\) is transitive and thus it is true for any composition of expansions.

\[\square\]

**Lemma 8.3.21.** For any map \(g : A \rightarrow B\), the following diagram commutes:

\[
\begin{array}{ccc}
A & \xrightarrow{g} & B \\
\downarrow \zeta_A & & \downarrow \zeta_B \\
K_A & \xrightarrow{m_{A,B}} & K_B.
\end{array}
\]

*Proof.* First note that \(m_{A,B}\) is well-defined using Lemma 8.3.20. Then we proceed by case distinctions.

• First, \(g\) can be the identity, then it’s done.

• It can be an elementary arrow. In the case \(\delta_A\), the following diagram has to commute:

\[
\begin{array}{ccc}
LA & \xrightarrow{\delta_A} & L^2A \\
\downarrow \zeta_LA & & \downarrow \zeta_{L^2A} \\
K(LA) & \xrightarrow{m_{K(LA),K(L^2A)}} & K(L^2A).
\end{array}
\]

By definition \(K(L^2A) = K(LA)\), and from Lemma 8.3.15 \(m_{K(LA),K(LA)} = id_{K(LA)}\). Thus diagram becomes:

\[
\begin{array}{ccc}
LA & \xrightarrow{\delta_A} & L^2A \\
\downarrow \zeta_LA & \xrightarrow{(a)} & \downarrow \zeta_{L^2A} \\
\downarrow \zeta_LA & \xrightarrow{(b)} & \downarrow \zeta_{L^2A} \\
K(LA) & \xrightarrow{id} & K(LA),
\end{array}
\]

where (a) commutes by property of comonads, (b) commutes by definition, (c) by vacuity.
• In the case $\epsilon_A$, from Lemma 8.3.14 $A$ can be either a subvertex $U$ or of the form $LB$. In the case $A = LB$, using the same equality as above, the following diagram has to commute:

$$
\begin{array}{ccc}
L^2B & \xrightarrow{\epsilon_{LB}} & LB \\
\zeta_{L^2B} & \downarrow (a) & \downarrow \zeta_{LB} \\
K(LA) & \xrightarrow{\iota} & K(LA),
\end{array}
$$

and $(a)$ commutes by definition, and $(b)$ by vacuity. In the case $A = LU$, the diagram that needs to commute, using Lemma 8.3.14 and Lemma 8.3.15 is:

$$
\begin{array}{ccc}
LU & \xrightarrow{\epsilon_U} & U \\
\zeta_{LU} & \downarrow \zeta_U & \downarrow \zeta_U \\
L(KU) & \xrightarrow{\iota} & L(KU),
\end{array}
$$

which commutes by definition of $\zeta_{LU}$.

• Then it can be the expansion $f$ of an elementary arrow: we prove this case by induction on such an expansion. The base case has been done, it remains to show the general case. Suppose then that the result is true for some expansion $f : A \rightarrow B$. Then we need to show that the diagram commute for $LF$, for $f \times X$ and for $X \times f$.

• Finally, it can be a composition $f$ of such expansions. By induction on the number of expansions composed, we show that the diagram commutes for $f$. What remains to be shown is the iterative case: Suppose $f : A \rightarrow B$ is such a composition, and $g : B \rightarrow C$ is an expansion, we want the following diagram to commute:

$$
\begin{array}{ccc}
A & \xrightarrow{f} & B \\
\zeta_A & \downarrow & \zeta_B \\
KA & \xrightarrow{m_{KAKB}} & KB \\
\end{array}
\quad \begin{array}{ccc}
B & \xrightarrow{g} & C \\
\zeta_B & \downarrow & \zeta_C \\
KB & \xrightarrow{m_{KBKC}} & KC \\
\end{array}
$$

The upper triangle commutes by definition, the lower one using Lemma 8.3.15, the left square by induction hypothesis and the right square because we just proved the result for expansions.

\[ \square \]

**Corollary 8.3.22.** In $D$, any two arrows $f, g : A \rightarrow B$ are equal.

*Proof.* From Lemma 8.3.21 they are both equal to $\zeta_A; m_{KAKB}; \zeta_B^{-1}$. \[ \square \]

*Proof of Theorem.* We prove the theorem when $(F_1, \ldots, F_N)$ consists only of one functor $\times : C^{op} \times C \rightarrow C$.

Consider the free category $D$ with an idempotent comonad described in Definition 8.3.4. There is an inclusion $I : D \rightarrow C^{G}$, and any arrow in $C^{G}$ is the image of an arrow in $D$. Let $f, g : A \rightarrow B$ be arrows of $C^{G}$. There exists $f'$ and $g'$ in $D$ such that $I(f') = f$ and $I(g') = g$. From Corollary 8.3.22 $f' = g'$. Thus $f = g$. \[ \square \]
8.3.3 Duplicable Pairs and Pairs of Duplicable Elements

Unlike the work of [Benton et al. 1993; Bierman 1993], we want to think of the type $!A \otimes B$ as a type of pairs of elements of type $A$ and $B$: we want to use the same operation to access the components as one would use for a pair of type $A \otimes B$, without having to use a dereliction operation.

This immediately raises a concern: consider a pair of elements $\langle x, y \rangle$ of type $!A \otimes B$. Are $x$ and $y$ duplicable? In the usual linear logic interpretation, they are not. Having an infinite supply of pair of shoes does not mean one has an infinite supply of right shoes: we cannot discard the left shoes. On the other hand, in our interpretation of “classical” data as residing in “classical” memory and therefore being duplicable, if the string $\langle x, y \rangle$ is duplicable, then so should be the elements $x$ and $y$. In other words, we want the duplication to “permeate” the pairing.

The choice of such a “permeable” pairing is more or less forced on us by our desire to have no explicit term syntax for structural rules. Consider the following untyped terms, which can be typed if $t$ is of type $!(!B \otimes C)$:

\[
\begin{align*}
\text{let } \langle x, u \rangle &= t \text{ in let } \langle y, z \rangle &= u \text{ in } \langle \langle z, y \rangle, x \rangle, \\
\text{let } \langle x, u \rangle &= t \text{ in } \langle \langle \langle \langle z, y \rangle, x \rangle, u \rangle \rangle.
\end{align*}
\]

First, we expect these two terms to be axiomatically equal. Term (8.3.5) should be of type $!(!C \otimes B) \otimes A)$, regardless of the permeability of the pairing: if $\langle y, z \rangle$ is duplicable, so should be $\langle z, y \rangle$. Now, consider the term (8.3.4) with a non-permeable pairing. In the naive type system, $u$ ends up being of type $B \otimes C$, and the variables $y$ and $z$ in the final recombination end up being respectively of type $B$ and $C$. It is not possible to make $\langle z, y \rangle$ of the duplicable type $!C \otimes B$.

We therefore choose a permeable pairing. Is will be reflected, albeit subtly, in the typing rules $(\otimes. I)$ and $(\otimes. E)$ of Table 9.1. To enforce it in the semantics, we will pick a strong monoidal comonad to model the operator “$!$”.

8.4 Linear Category for Duplication

We now have enough background to define a candidate for the categorical model of the quantum lambda-calculus. As it was advertised, the structure of the categorical semantics will closely follow the one proposed by [Bierman 1993], but with the added twist of a computational monad à la [Moggi 1991].

Definition 8.4.1. A linear category for duplication is a category $C$ with the following structure:

- a symmetric monoidal structure $(\otimes, \top, \alpha, \lambda, \rho, \sigma)$;
- an idempotent, strongly monoidal, linear exponential comonad $(L, \delta, \epsilon, d^L, d^L, \Diamond, \triangle)$;
- a strong monad $(T, \mu, \eta, t)$;
- a Kleisli exponential $\Rightarrow$, that is, a natural bijective map of arrows

$$
\Phi : C(A \otimes B, TC) \to C(A, B \Rightarrow C)
$$

and a bifunctor

$$
\Rightarrow : C^{op} \times C \to C.
$$

The computational linear category is defined as the Kleisli category $C_T$.

Categories with a combination of a monad and a comonad have been studied by many authors [Barr 1990; Hyland and Schalk 2003; Power and Watanabe 2003]. However, none of them are equivalent to linear categories for duplication: either the monad and the comonad commute or the comonad is not idempotent.
Example 8.4.2. Consider the category $\text{ClasObj}$ of Examples 8.3.3 and 8.3.6. It is symmetric monoidal closed, and it has a idempotent, monoidal comonad $(L, \epsilon, \delta, d^L, d^L_L)$. Note that the coherence maps $d^L_L$ and $d^L_{A,B}$ are isomorphisms: the comonad $L$ is strongly monoidal. We can make $\text{ClasObj}$ into a non-trivial linear category for duplication, as follows:

- For every object $(E, B_E)$, define the arrows $\diamond_{(E, B_E)} : L(E, B_E) \rightarrow \top$ and $\triangle_{(E, B_E)} : L(E, B_E) \rightarrow L(E, B_E) \otimes L(E, B_E)$

\[
\diamond_{(E, B_E)} : \sum_{e \in B_E} \langle v| e \rangle, \quad \quad \quad \triangle_{(E, B_E)} : \sum_{e \in B_E} \langle v| e \rangle \otimes e.
\]

Equipped with these two maps, the comonad $L$ becomes an idempotent, strongly monoidal, linear exponential comonad.

- Consider the functor $T : \text{ClasObj} \rightarrow \text{ClasObj}$ sending the object $(E, B_E)$ to $(0, \emptyset)$ and the morphism $f : (E, B_E) \rightarrow (F, B_F)$ to $Tf = id_{(0, \emptyset)}$. Together with the maps

\[
\eta_{(E, B_E)} : (E, B_E) \rightarrow T(E, B_E) = x \mapsto 0,
\]

\[
\mu_{(E, B_E)} : TT(E, B_E) \rightarrow T(E, B_E) = id_{(0, \emptyset)},
\]

\[
t_{(E, B_E), (F, B_F)} : (E, B_E) \otimes T(F, B_F) \rightarrow T((E, B_E) \otimes (F, B_F)) = id_{(0, \emptyset)},
\]

the functor $T$ is a strong monad.

- Remember that $\text{ClasObj}$ is a symmetric monoidal category. At first sight, the functor $(-) \otimes (F, B_F)$ does not have a right adjoint. However, if one defines the object $(E, B_E) \rightarrow (F, B_F) = (0, \emptyset)$, there is a natural isomorphism

\[
\text{ClasObj}((E, B_E) \otimes (F, B_F), T(G, B_G)) \xrightarrow{\sim} \text{ClasObj}((E, B_E), (F, B_F) \rightarrow (G, B_G)),
\]

making $\rightarrow$ into a Kleisli exponential.

With the structure defined above, $\text{ClasObj}$ is a linear category for duplication.

Remark 8.4.3. A linear category for duplication gives rise to a double adjunction

\[
\begin{array}{ccc}
\text{C}_L & \xleftarrow{\perp} & \text{C} \\
\downarrow & & \downarrow \\
F_T & \xleftarrow{\perp} & \text{C}_T
\end{array}
\]

Here the left adjunction arises from the co-Kleisli category $\text{C}_L$ of the comonad $L$. It is as in the linear-non-linear models of [Benton, 1994], and $\text{C}_L$ is a category of classical (non-quantum) values. The right adjunction arises from the Kleisli category $\text{C}_T$ of the computational monad $T$, as in [Moggi, 1991]. Here $\text{C}_T$ is a category of (effectful) quantum computations.

Example 8.4.4. In Example 8.4.2 the category quantum values is the category $\text{ClasObj}$ and the category of classical values $\text{ClasObj}_L$ is equivalent to the category $\text{FinSet}$. Since the image of any object by $T$ is $(0, \emptyset)$, one can say that the category of computations $\text{ClasObj}_T$ contains only one element: the diverging computation.
Chapter 9

A Computational Lambda Calculus for Duplication

In this chapter, we describe a subset of the lambda calculus for quantum computation of Chapter 6. We focus primarily on the type system and language, and not on the structure of the actual “built-in” quantum operations (such as unitary operators and measurements). The language will be a generic call-by-value linear lambda calculus, which is parametric on some primitive operations that are not further explained. It should be understood, however, that the need to support primitive quantum operations motivates particular features of the type system.

9.1 An Indexed Lambda Calculus

We describe a linear typed lambda calculus with higher-order functions and pairs. The language is designed to manipulate both classical data, which is duplicable, and quantum data, which is non-duplicable. For simplicity, we assume the language is strictly linear, and not affine linear as in Chapter 6. This means duplicable values are both copyable and discardable, whereas non-duplicable values must be used once, and only once.

9.1.1 Type System

We define a type system together with a subtyping relation similar as in Section 6.3. The only difference is that type constants are parametrized.

**Definition 9.1.1.** The set of types is

$$\text{Type } A, B ::= \alpha \mid (A \rightarrow B) \mid (A \otimes B) \mid \top \mid !A,$$

where $\alpha$ ranges over type constants. While the remainder of this thesis does not depend on the choice of type constants, in our main application this is intended to include a type $qbit$ of quantum bits, and a type $bit$ of classical bits. The meaning of the types is the same as in Section 6.3. The type system comes also with the subtyping relation of Definition 6.3.3.

**Remark 9.1.2.** The type system satisfies Lemmas 6.3.4 and 6.3.5.
9.1.2 Terms

The language is a computational language: it is divided into values on the one hand, and general terms, or computations, on the other.

As in Chapter 6, our goal is to find a semantics for the language. The semantics has to be compositional, that is, the semantics of a given term has to be derived from the semantics of its subterm. We are focusing on a language close to the one of Chapter 6. In this language, consider the following two typing judgements, where $V$ and $W$ are values:

$$x : A \triangleright V : !B,$$

$$y : B \triangleright W : C.$$

The typing judgement $x : A \triangleright let y = V in W : C$ has two valid typing derivations:

$$x : A \triangleright V : !B \quad y : B \triangleright W : C \quad x : A \triangleright let y = V in W : C,$$

$$x : A \triangleright V : B \quad y : B \triangleright W : C \quad x : A \triangleright let y = V in W : C.$$

In order to be able to define a semantics for judgements like $x : A \triangleright let y = V in W : C$, we will use indexed terms. The purpose of the indexing will be to make sure that each valid typing judgement has a reasonably unique typing tree.

**Definition 9.1.3.** We define the notion of core values, extended values and terms as follows:

- CoreValue $U, U' ::= x^A | c^A | s^n | \lambda^n x^A. M | \langle U, U' \rangle^n$,
- ExtValue $V, W ::= U | \langle V, W \rangle^n | let \ x^A = V in W | let \ \langle x^A, y^B \rangle^n = V in W | let \ s^* = V in W$,
- Term $M, N ::= U | \langle M, N \rangle^n | (M N) | let \ \langle x^A, y^B \rangle^n = M in N | let \ s^* = M in N$.

where $n$ is an integer, $c$ ranges over a set of constant terms, $x$ over a set of term variables and $\alpha$ over a set of constant types. We abbreviate $(\lambda^n x^A. M) N$ by let $x^A = N in M$, $\lambda^n x^m \top$, let $s^* = x^T in M$ by $\lambda^n x^m. M$ and we omit numerical indexes when they are null.

Core values and extended values are often simply called values when the context is clear. General terms are also called computations.

**Remark 9.1.4.** A technicality we introduced is the distinction between core values and extended values. The core values have the same structure as the values of Section 9.1.3. The need for extended values comes from the fact that we wish to have a destructor for the unit type $\top$ in the category of values, so that we are allowed to write the judgement

$$x : A, y : \top \triangleright let \ s^* = y^\top in x^A : A.$$

However, as it will be further developed in Remark 9.1.9, this term will cause problem with substitution in certain situations. We are forced to restrict the substitution to core values.

**Convention 9.1.5.** We use the notations $\square, \square$ and $\square$ as placeholders for $x^A$, $\ast$ and $\langle x^A, y^B \rangle^n$ (not respectively). We use typed version of them when we want to enforce the underlying type. For example, $\square$ can stand for $x^A$, $\ast$ if $A = \top$, or $\langle x^B, y^C \rangle^n$ if $A = \nu^n (B \odot C)$.

We also define a notion of untyped terms as terms with no index:

$$PureTerm M, N ::= x | c | \lambda x. M | (M N) | \langle M, N \rangle | let \ \langle x, y \rangle = M in N | let \ s^* = M in N.$$

The erasure operation $Erase : Term \rightarrow PureTerm$ is defined as the operation of removing the types and integers attached to a given indexed term. If $M = Erase(M)$, we say that $M$ is an indexation of $M$. We define notions of free variables and $\alpha$-equivalence on terms, denoted by $=_{\alpha}$, in the usual way (see for example Section 6.1).
9.1. An Indexed Lambda Calculus

Definition 9.1.8. The indexation of terms induces a partial map from terms to types, called the raw type of $M$. We write $M : A$ to say that $A$ is the raw type of $M$, or that $M$ is raw-typed, and we define it inductively as follows:

- $x^A : A$,
- $c^B : B$,
- $\ast^n : \ast^n\top$,
- $\lambda^n x^A.\{M : A \to B\} : \ast^n(A \to B)$,
- $\lambda^n x^A.\{M : B\} : \ast^n(A \to B)$,
- $\lambda^n x^A.\{M : B\} : \lambda^n(A \to B)$,
- $\lambda^n x^A.\{M : B\} : \lambda^n(A \to B)$,
- $\lambda^n x^A.\{M : B\} : \lambda^n(A \to B)$,
- $\lambda^n x^A.\{M : B\} : \lambda^n(A \to B)$,
- $\lambda^n x^A.\{M : B\} : \lambda^n(A \to B)$.

We use the notation $\{M : A\}$ to say “the term $M$ when it is of type $A$”.

9.1.3 Typing Judgements

Definition 9.1.9. We define the notion of typing context of Definition [25, 60].

A typing judgement is a tuple $\Delta \triangleright M : A$, where $M$ is an indexed term, $A$ is a type, and $\Delta$ is a typing context. As before, to each constant term $c$ we assign a type $!A_c$. A valid typing judgement is a typing judgement that can be derived from the typing rules in Table 9.1.

Definition 9.1.10. We define a special derived rule called (let);

\[
\begin{align*}
!\Delta, \Gamma_1, x : A &\triangleright M : B \quad !\Delta, \Gamma_2 &\triangleright N : A \\
!\Delta, \Gamma_1, \Gamma_2 &\triangleright \text{let } x^A = N \text{ in } M : B
\end{align*}
\]

It is derived using Definition [24, 3] stating that let $x^A = N$ in $M$ is the term $(\lambda^0 x^A.M)N$:

\[
\begin{align*}
!\Delta, \Gamma_1, x : A &\triangleright M : B \\
!\Delta, \Gamma_1 &\triangleright \lambda^n x^A.\{M : B\} \quad (\lambda_1) \\
!\Delta, \Gamma_1, \Gamma_2 &\triangleright (\lambda^n x^A.M)N : B
\end{align*}
\]
Lemma 9.1.11. Consider a valid typing judgement $\Delta \vdash M : A$. Then $\text{FV}(M) \subseteq |\Delta|$. 

Proof. Proof by induction on the size of $M$.

Case $M \equiv x^B$. The only applicable rule is $(ax_1)$. Then the typing judgement is of the form $!\Delta, x : A \vdash x^B : B$. We thus have $\text{FV}(M) = \{x\} \subseteq |\Delta| \cup \{x\} = |\Delta, x : A|$. 

Cases $M \equiv c^B$ and $M \equiv s^n$. In both cases $\text{FV}(M) = \emptyset$. The result is true by vacuity.

Case $M \equiv \lambda^n x^A . N$. The only applicable rule is $(\lambda_i)$, with $i = 1$ if $n = 0$, $i = 2$ otherwise. Thus there exists some $B$ such that the typing judgement is of the form $\Delta \vdash \lambda^n x^A . N : !^n(A \rightarrow B)$ with $\Delta, x : A \vdash N : B$. By induction hypothesis, we have $\text{FV}(N) \subseteq |\Delta, x : A| = |\Delta| \cup \{x\}$. By definition of free variables, $\text{FV}(\lambda^n x^A . N) = \text{FV}(N) \setminus \{x\} \subseteq |\Delta|$. 

Case $M \equiv NP, (N, P)^n$ and let $* = M$ in $N$. The applicable rules are respectively $(app)$, $(\otimes, I)$ and $(\top, E)$. The context $\Delta$ splits into $|\Delta'$, $\Gamma_1, \Gamma_2$ such that $|\Delta', \Gamma_1 \vdash N : B$ and $|\Delta', \Gamma_2 \vdash P : C$ for some types $B$ and $C$. By induction hypothesis, $\text{FV}(N) \subseteq |\Delta', \Gamma_1|$ and $\text{FV}(P) \subseteq |\Delta', \Gamma_2|$. Since $\text{FV}(M) = \text{FV}(N) \cup \text{FV}(P)$, we have $\text{FV}(M) \subseteq |\Delta|$. Thus $\text{FV}(M) \subseteq |\Delta|$. 

And this closes the proof of Lemma 9.1.11. □

Lemma 9.1.12. Consider a term $M$ and any valid typing judgement $\Delta \vdash M : B$. Then the raw type $A$ of $M$ exists, and the type $B$ is equal to $A$.

Proof. Proof by structural induction on $M$, looking at a typical typing judgement $\Delta \vdash M : A$.

Case $M \equiv x^B$. The only applicable rule is $(ax_1)$. Then the typing judgement is of the form $!\Delta, x : A \vdash x^B : B$. We thus have $M : B$.

Case $M \equiv c^B$. The only applicable rule is $(ax_2)$. Thus the typing judgement is of the form $!\Delta \vdash c^B : B$. We thus have $M : B$.

Case $M \equiv s^n$. The only applicable rule is $(\top, I)$. The type $B$ is of the form $!^n \top$ and the typing judgement is of the form $!\Delta \vdash s^n : !^n \top$. Thus we have $M : B$.

Case $M \equiv \lambda^n x^A . N$. The only applicable rule is $(\lambda_1)$. Thus there exists some $B$ such that the typing judgement is of the form $\Delta \vdash \lambda^n x^A . N : A \rightarrow B$ with $\Delta, x : A \vdash N : B$. By induction hypothesis, we have $N : B$. Thus $\lambda^n x^A . N : A \rightarrow B$.

Case $M \equiv \lambda^{n+1} x^A . N$. The only applicable rule is $(\lambda_2)$. Then $\Delta = !\Delta'$, and there exists some $B$ such that the typing judgement is of the form $!\Delta' \vdash \lambda^{n+1} x^A . N : !^{n+1}(A \rightarrow B)$. The judgement is inferred from $!\Delta', x : A \vdash N : B$. By induction hypothesis, we have $N : B$. This makes $\lambda^{n+1} x^A . N : !^{n+1}(A \rightarrow B)$.

Case $M \equiv NP$. The only applicable rule is $(app)$. The context $\Delta$ splits into $!\Delta'$, $\Gamma_1, \Gamma_2$ such that $!\Delta', \Gamma_1 \vdash N : A \rightarrow B$ and $!\Delta', \Gamma_2 \vdash P : A$. By induction hypothesis, $N : A \rightarrow B$ and $M : A$. Thus $NP : B$. 

9.1. An Indexed Lambda Calculus 112
Case $M \equiv \langle N, P \rangle^n$. The only applicable rule is $\otimes, I$. This means $\Delta \triangleright (N, P)^n : !^n(A \otimes B)$, and the context $\Delta$ splits into $!\Delta', \Gamma_1, \Gamma_2$ such that $!\Delta', \Gamma_1 \triangleright N : !^n A$ and such that $!\Delta', \Gamma_2 \triangleright P : !^n B$. By induction hypothesis, $N : !^n A$ and $P : !^n B$. Thus $l(N, P)^n : !^n(A \otimes B)$.

Case $M \equiv \text{let } * = N \text{ in } P$. The only applicable rule is $\otimes, E$. For some type $A$, the typing judgement $\Delta \triangleright \text{let } * = N \text{ in } P : A$ is valid and $\Delta$ splits into $!\Delta', \Gamma_1, \Gamma_2$ such that $!\Delta', \Gamma_1 \triangleright N : \top$ and such that $!\Delta', \Gamma_2 \triangleright P : A$. By induction hypothesis, $P : A$. This means let $* = N \text{ in } P : A$.

Case $M \equiv \text{let } (x^n, y^A)^n = N \text{ in } P$. The only applicable rule is $\otimes, E$. We have the valid typing judgement $\Delta \triangleright \text{let } (x^n, y^A)^n = N \text{ in } P : B$ for some type $B$ and $\Delta$ splits into $!\Delta', \Gamma_1, \Gamma_2$ such that $!\Delta', \Gamma_1 \triangleright N : !^n(A \otimes A_n)$ and such that $!\Delta', \Gamma_2, x : !^n A_1, y : !^n A_2 \triangleright P : B$. By induction hypothesis, $P : B$. We deduce that $M : B$.

Therefore, regardless of the chosen typing judgement $\Delta \triangleright M : A$, $A$ is always equal to the raw type of $M$.

Remark 9.1.13. Although the type $A$ in a valid typing judgement $\Delta \triangleright M : A$ is only dictated by $M$, the context $\Delta$ cannot be chosen arbitrarily for the judgement to be valid. Consider the term $f^A\to A(f^A\to AxA) : A$. The context $(f : !(A \to A), x : A)$ makes a valid typing judgement, but $(f : A \to A, x : A)$ does not.

Lemma 9.1.14. Consider a valid typing judgement $\Delta \triangleright M : B$ and the subtyping relation $\Delta' \ll \Delta$. Then $\Delta' \triangleright M : B$ is valid.

Proof. Let $\pi$ be a derivation of $\Delta \triangleright M : B$ The proof is done by induction on the size of $\pi$.

Case (ax1). The typing judgement is of the form $!\Delta, x : A \triangleright x^B : B$ with $A < B$. We have then $\Delta = (x^n, x : A)$. Since $\Delta' \ll \Delta$, the context $\Delta'$ is of the form $(!\Delta', x : !A')$, where $!\Delta' \ll !\Delta$ and $A' < A$. Note that $!\Delta'$ is duplicable from Lemma 9.1.3. Since $A < B$, we have $A' < B$. Applying rule (ax1), we have that $!\Delta', x : A' \triangleright x^B : B$ is valid.

Cases (ax2) and (T, I). The typing judgements are respectively of the form $!\Delta \triangleright c^B : B$ and $!\Delta \triangleright *^n : !^n \top$, where $\Delta = !\Delta$. Since $\Delta' \ll \Delta$, from Lemma 9.1.3 $\Delta' = !\Delta'$. We can therefore apply (ax2) (respectively (T, I)) and get that $!\Delta' \triangleright M : B$ is valid.

Case (λI). There exists some $B$ such that the typing judgement is of the form $\Delta \triangleright \lambda^0 x^A : A \to B$ with $\Delta, x : A \triangleright N : B$. Since $\Delta' \ll \Delta$, by induction hypothesis, we have $\Delta', x : A \triangleright N : B$. Applying (λ1), we get that $\Delta \triangleright \lambda^0 x^A : A \to B$ is valid.

Case (λ2). We have $\Delta = !\Delta$, and there exists some $B$ such that the typing judgement is of the form $!\Delta \triangleright \lambda^0 x^A : A \to B$. The judgement is inferred from $!\Delta, x : A \triangleright N : B$. Since $\Delta' \ll !\Delta$, from Lemma 9.1.3 $\Delta' = !\Delta'$. We have therefore $(!\Delta', x : A) \ll (!\Delta, x : A)$. By induction hypothesis, we have $!\Delta', x : A \triangleright N : B$. Finally, we can apply the typing rule (λ2) in order to get that $!\Delta' \triangleright \lambda^0 x^A : A \to B$ is valid.

Cases (app), (⊗, I), (T, E) and (⊗, E). The term $M$ is respectively of the form

$$M \equiv NP,$$

$$M \equiv \langle N, P \rangle^n,$$

$$M \equiv \text{let } * = N \text{ in } P,$$

$$M \equiv \text{let } (x^{D_1}, y^{D_2})^n = N \text{ in } P.$$

The typing context $\Delta$ splits into $(!\Delta, \Gamma_1, \Gamma_2)$ such that $!\Delta, \Gamma_1 \triangleright N : B_1$ and such that $!\Delta, \Gamma_2, \Lambda \triangleright P : B_2$, for some types $B_1, B_2$ and a context $\Lambda$ that is empty for the first three cases and equal to $(x : D_1, y : D_2)$ in the last case.
Since \( \Delta' \subset \Delta \), the context \( \Delta' \) splits into \( (\Delta', \Gamma_1', \Gamma_2') \), where \( \Delta' \subset \Delta \), \( \Gamma_1' \subset \Gamma_1 \) and \( \Gamma_2' \subset \Gamma_2 \).

By induction hypothesis, in each of the four cases, \( \Delta', \Gamma_1' \vdash N : B_1 \) and \( \Delta', \Gamma_2' \vdash P : B_2 \) are valid. Applying the typing rule corresponding to the case considered, we get that \( \Delta' \vdash M : B \) is valid.

Thus, by induction the lemma is valid.

**Lemma 9.1.15.** Let \( \Gamma \vdash x : A \) be a valid typing judgement with \( x \) a core value. Then \( \Gamma = !\Delta \) for some context \( \Delta \).

**Proof.** The proof is done by induction on the structure of \( \Gamma \). Base cases:

**Case** \( \Gamma = x^B \). The only typing rule available for it is \((ax_1)\). The type \( B \) is therefore of the form \( !B' \).

The typing judgement is then of the form \( !\Delta, x : A \vdash x^B' : !B' \), with \( A < : !B \). By Lemma 6.3.5, the type \( A \) is of the form \( !A' \), and the context \( \Gamma \) is of the requested form.

**Case** \( \Gamma = e^B \). Here, the only typing rule available for it is \((ax_2)\). The typing judgement is therefore of the form \( !\Delta \vdash e^B : B \); the lemma is true in this case.

**Case** \( \Gamma = *^m \). The only typing rule available for it is \((\top I)\). The typing judgement is therefore of the form \( !\Delta \vdash *^{m+1} ; *^{m+1} \top \); as in the previous case, the lemma is true.

**Case** \( \Gamma = \lambda^m x^B.M \). There are two available typing rules for validating \( \Gamma \vdash \lambda^m x^B.M : !A \), but only one, that is \((\lambda_2)\), for a duplicable lambda abstraction. This means \( m = n + 1 \), and \(!A = !^{n+1}(B \rightarrow C)\). The fact that \( \Gamma \) is of the form \( !\Delta \) is one of the requirement of this typing rule.

The only inductive case is the following.

**Case** \( \Gamma = (V_1, V_2)^m \). The only available typing rule is \((\otimes I)\), stating that the type \(!A\) is of the form \( !^{n+1}(B \otimes C) \), that \( m = n + 1 \) and that \( \Gamma = !\Delta, \Gamma_1, \Gamma_2 \), with \( !\Delta \vdash V_1 : !^{n+1}B \) and \( !\Delta, \Gamma_1 \vdash V_2 : !^{n+1}C \) being valid typing judgements. Since \( V_1 \) and \( V_2 \) are core values, the induction hypothesis applies, and \( \Gamma_1 \) is of the form \( !\Gamma'_1 \) and \( \Gamma_2 \) is of the form \( !\Gamma'_2 \). This makes \( \Gamma \) of the requested form.

Hence, by structural induction on \( \Gamma \), Lemma 9.1.15 is valid.

**Lemma 9.1.16.** Consider the following valid typing judgement: \( \Delta, x : A \vdash M : B \). Then for every free instance \( x^{A'} \) in \( M \), \( A < : A' \).

**Proof.** Proof by induction on the size of the derivation of \( \Delta, x : A \vdash M : B \).

**Case** \((ax_1)\). The typing judgement is of the form \( !\Delta, y : A \vdash y^B : B \), with \( A < : B \). If \( y = x \), then the result is true since \( x^B \) is the only free variable. If \( y \neq x \), there is no instance of \( x \), so the result is true by vacuity.

**Case** \((ax_2)\). The typing judgement is of the form \( !\Delta \vdash c^B : B \); there is no free variable, the result is true by vacuity.

**Cases** \((\lambda_1)\) and \((\lambda_2)\). The typing judgement is of the form \( \Delta, x : A \vdash \lambda^ny^B.M : !^n(B \rightarrow C) \), with \( \Delta, x : A, y : B \vdash M : C \) and \( y \neq x \). By induction hypothesis, for every free instance \( x^{A'} \) of \( M \) one has \( A < : A' \). Thus it is also true for every free instance \( x^{A'} \) of \( \lambda^n y^B.M \).

**Cases** \((app), (\otimes I), (\top E)\) and \((\otimes E)\). The term \( M \) is respectively of the form

\[
M \equiv NP, \quad M \equiv \langle N, P \rangle^n, \quad M \equiv \text{let } * = N \text{ in } P, \quad M \equiv \langle y^D_1, z^D_2 \rangle^n = N \text{ in } P.
\]
The typing context $\Delta, x : A$ splits into $!\Delta, \Gamma_1, \Gamma_2$ such that $!\Delta, \Gamma_1 \triangleright N : B_1$ and such that $!\Delta, \Gamma_2, \Lambda \triangleright P : B_2$, for some types $B_1, B_2$ that depend on the rule used and for some context $\Lambda$, empty in the three first cases and equal to $(y : D_1, z : D_2)$ in the last case. Using $\alpha$-equivalence one can assume that $x$ is different from $y$ and $z$. There are three cases, depending on where $x : A$ is to be found.

Found in $\Gamma_1$. Then there is no free instance of $x$ in $P$, and by induction hypothesis every free instance $x^{A'}$ in $N$ are such that $A <: A'$. Then for all free instances $x^{A'}$ in $M$, $A <: A'$.

Found in $\Gamma_2$. Then there is no free instance of $x$ in $N$, and by induction hypothesis every free instance $x^{A'}$ in $P$ are such that $A <: A'$. Then for all free instances $x^{A'}$ in $M$, $A <: A'$.

Found in $!\Delta$. Applying induction hypothesis on both typing judgements, we get that or all free instances $x^{A'}$ in $N$ and in $P$, thus in $M$, $A <: A'$.

Thus by induction, the lemma is valid. $\square$

Lemma 9.1.17. Consider a valid typing judgement $\Delta \triangleright M : A$, where $M$ is of the form $(N, P)^n$, $NP$, let $x = N$ in $P$ or let $(y^C, z^D)^n = N$ in $P$. If $x : B$ ($x \neq y, z$) is in $\Delta$ such that $B$ is not of the form $\forall B'$ then $x$ cannot be free both in $N$ and in $P$.

Proof. The proof is done by case distinction on the first typing rule used in the derivation of $\Delta \triangleright M : A$. $\square$

Definition 9.1.18. In a typing judgement $\Delta \triangleright M : A$, a term variable $x \in |\Delta|$ is called dummy if $x \notin \text{FV}(M)$.

Lemma 9.1.19. Let $\Delta, x : A \triangleright M : B$ be a valid typing judgement where $x$ is a dummy variable for $M$. Then

1. the type $A$ is of the form $!A'$;
2. the typing judgement $\Delta \triangleright M : B$ is valid;
3. for any fresh variable $y$ and any type $C$, the typing judgement $\Delta, x : A, y : !C \triangleright M : B$ is valid.

Proof. We prove by induction on the size of the derivation of $\Delta, x : A \triangleright M : B$.

Case (ax1). The typing judgement is of the form $!\Delta', z : D \triangleright z^E : E$, with $D \triangleleft E$. If there is a dummy variable $x$, it is in $|\Delta'|$ and is therefore duplicable. If we write $!\Delta' = (!\Delta'', x : A)$, by rule (ax1) the typing judgement $!\Delta'', z : D \triangleright z^E : E$ is valid. Now, if for some fresh variable $y$ if we consider the typing judgement $!\Delta', z : D, y : !C \triangleright z^E : E$, it is valid by the same rule (ax1).

Case (ax2) and case (T.I). The typing judgement has a context of the form $!\Delta'$. If there is a dummy variable $x$, it is in $|\Delta'|$ and is therefore duplicable. Removing this variable from the context does not change the validity of the judgement. Now, if we extend the context to $!\Delta', y : !C$, the typing judgement is again valid by the same rule (ax2) or (T.I).

Case (l1) and case (l2). The typing judgement is $\Delta, x : A \triangleright \lambda^a z^D.N : !^n(D \rightarrow E)$, such that $\Delta, x : A, z : D \triangleright N : E$ is valid. If $x \notin \text{FV}(\lambda^a z^D.N)$, then $x$ is not in $\text{FV}(N)$ (since $x \neq z$), and by induction hypothesis its type is of the form $!A'$.

Now, consider some fresh $y$ and some type $C$. By induction hypothesis, $\Delta, y : !C, x : A, z : D \triangleright N : E$ and $\Delta, z : D \triangleright N : E$ are valid. If the rule applied was $(l_1)$, then the integer $n$ is 0, and one can again apply $(l_1)$ to get $\Delta, x : A, y : !C \triangleright \lambda^a z^D.N : !^n(D \rightarrow E)$ and $\Delta \triangleright \lambda^a z^D.N : !^n(D \rightarrow E)$. If the rule applied was $(l_2)$, then the integer $n$ is greater than 0 and the context $\Delta$ is of the form $!\Delta'$. Then rule $(l_2)$ can again be applied to get $!\Delta', x : A, y : !C \triangleright \lambda^a z^D.N : !^n(D \rightarrow E)$ and $!\Delta' \triangleright \lambda^a z^D.N : !^n(D \rightarrow E)$. 


9.1. An Indexed Lambda Calculus

Cases (app), (⊗I), (⊤E) and (⊗E). The term $M$ is respectively of the form

$$M \equiv NP,$$
$$M \equiv \text{let } * = N \text{ in } P,$$
$$M \equiv (N, P)^n,$$
$$M \equiv \text{let } \langle x^{D_1}, y^{D_2} \rangle^n = N \text{ in } P.$$

The typing context $(\Delta, x : A)$ splits into $\Delta', \Gamma_1, \Gamma_2$ such that $\Delta', \Gamma_1 \triangleright N : B_1$ and such that $\Delta', \Gamma_2 \triangleright P : B_2$, for some types $B_1, B_2$ that depend on the rule used.

First let us prove that $\Delta, x : A, y : !C \triangleright M : B$ is valid when $y$ is a fresh variable and $C$ is any type. By induction hypothesis, $\Delta', \Gamma_1, y : !C \triangleright N : B_1$ is valid. So applying back the rule that was used, we get that $\Delta, x : A, y : !C \triangleright M : B$ is valid.

Second, let us show that if $x \notin FV(M)$, then its type is of the form $!A'$ and $\Delta \triangleright M : B$ is valid. There are three cases, depending on where $x$ is to be found.

Found in $\Gamma_1$. Then since $x \in \{!\Delta', \Gamma_1\}$ but $x \notin FV(N)$, by induction hypothesis the type of $x$ in the context is of the form $!A$, and if $\Gamma_1 = (\Gamma_1', x : A)$, we have $\Delta, \Gamma_1' \triangleright N : B_1$ valid. Applying back the typing rule used, we get $\Delta \triangleright M : B$ valid.

Found in $\Gamma_2$. Then since $x \in \{!\Delta', \Gamma_2\}$ but $x \notin FV(P)$, by induction hypothesis the type of $x$ in the context is of the form $!A$, and if $\Gamma_2 = (\Gamma_2', x : A)$, we have $\Delta, \Gamma_2' \triangleright P : B_2$ valid. Applying back the typing rule used, we get $\Delta \triangleright M : B$ valid.

Found in $!\Delta'$. Then the type of $x$ is automatically of the form $!A$. By induction hypothesis, if $\Delta' = (\Delta'', x : A)$, both $\Delta'', \Gamma_1 \triangleright N : B_1$ and $\Delta'', \Gamma_2 \triangleright P : B_2$ are valid. Applying back the typing rule used, we get $\Delta \triangleright M : B$ valid.

And this ends the proof. □

9.1.4 Type Casting and Substitution Lemma

We now wish to define a set of call-by-value equations on typed terms. To be able to define a well-typed $\beta$-reduction rule, we first need a substitution lemma. Given a value $x : A \triangleright V : B$ and a term $y : B \triangleright M : C$, one expects $x : A \triangleright M[V/y] : C$ to be well typed. Unfortunately, in the presence of Church-style type indices and subtyping, the naive notion of substitution is not technically well-typed. For example, consider the following valid typing derivations:

$$\frac{A \rightarrow A <:\!A \rightarrow A}{x : A \rightarrow A \triangleright x^A \triangleright A : !A \rightarrow A,} \quad \frac{A <:\!A}{y : A \triangleright y^A : A \triangleright \lambda^0 y^A . y^A : A \rightarrow \rightarrow A.}
$$

Having two valid typing judgements of the form $x : B \triangleright M : C$ and $\triangleright V : B$, it is reasonable to expect to be able to substitute $V$ for $x$ in $M$. Doing so yields the following typing judgement:

$$\triangleright \lambda^0 y^A . y^A : !A \rightarrow A,$$

which is not technically valid due to the superscript on the bound variable. The standard trick to resolve this issue (see e.g. [Pierce 2002]) is to use implicit typecasting, or coercion. In this section we give a precise definition of substitution using typecasting.

Definition 9.1.20. Given $M : A$ and $A <:\!B$, we define the map $\text{TypeCast}(M, B)$, written $\{M : A <:\!B\}$, or simply $\{M <:\!B\}$, by induction on $M$ as follows:

- Base cases:
  $$\{x^A <:\!B\} = x^B, \quad \{c^A <:\!B\} = c^B, \quad \{s^n <:\!m^\top\} = s^m.$$
9.1. An Indexed Lambda Calculus

- Induction cases:
  \[
  \begin{align*}
  \{\lambda^n x^n. M <; !^k (A' \rightarrow B')\} &= \lambda^k x^{A'}. \{M <; B'\}, \\
  \{M \{N : A\} <; B\} &= \{M <; A \rightarrow B\}N, \\
  \{\langle M, N \rangle^n <; !^m (A \otimes B)\} &= \{\{M <; !^m A\}, \{N <; !^m B\}\}^m, \\
  \{\text{let } \langle x^A, y^B \rangle^n = M \text{ in } N <; C\} &= \{\langle x^A, y^B \rangle^n = M \text{ in } \{N <; C\}\}, \\
  \{\text{let } * = M \text{ in } N <; C\} &= \{\text{let } * = M \text{ in } \{N <; C\}\}.
  \end{align*}
\]

Lemma 9.1.21. Suppose \(N : A\) and \(A <; A'\). Then
\[
\{\text{let } x^B = M \text{ in } N <; A'\} = (\text{let } x^B = M \text{ in } \{N <; A'\}).
\]

Proof. Using Definition 9.1.20
\[
\{\text{let } x^B = M \text{ in } N <; A'\} = ((\lambda^0 x^B. N)M <; A')
\]
\[
= (\lambda^0 x^B. N <; A'\rightarrow A')M
\]
\[
= (\lambda^0 x^B. \{N <; A'\})M
\]
\[
= (\text{let } x^B = M \text{ in } \{N <; A'\}).
\]

And this ends the proof.

Lemma 9.1.22. Suppose that \(M : A\) and that \(A <; A' <; A''\). Then \(\{M <; A'\} <; A''\) = \(\{M <; A''\}\).

Proof. Proof by induction on the size of \(M\).

Lemma 9.1.23. Suppose that \(M : A\). then \(\{M <; A\} = M\).

Proof. Proof by structural induction on \(M\).

Lemma 9.1.24. If the raw type \(A\) of \(M\) exists and if \(A <; A'\), then \(\{M <; A'\}\) is well-defined and of raw type \(A'\).

Proof. Proof by structural induction on \(M\).

Lemma 9.1.25. Suppose \(M : A\) and \(A <; A'\). Then \(\text{Erase}(M) = \text{Erase}\{M <; A'\}\).

Proof. Proof by structural induction on \(M\).

Lemma 9.1.26. Consider a valid typing judgement \(\Delta \triangleright M : A\) and the subtyping relation \(A <; A'\). Then \(\Delta \triangleright \{M <; A'\} : A'\) is valid.

Proof. Proof by induction on the size of \(M\).

Lemma 9.1.27. Suppose \(M : A\) and \(A <; A'\). If \(M\) is a core value (respectively an extended value), so is \(M' = \{M <; A'\}\).

Proof. Lemma 9.1.26 states that \(\text{Erase}(M) = \text{Erase}(M')\) and Lemma 9.1.6 that since \(M\) is a (core, extended) value, so is \(M'\).
Definition 9.1.28. Given two valid typing judgements \( !\Delta, \Gamma_1 \triangleright V : A \) and \( !\Delta, \Gamma_2, x : A \triangleright M : B \) where \( V \) is a core value, where \( M \) is any term, and where \( |\Gamma_1| \cap |\Gamma_2| = \emptyset \), we define the substitution \( M[V/x] \) (capture avoiding) of \( x \) by \( V \) in \( M \) as follows: we replace each free instance \( x^A \) (where \( A : A' \) from Lemma 9.1.16 in \( M \) by \( V < : A' \)). This is formally defined as follows:

\[
\begin{align*}
(x^A)[V/x] &= \{ V < : A' \}, \\
(c^B)[V/x] &= c^B, \\
(*^n)[V/x] &= *^n, \\
(MN)[V/x] &= (M[V/x])(N[V/x]), \\
((M, N)^n)[V/x] &= (M[V/x], N[V/x])^n, \\
(let \ast = M \text{ in } N)[V/x] &= let \ast = (M[V/x]) \text{ in } (N[V/x]), \\
(\lambda^n x^C.M)[V/x] &= \lambda^n x^C.M, \\
(\lambda^n y^C.M)[V/x] &= \lambda^n y^C.(M[V/x]) \quad \text{(if } y \neq x \text{ and } y \not\in FV(V)) \end{align*}
\]

If \( y = x \) or \( z = x \):

\[
(let \ (y^C, z^D)^n = M \text{ in } N)[V/x] = (let \ (y^C, z^D)^n = (M[V/x]) \text{ in } N).\]

If \( y \neq x \) and \( z \neq x \) (and \( y, z \not\in FV(V) \)):

\[
(let \ (y^C, z^D)^n = M \text{ in } N)[V/x] = (let \ (y^C, z^D)^n = (M[V/x]) \text{ in } (N[V/x])).\]

Definition 9.1.29. Suppose that \( !\Delta, \Gamma_i \triangleright V_i : A_i \) are valid typing judgements, for \( i = 1 \ldots n \), where \( V_i \) is a core value. Suppose that \( !\Delta, \Gamma, x_1 : A_1, \ldots, x_n : A_n \triangleright M : B \) is valid. By \( \alpha \)-equivalence one can assume that the \( x_i \)’s are not free in any \( V_j \)’s. Assuming this, we define the multiple substitution \( M[V_1/x_1, \ldots, V_n/x_n] \) by induction on \( n \):

- If \( n = 0 \), we define \( M[] \) to be \( M \).
- If \( n \neq 0 \), we define \( M[V_1/x_1, \ldots, V_n/x_n] \) to be \( (M[V_1/x_1, \ldots, V_{n-1}/x_{n-1}])[V_n/x_n] \), using Definition 9.1.28.

Remark 9.1.30. Note that we only allow substitution by core values, and not by extended values nor by general terms. This is due to the strict linearity of the type system. Suppose that we are given the valid typing judgements

\[
x : \top \triangleright let \ast = x^\top \text{ in } ^1 \top, \quad y : \top \triangleright c : A_c.
\]

If we were allowed to substitute term variables by extended values, then one could substitute the former term in place of \( y \) in the latter term. However, doing so yield the non-valid typing judgement

\[
x : \top \triangleright c : A_c.
\]

The problem comes from the fact that an object of type \( \top \) can be weakened, although it is not explicitly duplicable.

Lemma 9.1.31. In Definition 9.1.28 if \( x \not\in FV(M) \), \( M[V/x] = M \).

Proof. One proves by structural induction on \( M \) that if \( x \not\in FV(M) \) then \( M[V/x] = M \).

Case \( M \equiv y^C \). Since \( x \not\in FV(M), y \leq x \). Then by definition \( M[V/x] = M \).
9.1. An Indexed Lambda Calculus

Cases $M \equiv \ast^n$ and $M \equiv \epsilon^B$. Again, by definition $M[V/x] = M$.

Case $M \equiv \lambda^n y^C.N$. If $y = x$ then by definition $(\lambda^n y^C.N)[V/x] = \lambda^n y^C.N$ and we are done. If $y \neq x$, the substitution $(\lambda^n y^C.N)[V/x]$ is equal to $\lambda^n y^C.(N[V/x])$. Then by induction hypothesis, $N[V/x] = N$, and thus $M[V/x] = M$.

Cases $M \equiv NP$, $M \equiv \langle N, P \rangle^n$ and $M \equiv \langle \text{let } \ast = N \text{ in } P \rangle$. By induction hypothesis, $N[V/x] = N$ and $P[V/x] = P$. Using the definition of substitution, $M[V/x] = M$.

Case $M \equiv \langle \langle y^C, z^D \rangle^n = N \text{ in } P \rangle$. Using $\alpha$-equivalence one can suppose that $y$ and $z$ are not free variables of $V$. Then, if $y = x$ or $z = x$, $M[V/x] = \langle \langle y^C, z^D \rangle^n = N[V/x] \text{ in } P \rangle$. By induction hypothesis $N[V/x] = N$ and thus $M[V/x] = M$. If $x \neq y$ and $x \neq z$, $M[V/x] = \langle \langle y^C, z^D \rangle^n = N[V/x] \text{ in } P[V/x] \rangle$. By induction hypothesis $N[V/x] = N$ and $P[V/x] = P$, thus making $M[V/x] = M$.

Then Lemma [9.1.31] is valid.

Lemma 9.1.32. In Definition 9.1.28, if $M$ is a core value, so is $M[V/x]$, and if $M$ is an extended value, so is $M[V/x]$.

Proof. By induction on the structure of $M$ we prove that if $M$ is a core value, so is $M[V/x]$. Base cases:

Case $M \equiv y^B$. If $y \neq x$, then $M[V/x] = M$, and it is a core value. If $y = x$, then $M[V/x] = \{V < : B\}$, which is a core value from Lemma [9.1.29].

Cases $M \equiv \ast^n$ and $M \equiv \epsilon^B$. In both cases $M[V/x] = M$, already in the form of a core value.

Case $M \equiv \lambda^n y^C.N$. In this situation $M[V/x]$ is of the form $\lambda \ldots$, which is a core value.

Induction case:

Case $M \equiv \langle V_1, V_2 \rangle^n$, with $V_1, V_2$ core values. We have $M[V/x] = \langle V_1[V/x], V_2[V/x] \rangle^n$. By induction hypothesis, $V_1[V/x]$ and $V_2[V/x]$ are core values, making the term $\langle V_1[V/x], V_2[V/x] \rangle^n = M[V/x]$ into a core value.

This exhausts the possibilities. Thus, by structural induction, $M[V/x]$ is a core value if $M$ is. We do the same for extended values, and prove that $M[V/x]$ is an extended value if $M$ is, by structural induction on $M$. The base cases are the same as the base cases for core values, replacing the word “core” by the word “extended”, and remembering that any core value is an extended value. The induction cases go as follows.

Cases $M \equiv \langle V_1, V_2 \rangle^n$ and (let $\ast = V_1$ in $V_2$). In this case, $V_1$ and $V_2$ are extended values. Write $f(V_1, V_2)$ for $M$. Then we have $M[V/x] = f(V_1[V/x], V_2[V/x])$. By induction hypothesis, $V_1[V/x]$ and $V_2[V/x]$ are extended values, making the term $M[V/x]$ into an extended value.

Case $M \equiv \langle \langle y^C, z^D \rangle^n = N \text{ in } P \rangle$ (respectively $M \equiv \langle \langle y^C, z^D \rangle^n = N \text{ in } P \rangle$). We write $f(V_1, V_2)$ in place of $M$. Modulo $\alpha$-equivalence one can assume that $y$ (resp. $y$ and $z$) is not a free variable of $V$. If $y = x$ (resp. $y = x$ or $z = x$) then $M[V/x] = f(V_1[V/x], V_2)$. If $y \neq x$ (resp. $y \neq x$ and $z \neq x$), $M[V/x] = f(V_1[V/x], V_2[V/x])$. By induction hypothesis, $V_1[V/x]$ (and $V_2[V/x]$ in the second case) are extended values. This means that in both cases, $M[V/x]$ is an extended value.

This exhausts the possible cases, and the proof by induction is done: if $M$ is an extended value, so is $M[V/x]$. □
Lemma 9.1.33 (Substitution Lemma). Given two valid typing judgements \( \Delta, \Gamma_1 \triangleright V : A \) and \( \Delta, \Gamma_2, x : A \triangleright M : B \) such that \( V \) is a core value and such that \( |\Gamma_1| \cap |\Gamma_2| = \emptyset \), the typing judgement \( \Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : A \) is valid.

Proof. The proof is done by induction on the size of \( M \).

Case \( M \equiv yB \). Then the derivation of the typing judgement of \( M \) starts with the rule \((ax_1)\). This implies that \( \Gamma_2 = \Gamma'_2 \) and the typing judgement of \( M \) is \( \Delta, \Gamma'_2, x : A \triangleright yB : B \).

If \( y = x \), then \( A <: B \) and \( M[V/x] = \Delta, \Gamma_1 \triangleright \{V <: B\} \) is well-defined. From Lemma 9.1.26, \( \Delta, \Gamma_1 \triangleright M[V/x] : B \) is well-typed. Using Lemma 9.1.26 \( \Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B \) is valid.

If \( y \neq x \), then \( M[V/x] = yB \). Since \( \Delta, \Gamma_2, x : A \triangleright yB : B \) is valid, \( A \) has to be of the form \( !A' \). Then by Lemma 9.1.15 \( \Gamma_1 \) is of the form \( !\Gamma'_1 \). From Lemma 9.1.10 \( \Delta, !\Gamma'_1, \Gamma'_2 \triangleright yB : B \) is valid.

Case \( M \equiv c^B \) (resp. \( M \equiv *^n \)). A derivation of \( M \) starts with rule \((ax_2)\) (resp. rule \((\top. I)\)). In both cases the context \( \Gamma_2, x : A \) is duplicable: \( \Gamma_2 = \Gamma'_2 \) and \( A = !A' \). The core value \( V \) is then in fact of type \( !A' \). By Lemma 9.1.15 \( \Gamma_1 \) is of the form \( !\Gamma'_1 \) for some context \( \Gamma'_1 \).

In both cases, the substitution \( M[V/x] \) yields back \( M \). Thus, \( \Delta, !\Gamma'_1, \Gamma'_2 \triangleright M[V/x] : B \) is valid by applying the typing rule \((ax_2)\) (resp. \((\top. I)\)).

Case \( M \equiv \lambda^n y^C.N \). The typing judgement for the term \( M \) is of the form \( \Delta, \Gamma_1, x : A \triangleright \lambda^n y^C.N : !^n (C \rightarrow D) \), and any typing derivation starts with either \((\lambda_1)\) or \((\lambda_2)\) depending on the value \( n \). In both cases, the judgement \( \Delta, \Gamma_2, x : A, y : C \triangleright N : D \) is valid.

Note that \( x \) cannot be equal to \( y \) in this situation. Using \( \alpha \)-equivalence, one can furthermore choose \( y \) such that it does not belong to \( \Gamma_1 \). The induction hypothesis can be applied on the judgements \( \Delta, \Gamma_1 \triangleright V : B \) and \( \Delta, \Gamma_2, y : C, x : A \triangleright N : D \), and we get the valid typing judgement

\[
\Delta, \Gamma_1, \Gamma_2, y : C \triangleright N[V/x] : D. \tag{9.1.1}
\]

Note that since \( x \neq y \), the substitution \((\lambda^n y^C.N)[V/x] = \lambda^n y^C.(N[V/x]) \). There are two cases, depending on \( n \).

If \( n = 0 \), applying \((\lambda_1)\) on Equation 9.1.1 gives that \( \Delta, \Gamma_1, \Gamma_2 \triangleright (\lambda^0 y^C.N)[V/x] : C \rightarrow D \) is valid.

If \( n \neq 0 \), then \((\lambda_2)\) is the first rule for the derivation of \( M \), and \( (\Gamma_2, x : A) \) is of the form \( (\Gamma'_2, x : !A') \). Since \( A = !A' \) is duplicable and since \( V \) is a core value, by Lemma 9.1.15 the context \( \Gamma_1 \) is of the form \( !\Gamma'_1 \). The rule \((\lambda_2)\) can then be applied on Equation 9.1.1 to get the valid typing judgement \( \Delta, \Gamma_1, \Gamma_2 \triangleright (\lambda^0 y^C.N)[V/x] : !^n (C \rightarrow D) \).

Cases \( M \equiv N_1 N_2, \langle N_1, N_2 \rangle^n, (let \ast = N_1 in N_2) \) and (let \( y^C, z^D \)^n = N_1 in N_2).

A typing derivation of \( \Delta, \Gamma_2, x : A \triangleright M : B \) starts with

\[
\Delta', \Gamma_{21} \triangleright N_1 : B_1, \Delta', \Gamma_{22} \triangleright N_2 : B_2 \quad (X),
\]

where \( (\Delta', \Gamma_{21}, \Gamma_{22}) = (\Delta, \Gamma_2, x : A) \), where \( M = f(N_1, N_2) \), where

\[
\begin{array}{llllllll}
\text{if } f(N_1, N_2) \text{ is } & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
N_1 N_2 & \langle N_1, N_2 \rangle^n & (\top. I) & (app) & D & C \rightarrow D & C & C \\
(let \ast = N_1 \text{ in } N_2) & (\top. E) & C \otimes D & D \quad & C & C & C & C \\
let \langle y^C, z^D \rangle^n = N_1 \text{ in } N_2 & E & C \otimes D & C & C & C & C & C \\
\end{array}
\]
for some types $C$, $D$ and $E$, and where $\Lambda$ is empty in the three first cases and equal to $(y : \!^m C, z : \!^m D)$ in the last case.

By $\alpha$-equivalence one can assume that in the last case, the variables $y$ and $z$ in $\Lambda$ does not occur anywhere in the other contexts.

From Lemma 9.1.19 one can assume that $\Gamma_21$ and $\Gamma_22$ do not contain any duplicate variables, implying that $|\Delta| \subseteq |\Delta'|$. Thus there exists some context $\Psi$ such that $\Delta' = (\Psi, \Delta)$. In particular, this means that

$$(\Psi, \Gamma_21, \Gamma_22) = (\Gamma_2, x : A).$$

(9.1.2)

This equation is the main result that will allow us to apply the induction hypothesis, and thus finish the proof.

The variable $x : A$ belongs to either $\Psi$, $\Gamma_21$ or $\Gamma_22$. We study each one of these three possibilities.

$x \in |\Psi|$. In this case, $A = A'$ and $\Psi = (\Psi', x : A')$. The contexts $(\Gamma_2', \Gamma_21)$ and $(\Gamma_2', \Gamma_22, \Lambda)$ becomes respectively $(\Psi', \Delta, \Gamma_21, x : A')$ and $(\Psi', \Delta, \Gamma_22, \Lambda, x : A')$, meaning that the typing judgements $\Psi', \Delta, \Gamma_21, x : A' \triangleright N_1 : B_1$ and $\Psi', \Delta, \Gamma_22, \Lambda, x : A' \triangleright N_2 : B_2$ are valid.

Since $\Psi = (\Psi', x : A')$, Equation (9.1.2) implies that $(\Psi', \Gamma_21, \Gamma_22) = \Gamma_2$. Since $|\Gamma_1| \cap |\Gamma_2| = \emptyset$, the set $|\Gamma_1|$ does not intersect $|\Psi'|$, $|\Gamma_21|$ nor $|\Gamma_22|$. This means that the typing context $(\Psi', \Delta, \Gamma_21, \Gamma_22, \Gamma_1)$ is valid. Together with the fact that $\Lambda$ does not intersect anything, the requirements on the contexts in the following are satisfied.

By Lemma 9.1.13 the typing judgement $\Psi', \Delta, \Gamma_1 \triangleright V : A'$ is valid.

One can apply the induction hypothesis and get that $\Psi', \Delta, \Gamma_21, \Gamma_1 \triangleright N_1[V/x] : B_1$ and $\Psi', \Delta, \Gamma_22, \Lambda, \Gamma_1 \triangleright N_2[V/x] : B_2$ are valid.

Since $A = A'$, from Lemma 9.1.13 the context $\Gamma_1$ is of the form $\!^m \Gamma_1$.

Using the fact that $f(N_1[V/x], N_2[V/x]) = f(N_1, N_2)[V/x]$ and applying the typing rule $(X)$, one gets that $\Psi', \Delta, \Gamma_21, \Gamma_22, \!^m \Gamma_1' \triangleright f(N_1, N_2)[V/x] : B$ is valid. One can rewrite this typing judgement as $\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B$.

Case $x \in |\Gamma_21|$ and case $x \in |\Gamma_22|$. The two cases are symmetrical (modulo the use of $\Lambda$, which does not bring any difficulty), by inverting the role of 1 and 2 in the indices. We provide the proof for 2.

The context $\Gamma_22$ is of the form $(\Gamma_2', x : A)$. This means that $\Psi, \Delta, \Gamma_22, \Lambda, x : A \triangleright N_2 : B_2$ is valid.

Equation (9.1.2) implies that $(\Psi, \Gamma_21, \Gamma_22) = \Gamma_2$. Since $|\Gamma_1| \cap |\Gamma_2| = \emptyset$, the set $|\Gamma_1|$ does not intersect $|\Psi'|$, $|\Gamma_21|$ nor $|\Gamma_22|$. This means that the typing context $(\Psi', \Delta, \Gamma_21, \Gamma_22, \Gamma_1)$ is valid. Together with the fact that $\Lambda$ does not intersect anything, the requirements on the contexts in the following are satisfied.

By Lemma 9.1.13 the typing judgement $\Psi, \Delta, \Gamma_1 \triangleright V : A$ is valid.

One can apply the induction hypothesis: the judgement $\Psi, \Delta, \Gamma_22, \Lambda, \Gamma_1 \triangleright N_2[V/x] : B_2$ is valid.

Since $(\Psi, \Delta) = \Delta'$, the typing judgement $\Psi, \Delta, \Gamma_21 \triangleright N_1 : B_1$ is valid.

Since $x \in |\Gamma_22|$, $x$ is not a free variable of $N_1$. From Lemma 9.1.32 this means that $f(N_1, N_2)[V/x] = f(N_1, N_2)[V/x]$. (In the case $x \in |\Gamma_21|$, the substitution $f(N_1, N_2)[V/x]$ is equal to $f(N_1[V/x], N_2)$.)

One can then apply the typing rule $(X)$ and get that $\Psi', \Delta, \Gamma_21, \Gamma_22, \Gamma_1 \triangleright f(N_1, N_2)[V/x] : B$ is valid.

By definition, the context of this typing judgement is $(\Delta, \Gamma_1, \Gamma_2)$. This makes $\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B$ a valid typing judgement.
Hence, by structural induction on the typing derivation of $!\Delta, \Gamma_2, x : A \triangleright M : B$, Lemma 9.1.34 is valid.

**Lemma 9.1.34.** Suppose that $!\Delta, \Gamma_1 \triangleright V : A$ and $!\Delta, \Gamma_2, x : A \triangleright M : B$ such that $V$ is a core value and such that $|\Gamma_1| \cap |\Gamma_2| = \emptyset$. Suppose moreover that $B < : B'$. Then $\{M[V/x] : B'\} = \{M < : B'\}[V/x]$. 

**Proof.** By induction on the size of $M$.

*Case $M \equiv x^B$.* In this case, the typing judgement $!\Delta, \Gamma_2, x : A \triangleright M : B$ comes from the rule $(ax_1)$, meaning that $A < : B$.

By definition, $\{M[V/x] : B'\} = \{\{V < : B\} : B'\}$ and $\{M < : B'\}[V/x] = \{V < : B'\}$. They are equivalent from Lemma 9.1.22.

*Case $M \equiv c^B$, $s^n$ and $y^B$, with $y \neq x$.*** In these three cases, $M[V/x] = M$ and $\{M < : B'\}[V/x] = \{M < : B'\}$, making the result true by vacuity.

*Case $M \equiv \lambda^n y^C . N$.*** By $\alpha$-equivalence one can assume that $y \neq x$. Since $M$ is well-typed of type $B$, there exists a type $D$ such that $B = !^n (C \rightarrow D)$. Therefore, $B' = !^n (C' \rightarrow D')$, with $D < : D'$, $C' < : C$ and $m = 0$ if $n = 0$. One has

$$\{M[V/x] : B'\} = \{\{\lambda^n y^C . N[V/x]\} < : !^m (C' \rightarrow D')\} = \{\lambda^m y^C . \{N[V/x] \} < : D'\}$$

which is by induction hypothesis $\lambda^m y^C . \{N < : D'\}[V/x]$. Therefore

$$\{M[V/x] : B'\} = \lambda^m y^C . \{N < : D'\}[V/x] = \{\lambda^m y^C . \{N < : D'\}\}[V/x]$$

$$= \{\lambda^m y^C . \{N < : !^m (C' \rightarrow D')\}\}[V/x] = \{M < : B'\}[V/x].$$

*Case $M \equiv NP$.*** Since $M$ is of type $B$, there exists a type $C$ such that $N : C \rightarrow B$ and such that $P : C$. We have therefore:

$$\{M[V/x] : B'\} = \{\{N[V/x]\} \{P[V/x]\} < : B'\}$$

$$= \{N[V/x]\} \{P[V/x] < : C \rightarrow B'\}$$

$$= \{N[V/x]\} \{\{P < : C \rightarrow B'\}[V/x]\}$$

$$= \{N\{P < : C \rightarrow B'\}\}[V/x]$$

$$= \{NP < : B'\}[V/x]$$

$$= \{M < : B'\}[V/x].$$

*Case $M \equiv (let \ (y^C, z^D)^n = N in P)$.*** By $\alpha$-equivalence one can assume that $x \neq y, z$. Since $M$ is of type $B$, in $P$. We have therefore:

$$\{M[V/x] : B'\} = \{let \ (y^C, z^D)^n = N[V/x] in P[V/x] < : B'\}$$

$$= let \ (y^C, z^D)^n = N[V/x] in \{P[V/x] < : B'\}$$

$$= let \ (y^C, z^D)^n = N[V/x] in \{P < : B'\}[V/x]$$

(by ind. hyp.)

$$= \{let \ (y^C, z^D)^n = N in \{P < : B'\}\}[V/x]$$

$$= \{let \ (y^C, z^D)^n = N in P < : B'\}[V/x]$$

$$= \{M < : B'\}[V/x].$$

The remaining cases are similar. □
9.2 Equational Logic of Typed Terms

Definition 9.2.1. We define an equivalence relation on (indexed) typing judgements. We write \( \Delta \vdash M \approx_{\text{ax}} M' : A \), or simply \( M \approx_{\text{ax}} M' \), to indicate that \( \Delta \vdash M : A \) and \( \Delta \vdash M' : A \) are equivalent. Axiomatic equivalence is defined as the reflexive, symmetric, transitive, and congruence closure of

- the rules from Tables 9.2, Table 9.3, so long as both sides of the equivalences are well-typed.
- the following two rules:
  
  Rule (app\_<\_). Suppose that \( A \prec B \). If \( \Delta = (\Delta', \Gamma_1, \Gamma_2) \) such that
  
  \[
  \Delta', \Gamma_1 \triangleright N : B \Rightarrow C, \quad \Delta', \Gamma_2 \triangleright P : A,
  \]
  
  then \( \Delta \triangleright \{ P \prec B \} \approx_{\text{ax}} \{ N \prec A \Rightarrow C \} P : C \).
  
  Rule (let\_{\_\_<\_}). Suppose that \( A \prec A', B \prec B' \), and \( n' = 0 \) if \( n = 0 \).
  
  If \( \Delta = (\Delta', \Gamma_1, \Gamma_2) \) such that
  
  \[
  \Delta', \Gamma_1 \triangleright N : !^n (A \otimes B), \quad \Delta', \Gamma_2, x : !^{n'} A', y : !^{n'} B' \triangleright P : E,
  \]
  
  then
  
  \[
  \Delta \triangleright \left( \text{let}\langle x^A, y^{B'} \rangle^{n'} = \{ N \prec : !^{n'} (A' \otimes B') \} \text{ in } P \right) \approx_{\text{ax}} \left( \text{let}\langle x^A, y^{B'} \rangle^n = N \text{ in } P \right) : E.
  \]

Lemma 9.2.2. Suppose that \( A \prec A' \). If \( \Delta = (\hat{\Delta}, \Gamma_1, \Gamma_2) \) such that

\[
\hat{\Delta}, \Gamma_1 \triangleright N : A, \quad \hat{\Delta}, \Gamma_2, x : A' \triangleright P : E,
\]

then

\[
\Delta \triangleright \left( \text{let}\ x^A = \{ N \prec : A' \} \text{ in } P \right) \approx_{\text{ax}} \left( \text{let}\ x^A = N \text{ in } P \right) : E.
\]

Proof. By definition. \( \square \)

Convention 9.2.3. The rule stated in Lemma 9.2.2 will be referred to using the symbol (let\_<\_<\_). \( \square \)

Lemma 9.2.4. The equivalences of Table 9.4 are derivable.

Proof. By case distinction. \( \square \)

Lemma 9.2.5. Suppose that \( \Delta \vdash M \approx_{\text{ax}} M' : A \) and that \( \Delta' \prec \Delta \) and \( A \prec A' \). Then \( \Delta' \vdash \{ M \prec : A' \} \approx_{\text{ax}} \{ M' \prec : A' \} : A' \).

Proof. We have \( \Delta \vdash M : A \) and \( \Delta \vdash M' : A \). From Lemma 9.1.14 \( \Delta' \vdash M : A \) and \( \Delta' \vdash M' : A \) are valid. From Lemma 9.1.26 \( \Delta' \vdash \{ M \prec : A' \} : A' \) and \( \Delta' \vdash \{ M' \prec : A' \} : A' \) are valid. It remains to prove that they are axiomatically equivalent. We prove it by induction on the size of the derivation of \( \Delta \vdash M \approx_{\text{ax}} M' : A \).

Base case: \( \beta\eta \)-rules.
(\beta_\lambda) \Delta \triangleright let x = V in M 
\approx_{\text{ax}} M[V/x] : A

(\beta_\eta) \Delta \triangleright let (x, y)^n = (V, W)^n in M 
\approx_{\text{ax}} M[V/W, W/Y] : A

(\beta_\delta) \Delta \triangleright let * = * in M 
\approx_{\text{ax}} M : A

(\eta_\lambda) \Delta \triangleright \lambda^n x^A \cdot \{V < A \mapsto B\} x^A 
\approx_{\text{ax}} V : \text{!}n(A \mapsto B).

(\eta_\delta) \Delta \triangleright let (x^A, y^B)^n = N in (x^n, y^n B)^n 
\approx_{\text{ax}} N : \text{!}n(A \otimes B).

(\eta_\delta^\prime) \Delta \triangleright let * = \{N < T\} in *^n 
\approx_{\text{ax}} N : \text{!}n T.

Table 9.2: Axiomatic equivalence axioms: beta-eta-rules

| (let_1) | \Delta \triangleright let \textbf{[]} = 
\begin{align*}
\text{(let \textbf{[]} = M in N) in P} & \\
\text{\approx_{ax} let \textbf{[]} = M in P} & \\
\text{\approx_{ax} let \textbf{[]} = N in P} & : A
\end{align*}

| (let_2) | \Delta \triangleright let \textbf{[]} = V in 
\begin{align*}
\text{let \textbf{[]} = W in M} & \\
\text{\approx_{ax} let \textbf{[]} = W in M} & : A
\end{align*}

| (let_app) | \Delta \triangleright let x^{A\otimes B} = M in 
\begin{align*}
\text{let y^A = N in xy} & \\
\text{\approx_{ax} MN} & : B
\end{align*}

| (let^\prime) | \Delta \triangleright let x^D = V in \lambda^n y^A.M 
\begin{align*}
\text{let y^A = N in x^D} & \\
\text{\approx_{ax} \lambda^n y^A.let x^D = V in M: !n(A \otimes B)} & : !n(A \otimes B)
\end{align*}

| (let^0) | \Delta \triangleright let x^{n\cdot A} = M in 
\begin{align*}
\text{let y^n \cdot B = N in (x^n\cdot A, y^n B)^n} & \\
\text{\approx_{ax} (M, N)^n} & : !n(A \otimes B)
\end{align*}

Table 9.3: Axiomatic equivalence axioms: commutation rules

Case (\beta_\lambda). In this case, M = (let x^B = V in N) and M’ = N[V/x]. From Lemma \[9.1.21\] \{M < A’\} = (let x^B = V in (N < A’)). From Lemma \[9.1.30\] (M’ < A’) = (N < A’)[V/x]. Using equivalence rule (\beta_\lambda), \Delta’ \triangleright let x^B = V in \{N < A’\}[V/x] : A’.

Case (\beta_\eta). We have M = (let (x^B, y^C)^n = (V, W)^n in N) and M’ = N[V/x, W/Y]. From Definition \[9.1.21\] \{M < A’\} = (let (x^B, y^C)^n = (V, W)^n in (N < A’)). From Lemma \[9.1.30\] (M’ < A’) = (N < A’)[V/x, W/Y]. Using equivalence rule (\beta_\eta), \Delta’ \triangleright let (x^B, y^C)^n = (V, W)^n in \{N < A’\}[V/x, W/Y] : A’.

Case (\beta_\delta). In this case, the term * is let \* = * in N and M’ = N. From Definition \[9.1.20\] (M < A’) = (let * = * in \{N < A’\}). Using equivalence rule (\beta_\delta), we obtain the axiomatic equivalence \Delta’ \triangleright let * = * in \{N < A’\} = \{N < A’\}.

Case (\eta_\lambda). In this case, M = \lambda^n x^B \cdot \{V < B \mapsto C\} x^B, M’ = V, A = !n(B \mapsto C) and A’ = !n(B’ \mapsto C’) with m = 0 if n = 0, C < C’ and B’ < B.

From Definition \[9.1.20\] (M < A’) = \lambda^n x^B \cdot \{V < B \mapsto C\} x^B < C’ = \lambda^n x^B \cdot \{V < B \mapsto C\} x^B. From Lemma \[9.1.21\] \{M < A’\} = \lambda^n x^B \cdot \{V < B \mapsto C\} x^B.

We have \Delta’ \triangleright V : !n(B \mapsto C). Since x \notin \Delta and since x^B = \{x^B < B\}, by equivalence rule (app_x), and by Lemma \[9.1.22\] we have \Delta’ : x : B’ \triangleright \{V < B \mapsto C’\} x^B \approx_{ax} \{V < B’ \mapsto C’\} x^B : C’. Thus by congruence \Delta’ \triangleright \{M < A’\} = !n x^B \cdot \{V < B’ \mapsto C’\} x^B : !n(B’ \mapsto C’).

From Lemma \[9.1.22\] \{V < B’ \mapsto C’\} = \{V < !n(B’ \mapsto C’)\} < B’ \mapsto C’. Using rule (\eta_\delta), by transitivity we get that \Delta’ \triangleright \{M < !n(B’ \mapsto C’)\} = \{V < !n(B’ \mapsto C’)\} : !n(B’ \mapsto C’).

Case (\beta_\delta^\prime). In this case, M = (let x^A = N in x^A) and M’ = N.

From Definition \[9.1.20\] \{M < A’\} = (let x^A = N in x^A). From Lemma \[9.1.21\] \Delta’ \triangleright (let x^A = N in x^A) \approx_{ax} \{N < A’\} in x^A : A’. Applying rule (\beta_\delta^\prime), we get that this is equivalent to \Delta’ \triangleright \{N < A’\} : A’.

Case (\eta_\delta). In this case, M = (let (x^B, y^C)^n = N in (x^n, y^n B)^n) and M’ = N. We also have
9.2. Equational Logic of Typed Terms

\((\alpha_{\text{let}}) \Delta, y : A \triangleright M : B \approx_{\text{ax}} \Delta, x : A \triangleright \text{let } y^A = x^A \text{ in } M : B\)

\((\lambda_{\text{let}}) !\Delta \triangleright \text{let } x^{C_1} = V \in \lambda y.M \approx_{\text{ax}} \lambda^{n+1} y. \text{let } x^{C_1} = V \in M : !^{n+1}(A \rightarrow B)\)

\((\text{let}^{\ominus}) \Delta \triangleright \langle V, \text{let } \varnothing = M \text{ in } N \rangle \approx_{\text{ax}} \text{let } \varnothing = M \text{ in } (V, N) : !^n(A \odot B)\)

\((\text{let}^\ominus_1) \Delta \triangleright \langle \text{let } = M \text{ in } N, V \rangle \approx_{\text{ax}} \text{let } = M \text{ in } (N, V) : !^n(A \odot B)\)

\((\text{let}^\ominus_2) \Delta \triangleright \langle \text{let } \varnothing = M \text{ in } N, V \rangle \approx_{\text{ax}} \text{let } \varnothing = M \text{ in } NV \quad : B\)

\((\text{let}^{\ominus_{\text{app}}}) \Delta \triangleright \langle \text{let } \varnothing = M \text{ in } N \rangle V \approx_{\text{ax}} \text{let } \varnothing = M \text{ in } NV \quad : B\)

Table 9.4: Axiomatic equivalence: derived rules

\(A = !^m(B \odot C)\), and since \(A <: A', A' = !^m(B' \odot C')\) with \(m = 0\) if \(n = 0\), with \(B <: B'\) and with \(C <: C'\).

From Definition 9.1.20 \(\{M <: A'\} = \{\text{let } (x^B, y^C)^n = N \text{ in } (x'^{B'}, y'^{C'})^n\} \). From equivalence rule (\(\text{let}^{\ominus_{\text{app}}}\)),

\(\Delta' \triangleright \{x^B, y^C\}^n = N \text{ in } (x'^{B'}, y'^{C'})^n \approx_{\text{ax}} \text{let } (x^B, y^C)^m = \{N <: !^m(B' \odot C')\} \text{ in } (x'^{B'}, y'^{C'})^n : A'\).

Applying rule (\(\eta_\ominus\)), we get that this is equivalent to \(\Delta' \triangleright \{N <: A'\} : A'\).

Case (\(\eta\)). In this case, \(M = \{\text{let } = \{N <: \top\} \text{ in } *^n\}, M' = N \text{ and } A = !^n \top\). Since \(A <: A', A' = !^m \top\) with \(m = 0\) if \(n = 0\).

From Definition 9.1.20 \(\{N <: \top\} = \{\{N <: !^m \top\} <: \top\}\).

From Lemma 9.1.21 \(\{N <: \top\} \approx_{\text{ax}} \{N <: !^m \top\} \approx_{\text{ax}} \{N <: !^n \top\}\).

And this prove the lemma for the \(\beta\) rules of Table 9.2.

Next base cases: commutation rules.

Case (\(\text{let}^1\)). Here, \(M = \{\text{let } -1 = \{\text{let } -2 = N_2 \text{ in } N_1 \} \text{ in } P\} \) and \(M' = \{\text{let } -2 = N_2 \text{ in } \text{let } -1 = N_1 \text{ in } P\} \).

Using Definition 9.1.20 \(\{M <: A'\} = \{\text{let } -1 = \{\text{let } -2 = N_2 \text{ in } N_1 \} \text{ in } \{P <: A'\}\} \) and \(\{M' <: A'\} = \{\text{let } -2 = N_2 \text{ in } \text{let } -1 = N_1 \text{ in } \{P <: A'\}\} \).

Using equivalence rule (\(\text{let}^1\)), we find that \(\Delta' \triangleright \{M <: A'\} \approx_{\text{ax}} \{M' <: A'\} : A'\).

Case (\(\text{let}^2\)). This case is similar to the previous one. Here, \(M = \{\text{let } -1 = V \text{ in } \text{let } -2 = W \text{ in } N\} \) and \(M' = \{\text{let } -2 = W \text{ in } \text{let } -1 = V \text{ in } N\} \).

Using Definition 9.1.20 \(\{M <: A'\} = \{\text{let } -1 = V \text{ in } \text{let } -2 = W \text{ in } \{N <: A'\}\} \) and \(\{M' <: A'\} = \{\text{let } -2 = W \text{ in } \text{let } -1 = V \text{ in } \{N <: A'\}\} \).

Using equivalence rule (\(\text{let}^2\)), we find that \(\Delta' \triangleright \{M <: A'\} \approx_{\text{ax}} \{M' <: A'\} : A'\).

Case (\(\text{let}^{\ominus_{\text{app}}}\)). Here, \(M = \{\text{let } x^{B \ominus A} = N \text{ in } \text{let } y^B = P \text{ in } (x^{B \ominus A} y^B)\} \) and \(M' = NP\).

Using Definition 9.1.20 \(\{M <: A'\} = \{\text{let } x^{B \ominus A} = N \text{ in } \text{let } y^B = P \text{ in } (x^{B \ominus A} y^B)\} \).

From Lemma 9.1.22 \(\Delta' \triangleright \text{let } x^{B \ominus A} = N \text{ in } \text{let } y^B = P \text{ in } (x^{B \ominus A} y^B) \approx_{\text{ax}} \text{let } x^{B \ominus A'} = \{N <: B \ominus A'\} \text{ in } y^B = P \text{ in } (x^{B \ominus A'} y^B) \approx_{\text{ax}} \{N <: B \ominus A'\} \times A' \).
Rule (let^app) says that this is equivalent to $\Delta' \triangleright \{ N <; B \rightarrow A' \} : A'$. The result is obtained by using Definition 9.1.20: $\{ M' : A' \} = (N <; B \rightarrow A') P$.

**Case (let^app).** In this case, $M = (let \ x^D = V \ in \ \lambda^n y^B.N)$, $M' = \lambda^n y^B.(let \ x^D = V \ in \ N)$, $A = !^n (B \rightarrow C)$ and $A' = !^m (B' \rightarrow C')$ with $m = 0$ if $n = 0$, $C <; C'$ and $B' <; B$.

Using Definition 9.1.20: $\{ M : A' \} = (let \ x^D = V \ in \ \lambda^n y^B.N)$ and $\{ M' : A' \} = \lambda^n y^B.(let \ x^D = V \ in \ N <; C')$.

Using equivalence rule (let^app), we have as required that $\Delta' \triangleright \{ M : A' \} \approx_{ax} \{ M' : A' \} : A'$.

**Case (let^ax).** In this situation, $M = (let \ x^m B = N \ in \ let \ y^n C = P \ in (x^m B', y^n C')^m)$, $M' = (N, P)^n$, $A = !^n (B \rightarrow C)$ and $A' = !^m (B' \rightarrow C')$ with $m = 0$ if $n = 0$, $C <; C'$ and $B <; B'$.

Using Definition 9.1.20: $\{ M : A' \} = (let \ x^m B = N \ in \ let \ y^n C = P \ in (x^m B', y^n C')^m)$ and $\{ M' : A' \} = \{ (N <; !^m C'), (P <; !^n D)^m \}$.

Using twice Lemma 9.1.22 we get that $\Delta' \triangleright \{ M : A' \} \approx_{ax} \{ let \ x^m B' = N <; !^m B' \} = \{ P <; !^m C' \} : A'$.

Applying (let^ax) and transitivity gives the desired result.

**Rules concerning typecasting.** For these two rules, we use in each case the notation of Definition 9.1.21. In both cases, we have $\Delta \triangleright M \approx_{ax} M' : C$. We want to show that if $\Delta' \triangleright \Delta$ and $C <; C'$ then $\Delta' \triangleright \{ M <; C' \} \approx_{ax} \{ M' <; C' \} : C'$.

The fact that $\Delta' \triangleright \{ M <; C' \} : C'$ and that $\Delta' \triangleright \{ M' <; C' \} : C'$ is proved using Lemma 9.1.13 and 9.1.20.

The fact that they are axiomatically equivalent is proved as follows.

**Case (app^c).** We have $N = N.P^B$ and $M' = \{ N <; A \rightarrow C \} P$.

Consider $\{ M : C' \}$. From Definition 9.1.20 it is equal to $\{ N <; B \rightarrow C' \} \{ P <; B \}$.

Using rule (app^c), we infer that $\Delta \triangleright \{ M : C' \} \approx_{ax} \{ (N <; B \rightarrow C') P <; B \} : C'$. Using Lemma 9.1.22 we deduce that $\Delta' \triangleright \{ N <; A \rightarrow C' \} P : C'$. We conclude by noticing that from Definition 9.1.20 $\{ M' <; C' \} = \{ N <; A \rightarrow C' \} P$.

**Case (let^ax).** Here, $M = (let \ x^A \ y^B \ n^m = \{ N <; !^m (A \rightarrow B') \} \ in \ P)$ and $M' = (let \ x^A \ y^B \ n^m = N \ in \ P)$.

The term $\{ M : C' \}$ is equal to let $\{ x^A, y^B \}^m = \{ N <; !^m (A \rightarrow B') \}$ in $P <; C'$ by Definition 9.1.20. The term $\{ M' : C' \}$ is equal from the same definition to the term let $\{ x^A, y^B \}^m = N \ in \ P <; C'$.

They are axiomatically equivalent because of the equivalence rule (let^ax).

And this ends the base cases.

**Rules specific to equivalence relations.** The reflexivity, the symmetry, the transitivity are trivial.

**Reflexivity.** In this case, $M = M'$. Therefore $\{ M <; A' \} = \{ M' <; A' \}$. One has that $\Delta' \triangleright \{ M <; A' \} \approx_{ax} \{ M' <; A' \} : A'$ by reflexivity.

**Symmetry.** Here, $\Delta \triangleright M \approx_{ax} M' : A'$ was proved by first proving $\Delta \triangleright M' \approx_{ax} M : A'$. By induction hypothesis, we conclude that $\Delta' \triangleright \{ M' <; A' \} \approx_{ax} \{ M <; A' \} : A'$. We then apply again symmetry.

**Transitivity.** Here, $\Delta \triangleright M \approx_{ax} M' : A'$ was proved introducing a third term $M''$ such that $\Delta \triangleright M \approx_{ax} M'' : A$ and $\Delta \triangleright M'' \approx_{ax} M' : A$. By induction hypothesis, this means that $\Delta' \triangleright \{ M <; A' \} \approx_{ax} \{ M'' <; A' \} : A'$ and $\Delta \triangleright \{ M'' <; A' \} \approx_{ax} \{ M' <; A' \} : A'$. We can then apply transitivity to get the result.
9.2. Equational Logic of Typed Terms

Congruence rules. There is one congruence rule per term constructor. We list the cases in the following.

Case $M \equiv \lambda^n x^C.N$. Then $M' \equiv \lambda^n x^C.N'$, the type $A$ is of the form $!^n(C \rightarrow D)$ and the type $A'$ is of the form $!^n(C \rightarrow D)$ with $D <: D'$. We have $\Delta, x : C \vdash N \approx_{ax} N' : D$, and $\Delta$ is of the form $!^\Delta$ if $n \neq 0$.

By induction hypothesis, $\Delta', x : C \vdash \{N <: D'\} \approx_{ax} \{N' <: D'\} : D'$. It is then possible to apply the congruence rule for $(\lambda)$ to get the result.

Cases $M \equiv N_1N_2$, $(N_1, N_2)^n$, (let $* = N_1$ in $N_2$) and (let $(y^C, z^D)^n = N_1$ in $N_2$).

In each case we use the shortcut notation $f(N_1, N_2)$ in place of the corresponding construct.

Then $M' \equiv f(N'_1, N'_2)$ for some terms $N'_1$ and $N'_2$, and the typing judgements $\Delta \vdash M, M' : A$ comes from

\[
\begin{array}{ll}
!^\Delta, \Gamma_1 \vdash N_1 : B_1, & !^\Delta, \Gamma_2, \Lambda \vdash N_2 : B_2, \\
!^\Delta, \Gamma_1 \vdash N'_1 : B_1, & !^\Delta, \Gamma_2, \Lambda \vdash N'_2 : B_2,
\end{array}
\]

and from the application of a typing rule $(X)$, where $\Delta = (!^\Delta, \Gamma_1, \Gamma_2)$, and where

| $N_1 N_2$ | $(\text{app})$ | $D$ | $C \rightarrow D$ | $C$ \\
| $(N_1, N_2)^n$ | $(\otimes, I)$ | $!^n(C \otimes D)$ | $!^n C$ | $!^n D$ \\
| let $* = N_1$ in $N_2$ | $(\top, E)$ | $D$ | $\top$ | $D$ \\
| let $(y^C, z^D)^n = N_1$ in $N_2$ | $(\otimes, E)$ | $E$ | $!^n(C \otimes D)$ | $E$

for some types $C$, $D$ and $E$, and where $\Lambda$ is empty in the three first cases and equal to $(y : !^nC, z : !^nD)$ in the last case. By $\alpha$-equivalence one can assume that in the last case, the variables $y$ and $z$ in $\Lambda$ does not occur anywhere in the other contexts.

The type $A'$ is such that $A <: A'$. Its structure is then governed by the one of $A$. Let us construct $B'_1$ and $B'_2$ as follows:

| $N_1 N_2$ | $D'$ | $C \rightarrow D'$ | $C$ \\
| $(N_1, N_2)^n$ | $!^m(C' \otimes Dm)$ | $!^m C'$ | $!^m D'$ \\
| let $* = N_1$ in $N_2$ | $D'$ | $\top$ | $D'$ \\
| let $(y^C, z^D)^n = N_1$ in $N_2$ | $E'$ | $!^n(C \otimes D)$ | $E'$

for some types $C'$, $D'$ and $E'$ such that $C <: C'$, $D <: D'$ and $E <: E'$.

Note that in each case, $B_1 <: B'_1$ and $B_2 <: B'_2$. One can therefore apply the induction hypothesis, and get that

\[
\begin{array}{ll}
!^\Delta', \Gamma'_1 \vdash \{N_1 <: B'_1\} : B'_1, & !^\Delta', \Gamma'_2, \Lambda \vdash \{N_2 <: B'_2\} : B'_2, \\
!^\Delta', \Gamma'_1 \vdash \{N'_1 <: B'_1\} : B'_1, & !^\Delta', \Gamma'_2, \Lambda \vdash \{N'_2 <: B'_2\} : B'_2,
\end{array}
\]

are valid, where $\Delta' = (!^\Delta', \Gamma'_1, \Gamma'_2)$, where $!^\Delta' <: !^\Delta$, $\Gamma'_1 <: \Gamma_1$ and $\Gamma'_2 <: \Gamma_2$.

One can apply the congruence rule for $(X)$ and get that

$\Delta' \vdash \{N_1 <: B'_1\} \{N_2 <: B'_2\} \approx_{ax} \{N'_1 <: B'_1\} \{N'_2 <: B'_2\}' : A'$.

Using Definition 9.1.20 and Lemma 9.1.26 this means that $\Delta' \vdash \{M <: A'\} \approx_{ax} \{M' <: A'\} : A'$.

and this closes the list of possible cases. \qed
9.3. The Category $C_\lambda$

**Lemma 9.2.6.** If $\Delta, \Gamma_1, x : A \triangleright M : B$ is a valid typing judgement, and if $\Delta, \Gamma_2 \triangleright V \approx_{ax} V' : A$ are axiomatically equivalent core values such that $|\Gamma_1| \cap |\Gamma_2| = \emptyset$, then $\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] \approx_{ax} M[V'/x] : B$ are axiomatically equivalent typing judgements.

**Proof.** First note that $\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B$ and $\Delta, \Gamma_1, \Gamma_2 \triangleright M[V'/x] : B$ are valid typing judgements from Lemma 9.1.33. We prove that they are axiomatically equivalent by induction on the size of $M$.

**Cases $M \equiv y^C, c^C$ and $^n$.** In these cases, either $M[V/x] = M$ and $M[V'/x] = M$ or $M[V/x] = \{V : A <: C\}$ and $M[V'/x] = \{V' : A <: C\}$. In the former case we have axiomatic equivalence by reflexivity of the relation and in the latter case we have it using Lemma 9.2.5.

**Case $M \equiv \lambda^n y^C . N$.** Note that since $\Delta, \Gamma_1, x : A \triangleright M : B$ is well-typed, we have $x \neq y$ and $B = \lambda^n (C \rightarrow D)$ for some type $D$. In particular, $M[V/x] = \lambda^n y^C . (N[V/x])$ and $M[V'/x] = \lambda^n y^C . (N[V'/x])$.

The typing tree of the typing judgement starts with typing rule (\(\lambda_i\)) and has for hypothesis $\Delta, \Gamma_1, x : A, y : C \triangleright N : D$. The integer $i = 1$ if $n = 0$, otherwise $i = 2$. In the latter, $(\Gamma_1, x : A)$ is of the form $(!\Gamma_1, x : !A)$.

By induction hypothesis, $\Delta, \Gamma_1, y : C \triangleright N[V/x] \approx_{ax} N[V'/x] : D$. Applying congruence, we get the result.

The remaining cases are done similarly.

The following result stipulates that all the indexations of a given erasure live in the same axiomatic class. In other words, the axiomatic equivalence class of a term is independent of its indexation.

**Theorem 9.2.7.** If $\text{Erase}(M) = \text{Erase}(M')$ and if $\Delta \triangleright M, M' : A$ are valid typing judgements, then $M \approx_{ax} M'$.

The proof of Theorem 9.2.7 requires some machinery that we develop in Chapter 10.

9.3 The Category $C_\lambda$

The language of Section 9.1 together with its axiomatic equivalence can be made into a syntactic category.

**Lemma 9.3.1.** We can define a category as follows: Objects are types, and arrows $A \rightarrow B$ are axiomatic classes of valid typing judgements of the form $x : A \triangleright V : B$, where $V$ is an extended value. We define the composition of arrows $x : A \triangleright V : B$ and $y : B \triangleright W : C$ to be $x : A \triangleright let y = V in W : C$. The identity on $A$ is set to be the arrow $x : A \triangleright x : A$.

**Proof.** The composition of two arrows yields an arrow axiomatically equivalent to a value due to Axiom (\(\beta_\lambda\)) and Lemmas 9.1.32 and 9.1.33. Composition is associative due to Axiom (let1). The arrow $x : A \triangleright x : A$ is indeed the identity on $A$ due to axioms (\(\alpha_{let}\)) and (\(\beta_\lambda\)).

**Definition 9.3.2.** The category described in Lemma 9.3.1 will be referred as $C_\lambda$. 
9.3.1 Monoidal Structure

Lemma 9.3.3. If we equip $\mathcal{C}_\lambda$ with the arrows

\[
\begin{align*}
\alpha_{A,B,C} &= x : A \otimes (B \otimes C) \triangleright \text{let } (y, z) = x \text{ in } (t, u) = z \text{ in } (y, t, u) : (A \otimes B) \otimes C, \\
\lambda_A &= x : \top \otimes A \triangleright \text{let } (y, z) = x \text{ in } * = y \text{ in } z : A, \\
\rho_A &= x : A \otimes \top \triangleright \text{let } (y, z) = x \text{ in } * = z \text{ in } y : A, \\
\sigma_{A,B} &= x : A \otimes B \triangleright \text{let } (y, z) = x \text{ in } (z, y) : B \otimes A,
\end{align*}
\]

and with the arrows $(x : A \triangleright V : B) \otimes (y : C \triangleright W : D) = z : A \otimes B \triangleright \text{let } (x, y) = z \text{ in } (V, W) : C \otimes D$, then $(\mathcal{C}_\lambda, \otimes, \top)$ is a symmetric monoidal category.

Proof. We need to show that $\alpha$, $\lambda$, $\rho$ and $\sigma$ are isomorphic natural transformations, that they satisfy Equations $\triangleright (\triangleright 1)\triangleright$, $\triangleright (\triangleright 2)\triangleright$, $\triangleright (\triangleright 3)\triangleright$, $\triangleright (\triangleright 4)\triangleright$, $\triangleright (\triangleright 5)\triangleright$ and $\triangleright (\triangleright 6)\triangleright$, and that the constructor $\otimes$ is a bifunctor.

Isomorphisms. Let’s show that $\alpha$, $\lambda$, $\rho$ and $\sigma$ are isomorphisms. It is sufficient to exhibit their inverse:

\[
\begin{align*}
\alpha^{-1}_{A,B,C} &= x' : (A \otimes B) \otimes C \triangleright \text{let } (y', z') = x' \text{ in } (t', u') = y' \text{ in } (t', (u', z')) : A \otimes (B \otimes C), \\
\lambda^{-1}_A &= x' : A \otimes (\ast \cdot x') : \top \otimes A, \\
\rho_A &= x' : A \otimes (\ast \cdot x') : A \otimes \top, \\
\sigma_{A,B} &= x' : B \otimes A \triangleright \text{let } (y', z') = x' \text{ in } (z', y') : A \otimes B.
\end{align*}
\]

Naturality of $\alpha$, $\lambda$, $\rho$ and $\sigma$. Let’s consider $\lambda$. Let $f = a : A \triangleright V : A'$ be any arrow $A \rightarrow A'$. We need to show that the diagram

\[
\begin{array}{ccc}
\top \otimes A & \xrightarrow{\lambda_A} & A \\
\downarrow f & & \downarrow f \\
\top \otimes A & \xrightarrow{\lambda_{A'}} & A'
\end{array}
\]

commutes, i.e. $\lambda_A : f \equiv (\top \otimes f) ; \lambda_{A'}$. The arrow $\top \otimes f$ is $b : \top \otimes A \triangleright \text{let } (t, a) = b \text{ in } (t, V) : \top \otimes A'$. So the left hand side of the equation is

\[
\begin{align*}
b : \top \otimes A \triangleright \text{let } x = \text{let } (t, a) = b \text{ in } (t, V) \otimes x = \text{let } * = y \text{ in } z \\
\approx_{ax} \text{let } (t, a) = b \text{ in } \text{let } x = (t, V) \otimes \text{let } * = y \text{ in } z \\
\approx_{ax} \text{let } (t, a) = b \text{ in } \text{let } y, z = (t, V) \otimes \text{let } * = y \text{ in } z \quad \text{by (let)} \\
\approx_{ax} \text{let } (t, a) = b \text{ in } \text{let } y, z = (t, V) \otimes \text{let } * = y \text{ in } z \quad \text{by } (\beta_\otimes).
\end{align*}
\]

The right hand side of the equation is

\[
\begin{align*}
x : \top \otimes A \triangleright \text{let } a = \text{let } (y, z) = x \text{ in } \text{let } * = y \text{ in } z \otimes \text{let } a = z \text{ in } V : A' \\
\approx_{ax} \text{let } (y, z) = x \text{ in } \text{let } * = y \text{ in } \text{let } a = z \text{ in } V : A' \quad \text{by (let)} \\
\approx_{ax} \text{let } (y, a) = x \text{ in } \text{let } * = y \text{ in } V : A' \quad \text{by } (\alpha_{let}).
\end{align*}
\]
Both sides are then alpha equivalent, which proves that $\lambda$ is a natural transformation.

The proof of naturality of $\rho$ is exactly similar.

For $\sigma_{A,B}$, consider two arrows $f = a : A \triangleright V : A'$ and $g = b : B \triangleright W : B'$. We aim to show that

$$\begin{align*}
A \otimes B & \xrightarrow{\sigma_{A,B}} B \otimes A \\
\alpha & \downarrow \quad \downarrow \alpha' \\
A' \otimes B' & \xrightarrow{\sigma'_{A',B'}} B' \otimes A'
\end{align*}$$

commutes, i.e. $\sigma_{A,B}((g \otimes f)) = ((f \otimes g)) \sigma'_{A',B'}$. The left hand side is the arrow:

$$x : A \otimes B \triangleright let \ c = (let \ (y, z) = x \ in \ (z, y)) \ in \\
\approx_{ax} let \ (y, z) = x \ in \ let \ c = (z, y) \ in \\
\approx_{ax} let \ (a, b) = x \ in \ let \ c = (V, W) \ in \\
\approx_{ax} let \ (a, b) = x \ in \ let \ c = (W, V) \ in$$

The right hand side is

$$x : A \otimes B \triangleright let \ c = (let \ (a, b) = x \ in \ (V, W)) \ in \\
\approx_{ax} let \ (y, z) = x \ in \ let \ c = (V, W) \ in \\
\approx_{ax} let \ (a, b) = x \ in \ let \ c = (W, V) \ in$$

which is axiomatically equivalent to the left hand side.

To prove the naturality of $\alpha$, let’s choose three functions $f = a : A \triangleright V : A'$, $g = b : B \triangleright W : B'$ and $h : c : C \triangleright U : C'$. It is enough to show that

$$\begin{align*}
A \otimes (B \otimes C) & \xrightarrow{\alpha_{A,B,C}} (A \otimes B) \otimes C \\
f \otimes (g \otimes h) & \downarrow \quad \downarrow (f \otimes g) \otimes h \\
A' \otimes (B' \otimes C') & \xrightarrow{\alpha'_{A',B',C'}} (A' \otimes B') \otimes C'
\end{align*}$$

commutes, i.e. that $\alpha_{A,B,C}(\langle f \otimes g \rangle \otimes h) = \langle f \otimes (g \otimes h) \rangle); \alpha'_{A',B',C'}$. Let’s write down the components of this equation:

$$\begin{align*}
f \otimes g & = \quad d : A \otimes B \triangleright let \ (a, b) = d \ in \ (V, W) : A' \otimes B' \\
g \otimes h & = \quad i : B \otimes C \triangleright let \ (b, c) = i \ in \ (W, U) : B' \otimes C'
\end{align*}$$

Then

$$(f \otimes g) \otimes h \quad c : (A \otimes B) \otimes C \triangleright$$
9.3. The Category $C_\lambda$

\[ \text{let } \langle d, e \rangle = e \text{ in } \left\{ \text{let } \langle a, b \rangle = d \text{ in } \langle V, W \rangle, U \right\} \]

\[ \cong_{ax} \text{let } \langle d, e \rangle = e \text{ in } \text{let } \langle a, b \rangle = d \text{ in } \]

\[ \langle \langle V, W \rangle, U \rangle : (A' \otimes B') \otimes C' \text{ \quad by (let}_0^2) \]

\[ f \otimes (g \otimes h) \quad j : A \otimes (B \otimes C) \triangleright \]

\[ \text{let } \langle a, i \rangle = j \text{ in } \langle V, \text{let } \langle b, c \rangle = i \text{ in } \rangle (W, U) \]

\[ \cong_{ax} \text{let } \langle a, i \rangle = j \text{ in } \text{let } \langle b, c \rangle = i \text{ in } \]

\[ \langle V, \langle W, U \rangle \rangle : A' \otimes (B' \otimes C') \text{ \quad by (let}_1^0 \rangle. \]

Finally,

\[ \alpha_{A, B, C} = \quad x : A \otimes (B \otimes C) \triangleright \text{let } \langle x_1, x_2 \rangle = x \text{ in } \]

\[ \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \]

\[ \langle \langle x_1, x_3 \rangle, x_4 \rangle : (A \otimes B) \otimes C \]

\[ \alpha_{A', B', C'} = \quad y : A' \otimes (B' \otimes C') \triangleright \text{let } \langle y_1, y_2 \rangle = y \text{ in } \]

\[ \text{let } \langle y_3, y_4 \rangle = y_2 \text{ in } \]

\[ \langle \langle y_1, y_3 \rangle, y_4 \rangle : (A' \otimes B') \otimes C' \]

Now let’s consider the left hand side of the equation:

\[ x : A \otimes (B \otimes C) \triangleright \text{let } e = \left( \begin{array}{c}
\text{let } \langle x_1, x_2 \rangle = x \text{ in } \\
\text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \\
\langle \langle x_1, x_3 \rangle, x_4 \rangle
\end{array} \right) \text{ in } \]

\[ \text{let } \langle d, e \rangle = e \text{ in } \]

\[ \text{let } \langle a, b \rangle = d \text{ in } \langle \langle V, W \rangle, U \rangle \]

\[ \cong_{ax} \text{let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \]

\[ \text{let } e = \langle \langle x_1, x_3 \rangle, x_4 \rangle \text{ in } \]

\[ \text{let } \langle d, c \rangle = e \text{ in } \]

\[ \text{let } \langle a, b \rangle = d \text{ in } \langle \langle V, W \rangle, U \rangle \text{ \quad by (let}_1 \rangle. \]

\[ \cong_{ax} \text{let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \]

\[ \text{let } \langle d, c \rangle = \langle \langle x_1, x_3 \rangle, x_4 \rangle \text{ in } \]

\[ \text{let } \langle a, b \rangle = d \text{ in } \langle \langle V, W \rangle, U \rangle \text{ \quad by (} \beta_\lambda) \]

\[ \cong_{ax} \text{let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \]

\[ \text{let } \langle a, b \rangle = \langle x_1, x_3 \rangle \text{ in } \langle \langle V, W \rangle, U \rangle[x_4/c] \text{ \quad by (} \beta_0) \]

\[ \cong_{ax} \text{let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \]

\[ \langle \langle V, W \rangle, U \rangle[x_1/a, x_3/b, x_4/c] \text{ \quad by (} \beta_0) \]

\[ \cong_{ax} \text{let } \langle a, x_2 \rangle = x \text{ in } \text{let } \langle b, c \rangle = x_2 \text{ in } \]

\[ \langle \langle V, W \rangle, U \rangle : (A' \otimes B') \otimes C' \text{ \quad by (a}_{let}) \].
9.3. The Category $C_\lambda$

The right hand side is

$$j : A \otimes (B \otimes C) \triangleright \text{let } y = \left( \begin{array}{l}
\text{let } (a,i) = j \text{ in } \langle V, \langle W,U \rangle \rangle \\
\text{let } (b,c) = i \text{ in } \langle V, \langle W,U \rangle \rangle
\end{array} \right) \text{ in }$$

$$\langle V, \langle W,U \rangle \rangle$$

$$\text{let } (y_1, y_2) = y \text{ in }$$

$$\text{let } (y_3, y_4) = y_2 \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle$$

$$\cong_{ax} \text{let } (a,i) = j \text{ in } \langle b,c \rangle = i \text{ in }$$

$$\langle \langle y_1, y_3 \rangle, y_4 \rangle$$

$$\cong_{ax} \text{let } (a,i) = j \text{ in } \langle b,c \rangle = i \text{ in }$$

$$\langle V, \langle W,U \rangle \rangle$$

which is axiomatically equivalent to the left hand side. Thus $\alpha$ is a natural transformation.

**Bifunctoriality of $\otimes$.** First we need to show that $\otimes$ is well defined on equivalence classes. Given $x : A \triangleright V \cong_{ax} V' : B$ and $x : C \triangleright W \cong_{ax} W' : D$:

$$z : A \otimes C \triangleright \text{let } \langle x,y \rangle = z \text{ in } \langle V, W \rangle$$

$$\cong_{ax} \text{let } \langle x,y \rangle = z \text{ in } \langle V', W' \rangle : B \otimes D$$

since $\langle V, W \rangle \cong_{ax} \langle V', W' \rangle$.

Then we need to show that $(- \otimes B)$ and $(A \otimes -)$ are functors, and that they verify

$$(f \otimes B); (A' \otimes g) = (A \otimes g); (f \otimes B')$$

for every $f : A \rightarrow A'$ and every $g : B \rightarrow B'$.

$(id_A \otimes id_B)$ is $(x : A \triangleright x : A) \otimes (y : B \triangleright y : B)$, which is by definition $z : A \otimes B \triangleright \text{let } \langle x,y \rangle = z \text{ in } \langle x,y \rangle : A \otimes B$, which is by Rule $(\eta_\otimes)$ the identity $z : A \otimes B \triangleright z : A \otimes B$. So both $(- \otimes B)$ and $(A \otimes -)$ preserves identities.

Let $f$ be $x : A \triangleright V : A'$ and $f'$ be $y : A' \triangleright W : A''$. $((f, f') \otimes B)$ is

$$z : A \otimes B \triangleright \text{let } \langle x,t \rangle = z \text{ in } \langle V, W, t \rangle$$

$$\cong_{ax} \text{let } \langle x,t \rangle = z \text{ in } \langle y, V \text{ in } W, t \rangle$$

by $(let_2^\otimes)$.

On another hand, $(f \otimes B); (f' \otimes B)$ is

$$z : A \otimes B \triangleright u = \langle \text{let } \langle x,t \rangle = z \text{ in } \langle V, t \rangle \rangle \text{ in }$$

$$\langle y, V \text{ in } W, v \rangle$$

$$\cong_{ax} \text{let } \langle x,t \rangle = z \text{ in } u = \langle \text{let } \langle y, V \text{ in } W, t \rangle \rangle \text{ in }$$

$$\langle y, V \text{ in } W, v \rangle$$

by $(let_1)$

$$\cong_{ax} \text{let } \langle x,t \rangle = z \text{ in } \langle y, V \text{ in } W, v \rangle$$

by $(\beta_\otimes)$

$$\cong_{ax} \text{let } \langle x,t \rangle = z \text{ in } y = V \text{ in }$$
9.3. The Category $C_\lambda$

\[ \text{let } v = t \text{ in } (W, v) \]
\[ \approx_{ax} \text{let } (x, t) = z \text{ in } y = V \text{ in } (W, t) \]

by $(\beta_\otimes), (\beta_\lambda)$

and thus the two morphisms are equals.

Using $(\text{let}_1^\otimes)$ in place of $(\text{let}_2^\otimes)$, we can show that $(A \otimes (g, g')) = (A \otimes g); (A \otimes g')$, for $g : B \to B'$ and $g' : B' \to B''$.

Finally, we must show that if $f$ is $x : A \triangleright V : A'$ and $g$ is $y : B \triangleright U : B'$, then $(f \otimes B); (A' \otimes g) = (A \otimes g); (f \otimes B')$. The left hand side is

\[
\begin{align*}
z & : A \otimes B \triangleright \text{let } z' = (\text{let}(x, t) = z \text{ in } \langle V, t \rangle) \text{in} \\
& \approx_{ax} \text{let } (x, t) = z \text{ in } \text{let } z' = \langle V, t \rangle \text{ in} \\
& \approx_{ax} \text{let } (x, t) = z \text{ in } \text{let } y = V \text{ in} \\
& \approx_{ax} \text{let } (x, y) = z \text{ in } \langle V, W \rangle
\end{align*}
\]

by $(\text{let}_1)$, $(\beta_\lambda), (\beta_\otimes), (\beta_\lambda)$ and $(\alpha_{let})$, this is axiomatically equivalent to $\text{let}(x, y) = z \text{ in } \langle V, W \rangle$. Thus the required equality hold, and $\otimes$ is bifunctorial.

Commutative diagrams. We need to verify Equations 2.7.1, 2.7.2, 2.7.3, 2.7.4, 2.7.5, and Equation 2.7.6.

Equations 2.7.1. We want to prove that

\[
\alpha_{A, B, C \otimes D}; \alpha_{A \otimes B, C, D} = (\text{id}_A \otimes \alpha_{B, C \otimes D}); \alpha_{A, B \otimes C, D}; \alpha_{A, B \otimes C} \otimes \text{id}_D.
\]

For the left hand side, the components are as follows:

\[
\begin{align*}
\alpha_{A, B, C \otimes D} & = x : A \otimes (B \otimes (C \otimes D)) \triangleright \text{let } \langle x_1, x_2 \rangle = x \text{ in} \\
& \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in} \\
& \langle \langle x_1, x_3 \rangle, x_4 \rangle : (A \otimes B) \otimes (C \otimes D), \\
\alpha_{A \otimes B, C, D} & = y : (A \otimes B) \otimes (C \otimes D) \triangleright \text{let } \langle y_1, y_2 \rangle = y \text{ in} \\
& \text{let } \langle y_3, y_4 \rangle = y_2 \text{ in} \\
& \langle \langle y_1, y_3 \rangle, y_4 \rangle : ((A \otimes B) \otimes C) \otimes D).
\end{align*}
\]

Then

\[
\begin{align*}
\alpha_{A, B, C \otimes D}; \alpha_{A \otimes B, C, D} & = x : A \otimes (B \otimes (C \otimes D)) \triangleright \\
& \text{let } y = \left( \begin{array}{l}
\text{let } \langle x_1, x_2 \rangle = x \text{ in} \\
\text{let } \langle x_3, x_4 \rangle = x_2 \text{ in} \\
\langle \langle x_1, x_3 \rangle, x_4 \rangle
\end{array} \right) \text{ in} \left( \begin{array}{l}
\text{let } \langle y_1, y_2 \rangle = y \text{ in} \\
\text{let } \langle y_3, y_4 \rangle = y_2 \text{ in} \\
\langle \langle y_1, y_3 \rangle, y_4 \rangle
\end{array} \right)
\end{align*}
\]

\[\approx_{ax} \text{ let } (x_1, x_2) = x \text{ in let } (x_3, x_4) = x_2 \text{ in}
\]
\[\text{let } y = \langle \langle x_1, x_3 \rangle, x_4 \rangle \text{ in let } \langle y_1, y_2 \rangle = y \text{ in } \text{by } (\text{let}_1)\]
\[\text{let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle\]
\[\approx_{ax} \text{ let } (x_1, x_2) = x \text{ in let } (x_3, x_4) = x_2 \text{ in}
\]
\[\text{let } \langle y_1, y_2 \rangle = \langle \langle x_1, x_3 \rangle, x_4 \rangle \text{ in } \text{by } (\text{let}_1)\]
\[\text{let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle\]
\[\approx_{ax} \text{ let } (x_1, x_2) = x \text{ in let } (x_3, x_4) = x_2 \text{ in } \text{by } (\beta_{\otimes})\]
\[\text{let } \langle y_3, y_4 \rangle = x_4 \text{ in } \langle \langle x_1, x_3 \rangle, y_3 \rangle, y_4 \rangle\]

For the right hand side, the components are as follows:

\[\alpha_{B,C \otimes D} = x : B \otimes (C \otimes D) \triangleright \text{let } (x_1, x_2) = x \text{ in}
\]
\[\text{let } (x_3, x_4) = x_2 \text{ in } \langle \langle x_1, x_3 \rangle, x_4 \rangle : (B \otimes C) \otimes D,\]
\[\alpha_{A,B \otimes C} = y : A \otimes (B \otimes C) \triangleright \text{let } \langle y_1, y_2 \rangle = y \text{ in}
\]
\[\text{let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle : (A \otimes B) \otimes C,\]
\[\alpha_{A,B \otimes C, D} = z : A \otimes ((B \otimes C) \otimes D) \triangleright \text{let } \langle z_1, z_2 \rangle = z \text{ in}
\]
\[\text{let } (z_3, z_4) = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle : (A \otimes (B \otimes C)) \otimes D,\]

and \(id_A = t : A \triangleright t : A, \ id_D = u : D \triangleright u : D\). Then we compute \(id_A \otimes \alpha_{B,C \otimes D}\) and \(\alpha_{A,B \otimes C} \otimes id_D\):

\[id_A \otimes \alpha_{B,C \otimes D} = a : A \otimes (B \otimes (C \otimes D)) \triangleright
\]
\[\text{let } (t, x) = a \text{ in } \begin{cases} \text{let } (x_1, x_2) = x \text{ in } (t, \text{let } (x_3, x_4) = x_2 \text{ in } \langle \langle x_1, x_3 \rangle, x_4 \rangle) \end{cases} \text{by } (\text{let}_1)\]
\[\alpha_{A,B \otimes C} \otimes id_D = b : (A \otimes (B \otimes C)) \otimes D \triangleright
\]
\[\text{let } \langle y, u \rangle = b \text{ in } \begin{cases} \text{let } \langle y_1, y_2 \rangle = y \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle, u \end{cases} \text{by } (\text{let}_2)\]

Let's compose \((id_A \otimes \alpha_{B,C \otimes D}) ; \alpha_{A,B \otimes C, D}\):

\[(id_A \otimes \alpha_{B,C \otimes D}) ; \alpha_{A,B \otimes C, D} = \]
\[a : A \otimes (B \otimes (C \otimes D)) \triangleright \text{let } z = \begin{cases} \text{let } (t, x) = a \text{ in let } (x_1, x_2) = x \text{ in } \text{let } (x_3, x_4) = x_2 \text{ in } (t, \langle \langle x_1, x_3 \rangle, x_4 \rangle) \end{cases} \text{in}
\]
\[\text{let } (z_1, z_2) = z \text{ in let } (z_3, z_4) = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle\]

\[\text{in } \langle \langle x_1, x_3 \rangle, x_4 \rangle, y_3 \rangle, y_4 \rangle\]

\[\approx_{ax} \text{ let } (x_1, x_2) = x \text{ in let } (x_3, x_4) = x_2 \text{ in}
\]
\[\text{let } y = \langle \langle x_1, x_3 \rangle, x_4 \rangle \text{ in let } \langle y_1, y_2 \rangle = y \text{ in } \text{by } (\text{let}_1)\]
\[\text{let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle\]
\[\approx_{ax} \text{ let } (x_1, x_2) = x \text{ in let } (x_3, x_4) = x_2 \text{ in } \text{by } (\beta_{\otimes})\]
\[\text{let } \langle y_3, y_4 \rangle = x_4 \text{ in } \langle \langle x_1, x_3 \rangle, y_3 \rangle, y_4 \rangle\]

For the right hand side, the components are as follows:

\[\alpha_{B,C \otimes D} = x : B \otimes (C \otimes D) \triangleright \text{let } (x_1, x_2) = x \text{ in}
\]
\[\text{let } (x_3, x_4) = x_2 \text{ in } \langle \langle x_1, x_3 \rangle, x_4 \rangle : (B \otimes C) \otimes D,\]
\[\alpha_{A,B \otimes C} = y : A \otimes (B \otimes C) \triangleright \text{let } \langle y_1, y_2 \rangle = y \text{ in}
\]
\[\text{let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle : (A \otimes B) \otimes C,\]
\[\alpha_{A,B \otimes C, D} = z : A \otimes ((B \otimes C) \otimes D) \triangleright \text{let } \langle z_1, z_2 \rangle = z \text{ in}
\]
\[\text{let } (z_3, z_4) = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle : (A \otimes (B \otimes C)) \otimes D,\]

and \(id_A = t : A \triangleright t : A, \ id_D = u : D \triangleright u : D\). Then we compute \(id_A \otimes \alpha_{B,C \otimes D}\) and \(\alpha_{A,B \otimes C} \otimes id_D\):

\[id_A \otimes \alpha_{B,C \otimes D} = a : A \otimes (B \otimes (C \otimes D)) \triangleright
\]
\[\text{let } (t, x) = a \text{ in } \begin{cases} \text{let } (x_1, x_2) = x \text{ in } (t, \text{let } (x_3, x_4) = x_2 \text{ in } \langle \langle x_1, x_3 \rangle, x_4 \rangle) \end{cases} \text{by } (\text{let}_1)\]
\[\alpha_{A,B \otimes C} \otimes id_D = b : (A \otimes (B \otimes C)) \otimes D \triangleright
\]
\[\text{let } \langle y, u \rangle = b \text{ in } \begin{cases} \text{let } \langle y_1, y_2 \rangle = y \text{ in } \langle \langle y_1, y_3 \rangle, y_4 \rangle, u \end{cases} \text{by } (\text{let}_2)\]

Let's compose \((id_A \otimes \alpha_{B,C \otimes D}) ; \alpha_{A,B \otimes C, D}\):

\[(id_A \otimes \alpha_{B,C \otimes D}) ; \alpha_{A,B \otimes C, D} = \]
\[a : A \otimes (B \otimes (C \otimes D)) \triangleright \text{let } z = \begin{cases} \text{let } (t, x) = a \text{ in let } (x_1, x_2) = x \text{ in } \text{let } (x_3, x_4) = x_2 \text{ in } (t, \langle \langle x_1, x_3 \rangle, x_4 \rangle) \end{cases} \text{in}
\]
\[\text{let } (z_1, z_2) = z \text{ in let } (z_3, z_4) = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle\]
9.3. The Category $\mathcal{C}_\lambda$

\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in} \]
\[ \text{ let } \langle x_3, x_4 \rangle = x_2 \text{ in let } z = \langle t, \langle x_1, x_3 \rangle, x_4 \rangle \text{ in} \]
\[ \text{ let } (z_1, z_2) = z \text{ in let } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle (z_1, z_3), z_4 \rangle \text{ by (let1)} \]
\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in} \]
\[ \text{ let } (z_1, z_2) = \langle t, \langle x_1, x_3 \rangle, x_4 \rangle \text{ in} \]
\[ \text{ let } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle (z_1, z_3), z_4 \rangle \text{ by (let1)} \]
\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in} \]
\[ \text{ let } \langle z_1, z_2 \rangle = \langle t, \langle x_1, x_3 \rangle, x_4 \rangle \text{ in} \]
\[ \text{ let } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle (z_1, z_3), z_4 \rangle \text{ by (let1)} \]
\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in} \]
\[ \text{ let } \langle z_1, z_2 \rangle = \langle t, \langle x_1, x_3 \rangle, x_4 \rangle \text{ in} \]
\[ \text{ let } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle (z_1, z_3), z_4 \rangle \text{ by (let1)} \]

Now we can compute the right hand of the equation:

\[ (\text{id}_A \otimes \alpha_{A,B\otimes C,D}) \circ \alpha_{A,B\otimes C,D} \circ (\alpha_{A,B\otimes C} \otimes \text{id}_D) = \]

\[ a : A \otimes (B \otimes (C \otimes D)) \triangleright \text{ let } b = \left( \begin{array}{c}
\text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in } \langle (t, (x_1, x_3)), x_4 \rangle, x_4 \end{array} \right) \text{ in} \]
\[ \text{ let } (y, u) = b \text{ in let } \langle y_1, y_2 \rangle = y \text{ in let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle (y_1, y_3), y_4 \rangle, u \rangle \]

\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in let } \langle y_1, y_2 \rangle = y \text{ in let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle (y_1, y_3), y_4 \rangle, u \rangle \text{ by (let1)} \]

\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in let } \langle y_1, y_2 \rangle = y \text{ in let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle (y_1, y_3), y_4 \rangle, u \rangle \text{ by (let1)} \]

\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in let } \langle y_1, y_2 \rangle = y \text{ in let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle (y_1, y_3), y_4 \rangle, u \rangle \text{ by (let1)} \]

\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in let } \langle y_1, y_2 \rangle = y \text{ in let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle (y_1, y_3), y_4 \rangle, u \rangle \text{ by (let1)} \]

\[ \cong_{ax} \text{ let } (t, x) = a \text{ in let } \langle x_1, x_2 \rangle = x \text{ in let } \langle x_3, x_4 \rangle = x_2 \text{ in let } \langle y_1, y_2 \rangle = y \text{ in let } \langle y_3, y_4 \rangle = y_2 \text{ in } \langle \langle (y_1, y_3), y_4 \rangle, u \rangle \text{ by (let1)} \]

and this is alpha-equivalent to $\alpha_{A,B\otimes C,D} \circ \alpha_{A,B\otimes C,D}$. Thus Equation (4.7) is verified.

**Equation (4.7)**. We want to show that $\alpha_{A,T,C} : (\beta_{A \otimes C}) = \text{id}_A \otimes \lambda_C$. We first write down the components of the equation:

\[ \alpha_{A,T,C} = \]
\[ a : A \otimes (T \otimes C) \triangleright \text{ let } \langle a_1, a_2 \rangle = a \text{ in let } \langle a_3, a_4 \rangle = a_2 \text{ in } \langle \langle a_1, a_3 \rangle, a_4 \rangle : (A \otimes T) \otimes C, \]
\[ \lambda_C = \]
\[ b : T \otimes C \triangleright \text{ let } \langle b_1, b_2 \rangle = b \text{ in let } * = b_1 \text{ in } b_2 : C, \]
\[ \rho_A = \]
\[ c : A \otimes T \triangleright \text{ let } \langle c_1, c_2 \rangle = c \text{ in let } * = c_2 \text{ in } c_1 : A, \]

and $\text{id}_A = u : A \triangleright u : A$ and $\text{id}_C = v : C \triangleright v : C$. Then we construct $\rho_A \otimes \text{id}_C$ and $\text{id}_A \otimes \lambda_C$:

\[ \rho_A \otimes \text{id}_C = \]
\[ x : (A \otimes T) \otimes C \triangleright \text{ let } \langle c, v \rangle = x \text{ in let } * = c_2 \text{ in } c_1 : A, v \]
We then compose $\alpha_{A,T,C}$ and $\rho_A \otimes \lambda_C$:

$$a : A \otimes (T \otimes C) \triangleright let \ x = \left( \begin{array}{l} \text{let } (a_1, a_2) = a \text{ in } \langle a_1, a_3 \rangle = a_2 \text{ in } \langle a_1, a_3, a_4 \rangle \end{array} \right) \text{ in }$$

$$\approx_{ax} \text{ let } (c, v) = x \text{ in let } (c_1, c_2) = c \text{ in }$$

$$\text{let } * = c_2 \text{ in } \langle c_1, v \rangle$$

$$\approx_{ax} \text{ let } \langle a_1, a_2 \rangle = a \text{ in let } \langle a_3, a_4 \rangle = a_2 \text{ in }$$

$$\text{let } x = \langle \langle a_1, a_3 \rangle, a_4 \rangle \text{ in }$$

$$\text{let } (c, v) = x \text{ in let } (c_1, c_2) = c \text{ in }$$

$$\text{let } * = c_2 \text{ in } \langle c_1, v \rangle \text{ by (let}_1)$$

$$\approx_{ax} \text{ let } \langle a_1, a_2 \rangle = a \text{ in let } \langle a_3, a_4 \rangle = a_2 \text{ in }$$

$$\text{let } \langle c, v \rangle = \langle \langle a_1, a_3 \rangle, a_4 \rangle \text{ in let } \langle c_1, c_2 \rangle = c \text{ in }$$

$$\text{let } * = c_2 \text{ in } \langle c_1, v \rangle \text{ by } \beta_{\alpha}$$

$$\approx_{ax} \text{ let } \langle a_1, a_2 \rangle = a \text{ in let } \langle a_3, a_4 \rangle = a_2 \text{ in }$$

$$\text{let } \langle c_1, c_2 \rangle = \langle a_1, a_3 \rangle \text{ in let } * = c_2 \text{ in } \langle c_1, a_4 \rangle \text{ by } (\beta_{\beta})$$

$$\approx_{ax} \text{ let } \langle a_1, a_2 \rangle = a \text{ in let } \langle a_3, a_4 \rangle = a_2 \text{ in }$$

$$\text{let } \langle c_1, a_2 \rangle = a \text{ in let } c_1, a_4 : A \otimes C \text{ by } (\beta_{\beta})$$

which is alpha-equivalent to $\lambda_C \otimes \beta_C$: Diagram 2.7.4 commutes.

**Equation (2.7.4)**. We want to show $\lambda_T = \rho_T$.

$$\lambda_T = \ x : T \otimes T \triangleright let \ \langle y, z \rangle = x \text{ in let } * = y \text{ in } z$$

$$\approx_{ax} \text{ let } \langle y, z \rangle = x \text{ in }$$

$$\text{let } * = y \text{ in let } * = z \text{ in } * : T \text{ by } (\eta_*)$$

$$\rho_T = \ x : T \otimes T \triangleright let \ \langle y, z \rangle = x \text{ in let } * = z \text{ in } y$$

$$\approx_{ax} \text{ let } \langle y, z \rangle = x \text{ in }$$

$$\text{let } * = z \text{ in let } * = y \text{ in } * \text{ by } (\eta_*)$$

$$\approx_{ax} \text{ let } \langle y, z \rangle = x \text{ in }$$

$$\text{let } * = y \text{ in let } * = z \text{ in } * : T \text{ by (let}_2).$$

Thus the equation holds.

**Equation (2.7.5)**. We want $\sigma_{A,B} ; \sigma_{B,A} = id_{A \otimes B}$.

$$\sigma_{A,B} = x : A \otimes B \triangleright let \ \langle x_1, x_2 \rangle = x \text{ in } (x_2, x_1) : B \otimes A$$

$$\sigma_{B,A} = y : B \otimes A \triangleright let \ \langle y_1, y_2 \rangle = y \text{ in } (y_2, y_1) : A \otimes B$$

$$\sigma_{A,B} ; \sigma_{B,A} = x : A \otimes B \triangleright let \ y = (let \ \langle x_1, x_2 \rangle = x \text{ in } (x_2, x_1)) \text{ in }$$

$$\text{let } \langle y_1, y_2 \rangle = y \text{ in } (y_2, y_1)$$
9.3. The Category $C_\Lambda$

\[ \approx_\text{ax} \text{ let } \langle x_1, x_2 \rangle = x \text{ in } \]
\[ \text{let } y = \langle x_2, x_1 \rangle \text{ in } \]
\[ \text{let } \langle y_1, y_2 \rangle = y \text{ in } \langle y_2, y_1 \rangle \]
\[ \approx_\text{ax} \text{ let } \langle x_1, x_2 \rangle = x \text{ in } \]
\[ \text{let } \langle y_1, y_2 \rangle = \langle x_2, x_1 \rangle \text{ in } \langle y_2, y_1 \rangle \]
\[ \approx_\text{ax} \text{ let } \langle x_1, x_2 \rangle = x \text{ in } (x_1, x_2) \]
\[ \approx_\text{ax} \text{ x : A } \otimes B \]

and thus $\text{(2.7.4)}$ holds.

**Equation (2.7.5).** We want $\sigma_{B, T} ; \lambda_B = \rho_B$. The components are:

\[ \sigma_{B, T} = x : B \otimes T \triangleright \text{let } \langle x_1, x_2 \rangle = x \text{ in } \langle x_2, x_1 \rangle : B \otimes T, \]
\[ \lambda_B = y : T \otimes B \triangleright \text{let } \langle y_1, y_2 \rangle = y \text{ in } \text{let } * = y_1 \text{ in } y_2 : B, \]
\[ \rho_B = x : B \otimes T \triangleright \text{let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } * = x_2 \text{ in } x_1 : B. \]

Let’s compute the left hand side:

\[ \sigma_{B, T} ; \lambda_B = x : B \otimes T \triangleright \text{let } y = \langle \text{let } \langle x_1, x_2 \rangle = x \text{ in } \langle x_2, x_1 \rangle \rangle \text{ in } \]
\[ \text{let } \langle y_1, y_2 \rangle = y \text{ in } \text{let } * = y_1 \text{ in } y_2 \]
\[ \approx_\text{ax} \text{ let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } y = \langle x_2, x_1 \rangle \text{ in } \]
\[ \text{let } \langle y_1, y_2 \rangle = y \text{ in } \text{let } * = y_1 \text{ in } y_2 \]
\[ \approx_\text{ax} \text{ let } \langle x_1, x_2 \rangle = x \text{ in } \]
\[ \text{let } \langle y_1, y_2 \rangle = \langle x_2, x_1 \rangle \text{ in } \text{let } * = y_1 \text{ in } y_2 \]
\[ \approx_\text{ax} \text{ let } \langle x_1, x_2 \rangle = x \text{ in } \text{let } * = x_2 \text{ in } x_1 : B \]

which is exactly $\rho_B$.

**Equation (2.7.6).** We want

\[ \alpha_{A, B, C} ; \sigma_{A \otimes B, C} ; \alpha_{C, A, B} = \langle \text{id}_A \otimes \sigma_{B, C} ; \alpha_{A, C, B} ; (\sigma_{A, C} \otimes \text{id}_B) \rangle. \]

Left hand side. Let’s write down the components:

\[ \alpha_{A, B, C} = x : A \otimes (B \otimes C) \triangleright \text{let } \langle x_1, x_2 \rangle = x \text{ in } \]
\[ \text{let } \langle x_3, x_4 \rangle = x_2 \text{ in } \langle \langle x_1, x_3 \rangle, x_4 \rangle : (A \otimes C) \otimes B, \]
\[ \sigma_{A \otimes B, C} = y : (A \otimes B) \otimes C \triangleright \text{let } \langle y_1, y_2 \rangle = y \text{ in } \langle y_2, y_1 \rangle : C \otimes (A \otimes B), \]
\[ \alpha_{C, A, B} = z : C \otimes (A \otimes B) \triangleright \text{let } \langle z_1, z_2 \rangle = z \text{ in } \]
\[ \text{let } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle : (C \otimes A) \otimes B. \]

Let’s compose $\alpha_{A, B, C}$ and $\sigma_{A \otimes B, C}$:

\[ \alpha_{A, B, C} ; \sigma_{A \otimes B, C} = x : A \otimes (B \otimes C) \triangleright \]
9.3. The Category $C_A$

Let $y = \left( \begin{array}{l} \text{let } (x_1, x_2) = x \text{ in } \\ \text{let } (x_3, x_4) = x_2 \text{ in } \langle x_1, x_3 \rangle, x_4 \rangle \end{array} \right) \text{ in }$

$\approx_{ax} \text{ let } y = \langle x_1, x_3 \rangle, x_4 \rangle \text{ in }$

Let $\alpha_{A,B,C} : \sigma_{A \otimes B,C} : \alpha_{C,A,B} = x : A \otimes (B \otimes C) \triangleright$

$\approx_{ax} \text{ let } (z_1, z_2) = z \text{ in } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle$

Let $\beta_{A,B,C} : \approx_{ax} \text{ let } (x_1, x_2) = x \text{ in } \langle x_3, x_4 \rangle = x_2 \text{ in }$

$\approx_{ax} \text{ let } (z_3, z_4) = z_2 \text{ in } \langle \langle z_1, z_3 \rangle, z_4 \rangle \text{ by (let}_1)$

$\approx_{ax} \text{ let } (x_1, x_2) = x \text{ in } \langle x_3, x_4 \rangle = x_2 \text{ in }$

Right hand side: let’s right down the components:

$id_A = a : A \triangleright a : A,$

$\sigma_{B,C} = b : B \otimes C \triangleright \text{ let } \langle b_1, b_2 \rangle = b \text{ in } \langle b_2, b_1 \rangle : C \otimes B,$

thus

$id_A \otimes \sigma_{B,C} = \langle a, b \rangle = x \text{ in }$

$\approx_{ax} \text{ let } \langle a, b \rangle = x \text{ in }$

$\approx_{\beta} \text{ let } \langle b_1, b_2 \rangle = b \text{ in } \langle b_2, b_1 \rangle \text{ by (let}_1^\otimes).$

$id_B = c : B \triangleright c : B,$

$\sigma_{A,C} = d : A \otimes C \triangleright \text{ let } \langle d_1, d_2 \rangle = d \text{ in } \langle d_2, d_1 \rangle : C \otimes A,$

thus

$\sigma_{A,C} \otimes id_B = y : (A \otimes C) \otimes B \triangleright \text{ let } \langle d, c \rangle = y \text{ in }$
\[ \langle \text{let } (d_1, d_2) = d \text{ in } (d_2, d_1), c \rangle \]
\[ \cong_{\text{ax}} \text{let } (d, c) = y \text{ in } \text{let } (d_1, d_2) = d \text{ in } \langle (d_2, d_1), c \rangle : (B \otimes A) \otimes B \]

Finally,
\[ \alpha_{A,C,B} = z : A \otimes (C \otimes B) \triangleright \text{let } \langle z_1, z_2 \rangle = z \text{ in } \text{let } \langle z_3, z_4 \rangle = z_2 \text{ in } \langle (z_1, z_3), z_4 \rangle : (A \otimes C) \otimes B. \]

Let's compose \( id_A \otimes \sigma_{B,C} \) and \( \alpha_{A,C,B} \):
\[ (id_A \otimes \sigma_{B,C}) ; \alpha_{A,C,B} = x : A \otimes (B \otimes C) \triangleright \]
\[ \text{let } z = \left( \begin{array}{c}
\text{let } (a, b) = x \text{ in } \text{let } (b_1, b_2) = b \text{ in } \langle (a, b_2), b_1 \rangle \\
\text{let } (z_1, z_2) = z \text{ in } \text{let } (z_3, z_4) = z_2 \text{ in } \langle (z_1, z_3), z_4 \rangle \\
\end{array} \right) \text{ in } \langle (a, b_2), b_1 \rangle : (A \otimes C) \otimes B \]

Let's compute the complete right hand side:
\[ (id_A \otimes \sigma_{B,C}) ; \alpha_{A,C,B} ; (\sigma_{A,C} \otimes id_B) = x : A \otimes (B \otimes C) \triangleright \]
\[ \text{let } y = \left( \begin{array}{c}
\text{let } (a, b) = x \text{ in } \text{let } (b_1, b_2) = b \text{ in } \langle (a, b_2), b_1 \rangle \\
\text{let } (d, c) = y \text{ in } \text{let } (d_1, d_2) = d \text{ in } \langle (d_2, d_1), c \rangle \\
\end{array} \right) \text{ in } \langle (b_2, a), b_1 \rangle : (C \otimes A) \otimes B \]

And this is alpha equivalent to the left hand side. So the equation holds.

\[ \square \]

9.3.2 Monadic Structure

**Definition 9.3.4.** We define the term \( \lambda. M \) to be \( \lambda y. (\text{let } * = y \text{ in } M) \), for \( M \) a fresh variable.
Lemma 9.3.5. The following axiomatic relations are satisfied:

\[ \Delta \triangleright (\lambda x.M) \approx_{ax} M : A \quad (\beta_*) \]
\[ \Delta \triangleright (\lambda x.V) \approx_{ax} V : T \rightarrow A \quad (\eta_*) \]

Proof. Proof by simple calculation. \(\square\)

Lemma 9.3.6. Let \(\eta_A\) be the arrow \(x : A \triangleright \lambda x.x : T \rightarrow A\) and

\[ (x : A \triangleright V : T \rightarrow B) \overset{\alpha}{=} y : T \rightarrow A \triangleright \lambda x. \text{let } x = (y*) \text{ in } (V*) : T \rightarrow B. \]

Then \((T \rightarrow \alpha, \eta, -^\ast)\) is a Kleisli triple in \(C_\lambda\).

Proof. We need to check Equations (2.5.3), (2.5.4) and (2.5.5).

Equation (2.5.3). We want \(\eta_A^* = id_{T \rightarrow A}\):

\[ \eta_A = x : A \triangleright \lambda x.x : T \rightarrow A \]
\[ \eta_A^* = y : T \rightarrow A \triangleright \lambda x. \text{let } x = (y*) \text{ in } ((\lambda x.x)^*) : T \rightarrow A \]
\[ = y : T \rightarrow A \triangleright \lambda x. \text{let } x = (y*) \text{ in } x \triangleright T \rightarrow A \]
\[ = y : T \rightarrow A \triangleright \lambda x.y^* : T \rightarrow A \quad \text{by } (\beta_*) \]
\[ = y : T \rightarrow A \triangleright y : T \rightarrow A \quad \text{by } (\eta_*) \]
\[ = id_{T \rightarrow A}. \]

Equation (2.5.4). We want \(\eta_A; f^* = f\):

\[ \eta_A = x : A \triangleright \lambda x.x : T \rightarrow A, \quad f = y : A \triangleright V : T \rightarrow B, \]

thus

\[ f^* = z : T \rightarrow A \triangleright \lambda x. \text{let } y = (z*) \text{ in } (V*) : T \rightarrow B, \quad \text{and} \]
\[ \eta_A; f^* = x : A \triangleright \lambda x.x \overset{\text{let } z = (\lambda x.x)^*}{=} \]
\[ \lambda x. \text{let } y = (z*) \text{ in } (V*) : T \rightarrow B \]
\[ = x : A \triangleright \lambda x. \text{let } y = ((\lambda x.x)^*) \text{ in } (V*) : T \rightarrow B \quad \text{by } (\beta_*) \]
\[ = x : A \triangleright \lambda x. \text{let } y = x \text{ in } (V*) : T \rightarrow B \quad \text{by } (\beta_*) \]
\[ = y : A \triangleright \lambda x.y^* : T \rightarrow B \quad \text{by } (\alpha_{let}) \]
\[ = y : A \triangleright V : T \rightarrow B \quad \text{by } (\eta_*) \]
\[ = f. \]

Equation (2.5.5). We want \(f^*; g^* = (f; g^*)^*\). If

\[ f = x : A \triangleright V : T \rightarrow B, \quad g = y : B \triangleright W : T \rightarrow C, \]

then

\[ f^* = z : T \rightarrow A \triangleright \lambda x. \text{let } x = (z*) \text{ in } (V*) : T \rightarrow B, \]
\[ g^* = t : T \rightarrow B \triangleright \lambda x. \text{let } y = (t*) \text{ in } (W*) : T \rightarrow C. \]

Let’s compute

\[ f^*; g^* = z : T \rightarrow A \triangleright \text{let } t = (\lambda x. \text{let } x = (z*) \text{ in } (V*)). \]
in \( \lambda \). let \( y = (t*) \in (W*) : \top \rightarrow C \),
\[ = z : \top \rightarrow A \triangleright \lambda \), let \( y = (z*) \in (W*) : \top \rightarrow C \]
by \( (\beta_\lambda) \)
\[ = z : \top \rightarrow A \triangleright \lambda \), let \( y = (z*) \in (W*) \)
in \( (W*) : \top \rightarrow C \)
by \( (\beta_\lambda) \), \( (\beta_\lambda) \)
\[ = z : \top \rightarrow A \triangleright \lambda \), let \( x = (z*) \in \]
let \( y = (V*) \in (W*) : \top \rightarrow C \)
by \( (\text{let}_1) \).

Next we need
\[ f; g^* = x : A \triangleright \text{let} \ t = \ V \]
in \( \lambda \), let \( y = (t*) \in (W*) : \top \rightarrow C \),
\[ = x : A \triangleright \lambda \), let \( t = \ V \in \]
let \( y = (t*) \in (W*) : \top \rightarrow C \)
by \( (\text{let}^\lambda) \),
and thus
\[ (f; g^*)^* = z : \top \rightarrow A \triangleright \lambda \), let \( x = (z*) \in \]
\( \lambda \). (let \( t = \ V \ in \ let \ y = (t*) \in (W*)^* : \top \rightarrow C \)
\[ = z : \top \rightarrow A \triangleright \lambda \), let \( x = (z*) \in \]
let \( t = \ V \ in \ let \ y = (t*) \in (W*) : \top \rightarrow C \)
by \( (\beta_\lambda) \), \( (\beta_\lambda) \)
\[ = z : \top \rightarrow A \triangleright \lambda \), let \( x = (z*) \in \]
let \( y = (V*) \in (W*) : \top \rightarrow C \)
by \( (\beta_\lambda) \)
\[ = f^* : g^* \]
This ends the proof. \( \square \)

**Lemma 9.3.7.** Define the following maps:

\[
\eta_A = x : A \triangleright \lambda, x : \top \rightarrow A,
\mu_A = x : \top \rightarrow (\top \rightarrow A) \triangleright \lambda, (x*) \in (W*) : \top \rightarrow A,
\]
\( t_{A,B} = z : A \otimes (\top \rightarrow B) \triangleright \lambda, (x,y) = z \in \lambda, (x,y) : \top \rightarrow (A \otimes B), \)
and define \( \top \rightarrow (x : A \triangleright \top : V : B) \) as the arrow
\[ y : \top \rightarrow A \triangleright \lambda, \text{let} \ x = (y*) \in V : \top \rightarrow B. \]

Then \( (\top \rightarrow - , \eta, \mu, t) \) is a strong monad.

**Proof.** The fact that \( (\top \rightarrow - , \eta, \mu) \) is a monad is a corollary of Lemmas 9.3.6 and 2.5.3. It remains to show that \( t_{A,B} \) is natural in \( A \) and \( B \), and that it satisfies Equations (9.3.1), (9.3.4) and (9.3.5).

**Naturality of** \( t_{A,B} \) **in** \( A \). We must show that given a map \( f : A \rightarrow A' \), the diagram

\[
\begin{array}{ccc}
A \otimes (\top \rightarrow B) & \xrightarrow{t_{A,B}} & \top \rightarrow (A \otimes B) \\
\downarrow \cong & & \downarrow \cong \\
(f \otimes \top \rightarrow B) & \xrightarrow{t_{A',B}} & \top \rightarrow (A' \otimes B)
\end{array}
\]
9.3. The Category $C_\lambda$

commutes. If $f$ is the value $t : A \triangleright V : A'$, then

$$f \otimes (\top \rightarrow B) : u : A \otimes (\top \rightarrow B) \triangleright \langle t, u' \rangle = u \in (V, u') : A' \otimes (\top \rightarrow B),$$

$$\top \rightarrow (f \otimes B) : v : \top \rightarrow (A \otimes B) \triangleright \lambda^* \begin{pmatrix} \text{let } v' = v^* \in \langle t, v'' \rangle = v' \in (V, v'') \end{pmatrix} : \top \rightarrow (A' \otimes B).$$

One needs to show that $f \otimes (\top \rightarrow B) ; t_{A' ; B} = t_{A ; B} ; \top \rightarrow (f \otimes B)$. The right hand side is

$$z : A \otimes (\top \rightarrow B) \triangleright \text{let } v = \begin{pmatrix} \text{let } \langle x, y \rangle = z \in \langle x, y^* \rangle \\
\text{let } v' = v^* \in \langle t, v'' \rangle = v' \in (V, v'') \\
\text{in } \lambda^* \end{pmatrix} \text{ by (let*) and (let_1)}$$

$$\approx_{ax} \lambda^* \begin{pmatrix} \text{let } \langle x, y \rangle = z \in \\
\text{let } v = \lambda^* . \langle x, y^* \rangle \\
\text{let } v' = v^* \in \langle t, v'' \rangle = v' \in (V, v'') \\
\text{in } \langle t, v'' \rangle = v' \in (V, v'') \end{pmatrix} \text{ by (let_1)}$$

$$\approx_{ax} \lambda^* \begin{pmatrix} \text{let } \langle x, y \rangle = z \in \\
\text{let } v = \lambda^* . \langle x, y^* \rangle \\
\text{let } v' = \langle x, y^* \rangle \text{ in } \langle t, v'' \rangle = v' \in (V, v'') \\
\text{in } \langle t, v'' \rangle = v' \in (V, v'') \end{pmatrix} \text{ by (let_1)}$$

The left hand side is

$$u : A \otimes (\top \rightarrow B) \triangleright \text{let } z = \langle t, u' \rangle = u \in (V, u') \text{ in }$$

$$\approx_{ax} \text{let } \langle t, u' \rangle = u \text{ in }$$

$$\text{let } z = \langle V, u' \rangle \text{ in }$$
This ends the proof.

We define $\Phi$. With these structures, $C_{\lambda}$ is a category.

**Naturality of $t_{A,B}$ in $B$.** One needs to show that given a map $f : B \to B'$, the diagram

$$
\begin{array}{ccc}
A \otimes (T \to B) & \xrightarrow{t_{A,B}} & T \to (A \otimes B) \\
\bigg\downarrow A \otimes (T \to f) & & \bigg\downarrow T \to (A \otimes f) \\
A' \otimes (T \to B') & \xrightarrow{t_{A',B'}} & T \to (A \otimes B')
\end{array}
$$

commutes. That is to say, $t_{A,B} \circ (T \to f) = A \otimes (T \to f) \circ t_{A,B'}$. If $f$ is the value $t : B \triangleright V : B'$, one has:

$$
A \otimes (T \to f) : \quad u : A \otimes (T \to B) \triangleright \langle t', u' \rangle = u \text{ in } \langle t', \lambda \ast \rangle \text{, let } \lambda \ast \text{ in } V : A \otimes C,
$$

$$
T \to \circ (A \otimes f) : \quad v : T \to (A \otimes B) \triangleright \lambda \ast . \left( \begin{array}{c}
\text{let } v' = v \ast \text{ in } \\
\text{let } \langle t', t \rangle = v' \triangleright t', V \end{array} \right) : T \to (A \otimes B').
$$

This ends the proof.

**Lemma 9.3.8.** Consider the functor $\circ : C_{\lambda} \times C_{\lambda} \to C_{\lambda}$, defined on arrows by

$$
\begin{array}{ccc}
A \to (x : B \triangleright V : C) & = & y : A \to B \triangleright \lambda z \text{, let } x = y z \text{ in } V : A \to C, \\
(x : A \triangleright V : B) \to C & = & y : B \to C \triangleright \lambda x , y V : A \to C,
\end{array}
$$

and the map $\Phi_{A,B,C} : C_{\lambda}(A,B \to C) \to C_{\lambda}(A \otimes B, T \to C)$:

$$
\langle x : A \triangleright V : B \to C \rangle \quad \mapsto \quad (t : A \otimes B \triangleright \lambda \ast \text{, let } \langle x, y \rangle = t \text{ in } V y : \lambda \ast \to C).
$$

With these structures, $C_{\lambda}$ has $(\triangleright \circ \to \circ)$-exponentials, as in Definition 5.4.8.

**Proof.** We define $\Phi_{A,B,C}^{-1}$:

$$
(t : A \otimes B \triangleright V : T \to C) \quad \mapsto \quad (x : A \triangleright \lambda y , \text{ let } t = \langle x, y \rangle \text{ in } V y : B \to C),
$$

and we show that they are inverse of each other. Given $x : A \triangleright V : B \to C$, let’s apply $\Phi_{A,B,C}^{-1} \circ \Phi_{A,B,C}$. We get

$$
x : A \triangleright \lambda y , \text{ (let } t = \langle x, y \rangle \text{ in } \lambda \ast \text{, let } \langle x, y \rangle = t \text{ in } V y) \ast
$$

$\approx_{ax} \lambda y , \text{ (let } t = \langle x, y \rangle \text{ in let } \langle x, y \rangle = t \text{ in } V y) \ast$

by $(\beta_{\ast})$

$\approx_{ax} \lambda y , \text{ (let } x = \langle x, y \rangle \text{ in } V y) \ast$

by $(\beta_{\ast})$

$\approx_{ax} V : B \to C$

by $(\eta_{\ast})$.

Similarly, if we apply $\Phi_{A,B,C} \circ \Phi_{A,B,C}^{-1}$ to $t : A \otimes B \triangleright V : T \to C$, we get

$$
t : A \otimes B \triangleright \lambda \ast , \text{ let } \langle x, y \rangle = t \text{ in } (\lambda y , \text{ let } t = \langle x, y \rangle \text{ in } V) y
$$
Let’s compute \((f, g)\) and thus \(\Phi\) is bifunctorial. We need to show that \((-\circ-\circ-\circ-)\) are functors, and that \((f \circ \circ - B); (A \circ g) = (A' \circ - g); (f \circ \circ - B)\), if \(f : A \rightarrow A'\) and \(g : B \rightarrow B'\).

The composition of these two functions is \(f\) which is precisely \((f' \circ \circ - B)\).

First consider \(id_A \circ id_B\): If \(id_A = a : A\) and \(id_B = b : B \circ b : B\) it is the arrow

\[
x : A \rightarrow B \circ \lambda x.\ let\ b = xa\ in\ b
\]

\[
\approx_{ax}\lambda a.xa
\]

\[
\approx_x x
\]

Then consider \(f = x : A \rightarrow V : A'\) and \(f' = y : A' \circ \circ - W : A''\). We compute \((f \circ \circ - B)\) and \((f' \circ \circ - B)\):

\[
t : A'' \rightarrow B \circ \lambda y.\ let\ b = tW\ in\ b
\]

\[
\approx_{ax}\lambda y.tW : A' \circ \circ - B
\]

\[
z : A' \rightarrow B \circ \lambda x.\ let\ b = zV\ in\ b
\]

\[
\approx_{ax}\lambda x.zV : A \rightarrow B
\]

The composition of these two functions is

\[
t : A'' \rightarrow B \circ \circ - B \circ \circ \circ \circ - z = \lambda y.tW\ in\ \lambda x.zV
\]

\[
\approx_{ax}\lambda x.(\lambda y.tW)V
\]

\[
\approx_{ax}\lambda x.\ let\ y = V\ in\ tW : A \rightarrow B
\]

by definition.

Let’s compute \((f'; f) \circ \circ - B)\:

\[
t : A'' \rightarrow B \circ \circ \circ \circ - B \circ \circ \circ \circ - b = t(\circ \circ - let\ y = V\ in\ W)\ in\ b
\]

\[
\approx_{ax}\lambda x.t(\circ \circ - let\ y = V\ in\ W)
\]

\[
\approx_{ax}\lambda x.\ let\ y = V\ in\ tW : A \rightarrow B
\]

by \((\beta_\circ^1)\),

which is precisely \((f' \circ \circ - B); (f \circ \circ - B)\).

Now consider \(g = x : B \circ \circ V : B'\) and \(g' = y : B' \circ \circ \circ \circ - W : B''\). We compute \((A \circ \circ - g)\) and \((A \circ \circ - g')\):

\[
t : A \rightarrow B \circ \lambda a.\ let\ x = ta\ in\ V : A \rightarrow B',
\]

\[
z : A \rightarrow B' \circ \circ \circ \circ \circ - \lambda a'.\ let\ y = za'\ in\ W : A \rightarrow B''
\]

The composition of these two functions is

\[
t : A \rightarrow B \circ \circ \circ \circ - let\ z = (\lambda a.\ let\ x = ta\ in\ V)\ in\ V
\]

\[
\lambda a'.\ let\ y = za'\ in\ W
\]

\[
\approx_{ax}\lambda a'.\ let\ y = (\lambda a.\ let\ x = ta\ in\ V)a'\ in\ W
\]

by \((\beta_\circ_1)\),

\[
\approx_{ax}\lambda a'.\ let\ y = (\circ \circ - let\ x = ta\ in\ V)\ in\ W
\]

by \((\beta_\circ)\).
which is exactly \( A \to (g; g') \).

Remains to show that if \( f = x : A \vdash U : A' \) and \( g = y : B \vdash V : B' \) then \( (A' \to g); (f \to B') = (f \to B); (A \to g) \).

Let us write down the components of the equations:

\[
\begin{align*}
A' \to g &= \quad x_1 : A' \to B \vdash \lambda x_2. \text{let } y = x_1 x_2 \text{ in } V : A' \to B' \\
A \to g &= \quad y_1 : A \to B \vdash \lambda y_2. \text{let } y = y_1 y_2 \text{ in } V : A \to B' \\
f \to B' &= \quad z : A' \to B' \vdash \lambda x. z U : A \to B' \\
f \to B &= \quad t : A' \to B \vdash \lambda x. t U : A \to B
\end{align*}
\]

Computing \((A' \to g); (f \to B')\), we get

\[
\begin{align*}
x_1 : A' \to B \vdash \text{let } z = (\lambda x_2. \text{let } y = x_1 x_2 \text{ in } V) \text{ in } \lambda x. z U \\
&\approx_{\text{ax}} \lambda x. (\lambda x_2. \text{let } y = x_1 x_2 \text{ in } V) U \\
&\approx_{\text{ax}} \lambda x. \text{let } y = x_1 U \text{ in } V : A \to B' \\
&\quad \text{by } (\beta_\lambda)
\end{align*}
\]

The other hand of the equation, \((f \to B); (A \to g)\), is

\[
\begin{align*}
t : A' \to B \vdash \text{let } y_1 = \lambda x. t U \text{ in } \lambda y_2. \text{let } y = y_1 y_2 \text{ in } V \\
&\approx_{\text{ax}} \lambda y_2. \text{let } y = (\lambda x. t U) y_2 \text{ in } V \\
&\approx_{\text{ax}} \lambda x. \text{let } y = t U \text{ in } V \\
&\quad \text{by } (\beta_\lambda)
\end{align*}
\]

and thus the equation holds. \(\square\)

### 9.3.3 Comonadic Structure

The goal is to have the type constructor \( ! \) to be a comonad. Ideally, we would like the following to be a derived typing rule:

\[
\begin{array}{c}
\frac{\Delta \vdash V : A}{! \Delta \vdash V : ! A} \\
\Delta \vdash V : ! A.
\end{array}
\]

However, in the context of indexed terms, the correct statement of this property is slightly more technical. It means there should be another indexation of the original value, well-typed of duplicable type.

**Definition 9.3.9.** We define a map \( ! \) on extended values in the following way:

\[
\begin{align*}
!(x^B) &= x^B^! \\
!(c^B) &= c^B^! \\
!(*)^n &= *^n^! \\
!(\lambda^x^B. M) &= \lambda^{n+1}x^B. M^! \\
!(\langle x^A, y^B \rangle)^n &= \langle x^A, y^B \rangle^{n+1}^{!} = V^{!} \quad \text{in } W
\end{align*}
\]

We call \( ! V \) the promotion of \( V \).

**Lemma 9.3.10.** Given a valid typing judgement \( ! \Delta \vdash V : A \), the typing judgement \( ! \Delta \vdash V : ! A \) is valid.

**Proof.** Proof by structural induction on the derivation of \( ! \Delta \vdash V : A \) \(\square\)
9.3. The Category $\mathcal{C}_\lambda$

**Lemma 9.3.11.** Consider an extended value $V$. Then $\text{Erase}(V) = \text{Erase}(V)$.

*Proof.* Proof by structural induction on $V$.

**Lemma 9.3.12.** If $\Delta \triangleright V \approx ax W : A$, then $\triangleright !V \approx ax !W : !A$

*Proof.* By structural induction on the derivation of $\Delta \triangleright V \approx ax W : A$. We fully develop the case $(\beta_\lambda)$.

Suppose that $\triangleright let x^A = V in W : B$ is valid. Then so is $\triangleright W[V/x] : B$. We have $!(let x^A = V in W) = (let x^A = V in !W)$. From Lemma 9.3.11

$$\triangleright x^A = !V in !W : !B$$

is valid. We therefore have $\triangleright !W[V/x] : !B$, and by rule $(\beta_\lambda)$, they are axiomatically equivalent.

Now, using Lemma 9.3.11 and Theorem 9.2.7, we have $\triangleright !(W[V/x]) \approx ax !W[V/x] : !B$. Therefore,

$$\triangleright !(let x^A = V in W) \approx ax !(W[V/x]) : !B.$$

The other cases are handled similarly.

**Lemma 9.3.13.** Given the map $!$ sending $A$ to $!A$ and sending $x : A \triangleright V : B$ to $x : !A \triangleright !V : !B$, and given the arrows

$$\epsilon_A = x : !A \triangleright x^A : A, \quad \delta_A = x : !A \triangleright x^{\triangleright A} : !^2A,$$

$(!, \epsilon, \delta)$ is an idempotent comonad in $\mathcal{C}_\lambda$.

*Proof.* We show the functoriality of $!$.

First, from Lemma 9.3.12 $!$ is a well-defined function on equivalence class of arrows.

Then, given $x : A \triangleright V : B$ and $y : B \triangleright W : C$ we need to show that

$$!(x : A \triangleright V : B); !(y : B \triangleright W : C) \approx ax !(x : A \triangleright let y = V in W : C).$$

Both hands end up being by definition of composition equal to

$$x : !A \triangleright let y^{\triangleright B} = !V in !W : !C.$$

Finally, the image $!id_A$ of the unit on $A$ is $!(x : A \triangleright x : A)$, which is by definition $x : !A \triangleright x : !A$, namely the unit on $!A$.

Using Lemmas 9.1.29 9.3.11 and Theorem 9.2.7, the equations

$$x : !A \triangleright \{V <: B\} \approx ax V : B,$$

$$x : !A \triangleright \{V <: !B\} \approx ax \{W : !!B\} : !!B$$

are valid, making $\epsilon$ and $\delta$ natural transformations.

The facts that $\delta_A$ is an isomorphism and that $\delta$ and $\epsilon$ verify Equation 2.6.1 and 2.6.2 use similar techniques.
Lemma 9.3.14. Given the maps

\[
\alpha_{A,B,C} = x : A \otimes (B \otimes C) \triangleright \text{let}(y,z) = x \text{in} \text{let}(t,u) = z \in \langle(y,t),u\rangle : (A \otimes B) \otimes C
\]

\[
\lambda_A = x : \top \otimes A \triangleright \text{let}(y,z) = x \text{in} \text{let} *= y \text{in} z : A
\]

\[
\rho_A = x : A \otimes \top \triangleright \text{let}(y,z) = x \text{in} \text{let} *= z \text{in} y : A
\]

\[
\sigma_{A,B} = x : A \otimes B \triangleright \text{let}(y,z) = x \text{in} \langle z,y \rangle : B \otimes A
\]

\[
\eta_A = x : A \triangleright \lambda *, x : \top \rightarrow A
\]

\[
\mu_A = x : \top \rightarrow (\top \rightarrow A) \triangleright \lambda *.(x*)* : \top \rightarrow A
\]

\[
\iota_{A,B} = z : A \otimes (\top \rightarrow B) \triangleright \text{let}(x,y) = z \text{in} \lambda *. (x,y) : \top \rightarrow (A \otimes B)
\]

\[
\varepsilon_A = x : !A \triangleright x^A : A
\]

\[
\delta_A = x : !A \triangleright x^A : !^2 A
\]

\[
d^A_{A,B} = z : !A \otimes !B \triangleright \text{let}(x,y) = z \text{in} \langle x,y \rangle : (!A \otimes B)
\]

\[
d^*_A = z : \top \triangleright \text{let} *= z \text{in} * : !\top
\]

\[
\triangle_A = x : !A \triangleright \langle x,x \rangle : !A \otimes !A
\]

\[
\Diamond_A = x : !A \triangleright * : \top
\]

\[
(x : A \triangleright V : B) \otimes (y : C \triangleright W : D) =
\]

\[
z : A \otimes B \triangleright \text{let}(x,y) = z \text{in} \langle V,W \rangle : C \otimes D
\]

\[
(x : A \triangleright V : B) \rightarrow (y : C \triangleright W : D) =
\]

\[
z : B \rightarrow C \triangleright \lambda x.\text{let}(y = zV \text{in} W) : A \rightarrow D
\]

\[
(x : A \triangleright V : \top \rightarrow B)^* =
\]

\[
y : \top \rightarrow A \triangleright \lambda *, \text{let } x = (y*) \text{in} (V*) : \top \rightarrow B
\]

\[
\Phi_{A,B,C}(x : A \triangleright V : B \rightarrow C) =
\]

\[
t : A \otimes B \triangleright \lambda *, \text{let } \langle x,y \rangle = t \text{in} V y : \top \rightarrow C
\]

Table 9.5: Definitions of maps and operations on maps in \(C_\lambda\)

9.3.4 The Category \(C_\lambda\) is a Linear Category for Duplication

Lemma 9.3.14. Given the maps

\[
d^A_{A,B} = z !A \otimes !B \triangleright \text{let}(x^A,y^B)^1 = z \text{ in} \langle x^A,y^B \rangle^1 : (!A \otimes B),
\]

\[
d^*_A = z : \top \triangleright \text{let} *= z^* \text{in} * : !\top,
\]

\((!,\delta,\epsilon,d,d)\) is a strong symmetric monoidal comonad in \(C_\lambda\).

Proof. The proof makes heavy use of Lemmas \(9.1.25\) and Theorem \(9.1.24\). \(\square\)

Lemma 9.3.15. If we define the maps

\[
\triangle_A = x : !A \triangleright \langle x^A,x^A \rangle^0 : !A \otimes !A,
\]

\[
\Diamond_A = x : !A \triangleright e^0 : \top,
\]

the functor \(\triangleright^\ast\) is a linear exponential comonad.

Proof. Again, the proof makes heavy use of Lemmas \(9.1.25\) and Theorem \(9.1.24\). \(\square\)

Theorem 9.3.16. Together with the maps and the operations on maps defined in Table \(9.5\), \(C_\lambda\) is a linear category for duplication.

Proof. This is a corollary of Lemmas \(9.1.14\) and Theorem \(9.1.14\). \(\square\)
Chapter 10

Proof of Theorem 9.2.7

In Chapter 9, we defined a typed lambda-calculus with explicit indexing. The type system is based on linear logic: a special type operator “!” is used to distinguish duplicable and non-duplicable terms. We recall Definitions 9.1.1 and 9.1.3 giving the set of terms and types:

\[
\text{Type } A, B ::= \alpha | (A \rightarrow B) | (A \otimes B) | \top | !A, \\
\text{CoreValue } U, U' ::= x^A | c^A | \ast^n | \lambda^nx^A.M | (U, U')^n, \\
\text{ExtValue } V, W ::= U | \langle V, W \rangle^n | \text{let } x^A = V \text{ in } W | \langle x^A, y^B \rangle^n = V \text{ in } W | \langle \ast^n = V \rangle, \\
\text{Term } M, N ::= U | \langle M, N \rangle^n | (MN) | \text{let } x^A, y^B \text{ in } M = V \text{ in } N | \text{let } \ast^n = M \text{ in } N.
\]

Terms are separated into core values, extended values and computations. We introduced the notion of valid typing judgement in Section 9.1.2 and we provided an axiomatic equivalence relation on valid typing judgement in Section 9.2. For the axiomatic equivalence relation a notion of substitution is required. This does not come for free due to the subtyping relation: an explicit typecasting is needed and is introduced in Definition 9.1.20: If a term \(M\) is valid with type \(A\) and if \(A <: B\), we inductively define a valid term \(\{M <: B\}\) of type \(B\).

One of the desired properties derived from the axiomatic equivalence relation is Theorem 9.2.7 stating that if any two terms are equal without their indexation, they are axiomatically equivalent. In other words, the meaning of a term is not related to the indexation of the term. This chapter is uniquely devoted to the proof of Theorem 9.2.7. The main technical obstacle to a direct proof of this theorem, is the possible “wild” use of the application inside any given term. In order to restrict this, we use the notion of “normalization by evaluation” of Filinski (2001) and Ohori (1999): we work with a subclass of terms that we call neutral terms on which the result can be proved directly. We then define two rewriting systems yielding in two steps neutral terms from any given terms. A weak normalization result is shown, adapting the techniques described in Girard et al. (1990, Ch. 4): we show that terms admit a well-founded measure of convergence, which decreases with each reduction of the rewriting system.

10.1 A Handy Tool: Neutral Terms

In this section we define the notion of neutral terms. We prove some useful results and then that any term is axiomatically equivalent to a neutral term.
Definition 10.1.1. Following Filinski [2001] and Ohori [1999], the notions of neutral value and neutral term are defined as follows (omitting indexes for readability when possible):

\[
\begin{align*}
NValue & \quad V, W ::= x \mid c \mid * \mid \lambda x.M \mid (V, W), \\
NTerm & \quad M, N ::= V \mid xV \mid cV \mid \text{let } \Box = xV \text{ in } M \mid \text{let } \Box = cV \text{ in } M \\
& \quad \text{let } (y, z) = x \text{ in } M \mid \text{let } (y, z) = c \text{ in } M \\
& \quad \text{let } * = x \text{ in } M \mid \text{let } * = c \text{ in } M,
\end{align*}
\]

where \( \Box \) means any of \( x, (x, y) \) or \( * \). A neutral typing judgement is a typing judgement \( \Delta \vdash M : A \) where \( M \) is a neutral term.

Lemma 10.1.2. Suppose that \( M \) is a neutral term and that \( \text{Erase}(M) = \text{Erase}(N) \). Then \( N \) is a neutral term.

Proof. Proof by structural induction on \( M \).

Lemma 10.1.3. For valid neutral terms \( \Delta \vdash M, M' : B \), if \( \text{Erase}(M) = \text{Erase}(M') \) then \( \Delta \vdash M \approx_{ax} M' : B \).

Proof. We prove it by induction on \( \hat{M} = \text{Erase}(M) \). Note that we also have \( \hat{M} = \text{Erase}(M') \). Base cases:

Cases \( \hat{M} \equiv x, c \) or \( * \). In each case, there is only one typing rule available for \( \Delta \vdash M : B \) and \( \Delta \vdash M' : B \), namely \((ax_1)\) in the first case, \((ax_2)\) in the second case and \((\top . I)\) in the last case. The terms \( M \) and \( M' \) are determined by the type \( B \), so \( M = M' \) and the result is correct by reflexivity.

Induction cases:

Case \( \hat{M} \equiv \lambda x.\hat{N} \). In this case, \( M \) is of the form \( \lambda^n x^A.N \) and \( M' \) is of the form \( \lambda^{n'} x^{A'} .N' \), where \( \hat{N} = \text{Erase}(N) = \text{Erase}(N') \). Since they are well-typed, from Lemma 9.1.12, the type \( B \) is the raw type of \( M \) and \( M' \) (as in Definition 9.1.8). This implies that \( A = A', n = n' \) and that \( B \) is equal to \( !^n(A \rightarrow C) \) for some type \( C \). Typing trees for \( \Delta \vdash M : B \) and \( \Delta \vdash M' : B \) start with a typing rule of the form \((\lambda_i)\), with \( i = 1 \) if \( n = 0 \), \( i = 2 \) otherwise, and we have the following valid typing judgements:

\[
\Delta, x : A \vdash N : C, \quad \Delta, x : A \vdash N' : C.
\]

By induction hypothesis, \( \Delta, x : A \vdash N \approx_{ax} N' : C \). By congruence, we have \( \Delta \vdash M \approx_{ax} M' : B \).

Case \( \hat{M} \equiv (\hat{U}, \hat{V}) \). \( M \) is of the form \( (N, P)^n \) and \( M' \) is of the form \( (N', P'^{n'}) \), where \( \hat{U} = \text{Erase}(N) = \text{Erase}(N') \) and \( \hat{V} = \text{Erase}(P) = \text{Erase}(P') \). Since \( M \) and \( M' \) are well-typed, from Lemma 9.1.12, the type \( B \) is the raw type of \( M \) and \( M' \). This implies that \( n = n' \) and that \( B \) is equal to \( !^n(A \otimes C) \) for some types \( A \) and \( C \). Typing trees for \( \Delta \vdash M : B \) and \( \Delta \vdash M' : B \) start with a typing rule of the form \((\otimes . I)\), and we have the following valid typing judgements:

\[
!\Delta, \Gamma_1 \vdash N : !^n A, \quad !\Delta', \Gamma'_1 \vdash N' : !^n A, \quad !\Delta, \Gamma_2 \vdash P : !^n C, \quad !\Delta', \Gamma'_2 \vdash P' : !^n C,
\]

where \( \Delta = (\hat{\Delta}, \Gamma_1, \Gamma_2) = (\hat{\Delta}', \Gamma'_1, \Gamma'_2) \).

From Lemma 9.1.14 it is possible to have \( \Gamma_1, \Gamma_2, \Gamma'_1 \) and \( \Gamma'_2 \) containing only non-duplicable variables. From Lemma 9.1.17 the sets \( |\Gamma_1| \cap |\Gamma_2| \) and \( |\Gamma'_1| \cap |\Gamma'_2| \) are empty. From Lemma 9.1.7,

\[
\text{FV}(N) = \text{FV}(N') \quad \text{and} \quad \text{FV}(P) = \text{FV}(P').
\]
This means that \( \Gamma_1 = \Gamma'_1, \Gamma_2 = \Gamma'_2 \) and \(!\Delta = !\Delta'\). We can then apply the induction hypothesis twice and deduce

\[
!\Delta, \Gamma_1 \triangleright N \triangleq \alpha N' : !^n A, \quad !\Delta, \Gamma_2 \triangleright P \triangleq \alpha P' : !^n C.
\]

Finally, using the congruence rule of the axiomatic equivalence relation, we get that \( \Delta \triangleright M \triangleq \alpha M' : B \).

**Case** \( \hat{M} \equiv x\hat{V} \). Since \( M \) and \( M' \) are well-typed, from Lemma 9.1.24 the type \( B \) is the raw type of \( M \) and \( M' \). This implies that \( M \) is of the form \( x^{A_1 \to B} V_1 \) and \( M' \) is of the form \( x^{A_2 \to B} V_2 \), for some neutral values \( V_1 \) and \( V_2 \) and some types \( A_1 \) and \( A_2 \). The typing judgements \( \Delta \triangleright M, M' : B \) comes from the typing rule (\( \text{app} \)): the variable \( x \) occurs in \( \Delta \). The context \( \Delta \) can thus be split into \( (\hat{\Delta}, x : !^n(A' \to B')) \) for some integer \( n \) where, from Lemma 9.1.16, \( A_1 \triangleq A', A_2 \triangleq A' \) and \( B' \triangleq B \).

Using Lemma 9.1.19 one can construct the typing derivations \( \pi \) and \( \pi' \) of the form

\[
\frac{x : !^n(A' \to B') \triangleright x^{A_1 \to B} (ax_1) \quad \hat{\Delta}_1 \triangleright V_i : A_i}{\hat{\Delta}, x : !^n(A' \to B') \triangleright x^{A_1 \to B} V_i : B} \quad (\text{app})
\]

where

\[
\hat{\Delta}_i = \begin{cases} 
\hat{\Delta}, x : !^n(A'' \to B') & \text{if } x \in FV(V_i), \\
\hat{\Delta} & \text{otherwise},
\end{cases}
\]

for \( i = 1, 2 \). From Lemma 9.1.7 \( FV(V_1) = FV(V_2) \). This makes \( \hat{\Delta}_1 = \hat{\Delta}_2 \). Call it \( \hat{\Delta}' \). From Lemma 9.1.20 \( \Delta' \triangleright \{V_1 \triangleq A'\} : A' \) and \( \hat{\Delta}' \triangleright \{V_2 \triangleq A'\} : A' \) are valid. Applying induction hypothesis, they are axiomatically equivalent.

Finally, applying axiomatic equivalence rule (\( \text{app}_{\triangleq} \)), the axiomatic equivalence

\( \Delta \triangleright x^{A_1 \to B} V_i \triangleq \alpha x^{A' \to B} \{V_i \triangleq A'\} : B \),

is valid for \( i = 1, 2 \). This makes \( \Delta \triangleright M \triangleq \alpha M' : B \) valid by congruence and by transitivity.

**Case** \( \hat{M} \equiv c\hat{V} \). This case is similar to the previous case.

Since \( M \) and \( M' \) are well-typed, from Lemma 9.1.12 the type \( B \) is the raw type of \( M \) and \( M' \). This implies that \( M \) is of the form \( c^{A_1 \to B} V_1 \) and \( M' \) is of the form \( c^{A_2 \to B} V_2 \), for some neutral values \( V_1 \) and \( V_2 \) and some types \( A_1 \) and \( A_2 \) such that \( !A_c \triangleq A_1 \to B \) and \( !A_c \triangleq A_2 \to B \). The type \( A_c \) is therefore of the form \( A' \to B' \), with \( A_1 \triangleq A', A_2 \triangleq A' \) and \( B' \triangleq B \).

The typing judgements \( \Delta \triangleright M, M' : B \) comes from the typing rule (\( \text{app} \)). Using Lemma 9.1.19, one can construct the typing derivation

\[
\frac{\triangleright c^{A_1 \to B} (ax_1) \quad \Delta \triangleright V_i : A_i}{\Delta \triangleright c^{A_1 \to B} V_i : B} \quad (\text{app})
\]

for \( i = 1, 2 \). From Lemma 9.1.20 \( \Delta \triangleright \{V_1 \triangleq A'\} : A' \) and \( \Delta \triangleright \{V_2 \triangleq A'\} : A' \) are valid. Applying induction hypothesis, they are axiomatically equivalent.

Finally, applying axiomatic equivalence rule (\( \text{app}_{\triangleq} \)), the axiomatic equivalence

\( \Delta \triangleright c^{A_1 \to B} V_i \triangleq \alpha c^{A' \to B} \{V_i \triangleq A'\} : B \),

is valid for \( i = 1, 2 \). This makes \( \Delta \triangleright M \triangleq \alpha M' : B \) valid by congruence and by transitivity.
Case $\bar{M} \equiv (let \quad = \quad \bar{V} \ in \ P)$. The placeholder $\square$ stands for $\,*$, $y$ or $\langle y, z \rangle$, and the placeholder $\bar{\square}$ for $x$ or $c$. By $\alpha$-equivalence one can assume that when relevant, the variables $y, z$ are different from $x$ and are not free in $\bar{V}$.

The terms $M$ and $M'$ are well-typed. From Lemma 10.1.2 the type $B$ is the raw type of $M$ and $M'$. One can conclude that $M = (let \quad = \quad \bar{V}_1 \ in \ P_1)$ and $M' = (let \quad = \quad \bar{V}_2 \ in \ P_2)$, where $V = \text{Erase}(V_1) = \text{Erase}(V_2)$, $P = \text{Erase}(P_1) = \text{Erase}(P_2)$, and where $(i$ being $1, 2)$

<table>
<thead>
<tr>
<th>$\square$ is...</th>
<th>*</th>
<th>*</th>
<th>$\uparrow$</th>
<th>$\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>$y^{c_1}$</td>
<td>$(y^{c_1}, z^{D_1})^{\gamma_1}$</td>
<td>$C_i$</td>
<td>$\uparrow^{\gamma_1}(C_i \otimes D_i)$</td>
</tr>
</tbody>
</table>

for some types $C_i$ and $D_i$ and for some integers $n_i$.

In the case $\square = x$, the variable $x$ is free in $M$ and in $M'$. From Lemma 10.1.4 it is in $\Delta$. From Lemma 10.1.10 the type of $x$ in $\Delta$ is of the form $\uparrow^n(A' \to B')$, with $A_1 < : A'$, $A_2 < : A'$, $B' < : B_1$ and $B' < : B_2$.

In the case $\square = c$, we have $\uparrow n_{A_c} < : (A_2 \to B_1)$ and $\uparrow n_{A_c} < : (A_2 \to B_2)$. Let $\uparrow^n(A' \to B')$ be $A_c$. We then have $\uparrow A_1 < : A'$, $A_2 < : A'$, $B' < : B_1$ and $B' < : B_2$.

In both cases, this means that if $\square$ is $\uparrow^n$, then $B'$ is $\{\top^c \quad \}$ where $C' < : C_i$, $D' < : D_i$, and $n' \geq n_i$. Using respectively the identity, the axiomatic equivalence rule ($let^{<}_{<}$) and the rule ($let^{<}_{<}$), we get that

$$\Delta \triangleright let \quad = \quad \bar{V}_i \ in \ P_i \approx_{ax} let \quad = \quad \bar{V}_i \ in \ P_i : B, \quad (10.1.1)$$

such that if $\square$ is $\uparrow^n$, then $\square$ is $\{\top^c \quad \}$.

One can split $\Delta$ as $(\bar{\Delta}, \Gamma_1, \Gamma_2)$ such that $\Gamma_1$ contains the free variables of $\bar{V}$ that are not duplicable and $\Gamma_2$ contains the free variables of $\bar{P}$ that are not duplicable: from Lemma 10.1.4 and $\Gamma_2$ do not intersect. Invoking Lemma 10.1.7 and Lemma 10.1.9 for $i = 1, 2$ one can write a typing derivation

$$\begin{align*}
\vdots & \\
\bar{\Delta}, \Gamma_1 \triangleright & \quad \bar{V}_i : B' \\
\bar{\Delta}, \Gamma_2 \triangleright & let \quad = \quad \bar{V}_i \ in \ P_i : B \\
\bar{\Delta}, \Gamma_1, \Gamma_2 \triangleright & let \quad = \quad \bar{V}_i \ in \ P_i : B
\end{align*} \quad (X)$$

where if $\square$ is $\{\top^c \quad \}$ then $(X)$ is $\{\top^c \quad \}$ and $\Lambda$ is $\{\top^c \quad \}$.

One can apply the induction hypothesis on each branch of the derivation. Using the congruence rules of the axiomatic relation and Equation (10.1.1) we obtain finally that $\Delta \triangleright M \approx_{ax} M' : B$.

Case $\bar{M} \equiv (let \quad = \quad \bar{P})$. The placeholder $\square$ stands for $\,*$ or $\langle y, z \rangle$, and the placeholder $\bar{\square}$ for $x$ or $c$. By $\alpha$-equivalence one can assume that when relevant, the variables $y, z$ are different from $x$. 

10.1. A Handy Tool: Neutral Terms
The terms $M$ and $M'$ are well-typed. From Lemma 9.1.12 the type $B$ is the raw type of $M$ and $M'$. One can conclude that $M = (\text{let } x = B_1 \text{ in } P_1)$ and $M' = (\text{let } x = B_2 \text{ in } P_2)$, where $P = \text{Erase}(P_1) = \text{Erase}(P_2)$ and where (i being 1, 2)

\[
\begin{array}{c|c|c|c}
 & \text{if } \Box \text{ is...} & \text{is...} & \text{is...} \\
\hline
\star & y & y^{C_i} & \downarrow^{C_i} \\
\langle y, z \rangle & \langle y^{C_i}, z^{D_i} \rangle^{n_i} & \uparrow^{n_i}(C_i \otimes D_i) \\
\end{array}
\]

for some types $C_i$ and $D_i$ and for some integers $n_i$.

In the case $\Box = x$, the variable $x$ is free in $M$ and in $M'$. From Lemma 9.1.11 it is in $\Delta$. From Lemma 9.1.15 the type $A'$ of $x$ in $\Delta$ is such that $A' :<: B_1$ and $A' :<: B_2$.

In the case $\Box = c$, we have $!A_c :<: B_1$ and $!A_c :<: B_2$. Let $A'$ be $!A_c$.

In both cases, this means that if $\Box$ is \(\begin{cases} \star \\ y \langle y, z \rangle \end{cases}\) then $A'$ is \(\begin{cases} \top \\ C' \\ \uparrow^{n'_i}(C'_i \otimes D'_i) \end{cases}\) for some types $C_i$ and $D_i$ and for some integers $n_i$.

Using respectively the identity, the axiomatic equivalence rule (\text{let}^=_{<}) and the rule (\text{let}^\otimes_{<}) we get that

\[
\Delta \triangleright \text{let } x = \Box^{B_i} \text{ in } P_i \approx_{\text{ax}} \text{let } x = \Box^{A'} \text{ in } P_i : B,
\]

such that if $\Box$ is \(\begin{cases} \star \\ y \langle y, z \rangle \end{cases}\) then $\Box$ is \(\begin{cases} \star \\ y^{C'_i} \langle y^{C'_i}, z^{D'_i} \rangle^{n'_i} \end{cases}\).

Invoking Lemma 9.1.16 for $i = 1, 2$ one can write a typing derivation

\[
\begin{array}{c}
\Gamma \triangleright \Box^{A'} : A' \\ \Delta \triangleright \text{let } x = \Box^{B_i} \text{ in } P_i : B \end{array} \quad \Delta \triangleright \text{let } x = \Box^{A'} \text{ in } P_i : B \quad (X)
\]

where, if $\Box = c$, $\Gamma = ()$ and if $\Box = x$, $\Gamma = (x : A')$.

One can apply the induction hypothesis on each branch of the derivation. Using the congruence rules of the axiomatic relation and Equation (10.1.2) we obtain finally that $\Delta \triangleright M \approx_{\text{ax}} M' : B$.

This ends the list of cases and proves Lemma 10.1.3.

\[ \square \]

10.2 Term Rewriting System Number One

We define a rewriting system of valid terms yielding neutral terms. We want the denotation to stay constant through the rewriting. The rewriting will be two-fold. First we get rid of “complex computations”, and then we get rid of beta-reductions. Note that we are not interested in confluence: the only thing we want is normalization.
Definition 10.2.1. We define the notion of intermediate neutral value and of intermediate neutral term as follows:

\[
\begin{align*}
INValue \ V, W & \ ::= \ x^A \ | \ c^A \ | \ *^n \ | \ \lambda^n x^A.M \ | \ \langle V, W \rangle^n, \\
INTerm \ M, N & \ ::= \ V \ | \ x^{A \rightarrow B}V \ | \ c^{A \rightarrow B}V \ | \ \text{let } \boxed{=} = M \ in \ N,
\end{align*}
\]

where \boxed{=} is a placeholder for any of \(*, x^A\) or \(\langle x^A, y^B \rangle^n\).

Lemma 10.2.2. A neutral term is in intermediate neutral form.

Proof. The proof is done by inspection of the cases. \(\square\)

Remark 10.2.3. A term in intermediate neutral form is not automatically in neutral form. The two main examples of terms that are intermediate neutral but not neutral are

\[
\begin{align*}
\text{let } x^A &= (\text{let } y^B = M \ in \ N) \ in \ P, \quad \text{let } x^A = V \ in \ M,
\end{align*}
\]

where \(V\) is a core value.

Lemma 10.2.4. Suppose that \(\Delta \triangleright M, M' : B\) are valid typing judgements. Suppose moreover that \(\text{Erase}(M) = \text{Erase}(M')\) and that \(M\) is an intermediate neutral term (respectively value). Then \(M'\) is also an intermediate neutral form (respectively value).

Proof. Proof by induction on the size of \(M\). \(\square\)

Lemma 10.2.5. Suppose that \(\Delta = (\Delta', \Gamma_1, \Gamma_2)\), that \(\Gamma' \triangleright V : A\) and \(\Gamma_1, x : A \triangleright M : B\). If \(V\) is an intermediate neutral value and if \(M\) is an intermediate neutral term (respectively intermediate neutral value), then \(M[V/x]\) is an intermediate neutral term (respectively an intermediate neutral value).

Proof. Proof by induction on the size of \(M\). The interesting case is when \(M\) is of the form \(x^{A \rightarrow B}W\): in this situation, the value \(V\) is of type \(A \rightarrow B\), that is, of the form \(\lambda^n y^A.N\) where \(N\) is an intermediate neutral term of type \(B\). Then \(M[V/x] = (\lambda^n y^A.N)W\) which is by definition (let \(y^A = V \ in \ N\)), an intermediate neutral term. Note that in this case, \(M\) cannot be a value. \(\square\)

Definition 10.2.6. We define the rewrite system number one of terms as the reflexive closure of the relation defined as follows. Given \(x, y\) fresh variables, provided that \(MN : B\), that \(M : A \rightarrow B\) is not a lambda-abstraction, and that \(M\) and \(N\) are not both term variables,

\[
MN \rightarrow_{1} \text{let } x^{A \rightarrow B} = M \ in \ let \ y^A = N \ in \ x^{A \rightarrow B} y^A.
\]

Provided that \(\langle M, N \rangle^n : !^n (A \otimes B)\), and that \(M\) and \(N\) are not both term variables,

\[
\langle M, N \rangle^n \rightarrow_{1} \text{let } x^{n\ A} = M \ in \ let \ y^{n\ B} = N \ in \ \langle x^{n\ A}, y^{n\ B} \rangle^n.
\]

Finally, we add one congruence rule for each term constructor in the following way: for each rule, \(M\) and \(N\) are terms such that \(M \rightarrow_1 M'\) and \(N \rightarrow_1 N'\).

\[
\begin{align*}
\lambda^n x^A.M & \rightarrow_1 \lambda^n x^A.M', \\
MN & \rightarrow_1 M' N', \\
\langle M, N \rangle^n & \rightarrow_1 \langle M', N' \rangle^n, \\
\text{let } \langle x^A, y^B \rangle^n & = M \ in \ N \rightarrow_1 \text{let } \langle x^A, y^B \rangle^n = M' \ in \ N', \\
\text{let } * & = M \ in \ N \rightarrow_1 \text{let } * = M' \ in \ N'.
\end{align*}
\]
Lemma 10.2.7. Suppose that $\Delta \triangleright M : C$ is a valid typing judgement, where $M \rightarrow_1 M'$. Then $\Delta \triangleright M' : C$ is valid and $\Delta \triangleright M \simeq_{ax} M' : C$.

Proof. The proof is done by induction of the reduction $M \rightarrow_1 M'$.

If reflexivity was used. Then $M = M'$ and the result is immediate.

Case $\triangleright_2 1$. Here, $M \equiv NP$. From Lemma 10.2.2, $B = C$. Suppose that $\Delta \triangleright NP : B$, where $N : A \rightarrow B$. Then one can split $\Delta$ into $(\Delta_1, \Gamma_1, \Gamma_2)$ where

\[
\begin{align*}
\pi_1 : & !\Delta_1, \Gamma_1 \triangleright N : A \rightarrow B \\
\pi_2 : & !\Delta_2, \Gamma_2 \triangleright P : A
\end{align*}
\]

is a valid typing derivation. Then if $x$ and $y$ are fresh variables, the derivation

\[
\begin{align*}
\pi_1 & : x : A \rightarrow B \triangleright x^A \rightarrow B \\
\pi_2 & : y : A \triangleright y^A : A
\end{align*}
\]

\]

is valid. The typing judgements $\Delta \triangleright NP : B$ and $\Delta \triangleright \text{let } x^A \rightarrow B = N \text{ in } y^A = P \text{ in } x^A \rightarrow B y^A : B$ are axiomatically equivalent by rule $(\triangleright_2, ax)$.}

Proof. The proof is done by induction of the reduction $M \rightarrow_1 M'$.

If reflexivity was used. Then $M = M'$ and the result is immediate.

Case $\triangleright_2 1$. Here, $M \equiv \langle N, P \rangle^n$. From Lemma 10.2.1, $C = !^n (A \otimes B)$. One can split $\Delta$ into $(\Delta_1, \Gamma_1, \Gamma_2)$ where

\[
\begin{align*}
\pi_1 & : !\Delta_1, \Gamma_1 \triangleright N : !^n A \\
\pi_2 & : !\Delta_2, \Gamma_2 \triangleright P : !^n B
\end{align*}
\]

is a valid typing derivation. Then if $x$ and $y$ are fresh variables, the derivation

\[
\begin{align*}
\pi_1 & : x : !^n A \triangleright x^A : !^n A \\
\pi_2 & : y : !^n B \triangleright y^A : !^n B
\end{align*}
\]

\]

is valid. The two typing judgements are axiomatically equivalent by rule $(\triangleright_2, ax)$.}

Proof. The proof is done by induction of the reduction $M \rightarrow_1 M'$.

If reflexivity was used. Then $M = M'$ and the result is immediate.

Case $\triangleright_2 1$. We have $M \equiv \lambda^n x^A N$. Suppose that $\Delta \triangleright \lambda^n x^A N : C$. Then $C = !^n (A \rightarrow B)$ for some type $B$, the judgement comes from $\Delta, x : A \triangleright N : B$ using rule $(\lambda_i)$, where $i = 1$ if $n = 0$, $i = 2$ otherwise. By induction hypothesis, if $N \rightarrow_1 N'$ then $\Delta, x : A \triangleright N' : B$ is valid. Then, if $n$ is non null, $\Delta$ is duplicable, and one can apply $(\lambda_2)$ to get $\Delta \triangleright \lambda^n x^A N' : C$. Otherwise, $n = 0$ and we apply $(\lambda_1)$ to get $\Delta \triangleright \lambda^n x^A N' : C$.

By induction hypothesis, $\Delta, x : A \triangleright N \simeq_{ax} N' : B$. Therefore, by congruence $\Delta \triangleright \lambda^n x^A N \simeq_{ax} \lambda^n x^A N' : C$. 


Cases \( \text{rw1.c1} \), \( \text{rw1.c2} \), \( \text{rw1.c3} \) and \( \text{rw1.c4} \). We have \( M = f(N, P) \), where \( f(N, P) \) is respectively \( NP \), \( \langle N, P \rangle \), \( \langle x^A, y^B \rangle \) and \( \langle x^A, y^B \rangle \). Suppose that \( \Delta \vdash f(N, P) : C \). Then \( \Delta \) split into \( (\Delta, \Gamma_1, \Gamma_2) \) and the typing judgement is originated from the typing rule \((X)\) and the judgements \(!\Delta, \Gamma_1 \vdash N : C_1 \) and \(!\Delta, \Gamma_2, \Lambda \vdash P : C_2 \), where

\[
\begin{array}{c|c|c|c|c}
\text{if } f(N, P) \text{ is } & (X) \text{ is } & C_1 \text{ is } & C_2 \text{ is } & C \text{ is } \\
NP & (\text{app}) & A \rightarrow B & A & B \\
\langle N, P \rangle \alpha & (\otimes. I) & \alpha^\top A & \alpha^\top B & \alpha^\top (A \otimes B) \\
\text{let } * = N \text{ in } P & (\top. E) & \top & B & B \\
\text{let } \langle x^A, y^B \rangle \alpha = N \text{ in } P & (\otimes. E) & \alpha^\top (A \otimes B) & D & D,
\end{array}
\]

with some types \( A, B \) and \( D \), and such that \( \Lambda \) is empty in the three first cases and equal to \( (x : \alpha^\top A, y : \alpha^\top B) \) in the last case.

By induction hypothesis, if \( N \rightarrow N' \) and \( P \rightarrow P' \), then \( !\Delta, \Gamma_1 \vdash N' : C_1 \) and \( !\Delta, \Gamma_2, \Lambda \vdash P' : C_2 \) are well-typed. Then, applying \((X)\) we get that \( \Delta \vdash f(N', P') : C \) is valid.

By induction hypothesis again, \( !\Delta, \Gamma_1 \vdash N \approx_{ax} N' : C_1 \) and \( !\Delta, \Gamma_2, \Lambda \vdash P \approx_{ax} P' : C_2 \). Therefore, by congruence \( \Delta \vdash f(N, P) \approx_{ax} f(N', P') : C \).

This closes the proof of Lemma 10.2.8.

\[ \square \]

Lemma 10.2.8. Suppose that \( M, N : C \) and that \( \text{Erase}(M) = \text{Erase}(N) \). If \( M \rightarrow M' \), then there exists \( N' \) such that \( \text{Erase}(M') = \text{Erase}(N') \) and such that \( N \rightarrow N' \).

Proof. The proof is done by induction on the reduction \( M \rightarrow M' \).

If reflexivity was used. Then \( M = M' \) and the result is immediate by setting \( N = N' \).

Case \( \text{rw1.c1} \). The term \( M \) is of the form \( M_1 M_2 \), where \( M_1 : A \rightarrow B \) and where \( M_1 M_2 : B \). In this case, \( M' = (\text{let } x^A \rightarrow B = M_1 \text{ in } y^A = M_2 \text{ in } x^A \rightarrow B y^A) \). If \( N \) is such that \( \text{Erase}(M) = \text{Erase}(N) \), then \( N = N_1 N_2 \), where \( \text{Erase}(N_1) = \text{Erase}(M_1) \) and \( \text{Erase}(N_2) = \text{Erase}(M_2) \).

Since the raw type of \( N \) is \( B \), there exists a term \( A' \) such that \( N_1 : A' \rightarrow B \). If we define \( N' = (\text{let } x^A \rightarrow B = N_1 \text{ in } y^A = N_2 \text{ in } x^A \rightarrow B y^A) \), then by reduction rule \( \text{rw1.c1} \), we have \( N \rightarrow N' \), and by definition of the map \( \text{Erase} \), the terms \( M' \) and \( N' \) have the same erasure.

Case \( \text{rw1.c2} \). The term \( M \) is of the form \( \langle M_1, M_2 \rangle \), where \( \langle M_1, M_2 \rangle : \alpha^\top (A \otimes B) \). In this case, \( \langle M_1, M_2 \rangle = (\text{let } x^A = M_1 \text{ in } y^B = M_2 \text{ in } \langle x^A, y^B \rangle) \). If \( N \) is such that \( \text{Erase}(M) = \text{Erase}(N) \), then \( N = \langle N_1, N_2 \rangle \), where \( \text{Erase}(N_1) = \text{Erase}(M_1) \) and \( \text{Erase}(N_2) = \text{Erase}(M_2) \).

Since \( N : \alpha^\top (A \otimes B) \), the raw type of \( N \) is \( \alpha^\top (A \otimes B) \). If we define \( N' = (\text{let } x^A \rightarrow B = N_1 \text{ in } y^A = N_2 \text{ in } \langle x^A, y^B \rangle) \), then by reduction rule \( \text{rw1.c2} \), we have \( N \rightarrow N' \), and by definition of the map \( \text{Erase} \), the terms \( M' \) and \( N' \) have the same erasure.

Case \( \text{rw1.c3} \). The term \( M \) is of the form \( \lambda^x A. M \). In this case, \( M' = \lambda^x A. M' \), where \( M \rightarrow M' \).

Since \( \text{Erase}(M) = \text{Erase}(N) \), the term \( N \) is also of the form \( \lambda^x A. N \), with \( \text{Erase}(M) = \text{Erase}(N) \). By induction hypothesis there exists \( N' \) such that \( N \rightarrow N' \). Let \( N' = \lambda^x A. N' \). One has \( \text{Erase}(N) = \text{Erase}(N') \). Using congruence rule \( \text{rw1.c3} \), we deduce that \( N \rightarrow N' \).

Cases \( \text{rw1.c4} \), \( \text{rw1.c5} \), \( \text{rw1.c6} \) and \( \text{rw1.c7} \). In these cases, the term \( M \) is equal to \( f(M_1, M_2) \), where \( f(M_1, M_2) \) is respectively the form \( M_1 M_2 \), \( (M_1, M_2) \), \( (x^A, y^B) \), \( M_1 \) in \( M_2 \), and of the form \( \langle x^A, y^B \rangle = M_1 \), with \( M_1 \rightarrow M_1' \) and \( M_2 \rightarrow M_2' \). We have that \( M \rightarrow M' \) where \( M' = f(M_1', M_2') \).
Since $\text{Erase}(M) = \text{Erase}(N)$, the term $N = g(N_1, N_2)$, where $g(N_1, N_2)$ is respectively of the form $N_1N_2$, $(N_1, N_2)^m$, (let $* = N_1$ in $N_2$) and (let $(x^A, y^B)^m = N_1$ in $N_2$).

By induction hypothesis, $N_1 \rightarrow_1 N_1'$ and $N_2 \rightarrow_2 N_2'$ such that $\text{Erase}(N_1') = \text{Erase}(M_1')$ and $\text{Erase}(N_2') = \text{Erase}(M_2')$. This means, using respectively rules (rw1.ξ), (rw1.ξ), (rw1.ξ) and (rw1.ξ), that $g(N_1, N_2) \rightarrow_1 g(N_1', N_2')$.

Since $\text{Erase}(g(N_1', N_2')) = \text{Erase}(f(M_1', M_2'))$, we can define $N' = g(N_1', N_2')$ and we have that $\text{Erase}(M') = \text{Erase}(N')$ and $N \rightarrow_1 N'$.

This closes the proof of Lemma 10.2.9.

**Definition 10.2.9.** We define the non-negative integer measure $m(M)$ of a typed term $M$ by induction on $M$ as follows (we omit indices for legibility):

- $m(x) = 0$,
- $m(c) = 0$,
- $m(*) = 0$,
- $m(xy) = 0$,
- $m((x,y)) = 0$.

In the case $m(NP)$, we assume that $N$ is not a lambda-abstraction. In the cases $m(NP)$ and $m((N,P))$, we assume that $N$ and $P$ are not both term variables.

**Lemma 10.2.10.** Suppose that $M \rightarrow_1 M'$. Then $M$ is a lambda abstraction if and only if $M'$ is.

**Proof.** The proof is done by case inspection: there are only two possible reductions of the form $\lambda^a.A.N \rightarrow M'$, namely (rw1.ξ), and the identity. In both cases, $M'$ is of the form $\lambda^a.A.N'$.

Similarly, the only rules of the form $M \rightarrow_1 \lambda^a.A.N'$ are (rw1.ξ), and the trivial one. Again, $M$ is of the form $\lambda^a.A.N$ in both cases.

**Lemma 10.2.11.** Suppose that $M \rightarrow_1 M'$ where $M \neq M'$. Then $m(M') < m(M)$.

**Proof.** Proof by induction on the reduction. Note that reflexivity is not a possibility since we want $M \neq M'$.

- **Case (rw1.ξ).** In this case, $M' = \lambda^a.A. = N \in \lambda^a.A.B$ and $y^B = P$ for some types $A$ and $B$, where $N$ is not a lambda-abstraction and $N$ and $P$ are not both term variables. Since $m(M) = 1 + m(N) + m(P)$ and $m(M') = m(N) + m(P)$, the strict inequality $m(M') < m(M)$ is valid.

- **Case (rw1.ξ).** In this case, $M' = \lambda^a.A. = N \in \lambda^a.A. = P$ and $y^B = P$ for some types $A$ and $B$, where $N$ and $P$ are not both term variables. Since $m(M) = 1 + m(N) + m(P)$ and $m(M') = m(N) + m(P)$, the inequality $m(M') < m(M)$ is valid.

- **Case (rw1.ξ).** In this case, $M = \lambda^a.A.B$ and $M' = \lambda^a.A.B$ where $N \rightarrow_1 N'$. By induction hypothesis, $m(N') < m(N)$. By definition, we then have that $m(M') = m(N') < m(N) = m(M)$.
10.2. Term Rewriting System Number One 157

Case [rw1.ξ1]. We have $M = NP$, where $N \rightarrow N'$ and $P \rightarrow P'$. The terms $N$ and $P$ therefore cannot be both a term variable. By induction hypothesis, $m(N') < m(N)$ or $N' = N$, and $m(P') < m(P)$ or $P' = P$, but equality do not hold for both.

If $N$ is a lambda abstraction, $m(NP) = m(N) + m(P)$. Using Lemma 10.2.10, the term $N'$ is a lambda abstraction: we have $m(N'P') = m(N') + m(P')$. One can conclude that $m(M') = m(N') + m(P') < m(N) + m(P) = m(M)$.

If $N$ is not a lambda abstraction, then we have $m(NP) = 1 + m(N) + m(P)$. By Lemma 10.2.10 $N'$ is not a lambda-abstraction: $m(N'P') = 1 + m(N') + m(P')$. Therefore $m(M') = 1 + m(N') + m(P') < 1 + m(N) + m(P) = m(M)$.

Case [rw1.ξ2]. We have $M = \langle N, P \rangle^n$, where $N \rightarrow N'$ and $P \rightarrow P'$. The terms $N$ and $P$ therefore cannot be both a term variable. By induction hypothesis, $m(N') < m(N)$ or $N' = N$, and $m(P') < m(P)$ or $P' = P$ (but equality do not hold for both).

We have $m(\langle N, P \rangle^n) = 1 + m(N) + m(P)$ and $m(\langle N', P' \rangle^n) = 1 + m(N') + m(P')$. This is enough to say that $m(M') = 1 + m(N') + m(P') < 1 + m(N) + m(P) = m(M)$.

Cases [rw1.ξ4] and [rw1.ξ5]. We have $M$ respectively of the form (let $* = N$ in $P$) and of the form (let $(x^A, y^B)^n = N$ in $P$), with $N \rightarrow N'$ and $P \rightarrow P'$. We have that $M \rightarrow M'$ where $M'$ is respectively (let $* = N'$ in $P$) and (let $(x^A, y^B)^n = N'$ in $P'$)

By induction hypothesis, $m(N') < m(N)$ or $N' = N$, and $m(P') < m(P)$ or $P' = P$ (but not $N' = N$ and $P' = P$). Since in both cases, $m(M) = m(N) + m(P)$ and $m(M') = m(N') + m(P')$, we have $m(M') < m(M)$.

This ends the induction, and proves the lemma. \[\square\]

Lemma 10.2.12. Rewrite system number one is normalizing: For any term $M$ and any sequence $(M_i)_i$ such that $M = M_0$ and such that for all $i \geq 0$, $M_i \rightarrow M_{i+1}$, there exists a number $n$ such that for all $i \geq n$, $M_i = M_{i+1}$.

Proof. We prove the lemma by contradiction. Suppose there exists a term $M$ and a sequence $(M_i)_i$ such that $M = M_0$ and such that for all $i \geq 0$, $M_i \rightarrow M_{i+1}$ and such that for all $n$ there exists a number $n' > n$ such that $M_{n'} \neq M_n$. Define a sequence $(n(i))$, as follows. First, $n(0) = 0$. Then, for all $i \geq 0$, define $n(i+1)$ as the first number $n' > n(i)$ such that $M_{n'} \neq M_{n(i)}$. Note that in this case, for all $i$ we have $M_{n(i)} = M_{n(i)+1}$ and $M_{n(i)+1} \neq M_{n(i+1)}$. By Lemma 10.2.11, this means that $m(M_{n(i)+1}) \leq m(M_{n(i)}) - 1$. We prove by induction on $i$ that for all $i$, $m(M_{n(i)}) \leq m(M) - i$:

Case $i = 0$. In this case, $M_{n(0)} = M_0 = M$, therefore $m(M_{n(0)}) = m(M) - 0$.

Case $i > 0$. Suppose that $m(M_{n(i)}) \leq m(M) - i$. Since $m(M_{n(i)+1}) < m(M_{n(i)+1} - 1)$, we have $m(M_{n(i)+1}) \leq m(M_{n(i)}) - 1 \leq m(M) - i - 1 \leq m(M) - (i + 1)$.

Therefore, by induction on $i$, $m(M_{n(i)}) \leq m(M) - i$. But for $i > m(M)$, the value $m(M_i)$ is negative, contradicting the non-negativity of the measure. Therefore there exists a number $n$ such that for all $i \geq n$, $M_i = M_{i+1}$.

\[\square\]

Lemma 10.2.13. Suppose that $\Delta \triangleright M : B$ is valid and that for all $M'$ such that $M \rightarrow M'$, $M = M'$. Then $M$ is in intermediate neutral form.

Proof. We prove the lemma by induction on the size of $M$.

Case $M \equiv x^A$, $e^A$ and $*^n$. Then $M$ is in intermediate neutral form.
10.3. Term Rewriting System Number Two

Case $M \equiv \lambda^n x^A. N$. The only applicable rule is \texttt{rw1.ξ1}, meaning that the set of possible $M'$ is the set of terms of the form $\lambda^n x^A. N'$, where $N \to_1 N'$. Since $M = M'$, we have $N = N'$. By induction hypothesis, $N$ is in intermediate neutral form. This makes $M$ in intermediate neutral form.

Case $M \equiv NP$. There are three possible subcases:

Case $N$ and $P$ are both term variables. Then $M$ is in intermediate neutral form.

Case $N \equiv \lambda^n x^A. \bar{N}$. The only applicable rewriting rule is \texttt{rw1.ξ2}. Since $\Delta \triangleright M : B$ is valid, $n = 0$ and $M \equiv \text{let } x^A = P \text{ in } \bar{N}$.

In this case, the set of $M'$ such that $M \to_1 M'$ is spanned by all the terms of the form $\text{let } x^A = P' \text{ in } \bar{N}'$, when $\bar{N} \to_1 \bar{N}'$ and $P \to_1 P'$. Since $M' = M$, one has $\bar{N}' = \bar{N}$ and $P' = P$. By induction hypothesis, this means that $\bar{N}$ and $P$ are in intermediate neutral form.

The term $M$ is thus in intermediate neutral form.

Otherwise. It is always possible to apply rule \texttt{rw1.ξ1}, in which case $M'$ is not equal to $M$. This case is not to be considered.

Case $M \equiv \langle N, P \rangle^n$. There are two possible subcases:

Case $N$ and $P$ are both term variables. Then $M$ is in intermediate neutral form.

Otherwise. It is always possible to apply rule \texttt{rw1.ξ2}, in which case $M'$ is not equal to $M$. This case is not to be considered.

Cases $M \equiv (\text{let } * = N \text{ in } P)$ and $(\text{let } \langle x^A, y^B \rangle^n = N \text{ in } P)$. For these two cases, the only applicable rules are respectively \texttt{rw1.ξ3} and \texttt{rw1.ξ4}, which means that $M'$ is respectively equal to $(\text{let } * = N' \text{ in } P')$ and $(\text{let } \langle x^A, y^B \rangle^n = N' \text{ in } P')$, with $\bar{N} \to_1 \bar{N}'$ and $P \to_1 P'$. Since $M = M'$, we have $N = N'$ and $P = P'$. By induction hypothesis, $N$ and $P$ are in intermediate neutral form. This makes $M$ in intermediate neutral form.

By exhaustion of the cases, the lemma is verified. \hfill \Box

10.3 Term Rewriting System Number Two

We now define a rewriting system that takes intermediate neutral terms and returns neutral terms. The proof of the convergence of the rewriting system requires several notions of measures on terms.

10.3.1 The Rewriting System

Definition 10.3.1. We define the rewrite system number two of valid typed intermediate neutral terms. We write $M \to_2 N$ if either $M = N$, or $M \to_{2β} N$ or $M \to_{2c} N$, where ($\to_{2β}$) and ($\to_{2c}$) are defined as follows. First,

\[
\begin{align*}
\text{let } x^A &= V \text{ in } M \to_{2β} M[V/x], & (\texttt{rw2.β1}) \\
\text{let } \langle x^A, y^B \rangle^n &= \langle V, W \rangle^n \text{ in } M \to_{2β} M[V/x, W/y], & (\texttt{rw2.β2}) \\
\text{let } * &= * \text{ in } M \to_{2β} M, & (\texttt{rw2.β3})
\end{align*}
\]

(provided that $x, y$ are fresh variables). Then,

\[
\begin{align*}
\text{let } \Box &= (\text{let } \Box = M \text{ in } N) \text{ in } P \to_{2c} & (\texttt{rw2.c1})
\end{align*}
\]
Proof by structural induction on \( M \) and \( N \) are terms such that \( M \to_{2i} M' \) and \( N \to_{2i} N' \), where \( i \) is \( \beta \) or \( c \).

\[
\begin{align*}
\lambda^n x^A.M & \to_{2i} \lambda^n x^A.M', \\
M N & \to_{2i} M' N', \\
\langle M, N \rangle^n & \to_{2i} \langle M', N' \rangle^n, \\
\langle x^A, y^B \rangle^n & \to_{2i} \langle x^A, y^B \rangle^n, \\
* & \to_{2i} *.
\end{align*}
\]

If \( M \to_{2i} N \) if and only if \( M = N \), we say that \( M \) is in \( \to_{2i} \)-normal form.

**Lemma 10.3.2.** Let \( M \) be an intermediate neutral term (respectively intermediate neutral value). If \( M \to N \), then \( N \) is an intermediate neutral term (respectively intermediate neutral value).

**Proof.** Proof by structural induction on \( M \). \( \square \)

**Lemma 10.3.3.** If \( \text{Erase}(M) = \text{Erase}(N) \) and if \( M \to_{2i} M' \) there exists \( N' \) such that \( \text{Erase}(M') = \text{Erase}(N') \) and such that \( N \to_{2i} N' \).

**Proof.** Proof by induction on the reduction \( M \to_{2i} M' \). \( \square \)

**Lemma 10.3.4.** Suppose that \( \Delta \vdash M : A \) is a valid typing judgements such that \( M \to_{2i} N \). Then \( \Delta \vdash N : A \) is valid and \( \Delta \vdash M \approx_{ax} N : A \).

**Proof.** Proof by induction on the reduction \( M \to_{2i} N \). \( \square \)

**Lemma 10.3.5.** Let \( M \) be an intermediate neutral term. If \( M \) is in \( \to_{2\beta} \)-normal form and in \( \to_{2\alpha} \)-normal form then it is a neutral term.

**Proof.** Proof by structural induction on \( M \). \( \square \)

**Lemma 10.3.6.** Suppose that \( V \) is an intermediate neutral value such that \( V \to V' \). Suppose that \( M \to_{2i} M' \). Then \( M[V/x] \to_{2i} M[V'/x] \) and \( M[V/x] \to_{2i} M'[V/x] \).

**Proof.** Proof by structural induction on \( M \). \( \square \)

**Lemma 10.3.7.** Let \( V \) and \( W \) be intermediate neutral values and let \( M \) be a neutral term.

- If \( \langle x^A, y^B \rangle^n = \langle V, W \rangle^n \) in \( N \), then \( N[V/x, W/y] \to_{2i} M \);
- If \( \langle x^A, y^B \rangle^n = \langle V, W \rangle^n \) in \( N \), then \( N[V/x, W/y] \to_{2i} M \);
10.3. Term Rewriting System Number Two

• If \((let * = * in N) \rightarrow^*_2 M\) then \(N \rightarrow^*_2 M\);

Proof. We prove the result for all the cases in parallel. We consider a term \(P\) of the form \((let x^A = V in N)\), \((let (x^A,y^B)^n = ⟨V,W⟩^n in N)\) or \((let * = * in N)\) reducing to a neutral term \(M\) in \(n\) steps and proceed be induction on \(n\).

In each case, there is at least one reduction possible (using respectively \([\text{rw}2.β]\), \([\text{rw}2.β]\) and \([\text{rw}2.β]\), and then the reduction \(P \rightarrow^*_2 M\) can be decomposed into \(P \rightarrow^*_2 M' \rightarrow^*_2 M\). The first reduction can be one of the following cases:

\([\text{rw}2.β]\), \([\text{rw}2.β]\) and \([\text{rw}2.β]\). Then \(M'\) is respectively of the form \(N[V/x], N[V/x,W/y]\) and \(N\), and we are done.

\([\text{rw}2.β]\). Then \(M'\) is of the form \(let x^A = V in N'\). By induction hypothesis, \(N'[V/x]\) reduces to \(M\). Since \(N[V/x] = N'[V/x]\), we have the result.

\([\text{rw}2.β]\). This case is similar to the previous one. \(M'\) is of the form

\[let x_1^A = V_1 in let x_2^A = V_2 in N',\]

and \(V = (V_1,V_2)^n\). Applying the induction hypothesis twice and noticing that \(N[V/x] \rightarrow^*_2 N'[V_1/x_1,V_2/x_2]\), we obtain that \(N[V/x] \rightarrow^*_2 M\)

\([\text{rw}2.β]\). In this case, \(V = *^n\) and \(M'\) is of the form \(let x^A = *^n in N'\). Applying induction hypothesis, we get that \(N'[V/x] \rightarrow^*_2 M\). Since \(N[V/x] \rightarrow^*_2 N'[V/x]\), we get the result.

Confluence cases. These cases are taken care of by using the induction hypothesis, Lemma 10.3.8 and Lemma 10.3.9.

\[\square\]

10.3.2 Height of Terms

Definition 10.3.8. Let \(h(M)\) be the height of the term \(M\) to be the positive integer defined as follows:

\[h(x^A) = h(c^A) = h(*^n) = 1,\]

\[h(\lambda^x x^A, M) = h(M) + 1\]

\[h((V,W)^n) = h(V) + h(W) + 1,\]

\[h(c^A \rightarrow^B V) = h(V) + 1,\]

\[h(x^A \rightarrow^B V) = h(V) + 1,\]

\[h(let M = M in N) = 2^{h(M)} + h(N).\]

Lemma 10.3.9. Suppose that \(M : A\) and that \(A <: B\). Then \(h\{M <: B\} = h(M)\).

Proof. Proof by structural induction on \(M\).

Lemma 10.3.10. For all intermediate neutral terms \(P, Q\) and \(R\), the following equality is valid:

\[h(let M = P in let M = Q in R) < h(let M = (let M = P in Q) in R).\]

Proof. We have

\[h(let M = (let M = P in Q) in R) = 2^{h(P)} + h(Q) + h(R)\]

\[= 2^{h(P)} \cdot 2^{h(Q)} + h(R)\]

\[\square\]
and
\[ h(\text{let } \Box = (\text{let } \Box = P \text{ in } Q) \text{ in } R) = 2^{h(P)} + 2^{h(Q)} + h(R). \]
Since for all \( x \geq 1 \) and for all \( y, z \geq 2 \), the inequalities \( 2^x < 2^{2^x} \) and \( y + z \leq y \cdot z \) hold, the lemma is valid.

10.3.3 Degree of Terms

**Definition 10.3.11.** We define the *degree* \( \delta(A) \) of a type \( A \) as follows:

\[
\begin{align*}
\delta(\top) &= 1, \\
\delta(\alpha) &= 1, \\
\delta(!A) &= \delta(A), \\
\delta(A \otimes B) &= \delta(A \to B) = \max(\delta(A), \delta(B)) + 1.
\end{align*}
\]

The degree of an intermediate neutral term is defined as follows:

\[
\begin{align*}
d(x^A) &= 0, & d(x^{A \to B}M) &= d(M), \\
d(c^A) &= 0, & d(c^{A \to B}M) &= d(M), \\
d(\ast^n) &= 0, & d(\lambda^n x^A.M) &= d(M), \\
d(\langle V, W \rangle^n) &= \max(d(V), d(W)) \\
d(\text{let } x^A = V \text{ in } N^C) &= \max(d(V), d(N), \delta(A \to C)) \\
d(\text{let } x^A, y^B = \langle V, W \rangle^n \text{ in } N^C) &= \max(d(V), d(W), d(N), \delta(A \to C), \delta(B \to C)) \\
d(\text{let } \ast = \ast \text{ in } N^C) &= \max(d(N), d(\top \to C)) \\
d(\text{let } \Box = (\text{let } \Box = M \text{ in } N) \text{ in } P) &= d(\text{let } \Box = M \text{ in } \text{let } \Box = N \text{ in } P)
\end{align*}
\]

and in all other cases,

\[ d(\text{let } \Box = M \text{ in } N) = \max(d(M), d(N)) \]

**Lemma 10.3.12.** For all types \( A \), \( \delta(A) \geq 0 \).

*Proof.* Proof by structural induction on \( A \). □

**Lemma 10.3.13.** For all terms \( M \), \( d(M) \geq 0 \).

*Proof.* Proof by structural induction on \( M \), using Lemma 10.3.12. □

**Lemma 10.3.14.** If \( A <: B \), then \( \delta(A) = \delta(B) \).

*Proof.* By structural induction on a derivation of \( A <: B \). □

**Lemma 10.3.15.** Suppose that \( M : A \) and that \( A <: B \). Then \( d(M) = d\{M <: B\} \).

*Proof.* By structural induction on \( M \). □

**Lemma 10.3.16.** Suppose that \( M \) is a neutral term. Then \( d(M) = 0 \).

*Proof.* Proof by structural induction on \( M \). □
Lemma 10.3.17. The following inequalities are valid:
\[
\begin{align*}
d(\text{let } (x^A, y^B)^n = M \text{ in } N^C) & \leq \max(\delta(A \rightarrow C), \delta(B \rightarrow C), d(M), d(N)), \\
d(\text{let } x^A = M \text{ in } N^C) & \leq \max(\delta(A \rightarrow C), d(M), d(N)), \\
d(\text{let } * = M \text{ in } N^C) & \leq \max(\delta(T \rightarrow C), d(M), d(N)), \\
d(\text{let } \square^A = M \text{ in } N) & \geq \max(d(M), d(N)).
\end{align*}
\]

Proof. The proof is done by induction on the height of \( \square^A = M \text{ in } N^C \). We proceed by case distinction on \( \square^A = M \).

Cases \((x^A = V), ((x^A, y^B)^n = \langle V, W \rangle^n) \) and \((* = *)\). In these three cases, the result is true by the definition of the degree.

Case \( \square^A = R \text{ in } P \). By definition:
\[
\begin{align*}
d(\text{let } \square^A = M \text{ in } N) &= d(\text{let } \square^B = P \text{ in } \square^A = Q \text{ in } N), \\
\max(\delta(A \rightarrow C), d(M), d(N)) &= \\
\max(\delta(A \rightarrow C), d(M), d(N)) &= \\
\max(d(M), d(N)) &= \max(d(\text{let } \square^B = P \text{ in } Q), d(N)), \\
\max(d(M), d(N)) &= \max(d(\text{let } \square^B = P \text{ in } Q), d(N)).
\end{align*}
\]

We prove the equality by case distinction on \( \square^B = P \):

Subcase \((x^B = V)\). The terms in Equations (10.3.1)-(10.3.3) become
\[
\begin{align*}
d(\text{let } \square^A = M \text{ in } N) &= \max(d(P), d(\text{let } \square^A = Q \text{ in } N), \delta(B \rightarrow C)), \\
\max(\delta(A \rightarrow C), d(M), d(N)) &= \max(d(P), d(Q), d(N), \delta(A \rightarrow C), \delta(B \rightarrow C)), \\
\max(d(M), d(N)) &= \max(d(P), d(Q), d(N), \delta(B \rightarrow C)).
\end{align*}
\]

By induction hypothesis,
\[
\max(d(Q), d(N)) \leq d(\text{let } \square^A = Q \text{ in } N) \leq \max(d(Q), d(N), \delta(A \rightarrow C)).
\]

Therefore,
\[
\max(d(P), d(Q), d(N), \delta(B \rightarrow C)) \\
\leq \max(d(P), d(Q), d(N), \delta(B \rightarrow C)) \\
\leq \max(d(P), d(Q), d(N), \delta(A \rightarrow C), \delta(B \rightarrow C)).
\]

Subcases \((x^B, y^B)^n = \langle V, W \rangle^n) \) and \((* = *)\). These cases are similar to the previous one.

Subcase \( \square^B = R \text{ in } S \). Then
\[
\begin{align*}
\max(\delta(A \rightarrow C), d(M), d(N)) &= \max(\delta(A \rightarrow C), d(\text{let } \square^B = R \text{ in } \square^A = S \text{ in } Q), d(N)), \\
\max(d(M), d(N)) &= \max(d(\text{let } \square^B = R \text{ in } \square^A = S \text{ in } Q), d(N)).
\end{align*}
\]

By definition,
\[
d(\text{let } \square^A = \text{let } \square^B = \text{let } \square^D = \text{let } R \text{ in } S \text{ in } Q) \text{ in } N)
\]
10.3. Term Rewriting System Number Two

10.3.4 Last Notion of Measure: Number of a Term

**Definition 10.3.18.** Let $e_{A_1,\ldots,A_n}^i$ be 1 when $i = \max(\delta(A_1), \ldots, \delta(A_n))$ and 0 otherwise. We define the $i$-th number $n^i(M)$ of $M$ to be a non-negative integer as follows:

\[
\begin{align*}
n^i(x^A) &= 0, & n^i(x^{A \rightarrow B} M) &= n^i(M), \\
n^i(c^A) &= 0, & n^i(c^{A \rightarrow B} M) &= n^i(M), \\
n^i(n) &= 0, & n^i(\lambda^n x^A. M) &= n^i(M), \\
n^i((V, W)^n) &= n^i(V) + n^i(W), & n^i(\text{let } x^A = V \text{ in } N^C) &= n^i(V) + n^i(N) + e_{A \rightarrow C}^i, \\
n^i(\text{let } x^A, y^B)^n = (V, W)^n \text{ in } N^C) &= n^i((V, W)^n) + n^i(N) + e_{A \rightarrow C, B \rightarrow C}^i, \\
n^i(\text{let } * = * \text{ in } N^C) &= n^i(N) + e_{T \rightarrow C}^i, \\
n^i(\text{let } \Box^A = (\text{let } \Box^B = M \text{ in } N) \text{ in } P) &= n^i(\text{let } \Box^B = M \text{ in } \text{let } \Box^A = N \text{ in } P).
\end{align*}
\]

and in all other cases,

\[
n^i(\text{let } \Box = M \text{ in } N) = n^i(M) + n^i(N)
\]

We define $n(M) = n^{d(M)}(M)$, and call it the **maximal number of $M$**.
10.3. Term Rewriting System Number Two

Lemma 10.3.19. Suppose that \( M : A \) and that \( A \prec B \). Then \( n^i(M) = n^i\{M \prec B\} \).

Proof. Proof by structural induction on \( M \).

Lemma 10.3.20. For all \( i \) we have \( n^i(\text{let } \square = M \text{ in } N) \geq \max(n^i(M), n^i(N)) \).

Proof. Proof by induction on \( M \).

Lemma 10.3.21. Suppose that \( M : C \) and that \( i > d(M) \). Then \( n^i(M) = 0 \).

Proof. The proof is done by double induction on the size and on the height of \( M \).

Case \( M \equiv (V, W)^n \). If \( i > d(M) \), then \( i > d(V) \) and \( i > d(W) \). By induction hypothesis (the size of \( V \) and the size of \( W \) being strictly smaller than the size of \( M \)), we have \( n^i(V) = n^i(W) = 0 \). Since \( n^i(M) = n^i(V) + n^i(W) \), it is also null.

Case \( M \equiv (\text{let } \langle x^A, y^B \rangle^n = \langle V, W \rangle^n \text{ in } N) \). Here, \( d(M) = \max(d(V), d(W), d(N), \delta(A \prec C), \delta(B \prec C)) \), and \( n^i(M) = n^i(V) + n^i(W) + n^i(N) + e_{A \prec C, B \prec C}^i \). If \( i > d(M) \), by induction hypothesis, \( n^i(V) = n^i(W) = n^i(N) = 0 \). Also, by definition \( e_{A \prec C, B \prec C}^i = 0 \). Thus \( n^i(M) = 0 \).

Case \( M \equiv (\text{let } \square = (\text{let } \square = N \text{ in } P) \text{ in } Q) \). In this case, \( M \rightarrow_{2c} M' \) where \( M' = \text{let } \square = N \text{ in } P \) in \( Q \). By Lemma 10.3.21 \( h(M') < h(M) \). Since \( d(M) = d(M') \) and \( n^i(M) = n^i(M') \), by induction hypothesis \( n^i(M) = 0 \).

The remaining cases are similar.

10.3.5 Putting Everything Together

The degree, the number and the height of a term are related: if \( M \) is not neutral, there exists \( N \) such that \( M \rightarrow_2 N \) and such that the triple \( (d(N), n(N), h(N)) \) is smaller with the lexicographic order to the triple \( (d(M), n(M), h(M)) \).

Definition 10.3.22. We define a function \( \mu \) from intermediate neutral terms to \( \mathbb{N}^3 \) as follows:

\[
\mu(M) = (d(M), n(M), h(M)).
\]

We consider the lexicographic ordering on \( \mathbb{N}^3 \): \( (a, b, c) < (a', b', c') \) if and only if

\[
(a < a') \vee \left( (a = a') \land (b < b') \right) \vee \left( (a = a') \land (b = b') \land (c < c') \right).
\]

Lemma 10.3.23. Suppose that \( (a_n, b_n, c_n)_n \) is a sequence in \( \mathbb{N}^3 \) such that for all \( n \), the inequality \( (a_{n+1}, b_{n+1}, c_{n+1}) \leq (a_n, b_n, c_n) \) holds. Then there exists an integer \( n_0 \) such that for all \( n \geq n_0 \), \((a_n, b_n, c_n) = (a_{n_0}, b_{n_0}, c_{n_0})\).

Proof. Consider the sequence \( (a_n)_n \). By definition, \( a_{n+1} \leq a_n \). Since this is a decreasing sequence of natural numbers: there exists \( n_a \) such that \( a_n = a_{n_a} \) for all \( n \geq n_a \). Consider the sequence \( (b_n)_{n \geq n_a} \). By definition, it is a decreasing sequence of natural numbers. Again, there exists some \( n_b \geq n_a \) such that for all \( n \geq n_b \), \( b_n = b_{n_b} \). Finally, consider the sequence \( (c_n)_{n \geq n_b} \). Again by definition, it is a decreasing sequence of natural numbers: there exists some \( n_c \geq n_b \) such that or all \( n \geq n_c \), \( c_n = c_{n_c} \). Thus, for all \( n \geq n_0 \), \( (a_n, b_n, c_n) = (a_{n_0}, b_{n_0}, c_{n_0}) \).

Lemma 10.3.24. Suppose that \( x^A = V \) in \( M : B \), where \( V \) is an intermediate neutral value and \( M \) an intermediate neutral term. Suppose moreover that let \( x^A = V \) in \( M \) is in normal form with respect to \( \rightarrow_{2c} \). Then
1. \( d(M[V/x]) \leq \max(d(M), d(V), \delta(A)) \);

2. for all \( i \geq \max(d(M), d(V), \delta(A) + 1) \), \( n^i(M[V/x]) = n^i(M) + \xi_M \cdot n^i(V) \), where \( \xi_M \) is the number of free occurrences of \( x \) in \( M \).

**Proof.** First note that is let \( x^A = V \) in \( M \) is in normal form with respect to \( \rightarrow_{2c} \), then any subterm of \( V \) and \( M \) are also in normal form with respect to \( \rightarrow_{2c} \). The proof is done by induction on the size of \( M \) plus the size of \( V \). Cases where \( M \) is an intermediate neutral value:

**Case** \( M \equiv x^B \). Then \( A \prec B \), and \( M[V/x] = \{ V \prec B \} \).

1. From Lemma 10.3.15, \( d(V \prec B) = d(V) \leq \max(d(M), d(V), \delta(A) + 1) \).
2. From Lemma 10.3.19, for all integers \( i \), \( n^i(V \prec B) = n^i(V) = n^i(M) + 1 \cdot n^i(V) \).

**Cases** \( M \equiv y^B \), \( M \equiv x^B \) and \( M \equiv 0^a \). In these three cases, \( M[V/x] = M \).

1. \( d(M[V/x]) = d(M) \leq \max(d(M), d(V), \delta(A) + 1) \).
2. Since \( \xi_M = 0 \), for all integers \( i \), \( n^i(M[V/x]) = n^i(M) + \xi_M \cdot n^i(V) \).

**Case** \( M \equiv \lambda^a y^C \cdot N \). The type \( B \) is of the form \( !^c(C \rightarrow D) \), with \( N : D \).

Using \( \alpha \)-equivalence one can assume that \( y \neq x \). Therefore, the term \( M[V/x] \) is equal to \( \lambda^a y^C \cdot (N[V/x]) \). By induction hypothesis,

\[
\forall i \geq \max(d(N), d(V), \delta(A) + 1), \quad n^i(N[V/x]) = n^i(N) + \xi_N \cdot n^i(V).
\]

Note that \( M[V/x] = \lambda^a y^C \cdot N[V/x] \). Therefore:

1. Since \( d(M) = d(N) \) and \( d(M[V/x]) = d(N[V/x]) \),

\[
d(M[V/x]) \leq \max(d(N), d(V), \delta(A) + 1) = \max(d(M), d(V), \delta(A) + 1).
\]

2. Since \( x \neq y \), we have \( \xi_M = \xi_N \). Finally, since \( n^i(M) = n^i(N) \) and since \( d(M) = d(N) \), for all \( i \geq \max(d(M), d(V), \delta(A) + 1) \),

\[
n^i(M[V/x]) = n^i(N[V/x]) = n^i(N) + \xi_N \cdot n^i(V) = n^i(M) + \xi_M \cdot n^i(V).
\]

**Case** \( M \equiv (W_1, W_2)^n \). The substitution \( M[V/x] \) is equal to \( (W_1[V/x], W_2[V/x])^n \), and the type \( B \) is of the form \( !^c(B_1 \otimes B_2) \) with \( W_1 : !^n B_1 \) and \( W_2 : !^n B_2 \).

By induction hypothesis, for \( j = 1, 2 \)

\[
\forall i \geq \max(d(W_j), d(V), \delta(A) + 1), \quad n^i(W_j[V/x]) = n^i(W_j) + \xi_{W_j} \cdot n^i(V).
\]

Note that \( M[V/x] = (W_1[V/x], W_2[V/x])^n \).

1. The degree of \( M[V/x] \) satisfies the following:

\[
d(M[V/x]) = \max(d(W_1[V/x]), d(W_2[V/x]))
\leq \max(d(W_1), d(W_2), d(V), \delta(A))
= \max(d(M), d(V), \delta(A)).
\]
2. Since $\xi_M = \xi_{W_1} + \xi_{W_2}$ and since for $j = 1, 2$ we have $d(M) \geq d(W_j)$, then for all $i$ greater or equal to $\max(d(M), d(V), \delta(A) + 1)$,

$$n'(M[V/x]) = n'(W_1[V/x]) + n'(W_2[V/x])$$
$$= n'(W_1) + \xi_{W_1} \cdot n'(V) + n'(W_2) + \xi_{W_2} \cdot n'(V)$$
$$= n'(M) + \xi_M \cdot n'(V).$$

The cases where $M$ is not a value are the following.

**Case** $M \equiv x^{C \rightarrow B} W$. Here, $W$ is an intermediate neutral value with $W : C$. From Lemma 9.1.16, $A <: C \rightarrow B$. This means that $A = !n(C' \rightarrow B')$, where $B' <: B$ and $C <: C'$.

Since $V$ is of type $!n(C' \rightarrow B')$, it is of the form $\lambda^p y^{C'}.N$, and then

$$M[V/x] = \{\lambda^p y^{C'.N} <; C \rightarrow B\}(W[V/x])$$
$$= (\lambda^p y^{C'.N} <; B\})(W[V/x])$$
$$= (\text{let } y^C = W[V/x] \in \{N <: B\}).$$

By induction hypothesis,

$$d(W[V/x]) \leq \max(d(W), d(V), \delta(A)),$$
$$\forall i \geq \max(d(W), d(V), \delta(A) + 1), \quad n'(W[V/x]) = n'(W) + \xi_W \cdot n'(V).$$

From Lemma 10.3.16, $W[V/x]$ is an intermediate neutral value. Note that $M[V/x] = (\text{let } y^C = W[V/x] \in \{N <: B\}).$

1. Using Lemma 10.3.14, $\delta(A) = \delta(C' \rightarrow B') = \delta(C \rightarrow B)$. From Lemma 10.3.16, we have $d\{N <: B\} = d(N)$. Finally, $d(N) = d(V)$ and $d(M) = d(W)$. Thus,

$$d(M[V/x]) = \max(d(W[V/x]), d\{N <: B\}, \delta(C \rightarrow B))$$
$$\leq \max(d(W), d(V), \delta(A), d\{N <: B\}, \delta(C \rightarrow B))$$
$$\leq \max(d(M), d(V), \delta(A)).$$

2. From Lemma 10.3.19, $n'(N <: B) = n'(N)$. By definition, $n'(N) = n'(V)$ and $n'(W) = n'(M)$. Finally, $\xi_M = 1 + \xi_W$. Since $d(M) = d(W)$, for all $i \geq \max(d(M), d(V), \delta(A) + 1)$, we have $c_{C \rightarrow B} = 0$ and thus

$$n'(M[V/x]) = n'(W[V/x]) + n'(N <: B) = n'(W) + \xi_W \cdot n'(V) + n'(N <: B)$$
$$= n'(M) + \xi_W \cdot n'(V) + n'(V) = n'(M) + \xi_M \cdot n'(V).$$

**Cases** $M \equiv x^{C \rightarrow B} W$ and $M \equiv y^{C \rightarrow B} W$, with $y \neq x$. The term $W$ is an intermediate neutral value such that $W : C$ and the term $M[V/x]$ is respectively $x^{C \rightarrow B}(W[V/x])$, $y^{C \rightarrow B}(W[V/x])$. From Lemma 10.2.1, $W[V/x]$ is an intermediate neutral value. By induction hypothesis,

$$d(W[V/x]) \leq \max(d(W), d(V), \delta(A)),$$
$$\forall i \geq \max(d(W), d(V), \delta(A) + 1), \quad n'(W[V/x]) = n'(W) + \xi_W \cdot n'(V).$$

Note that $M[V/x] = x^{C \rightarrow B}(W[V/x])$.

1. The degree of $M[V/x]$ is equal to $d(W[V/x])$. Thus $d(M[V/x]) \leq \max(d(W), d(V), \delta(A)) = \max(d(M), d(V), \delta(A))$. 

10.3. Term Rewriting System Number Two 166
2. Since $d(M) = d(W)$ and $\xi_M = \xi_W$, for all $i \geq \max(d(M), d(V), \delta(A) + 1)$,

$$n^i(M[V/x]) = n^i(W[V/x]) = n^i(W) + \xi_W \cdot n^i(V) = n^i(M) + \xi_M \cdot n^i(V).$$

**Case** $M \equiv (\text{let } y^C = W \text{ in } N)$. Here, the term $W$ is an intermediate neutral value, and $W : C$ and $N : B$. Using $\alpha$-equivalence, one can assume that $y \neq x$. We have $M[V/x] = (\text{let } y^C = W[V/x] \text{ in } N[V/x])$.

By induction hypothesis,

$$d(W[V/x]) \leq \max(d(W), d(V), \delta(A)),
\quad d(N[V/x]) \leq \max(d(N), d(V), \delta(A)),
\forall i \geq \max(d(W), d(V), \delta(A) + 1),
\quad n^i(W[V/x]) = n^i(W) + \xi_W \cdot n^i(V),
\forall i \geq \max(d(N), d(V), \delta(A) + 1),
\quad n^i(N[V/x]) = n^i(N) + \xi_N \cdot n^i(V).$$

From Lemma 10.2.5 $W[V/x]$ is an intermediate neutral value. Note that the degree $d(M)$ is equal to $\max(d(W), d(N), \delta(C \rightarrow B))$:

1. The degree of $M[V/x]$ satisfies

$$d(M[V/x]) = \max(d(W[V/x]), d(N[V/x]), \delta(C \rightarrow B))
\leq \max(d(W), d(V), \delta(A), d(N), d(V), \delta(A), \delta(C \rightarrow B))
= \max(d(M), d(V), \delta(A)).$$

2. Note that $\xi_M = \xi_W + \xi_N$, that $d(M) \geq d(W)$ and $d(M) \geq d(N)$, and that $n^i(M) = n^i(W) + n^i(N) + e^i_{C \rightarrow B}$. For all $i \geq \max(d(M), d(V), \delta(A) + 1)$, we have

$$n^i(M[V/x]) = n^i(W[V/x]) + n^i(N[V/x]) + e^i_{C \rightarrow B}
= n^i(W) + \xi_W \cdot n^i(V) + n^i(N) + \xi_N \cdot n^i(V) + e^i_{C \rightarrow B}
= n^i(M) + \xi_M \cdot n^i(V).$$

**Case** $M \equiv (y^C, z^D)^n = (W_1, W_2)^n$ in $N$.

Idem.

**Case** $M \equiv \ast = \ast$ in $N$.

Idem.

**Case** $M \equiv \Box = (\text{let } \Box = N \text{ in } P) \text{ in } Q$. In this case,

$$\hat{M} = (\text{let } \Box = N \text{ in let } \Box = P \text{ in } Q)$$

is such that $M \rightarrow_2 \hat{M}$ but $M \neq \hat{M}$. Therefore $M$ is not in $\rightarrow_2$-normal form: We do not consider this case.

**Catch-all case**: $M \equiv \Box = N$ in $P$. Here, $\Box$ is respectively $\ast$, $y^C$ and $(y^C, z^D)^n$. By $\alpha$-equivalence, one can assume that $x \neq y$ and $x \neq z$. Then by definition, the term $M[V/x]$ is equal to $(\text{let } \Box = N[V/x] \text{ in } P[V/x])$. We claim that if we are indeed in the catch-all case, then so is $(\text{let } \Box = N[V/x] \text{ in } P[V/x])$.

We study the four other possible cases for $\Box = N[V/x]$

**Case** $(y^C = W)$. Here, $W$ is a value. For $W = N[V/x]$ to be a value, $N$ must be a value: we are not in the catch-all case.
10.3. Term Rewriting System Number Two

Case \( ((y^C, z^D)^n = (V_1, V_2)^n) \). This situation occurs either when \( N \) is a core value of the form \( (W_1, W_2)^n \), not possible since we are in the catch-all case, or when \( N \) is the variable \( x^{n(C \otimes D)} \) and \( V = (V_1, V_2)^n \). We have

\[
\begin{align*}
& (\text{let } x^A = V \text{ in } M) \\
& = (\text{let } x^{n(C \otimes D)} = (W_1, W_2)^n \text{ in let } (y^C, z^D)^n = x^{n(C \otimes D)} \text{ in } P) \\
& \rightarrow 2c: \text{let } x^{n(C \otimes D)} = (W_1, W_2)^n \text{ in let } y^C = W_1 \text{ in let } z^D = W_2 \text{ in } P.
\end{align*}
\]

which contradicts the hypothesis of the lemma. Therefore this case is not to be considered.

Case \( (* = *) \). This case is obtained when either \( N = * \), not possible since we are in the catch-all case, or when \( N = x^\top \) and \( V = *^n \). In this situation, \( A = *^n \). We have

\[
\begin{align*}
& (\text{let } x^A = V \text{ in } M) = (\text{let } x^{n^\top} = *^n \text{ in let } * = x^\top \text{ in } P) \\
& \rightarrow 2c: \text{let } x^{n^\top} = *^n \text{ in } P
\end{align*}
\]

which contradicts the hypothesis of the lemma. Therefore this case is not to be considered.

Case \( [C^\top = (\text{let } D^\top = Q \text{ in } R)] \). There are two subcases. Either \( N \subseteq \text{let } D^\top = S \text{ in } T \), but then we are not in the catch-all case, or we have \( N \equiv x^{D \rightarrow C} S \) with \( S \) a value and \( V \equiv \lambda^n y^D.R \). In this case,

\[
\begin{align*}
& (\text{let } x^A = V \text{ in } M) = (\text{let } x = \lambda y.R \text{ in let } C^\top = (\text{let } y^D = S \text{ in } R) \text{ in } P) \\
& \rightarrow 2c: \text{let } x = \lambda y.R \text{ in let } y^D = S \text{ in let } C^\top = R \text{ in } P,
\end{align*}
\]

which contradicts the hypothesis of the lemma. Therefore this case is not to be considered.

Thus the term \( M[V/x] = (\text{let } C^\top = N[V/x] \text{ in } P[V/x]) \) is in the catch-all form. By induction hypothesis,

\[
\begin{align*}
& d(N[V/x]) \leq \max(d(N), d(V), \delta(A)), \\
& d(P[V/x]) \leq \max(d(P), d(V), \delta(A)), \\
& \forall i \geq \max(d(N), d(V), \delta(A) + 1), \ n^i(N[V/x]) = n^i(N) + \xi_N \cdot n^i(V), \\
& \forall i \geq \max(d(P), d(V), \delta(A) + 1), \ n^i(P[V/x]) = n^i(P) + \xi_P \cdot n^i(V).
\end{align*}
\]

Now:

1. Since the degrees \( d(M) \) and \( d(M[V/x]) \) are respectively equal to \( \max(d(P), d(N)) \) and equal to \( \max(d(P[V/x]), d(N[V/x])) \), we have \( d(M[V/x]) \leq \max(d(M), d(V), \delta(A)) \).

2. We have \( \xi_M = \xi_N + \xi_P, \ d(M) \geq d(N) \) and \( d(M) \geq d(P) \). Therefore, for all integer

\[
\begin{align*}
n^i(M[V/x]) &= n^i(N[V/x]) + n^i(P[V/x]) \\
&= n^i(N) + \xi_N \cdot n^i(V) + n^i(P) + \xi_P \cdot n^i(V) \\
&= n^i(M) + \xi_M \cdot n^i(V).
\end{align*}
\]

Thus, by induction the degree of \( M[V/x] \) is smaller or equal to \( \max(d(M), d(V), \delta(A)) \), and for all integers \( i \) bigger or equal to \( \max(d(M), d(V), \delta(A) + 1), \ n^i(M[V/x]) = n^i(M) + \xi_M \cdot n^i(V). \)

**Lemma 10.3.25.** Suppose that \( M : B \) is in \( \rightarrow 2c \)-normal form. Then one of the following three statements is valid:
1. $M$ is in $\rightarrow_{2\beta}$-normal form.

2. $n(M) = 0$ and there exists $N \neq M$ such that $M \rightarrow_2 N$ with $d(N) < d(M)$.

3. $n(M) \neq 0$ and there exists $N \neq M$ such that $M \rightarrow_2 N$ with $d(M) < d(N)$ or $d(M) = d(N)$ and $n(N) < n(M)$.

Proof. Suppose that $M$ is not in $\rightarrow_{2\beta}$-normal form. We prove the result by induction on $M$.

Case $M \equiv (\text{let } x = P \text{ in } Q) \in R)$. In this case, $M$ is not in $\rightarrow_{2\beta}$-normal form. Therefore this case is not to be considered.

Case $M \equiv (\text{let } x^A = V \text{ in } P)$. The degree of $M$ is equal to $\max(d(V), d(P), \delta(A \rightarrow B))$, and $n(M) = n(d(M))(P) + n(d(M))(V) + \epsilon_{A \rightarrow B}$. Note that since $M$ is in $\rightarrow_{2\beta}$-normal form, so are $V$ and $P$.

Subcase $n(M) = 0$. We construct a term $N$ such that $d(N) < d(M)$ and $M \rightarrow_{2\beta} N$.

From Lemma [10.3.18] $n_{d(M)}(P) = n_{d(M)}(V) = \epsilon_{A \rightarrow B} = 0$. Since $d(M) \geq \delta(A \rightarrow B)$, this means $d(M) > \delta(A \rightarrow B)$. Note that $d(V), d(P) \leq d(M)$.

Suppose that $d(V) = d(M)$. Then $n(V) = n_{d(M)}(V) = 0$. Since $M$ is in $\rightarrow_{2\beta}$-normal form, so is $V$. Then by induction hypothesis, there exists $V'$ such that $V \rightarrow_{2\beta} V'$ with $d(V') < d(V)$. From Lemma [10.3.18] the term $V'$ is an intermediate neutral value. If $d(V') < d(M)$, define $V'$ to be $V$. Therefore, in both cases we obtain an intermediate neutral value $V'$ such that $V \rightarrow_{2\beta} V'$ with $d(V') < d(M)$.

Suppose that $d(P) = d(M)$. Then $n(P) = n_{d(M)}(P) = 0$. By induction hypothesis, there exists $P'$ such that $P \rightarrow_{2\beta} P'$ with $d(P') < d(P)$. If $d(P') < d(M)$, define $P'$ to be $P$. Therefore, in both cases we obtain a term $P'$ such that $P \rightarrow_{2\beta} P'$ with $d(P') < d(M)$.

Define $N$ to be $(\text{let } x^A = V' \text{ in } P')$. Since $V'$ is an intermediate neutral value, $d(N) = \max(d(V'), d(P'), \delta(A \rightarrow B))$.

Since $d(V'), d(P'), \delta(A \rightarrow B) < d(M)$, we also have $d(N) < d(M)$.

Subcase $n(M) \neq 0$. We construct a term $N$ such that $M \rightarrow_{2\beta} N$ with $d(N) < d(M)$ or with $d(M) = d(N)$ and $n(N) < n(M)$.

Suppose that $n_{d(M)}(P) \neq 0$. From Lemma [10.3.21] and since $d(P) \leq d(M)$, the degree of $P$ is equal to the degree of $M$. Therefore, $n(P) \neq 0$ and we can apply the induction hypothesis: there exists a term $P'$ such that $P \rightarrow_{2\beta} P'$, $d(P') = d(P)$ and $n(P') < n(P)$. Let $N$ be $(\text{let } x^A = V \text{ in } P')$. We have $M \rightarrow_{2\beta} N$ with $d(N) = \max(d(V'), d(P'), \delta(A \rightarrow B)) = d(M)$ and $n(N) = n_{d(M)}(P') + n_{d(M)}(V') + \epsilon_{A \rightarrow B} < n(M)$.

Now, suppose that $n_{d(M)}(V) \neq 0$. Similarly, from Lemma [10.3.21] and since $d(V) \leq d(M)$, the degree $d(V) = d(M)$. This means $n(V) \neq 0$ and we can apply the induction hypothesis: there exists a term $V'$ such that $V \rightarrow_{2\beta} V'$, $d(V') = d(V)$ and $n(V') < n(V)$. Let $N$ be $(\text{let } x^A = V' \text{ in } P)$. We have $M \rightarrow_{2\beta} N$ with $d(N) = \max(d(V'), d(P), \delta(A \rightarrow B)) = d(M)$ and $n(N) = n_{d(M)}(P) + n_{d(M)}(V') + \epsilon_{A \rightarrow B} < n(M)$.

Finally, suppose that both $n_{d(M)}(V)$ and $n_{d(M)}(P)$ are null. Then $n(M) = \epsilon_{A \rightarrow B} = 1$. Let $N$ be $P[V/x]$. Since $V$ is an intermediate neutral value, $M \rightarrow_{2\beta} N$. By Lemma [10.3.24] $d(N) \leq \max(d(P), d(V), \delta(A)) \leq d(M)$ and for all $i \geq \max(d(P), d(V), \delta(A)+1), n_i(N) = n_i(P)+\epsilon_{P}n_i(V)$. If $d(N) < d(M)$, we are done. If $d(N) = d(M)$, then $d(M) \geq \max(d(P), d(V), \delta(A)+1)$, $n_i(N) = n_i(P)+\epsilon_{P}n_i(V)$. If $d(N) < d(M)$, we are done. Hence $d(N) < d(M)$, and we are also done.

The two cases $M \equiv (\text{let } x^A, y^B)^n = (V, W)^n \text{ in } P)$ and $N \equiv (\text{let } * = * \text{ in } M)$ are done similarly. Finally, the congruence cases are trivial.
10.3. Term Rewriting System Number Two

Definition 10.3.26. We say that an intermediate term $M$ satisfies reduction if the following is valid:

“if $M \rightarrow_{2c} N$, then there exists $N'$ such that $M \rightarrow_{2}^* N'$ and $\mu(N') < \mu(M)$.”

Lemma 10.3.27. Let $M$ be an intermediate neutral term satisfying reduction, then either $M$ is a neutral term or there exists $N$ such that $\mu(N) < \mu(M)$ and $M \rightarrow_{2}^* N$.

Proof. Consider an intermediate neutral term $M$ satisfying reduction. If it is already a neutral term, we are done. If it is not, from Lemma 10.3.5 then it is not in $\rightarrow_{2c}$-normal form. There are two cases:

1. $M$ is not in $\rightarrow_{2c}$-normal form. By the reduction property of $M$ there is $N$ such that $M \rightarrow_{2}^* N$ and $\mu(N) < \mu(M)$.

2. $M$ is in $\rightarrow_{2c}$-normal form. It is not in $\rightarrow_{2\beta}$-normal form, and from Lemma 10.3.25 there exists $N$ such that $M \rightarrow_{2}^\beta N$ with $\mu(N) < \mu(M)$.

In both cases, we have as requested the existence of some intermediate neutral term $N$ such that $M \rightarrow_{2}^* N$ with $\mu(N) < \mu(M)$.

Corollary 10.3.28. Let $M$ be an intermediate neutral term satisfying reduction. Then there exists a neutral term $N$ such that $M \rightarrow_{2}^* N$.

Proof. Let $S$ be the set of terms $M$ such that $M \rightarrow_{2}^* M'$. Let $I = \mu(S)$. From Lemma 10.3.23 this set has a minimum element. Let $N \in S$ be a term such that $\mu(N)$ is this minimum. Lemma 10.3.27 states that $N$ is neutral.

Lemma 10.3.29. Every intermediate term satisfies reduction, as stated in Definition 10.3.26.

Proof. Suppose that $M \rightarrow_{2c} N$. We prove the result by induction on the size $M$. Case distinction on the last rule used in the derivation of $M \rightarrow_{2c} N$:

Case [rw2.c1]. In this case, by definition, the two sides of the relation have the same degree and the same maximal number, but a smaller height by Lemma 10.3.10. Choose $N'$ to be equal to $N$ and this satisfies the lemma.

Case [rw2.c2]. Suppose that $P : B$. The left hand side is

$M = (let \ x^n(D \rightarrow C) = \lambda^n y^D.R \ in \ let \ x^{D'} = x^{D' \rightarrow C'} S \ in \ P)$

Its degree is $d(M) = \max(d(R), \delta((D \rightarrow C) \rightarrow B), d(S), d(P))$. Note that $S$ is an intermediate value.

The right hand side is

$N = (let \ x^n(D \rightarrow C) = \lambda^n y^D.R \ in \ let \ y^{D'} = S \ in \ let \ x^{C'} = \{R <: C'\} \ in \ P)$

Its degree is

$d(N) = \max(d(R), \delta((D \rightarrow C) \rightarrow B), d(S), \delta(D' \rightarrow B), d(let \ x^{C'} = \{R <: C'\} \ in \ P))$.

From Lemma 10.3.17 we get $d(N) = d(M)$.

We proceed by case distinction on $d(M)$ to find the term $N'$.
Suppose that \( d(M) > \delta(D \rightarrow C) \rightarrow B \). Consider \( X \) being the terms among \( R, S \) and \( P \) such that \( d(X) = d(M) \). Whether \( X \) are or not in \( \rightarrow_{2c} \)-normal form, one of Lemma 10.3.32 or the induction hypothesis applies: Either \( X \) is in \( \beta \)-normal form, and thus by Lemma 10.3.16 \( d(X) = 0 \), or \( X \) reduces to \( X' \) with \( d(X') < d(X) \) or with \( d(X') = d(X) = d(M) \) and \( n(X') < n(M) \). In this last case, since \( d(X') = d(M) \), this means that \( n(N') < n(M) \) and \( d(N') = d(M) \) if \( N' \) is \( M \) with \( X' \) in place of \( X \).

Otherwise, we have a term \( N' \) with \( M \rightarrow_{2c} N' \) and \( d(N') < d(M) \).

Suppose that \( d(M) = 1 + \delta(D \rightarrow C) \rightarrow B \) and \( d(M) > 1 + \delta(B) \). The terms \( R \) and \( S \) are smaller than \( M \): By induction hypothesis they satisfy the lemma. Then by Lemma 10.3.28 there exist neutral terms \( R' \) and \( S' \) such that \( R \rightarrow_{2c} R' \) and \( S \rightarrow_{2c} S' \). By Lemma 10.3.16 we have \( d(R') = 0 \) and \( d(S') = 0 \).

Since \( \lambda x^n(D \rightarrow C) = \lambda x^n y^D \cdot R \) in \( P \) is smaller than \( M \), by the induction hypothesis one can apply Lemma 10.3.28 and it reduces to a neutral term \( P' \). From Lemma 10.3.7 \( P[V/x] \) reduces to the same neutral term \( P' \). Therefore \( M \) reduces to

\[
M \rightarrow_{2c} 2 \text{ let } y^D = S'[\lambda x^n y^D \cdot R'/x] \text{ in let } \{ R' < C' \} \text{ in } P[\lambda x^n y^D \cdot R'/x] \\
\rightarrow_{2c} 2 \text{ let } y^D = S'[\lambda x^n y^D \cdot R'/x] \text{ in let } \{ R' < C' \} \text{ in } P'.
\]

Let \( N' \) be this last term. From since Lemma 10.3.22 \( S'[\lambda x^n y^D \cdot R'/x] \) is a value, the term \( x^{n \rightarrow C} = \lambda x^n y^D \cdot R' \) in \( S' \) is in \( \rightarrow_{2c} \)-normal form. Thus Lemma 10.3.22 applies:

\[
d(N') = \max(\delta(D \rightarrow B), d(S'[\lambda x^n y^D \cdot R'/x]), d(\{ R' < C' \} \text{ in } P')) \\
\leq \max(\delta(D \rightarrow B), d(S'), d(R'), \delta(D \rightarrow C), d(\{ R' < C' \} \text{ in } P')) \\
\leq \max(\delta(D \rightarrow B), d(S'), d(R'), \delta(D \rightarrow C), d(\{ R' < C' \} \text{ in } P'), \delta(C' \rightarrow B)).
\]

Since \( d(S') = d(R') = d(P') = 0 \) (they are neutral), and since \( d(M) > \delta(D) + 1 \), \( d(M) > \delta(C) + 1 \) and \( d(M) > \delta(B) + 1 \), we have \( d(N') < d(M) \).

Suppose that \( d(M) = 1 + \delta(B) \). Let \( Q = (\text{let } y'^D = S \text{ in let } \{ R' < C' \} \text{ in } P) \). The term \( Q \) is smaller than \( M \), so induction hypothesis applies. Suppose that \( Q \) is not in \( \rightarrow_{2c} \)-normal form. Then there exists \( Q' \) such that \( Q \rightarrow_{2c} Q' \) with \( \mu(Q') < \mu(Q) \). Let \( N' = (\text{let } x^{n \rightarrow C} = \lambda x^n y^D \cdot R \text{ in } Q') \). Note that \( M \rightarrow_{2c} N' \). We have \( d(N') = \max(\delta(D \rightarrow C) \rightarrow B, d(R), d(Q')) = 1 + \delta(B) = d(M) \), and \( n(N') \leq n(M) \).

Finally, since \( h(Q') < h(Q) \), we have \( h(N') < h(M) \). Therefore \( \mu(N') < \mu(M) \).

Now, suppose that \( Q \) is in \( \rightarrow_{2c} \)-normal form. Then so is \( N \), and we have the result using Lemma 10.3.22.

Case \([\text{rw2.c}] \). Suppose that \( P : B \). The left hand side is

\[
M = (\text{let } x^{n \rightarrow C \otimes D} = (V, W)^n \text{ in let } (y^C, z^D)n = x^{m \rightarrow C \otimes D} \text{ in } P)
\]

Its degree is \( d(M) = \max(d(V), d(W), \delta(D \otimes C) \rightarrow B, d(P)) \). The right hand side is

\[
N = (\text{let } x^{m \rightarrow C \otimes D} = (V, W)^n \text{ in } y^{m C'} = \{ V < :!^m C' \} \text{ in } z^{m D'} = \{ W < :!^m D' \} \text{ in } P)
\]

and its degree is the same as the one of \( M \).

We proceed by case distinction on \( d(M) \) to find the term \( N' \). The cases are treated similarly as in the previous case.
10.4. Proof of Theorem 9.2.7

Suppose that \( d(M) > \delta((D \otimes C) \rightarrow B) \). Using Lemma 10.3.14, one can find \( V', W' \) and \( P' \) such that
\[
\delta(\langle V' \rangle, \langle W' \rangle) = \mu(M) \quad \text{for all intermediate neutral term}
\]
From Lemma 10.3.29, every intermediate term satisfies reduction. Then from Lemma 10.3.27,
\[
\text{Proof.}
\]
Lemma 10.3.30. Suppose that \( d(M) = 1 + \delta(D \otimes C) \) and \( d(M) > 1 + \delta(B) \). Since \( V', W, \) and let \( x = \langle V, W \rangle \) in \( P \) are strictly smaller than \( M \), the induction hypothesis applies and they satisfy the lemma. Thus Lemma 10.3.28 applies and they reduces to some neutral terms \( V', W' \) and \( P' \). From Lemma 10.3.4, \( P'(\langle V, W \rangle) \) also reduces to \( P' \). Therefore \( M \) reduces to
\[
N' = (\text{let } y^C, z^{D'})^m = \langle V', W' \rangle^n \text{ in } P'.
\]
We have
\[
d(N') = \max(d(V'), d(W'), d(P'), \delta(C' \rightarrow B), \delta(D' \rightarrow B)).
\]
From Lemma 10.3.14, \( d(N') < d(M) \).
Suppose that \( d(M) = 1 + \delta(B) \). This time, we choose
\[
Q = (\text{let } y^m C' = \langle V <: !^m C' \rangle \text{ in let } z^{mD'} = \langle W <: !^m D' \rangle \text{ in } P)
\]
The term \( Q \) being smaller than \( N \), the induction hypothesis applies.
Suppose that it is not in \( \rightarrow_{2,\text{-normal}} \) form. Then there exists \( Q' \) such that \( Q \rightarrow^* Q' \) with \( \mu(Q') < \mu(Q) \). Let \( N' = (\text{let } x^m (C \otimes D) = \langle V', W' \rangle^n \text{ in } Q') \). We have \( M \rightarrow^* N' \), with
\[
d(N') = d(M), n(N') \leq n(M), \text{ and } h(N') < h(M). \text{ Thus } \mu(N') < \mu(M).
\]
Now, suppose that it is in \( \rightarrow_{2,\text{-normal}} \) form. Then so is \( N \), and we have the result by Lemma 10.3.4.

Case (1w2.2). In this case,
\[
M = (\text{let } x^m \tau = \ast^n \text{ in let } \ast = x^\tau \text{ in } P), \quad N = (\text{let } x^m \tau = \ast^n \text{ in } P).
\]
Note that \( d(M) = d(N) \), that \( n(N) = n(P) \leq n(M) \) and that \( h(N) < h(M) \). Therefore \( \mu(N) < \mu(M) \). Choose \( N' = N \).

Cases (1w2.4) - (1w2.5). One can assume that \( M \) is not in one of the form requested by the previous cases. The induction hypothesis can therefore be applied and the result is obtained directly.

By case exhaustion, the lemma is then true.

Lemma 10.3.30. Suppose that \( M \) is an intermediate neutral term. Then either \( M \) is a neutral term or there exists \( N \) such that \( \mu(N) < \mu(M) \) and \( M \rightarrow^* N \).

Proof. From Lemma 10.3.29 every intermediate term satisfies reduction. Then from Lemma 10.3.27, for all intermediate neutral term \( M \), either \( M \) is a neutral term or there exists \( N \) such that \( \mu(N) < \mu(M) \) and \( M \rightarrow^* N \).

10.4 Proof of Theorem 9.2.7

The proof of Theorem 9.2.7 is done in two steps, by reducing first to intermediate neutral terms and then to neutral terms. We make these steps explicit in Lemma 10.3.14 and Lemma 10.3.4.
Lemma 10.4.1. Suppose that $\Delta \triangleright M, N : A$ are two valid typing judgements such that $\text{Erase}(M) = \text{Erase}(N)$. Then there exists two intermediate neutral terms $M'$ and $N'$ such that $\Delta \triangleright M' : A$ and $\Delta \triangleright N' : A$ are valid, such that $\text{Erase}(M') = \text{Erase}(N')$ and such that $\Delta \triangleright M \approx_{ax} M' : A$ and $\Delta \triangleright N \approx_{ax} N' : A$.

Proof. Let $S$ be the set of all possible sequences of terms $(M_i)$, such that $M_0 = 0$ and $M_i \rightarrow_{1} M_{i+1}$.

From Lemma 10.2.12, one can define the function $f$ from $S$ to the integers picking out the smallest $n$ such that for all $i \geq n$, $M_i = M_{n0}$. Consider the function $g$ mapping a sequence $(M_i)_i$ in $S$ to the value $m(M_f((M_i)_i))$. Since the measure yields non-negative integers, the set $g(S)$ is a subset of the non-negative integers. It contains then an attainable minimal element $m_0$. Let $(M_i)_i$ be a sequence of $S$ such that $g((M_i)_i) = m_0$.

Let $n_0 = f((M_i)_i)$. We claim that $M_{n_0}$ is an intermediate neutral term. Indeed, suppose otherwise. Then by Lemma 10.2.13, there exists a term $M'$ such that $M_{n_0} \rightarrow_1 M'$ and $M' \neq M_{n_0}$. We can now construct a sequence of terms $(M_i)_i$ as follows: $M_i = M_i$ for all $i = 0 \ldots n_0$, $M_i = M'$ for all $i > n_0$. Note that $(M_i)_i$ belongs in $S$. However, using Lemma 10.2.11, we deduce that $g((M_i)_i) < m_0$, contradicting the minimality of $m_0$.

Now, using Lemma 10.2.8 one can construct a sequence of terms $(N_i)_i$, such that $N_0 = N$ and such that for all $i$, $\text{Erase}(N_i) = \text{Erase}(M_i)$ and $N_i \rightarrow_1 N_{i+1}$. Using Lemma 10.2.7, for all $i$ the typing judgements $\Delta \triangleright M_{n_0}, N_{n_0} : B$ are valid, and $\Delta \triangleright M \approx_{ax} M_{n_0} : B$ and $\Delta \triangleright N \approx_{ax} N_{n_0} : B$. Using Lemma 10.2.4, the term $N_{n_0}$ is in intermediate neutral form. This makes the terms $M' = M_{n_0}$ and $N' = N_{n_0}$ fulfill the requirements of the lemma.

Lemma 10.4.2. Suppose that $\Delta \triangleright M, N : A$ are two valid typing judgements, where $M$ and $N$ are intermediate neutral terms such that $\text{Erase}(M) = \text{Erase}(N)$. Then there exists two neutral terms $M'$ and $N'$ such that $\text{Erase}(M') = \text{Erase}(N')$ and such that $\Delta \triangleright M \approx_{ax} M' : A$ and $\Delta \triangleright N \approx_{ax} N' : A$.

Proof. Consider two valid typing judgements $\Delta \triangleright M, N : A$ such that $M$ and $N$ are intermediate neutral terms with $\text{Erase}(M) = \text{Erase}(N)$.

Using Lemma 10.3.31, it is possible to construct sequences of terms $(M_i)_i$ such that $M_0 = M$, and such that for all $i$, either $M_i$ is a neutral value or $M_i \rightarrow_{1} M_{i+1}$ with $\mu(M_{i+1}) < \mu(M_i)$. From Lemma 10.3.28, there exists $n_0$ such that for all $n \geq n_0$, $\mu(M_n) = \mu(M_{n_0})$. Using the definition of the sequence, this means that for all $n \geq n_0$, $M_n = M_{n_0}$ is a neutral term. Using Lemma 10.3.36, one can build a sequence $(N_n)_n$ such that for all $n$, $\text{Erase}(N_n) = \text{Erase}(M_n)$ and such that $N_n \rightarrow_{2} N_{n+1}$. From Lemma 10.2.4, the typing judgements $\Delta \triangleright M_{n_0}, N_{n_0} : A$ are valid, and $\Delta \triangleright M \approx_{ax} M_{n_0} : A$ and $\Delta \triangleright N \approx_{ax} N_{n_0} : A$.

The requested terms $M'$ and $N'$ can be respectively taken as $M_{n_0}$ and $N_{n_0}$.

We are now ready to prove Theorem 9.2.7.

Proof of Theorem 9.2.7. Suppose that $M$ and $N$ are two terms such that $\text{Erase}(M) = \text{Erase}(N)$ and such that $\Delta \triangleright M, N : A$ are valid typing judgements.

Using Lemma 10.1.1, there exists two intermediate neutral terms $M'$ and $N'$ such that $\Delta \triangleright M', N' : A$ and such that $\Delta \triangleright M \approx_{ax} M' : A$ and $\Delta \triangleright N \approx_{ax} N' : A$ and such that $\text{Erase}(M') = \text{Erase}(N')$. Using Lemma 10.1.2, there exists two neutral terms $M''$ and $N''$ such that $\text{Erase}(M'') = \text{Erase}(N'')$ and such that $\Delta \triangleright M'' \approx_{ax} M' : A$ and $\Delta \triangleright N'' \approx_{ax} N' : A$. Using Lemma 10.1.3, we have $\Delta \triangleright M'' \approx_{ax} N'' : A$.

We obtain $\Delta \triangleright M \approx_{ax} N : A$ using the transitivity of the axiomatic relation.
Chapter 11

Categorical Semantics

In Chapter 8, we described the categorical structure believed to be required for interpreting the language described in Chapter 9. In this chapter, we show that one can indeed interpret the language in a linear category for duplication, and we prove a soundness and completeness result.

11.1 Denotational Semantics

The ambient category this section is concerned with is a linear category for duplication \( C \), with the notation of Definition 8.4.1.

Convention 11.1.1. For notation purposes, we assume the category to be strict monoidal. This can be safely assumed, thanks to Theorem 2.8.2.

11.1.1 Interpretation of the Type System

Definition 11.1.2. We define an interpretation of the type system to be a map \( \Theta \) that assigns to each constant type \( \alpha \) an object \( \Theta(\alpha) \). Each type \( A \) is interpreted as an object of \( C \):

\[
\begin{align*}
\Theta(\alpha) &= \Theta(\alpha), \\
\Theta(\top) &= \top, \\
\Theta(\!A!) &= \Theta(A), \\
\Theta(A \otimes B) &= \Theta(A) \otimes \Theta(B), \\
\Theta(A \rightarrow B) &= \Theta(A) \rightarrow \Theta(B).
\end{align*}
\]

Consider the alphabet \( \{ \alpha | \alpha \) is a constant type\} \), and the set \( \mathcal{A} \) of types. If \( G(A) = [A]_\Theta \), we are in the context of Theorem 8.3.8, with the list of \( F_i \)'s being \( (\otimes, \rightarrow) \): any two maps \( A \rightarrow B \) in \( C^G \) are equal.

Definition 11.1.3. Given a valid subtyping \( A <: B \), there exists a unique arrow \( [A]_\Theta \rightarrow [B]_\Theta \) in \( C^G \), as defined in Theorem 8.3.8. We extend the map \( \Theta \) to interpret \( A <: B \) as this unique arrow and we denote it by \( I_{A,B} \).

11.1.2 Interpretation of the Language

Convention 11.1.4. Typing contexts were defined as sets. In the case we are interested in, we need them to be lists. We will therefore consider them to be lists in this section, the typing rules of Table 9.1 being physically unchanged. In order to still be able to reorder typing contexts, we add a set of new rules \( (X_s) \), one for each permutation \( s \) on \( \{1, \ldots, n\} \):

\[
\frac{x_1 : A_1, \ldots, x_n : A_n \triangleright M : A}{x_{s(1)} : A_{s(1)}, \ldots, x_{s(n)} : A_{s(n)} \triangleright M : A} \ (X_s)
\]
Lemma 11.1.7. Suppose that \( \Delta = (x_1 : A_1, \ldots, x_n : A_n) \) by the following:

\[
[\Delta] = [A_1] \otimes \cdots \otimes [A_n].
\]

Definition 11.1.6. We use the following shortcut definitions, where \( A, B, A_1, \ldots, A_n \) are types and where \( \Delta, \Gamma_1 \) and \( \Gamma_2 \) are typing contexts:

- Given a natural transformation \( n : F \Rightarrow G \), we write \( n_A \) in place of \( n[A] \). If \( \Delta = (x_1 : A_1, \ldots, x_n : A_n) \) we define \( n_\Delta = n_{A_1} \otimes \cdots \otimes n_{A_n} \).
- For the coherence map \( d^L_{A,B} : LA \otimes LB \to L(A \otimes B) \), if \( \Delta = (x_1 : A_1, \ldots, x_n : A_n) \), we define \( d^L_\Delta \) to be the map

\[
L[A_1] \otimes \cdots \otimes L[A_n] \xrightarrow{id \otimes d^L_{A_n-1,A_n} id} L[A_1] \otimes \cdots \otimes L[A_{n-1} \otimes A_n] \xrightarrow{\sim} L[A_1 \otimes \cdots \otimes A_n].
\]
- We define \( Split_{\Delta, \Gamma_1, \Gamma_2} \) to be the canonical map from the monoidal structure and using the map \( \Delta_A : LA \to LA \otimes LA \) from Definition 11.1.5.

\[
Split_{\Delta, \Gamma_1, \Gamma_2} : [\Delta] \otimes [\Gamma_1] \otimes [\Gamma_2] \to [\Delta] \otimes [\Gamma_1] \otimes [\Delta] \otimes [\Gamma_2].
\]
- Given \( f : [\Delta] \otimes [\Gamma_1] \to [A] \) and \( g : [\Delta] \otimes [\Gamma_2] \to [B] \), we define the map \( f \otimes_{\Delta} g : [\Delta] \otimes [\Gamma_1] \otimes [\Gamma_2] \to [A] \otimes [B] \) as follows:

\[
[\Delta] \otimes [\Gamma_1] \otimes [\Gamma_2] \xrightarrow{Split_{\Delta, \Gamma_1, \Gamma_2}} [\Delta] \otimes [\Gamma_1] \otimes [\Delta] \otimes [\Gamma_2] \xrightarrow{f \otimes g} [A] \otimes [B].
\]

Lemma 11.1.7. Suppose that \( f : [\Delta, \Gamma_2] \to [A] \). Suppose that \( \Gamma_1 \) is a context of fresh variables. Then

\[
Split_{\Delta, \Gamma_1, \Gamma_2} (\Diamond \Delta, \Gamma_1, f) = (\sigma_{[\Delta], [\Gamma_1]} \otimes id_{\Gamma_2}) (\Diamond \Gamma_1, f) : [\Delta, \Gamma_1, \Gamma_2] \to [A].
\]

Proof. Using the equations of comonoids.

Lemma 11.1.8. Suppose that \( f : [\Delta, \Gamma_1] \to [A] \) and that \( g : [\Delta, \Gamma_2] \to [B] \). Suppose moreover that \( !\Delta \) is a context of fresh variables. Then

\[
(\Diamond !\Delta, f \otimes f) \otimes_{!\Delta, !\Delta} (\Diamond !\Delta \otimes g) = \Diamond !\Delta \otimes (f \otimes g) : [!\Delta, !\Delta, \Gamma_1, \Gamma_2] \to [A] \otimes [B].
\]

Proof. It is enough to show that the maps

\[
[!\Delta, !\Delta, \Gamma_1, \Gamma_2] \xrightarrow{Split_{\Delta, \Gamma_1, \Gamma_2}} [!\Delta, !\Delta, \Gamma_1, !\Delta, !\Delta, \Gamma_2] \xrightarrow{\Diamond !\Delta \otimes f \otimes !\Delta \otimes g} [A_1] \otimes [A_2]
\]

and

\[
[!\Delta, !\Delta, \Gamma_1, \Gamma_2] \xrightarrow{\Diamond !\Delta \otimes id} [!\Delta, \Gamma_1, \Gamma_2] \xrightarrow{Split_{\Delta, \Gamma_1, \Gamma_2}} [!\Delta, !\Delta, \Gamma_1, \Gamma_2] \xrightarrow{f \otimes g} [A_1] \otimes [A_2]
\]

are equal. This is done using the comonoid structure of \((!\Delta, \Delta_{\Delta !\Delta}, \Diamond !\Delta)\).

Lemma 11.1.9. Suppose that \( f : [!\Delta, !\Delta, \Gamma_1] \to [A] \) and that \( g : [!\Delta, \Gamma_2] \to [B] \). Then,

\[
f \otimes_{!\Delta, !\Delta} (\Diamond !\Delta \otimes g) = (id_{!\Delta} \otimes Split_{\Delta, \Gamma_1, \Gamma_2}) (f \otimes g) : [!\Delta, !\Delta, \Gamma_1, \Gamma_2] \to [A] \otimes [B].
\]

Proof. Proof using the comonoid structure of \((!\Delta, \Delta_{!\Delta !\Delta}, \Diamond !\Delta)\).
Lemma 11.1.10. Suppose that $f : [\Gamma_1] \to [A]$ and that $g : [A] \to [B]$. Then, if $!\Delta$ is a context of fresh variables,

$$(\Diamond_\Delta \otimes f); g = \Diamond_\Delta \otimes (f; g) : [!\Delta, \Gamma_1] \to [B].$$

Proof. In our strict monoidal category, the map $(\Diamond_\Delta \otimes f); g$ is

$$(\Diamond_\Delta \otimes f); g = (\Diamond_\Delta \otimes f); \lambda_A; g$$

$$(\Diamond_\Delta \otimes f); (id_T \otimes g); \lambda_B$$

by naturality of $\lambda$,

$$(\Diamond_\Delta \otimes (f; g)); \lambda_B$$

by bifunctoriality of $\otimes$.

And this map is precisely $\Diamond_\Delta \otimes (f; g)$. \hfill $\Box$

Lemma 11.1.11. Suppose that $f : [!\Delta, \Gamma_1] \to [A]$, $g : [!\Delta, \Gamma_2] \to [B]$, $h : [A] \to [C]$ and $k : [B] \to [D]$. Then

$$(f; h) \otimes_! \Delta, r_1.r_2 (g; k) = (f \otimes_! \Delta, r_1.r_2 g); (h \otimes k) : [!\Delta, \Gamma_1, \Gamma_2] \to [C] \otimes [D].$$

Proof. The proof uses the bifunctoriality of $\otimes$. \hfill $\Box$

Lemma 11.1.12. Suppose that $f : [!\Delta, \Gamma_1] \to [A]$, $g : [!\Delta, \Gamma_2] \to [B]$. Then

$$(f \otimes_! \Delta g); \sigma_{A,B} = (id_! \otimes \sigma_{[\Gamma_1],[\Gamma_2]}); (g \otimes_! \Delta f) : [!\Delta, \Gamma_1, \Gamma_2] \to [B \otimes A].$$

Proof. The proof is done using the correspondence between the symmetry and the comonoid multiplication. \hfill $\Box$

Definition 11.1.13. The map $\Theta$ as in Definition 11.1.2 is said to be an interpretation of the language if moreover it assigns to each constant term $c : !A_c$ an arrow $\Theta(c) : \top \to [!A_c]$ in $C$.

Given a linear category for duplication $C$, it is possible to interpret the typing derivation of a well-typed value as a map in $C$ and the typing derivation of a valid computation as a map in the Kleisli category $C_T$.

- If $x_1 : A_1, \ldots, x_n : A_n |\triangleright V : B$ is a value with typing derivation $\pi$, its value interpretation $[\pi]_\Theta^V$ is an arrow $[A_1] \otimes \ldots \otimes [A_n] \to [B]$;

- if $x_1 : A_1, \ldots, x_n : A_n |\triangleright M : A$ is a term with typing derivation $\pi$, its computational interpretation $[\pi]_\Theta^V$ is an arrow $[A_1] \otimes \ldots \otimes [A_n] \to \Gamma T([B])$.

We define them inductively in Tables 11.1 [11.1.2] and 11.3. Moreover, considering Convention 11.1.4 if $s$ is a permutation on $\{1, \ldots, n\}$:

$$[x_{s(1)} : A_{s(1)}, \ldots, x_{s(n)} : A_{s(n)} |\triangleright M : A]_\Theta^V =$$

$$\sigma_{[A_1, \ldots, A_n], [A_{s(1)}, \ldots, A_{s(n)}]} ; [x_1 : A_1, \ldots, x_n : A_n |\triangleright M : A]_\Theta^V,$$

$$[x_{s(1)} : A_{s(1)}, \ldots, x_{s(n)} : A_{s(n)} |\triangleright M : A]_\Theta^V =$$

$$\sigma_{[A_1, \ldots, A_n], [A_{s(1)}, \ldots, A_{s(n)}]} ; [x_1 : A_1, \ldots, x_n : A_n |\triangleright M : A]_\Theta^V.$$

Convention 11.1.14. We say that two morphisms

$$f : [x_1 : A_1, \ldots, x_n : A_n] \to [A], \quad g : [x_{s(1)} : A_{s(1)}, \ldots, x_{s(n)} : A_{s(n)}] \to [A]$$

are equal whenever $f = \sigma_{[A_1, \ldots, A_n], [A_{s(1)}, \ldots, A_{s(n)}]} ; g$. 

...
11.1. Denotational Semantics

Given a valid typing judgement

\[ \Gamma, x : A \vdash x : B \]

\[ \Delta \vdash c : B \]

\[ \Delta \vdash \ast : !n \top \]

Proof.

exists \( \Lambda \), Lemma 11.1.15. Suppose that \( \Delta \vdash M : B \). Moreover, if \( \Delta \vdash M : B \), then \( \Delta \vdash x : A \vdash P : B \).

Case \( \pi \),: \( \pi \vdash \Delta \vdash M : B \).

\[ \Delta \vdash \lambda^x A, \Delta \vdash !n + 1 A : B \]

\[ \Delta \vdash \lambda^x A, \Delta \vdash !n + 1 (A \rightarrow B) \]

\[ \Delta \vdash \lambda^x A, \Delta \vdash !n + 1 (A \rightarrow B) \]

\[ \Gamma, \Delta, \Gamma_1 \vdash \text{let } x^A = N \text{ in } P : B \]

\[ \Gamma, \Delta, \Gamma_1, \Gamma_2 \vdash x^A \]

\[ \Gamma, \Delta, \Gamma_1 \vdash N : A \]

\[ \Gamma, \Delta, \Gamma_2 \vdash x : A \vdash P : B \]

\[ \Gamma, \Delta, \Gamma_1, \Gamma_2 \vdash \text{let } x^A = N \text{ in } P : B \]

\[ \Gamma, \Delta, \Gamma_1, \Gamma_2 \vdash \lambda^x A, \Delta \vdash (\lambda^x A, P) : B \]

Table 11.1: Interpretation of core values.

Lemma 11.1.15. Suppose that

\[ f = \Gamma, \Delta, \Gamma_1 \vdash N : A \]

\[ g = \Gamma, \Delta, \Gamma_2, x : A \vdash P : B \]

Then

\[ \Gamma, \Delta, \Gamma_1, \Gamma_2 \vdash \text{let } x^A = N \text{ in } P : B \]

Proof. The proof is done by computing \( \Gamma, \Delta, \Gamma_1, \Gamma_2 \vdash (\lambda^x A, P) : B \).

Lemma 11.1.16. Given a valid typing judgement \( \Delta \vdash M : B \) with no dummy variables, there exists \( f_c : [\Delta] \rightarrow [B] \) such that for any dummy context \( \Lambda \) and for any typing derivation \( \pi \) of !\( \Lambda \), \( \Delta \vdash M : B \),

\[ [\pi]^c = [\Lambda] \otimes [\Delta] \rightarrow [B] \]

Moreover, if \( M = V \) is a value, there exists some \( f_c : [\Delta] \rightarrow [B] \) such that for any dummy context !\( \Lambda \) and for any typing derivation \( \pi \) of !\( \Lambda \), \( \Delta \vdash V : B \),

\[ [\pi]^c = [\Lambda] \rightarrow [B] \]

Proof. Consider a typing judgement \( \Delta \vdash M : A \) where \( |\Delta| = FV(M) \). We prove the lemma by structural induction on \( M \).

Case \( M = \lambda^x A, N \). In this case, \( A = B \rightarrow C \) and \( \Delta \vdash \lambda^x A, N : B \rightarrow C \) is valid, then so is \( \Delta, x : B \vdash N : C \). By hypothesis, \( |\Delta| = FV(M) \). Thus \( |\Delta| = FV(N) \setminus \{x\} \).

Let \( \Delta \) be the subcontext of \( (\Delta, x : B) \) consisting of the free variables of \( N \).

The induction hypothesis states that there exists a function \( f_c : [\Delta] \rightarrow [B] \) such that for all dummy typing contexts !\( \Lambda \) and for all typing derivations \( \pi \) the denotation \( [\pi]^c \) is \( \otimes f_c \).

There are two cases.

If \( x \in FV(N) \). Then \( |\Delta| = |\Delta, x : B| \). We define the requested \( f_c \) and \( f_c \) respectively by \( \Phi^{-1}(f_c) \) and \( \Phi^{-1}(f_c) \).

If \( x \notin FV(N) \). In this case, from Lemma 9.1.19 the type \( B \) is of the form !\( \hat{B} \). The context \( |\Delta| = |\Delta|, x \text{ is a dummy variable. We define the requested } f_c \) and \( f_c \) respectively by \( \Phi^{-1}(\otimes \hat{B}) \) and \( \Phi^{-1}(\otimes \hat{B}) \).
Consider a dummy context !Λ for M and a typing derivation π of ![\Delta, \Gamma \triangleright M : A]\_\text{11.1. Denotational Semantics}$. It starts with

$$
[\Delta, x : B \triangleright N : C]_{\text{11.1. Denotational Semantics}}
$$

for some derivation π'. By definition,

$$
[\pi']^v = \Phi^{-1}([\pi']^v) \quad \quad [\pi']^c = \Phi^{-1}([\pi']^c) : \eta_{B \rightarrow C}.
$$

We aim at showing that $[\pi']^v = \diamond \gamma A \otimes f_v$ and that $[\pi']^c = \diamond \gamma A \otimes f_c$.

There are two cases.

**If** $x \in \text{FV}(N)$. By induction hypothesis, $[\pi']^c = \diamond \gamma A \otimes f'_c$. By naturality of $\Phi$,

$$
[\pi']^v = \Phi^{-1}(\diamond \gamma A \otimes f'_c) : \eta_{B \rightarrow C} = \diamond \gamma A \otimes \Phi^{-1}(f'_c) : \eta_{B \rightarrow C},
$$

$$
[\pi']^v = \Phi^{-1}(\diamond \gamma A \otimes f'_c)
$$

**If** $x \notin \text{FV}(N)$. By induction hypothesis, $[\pi']^c = \diamond \gamma A \otimes \diamond \gamma B \otimes f'_c$. By naturality of $\Phi$,

$$
[\pi']^v = \Phi^{-1}(\diamond \gamma A \otimes \diamond \gamma B \otimes f'_c) : \eta_{B \rightarrow C} = \diamond \gamma A \otimes \Phi^{-1}(\diamond \gamma B \otimes f'_c) : \eta_{B \rightarrow C},
$$

$$
[\pi']^v = \Phi^{-1}(\diamond \gamma A \otimes \diamond \gamma B \otimes f'_c)
$$
First, if $U$ is a core value, $\sem{\Delta \triangleright U : A}_\theta = [\Delta \triangleright U : A]_\theta^\eta_A$.

$$
\begin{align*}
[\Delta, \Gamma_1 \triangleright M : A \rightarrow B]_\theta^c &= \sem{\Delta, \Gamma_1} \xrightarrow{L} T([A \rightarrow B]) \\
[\Delta, \Gamma_2 \triangleright N : A]_\theta^c &= \sem{\Delta, \Gamma_2} \xrightarrow{g} T([A]) \\
[\Delta, \Gamma_1, \Gamma_2 \triangleright MN : B]_\theta^c &= \sem{\Delta, \Gamma_1, \Gamma_2} \xrightarrow{\text{fsd}^g_{ \Gamma_2 \triangleright [A \rightarrow B]} T[A \rightarrow B] \otimes T[A]} T[B] \\
[\Delta, \Gamma_1 \triangleright M : !^n(A_1 \otimes A_2)]_\theta^c &= \sem{\Delta, \Gamma_1} \xrightarrow{L} T[L^n[A_1 \otimes A_2]] \\
[\Delta, \Gamma_2, x : !^nA_1, y : !^nA_2 \triangleright N : C]_\theta^c &= \sem{\Delta, \Gamma_2} \otimes T[L^n[A_1] \otimes L^n[A_2]] \xrightarrow{g} T[C] \\
[\Delta, \Gamma_1, \Gamma_2 \triangleright \text{let } * = M \text{ in } N : C]_\theta^c &= \sem{\Delta, \Gamma_1, \Gamma_2} \xrightarrow{\text{fsd}^g_{ \Gamma_2 \triangleright [A \rightarrow B]} T[L^n[A_1] \otimes L^n[A_2]]} T[C] \\
[\Delta, \Gamma_1 \triangleright M : \top]_\theta^c &= \sem{\Delta, \Gamma_1} \xrightarrow{L} T(\top) \\
[\Delta, \Gamma_2 \triangleright N : C]_\theta^c &= \sem{\Delta, \Gamma_2} \xrightarrow{g} T[C] \\
[\Delta, \Gamma_1, \Gamma_2 \triangleright \text{let } * = M \text{ in } N : C]_\theta^c &= \sem{\Delta, \Gamma_1, \Gamma_2} \xrightarrow{\text{fsd}^g_{ \Gamma_2 \triangleright [A \rightarrow B]} T(\top) \otimes T[C]} \xrightarrow{\Psi_1 Td^\eta_{L^n[B]}} T[C] \\
[\Delta, \Gamma_1 \triangleright M : !^nA]_\theta^c &= \sem{\Delta, \Gamma_1} \xrightarrow{L} T[L^n[A]] \\
[\Delta, \Gamma_2 \triangleright N : !^nB]_\theta^c &= \sem{\Delta, \Gamma_2} \xrightarrow{g} T[L^n[B]] \\
[\Delta, \Gamma_1, \Gamma_2 \triangleright \text{let } * = M \text{ in } N : !^n(A \otimes B)]_\theta^c &= \sem{\Delta, \Gamma_1, \Gamma_2} \xrightarrow{\text{fsd}^g_{ \Gamma_2 \triangleright [A \rightarrow B]} T[L^n[A] \otimes T[L^n[B]]} T[L^n[A \otimes B]]
\end{align*}
$$

Table 11.3: Interpretation of computations.

**Case** $M \equiv \lambda^{n+1}x^B . N$. In this case, $A = !^{n+1}(B \rightarrow C)$, and if $\Delta \triangleright \lambda^{n+1}x^B . N : B \rightarrow C$ is valid, then so is $\Delta, x : B \triangleright N : C$. By hypothesis, $|\Delta| = FV(M)$. Thus $|\Delta| = FV(N) \setminus \{x\}$.

Let $\Delta$ be the subcontext of $(x, B, \Delta)$ consisting of the free variables of $N$.

The induction hypothesis states that there exists a function $f_c' : [\Delta] \rightarrow T[C]$ such that for all dummy typing contexts $!A$ and for all typing derivations $\pi'$ the denotation $[\pi']^c$ is $\otimes f_c'$. Note $\Delta$ is of the form $!\Delta$. There are two cases.

**If** $x \in FV(N)$. Then $|\Delta| = [x : B, \Delta]$. We define the requested $f_v$ and $f_c$, respectively by

$$
\begin{align*}
f_v &= \delta_\Delta ; d^L_{\Delta} : L \Phi^{-1}(f'_c) ; I_{(B \rightarrow C)^{\eta^{n+1}(B \rightarrow C)}} \\
f_c &= \delta_\Delta ; d^L_{\Delta} : L \Phi^{-1}(f'_c) ; I_{(B \rightarrow C)^{\eta^{n+1}(B \rightarrow C)}}
\end{align*}
$$

**If** $x \not\in FV(N)$. In this case, from Lemma [11.40] the type $B$ is of the form $!\hat{B}$. The context $|\Delta| = |\Delta|$, and $x$ is a dummy variable. We define the requested $f_v$ and $f_c$, respectively by

$$
\begin{align*}
f_v &= \delta_\Delta ; d^L_{\Delta} : L \Phi^{-1}(\otimes_{!B} f_c') ; I_{(B \rightarrow C)^{\eta^{n+1}(B \rightarrow C)}} \\
f_c &= \delta_\Delta ; d^L_{\Delta} : L \Phi^{-1}(\otimes_{!B} f_c') ; I_{(B \rightarrow C)^{\eta^{n+1}(B \rightarrow C)}}
\end{align*}
$$
Consider a dummy context !Λ for M and a typing derivation π of !Λ, Δ ▷ M : A. It starts with

\[ \pi' \]

\[ !\Lambda, \Delta, x : B \triangleright N : C \]

\[ !\Lambda \triangleright \lambda^{n+1} x^B : N : B \rightarrow C \]

(λ2)

for some derivation π'. By definition,

\[ [\pi]^{\nu} = \delta_{!\Lambda, \Delta} \cdot d_{!\Lambda, \Delta}^{\nu} : L\Phi^{-1}([\pi]^{\nu}) ; I(\lambda^{B \rightarrow C}) \cdot \iota^{n+1}(B \rightarrow C), \]

\[ [\pi]^{c} = \delta_{!\Lambda, \Delta} \cdot d_{!\Lambda, \Delta}^{c} : L\Phi^{-1}([\pi]^{c}) ; I(\lambda^{B \rightarrow C}) \cdot \iota^{n+1}(B \rightarrow C). \]

We aim at showing that \([\pi]^{\nu} = \Diamond_{!\Lambda} \circ f_c\) and that \([\pi]^{c} = \Diamond_{!\Lambda} \circ f_c\).

One can show that for all g : \([\Delta] \rightarrow [D]\),

\[ \delta_{!\Lambda, \Delta} \cdot d_{!\Lambda, \Delta}^{i} : L(\Diamond_{!\Lambda} \circ \iota^{\Delta} ; L(id \circ g) = \Diamond_{!\Lambda} \circ (\delta_{!\Lambda} ; d_{!\Lambda}^{i} ; L(id \circ Lg)) \] (11.1.1)

by using the monoidality of \(d_L\) and \(\delta\), and by using the fact that the category is strict. There are two cases.

**If** \(x \in FV(N)\). By induction hypothesis, \([\pi]^{\nu} = \Diamond_{!\Lambda} \circ f_c\). By naturality of \(\Phi\), \(\Phi^{-1}(\Diamond_{!\Lambda} \circ f_c) = \Diamond_{!\Lambda} \circ \Phi^{-1}(f_c)\). One can then conclude, by using the functoriality of \(L\) and \(\circ\) and replacing \(g\) with \(\Phi^{-1}(f_c)\) in Equation 11.1.1.

**If** \(x \not\in FV(N)\). By induction hypothesis, \([\pi]^{c} = \Diamond_{!\Lambda} \circ \Diamond_{!\Gamma} \circ f_c\).

By naturality of \(\Phi\), \(\Phi^{-1}(\Diamond_{!\Lambda} \circ \Diamond_{!\Gamma} \circ f_c) = \Diamond_{!\Lambda} \circ \Phi^{-1}(\Diamond_{!\Gamma} \circ f_c)\). One can then conclude, by using the functoriality of \(L\) and \(\circ\) and replacing \(g\) with \(\Phi^{-1}(\Diamond_{!\Gamma} \circ f_c)\) in Equation 11.1.1.

**Case** \(M \equiv NP\). In this case, one can split \(\Delta\) into \((!\Delta, \Gamma_1, \Gamma_2)\), where \(\Gamma_1\) consists of all the free variables in \(N\) but not in \(P\), \(\Gamma_2\) of all the free variables in \(P\) but not in \(N\), and \(!\Delta\) of all the free variables both in \(N\) and in \(P\). Note that these are the only possibilities, since \(\Delta\) does not contain any dummy variables.

Using Lemma [11.1.14] If \(\Delta \triangleright NP : A\), then there exists some \(B\) such that \(!\Delta, \Gamma_1 \triangleright N : B \rightarrow A\) and \(!\Delta, \Gamma_2 \triangleright P : B\).

By induction hypothesis, there exists \(f_c' : \([!\Delta] \circ [\Gamma_1] \rightarrow T[B \rightarrow A]\) such that for all dummy contexts ! for \(N\) and for all typing derivations \(\pi_1\) for \(!\Lambda, !\Delta, \Gamma_1 \triangleright N : B \rightarrow A\) the denotation \([\pi_1]^{\nu} = \Diamond_{!\Lambda} \circ f_c'\).

Similarly, by induction hypothesis, there exists \(g_c' : \([!\Delta] \circ [\Gamma_2] \rightarrow T[B]\) such that for all dummy contexts ! for \(P\) and for all typing derivations \(\pi_2\) for \(!\Lambda, !\Delta, \Gamma_2 \triangleright P : B\) the denotation \([\pi_2]^{c} = \Diamond_{!\Lambda} \circ g_c'\).

Let \(f_c\) be the map

\[ \[!\Delta\] \circ [\Gamma_1] \circ [\Gamma_2] \xrightarrow{f_c' \circ \Delta} T[B \rightarrow A] \circ T[B] \xrightarrow{\Phi, \circ} T([B \rightarrow A] \circ [B]) \xrightarrow{\varepsilon_{A,B}} T[A]. \]

Consider a dummy context !\ for \(NP\) and a valid derivation π for \(!\Lambda, \Delta \triangleright NP : A\). This derivation starts with

\[ \pi_1 \]

\[ \cdots \]

\[ !\Lambda, \Delta, \Gamma_1 \triangleright N : B' \rightarrow A \]

\[ !\Lambda, \Delta, \Gamma_2 \triangleright P : B' \]

(app)
The context \((!\Lambda, \Delta)\) is equal to \((!\hat{\Delta}, \hat{\Gamma}_1, \hat{\Gamma}_2)\). From Lemma 11.1.12, \(B' = B\). From Lemma 11.1.11 we have \(!|\Delta| \subseteq |!\hat{\Delta}|, |!\Gamma_1| \subseteq |!\hat{\Gamma}_1|\) and \(!|\Gamma_2| \subseteq |!\hat{\Gamma}_2|\). Let us split \(!\hat{\Delta}, \hat{\Gamma}_1\) and \(\hat{\Gamma}_2\) into

\[
\begin{align*}
!\hat{\Delta} &= (!d^{NP} \hat{\Delta}, !d^{NP} \hat{\Gamma}_1, !d^{NP} \hat{\Gamma}_2), \\
\hat{\Gamma}_1 &= (\hat{\Gamma}_1^{dNP}, \hat{\Gamma}_1^{dNP}), \\
\hat{\Gamma}_2 &= (\hat{\Gamma}_2^{dNP}, \hat{\Gamma}_2^{dNP}),
\end{align*}
\]

where for each context of the form \(X^{dY_1 \ldots Y_n Z_1 \ldots Z_i}\), if \(x \in |X^{dY_1 \ldots Y_n Z_1 \ldots Z_i}|\), then \(x \in FV(Z_1) \cap \ldots \cap FV(Y_i)\) but \(x \notin FV(Y_1) \cup \ldots \cup FV(Y_k)\). We did not include the four contexts \(\hat{\Gamma}_1^{dNP}, \hat{\Gamma}_2^{dNP}, \hat{\Gamma}_1^{dNPN}\) and \(\hat{\Gamma}_2^{dNPN}\), which are empty by definition: otherwise the context \((\Delta, \hat{\Gamma}_1, \hat{\Gamma}_2)\) would have duplicates. We can rewrite the following contexts:

\[
\begin{align*}
!\Lambda &= (d^{NP}, !d^{NP}, !d^{NP}, !d^{NP}), \\
!\hat{\Delta} &= !d^{NP},
\end{align*}
\]

Note that \(\hat{\Gamma}_1^{dNP}, \hat{\Gamma}_1^{dNPN}\), and \(\hat{\Gamma}_2^{dNPN}\), are duplicable.

By definition, \([\pi]^c = ([\pi_1]^c \otimes !!\Delta) \otimes [\pi_2]^c\); \(\Psi_1; \varepsilon^A_{A,B}\). We want to show that \([\pi]^c = \otimes !\Lambda \otimes f_c\). We have from the induction hypothesis that if \(f'_c\) is the morphism \(\otimes !\Delta^{dNP}, \otimes !\Delta^{dNP}, \otimes !\Delta^{dNPN}\) and if \(g'_c\) is the morphism \(\otimes !\Delta^{dNP}, \otimes !\Delta^{dNP}, \otimes !\Delta^{dNPN}\), then \([\pi_1]^c\) and \([\pi_2]^c\) are respectively \(f'_c \otimes f_c\) and \(g'_c \otimes g_c\), and

\[
\begin{align*}
f'_c : [!d^{NP}, !d^{NP}, !d^{NP}, !d^{NP}] &\rightarrow T, \\
g'_c : [!d^{NP}, !d^{NP}, !d^{NP}, !d^{NP}] &\rightarrow T[B \rightarrow A], \\
f'_c : [!d^{NP}, !d^{NP}, !d^{NP}, !d^{NP}] &\rightarrow T[B].
\end{align*}
\]

Now, since from Lemma 11.1.3 and Lemma 11.1.10

\[
(\otimes !\Delta^{dNP} \otimes \otimes !\Delta^{dNP} \otimes \otimes !\Delta^{dNP} \otimes f'_c) \\
\otimes !\Delta^{dNP} \otimes !\Delta^{dNP} \otimes !\Delta^{dNP} \otimes g'_c
\]

the denotation \([\pi]^c\) is then equal to

\[
[\pi]^c = \left(\otimes !\Delta^{dNP} \otimes \otimes !\Delta^{dNP} \otimes \otimes !\Delta^{dNP} \otimes (f'_c \otimes !\Delta^{dNP} g'_c)\right) ; \Psi_1; \varepsilon^A_{A,B}
\]

One concludes using Lemma 11.1.10.

The remaining cases are handled similarly. 

**Theorem 11.1.17.** Given a valid typing judgement \(\Delta \triangleright M : B\) with two typing derivations \(\pi\) and \(\pi'\), for any interpretation \(\Theta\) we have \([\pi]^c = [\pi']^c\) (and \([\pi]^v = [\pi']^v\) if \(M\) is a value).

**Proof.** We consider denotations with respect to \(\Theta\). The context \(\Delta\) splits into \((!\Lambda, \Delta)\), where \(!\Lambda\) is a dummy context and where \(|\Delta| = FV(M)|. From Lemma 11.1.11 \(\Delta \triangleright M : A\) is valid. We are in the context of Lemma 11.1.11 there exists a function \(f_c : [\Delta] \rightarrow [T[A]]\) such that \([\pi]^v\) and \([\pi']^v\) are both equal to \(\otimes !\Lambda \otimes f_c\). If \(M\) is a value, then there exists a function \(f_v : [\Delta] \rightarrow [A]\) such that \([\pi]^v\) and \([\pi']^v\) are both equal to \(\otimes !\Lambda \otimes f_v\). 

11.2 Soundness of the Denotation

Definition 11.1.18. Given an interpretation $\Theta$ of the language in a category $\mathcal{C}$, we define the denotation of a valid typing judgement $\Delta \triangleright M : A$ with typing derivation $\pi$ to be $[\Delta \triangleright M : A]^c_{\Theta} = [\pi]_{\Theta}^c$ and $[\Delta \triangleright M : A]^v_{\Theta} = [\pi]_{\Theta}^v$ if $M$ is a value.

Remark 11.1.19. From Theorem 11.1.17, Definition 11.1.18 is valid.

11.2 Soundness of the Denotation

The axiomatic equivalence and the categorical semantics are two faces of the same coin. Indeed, as we will prove in this section, two terms in the same axiomatic equivalence class have the same denotation. A corollary is that the indexation of terms does not influence the denotation. This proves semantically the fact that it is safe to work with untyped terms. An alternate justification of this fact is of course the operational semantics, which was given in Section 6.

Lemma 11.2.1. Suppose that $\Delta' \triangleright \Delta$ and that $\Delta \triangleright M : A$. Then

$$[\Delta' \triangleright M : A]^c = I_{\Delta', \Delta}; [\Delta \triangleright M : A']^c.$$  

If $M = V$ is a value, we have moreover

$$[\Delta' \triangleright V : A]^v = I_{\Delta', \Delta}; [\Delta \triangleright V : A']^v.$$  

Proof. Proof by induction on the size of $M$.  

Lemma 11.2.2. Suppose that $\Delta \triangleright M : A$ and that $A \triangleright A'$. Then

$$[\Delta \triangleright \{ M <: A' \} : A']^c = [\Delta \triangleright M : A]^c; T(I_{A, A'}).$$  

If $M = V$ is a value, from Lemma 9.1.27, $(V <: A')$ is a value. Then

$$[\Delta \triangleright \{ V <: A' \} : A']^v = [\Delta \triangleright V : A]^v; I_{A, A'}.$$  

Proof. Proof by induction on the size of $M$.  

Lemma 11.2.3. Suppose that $\Delta \triangleright M : A$ is a valid typing judgement and that $!\Lambda$ is a context of dummy variables for $M$. Then

$$[!\Lambda, \Delta \triangleright M : A]^c = \diamond \Lambda \odot [\Delta \triangleright M : A]^c.$$  

If $M = V$ is a value,

$$[!\Lambda, \Delta \triangleright V : A]^v = \diamond \Lambda \odot [\Delta \triangleright V : A]^v.$$  

Proof. Proof by induction on the size of $M$.  

Lemma 11.2.4 (Substitution). Given two valid typing judgements $!\Delta, \Gamma_1, x : A \triangleright M : B$ and $!\Delta, \Gamma_2 \triangleright V : A$, the typing judgement $!\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B$ is valid. Let $h$ be the denotation $![\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B]^c$ and $h'$ be $![\Delta, \Gamma_1, \Gamma_2 \triangleright W[V/x] : B]^v$, in the case where $M = W$ is a value. Then they are defined by

$$![\Delta] \odot [\Gamma_1] \odot [\Gamma_2] \xrightarrow{\text{Split}_{\Delta, \Gamma_1, \Gamma_2}} T([B])$$  

$$![\Delta] \odot [\Gamma_1] \odot ![\Delta] \odot [\Gamma_2] \xrightarrow{\text{split}[\Delta, \Gamma_1, \Gamma_2 \triangleright V : A]^v} ![\Delta] \odot ![\Gamma_1] \odot [A].$$
Proof. Let \( v \) be the map \([!\Delta, \Gamma_2 \triangleright V : A]^v\). We do the proof by induction on the size of \( M \).

Cases where \( M \) is a core value. Case \( M \equiv x^B \). The term \( M \) is a core value; we have to show Equation (11.2.1) and Equation (11.2.2). We show the former, the latter being exactly similar.

The typing derivation \(!\Delta, \Gamma_1, x : A \triangleright x^B : B\) comes from typing rule \((ax_1)\). This means that \( \Gamma_1 \) is of the form \(!\Gamma'_1\), and that \( A \triangleright B \). Using Lemma 11.2.2, the map \( h \) is
\[
h = [!\Delta, \Gamma'_1, \Gamma_2 \triangleright \{V < : B\} : B]^v; \eta_A; T(I_{A,B}).
\]
Since \( V \) is a core value and using Lemma 11.2.3 and 11.1.7
\[
h = Split_{!\Delta, \Gamma'_1, \Gamma_2}; (\hat{\Delta}_{\Gamma'_1} \otimes [!\Delta, \Gamma_2 \triangleright V : A]^v; \eta_A; T(I_{A,B}))
\]
By definition
\[
[!\Delta, \Gamma'_1, x : A \triangleright x^B : B]^v = \hat{\Delta}_{\Gamma'_1} \otimes I_{A,B}; \eta_B.
\]
Thus
\[
Split_{!\Delta, \Gamma'_1, \Gamma_2}; (id \otimes [!\Delta, \Gamma_2 \triangleright V : A]^v; [!\Delta, \Gamma'_1, x : A \triangleright x^B : B]^v)
\]
\[
= Split_{!\Delta, \Gamma'_1, \Gamma_2}; (id \otimes [!\Delta, \Gamma_2 \triangleright V : A]^v; \hat{\Delta}_{\Gamma'_1} \otimes I_{A,B}; \eta_B)
\]
\[
= Split_{!\Delta, \Gamma'_1, \Gamma_2}; (\hat{\Delta}_{\Gamma'_1} \otimes ([!\Delta, \Gamma_2 \triangleright V : A]^v; I_{A,B}; \eta_B))
\]
(bifunc. of \( \otimes \))

which is equal to Formula (11.2.3), the naturality of \( \eta \).

Case \( M \equiv x^n, c^B \) and \( y^B \), where \( y \neq x \). Again in these cases, the term \( M \) is a core value: we have to show Equation (11.2.1) and Equation (11.2.2). We show the former, the latter being exactly similar.

The typing judgement \(!\Delta, \Gamma_1, x : A \triangleright M : B\) comes from respectively rules \((\Gamma.J)\), \((ax_2)\) and \((ax_1)\). In all of these cases, \( \Gamma_1 \) is of the form \(!\Gamma'_1\) and \( A \) of the form \(!A'\). From Lemma 9.1.16, the context \( \Gamma_2 \) is of the form \( !\Gamma'_2 \). Since the term \( M[V/x] \) is equal to \( M \) is all three cases, the maps \( h \) and \( h' \) are of the form \( \hat{\Delta}_{!\delta', \eta'} : g, \) for some map \( g \). The map \([!\Delta, \Gamma'_1, x : !A' \triangleright M : B]\) is of the form \( \hat{\Delta}_{\Gamma'_1, A'} : g \) for the same map \( g \). The bottom part of Diagram (11.2.1) is
\[
Split_{!\Delta, \Gamma'_1, \Gamma_2}; (id \otimes [!\Delta, \Gamma'_2 \triangleright V : A]^v; \hat{\Delta}_{\Gamma'_1, A'} : g) = \hat{\Delta}_{\Gamma'_1, \Gamma_2} : g
\]
using Equation (2.7.1). This is precisely the map \([!\Delta, \Gamma'_1, x : !A' \triangleright M : B]\), ending the proof for this case.

Case \( M \equiv x^B y^C.N \). In this case, \( B = C \rightarrow D \) for some type \( D \), and the typing judgement \(!\Delta, \Gamma_1, x : A \triangleright M : B\) comes from typing rule \((\lambda_1)\). That is, \(!\Delta, \Gamma_1, y : C, x : A \triangleright N : C\) is valid.

We have \( y \neq x \); this means that \( M[V/x] = x^B y^C.N[V/x] \). By induction hypothesis, the morphism \( k = [!\Delta, \Gamma_1, y : C, \Gamma_2 \triangleright N[V/x] : D]^v \) satisfies the following commutative diagram:

\[
[!\Delta, \Gamma_1, y : C, \Gamma_2] \xrightarrow{Split_{!\Delta, \Gamma_1, y : C, \Gamma_2}} [!\Delta, \Gamma_1, y : C, \Gamma_2 \triangleright V : A]^v
\]

\[
\xrightarrow{T([V/x])}
\]

\[
[!\Delta, \Gamma_1, y : C] \otimes [!\Delta, \Gamma_2] \xrightarrow{id \otimes ![\Delta, \Gamma_2 \triangleright V : A]^v} [!\Delta, \Gamma_1, y : C] \otimes [A],
\]
Now, by definition,
\[
\begin{align*}
[[!\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : C \rightarrow D]]^v &= \Phi^{-1}(k), \\
[[!\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : C \rightarrow D]]^c &= \Phi^{-1}(k) \cdot \mu_{C \rightarrow D}, \\
[[!\Delta, \Gamma_1 \triangleright M : C \rightarrow D]]^v &= \Phi^{-1}([[!\Delta, \Gamma_1, y : C \triangleright N : D]]^v), \\
[[!\Delta, \Gamma_1 \triangleright M : C \rightarrow D]]^c &= (\Phi^{-1}([[!\Delta, \Gamma_1, y : C \triangleright N : D]]^c) \cdot \mu_{C \rightarrow D}).
\end{align*}
\]

Equations (11.2.1) and (11.2.2) are derived using the naturality of \(\Phi\).

**Case** \(M \equiv \lambda^{\alpha+1} \cdot x^C : N\). This case is similar to the previous one. The only difference is in that the equations under consideration are encapsulated into
\[
f \mapsto \delta_{!\Delta}; d^!_{\Delta}; L(f); \Pi_{(A \rightarrow B), (\alpha+1)(A \rightarrow B)}.
\]

**Cases** \(M \equiv (let \ x = N in P)\). In all these cases, the typing judgement \(!\Delta, \Gamma_1, x : A \triangleright M : B\) comes from a typing rule \((X)\) with hypotheses
\[
!\Delta', \Gamma_{11} \triangleright N : C, \quad !\Delta', \Gamma_{12}, \Lambda \triangleright P : B,
\]
for some context \(\Lambda\), where \((!\Delta', \Gamma_{11}, \Gamma_{12}) = (!\Delta, \Gamma_1, x : A)\). If we call \(f\) and \(g\) the following maps:
\[
f = [[!\Delta', \Gamma_{11} \triangleright N : C]]^v, \quad g = [[!\Delta', \Gamma_{12}, \Lambda \triangleright P : B]]^c,
\]
we have
\[
[[!\Delta, \Gamma_1, x : A \triangleright M : B]]^v = (f \otimes_{!\Delta} id) \cdot (T(F) \otimes id); \sigma; t; g^* \quad \text{for a given map } F, \text{ depending on } \delta_{!\Delta}.
\]

If \(M\) is an extended value, so are \(N\) and \(P\). If we call \(f'\) and \(g'\) the following maps:
\[
f' = [[!\Delta', \Gamma_{11} \triangleright N : C]]^v, \quad g' = [[!\Delta', \Gamma_{12}, \Lambda \triangleright P : B]]^v,
\]
we have
\[
[[!\Delta, \Gamma_1, x : A \triangleright M : B]]^v = (f' \otimes_{!\Delta} id) \cdot (F \otimes id); \sigma; g'.
\]

Note that we might need to add a permutation in front of \(f, g, f'\) or \(g'\), depending on the order of their respective input.

By \(\alpha\)-equivalence one can assume that the variables occurring in \(\Lambda\) does not occur anywhere in the other contexts. From Lemma 11.2.14 one can assume that \(\Gamma_{11}\) and \(\Gamma_{12}\) do not contain any duplicable variables, implying that \(\|!\Delta\| \subseteq \|!\Delta'\|\). Thus there exists some context \(!\Psi\) such that \(!\Delta' = (!\Psi, !\Delta)\). In particular, this means that
\[
(!\Psi, \Gamma_{11}, \Gamma_{12}) = (\Gamma_1, x : A).
\]

This equation is the main point that will allow us to apply the induction hypothesis, and thus finish the proof. The variable \(x : A\) belongs to either \(!\Psi, \Gamma_{11}\) or \(\Gamma_{12}\). We study each one of these three possibilities.

If \(x \in \|!\Psi\|\). In this case, \(A = !A'\) and \(!\Psi = (\Psi', x : !A')\). The contexts \((!\Delta', \Gamma_{11})\) and \((!\Delta', \Gamma_{12}, \Lambda)\) becomes respectively \((!\Delta, !\Psi', \Gamma_{11}, x : !A')\) and \((!\Delta, !\Psi', \Gamma_{12}, \Lambda, x : !A')\), meaning that the typing judgements \(!\Delta, !\Psi', \Gamma_{11}, x : !A' \triangleright N : C\) and \(!\Delta, !\Psi', \Gamma_{12}, \Lambda, x : !A' \triangleright P : B\) are valid.
11.2 Soundness of the Denotation

Since $A = !\Delta$ and since $V$ is a value, by Lemma L1.13, the context $\Gamma_2$ is of the form $!\Gamma'_2$. We have $M[V/x] = (let \Gamma' = N[V/x] in P[V/x])$, and

$$M[V/x] : C, \; \Delta, !\Psi, \Gamma_1, \Gamma_2, !\Gamma'_2 : B$$

are valid typing judgements, such that by induction hypotheses their respective computational denotations $\hat{f}$ and $\hat{g}$ are

$$\hat{f} \rightarrow T[C]$$

$$\hat{g} \rightarrow T[B]$$

By definition, the denotation $h = [\Delta, !\Delta, !\Gamma'_2, !\Gamma'_2] : M[V/x] : B$ is equal to

$$[\Delta, !\Delta, !\Psi', \Gamma_1, \Gamma_2, !\Gamma'_2] \xrightarrow{sw_1; Split; sw_2} [\Delta, !\Delta, !\Psi', \Gamma_1, \Gamma_2, !\Gamma'_2]$$

Using Equations L1.25 and L1.26, one can rewrite $h$ as follows:

$$[\Delta, !\Delta, !\Psi', \Gamma_1, \Gamma_2, !\Gamma'_2] \xrightarrow{\sigma \circ \sigma} [\Delta, !\Psi', \Gamma_1, \Gamma_2, !\Delta, !\Psi', !\Gamma'_2, !\Gamma'_2]$$

where $sw_1$ and $sw_2$ are permutations. Note that the map $Split_{\Delta, !\Psi', \Gamma_2, !\Gamma'_2}$ is $\Delta_\Delta$ followed by a permutation. It is then possible to lift the second occurrence of $v$ to the top, thus getting $h$ in the required format.

The computation for $h' = [\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B]^v$ is similar.
If \( x \in \Gamma_{11} \). This time, the context \( \Gamma_{11} \) is of the form \((\Gamma'_{11}, x : A)\), and \( \Gamma_1 \) can be split into \((\Delta, !\Psi, \Gamma'_{11}, \Gamma_{12})\). We have \( M[V/x] = (\text{let } \Box^2 = N[V/x] \text{ in } P) \), and

\[
\begin{align*}
!\Delta, !\Psi, \Gamma_{11}, \Gamma_2 &\triangleright N[V/x] : C, \quad !\Delta, !\Psi, \Gamma_{12}, \Delta &\triangleright P : B
\end{align*}
\]

are valid typing judgements. By induction hypothesis, the computational denotation of the former is

\[
\begin{array}{c}
\begin{array}{c}
\!\Delta, \!\Psi, \Gamma_{11}, \Gamma_2 \quad \vdash \quad T[C]
\end{array}
\end{array}
\]

Using the Equation (11.2.7), this is equal to

\[
\begin{array}{c}
\begin{array}{c}
\!\Delta, \!\Psi, \Gamma_{11}, \Gamma_2 \rightarrow \bigotimes \!\Delta, \!\Psi, \Gamma_{11}, \Gamma_2 \rightarrow \bigotimes \!\Delta, \!\Psi, \Gamma_{12}
\end{array}
\end{array}
\]

By definition, the denotation \( h = \bigotimes \Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B \bigotimes \) is equal to

\[
\begin{array}{c}
\begin{array}{c}
\!\Delta, \!\Psi, \Gamma'_{11}, \Gamma_1, \Gamma_2, \Gamma_{12} \rightarrow \bigotimes \!\Delta, \!\Psi, \Gamma'_{11}, \Gamma_{12}
\end{array}
\end{array}
\]

Using the Equation (11.2.7), this is equal to

\[
\begin{array}{c}
\begin{array}{c}
\!\Delta, \!\Psi, \Gamma'_{11}, \Gamma_1, \Gamma_2, \Gamma_{12} \rightarrow \bigotimes \!\Delta, \!\Psi, \Gamma'_{11}, \Gamma_{12}
\end{array}
\end{array}
\]

where \( \text{perm} \) is a permutation. Using the coherence properties of the symmetry and the comonoid multiplication, the map

\[
\text{perm}; \text{Split}\,(\Delta, \!\Psi, \Gamma'_{11}, \Gamma_1, \Gamma_{12}) ; \text{Split}\,(\Delta, \!\Psi, \Gamma'_{11}, \Gamma_{12}) ; (\text{id} \otimes \text{id}) ; (\text{id} \otimes \text{id}) \]

\[
\begin{array}{c}
\begin{array}{c}
\!\Delta, \!\Psi, \Gamma'_{11}, \Gamma_{12} \rightarrow \bigotimes \!\Delta, \!\Psi, \Gamma'_{11}, \Gamma_{12} \rightarrow \bigotimes \!\Delta, \!\Psi, \Gamma_{12}
\end{array}
\end{array}
\]

that is, \( (\text{id} \otimes \text{id}) \): the map \( h \) is in the requested form.

Again, the computation for \( h' = \bigotimes \Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B \bigotimes \) is similar.

Case \( x \in \Gamma_{12} \). The context \( \Gamma_{12} \) is of the form \((\Gamma'_{12}, x : A)\), and \( \Gamma_1 \) is equal to \((\Delta, !\Psi, \Gamma_{11}, \Gamma_{12})\). We have \( M[V/x] = (\text{let } \Box^2 = N \text{ in } P[V/x]) \), and

\[
\begin{array}{c}
\begin{array}{c}
!\Delta, !\Psi, \Gamma_{11} \triangleright N : C, \quad !\Delta, !\Psi, \Gamma_{12}, \Delta, \Gamma_2 \triangleright P[V/x] : B
\end{array}
\end{array}
\]
are valid typing judgements. By induction hypothesis, the computational denotation of the latter is

\[
\begin{align*}
\[
\Gamma', \Gamma_2 \triangleright M \] & \quad \text{by definition, the denotation}\quad \text{is valid typing judgements. By induction hypothesis, the computational denotation of the latter is} \\
\[
\Gamma', \Gamma_2 \triangleright M \] & \quad \text{Using the Equation (11.2.8), this is equal to} \\
\[
\Gamma', \Gamma_2 \triangleright M \] & \quad \text{Using the same technique as in the case where} \\
\end{align*}
\]

\[
\begin{align*}
\Gamma', \Gamma_2 \triangleright M \] & \quad \text{By definition, the denotation} \\
\end{align*}
\]

Using the Equation (11.2.8), this is equal to

\[
\begin{align*}
\Gamma', \Gamma_2 \triangleright M \] & \quad \text{Using the same technique as in the case where} \\
\end{align*}
\]

Using the same technique as in the case where \( x \in [\Gamma'] \), one can move up the occurrence of \( v \), and thus get \( h \) in the requested form.

Again, the computation for \( h' = [\Delta, \Gamma_1, \Gamma_2 \triangleright M[V/x] : B]' \) is similar.

**Cases** \( M \equiv NP \) and \( M \equiv (N, P)^n \). As in the previous case, the typing judgement \( !\Delta, \Gamma_1, x : A \triangleright M : B \) comes from a typing rule (X) with hypotheses

\[
\begin{align*}
!\Delta', \Gamma_1 \triangleright N : C, & \quad !\Delta', \Gamma_2 \triangleright P : D,
\end{align*}
\]

for some types \( C \) and \( D \), where \((!\Delta', \Gamma_1, \Gamma_2) = (!\Delta, \Gamma_1, x : A)\). If we call \( f \) and \( g \) the following maps:

\[
\begin{align*}
f = [!\Delta', \Gamma_1 \triangleright N : C]' & \quad \text{and} \quad g = [!\Delta', \Gamma_2 \triangleright P : D]'\quad \text{we have} \\
\end{align*}
\]

\[
\begin{align*}
[!\Delta, \Gamma_1, x : A \triangleright M : B]' &= (f \oplus_{\Delta} g); F
\end{align*}
\]
for a given map $F$, depending if $M$ is of the form $NP$ or of the form $(N, P)^n$. If $M$ is an extended value, so are $N$ and $P$. If we call $f'$ and $g'$ the following maps:

$$f' = [\Delta', \Gamma_1 \triangleright N : C]^v, \quad g' = [\Delta', \Gamma_1 \triangleright P : D]^v,$$

we have

$$[\Delta, \Gamma_1, x : A \triangleright M : B]^v = (f' \otimes \Delta, g); G,$$

for a given map $G$, depending if $M$ is of the form $NP$ or of the form $(N, P)^n$. From Lemma one can assume that $\Gamma_1$ and $\Gamma_1$ do not contain any duplicable variables, implying that $\|[\Delta]\| \subseteq \|[\Delta']\|$. Thus there exists some context $\Psi$ such that $\delta = (\delta, \delta)$. In particular, this means that

$$\delta = (\delta, \delta). \tag{11.2.9}$$

The variable $x : A$ belongs to either $\Psi$, $\Gamma_1$ or $\Gamma_1$. The study of each one of these three possibilities is similar as in the previous case.

This ends the list of cases: the lemma is verified. \hfill \Box

**Theorem 11.2.5 (Soundness).** If $\Delta \triangleright M \equiv M' : A$ then we have $[\Delta \triangleright M : A]^\Theta = [\Delta \triangleright M' : A]^\Theta$ (and $[\Delta \triangleright M : A]^\Theta = [\Delta \triangleright M' : A]^\Theta$ if $M$ and $M'$ are values) for every interpretation $\Theta$.

**Proof.** Proof by induction on the size of the derivation of $M \equiv M'$. Case distinction on the last rule used. The cases corresponding to reflexivity, symmetry, transitivity and congruence are trivial. We prove the result for the other cases:

**Rule ($\beta_\lambda$).** Here, we are considering $\Delta \triangleright \text{let } x = V \text{ in } M \equiv M[V/x] : A$. We have directly from Lemma 11.2.4 that $[M[V/x]]^\Theta = [\text{let } x = V \text{ in } M]^\Theta$, and $[M[V/x]]^\Theta = [\text{let } x = V \text{ in } M]^\Theta$ if they are extended values.

**Rule ($\beta_\circ$).** We are considering $\Delta \triangleright \text{let } (x^2, y^2) = (V, W)^n \text{ in } M \equiv M[V/x, W/y] : A$. By $\alpha$-equivalence one can assume that $x$ is not free in $W$: we have $M[V/x, W/y] = (M[W/y])[V/x]$. One can split $\Delta$ into $[\Delta', \Gamma_1, \Gamma_2, \Gamma_3]$, such that $[\Gamma_1]$ contains only the free variables of $V$, $[\Gamma_2]$ only the free variables of $W$ and $[\Gamma_3]$ only the ones of $M$. From Lemma 11.2.4 the typing judgements $\text{!}\Delta', \Gamma_1 \triangleright V : \text{!}^n C, \quad \text{!}\Delta', \Gamma_2 \triangleright W : \text{!}^n D, \quad \text{!}\Delta', \Gamma_3, x : \text{!}^n C, y : \text{!}^n D \triangleright M : A,$

are valid. Let $f$, $g$ and $h$ be the denotations

$$f = [\Delta', \Gamma_1 \triangleright V : \text{!}^n C]^v, \quad g = [\Delta', \Gamma_2 \triangleright W : \text{!}^n D]^v, \quad h = [\Delta', \Gamma_3, x : \text{!}^n C, y : \text{!}^n D \triangleright M : B]^v.$$

From Lemma 11.2.4

$$[\Delta', \Gamma_3, x : \text{!}^n C, \Gamma_2 \triangleright M[W/y] : A]^v$$

$$= (\text{id} \otimes \Delta', [\Delta', \Gamma_2 \triangleright W : \text{!}^n D]^v); [\Delta', \Gamma_3, x : \text{!}^n C, y : \text{!}^n D \triangleright M : A]^v,$$

thus, still from Lemma 11.2.4

$$[\Delta', \Gamma_3, \Gamma_2, \Gamma_1 \triangleright (M[W/y])[V/x] : A]^v.$$
\[ = (id_{\Delta'},\Gamma_3,\tau_2 \otimes \Delta; f); ![\Delta',\Gamma_2,\tau : \eta^\circ C \triangleright M[\eta / y] : A]^c \]
\[ = (id_{\Delta'},\Gamma_3,\tau_2 \otimes \Delta; f); (id_{\Delta'},\Gamma_3 \otimes \sigma); (id_{\Delta',\Gamma_3,\eta^\circ C} \otimes \Delta; g); h \]

On another hand, by definition we have

\[ [[\Delta',\Gamma_1,\Gamma_2 \triangleright (V,W)^n : \eta^\circ (C \otimes D)]^c \]
\[ = ((\delta;\eta^\circ C) \otimes \Delta', (g;\eta^\circ D)); \Psi_1; Td_{\Delta'}^n, \]
\[ [[\Delta',\Gamma_1,\Gamma_2 \triangleright \text{let } (x^C,y^D)^n = (V,W)^n \text{ in } M : A]^c \]
\[ = \left( \left( (f;\eta^\circ C) \otimes \Delta', (g;\eta^\circ D)); \Psi_1; d_{\Delta'}^n \otimes \Delta \text{id} \right) \left( T \left( d_{\Delta'}^n \right)^{-1} \otimes \text{id} \right) ; \sigma; t; h^* \right) \]
\[ = \left( (f;\eta^\circ (C \otimes D)) \otimes \Delta \text{id} \right); \left( T \left( d_{\Delta'}^n \right)^{-1} \otimes \text{id} \right) ; \sigma; t; h^* \]
\[ = ((f;\eta^\circ (C \otimes D)) \otimes \Delta \text{id}); \left( T \left( d_{\Delta'}^n \right)^{-1} \otimes \text{id} \right) ; \sigma; t; h^* \]
\[ = ((f;\eta^\circ (C \otimes D)) \otimes \Delta \text{id}); \left( T \left( d_{\Delta'}^n \right)^{-1} \otimes \text{id} \right) ; \sigma; t; h^* \]
\[ = ((f;\eta^\circ (C \otimes D)) \otimes \Delta \text{id}); \left( T \left( d_{\Delta'}^n \right)^{-1} \otimes \text{id} \right) ; \sigma; t; h^* \]

which makes the same map (modulo permutation of inputs) as in the previous calculation.

The case of the value denotation is similar.

**Rule \((\beta_1)\).** We have \(\Delta \triangleright \text{let } * = * \text{ in } M \equiv_{ax} M : A\).

The typing judgement \(\Delta \triangleright \text{let } * = * \text{ in } M : A\) comes from rule \((\top.E)\). The typing judgements \(\triangleright * : \top \text{ and } \Delta \triangleright M : A\) are valid, and since \([\triangleright * : \top]^c = id_{\top}\) if we call

\[ g = [\Delta \triangleright M : A]^v, \quad h = [\Delta \triangleright \text{let } * = * \text{ in } M : A]^v, \]

then \(h = (\eta_{\top} \otimes g) ; \Psi_1 = g\). Similarly, if \(M\) is an extended value, and if we call

\[ g' = [\Delta \triangleright M : A]^v, \quad h' = [\Delta \triangleright \text{let } * = * \text{ in } M : A]^v, \]

then \(h' = (id_{\top} \otimes g') = g'\).

**Rule \((\eta_\lambda)\).** We have \(\Delta \triangleright \lambda^x.x^A, \{V < : A \rightarrow B\} x^A \approx_{ax} V : \eta^\circ (A \rightarrow B)\). Both \(V\) and \(\lambda^x.x^A, \{V < : A \rightarrow B\} x^A\) are values.

Let \(f = [\Delta \triangleright V : \eta^\circ (A \rightarrow B)]^v\).

If \(n \neq 0\). Then from Lemma \([11.13.11]\) \(\Delta = ![\Delta']\).

Let \(g = [\Delta, x : A \triangleright \{V < : A \rightarrow B\} x^A : B]^c\). From Lemma \([11.13.22]\) we have

\[ g = ((f ; I_{\eta^\circ (A \rightarrow B), \eta A \rightarrow B}; \eta A) \otimes \eta A); \Psi_1 ; \varepsilon_{A,B} \]
\[ = ((f ; I_{\eta^\circ (A \rightarrow B), \eta A \rightarrow B}); \eta A \otimes \eta A); \Psi_1 \varepsilon_{A,B} \]
\[ = ((f ; I_{\eta^\circ (A \rightarrow B), \eta A \rightarrow B}); \eta A \otimes \eta A); \varepsilon_{A,B}. \]

By naturality of \(\Phi\),

\[ \Phi^{-1} \left( ((f ; I_{\eta^\circ (A \rightarrow B), \eta A \rightarrow B}); \varepsilon_{A,B}) = f ; I_{\eta^\circ (A \rightarrow B), \eta A \rightarrow B}; \Phi^{-1}(\varepsilon_{A,B}) \right) \]
11.2. Soundness of the Denotation

Thus, if $h = [\Delta \triangleright \lambda^nx^A, \{V: A \rightarrow B\}x^A : !^n(A \rightarrow B)]^v$, 
\[
h = \delta_\Delta; d_\Delta^L; L(\Phi^{-1}(g)); I_{(A \rightarrow B), !^n(A \rightarrow B)} = \delta_\Delta; d_\Delta^L; L(f); I_{(A \rightarrow B), !^n(A \rightarrow B)} = \delta_\Delta; d_\Delta^L; L(f); I_{(A \rightarrow B), !^n(A \rightarrow B)} = \delta_\Delta; d_\Delta^L; L(f); \epsilon_{!^n(A \rightarrow B)} = \delta_\Delta; d_\Delta^L; \varepsilon_\Delta; f = f.
\]

If $n = 0$. The computation is similar, although simpler since there is no reference to $!$ nor to $L$. Indeed, $\{V: A \rightarrow B\} = V$ from Lemma 11.1.15, and if $g = [\Delta, x: A \triangleright Vx^A : B]^v$, 
\[
g = (f; \eta_{A \rightarrow B}) \otimes (\eta_A) ; \Psi_1; \varepsilon_{A,B}^v = (f \otimes \text{id}_A) ; (\eta_{A \rightarrow B}) \otimes (\eta_A) ; \Psi_1; \varepsilon_{A,B}^v
\]
By naturality of $\Phi$, if $h = [\Delta \triangleright \lambda^nx^A, \{V: A \rightarrow B\}x^A : A \rightarrow B]^v$, 
\[
h = L(\Phi^{-1}(g)) = f
\]
Finally, in the case of the computational denotation, since both terms are core values, we have the result by definition.

Rule $\beta^2$. We have $\Delta \triangleright x^A = N$ in $x^A \approx_{ax} N : A$. By $\alpha$-equivalence one can assume that $x$ is not a free variable of $N$.
Suppose that $N$ is an extended value, and let $f$ be $[\Delta \triangleright N : A]^v$. Then if $h$ is the denotation $[\Delta \triangleright let x^A = N in x^A : A]^v$, we have 
\[
h = (f \otimes \text{id}_\top) ; \sigma_{A,\top} ; \text{id}_A = f.
\]
Now, suppose that $N$ is any term, and let $f'$ be $[\Delta \triangleright N : A]^v$. From Lemma 11.1.15 if $h' = [\Delta \triangleright let x^A = N in x^A : A]^v$, 
\[
h' = (f' \otimes \text{id}_\top) ; \sigma_{T[\top],\top} ; \text{id}_A = (f' \otimes \text{id}_\top) ; \sigma_{A,\top} ; \text{id}_A = f'.
\]

Rule $\eta_\otimes$. We have $\Delta \triangleright let \langle x^A, y^B \rangle^n = N$ in $\langle x^A, y^B \rangle^n \approx_{ax} N : !^n(A \otimes B)$. By $\alpha$-equivalence one can assume that neither $x$ nor $y$ are free variables of $N$.
Suppose that $N$ is an extended value, and let $f$ be $[\Delta \triangleright N : !^n(A \otimes B)]^v$. Then if $h$ is the denotation $[\Delta \triangleright let \langle x^A, y^B \rangle^n = N in \langle x^A, y^B \rangle^n : !^n(A \otimes B)]^v$, using the fact that the category is strongly monoidal and monoidal strict, 
\[
h = (f \otimes \text{id}_\top) ; (d^L_{[A],[B]}^{n})^{-1} \otimes \text{id}_\top) ; \sigma_{n(A \otimes B),\top} ; d^L_{[A],[B]}^{n} = f ; (d^L_{[A],[B]}^{n})^{-1} ; d^L_{[A],[B]}^{n} = f.
\]
11.2. Soundness of the Denotation

Now, suppose that $N$ is any term, and let $f'$ be $[\Delta \triangleright N : !^n(A \otimes B)]^\circ$. If $h'$ is the denotation $[\Delta \triangleright \langle x^A, y^B \rangle^n = N \text{ in } \langle x^A, y^B \rangle^n : !^n(A \otimes B)]^\circ$, we have

$$h' = (f' \otimes \text{id}_\top); \left( T\left( d_{[A]_B}^{f_n}\right) \right)^{-1} \otimes \text{id}_\top); \sigma_T[\pi_{\otimes(A \otimes B), \top}; T; \left( d_{[A]_B}^{f_n}; \eta \right)]^*$$

$$= f'; T\left( d_{[A]_B}^{f_n}\right)^{-1}; T d_{[A]_B}^{f_n}$$

$$= f'.$$

Rule ($\eta_\Box$). We have $\Delta \triangleright \text{let } * = \{ N <: \top \} \text{ in } *^n \approx_{\Box} N : !^n \top$.

Suppose that $N$ is an extended value, and let $f$ be $[[\Delta \triangleright N : !^n \top]]^\circ$. Then if $h$ is the denotation $[\Delta \triangleright \text{let } * = \{ N <: \top \} \text{ in } *^n : !^n \top]^\circ$, using Lemma [11.2.2] and Lemma 8.3.5,

$$h = ((f; I_{n\top, \top}) \otimes d_L^n) = f; I_{n\top, \top}; d_L^n = f.$$
There exists a map $F : \llbracket D \rrbracket \to \llbracket A_1 \rrbracket$ and a map $G : \llbracket C \rrbracket \to \llbracket A_2 \rrbracket$, respectively based upon the value of $\Box$ and $\Box$, such that

$$
g_1 = (f_M \otimes \iota_{\Delta'} \otimes id_{\Delta' \Gamma_2}); (TF \otimes id_{\Delta' \Gamma_2}); \sigma_{T\llbracket A_1 \rrbracket \llbracket A_2 \rrbracket }; t; f^*_N,
$$

$$
g_2 = (f_N \otimes \iota_{\Delta'} \otimes id_{\Delta' \Gamma_2}); (TG \otimes id_{\Delta' \Gamma_2}); \sigma_{T\llbracket A_1 \rrbracket \llbracket A_2 \rrbracket }; t; f^*_P,
$$

$$
h_1 = (g_1 \otimes \iota_{\Delta'} \otimes id_{\Delta' \Gamma_2}); (TG \otimes id_{\Delta' \Gamma_2}); \sigma_{T\llbracket A_1 \rrbracket \llbracket A_2 \rrbracket }; t; f^*_P,
$$

$$
h_2 = (f_M \otimes \iota_{\Delta'} \otimes id_{\Delta' \Gamma_2}); (TF \otimes id_{\Delta' \Gamma_2}); \sigma_{T\llbracket A_1 \rrbracket \llbracket A_2 \rrbracket }; t; T(id \otimes \sigma_{[\Gamma_1],[\Lambda_1]}); g^*_2.
$$

Using Lemma 11.1.3 one can rewrite $h_1$ as

$$
h_1 = \left( ((f_M; TF) \otimes \iota_{\Delta'} \otimes id_{\Delta' \Gamma_2}) \otimes \iota_{\Delta'} \otimes id_{\Delta' \Gamma_2} \right); ((\sigma; t; f^*_N; TG) \otimes id_{\Delta' \Gamma_2}); \sigma; t; f^*_N.
$$

Using Lemma 11.1.4 one can show that

$$
T[\llbracket A_1 \rrbracket \otimes \llbracket \Delta' \Gamma_2 \Gamma_3 \Lambda_1 \rrbracket] \xrightarrow{\sigma \otimes t} T[\llbracket \Delta' \Gamma_2 \Gamma_3 \Lambda_1 \rrbracket] \xrightarrow{T(id \otimes \sigma)} T[\llbracket \Delta' \Gamma_2 \Gamma_3 \Lambda_1 \rrbracket] \xrightarrow{T((f_N; TG) \otimes \iota_{\Delta'} \otimes id)} T(T[\llbracket A_2 \rrbracket \otimes \llbracket \Delta' \Gamma_3 \rrbracket])
$$

is equal to

$$
T[\llbracket A_1 \rrbracket \otimes \llbracket \Delta' \Gamma_2 \Gamma_3 \Lambda_1 \rrbracket] \xrightarrow{T(id \otimes \Delta') \otimes \iota_{\Delta'} \otimes id} T[\llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket] \xrightarrow{T(f_N; TG)} T([\llbracket \Delta' \Gamma_3 \rrbracket \otimes T[\llbracket A_2 \rrbracket]) \xrightarrow{T(\sigma)} T([\llbracket A_2 \rrbracket \otimes \llbracket \Delta' \Gamma_3 \rrbracket]).
$$

Now, using Equations 5.3.4 and 5.3.3, the following diagram is commutative:

$$
\begin{array}{c}
\llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \xrightarrow{id \otimes t} \llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \\
\llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \xrightarrow{T(\Delta' \iota \Lambda_1 \Gamma_2 \Gamma_3 \Lambda_1)} \llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \\
\llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \xrightarrow{T(\Delta' \iota \Lambda_1 \Gamma_2 \Gamma_3 \Lambda_1)} \llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \\
\llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \xrightarrow{T(\Delta' \iota \Lambda_1 \Gamma_2 \Gamma_3 \Lambda_1)} \llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \\
\llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \xrightarrow{T(\Delta' \iota \Lambda_1 \Gamma_2 \Gamma_3 \Lambda_1)} \llbracket \Delta' \Gamma_3 \Gamma_2 \Lambda_1 \rrbracket \\
\end{array}
$$

This is enough to say that $h_1 = h_2$.

The case where $M$, $N$ and $P$ are extended values is similar, without the need for the equations of strong monadicity.

Rule (let2). We have

$$
\Delta \triangleright \text{let } \Box = V \text{ in let } \Box = W \text{ in } M \approx_{ax} \text{ let } \Box = W \text{ in let } \Box = V \in M : A,
$$
where $V$ and $W$ are core values. Using Lemma 11.1.4, there exist contexts $!\Delta'$, $\Gamma_1$, $\Gamma_2$, $\Gamma_3$, $A_1$ and $A_2$ such that $A_1$ contains the variables in $\blacksquare$, $A_2$ the variables in $\blacksquare$ such that $\Delta = (!\Delta', \Gamma_1, \Gamma_2, \Gamma_3)$ and such that

$$
!\Delta', \Gamma_3, A_1, A_2 \triangleright M : A, \quad !\Delta', \Gamma_1 \triangleright V : C, \quad !\Delta', \Gamma_2 \triangleright W : D.
$$

Let $h_1$ and $h_2$ be

$$
h_1 = [|!\Delta', \Gamma_1, \Gamma_2, \Gamma_3 \triangleright \text{let } \blacksquare = V \text{ in let } \blacksquare = W \text{ in } M : A]^c, \quad h_2 = [|!\Delta', \Gamma_1, \Gamma_2, \Gamma_3 \triangleright \text{let } \blacksquare = V \text{ in let } \blacksquare = W \text{ in } M : A]^c.
$$

Let us define the following maps:

$$
\begin{align*}
f_M &= [|!\Delta', \Gamma_3, A_1, A_2 \triangleright M : A]^c, \\
f_V &= [|!\Delta', \Gamma_1 \triangleright V : C]^v, \\
f_W &= [|!\Delta', \Gamma_2 \triangleright W : D]^v, \\
g_1 &= [|!\Delta', \Gamma_1, \Gamma_3, A_2 \triangleright \text{let } \blacksquare = V \text{ in let } \blacksquare = W \text{ in } M : A]^c, \\
g_2 &= [|!\Delta', \Gamma_1, \Gamma_2, \Gamma_3 \triangleright \text{let } \blacksquare = V \text{ in let } \blacksquare = W \text{ in } M : A]^c.
\end{align*}
$$

There exists a map $F : [C] \rightarrow [A_1]$ and a map $G : [D] \rightarrow [A_2]$, respectively based upon the value of $\blacksquare$ and $\blacksquare$, such that

$$
\begin{align*}
g_1 &= ((f_V ; \eta_C ; TF) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{T[A_3] ; [!\Delta', \Gamma_3, A_2]} ; \top ; (id \otimes \sigma_{[A_2] ; [A_3]}) ; f_M, \\
g_2 &= ((f_W ; \eta_D ; TG) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_1}) ; \sigma_{T[A_3] ; [!\Delta', \Gamma_3, A_1]} ; \top ; f_M, \\
h_1 &= ((f_V ; \eta_C ; TF) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{T[A_3] ; [!\Delta', \Gamma_3, A_2]} ; \top ; g_1^v, \\
h_2 &= id \otimes \sigma_{[A_1] ; [A_2]} \otimes id ; ((f_W ; \eta_D ; TG) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_1}) ; \sigma_{T[A_3] ; [!\Delta', \Gamma_3, A_1]} ; \top ; g_1^v.
\end{align*}
$$

We now want to show that $h_1 = h_2$. Using Equation 11.1.2.1 stating that $\eta_{A \otimes B} = (id_A \otimes \eta_B) ; t_{A,B}$, we rewrite $h_1$:

$$
\begin{align*}
h_1 &= ((f_V ; \eta_C ; TF) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{T[A_3] ; [!\Delta', \Gamma_3, A_2]} ; \top ; g_2^v \\
&= ((f_V ; f) ; \eta_{[A_3]} \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{T[A_3] ; [!\Delta', \Gamma_3, A_2]} ; \top ; g_2^v \\
&= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; (id_{!\Delta', \Gamma_3, A_2} \otimes \eta_{[A_3]}) \top ; g_2^v \\
&= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; \eta ; g_2^v \\
&= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; \top ; g_2^v.
\end{align*}
$$

Similarly one can rewrite $g_1$, $g_2$ and $h_2$: $g_1$, $g_2$ and $h_2$:

$$
\begin{align*}
g_1 &= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; (id \otimes \sigma_{[A_2] ; [A_3]}) ; f_M, \\
g_2 &= ((f_W ; G) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_1}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; f_M, \\
h_2 &= id \otimes \sigma_{[A_1] ; [A_2]} \otimes id ; ((f_W ; G) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_1}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; g_1.
\end{align*}
$$

The term $h_1$ becomes:

$$
\begin{align*}
h_1 &= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; g_2 \\
&= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; ((f_W ; G) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_1}) ; \\
&\quad \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; f_M \\
&= ((f_V ; f) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_2}) ; (id_{A_3} \otimes ((f_W ; G) \otimes !\Delta' \otimes id_{!\Delta', \Gamma_3, A_1}) ; \\
&\quad \sigma_{[A_3] ; [!\Delta', \Gamma_3, A_2]} ; f_M
\end{align*}
$$
Rules (let\textsuperscript{app} and let\textsuperscript{⊗}). We have

\[
\begin{align*}
\Delta \triangleright let \ x^A \rightarrow B &= M \text{ in let } y^A = N \text{ in } x^A \rightarrow B y^A \equiv_{ax} MN : B, \\
\Delta \triangleright let \ x^n A &= M \text{ in let } y^n B = N \in \langle x^n A, y^n B \rangle^n \equiv_{ax} (M, N)^n : !n (A \otimes B).
\end{align*}
\]

Note that both of these equations are of the form

\[
\Delta \triangleright let \ x^C = M \text{ in let } y^D = N \in P \equiv_{ax} P(M, N) : E,
\]

for some types \(C, D\) and \(E\) and some term construct \(P(-, -)\).

We can compute the computational interpretation in both cases. The value interpretation can be done only in the second case. However, the proof being a subset of the one required for the computational interpretation, we omit it.

Using \(\alpha\)-equivalence, one can assume that \(x\) and \(y\) does not occur in \(M\) nor in \(N\). Using Lemma \ref{lemma1}, there exist contexts !\(\Delta', \Gamma_1\) and \(\Gamma_2\) such that \(\Delta = (!\Delta', \Gamma_1, \Gamma_2)\) and such that

\[
!\Delta', \Gamma_1 \triangleright M : C, \quad !\Delta', \Gamma_2 \triangleright N : D.
\]

Let \(h_1\) and \(h_2\) be

\[
\begin{align*}
h_1 &= [!\Delta', \Gamma_1, \Gamma_2 \triangleright let \ x^C = M \text{ in let } y^D = N \in P(x, y) : B]^c, \\
h_2 &= [!\Delta', \Gamma_1, \Gamma_2 \triangleright P(M, N) : B]^c.
\end{align*}
\]

Let us define the following maps:

\[
\begin{align*}
f_M &= [!]\Delta', \Gamma_1 \triangleright M : C]^c, \\
f_N &= [!]\Delta', \Gamma_2 \triangleright N : D]^c, \\
g &= [!]\Delta', \Gamma_2, x : C \triangleright let \ y^D = N \in P(x, y) : B]^c, \\
k &= [x : C, y : D \triangleright P(x, y) : B]^c.
\end{align*}
\]

By definition and from Lemma \ref{lemma1}, there exists a map \(F : [C \otimes D] \rightarrow T(B)\) such that

\[
\begin{align*}
k &= (\eta_C \otimes \eta_D); \Psi_1; F^*, \\
g &= (f_N \otimes id_C); \sigma_{T[D], [C]}; t; k^*, \\
h_1 &= (f_M \otimes \Delta^*; id_{!\Delta', \Gamma_2}); \sigma_{T[C], [!\Delta', \Gamma_2]}; t; g^*, \\
h_2 &= (f_M \otimes \Delta^*; f_N); \Psi_1; F^*.
\end{align*}
\]

Using Equation \ref{tensorial-strength} of the tensorial strength, let us rewrite \(k\):

\[
\begin{align*}
k &= (\eta_C \otimes \eta_D); \Psi_1; F^* \\
&= (\eta_C \otimes \eta_D); \sigma_{T[C], T[D]}; T_{T[D], [C]}; T_{T[D], [C]}; T_{T[D], [C]}; T_{T[D], [C]}; \mu; F^* \\
&= (\eta_D \otimes \eta_C); t_{T[D], [C]}; T_{T[D], [C]}; T_{T[D], [C]}; T_{T[D], [C]}; \mu; F^* \\
&= (\eta_D \otimes id_C); \eta_{T[D] \otimes [C]}; T_{T[D], [C]}; T_{T[D], [C]}; T_{T[D], [C]}; \mu; F^*
\end{align*}
\]
11.2. Soundness of the Denotation

Let us rewrite $g$:

$$g = (f_N \otimes id_C); \sigma T[D]; [C]; t; F^*$$

Let us rewrite $h_1$:

$$h_1 = (f_M \otimes !\Delta') \otimes \sigma T[C]; \sigma T[D]; [C]; [D]; t; g^*$$

Rule (let$^\lambda$). We have

$$!\Delta \triangleright let \; x^D = V \text{ in } \lambda^n y^A. M \approx_{\alpha} \lambda^n y^A. \text{let } \; x^D = V \text{ in } M : !^n(A \rightarrow B),$$

where $V$ is a core value. Using $\alpha$-equivalence, one can assume that $x$ does not occur in $M$ nor in $V$, and that $y \neq x$. Using Lemma 9.1.10, there exist contexts $!\Delta'$, $\Gamma_1$ and $\Gamma_2$ such that

$$!\Delta', \Gamma_1 \triangleright V : D,$$

Let $h_1$ and $h_2$ be

$$h_1 = [[!\Delta', \Gamma_1, \Gamma_2 \triangleright let \; x^D = V \text{ in } \lambda^n y^A. M : !^n(A \rightarrow B)]^c,$$

$$h_2 = [[!\Delta', \Gamma_1, \Gamma_2 \triangleright !\Delta' \triangleright V : D]$$

Let us define the following maps:

$$f_M = [[!\Delta', \Gamma_1, \Gamma_2 \triangleright x : D, y : A \triangleright M : B]^c,$$

$$f_V = [[!\Delta', \Gamma_1 \triangleright V : D]^c,$$

$$g_1 = [[!\Delta', \Gamma_1, \Gamma_2 \triangleright !\Delta' \triangleright V : D]$$

There are four cases: $n$ can be null or not, and $M$ can be an extended value or not (allowing us to compute the value interpretation). We fully develop the computational interpretation of the case $n = 0$ since the other three cases are done using similar calculations.

From Lemma 9.1.10, $\Gamma_1$ and $\Gamma_2$ are respectively of the form $!\Gamma'_1$ and $!\Gamma'_2$, and $D = !D'$. By definition, we have

$$g_1 = \delta_{!\Delta'; !\Gamma'_2; x : !D'} \otimes d_{!\Delta'; !\Gamma'_2; x : !D'} \otimes L\Phi^{-1}(f_M); \eta_{(A \rightarrow B); !^n(A \rightarrow B); \eta_{(A \rightarrow B)}$$

$$g_2 = ((f_V; \eta_D) \otimes !\Delta' \otimes id_{!\Delta'; !\Gamma'; !\Gamma'; A}); \sigma T[D]; [D]; [D]; [C]; [C]; [D]; t; T; (id \otimes \sigma A; D); f'_M,$$
11.2. Soundness of the Denotation

We are considering the relation $\Delta':\Gamma_1, \Gamma_2: N : B \rightarrow C$ and $\Delta':\Gamma_2: P : A'$.

We want to show that $h_1 = h_2$. Let us rewrite $g_2$, using Equation 5.4.4 of strong monadicity of $T$:

$$g_2 = (f_v; \eta_D) \odot \Delta' \cdot id_{\Delta', \Gamma_2, y: A}; \sigma_{T[D], [\Delta', \Gamma_2, y: A]}; t; f_M$$

By naturality of $\Phi$, $\Phi^{-1}(g_2) = (f_v \odot \Delta' \cdot id_{\Delta', \Gamma_2}; \sigma_{[D], [\Delta', \Gamma_2]}; \Phi^{-1}(f_M))$.

The map $h_2$ is reformatted using the naturality of $d_L$ and $\delta$:

$$h_2 = \delta_{\Delta', \Gamma_1', \Gamma_2'; y, y': D} \cdot d_L^{\Delta', \Gamma_2', y: y'; \Gamma_2, \Gamma'_1}; L\Phi^{-1}(g_2); I_{(A \rightarrow B), \eta^n(A \rightarrow B)}; \eta^n(A \rightarrow B)$$

$$= \delta_{\Delta', \Gamma_1', \Gamma_2'; y, y': D} \cdot d_L^{\Delta', \Gamma_2', y: y'; \Gamma_2, \Gamma'_1}; L(f_v \odot \Delta' \cdot id_{\Delta', \Gamma_2}; L\sigma_{[D], [\Delta', \Gamma_2]}; L\Phi^{-1}(f_M)); I_{(A \rightarrow B), \eta^n(A \rightarrow B)}; \eta^n(A \rightarrow B)$$

$$= (f_v \odot \Delta' \cdot id_{\Delta', \Gamma_2}; \sigma_{[D], [\Delta', \Gamma_2]}; \delta_{\Delta', \Gamma_2, y: y'; D}; d_L^{\Delta', \Gamma_2', y: y'; \Gamma_2, \Gamma'_1}; L\Phi^{-1}(f_M)); I_{(A \rightarrow B), \eta^n(A \rightarrow B)}; \eta^n(A \rightarrow B)$$

Note that $g_1$ is of the form $g_1^*: \eta^n(A \rightarrow B)$. Again using Equation 5.4.4:

$$h_1 = ((f_v; \eta_D) \odot \Delta' \cdot id_{\Delta', \Gamma_2}; \sigma_{T[D], [\Delta', \Gamma_2]}; t; g_1^*)$$

$$= (f_v \odot \Delta' \cdot id_{\Delta', \Gamma_2}; \sigma_{[D], [\Delta', \Gamma_2]}; \eta^n(A \rightarrow B); t; g_1^*)$$

The setting is the following: $A \prec B$, $\Delta' = \langle \Delta', \Gamma_1, \Gamma_2 \rangle$ and $\Delta: \Gamma_1, \Gamma_2: N : B \rightarrow C$.

We are considering the relation $\Delta: N \{P \prec B\} \equiv_{ax} \{N \prec A \rightarrow C\} P : C$.

In this situation, only the computational interpretation makes sense. If $f = [\Delta': \Gamma_1 \triangleright N : B \rightarrow C]^c$, $g = [\Delta': \Gamma_2 \triangleright P : A]^c$,
then, since $I_{B \to C,A \to C} = (I_{A,B}) \to \circ C$, we have

$$[\Delta \triangleright N \{P <: B \} : B]^{c} = (f \otimes \Delta' \triangleright (g; TI_{A,B}); \Psi_1; \varepsilon_{B,C}^{*})$$

$$= (f \otimes \Delta' g; (id \otimes TI_{A,B}); \Psi_1; \varepsilon_{B,C}^{*})$$

$$= (f \otimes \Delta' g; \Psi_1; ((id \otimes I_{A,B}); \varepsilon_{B,C})^{*})$$

and

$$[\Delta \triangleright \{N <: A \to C \} P : C]^{c} = ((f; TI_{B \to C,A \to C} \otimes \Delta' g); \Psi_1; \varepsilon_{A,C}^{*})$$

$$= (f \otimes \Delta' g; (TI_{B \to C,A \to C} \otimes id); \Psi_1; \varepsilon_{A,C}^{*})$$

$$= (f \otimes \Delta' g; \Psi_1; ((I_{B \to C,A \to C} \otimes id); \varepsilon_{A,C})^{*})$$

The two denotations are equal by naturality of $\Phi$. Indeed, since

$$\mathcal{C}(A \to C, A \to C) \xrightarrow{\phi} \mathcal{C}((A \to C) \otimes A, TC)$$

$$(I_{A,B} \to C); - \downarrow$$

we have $\Phi(I_{A,B} \to C) = ((I_{A,B} \to C) \otimes A); \varepsilon_{A,C}$. Since

$$\mathcal{C}(B \to C, B \to C) \xrightarrow{\phi} \mathcal{C}((B \to C) \otimes B, TC)$$

$$(B \to C, A \to C) \xrightarrow{\phi} \mathcal{C}((B \to C) \otimes A, TC),$$

we have $\Phi(I_{A,B} \to C) = ((B \to C) \otimes I_{A,B}); \varepsilon_{B,C}$. Thus,

$$((B \to C) \otimes I_{A,B}); \varepsilon_{B,C} = ((I_{A,B} \to C) \otimes A); \varepsilon_{A,C}.$$

Rule (let\(^*\)). In this case, $A <: A', B <: B'$, and $n' = 0$ if $n = 0$. We have $\Delta = (!\Delta', \Gamma_1, \Gamma_2)$ with

$$!\Delta', \Gamma_1 \triangleright N : !n'(A \otimes B), \quad !\Delta', \Gamma_2, x : !n' A', y : !n' B' \triangleright P : E,$$

and we consider the following relation:

$$\Delta \triangleright \left\{ \right. \left( let \; (x^{A'}, y^{B'})^{n'} = \{N <: !n'(A' \otimes B')\} \; in \; P \left. \right) \cong_{ax} \right\} \left( let \; (x^{A}, y^{B})^{n} = N \; in \; P \right) : E.$$

By $\alpha$-equivalence, one can assume that $x, y$ do not occur in $N$.

Let $f, g, h_1$ and $h_2$ be the following maps:

$$f = [[!\Delta', \Gamma_1 \triangleright N : !n'(A \otimes B)]^{c}],$$

$$g = [[!\Delta', \Gamma_2, x : !n' A', y : !n' B' \triangleright P : E]^{c}],$$

$$h_1 = [[!\Delta', \Gamma_1, \Gamma_2 \triangleright let \; (x^{A'}, y^{B'})^{n'} = \{N <: !n'(A' \otimes B')\} \; in \; P : E]^{c}],$$

$$h_2 = [[!\Delta', \Gamma_1, \Gamma_2 \triangleright let \; (x^{A}, y^{B})^{n} = N \; in \; P : E]^{c}].$$
By definition, we have
\[ h_1 = ((f; I_{n(\otimes B)}, l^\alpha(A', B')) \otimes i; id_{\Delta}; v); (T(d_{\Delta}^{\alpha})^{-1} \otimes id); \]
\[ \sigma_{T[x_{A', A''}]; t}; g^*, \]
\[ h_2 = (f \otimes id; id_{\Delta}; v); (T(d_{\Delta}^{\alpha})^{-1} \otimes id); \]
\[ \sigma_{T[x_{A', A''}]; t}; (I(i_{\Delta'}, \Delta' x; l^\alpha x_{A', A''} v); g^*). \]

These two maps are the same since
\[ I_{n(\otimes B), l^\alpha(A', B')} = L^n I_{A, A'} \otimes L^\alpha I_{B, B'}, \]
and
\[ I(i_{\Delta', \Delta' x; l^\alpha x_{A', A''} v}; g^*). \]

For the same reasons the value interpretation of the two terms in relation are the same.

This ends the list of possible cases. Theorem 11.2.5 is then verified.

Corollary 11.2.6. If \( \text{Erase}(M) = \text{Erase}(M') \) and if \( \Delta \triangleright \ M, M' : A \) are valid typing judgements, then \( [M]^v = [M']^v \) (and \( [M]^c = [M']^c \) if \( M \) and \( M' \) are values).

Proof. Corollary of Theorems 11.2.4 and 11.2.5.

11.3 Completeness

Recall the category \( C_L \) from Section 11.2. Since the category \( C_L \) is a linear category for duplication, one can interpret the language in it. This section states that the defined lambda calculus is an internal language of linear categories for duplication.

Since the category \( C_L \) is a monoidal category, one can w.l.o.g. generalize the notion of pairing to finite tensor products of terms.

Definition 11.3.1. We define the following shortcut notations for terms:
\[ \langle \rangle^n = \ast^n, \quad \langle x^A \rangle^n = x^{nA}, \quad \langle x_1, x_2 \ldots \rangle^n = \langle x_1, \langle x_2 \ldots \rangle^0 \rangle^n, \]
\[ (\text{let } \langle \rangle^n = M \text{ in } N) = (\text{let } \ast = \{M < : \top\} \text{ in } vN), \]
\[ (\text{let } \langle x^A \rangle^n = M \text{ in } N) = (\text{let } x^{nA} = M \text{ in } N), \]
\[ (\text{let } \langle x_1, x_2 \ldots \rangle^n = M \text{ in } N) = (\text{let } \langle x_1, z_2 \rangle^n = M \text{ in let } \langle x_2 \ldots \rangle^n = z_2 \text{ in } N), \]
and for types:
\[ A_1 \otimes A_2 \otimes \cdots \otimes A_n = A_1 \otimes (A_2 \otimes (\cdots \otimes A_n) \cdots). \]

Lemma 11.3.2. The following derived typing rules are correct:
\[ \frac{\Delta, \Gamma_1 \triangleright M : l^n(A_1 \otimes \cdots \otimes A_k)}{\Delta, \Gamma_1, \Gamma_2 \triangleright \langle x_1, x_2 \ldots \rangle^n = M \text{ in } N : B} \quad (11.3.1) \]
and

\[ !\Delta, \Gamma_1 \triangleright M_1 : !^n A_1 \]
\[ !\Delta, \Gamma_2 \triangleright M_2 : !^n A_2 \]
\[ \vdots \]
\[ !\Delta, \Gamma_k \triangleright M_k : !^n A_k \]
\[ !\Delta, \Gamma_1, \ldots, \Gamma_k \triangleright (M_1, \ldots, M_k) : !^n (A_1 \otimes \cdots \otimes A_k). \]  

(11.3.2)

Proof. Proof by induction on \( k \).

Lemma 11.3.3. In the setting of Equation (11.3.2) in Lemma 11.3.2, if for all \( i \) we have \( f_i = [![\Delta, \Gamma_i \triangleright M_i : !^n A_i]^c_\Delta]_\Psi \), then

\[ [\Gamma_1, \ldots, \Gamma_k \triangleright (M_1, \ldots, M_k)^n : !^n (A_1 \otimes \cdots \otimes A_k)]^c = (f_1 \odot \Delta \cdots \odot \Delta f_k); \Psi; Td_{A_1, \ldots, A_k}^n. \]

In the case where all the \( M_i \) are extended values, if \( f_i' \) is \( f_i = [![\Delta, \Gamma_i \triangleright M_i : !^n A_i]^c_\Delta]_\Psi \), then

\[ [\Gamma_1, \ldots, \Gamma_k \triangleright (M_1, \ldots, M_k)^n : !^n (A_1 \otimes \cdots \otimes A_k)]^c = (f_1 \odot \Delta \cdots \odot \Delta f_k); d_{A_1, \ldots, A_k}^T. \]

Proof. Proof by induction on \( k \).

Lemma 11.3.4. In the setting of Equation (11.3.2) in Lemma 11.3.2, if

\[ f = [![\Delta, \Gamma_1 \triangleright M : !^n (A_1 \otimes \cdots \otimes A_k)]^c_\Delta]_\Psi, \]
\[ g = [![\Delta, \Gamma_2, x_1 : !^n A_1, \ldots, x_k : !^n A_k \triangleright N : B]^c_\Delta]_\Psi, \]

then

\[ [![\Delta, \Gamma_1, \Gamma_2 \triangleright \text{let} \langle x_1, \ldots, x_k \rangle^n = M \text{ in } N : B]^c_\Delta]_\Psi = (f \odot \Delta \text{id}); (T(d_{A_1, \ldots, A_k}^T)^{-1} \otimes \text{id}); \sigma; t; g^*. \]

In the case where \( M \) and \( N \) are extended values, if

\[ f = [![\Delta, \Gamma_1 \triangleright M : !^n (A_1 \otimes \cdots \otimes A_k)]^c_\Delta]_\Psi, \]
\[ g = [![\Delta, \Gamma_2, x_1 : !^n A_1, \ldots, x_k : !^n A_k \triangleright N : B]^c_\Delta]_\Psi, \]

then

\[ [![\Delta, \Gamma_1, \Gamma_2 \triangleright \text{let} \langle x_1, \ldots, x_k \rangle^n = M \text{ in } N : B]^c_\Delta]_\Psi = (f \odot \Delta \text{id}); (T(d_{A_1, \ldots, A_k}^T)^{-1} \otimes \text{id}); \sigma; g. \]

Proof. Proof by induction on \( n \).

Lemma 11.3.5. In \( \mathcal{C}_\Lambda \), a valid typing judgement \( x_1 : A_1, \ldots, x_n : A_n \triangleright M : B \) has for computational denotation \( t : A_1 \otimes \cdots \otimes A_n \triangleright \text{let} \langle x_1, \ldots, x_n \rangle = t \exists \lambda d.M : \top \rightarrow B \). If \( M = \text{value} \) is a value, the value interpretation is \( t : A_1 \otimes \cdots \otimes A_n \triangleright \text{let} \langle x_1, \ldots, x_n \rangle = t \exists \sigma t \in V : B \).

Proof. Proof by structural induction on \( M \), using Lemmas 11.3.3 and 11.3.4. For each case, one writes the categorical morphism corresponding to the term and one interprets it in \( \mathcal{C}_\Lambda \).

Theorem 11.3.6 (Completeness). In \( \mathcal{C}_\Lambda \), \( \emptyset \) being the identity, one has \([x : A \triangleright M : B]_\emptyset \approx_{ax} (x : A \triangleright \lambda d.M : \top \rightarrow B)\) and \([x : A \triangleright V : B]_\emptyset \approx_{ax} (x : A \triangleright V : B)\).

Proof. Corollary of Lemma 11.3.5.
11.4 Towards a Denotational Model

In Chapters 8-11, we developed the categorical requirements for modeling a generic call-by-value linear lambda calculus, i.e., its type system (which includes subtyping) and equational laws. We have not yet specialized the language to a particular set of built-in operators, for example, those that are required for quantum computation.

However, since the quantum lambda calculus of Chapter 6 is the main motivation behind our work, we will comment very briefly on what additional properties would be required to interpret its primitives. The quantum lambda calculus of Chapter 6 is obtained by instantiating and extending the call-by-value language of Chapter 9 with the following primitive types, constants, and operations:

- **Types:** bit, qbit
- **Constants:** 0 : !bit, 1 : !bit, new : !(bit → qbit), U : !(qbit^n → qbit^n), meas : !(qbit → !bit)
- **Operations:** Γ_1, !Δ ⊲ P : bit, Γ_2, !Δ ⊲ M : A, Γ_2, !Δ ⊲ N : A, (if)

In the intended semantics, !bit ≡ bit, while !qbit is empty. new creates a new qubit, and meas measures a qubit.

The denotational semantics of these operations is already well-understood (see Section 4.3) in the absence of higher-order types. They can all be interpreted in the category Q of superoperators described in Section 4.3. The part that is not yet well-understood is how these features interact with higher-order types.

In light of our present work, we can conclude that a model of the quantum lambda calculus consists of a linear category for duplication (C, L, T, ⊸), such that the associated category of computations C_T contains the category Q as a full monoidal subcategory. To construct an actual instance of such a model is still an open problem.
Bibliography


1http://www.theory.caltech.edu/people/preskill/ph229/


Index

\[ Y^A_{x:y} \] 91
\[ Y^A_{x:0} \] 91

adjunction, 8
composition, 9
left/right adjoint, 8
unit,counit, 8
alpha-equivalence, 42
arrow, 6
axiomatic equivalence, 45, 86, 123

\[ B_\sigma \] 37
basis along \( x,y,z \), 23
Bell’s algorithm, 27
Bell’s inequalities, 27
Bloch sphere, 22, 27
BNF, 41
bound variable, 42, 55
bra, 23

\[ C_\lambda \] 128
call-by-name, 45, 74
call-by-value, 45, 57, 77, 239
canonical first-order representation, 11

\textbf{Cat.} 1

categorical semantics, 174
category, 6
cartesian category, 11
cartesian closed, 57
equivalence of, 8
functor category, 8
linear, 58
linear-non-linear, 58
opposite category, 9
product category, 9
small,locally small, 8
trivial category, 8
characteristic matrix, 20, 38
Church-Rosser theorem, 138
Church-style, 35
classical object, 31, 110

classical objects, 57
co-Eilenberg-Moore category, 12
coe-Kleisli category, 11
coalgebra, 12
coherence map, 11
commutation of diagrams, 8
commutative comonoid, 13
comonad, 11
  idempotent, 100
  coherence properties, 101
completely positive map, 26
completeness, 48, 199
composition, 6
calculation
  category of, 56
concurrent quantum computation, 34
consistency, 59
context, 43
contraction, 20, 59, 51
controlled gate, 24
controlled-not gate, 24
coproduct, 170
  binary, 170
copying, 26, 99, 100

\textbf{CPM} 37
Curry-Howard isomorphism, 46
Curry-style, 45
dagger category, 30
  compact-closed, 56
degree of
  a term, 161
  a type, 161
denotation, 82, 86
denotational equivalence, 90
denotational semantics, 83
dense-coding algorithm, 30
  density matrix, 18
Deutsch Algorithm, 64
Deutsch algorithm, 31
diagonal structure, 12

207