September 1995

Proof Theoretic Approach to Specification Languages

Jawahar Lal Chirimar
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/ircs_reports

http://repository.upenn.edu/ircs_reports/132


This paper is posted at ScholarlyCommons. http://repository.upenn.edu/ircs_reports/132
For more information, please contact libraryrepository@pobox.upenn.edu.
Proof-Theoretic Approach to Specification Languages

Abstract
In this thesis I study FORUM as a specification language. FORUM is a higher-order logic based on the logical connectives of Linear Logic. As an initial example, I demonstrate that FORUM is well suited for specifying concurrent computations by specifying the higher-order π calculus. Next, I focus on the problem of specifying programming languages with higher-order functions, and imperative features such as assignable variables, exceptions and first-class continuations. I provide a modular and declarative specification of an untyped programming language, UML, which contains the above mentioned features. Further, I use the proof theory of FORUM to study program equivalence for the functional core of UML, augmented with assignable variables. Using my compositional specifications in FORUM, I prove equivalence of programs that have been challenging for other specification languages. Finally I study the operation semantics of DLX, a prototypical RISC machine. I specify the sequential and pipelined operational semantics of DLX with important optimizations such as call-forwarding and early branch resolution, and prove them to be equivalent. Furthermore, I study the problem of code equivalence via the FORUM specification, and, in particular, analyze the problem of code rescheduling for DLX.

Comments
Proof Theoretic Approach to Specification Languages

Jawahar Lal Chirimar

University of Pennsylvania
3401 Walnut Street, Suite 400C
Philadelphia, PA 19104-6228

September 1995

Site of the NSF Science and Technology Center for Research in Cognitive Science

IRCS Report 95-21
PROOF THEORETIC APPROACH TO SPECIFICATION LANGUAGES

JAWAHAR LAL CHIRIMAR

A DISSERTATION
IN
COMPUTER AND INFORMATION SCIENCE

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy.
1995

Dale Miller
Supervisor of Dissertation

Mark Steedman
Graduate Group Chairman
© Copyright 1995

by

Jawahar Lal Chirimar
I would, first and foremost, like to thank my parents and my wife for their endless support and love during my life as a Ph.D. student.

My committee members, Dale Miller, Carl Gunter, Robert Harper, Jon Riecke, Val Tannen and Peter Buneman provided me with important feedback on this research. I would especially like to thank Dale Miller who encouraged me to think independently and in a scientifically disciplined manner. I learned not only about technical matters from Carl, my supervisor for the first three years, but also about the global picture and the way research is done. From Andre Scedrov, who was on my committee in all but name, I learned the value of rigor – without rigor there is no mathematics and mathematics is not all rigor. I must thank Anil Nerode specially for introducing me to proof theory, functional languages and type theories. I have benefited from various technical discussions with Jean Gallier, Doug Howe, Jim Lipton and Chet Murthy.

There are many people whom I met in my stay at University of Pennsylvania. The numerous discussions with Vijay Gehlot, Ramesh Subramaniam, Anuj Dawar, Chuck Liang, Sandeep Biswas, Jon Riecke, Teow Hin Ngair, and other fellow students helped shaped my concepts and opinions. Jon Riecke, Sandeep Biswas and Ernesto Pimentel helped with proof reading this dissertation. Nandini, my wife, helped me clean up the grammar and insert all the appropriate commas and articles, hopefully! The staff of the graduate students office and business office put up very patiently with my eccentric habits and got all my papers through in time.
ABSTRACT

PROOF THEORETIC APPROACH TO SPECIFICATION LANGUAGES

Jawahar Lal Chirimar

Advisor: Dale Miller

In this thesis I study FORUM as a specification language. FORUM is a higher-order logic based on the logical connectives of Linear Logic. As an initial example, I demonstrate that FORUM is well suited for specifying concurrent computations by specifying the higher-order \( \pi \) calculus. Next, I focus on the problem of specifying programming languages with higher-order functions, and imperative features such as assignable variables, exceptions and first-class continuations. I provide a modular and declarative specification of an untyped programming language, UML, which contains the above mentioned features. Further, I use the proof theory of FORUM to study program equivalence for the functional core of UML augmented with assignable variables. Using my compositional specifications in FORUM, I prove equivalence of programs that have been challenging for other specification languages. Finally I study the operation semantics of DLX, a prototypical RISC machine. I specify the sequential and pipelined operational semantics of DLX with important optimizations such as call-forwarding and early branch resolution, and prove them to be equivalent. Furthermore, I study the problem of code equivalence via the FORUM specification, and, in particular, analyze the problem of code rescheduling for DLX.
## Contents

1 Introduction .................................................. 1

2 FORUM ...................................................... 6
   2.1 FORUM — Logic programming with multiple heads ................. 7
   2.2 Specifying $HO\pi$ in FORUM .................................. 13

3 Specifying UML ............................................... 20
   3.1 $\Lambda_v$ — Functional Core of UML .............................. 21
   3.2 $\Lambda_{vs}$ — State in UML .................................... 30
   3.3 $\Lambda_{vc}$ — Exceptions in UML ................................. 39
   3.4 $\Lambda_{vc}$ — Continuations in UML .............................. 51
   3.5 UML — Putting it together .................................... 58

4 Program Equivalence for $\Lambda_{vs}$ in FORUM .................. 60
   4.1 Defining Observational Equivalence .............................. 61
   4.2 Reduction in $\Lambda_{vs}$ preserves Observational Equivalence .... 67
List of Figures

2.1 Proof Rules for Forum .................................................. 10

2.2 Specifying inequality of natural numbers in FORUM ..................... 11

2.3 Reduction semantics for $HO\pi$-calculus .................................. 15

3.1 Syntax for $\Lambda_v$ ......................................................... 21

3.2 Natural Semantics specification for $\Lambda_v$ ................................. 23

3.3 $\Sigma_v$, Signature for $\Lambda_v$ evaluator .................................. 24

3.4 $\phi_v : \text{Values}_v \rightarrow \text{vl}, \mathcal{H}_v : \Lambda_v \rightarrow \text{tm} ............................................ 25$

3.5 $\psi_v : \text{vl} \rightarrow \text{Values}_v, \mathcal{L}_v : \text{tm} \rightarrow \Lambda_v ............................................ 25$

3.6 Clauses in $\mathcal{E}_v$ - the evaluator for $\Lambda_v$ .............................. 27

3.7 One step in the evaluation of $(M \cdot N)$ .................................. 28

3.8 Syntax for $\Lambda_{vs}$ ......................................................... 31

3.9 Natural Semantics specification for the new constructs in $\Lambda_{vs}$ .......... 32

3.10 Constants for translating $\Lambda_{vs}$ terms ................................ 33

3.11 Specification in FORUM for new constructs in $\Lambda_{vs}$ .................. 34
3.12 Syntax for \( \Lambda_{vc} \) .......................................................... 40
3.13 Natural Semantics specification for the new constructs in \( \Lambda_{vc} \) ........ 42
3.14 Constants for translating \( \Lambda_{vc} \) terms ................................ 45
3.15 Translating the new construct of \( \Lambda_{vc} \) to FORUM ...................... 45
3.16 Constants for exception management in FORUM .......................... 47
3.17 Specification in FORUM for new constructs in \( \Lambda_{vc} \) .................. 48
3.18 Syntax for \( \Lambda_{vc} \) ......................................................... 53
3.19 Constants for translating \( \Lambda_{vc} \) terms ................................. 54
3.20 Specification in FORUM for new constructs in \( \Lambda_{vc} \) ............... 54
3.21 \( C_{vc} \), translation of \( EvCont_{vc} \) to FORUM terms of type \( \forall l \to o \) 56
3.22 Specification for \text{callcc} in the presence of exceptions ............... 57
4.1 Contexts in \( \Lambda_{vs} \) .......................................................... 63
4.2 \( C_{vs} \), translation of \( EvCont_{vs} \) to FORUM terms of type \( \forall l \to o \) 65
5.1 Semantics of example instructions in DLX .................................. 84
5.2 List of DLX instructions selected for specification ......................... 85
5.3 Block diagram for the connectivity of functional blocks in the DLX ...... 86
5.4 Signature for specification of DLX ...................................... 89
5.5 Grammar for DLX programs ............................................. 90
5.6 Sequential specification of DLX ........................................ 91
5.7 Example program in DLX .................................................. 92
A.5 Matching final configurations of computations in FORUM ............ 145
Chapter 1

Introduction

Ever since we have had programming languages, we have had to *specify the computational behaviors of the languages*. In the most naive sense, we specify the actions of if $M$ then $N$ else $P$ — a typical phrase in a programming language — by saying that first execute $M$, and if the result is *true* then evaluate $N$, otherwise evaluate $P$. Even in this simple example, there are many ambiguities inherent in the above specification. For example, it is not clear whether $P$ is to be computed only when $M$ evaluates to *false*, or even if $M$ evaluates to some other value, say 3. On the one hand, we would like a specification to be precise in as much as we would like different implementations of the same specification to have identical computational behavior. On the other hand, if a specification is as precise as an actual implementation of a programming language, then the entire purpose of a specification as a tool to understand the language independently of its implementation is defeated.

Rigorous specifications which define the grammar and meaning of programming languages are indispensable in the present context for a variety of reasons. These include verifying safety of programs, developing optimizing compilers, and maintaining programs as language design evolves. Unfortunately, very few widely used programming languages have such a rigorous specification, the only two I know of being [HMT89, BR90]. In this thesis, I use
FORUM [Mil94] as a specification language. Specifications in FORUM convert computations into proofs — formal objects amenable to logical analysis within the meta-theory of FORUM. I call this style of specification *proof-theoretic* for the key emphasis placed upon proofs and their analysis.

Precise specification provides the sound basis on which one builds implementations of the language and programs. To quote [HMT89]:

...for a robust program written in an insecure language is like a house built upon sand.

The main point is that a programming language is different from its implementation — specifications are a way of making precise the behavior that an implementation must exhibit in order to implement a given programming language. Specifications play a crucial role in the present software environment where different commercial vendors are implementing the same language. It will be highly undesirable if programs written in a programming language have different behaviors in two implementations of the same language!

I expect specification languages to play a variety of roles. The most important among these are:

- a specification language should be rich enough to specify imperative, functional and concurrency features modularly
- provide modular specification of different features of the programming language, resulting in better and easier understanding of the programming language itself, and
- should be able to use the meta-theory of the specification language to study the program transformations and verify the correctness of implementations of the specifications.

In this thesis I use FORUM to specify the operational semantics of programming languages. I specify the operational semantics of UML, a prototypical functional language, and the
sequential and pipelined operational semantics of the DLX machine [HP90]. I take UML – an untyped higher-order functional language with exceptions, state and \texttt{callcc}, \textit{i.e.} first-class continuations – as a prototypical functional language. UML is the same as the untyped core Standard ML (SML) [HMT89], excluding data-types and pattern matching, augmented with \texttt{callcc}. UML is a significant language in as much as it contains both the functional and imperative features of SML, and thus provides a simpler setting to study the problems that arise due to these features in SML. UML supports first-class continuations because of the rich programming paradigm they provide. The DLX machine is representative of the eminently successful and popular RISC architectures of the last decade including Intel i860, MIPS R2000/R3000, Motorola 88000, SPARC, PowerPC. The specification of DLX resolves data and structural hazards, and implements optimizations characteristic of modern pipelined machines. The specification of pipelined DLX underlines the fact that FORUM provides the appropriate framework to specify a variety of computational processes – from as high-level and abstract as UML to as low-level and concrete as DLX.

Many attempts have been made at specifying the operational semantics of fragments of UML [Lan64, FF86, WF91, HMT89, MP92, HM92, Han90]. In this thesis I am able to address many issues regarding the specification of UML which the above semantics could not address satisfactorily. Firstly, I am working in a rich meta-theory where concurrency, higher-order functions and imperative features can be specified. This is a mixture of features which has traditionally been very difficult to specify. Secondly, in my translation computations become proofs — formal objects — which I analyze using proof transformations. Using this analysis I am able to prove many program equivalences in [MS88, OT93, SF92, MT92]. Finally, declarative and modular descriptions of imperative features such as mutable state, first-class continuations, and exceptions has been an elusive goal for specification languages, which I believe is satisfactorily answered by the specifications in FORUM.

Although the extant formal presentations of pipelines specify the temporal behavior of the pipeline [TK93], they are unable to provide a concurrent computational specification of the pipelined operational semantics. The FORUM specification of the DLX pipeline is a concurrent executable logic program – one obtains a simulation tool for free. In my thesis, I specify
the sequential and pipelined operational semantics for DLX, with important optimizations such as call-forwarding and early branch resolution [HP90]. Since floating-point operations and interrupts introduce unilluminating details to the specification, I have excluded them from the DLX instruction set that I specify. I prove the crucial equivalence theorem asserting the equivalence of the sequential and pipelined specifications of DLX. Furthermore, I define notions of program equivalences for DLX programs, and prove the correctness of the code rescheduling typically done for RISC machines [HP90]. The declarative specification of pipelined DLX operational semantics and the proofs of correctness of code rescheduling underlines the richness of FORUM as a specification language.

The point behind the specifications of imperative features is to extract the logical essence of imperative extensions of programming languages. The specification of state in natural semantics, higher order logic or other similar meta-theories [HMT89, MT92] represents state by non-logical means, such as a finite function, making the extension of state non-modular. In FORUM state is represented by logical propositions and is maintained in the sequent by logical rules. Thus, one can reason about state variables using cut-elimination. The understanding of this logical nature of imperative features underlines the richness of FORUM and its meta-theory.

The thesis has five main parts. In the first part I explain the logic programming methodology of FORUM, and then specify Higher-Order π Calculus (HOπ-calculus), a typical calculus for concurrent processes, [MPW92a, MPW92b] substantiating the idea that FORUM provides an appropriate framework for specifying concurrent processes. In the second part, I specify UML in stages starting from the functional part of UML, and adding exceptions, mutable store and callcc modularly in separate steps. I prove that the FORUM specification of UML without callcc is the same as the specification in [HMT89]. In the third part, I define a notion of program equivalence induced by the translation into FORUM, and prove that it coincides with the standard definition of program equivalence. Furthermore, using FORUM specifications I prove several of the program equivalences involving mutable store in [MS88, Sie93, SF92, OT93]. In the fourth part, I specify the sequential and pipelined operational semantics for DLX. I also prove that the sequential and pipelined operational
semantics are equivalent. Furthermore, I formulate a notion of program equivalence for
DLX programs, and prove correctness of code-rescheduling for the DLX machine. In the
final section, I explain the extensions that I seek of my current research and how I intend
to carry those out.
Chapter 2

FORUM

In this chapter I introduce FORUM, a new meta-logic proposed in [Mi94]. FORUM can encode linear logic [Gir87] without using any non-logical constants. On the one hand, provability in FORUM is the same as provability in linear logic. On the other hand, in FORUM all right hand rules permute with each other — a property which is not true of the proof systems for linear logic in [Gir87]. The novelty of FORUM is in the choice of connectives which makes all right hand rules permute. Since uniform proofs are complete for FORUM, following [MNPS91] a logic programming language can be designed for FORUM. FORUM extends earlier work in designing logic programming languages from linear logic [HM91, AP90] in the sense that the provability in FORUM is the same as the provability in linear logic.

Clauses in FORUM can have multiple heads. I show some programming examples which exploit multiple heads to represent synchronization in FORUM. The basic intuition is that concurrent computations can be represented in FORUM. I substantiate this claim by translating a particular presentation of a fragment of Higher-Order π (HOπ-calculus) calculus [San92a] which is rich enough to encode Lazy Lambda calculus [San92b, San92a, Chi94]. The parallel combinator of HOπ-calculus is mapped to $\varnothing$, the multiplicative disjunction of Linear Logic. This identification of concurrency with proof search in a multiple conclusion
logic where all the right-hand rules permute with each other makes the intuitions technically precise. The handling of names using the restriction operator in $HO\pi$-calculus and the universal quantifier in FORUM are quite different, and consequently the translation into FORUM is sound but not complete. The computational mechanism of FORUM provides a new flavor of process theories which are very expressive, and the translation shows how computations maybe mapped from $HO\pi$-calculus to FORUM.

In the first section, I introduce FORUM and define its syntax and proof rules precisely. I illustrate the programming style in the presence of multiple heads via some examples. In the second section, I introduce the syntax, structural equivalence, and reduction semantics of $HO\pi$-calculus. I translate $HO\pi$-calculus into FORUM and prove that if process $P$ reduces to $Q$, then the translation of $Q$ entails the translation of $P$. Furthermore, I illustrate the difference between the restriction operator of $HO\pi$-calculus and the universal quantifier in FORUM.

2.1 FORUM — Logic programming with multiple heads

FORUM [Mil94], is best explained as a particular presentation for Linear Logic which gives us access to the entire Linear Logic as a logic programming language in the sense of [MNPS91, AP90]. Linear logic was introduced in [Gir87] as a new logic which decomposed the connectives of the familiar classical and intuitionistic logics. This finer analysis of connectives had immediate implications in the design of logic programming languages which analyzed connectives as directions for proof search [MNPS91]. LO[AP90] and Lolli [HM91] were two new logic programming languages which resulted from different sets of connectives of linear logic. However, the logical constants in neither of these languages were rich enough to encode the entire linear logic. In [Mil94] classical linear logic is encoded in FORUM without using any non-logical constants. The logical connectives in FORUM are $\neg$, $\Rightarrow$, $\odot$, $\&$, $\forall$, $\bot$ and $\top$. $\neg$, $\odot$, $\&$, $\forall$, $\bot$ and $\top$ are linear logic connectives as defined in [Gir87]. Instead of the modalities of linear logic, FORUM has $\Rightarrow$, the intuitionistic implication. In [Mil94] it was proved that the proof system of FORUM has the uniform
proof and focussing property [MNPS91, AP90]. FORUM can be thus thought of Linear Logic and Church's simple theory of types [Chu40] put together. I begin by defining the syntax of FORUM.

**Definition 2.1** [Types, Terms and Formulas in FORUM] Let $\Theta$ be a set of base types and $o \in \Theta$ the type of propositions. The set of well formed types is defined as

- if $\sigma \in \Theta$ then $\sigma$ is a type, and
- if $\tau_1$ and $\tau_2$ are types then so is $\tau_1 \rightarrow \tau_2$.

Let $\Sigma$ be a set of pairs whose first component is a term, and the second component is the type of the term, written as $f : \tau$, if $f$ is a term with type $\tau$ in $\Sigma$. $\neg: o \rightarrow o \rightarrow o$, $\Rightarrow: o \rightarrow o \rightarrow o$, $\varnothing: o \rightarrow o \rightarrow o$, $\&: o \rightarrow o \rightarrow o$, $\forall: (\tau \rightarrow o) \rightarrow o$, $\perp: o$ and $\top: o$ are the logical constants in $\Sigma$. $\Rightarrow$ denotes intuitionistic implication and the infix symbol $\neg$ denotes the converse of $\Rightarrow$. The set of terms over $\Sigma$ is defined as:

- If $c : \tau \in \Sigma$ then $c$ is a term of type $\tau$.
- If $f : \tau \rightarrow \sigma$ and $t : \tau$ then $(ft)$ is a term of type $\sigma$.
- If $x$ is a variable of type $\sigma$ and $t : \tau$ then $\lambda x. t$ is of type $\sigma \rightarrow \tau$.

Terms of type $o$ are defined to be formulas. The *order of a type* $\tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_n \rightarrow \tau_0$ is the one plus the max of the orders of $\tau_1 \ldots \tau_n$, and $\tau_0 \in \Theta$. The order of the elements of $\Theta$ is 0. For a non-logical constant $c : \tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_n \rightarrow \tau_0 \in \Sigma$, $\tau_0$ is a member of $\Theta$, and if $\tau_0$ is $o$ then $c$ is called a *predicate*.

I follow [Bar84] in conventions regarding free and bound variables and $\alpha$ conversion. As usual, $\rightarrow$ associates to the right and application to the left. The logical constants are written in the familiar infix form, and I write $\forall x. \lambda x. t$ as $\forall x : \tau. t$. Let $t = s$, for $\lambda$-terms $t$ and $s$ mean that $t$ and $s$ are $\alpha$ equivalent. If the variable $x$ and term $s$ are of the same type, then
\[ t[x := s] \] denotes the capture-free substitution of \( s \) for \( x \) in \( t \). Besides \( \alpha \) conversion, terms are also related by the following rules of \( \beta \) and \( \eta \) conversions:

- The term \( s_1 \) \( \beta \)-converts to the term \( s_2 \) if \( s_1 \) contains a subformula occurrence of the form \( ((\lambda x.t_1)t_2) \) and \( s_2 \) arises from replacing that subformula occurrence with \( t_1[x := t_2] \).

- The term \( s_1 \) \( \eta \)-converts to the term \( s_2 \) if \( s_1 \) contains a subformula occurrence of the form \( \lambda x.(t\,x) \), in which \( x \) is not free in \( t \), and \( s_2 \) arises from replacing that subformula occurrence with \( t \).

The proof system for FORUM as presented in figure 2.1 is a minor variation on the one in [Mil94]. The sequents comprise of five parts, \( \Sigma, \Psi, \Delta, B \) and \( \Gamma \). The signature of the terms in the sequent is given by \( \Sigma \). The intuitionistic part of the sequent, \( \Psi \), is treated like a set, \emph{i.e.}, contraction, weakening and exchange are allowed on formulas in \( \Psi \). The linear parts of the sequent, \( \Delta \) and \( \Gamma \), are treated as multisets of formulas, allowing only exchange on these parts of the sequent. In the sequent \( \Sigma : \Psi \,; \Delta \xrightarrow{B} \Gamma \), one applies left rules to the formula \( B \). By abuse of notation I write \( B, \Psi \) to mean \( \{B\} \cup \Psi \), and \( \Delta_1, \Delta_2 \) to stand for the multiset union of the multisets \( \Delta_1 \) and \( \Delta_2 \), and \( B, \Delta_2 \) to stand for the multiset union of the multisets \( \{B\} \) and \( \Delta_2 \). \( \Psi[c := t] \) denotes the capture-free substitution of \( t \) for \( c \) in all formulas in the set \( \Psi \) and \( \Delta[c := t] \) denotes the capture-free substitution of \( t \) for \( c \) in all formulas in the multiset \( \Delta \). I write \( A \equiv_{\Sigma} B \) as an abbreviation for the statement that both \( \Sigma : ; A \rightarrow B \) and \( \Sigma : ; B \rightarrow A \) are provable in FORUM. I say that \( A \) is logically equivalent to \( B \) when \( A \equiv_{\Sigma} B \). One can prove a cut-elimination theorem for FORUM, which states that \( \text{CutL}, \text{CutI}, \text{and CutS} \) are redundant in FORUM [Mil94]. Also note that if \( \Sigma : \Psi ; \Delta \rightarrow \Gamma \) is provable, \( \Sigma \subseteq \Sigma_1 \) and \( \Psi \subseteq \Psi_1 \), then \( \Sigma_1 : \Psi_1 ; \Delta \rightarrow \Gamma \) is also provable.

Following the line of reasoning in [MNPS91], logic programs can be viewed as collections of formulas specifying the meaning of non-logical constants, and computation is identified
The rule $\forall$-R has the proviso that $y$ is not declared in the signature $\Sigma$.  

Figure 2.1
with the search for uniform proofs. The key feature of FORUM is that the right-hand side of the sequent can now have more than one formula. *How does one interpret uniform proofs in the presence of more than one goal formula?* The answer in [Mil94] comes from the concept of permutabilities in proof theory [Kle64]. Informally, it is required that the order in which goal formulas are processed does not affect the success of the proof search. This novelty lets us represent concurrent computations in FORUM, as was exhibited in [Mil94] by specifying Algol-like implementations of CML, and First-order π-Calculus in [Mil93].

I illustrate the computational mechanism of FORUM with a simple example. I begin by specifying the natural numbers in FORUM. I introduce a new type in FORUM called nat and two new constants $z : \text{nat}$ and a function $s : \text{nat} \rightarrow \text{nat}$. The intended meaning is that nat is the type of natural numbers, $z$ denotes 0, and $s$ denotes the successor function. I now want to define a predicate $\text{neq} : \text{nat} \rightarrow \text{nat} \rightarrow o$, which should be provable of the two terms $m$ and $n$ of type nat, if $m$ is not equal to $n$. The meaning of $\text{neq}$ is specified by the universal closure of the clauses in figure 2.2, called $C_{\text{neq}}$. The collection of clauses is called the program for the non-logical constant $\text{neq}$. In the specification I use a new logical connective $1$. It can be defined as $\bot \rightarrow \bot$, and the proof rule can be derived correspondingly. I show the right-hand side rule below.

$$\Sigma : \Psi ; \frac{}{1} 1 \rightarrow R$$

To check $\text{neq}$ for $(sz)$ and $(ssz)$, I try to construct a proof of

$$z : \text{nat}, s : \text{nat} \rightarrow \text{nat} : C_{\text{neq}} ; \frac{}{(\text{neq} (sz) (ssz))}.$$
If the sequent is provable then the two numbers are not equal. For the example at hand, the proof is constructed below. The first rule I apply is the backchain rule. The backchain rule is an abbreviation, instead of constructing the following proof

\[
\begin{align*}
\Sigma : \Psi, C \vdash B ; \Delta \rightarrow C, \Gamma & \quad \Sigma : \Psi, C \vdash B ; \Delta \rightarrow B & \text{initial} \\
\Sigma : \Psi, C \vdash B ; \Delta & \rightarrow B, \Gamma & \rightarrow \rightarrow -L \\
\Sigma : \Psi, C \vdash (B_1 \varphi \ldots \varphi B_n) ; \Delta & \rightarrow (B_1 \varphi \ldots \varphi B_n), \Gamma_1, \Gamma_2, \ldots, \Gamma_n, \Gamma_{n+1} & \text{decide2} \\
\Sigma : \Psi, C \vdash (B_1 \varphi \ldots \varphi B_n) ; \Delta & \rightarrow \Gamma_1, B_1, \Gamma_2, \ldots, \Gamma_n, B_n, \Gamma_{n+1} & \text{exchange}, \varphi \rightarrow R
\end{align*}
\]

I abbreviate it as

\[
\begin{align*}
\Sigma : \Psi, C \vdash (B_1 \varphi \ldots \varphi B_n) ; \Delta & \rightarrow C, \Gamma_1, \Gamma_2, \ldots, \Gamma_n, \Gamma_{n+1} & \text{backchain} \\
\Sigma : \Psi, C \vdash (B_1 \varphi \ldots \varphi B_n) ; \Delta & \rightarrow \Gamma_1, B_1, \Gamma_2, \ldots, \Gamma_n, B_n, \Gamma_{n+1} & \text{backchain}
\end{align*}
\]

\(B\) is an abbreviation for \(B_1 \varphi \ldots \varphi B_n\), and \(\Gamma\) is an abbreviation for \(\Gamma_1, \ldots, \Gamma_{n+1}\).

The formula on the right hand side unifies the head of the clause \((\text{neq} (s X) (s Y)) \vdash (\text{neq} X Y)\). So I then have to prove \((\text{neq} z (s z))\). The goal now unifies with the head of the clause \((\text{neq} z (s X)) \vdash 1\), leaving me to prove \(1\). I complete the proof using the \(1\) rule. In this manner, proof search for cut-free proofs is identified with computation.

\[
\begin{align*}
\overline{z : \text{nat}, s : \text{nat} \rightarrow \text{nat} : C_{\text{neq}} ; \rightarrow 1} & \quad \text{backchain} \\
\overline{z : \text{nat}, s : \text{nat} \rightarrow \text{nat} : C_{\text{neq}} ; \rightarrow (\text{neq} z (s z))} & \quad \text{backchain} \\
\overline{z : \text{nat}, s : \text{nat} \rightarrow \text{nat} : C_{\text{neq}} ; \rightarrow (\text{neq} (s z) (s s z))} & \quad \text{backchain}
\end{align*}
\]

The novelty of FORUM lies in the fact that the clauses can contain multiple heads, \(i.e.\) formulas like \(A \varphi B\). I want to specify a predicate \(\text{inc} : (\text{nat} \rightarrow o) \rightarrow o\), which takes a predicate of type \(\text{nat} \rightarrow o\) as an argument, such that every time \(\text{inc}\) is executed in a proof a new number is returned. Let \(\text{ctr} : \text{nat} \rightarrow o\) be a non-logical constant, denoting a memory cell in the environment storing the next number to be used by \(\text{inc}\). Suppose \(\text{ctr}\) is initialized to some number, then the only clause required will be the universal closure of the following clause called \(\text{Inc}\).

\[
[(\text{inc} P) \varphi(\text{ctr} X)] \vdash [(P X) \varphi(\text{ctr} (s X))].
\]
Both \( \text{inc} \, P \) and \( \text{ctr} \, X \) must be on the right hand side of the sequent before one can backchain on the clause. In this sense, clauses with multiple heads enforce synchronization between various predicates. The last step in the uniform proof of

\[
\text{inc} : (\text{nat} \rightarrow o) \rightarrow o, \text{ctr} : \text{nat} \rightarrow o, P : \text{nat} \rightarrow o : \text{Inc}; \quad \rightarrow (\text{inc} \, P) \varphi (\text{ctr} \, (s \, z))
\]

is shown below.

\[
\begin{array}{c}
\text{ctr, inc, P : Inc;} \\
\rightarrow (P \, (s \, z)) \varphi (\text{ctr} \, (s \, s \, z)) \\
\end{array}
\]

To use the clause \( \text{Inc} \), I need to have both \( (\text{inc} \, P) \) and \( (\text{ctr} \, (s \, z)) \), or, in other words, \( (\text{inc} \, P) \) and \( (\text{ctr} \, (s \, z)) \) synchronize with each other. Now, by backchaining on \( \text{Inc} \) the goal becomes \( (P \, (s \, z)) \varphi (\text{ctr} \, (s \, s \, z)) \). Note that because \( \text{ctr} \) is linear it is destructively read by the \( \text{Inc} \) clause — the number stored in \( \text{ctr} \) is increased by one as a result of backchaining.

The examples above point towards a relationship between concurrency and proof search in multiple conclusion logic where the right hand side rules permute. The idea is that all the processes on the right hand side are free to compute concurrently and synchronize with each other. Some of these intuitions will be made precise in the next section, where I specify a particular presentation of a fragment of \( HO\pi \)-calculus in FORUM.

### 2.2 Specifying \( HO\pi \) in FORUM

In this section I specify a fragment of \( HO\pi \)-calculus as defined in [San92a] to make concrete my claim that FORUM can be used to represent both abstraction and concurrency. I begin with a brief presentation of \( HO\pi \)-calculus, and refer the interested reader to [Mil89, MPW92a, MPW92b, San92b, San92a] for a detailed introduction to, and analyses of \( HO\pi \)-calculus and \( \pi \)-calculus. The motivating idea of \( HO\pi \)-calculus is to provide higher order communication in the framework of synchronous mobile process algebras. The fragment of \( HO\pi \)-calculus that I consider can encode call-by-value and call-by-name lambda
calculus [Mil90, San92b, San92a]. I use \( x, y, \ldots \) possibly subscripted, to stand for \textit{Names} and their capitalized versions to range over \textit{Vars}. Moreover, \( K \) stands for a process or a name, and \( U \) stands for a variable or a name.

\textbf{Definition 2.2} [Syntax for HO\(\pi\)-calculus.] The processes in HO\(\pi\)-calculus are defined by \( P \) and the prefixes by \( \alpha \).

\[
P := 0 \\
| (X) \\
| (P \mid P) \\
| (\nu x(P)) \\
| (\alpha.P)
\]

\[
\alpha := x(U) \mid \pi(K)
\]

0 is the inactive process — not capable of any action or interaction. \( x(U).P \) accepts input for variable \( U \) along the channel \( x \), while \( \pi(K).P \) transmits \( K \) along the channel \( x \). \( \nu x(P) \) makes the name \( x \) private to the process \( P \), and \( P \mid Q \) places the two processes \( P \) and \( Q \) in parallel. The variable \( U \) is bound in \( x(U).P \), and \( x \) is bound in \( \nu x(P) \). I often abbreviate \( \alpha.0 \) as \( \alpha \). The reduction relation for HO\(\pi\)-calculus is divided in two parts, the structural equivalence and the reduction relation. We illustrate the reduction semantics and explain the syntax conceptually with some examples.

\textbf{Example 2.3} Examples of reductions in HO\(\pi\)-calculus.

1. \( \pi(Q).P \mid x(Y).R \rightarrow_\pi P \mid R[Y := Q] \)
2. \( \nu x(\pi(x).P \mid Q) \mid w(y).R \rightarrow_\pi \nu x(Q \mid R[y := x]) \mid P, x \not\in \text{FV}(P) \) and \( x \not\in \text{FV}(w(y).R) \).
3. \( \pi(Q).0 \mid x(Y)(Y \mid P) \rightarrow_\pi Q \mid P \)
\[
\begin{align*}
\pi(Q).P \mid x(Y).R \Rightarrow_P P \mid R[Y := Q] \quad &\text{ComP} \\
\pi(y).P \mid x(z).R \Rightarrow_P P \mid R[z := y] \quad &\text{ComN} \\
P \Rightarrow_P Q \quad &\nu x(P) \Rightarrow_P \nu x(Q) \quad &\nu R \\
&\quad P \Rightarrow_P P_1 \quad P_1 \Rightarrow_P Q \quad \Rightarrow_P P_1 \mid Q \quad &\text{Par} \\
&P \equiv_P P_1 \quad P_1 \Rightarrow_P Q_1 \quad Q_1 \equiv_P Q \quad &\Rightarrow_P P \quad &\text{Struct} \\
&P \Rightarrow_P Q \quad &\Rightarrow_P P \quad &\text{Refl} \\
&P \Rightarrow_P Q_1 \quad Q_1 \Rightarrow_P Q \quad &\Rightarrow_P P \quad &\text{Trans}
\end{align*}
\]

Figure 2.3: Reduction semantics for HO\(\pi\)-calculus

In example 2.3.1, \(\pi(Q).P\) transmits \(Q\) along the channel \(x\) and \(x(Y).R\) receives \(Q\) on the \(x\). The rule underlines one key feature of the calculus — the communications are synchronous, i.e. a process, e.g. \(\pi(Q).P\), that wants to send a message waits until there is a process, e.g. \(x(Y).P\), in the environment which will accept that message, hence the name synchronous message passing. This also ensures a flavor of sequencing in the process, e.g. in \(\alpha_1.\alpha_2.P\) the action corresponding to \(\alpha_1\) must happen before the action for \(\alpha_2\) can occur. Another basic feature of HO\(\pi\)-calculus is the capability of changing the connectivity amongst processes during computation. In example 2.3.2, the channel \(x\) in \(\nu x(\pi(x).P \mid Q)\) is a private channel between \(\pi(x).P\) and \(Q\). However, it is possible for \(x\) to extrude its scope and be sent to \(w(y).R\). In this sense, the connectivity of the processes can change during the computation, and hence the name mobile processes. Example 2.3.3 illustrates the novelty of Higher- order processes. \(x(Y),(Y \mid P)\) can be viewed as a process that will execute whatever process it receives from the environment in parallel with \(P\). Higher-order communications provides the familiar substitution of \(\lambda\)-calculus in the context of HO\(\pi\)-calculus. [San92b, San92a] gave a nice encoding of Lazy \(\lambda\)-calculus in HO\(\pi\)-calculus.
and showed the correspondence between $HO\pi$-calculus and First Order $\pi$. [Chi94] compares the encoding of Lazy $\lambda$-calculus in $HO\pi$-calculus to the Continuation-passing semantics.

Now, I define the structural equivalence and reduction semantics of $HO\pi$-calculus formally.

**Definition 2.4** [Structural Equivalence, $\equiv_\pi$] $\equiv_\pi$ is defined as the least congruence containing the following rules.

1. $P \parallel 0 \equiv_\pi P$.
2. $P \parallel Q \equiv_\pi Q \parallel P$.
3. $(P \parallel Q) \parallel R \equiv_\pi P \parallel (Q \parallel R)$.
4. $\nu x(P) \parallel Q \equiv_\pi \nu x(P \parallel Q), x \not\in FV(Q)$.
5. $\nu x(P) \equiv_\pi P, x \not\in FV(P)$.
6. $\nu x(\nu y(P)) \equiv_\pi \nu y(\nu x(P))$.

The reduction semantics for $HO\pi$-calculus is specified as an unlabeled system in terms of proof rules. One notable feature is that the separation of structural equivalence from reduction rules enables a concise presentation of the latter. It should be noted that reduction is built as a congruence for all term constructors except prefixing, as seen in the rules $\nu R$, $Par$, $agentR$ and $Struct$; this feature forces a strict order of evaluation on prefixed processes, as we saw in example 2.3.1.

**Definition 2.5** [Translation of $HO\pi$-calculus in FORUM] Let $i$ be a new basic type, the type of names, and $\Sigma_\pi$ be the set consisting of the following constants needed to describe the translation of $HO\pi$-calculus into FORUM.
\[
\begin{align*}
\text{send}_{nm} & : i \rightarrow i \rightarrow o \rightarrow o \\
\text{receive}_{nm} & : i \rightarrow (i \rightarrow o) \rightarrow o \\
\text{send}_{pr} & : i \rightarrow o \rightarrow o \rightarrow o \\
\text{receive}_{pr} & : i \rightarrow (o \rightarrow o) \rightarrow o
\end{align*}
\]

The translation \((\_)^o\) takes a process in \(HO\pi\)-calculus to a formula, \textit{i.e.} terms of type \(o\), in \textsc{Forum}.

\[
\begin{align*}
0^o &= \bot \\
X^o &= X \\
(P \mid Q)^o &= P^o \varphi Q^o \\
(\nu x(P))^o &= \forall x : i. P^o \\
(\text{receive}_{pr} x Q)^o &= \text{send}_{pr} x Q^o P^o \\
(x(Y).P)^o &= \text{receive}_{pr} x \lambda Y. P^o \\
(\text{send}_{nm} x y P)^o &= \text{send}_{nm} x y P^o \\
(x(z).P)^o &= \text{receive}_{nm} x \lambda z. P^o
\end{align*}
\]

Let \(E_x\) be the set consisting of the universal closure of the following clause, which describes the meaning of the non-logicals in \(\Sigma_x\).

\[
\begin{align*}
(\text{receive}_{nm} x R \varphi (\text{send}_{nm} x y P)) & \land \ (R y) \varphi P, \quad \text{Name Cl} \\
(\text{receive}_{pr} x R \varphi (\text{send}_{pr} x Q P)) & \land \ (R Q) \varphi P, \quad \text{Process Cl}
\end{align*}
\]

Synchronization for processes exchanging names, \textit{i.e.} terms of type \(i\), is specified by \textit{Name Cl}, while \textit{Process Cl} specifies synchronization for processes exchanging processes, \textit{i.e.} terms of type \(o\). For the clauses in \(E_x\) to make sense, I need to show that the translation commutes with substitution. The proof is a straightforward induction on the structure of the process \(P\).

Lemma 2.6 (Substitution Lemma for \((\_)^o\)) Let \(P\) and \(Q\) be processes, then
This translation is a simple extension of the one in [Mil93] to the case of $HO\pi$-calculus. Although I cover a small fragment of the calculus here, it is easy to extend these ideas along the line of reasoning in [Mil93] to handle richer versions of $HO\pi$-calculus. However, specifying $HO\pi$-calculus augmented with constants and agents, i.e., abstractions over variables or names in processes is rather subtle. The translation is fairly simple, and gives concrete intuition about the nature of concurrency in FORUM. Concurrency is identified with proof search in a multiple conclusion logic, where all the right-hand rules permute with each other. The usage of $\varnothing$ as the translation of the parallel combinator of $HO\pi$-calculus in the above translation makes this intuition technically precise. The fact that the intention of the $HO\pi$-calculus is captured by the translation is underlined by lemma 2.7 and theorem 2.8.

Lemma 2.7 Given two processes $P$ and $Q$ in $HO\pi$-calculus, if $P \equiv_\pi Q$ then $P^o \equiv_{\Sigma_\pi} Q^o$.

Theorem 2.8 Let $P$ and $Q$ be $HO\pi$-calculus processes, and $\Sigma$ contain the names and process variables in $P$ and $Q$.

$\text{if } P \Rightarrow_\pi Q \text{ then } \Sigma_\pi, \Sigma : \mathcal{E}_\pi ; Q^* \longrightarrow P^*$

is provable in FORUM.

The reductions in translated processes have a simple shape. One essentially reorganizes the shape of the process using structural equivalences until either NameCl or ProcessCl
is applicable. Although it is an interesting topic, I do not consider the issue of explaining interesting equivalence relations on $HO\pi$-calculus in FORUM, because the above translation makes my point — concurrent computations can be represented in FORUM. In [Mil93] bisimulations and trace equivalences for First Order $\pi$-calculus without $\nu$ are analyzed in the framework of a FORUM like language.

The translation of $HO\pi$-calculus into FORUM is proved sound by theorem 2.8, but is it complete? This question was answered affirmatively in [Mil93] for the $\pi$-calculus without the restriction operator. However, in FORUM

$$\Sigma ::= \forall x. \forall y. Q[e := x][d := y] \longrightarrow \forall z. Q[e := z][d := z]$$

is provable, where $Q$ is the translation of some process into FORUM. If the translation is complete, then for any process $Q$, this would imply that

$$\nu x(Q) \Rightarrow \nu x(\nu y(Q)).$$

This is false. Hence, the universal quantifier has a logical nature which is richer than that of the restriction operator in $HO\pi$-calculus. In the later chapters, I explore various specifications in the “process theory” obtained from FORUM, and its expressive power and abstraction mechanisms are made clear. Nonetheless, the question of a proof theoretic analogue of $HO\pi$-calculus remains, and there is some work in progress with new quantifiers which might shed more light on the restriction operator.
Chapter 3

Specifying UML

In this chapter I specify UML (Untyped ML) — untyped core SML excluding pattern matching and data-types, augmented with callcc — in FORUM. I prove that a program $P$ in UML without callcc evaluates to a value $V$ as per the specification in FORUM, if and only if $P$ evaluates to $V$ as per the specification in [HMT89]. The correspondence between the two operational semantics is restricted to core UML without callcc because [HMT89] does not provide a specification for callcc. The main point of this section is that FORUM allows us to specify imperative features — exceptions, mutable state, first-class continuations — in a modular and declarative way. In particular, I first specify the functional core of UML, and then specify exceptions, mutable state, and first-class continuations independently. If I need the specification for the functional core, and any combination of the imperative parts of UML, I just put the corresponding specifications together. For example, if I want the specification for the functional core of UML with exceptions, all I have to do is put my specification for the functional core of UML together with the specification for exceptions!

Modularity of specifications is as helpful in understanding the design of a programming language as it is in the design of the language itself. However, modularity is crucial not only for such esoteric purposes as ‘understanding’ and ‘designing’ languages, but also for many practical concerns including proving correctness of implementations [HM92], verifying
The syntax of $\Lambda_v$, the functional core of UML, is very similar to the functional part of
SML without data-types. The syntax of \( \Lambda_v \) is defined formally in figure 3.1. The language contains integers, \( \mathbb{Z} \), and booleans, \( \mathbb{B} \), as constants. Arithmetic operators and equality test for integers are included in the language as term constructors in the set \( \mathcal{O} \). Functional abstraction in the form of \( \lambda \) abstraction and application are represented by \( \lambda x. M \) and \( (M N) \) respectively. I include in the language \textbf{let val} \( x = M \) \textbf{in} \( N \) which is treated like \( ((\lambda x. N) M) \). \textbf{let fun} \( f x = M \) \textbf{in} \( N \) allows recursive definitions in \( \Lambda_v \). \( \bullet \) is a token, like the only value of type unit in SML [HMT89]. The evaluation contexts, \( EvCont \), in figure 3.1, are a way of parsing a given \( \Lambda_v \)-term to find out the next redex to be contracted during evaluation. One can write standard functional programs in this language. The following program, \texttt{exp}, calculates \( m \) raised to the power \( n \) for non-negative numbers \( m \) and \( n \). The constant \( - \) stands for subtraction and \( * \) for multiplication.

\[
\begin{align*}
\textbf{let fun} & \quad f \quad x \quad = \quad \lambda y, z, \quad \text{if} \quad (= \quad x \quad 1) \quad y \quad (f \quad (-x \quad 1) \quad (*y \quad z) \quad z) \\
& \quad \textbf{in} \\
& \quad \lambda y, x, \quad \text{if} \quad (= \quad x \quad 0) \quad 1 \quad (f \quad x \quad y \quad y)
\end{align*}
\]

The operational semantics for \( \Lambda_v \) as presented in figure 3.2 is culled out from the specification for SML in [HMT89] by using substitutions instead of environments. This style of presentation is called \textit{natural semantics} following [Kah87]. The evaluator is presented as a series of rules, all of which have a simple format — to evaluate an expression first evaluate the subexpressions and then put the results together as per the outermost term constructor. For example, to evaluate \( (M \quad N) \), first evaluate \( M \) to a function, \( \lambda x. P \), and \( N \) to a value \( U \), then finally evaluate \( P[x := U] \) to a value \( V \) — the value of \( (M \quad N) \). The evaluator thus specified is \textit{call-by-value}, as the argument to a function is evaluated before it is passed to the function. \( fVU \) denotes the constant in \( \Lambda_v \), which is obtained as a result of performing \( f \) viewed as an arithmetic operator on the numbers denoted by \( V \) and \( U \), \textit{e.g.} \( +5 \quad 4 \) denotes the constant 9.

\textit{The order of evaluation of the subterms is not a part of the syntax — it is an additional requirement that the hypothesis to the rules should be read from left-to-right.} In a language with partial arithmetic operations such as \( \div \), division, the order of evaluation is crucial.
\[
\begin{array}{c}
\frac{c}{c} \quad c \in \mathcal{Z} \cup \mathcal{B} \\
\frac{\lambda x. M}{\lambda x. M} \\
\frac{M \Downarrow \lambda x. P \quad N \Downarrow U \quad P[x := U] \Downarrow V}{M \downarrow N \downarrow V} \\
\frac{M \Downarrow U \quad N[x := U] \Downarrow V}{\text{let val } x = M \text{ in } N \Downarrow V} \\
\frac{N[f := \lambda x. \text{let fun } f x = M \text{ in } N] \Downarrow V}{\text{let fun } f x = M \text{ in } N \Downarrow V} \\
\frac{M \Downarrow V \quad N \Downarrow U \quad f \in \mathcal{O}}{f M \Downarrow f V U} \\
\frac{M \Downarrow \text{true}}{\text{if } M \quad N \Downarrow V} \\
\frac{M \Downarrow \text{false}}{\text{if } M \quad N \quad P \Downarrow V}
\end{array}
\]

Figure 3.2: Natural Semantics specification for $\Lambda_v$

For example, consider the program $(+ P (\div 5.0))$, where $P$ is any non-terminating program. In a language with exceptions, evaluating the program left-to-right will cause it to diverge. However, evaluating the program right to left may raise a division-by-0 exception. Using the given specification, left to right evaluation will result in an infinite search for a computation tree reflecting the non-termination of $P$, whereas right-to-left evaluation will cause a finite evaluation tree in which one can detect the division-by-0 error.

The specification of $\Lambda_v$ in FORUM requires that one translates $\Lambda_v$ terms into the higher order abstract syntax of FORUM. The translation reveals the binding structure of the language. Issues such as capture free substitution and $\alpha$-conversion in the object language, i.e. $\Lambda_v$, are taken care of by substitution and the binding mechanism of the meta-language, i.e. FORUM. The terms and types required to define the translation comprise the set $\Sigma_v$, and are defined in figure 3.3. I define two basic types, $\mathfrak{v}l$ and $\mathfrak{t}m$, along with a coercion function, $\langle \cdot \rangle$, mapping terms of type $\mathfrak{v}l$ to terms of type $\mathfrak{t}m$. $\mathfrak{v}l$ is the type of values, and $\mathfrak{t}m$ is the type of terms. I then introduce terms at appropriate types to encode the terms of $\Lambda_v$, 23
abs : (vl → tm) → vl
c : vl

(, ) : vl → tm
f : tm → tm → tm
app : tm → tm → tm
cond : tm → tm → tm → tm
letval : (vl → tm) → tm → tm
letfun : (vl → tm) → (vl → vl → tm) → tm
ifbr : vl → tm → tm → tm
apply : tm → tm → o
eval : tm → (vl → o) → o

c. g. app is a term construct which will be the target of application terms in the translation. For every integer and boolean, I introduce a constant of vl type. For every operator f, I introduce a constant f in the signature of type tm → tm → tm.

Following is the translation of exp into a term of type tm in FORUM.

letfun λf. (abs λy. (abs λx. cond (= x (0)) (1) (app (app (app f x) y) y)))

λf, x. (abs λy. (abs λz. cond (= x (1)) y (app (app (app f (– x 1)) (+ y z)) z)))

The first argument of letfun is the body of the let fun declaration, and the second argument is the function declaration. The body is parameterized over the function variable defined by the let fun as made explicit by the meta-level λ-binding of f in the first argument of letfun. The example also illustrates how the λ bindings in the object language get converted into λ-bindings in the meta-language. For example, λx. M is translated as (abs λx. M°), M° is the translation of M, and the λ binding in abs is at the meta-level. Other than this clean explanation of variable bindings, the translation, although heavy on usage of new syntax, is similar in spirit to parsing concrete terms into abstract syntax trees. The translation of Lω terms to FORUM is rather cumbersome, but it is crucial to the statement of the correspondence theorem between FORUM specification and the natural semantics.
\[ \begin{align*}
\phi_v(x) &= x \\
\phi_v(\lambda x. M) &= \text{abs } \lambda x : \text{vl}. \mathcal{H}_v(M) \\
\phi_v(c) &= c \quad c \in \mathbb{Z} \cup \mathbb{B} \cup \{\bullet\}
\end{align*} \]

\[ \begin{align*}
\mathcal{H}_v(V) &= \langle \phi_v(V) \rangle \\
\mathcal{H}_v(f M N) &= f \mathcal{H}_v(M) \mathcal{H}_v(N) \\
\mathcal{H}_v(\text{app } M N) &= \text{app } \mathcal{H}_v(M) \mathcal{H}_v(N) \\
\mathcal{H}_v(\text{if } M N P) &= \text{cond } \mathcal{H}_v(M) \mathcal{H}_v(N) \mathcal{H}_v(P) \\
\mathcal{H}_v(\text{let val } x = M \text{ in } N) &= \text{letval } (\lambda x. \mathcal{H}_v(N)) \mathcal{H}_v(M) \\
\mathcal{H}_v(\text{let fun } f x = M \text{ in } N) &= \text{letfun } (\lambda f. \mathcal{H}_v(N)) (\lambda f, x. \mathcal{H}_v(M))
\end{align*} \]

Figure 3.4: \( \phi_v : \text{Values}_v \rightarrow \text{vl}, \mathcal{H}_v : \Lambda_v \rightarrow \text{tm} \)

\[ \begin{align*}
\psi_v(x) &= x \\
\psi_v(\text{abs } \lambda x. M) &= \lambda x. \mathcal{L}_v(M) \\
\psi_v(c) &= c \quad c \in \mathbb{Z} \cup \mathbb{B} \cup \{\bullet\}
\end{align*} \]

\[ \begin{align*}
\mathcal{L}_v(\{V\}) &= \psi_v(V) \\
\mathcal{L}_v(f M N) &= f \mathcal{L}_v(M) \mathcal{L}_v(N) \quad f \in \mathbb{O} \\
\mathcal{L}_v(\text{app } M N) &= \mathcal{L}_v(M) \mathcal{L}_v(N) \\
\mathcal{L}_v(\text{if } b \text{ fr } V N P) &= \text{if } \psi_v(V) \mathcal{L}_v(N) \mathcal{L}_v(P) \quad b \in \mathbb{B} \\
\mathcal{L}_v(\text{cond } M N P) &= \text{if } \mathcal{L}_v(M) \mathcal{L}_v(N) \mathcal{L}_v(P) \\
\mathcal{L}_v(\text{let val } R x) &= \text{let val } x = \mathcal{L}_v(N) \text{ in } \mathcal{L}_v(R x) \quad x \text{ fresh} \\
\mathcal{L}_v(\text{let fun } R_1 R_2) &= \text{let fun } f x = \mathcal{L}_v(R_2 f x) \text{ in } \mathcal{L}_v(R_1 f) \quad f, x \text{ fresh}
\end{align*} \]

Figure 3.5: \( \psi_v : \text{vl} \rightarrow \text{Values}_v, \mathcal{L}_v : \text{tm} \rightarrow \Lambda_v \)

specification. I provide the details in figure 3.4. \texttt{iffr} and \texttt{apply} are constants which are not used in the translation \( \mathcal{H}_v \), but arise during evaluation of translated \( \Lambda_v \)-terms in FORUM.

There are some rather subtle issues in the translation in figure 3.4. For instance, on what basis do I choose \texttt{abs } \lambda x. x over \( ((\lambda u, u)(\text{abs } \lambda x. x)) \) as the encoding of \( \lambda x. x \)? The choice comes from the fact that there are unique \( \beta \eta \)-long normal forms in FORUM terms. So I pick the \( \beta \eta \)-long normal form as the encoding of the given \( \Lambda_v \) term. The translations from FORUM to \( \Lambda_v \), and vice versa, in figures 3.5 and 3.4 respectively, are straightforward recursions on the structure of the terms.

I would like to evaluate with the translated programs. Hence, I need to check whether
substitution commutes with the two translations $\mathcal{H}_v$ and $\mathcal{L}_v$. The following lemma states these identities precisely; proofs are deferred to the appendix.

**Lemma 3.1** Let $M \in \Lambda_v$, $V \in \text{Values}_v$, $N$ and $U$ be FORUM terms of types $\text{tm}$ and $\text{vl}$ respectively:

- $\mathcal{H}_v(M[x := V]) = \mathcal{H}_v(M)[x := \phi_v(V)]$.
- $\mathcal{L}_v(N[x := U]) = \mathcal{L}_v(N)[x := \psi_v(U)]$.

Armed with the precise definitions, I am now in a position to define the evaluator and the *correspondence theorem*. Keeping in mind that $\Lambda_v$ is to be evaluated left-to-right, and in fact, considering this as a part of the specification of $\Lambda_v$ operational semantics, I would like to specify the evaluator such that left-to-right evaluation is enforced on implementations complying with my specifications. Hence, a natural choice is to use a continuation-passing-style semantics for $\Lambda_v$ [Plo76, Rey72]. How is my specification then different from the standard continuation-passing-style semantics for $\Lambda_v$ [Plo76, Rey72]? The main point is that just the translation of programs to continuation-passing semantics does not yield the operational semantics. One also needs a strategy for executing the resulting programs. As my specifications are logic programs, search for cut-free proofs corresponds to computation — I get an evaluator for $\Lambda_v$ from very simple clauses.

The evaluator is presented as a set of universally quantified clauses defining the meaning of the non-logical constants I used in translating $\Lambda_v$ into FORUM. (eval $M \; K$) is a two place predicate, the first argument being of type $\text{tm}$ and the second being of type $\text{vl} \rightarrow \text{o}$, the type of continuations, with the intended meaning that the term $M$ is to be evaluated with $K$ being the continuation. The computational paradigm is that I evaluate $M$, and whatever is its result, I pass it to $K$ which then completes the evaluation. The evaluator $E_v$ is defined in figure 3.6. The order of evaluation for the terms in $\Lambda_v$ does not matter as far as one is concerned only with the values produced by the evaluations. In this sense, specifying the exact evaluation order may seem to be an overkill. However, the failure
(eval (V) K) ⊢ (KV)
(eval (app M N) K) ⊢ (eval M λv. (eval N λu. (apply v u K)))
(apply (abs R) U K) ⊢ (eval (RU) K)
(eval (f M N) K) ⊢ (eval M λv. (eval N λu. (K f v u)))  \text{if } f ∈ Ω
(eval (cond M N P) K) ⊢ (eval M λv. (eval (ifb r v N P) K))
(eval (ifb r true M N) K) ⊢ (eval M K)
(eval (ifb r false M N) K) ⊢ (eval N K)
(eval (letval R M) K) ⊢ (eval M λv. (eval (R v) K))
(eval (letfun R₁ R₂) K) ⊢ (eval (R₁ (abs λx. letfun (λf, R₂ f x) R₂)) K)

Figure 3.6: Clauses in $\mathcal{E}_v$ - the evaluator for $\Lambda_v$

to produce a value may be for two reasons. Firstly, the evaluator may get stuck, or the arithmetic operators may be undefined for some values, and secondly, the evaluation may never terminate. Suppose $P$ is a divergent program. The program $(+ P (÷ 5 0))$ will have an infinite evaluation tree under left-to-right evaluation, whereas under right-to-left evaluation it will result in a finite failure caused by a division by 0 error.

Now I have to define when a term evaluates to a value in FORUM. As computation corresponds to search for cut-free proofs, the definition will involve statements about existence of proofs of sequents in FORUM.

**Definition 3.2** [Evaluating $\Lambda_v$ terms in FORUM] $M : \text{tm}$ evaluates to $V : \text{vl}$, written as eval$_v(M, V)$, if

$$\Sigma_v : \mathcal{E}_v ; \quad \forall K : \text{vl} \rightarrow o. (KV) \vdash (\text{eval } M K)$$

is provable in FORUM.

I compare the evaluation of programs using $\mathcal{E}_v$ to the natural semantics evaluator to highlight some key aspects of $\mathcal{E}_v$. Let's look at the computation of $(\text{app } P Q)$ and $(M N)$ where $P$ and $Q$ are $\mathcal{H}_v(M)$ and $\mathcal{H}_v(N)$ respectively.

In figure 3.7 I use the backchain rule. This is essentially a composite rule in which I choose a clause from the evaluator clauses such that a right-hand side formula unifies with the head of
Figure 3.7: One step in the evaluation of \((MN)\)

The clause. The justification of the rule follows with use of \(-L\) followed by the observation that the right sub-proof of \(-L\) will trivially follow from \(initial\), as the head of the clause and right-hand side formulas unified. The key point of the above example is the role of the continuation. The natural semantic proof has three sub-proofs, whereas the \(E_v\) proof is linear — this corresponds to the idea that the evaluation order is completely specified. Furthermore, notice the structure of the continuation \(\lambda v. (eval \ Q \ \lambda u. (apply v u \ K))\) — this term encodes the fact that the value of \(P\), \((abs \ R)\), will be bound to \(v\), and then \(Q\) will be evaluated and its value, \(W\), will be bound to \(u\), and finally \((R \ W)\) will be evaluated. Hence, continuations provide notation within the syntax for the incomplete parts of natural semantics evaluation trees. This capability of representing incomplete proofs within the syntax plays a crucial role in the specification of exceptions and \texttt{callcc}. In fact, it seems that it is problematic to specify \texttt{callcc} in the natural semantics framework because of this deficiency in its syntax.

\(E_v\) and the natural semantics in figure 3.2 are two specifications of \(\Lambda_v\). I prove that the two specifications are identical to the extent that the values computed are identical. The complete proof is deferred to the appendix.

**Theorem 3.3 (Correspondence theorem for \(\Lambda_v\))** For all closed \(\Lambda_v\) terms \(M\) and values \(V\),

\[ M \Downarrow V \text{ if and only if } eval_v(H_v(M), \phi_v(V)) \]

The proof of theorem 3.3 is rather interesting. Going from left-to-right I induct on the evaluation tree for the term, \(i.e.,\) I build computations for larger terms using computations of the subterms. I illustrate the general strategy by showing the case for \(M = (NP)\).
Suppose $(N \; P) \Downarrow V$, then the only way this may happen is by the use of rule for application terms, which implies that

- $N \Downarrow \lambda x. Q$,
- $P \Downarrow U$ and
- $Q[x := U] \Downarrow V$.

The evaluation trees of $N$, $P$ and $Q[x := U]$ are smaller than the evaluation tree of $(N \; P)$. Let $N_1 = \text{def } \mathcal{H}_v(N)$, $P_1 = \text{def } \mathcal{H}_v(P)$, $Q_1 = \text{def } \mathcal{H}_v(Q)$, $U_1 = \text{def } \phi_v(U)$ and $V_1 = \text{def } \phi_v(V)$. By induction hypothesis I get proofs $\delta_1$, $\delta_2$ and $\delta_3$ in FORUM respectively for

- $\Sigma_v, K_1 : \mathcal{E}_v \vdash (K_1 (\text{abs } \lambda x. Q_1)) \to (\text{eval } N_1 K_1)$,
- $\Sigma_v, K_2 : \mathcal{E}_v \vdash (K_2 U_1) \to (\text{eval } P_1 K_2)$ and
- $\Sigma_v, K_3 : \mathcal{E}_v \vdash (K_3 V_1) \to (\text{eval } Q_1[x := U_1] K_3)$.

In the proof $\delta_3$ I use lemma 3.1 to rewrite $\mathcal{H}_v(Q[x := U])$ as $Q_1[x := U_1]$. Using the above proofs I need to construct a proof for the sequent

$$\Sigma_v, K : \mathcal{E}_v \vdash (K V_1) \to (\text{eval } \text{app } N_1 P_1 K).$$

I show below how to construct the required proof. To keep the proof readable, I do not write $\mathcal{E}_v$ in the intuitionistic context and the signature which is $\Sigma_v, K$ in all the sequents shown in the proof. In the proof let $C_1 = \text{def } \lambda u. (\text{eval } P_1 \lambda u. (\text{apply } v u K))$ and $C_2 = \text{def } \lambda u. (\text{apply } (\text{abs } \lambda x. Q_1) u K)$. It is interesting to note that the $\text{CutL}$ rules are needed exactly at those points in the computation when a term passes its values to its continuation. The $\text{CutL}$ passes the value of $N_1$ to $C_1$, the continuation of $N_1$.

$$\frac{(K \; V) \to (\text{eval } P_1 C_2) \quad (C_1 (\text{abs } \lambda x. Q_1)) \to (\text{eval } N_1 C_1) \quad \text{CutL}}{(K \; V) \to (\text{eval } N_1 C_1) \quad (K \; V) \to (\text{eval } \text{app } N_1 P_1 K) \quad \text{backchain}}$$
γ is constructed as shown below. The CutL passes the value of \( P_1 \) to \( C_2 \), the continuation of \( P_1 \).

\[
\begin{align*}
( K V ) \rightarrow ( \text{eval} \ Q_1 \left[ x := U_1 \right] \ K ) \\
( K V ) \rightarrow ( \text{apply} \ ( \text{abs} \ \lambda x. \ Q_1 ) \ U_1 \ K ) \quad \text{backchain} \\
( K V ) \rightarrow ( \text{eval} \ P_1 \ C_2 ) \\
( K V ) \rightarrow ( \text{eval} \ P_1 \ C_2 ) \quad \text{CutL}
\end{align*}
\]

Each \( \sigma_i, i \in [1, 3] \) is constructed from \( \delta_i \) using CutS. Note that

\[
\Sigma_v, K_1, K : \mathcal{E}_v; ( K_1 ( \text{abs} \ \lambda x. \ Q_1 ) ) \rightarrow ( \text{eval} \ N_1 \ K_1 )
\]

is provable by \( \delta_1 \) as \( \Sigma_v, K_1 \subseteq \Sigma_v, K_1, K \). Hence \( \sigma_1 \) can be completed by cutting on \( K_1 \) with \( C_1 \). Similarly, one can build \( \sigma_2 \) and \( \sigma_3 \).

This completes the proof for the case \( M = ( N \ P ) \). The important point is that proofs in FORUM are built using various cut rules rather than analyzing the structure of the proof, as one would expect for the left-to-right direction. Computation in FORUM is represented by search for cut-free proofs. However, since the proofs constructed above have cuts, I am implicitly using the cut-elimination theorem for FORUM. In the other direction, I analyze the proofs in FORUM and construct the natural semantics evaluation trees from FORUM proofs. The proofs are detailed in chapter A.

### 3.2 \( \Lambda_{\text{vs}} \) — State in UML

In this section I specify \( \Lambda_{\text{vs}}, \Lambda_v \) extended with state. \( \Lambda_v \) is a higher-order functional language where values are associated to variables via \( \lambda \) binding. For example, \textbf{let val} \( x = 2 \) in \( M \) associates the value 2 with the variable \( x \) in the term \( M \). The salient property of such variable bindings is that it \textit{cannot} be changed - \( x \) will remain bound to 2 throughout the evaluation of \( M \). Although in principle, one can program only with \( \lambda \) bindings, in practice there are many situations where one would like to \textit{update} the binding of a variable. For example, in \( \Lambda_v \), I cannot write a function, \textbf{inc}, such that it takes a dummy argument and
generates a new number every time it is called. Assignable variables introduce a notion of state in the programming language. The state of the computation is the current binding for the assignable variables. As I can update the binding for the assignable variables, the state may change during the computation.

The syntax for $\Lambda_{v3}$ is formally defined in figure 3.8. The definitions of $M$ and $E$ in figure 3.1 are extended with the clauses in figure 3.8 to obtain $\Lambda_{v3}$ and $EvCont_{v3}$, respectively. The definition for $V$ in figure 3.1 remains unchanged for $Values_{v3}$. However, the non-terminal $M$ ranges over the extended definition. The result of evaluating $(\text{ref } M)$ is a fresh location in the state which is bound to the value of $M$. $(\text{deref } M)$ evaluates $M$ to a location and then returns the contents of the state at that location. $(\text{asg } M \ N)$ assigns the value of $N$ to the location resulting from evaluating $M$, if that location is already defined in the current state. The inc function described above may be implemented in $\Lambda_{v3}$ as:

\[
\text{let val } x = \text{ref } 0 \text{ in } \lambda y. (\lambda z. \text{deref } x) (\text{asg } x (+ (\text{deref } x) 1))
\]

Evaluating $M$ in store $S$ may create a new location, say $l$. What should I do with $l$? One approach would be to extend $S$ with $l$, and let $l$ be in the state at the end of evaluating $M$. Another approach would be to treat $l$ as a variable local to $M$, which must not be in the state at the end of evaluating $M$. While the first approach results in a global view
of state as taken by SML [HMT89], the second approach is adopted by block-structured languages like ALGOL [Rey81a, Rey81b]. State is thought of as a finite function from \textit{Vars} to \textit{Values}. \textit{dom}(S) for a state \( S \) is the set of variables which are bound in the state, \textit{i.e.} the set of assignable variables currently defined. If \( l \in \text{dom}(S) \) then \( S[l \mapsto V] \) denotes the state which maps \( l \) to \( V \). If \( l \not\in \text{dom}(S) \) then \( S[l \mapsto V] \) denotes a new state, say \( S_1 \), such that \( \text{dom}(S_1) \) is the union of \( \text{dom}(S) \) and \( \{l\} \), and \( S_1(l) = V \).

I specify \( \Lambda_v \) in natural semantics in the style of [HMT89]. The specification for the new constructs is in figure 3.9. In the evaluation of \((\text{ref} \ M)\) the side condition \( l \not\in S_1 \) causes the creation of a new location in the state which is bound to \( V \). \texttt{deref} reads the value of the state at a location \( l \); the side condition \( l \in S_1 \) ensures that \( S_1 \) is defined for \( l \). \((\text{asg} \ M \ N)\) redefines the binding of \( S_2 \) at the location \( l \) to be \( V \), only if \( l \) is a variable already defined in \( S_2 \). Note that all currently defined assignable variables are treated as values. The changes to state are cumulative, \textit{e.g.}, in the rule for \((\text{asg} \ M \ N)\), evaluation of \( M \) results in the state \( S_1 \) which is then passed along as the starting state for the computation of \( N \). If I define some new locations in the process of evaluating \( M \), they will be in \( \text{dom}(S_1) \), and hence finally in \( S_2[l \mapsto V] \). Thus the rules specify a global view of state.

Unfortunately, the evaluator for \( \Lambda_v \) is not simply the union of rules in figure 3.9 and figure 3.2. Although no new rules are needed for the constructs of \( \Lambda_v \) due to the addition of state, the existing rules in figure 3.2 need to be modified. \textit{In this sense, the specification of \( \Lambda_v \) is not modularly extended to the specification of \( \Lambda_v \)}. However, the damage is mild.
cell : tm → tm
read : tm → tm
write : tm → tm → tm
get : vl → (vl → o) → o
set : vl → vl → (vl → o) → o

Figure 3.10: Constants for translating $\Lambda_{vs}$ terms

compared to the situation in section 3.3, where extending $\Lambda_v$ with exceptions will create new rules for the constructs in $\Lambda_v$. The modification to the rules in figure 3.2 is obtained by applying the state convention in section 3.3 to states. For example

$$
\begin{array}{c}
M \Downarrow \lambda x. P \\
N \Downarrow U \\
P[x := U] \Downarrow V
\end{array}
\Rightarrow
\begin{array}{c}
M \langle N, S_0 \rangle \Downarrow U \\
N \langle P, S_1 \rangle \Downarrow \langle U, S_2 \rangle \\
P[x := U] \langle V, S_3 \rangle \Downarrow \langle V, S_3 \rangle
\end{array}
$$

is considered to be an abbreviated form of

Thus the natural semantics specification for $\Lambda_{vs}$ is obtained by taking the rules in figure 3.9 and the rules obtained by applying the state convention to the rules in figure 3.2. I will freely use the abbreviated form of the natural semantics rules, because using the state convention I can always recover the full form.

I want to specify state in FORUM not as a finite function, but as some form of concurrent computation. I think of every location in the state as a separate process storing a value, which interacts with its environment only via read and write messages. Interaction on a read message causes the process to transmit the value it stores to the environment. Interaction on a write message causes the process to accept a value from the environment which replaces the value it stores. Adopting the paradigm used in section 2.1 to specify $HO\pi$-calculus in FORUM, I directly specify the above process-style reading of state in FORUM.

The signature for the translation, $\Sigma_{vs}$, is the union of $\Sigma_v$ and the constants in figure 3.10.
\[\text{eval} \ (\text{cell} \ M) \ K \ \vdash \ \text{eval} \ M \ \lambda v. \forall P, l. \text{getC}(P, l) \Rightarrow \text{setC}(P, l) \Rightarrow [(K \ l) \ @(P \ v)]\]
\[\text{eval} \ (\text{read} \ M) \ K \ \vdash \ \text{eval} \ M \ \lambda v. \ (\text{get} \ v \ K)\]
\[\text{eval} \ (\text{write} \ M \ N) \ K \ \vdash \ \text{eval} \ M \ \lambda v. \ (\text{eval} \ N \ \lambda u. \ (\text{set} \ u \ a \ K))\]

where
\[\text{getC}(P, l) = \text{def} \ \forall K, U. [(\text{get} \ l \ K) \ @(P \ U)] \Rightarrow [(K \ U) \ @(P \ U)]\]
\[\text{setC}(P, l) = \text{def} \ \forall K, V, U. [(\text{set} \ l \ V \ K) \ @(P \ U)] \Rightarrow [(K \ \bullet) \ @(P \ V)]\]

Figure 3.11: Specification in FORUM for new constructs in \(\Lambda_{vs}\)

I define translations \(\phi_{vs} : \text{Values}_{vs} \to \text{vl}, \mathcal{H}_{vs} : \Lambda_{vs} \to \text{tm}, \psi_{vs} : \text{vl} \to \text{Values}_{vs}\) and \(\mathcal{L}_{vs} : \text{tm} \to \Lambda_{vs}\) in the appendix. For example, \(\mathcal{H}_{vs}(\text{ref} \ M) = \text{cell} \ \mathcal{H}_{vs}(M)\). The following lemmas regarding the translations and substitution are proved in the appendix.

**Lemma 3.4** Let \(M \in \Lambda_{vs}, V \in \text{Values}_{vs}, N \) and \(U \) be a FORUM terms of types \(\text{tm}\) and \(\text{vl}\) respectively:

- \(\mathcal{H}_{vs}(M[x := V]) = \mathcal{H}_{vs}(M)[x := \phi_{vs}(V)].\)
- \(\mathcal{L}_{vs}(N[x := U]) = \mathcal{L}_{vs}(N)[x := \psi_{vs}(U)].\)

The evaluator for \(\Lambda_{vs}, \mathcal{E}_{vs}\), is the union \(\mathcal{E}_v\) and the clauses in figure 3.11. \((\text{cell} \ M)\) is evaluated by first evaluating \(M\) to a value \(V\). Next, a new predicate \(P\) is created and placed in the environment storing \(V\) as \((P \ V)\). The process identifier, \(l\), for the process \(P\), is passed to the continuation \(K\). Along with the creation of the process, two more clauses are introduced, namely \(\text{getC}\) and \(\text{setC}\). \(\text{getC}\) specifies the handling of read messages passed along the process identifier \(l\), while \(\text{setC}\) specifies the handling of write messages passed along the identifier \(l\). A read message, \((\text{get} \ l \ K)\), synchronizes with the process identified by \(l\), reads its value, and passes the value to the continuation \(K\). A write message, \((\text{set} \ l \ V \ K)\), synchronizes with the process identified by \(l\), sets its value to \(V\), and passes the token \(\bullet\) to the continuation \(K\), indicating the successful completion of the write message. \((\text{read} \ l)\) issues the read message to the process identified by \(l\), and \((\text{write} \ l \ V)\) issues the write message to the process identified by \(l\). I now prove the correspondence theorem between the
two specifications, namely $\mathcal{E}_{\text{vs}}$ and natural semantics, for $\Lambda_{\text{vs}}$. Before I can do this, I have to define when a term evaluates to a value in FORUM and translate state into FORUM.

**Remark 3.5** I define notation that I use in describing evaluation in FORUM. Let $S$ be a state, and $m$ be the number of elements in $\text{dom}(S)$.

- $P_S = \text{def } P_1, \ldots, P_m$.
- $l_S = \text{def } l_1, \ldots, l_m, l_i \in \text{dom}(S), i \in [1, m]$.
- $\Sigma_S = \text{def } \{P_1, \ldots, P_m\} \cup \{l_1, \ldots, l_m\}$.
- $\mathcal{C}L_S = \text{def } \{\text{get}(P_1, l_1), \ldots, \text{get}(P_m, l_m)\} \cup \{\text{set}(P_1, l_1), \ldots, \text{set}(P_m, l_m)\}$.
- $\Gamma_S = \text{def } (P_1 V_1) \bowtie \ldots \bowtie (P_m V_m)$.
- $V_i = \text{def } \phi_{\text{vs}}(S(l_i)), i \in [1, m]$.
- $\text{FV}(S(l_i)) \subset \text{dom}(S), i \in [1, m]$.
- A term $M$ is closed in a state $S$, written as $\text{close}(S, M)$, if $\text{FV}(S(l_i)) \subset \text{dom}(S), i \in [1, m]$ and $\text{FV}(M) \subset \text{dom}(S)$.

**Definition 3.6** [Translating state into FORUM] The translation of state $S$, written as $S^o$, is a FORUM term of type $\alpha \rightarrow \alpha$.

$$S^o = \text{def } \lambda u : \alpha \forall P_S, l_S. \text{get}(P_1, l_1) \Rightarrow \text{set}(P_1, l_1) \Rightarrow \cdots \Rightarrow \text{get}(P_m, l_m) \Rightarrow \text{set}(P_m, l_m) \Rightarrow [u \bowtie \Gamma_S]$$

The application of a term $M$ to $S^o$ would require that the free variables in $M$ be named apart from the bound variables in $S^o$ including $l_S$. I abuse notation because I want the location names to be “captured” by the substitution. Suppose $S$ is $\lambda u. \forall l. \text{get}(P, l) \Rightarrow \ldots \Rightarrow \text{set}(P, l) \Rightarrow [u \bowtie \Gamma_S]$. 

!
\[ \text{setC}(P, l) \Rightarrow \left[ u \triangleright \Gamma_s \right] \text{ and } M \text{ is } \langle l \rangle, \text{ then } (S\langle l \rangle) \Rightarrow \forall l. \text{getC}(P, l) \Rightarrow \text{setC}(P, l) \Rightarrow [\langle l \rangle \triangleright \Gamma_s]. \]

Note that \( l \) is free in \( \langle l \rangle \) but in \( (S\langle l \rangle) \) it gets captured by the universal quantification on \( l \).

This abuse of syntax comes in very handy, and should not be confusing.

**Definition 3.7** [Evaluating \( \Lambda_{vs} \) terms in FORUM] Let \( M \) be a \( \Lambda_{vs} \)-term, and \( S_0 \) be a state such that \( \text{close}(S_0, M) \). \( M \) with \( S_0 \) evaluates to \( V \) with \( S_1 \), written \( \text{eval}_{vs}(M, S_0, V, S_1) \), if

\[
\Sigma_{vs} : E_{vs} ; \rightarrow \forall K : \forall l \rightarrow \alpha. S_1^\circ(K V) \rightarrow S_0^\circ(\text{eval}_{vs}(M) K)
\]

is provable in FORUM. ■

**Theorem 3.8 (Correspondence theorem for \( \Lambda_{vs} \))** Let \( M \) be a \( \Lambda_{vs} \)-term, and \( S_0 \) a state such that \( \text{close}(S, M) \).

\( \langle M, S_0 \rangle \Downarrow \langle V, S_1 \rangle \) if and only if \( \text{eval}_{vs}(M, S_0, V, S_1) \)

The proof of theorem 3.8 is deferred to the appendix. It is along the same lines as the proof for theorem 3.3. Given the fact that \( \Lambda_{vs} \) has an imperative state, it is not a priori clear whether I can use Cut rules to compose proofs - the richer proof-theory of FORUM permits me to use Cut rules essentially because environments are maintained using logical constants.

I illustrate the proof strategy going left-to-right when \( \langle \text{asg} M \ N, S_0 \rangle \Downarrow \langle \bullet, S_2[l \leftarrow V] \rangle \). The last rule in the evaluation tree has to be for \( \text{asg} \), which implies that

- \( \langle M, S_0 \rangle \Downarrow \langle l, S_1 \rangle \),
- \( \langle N, S_1 \rangle \Downarrow \langle V, S_2 \rangle \), and
- \( l \in \text{dom}(S_2) \).

The evaluation trees of \( M \) and \( N \) are smaller than the evaluation tree of \( \langle \text{asg} M \ N \rangle \). Let \( M_1 = \text{def} \ H_{vs}(M), N_1 = \text{def} H_{vs}(N), S_3 = \text{def} S_2[l \leftarrow V] \) and \( V_1 = \text{def} \phi_{vs}(V) \). By induction hypothesis I get proofs \( \delta_1 \) and \( \delta_2 \) in FORUM respectively for
\* \* eval_{vs}(M, S_0, l, S_1), and \\
\* \* eval_{vs}(N, S_1, V, S_2).

Further, note that \( l \in \Sigma_{S_1} \). Using the above proofs, I need to construct a proof for the sequent

\[ \Sigma_v, K, \Sigma_{S_1} : E_{vs} ; S_0^{o}(K \bullet) \rightarrow S_0^{o}(\text{write } M_1 N_1 K) \].

Below, I construct the required proof. To keep the proof readable, I do not write \( E_{vs} \) in the intuitionistic context, and \( \Sigma_{vs}, K \) in the signature part of the sequent, as these parts are present in all the sequents in the part of the proof shown. Furthermore, I do not show the introductions of \( \Sigma_S \) and \( \text{CL}_S \) in the sequents below, because these can be deduced from the context. Let \( C_1 = \text{def} \lambda v. (\text{eval } N_1 \lambda u. (\text{set } v u K)) \) and \( C_2 = \text{def} \lambda u. (\text{set } l u K) \).

In the proof, I first do a bunch of right \( \forall \) introductions followed by a bunch of right \( \Rightarrow \) introductions to introduce \( \Sigma_{S_1} \) and \( \text{CL}_{S_1} \). Next, I backchain on the clause for \text{asg} \) clause, and then I do a \( \text{CutL} \). I am left with the construction of \( \gamma \) and \( \sigma_1 \).

\[
\begin{align*}
S_3^{o}(K \bullet) & \rightarrow \gamma^1 \rightarrow S_1^{o}(\text{eval } N_1 C_2) \rightarrow S_1^{o}(C_1 l) \rightarrow \sigma_1 \rightarrow \text{eval } M_1 C_1) \varphi \Gamma_{S_1} \\
S_3^{o}(K \bullet) & \rightarrow (\text{eval } M_1 C_1) \varphi \Gamma_{S_1} \\
S_3^{o}(K \bullet) & \rightarrow (\text{eval } (\text{write } M_1 N_1 K) \varphi \Gamma_{S_1} \backslash \text{backchain} \\
& \vdash \forall R, \Rightarrow R \\
S_3^{o}(K \bullet) & \rightarrow S_0^{o}(\text{write } M_1 N_1 K)
\end{align*}
\]

The construction of \( \gamma \) is given below. I first do a bunch of right \( \forall \) introductions followed by a bunch of right \( \Rightarrow \) introductions to introduce \( \Sigma_{S_1} \) and \( \text{CL}_{S_1} \). Next, I do a \( \text{CutL} \).

\[
\begin{align*}
S_3^{o}(K \bullet) & \rightarrow \gamma^1 \rightarrow S_2^{o}(\text{set } l V_1 K) \rightarrow S_2^{o}(C_2 V_1) \rightarrow \sigma_2 \rightarrow \text{eval } N_1 C_2) \varphi \Gamma_{S_1} \\
S_3^{o}(K \bullet) & \rightarrow (\text{eval } N_1 C_2) \varphi \Gamma_{S_1} \\
& \vdash \forall R, \Rightarrow R \\
S_3^{o}(K \bullet) & \rightarrow S_1^{o}(\text{eval } N_1 C_2)
\end{align*}
\]

37
Suppose \( \text{dom}(S_2) = \text{def } l_1, \ldots, l_n \). As \( l \in \text{dom}(S_2) \), for some \( i \in [1, n] \), \( l \) is \( l_i \). Construction of \( \gamma_1 \) is shown below. I start the proof with a bunch of right \( \forall \) introductions followed by a bunch of right \( \Rightarrow \) introductions to introduce \( \Sigma_{S_2} \) and \( \text{CL}_{S_2} \). Now I backchain over the \( \text{setC}(P_i, l_i) \) to change the value stored as \( P_i \) to be \( V_i \). I am left with the construction of \( \gamma_2 \).

\[
\begin{align*}
\gamma_2
\frac{S_3^\circ(K \bullet) \rightarrow (K \bullet) \varphi(P_1 V_1) \varphi(P_1 U_1) \cdots \varphi(P_n U_n)}{S_3^\circ(K \bullet) \rightarrow (\text{set } l_i V_i K) \varphi(P_1 U_1) \varphi(P_1 U_1) \cdots \varphi(P_n U_n)} \quad ; \forall R, \Rightarrow R \\
\frac{S_3^\circ(K \bullet) \rightarrow S_2^\circ(\text{set } l_i V_i K)}{}
\end{align*}
\]

The construction of \( \gamma_2 \) is rather interesting. I start a proof with a bunch of left \( \forall \) introductions. The purpose of these is to identify the location names in the final state with the locations created during the evaluation of the term. Next, I perform a bunch of left \( \Rightarrow \) introductions. The purpose of these is to constrain the substitution for the \( P \) variables. Essentially, the \( \text{getC} \) and \( \text{setC} \) clauses for each location are matched off against each other. I then reorganize the memory on the right-hand side to match the order in \( \Gamma_{S_3} \), and use identity.

\[
\begin{align*}
\frac{(K \bullet) \varphi \Gamma_{S_3} \rightarrow (K \bullet) \varphi \Gamma_{S_3}}{}
\frac{; \forall L, \Rightarrow L}{S_3^\circ(K \bullet) \rightarrow (K \bullet) \varphi \Gamma_{S_3}}
\end{align*}
\]

The construction of \( \sigma_1 \) and \( \sigma_2 \) from \( \delta_1 \) and \( \delta_2 \) is straightforward. Note that \( \text{CutL} \) is only used when terms pass values to their continuations. The novelty of this proof lies in the way the cell \( l \) is updated to store \( V_i \). The proof makes crucial use of multiple heads to synchronize between the \( \text{set} \) instruction and \( P_i \), the process identified by \( l_i \). Values can be updated by backchaining, because memory cells are linear objects in FORUM. When a cell synchronizes with a message from the environment, it is consumed, and thus needs to be refreshed. When refreshed, it may be updated as the \( \text{setC} \) clause does.
3.3 $\Lambda_{ve}$ — Exceptions in UML

In this subsection I specify $\Lambda_{ve}$ - $\Lambda_v$ with exceptions. Exceptions are a very important feature of any programming language, and especially indispensable in programs which accept inputs from users or external programs. The openin function of SML is a good example. Given the name of a file, the function opens a stream for reading the data in the file. However, if there does not exist a file with the specified name, then openin is faced with an exceptional situation. At this point openin has two acceptable strategies. The first strategy is that openin returns some value indicating the fact that the file does not exist - the path taken by C. The shortcoming of this strategy is that the programmer must check the value returned by openin to see whether the file was actually opened or not. If the programmer forgets to do so, one may get some obscure error in a possibly unrelated part of the program, and then have to trace the error back to the non-existence of the file. The second strategy is to send a signal to the function which invoked openin. Now, if the programmer does not check for the signal intentionally, it will cause the program to stop, and print an error message saying that the specified file did not exist. However, if not checking the signal was an oversight, then the error will be reported as being caused by the fact the specified file did not exist, and hence is easily detectable.

The SML exception mechanism makes possible the second of the two choices outlined above. There is no restriction on the number of exceptions that one may have in a SML program — exceptions can be created on the fly. Furthermore, exception handlers are scoped, i.e. I can declare a handler for an exception for any given sub-part of my program. These features make the exception mechanism of SML rich and elaborate. I begin by extending the syntax of $\Lambda_v$ for exceptions in figure 3.12. The definitions of $M$, $V$ and $E$ in figure 3.1 are extended with the clauses in figure 3.12 to obtain $\Lambda_{ve}$, $Values_{ve}$ and $EvCont_{ve}$ respectively. Declared exception names are also values. For this, I introduce a new countable syntactic class of ExnNames ranged over by $l$. The result of a computation now may not be a value, e.g. an uncaught exception. I introduce a new syntax class called answers for this purpose, and the natural semantic clause $M \Downarrow A$ will now read as $M$ evaluates to the answer $A$. [pk l V]
\( M := \ldots \)
| exception \( l \) \( M \)
| handle \( M \) \( M \) \( M \)
| raise \( M \) \( M \)

\( V := \ldots \)
| \( l \)

A ::= \( V \)
| [pk \( l \) \( V \)]

\( E ::= \ldots \)
| \( E[\text{raise} \ [\ ] \ M \] \)
| \( E[\text{raise} \ V \ [\ ] \] \)
| \( E[\text{exception} \ x \ [\ ] \] \)
| \( E[\text{handle} \ M \ [\ ] \ M \] \)
| \( E[\text{handle} \ M \ V \ [\ ] \] \)
| \( E[\text{handle} \ [\ ] \ V \ V \] \)

\( (\Lambda_{ve}) \)
\( l \in \text{ExnNames} \)

\( (\text{Values}_{ve}) \)
\( l \in \text{ExnNames} \)

\( (\text{Answers}_{ve}) \)
\( l \in \text{ExnNames} \)

\( (\text{EvCont}_{ve}) \)

Figure 3.12: Syntax for \( \Lambda_{ve} \)

is a called a packet, raising the exception \( l \) with the value \( V \).

(\text{exception} \( l \) \( M \)) binds the exception name \( l \), in the scope of \( M \). It is entirely conceivable to write a program such as \((\text{exception} \( l \) (\text{exception} \( l \) \( N \) \( M \))). This scoping of exceptions creates the need for renaming of exception names, \emph{i.e.} \( \alpha \)-conversion. This problem is handled in [HMT89] by evaluating the exception \( l \) to a new exception name, and carrying this binding around in the environment. (\text{raise} \( l \) \( V \)) indicates that the exceptional circumstance as indicated by \( l \) has occurred, and the function handling this exception should be called with the value \( V \). (\text{handle} \( M \ l \ N \)) declares that during the evaluation of \( M \), if the exception \( l \) is raised and uncaught within \( M \), then \( N \) will be the handling function. Furthermore, if the evaluation of \( M \) to a value is completed without raising \( l \), then \( N \) is removed as the function handling the exception \( l \). This semantics of installing handlers locally provides flexibility during programming.

I illustrate the point with the following example. Let \( x \) be some exception name resulting from the declaration of some exception, and \( P \) some function defined in the environment.
(\text{handle}(\text{if} \ (= \ x \ 0) \\
\hspace{1cm} (\text{handle} \ (\ P \ x) \ \text{xn} \ N_1)) \\
\hspace{1cm} M) \\
\hspace{1cm} (\text{xn} \ N_2)

Now, which handler is used for the exception \text{exn} depends upon the value of \( x \). Furthermore, if \( P \) is called and in \( P \) handlers are installed for \text{xn}, then those handlers take precedence over \( N_1 \). There are some elements of dynamic binding in the mechanism for determining which handler catches a raised exception. It is rather tricky to specify this exception mechanism because of the above mentioned considerations.

The specification of exceptions in [HMT89] is presented in a very slick manner. The specification proceeds in two stages. First, the rules needed for the new constructs are specified as in figure 3.13. The spirit of the rules remains the same as the rule for \( \Lambda_v \). The evaluation of a term is a result of the synthesis of the evaluations of its subterms. However, now I need to keep track of the exception names which have been declared thus far, which means that I need to carry along a state in the evaluation rules. The operational semantics as presented here differs from [HMT89] to the extent that I use substitution instead of maintaining closures.

Clearly, the mere addition of the above rules to \( \mathcal{E}_v \) is not enough to specify \( \Lambda_{ve} \). One problem is that the propagation of exception names has to be handled for all the existing rules for \( \Lambda_v \) terms. For this purpose, the state convention is adopted in [HMT89], which tells us how to restore exception states in a rule which omits them. According to this convention, if a rule is presented as

\[
\frac{M_1 \downarrow V_1 \ldots \ M_n \downarrow V_n}{M \downarrow V}
\]

then its full form is intended to be

41
\[
\langle l, \text{Ex} \rangle \downarrow \langle l, \text{Ex} \rangle \quad l \in \text{ExNames}
\]
\[
\langle M[x := l], \text{Ex}_0 \cup l \rangle \downarrow \langle V, \text{Ex}_1 \rangle \quad l \notin \text{Ex}_0
\]
\[
\text{exception } x M, \text{Ex}_0 \downarrow \langle V, \text{Ex}_1 \rangle
\]
\[
\langle M, \text{Ex}_0 \rangle \downarrow \langle l, \text{Ex}_1 \rangle \quad \langle N, \text{Ex}_1 \rangle \downarrow \langle V, \text{Ex}_2 \rangle
\]
\[
\langle \text{raise } MN, \text{Ex}_0 \rangle \downarrow \langle [\text{pk}\ l\ V], \text{Ex}_2 \rangle
\]
\[
\langle N, \text{Ex}_0 \rangle \downarrow \langle l, \text{Ex}_1 \rangle \quad \langle P, \text{Ex}_1 \rangle \downarrow \langle W, \text{Ex}_2 \rangle \quad \langle M, \text{Ex}_2 \rangle \downarrow \langle V, \text{Ex}_3 \rangle
\]
\[
\langle \text{handle } MN P, \text{Ex}_0 \rangle \downarrow \langle V, \text{Ex}_3 \rangle
\]
\[
\langle N, \text{Ex}_0 \rangle \downarrow \langle l_0, \text{Ex}_1 \rangle \quad \langle P, \text{Ex}_1 \rangle \downarrow \langle W, \text{Ex}_2 \rangle \quad \langle M, \text{Ex}_2 \rangle \downarrow \langle [\text{pk}\ l\ U], \text{Ex}_3 \rangle \quad l_0 \neq l_1, *
\]
\[
\langle \text{handle } MN P, \text{Ex}_0 \rangle \downarrow \langle [\text{pk}\ l_1\ U], \text{Ex}_3 \rangle
\]

Figure 3.13: Natural Semantics specification for the new constructs in \( \Lambda_{\text{w*}} \).

\[
\langle M_1, \text{Ex}_0 \rangle \downarrow \langle V_1, \text{Ex}_1 \rangle \quad \ldots \quad \langle M_n, \text{Ex}_{n-1} \rangle \downarrow \langle V_n, \text{Ex}_n \rangle
\]
\[
\langle M, \text{Ex}_0 \rangle \downarrow \langle V, \text{Ex}_n \rangle
\]

As I can always derive the full-form of a rule using the state convention, I will freely use the abbreviated versions of natural semantic rules from now on. The specification is still not complete. For example, what do I do if the evaluation of \( N \) in \( \text{raise } MN \) raises an exception? In fact, this question comes up in each and every rule specified in figure 3.2. Unfortunately, this leads to adding more clauses for all the rules. The description of the additional rules needed can be done concisely along the lines of [HMT89]. An exception convention defines new natural semantic rules based on the ones in figures 3.2 and 3.13, except for the rules labeled * in figure 3.13. Suppose the form of a rule is:

\[
M_1 \downarrow V_1 \quad \ldots \quad M_n \downarrow V_n
\]

Then for every \( k, 1 \leq k \leq n \), such that \( V_k \) is not a packet, we add another rule of the form
For example, the rule for application will now result in the following three new rules.

\[
\frac{M \downarrow [pk \ l \ V]}{MN \downarrow [pk \ l \ V]}
\]

\[
\frac{M \downarrow \lambda x. P \quad N \downarrow [pk \ l \ V]}{MN \downarrow [pk \ l \ V]}
\]

\[
\frac{M \downarrow \lambda x. P \quad N \downarrow U \quad P[x := U] \downarrow [pk \ l \ V]}{MN \downarrow [pk \ l \ V]}
\]

The natural semantics evaluator for \( \Lambda_{uc} \) is specified by the rules in figure 3.13, 3.2 and the ones created as a result of adopting the exception convention explained above. Thus, the exception convention causes the number of rules in the evaluator to increase from fourteen to thirty-two!

Let me present another way of looking at the operational semantics for exceptions. Suppose I am evaluating the term \( E[\text{handle} \ M \ l \ V] \) where \( (\text{handle} \ M \ l \ V) \) is the redex that I am reducing currently, and \( E \) is the current evaluation context or the part of the program that will take the value of \( (\text{handle} \ M \ l \ V) \) and complete the evaluation of the term, i.e. the current continuation. Further, let us suppose \( l \) and \( l_1 \) are two exception names, and no handler is installed for \( l_1 \). The evaluation of \( M \), if it terminates, will yield

1. a value \( U_1 \), or
2. a packet \([pk \ l \ U_2]\), or
3. a packet \([pk \ l_1 \ U_3]\).
In case 1, the computation continues with $E[U_1]$. In case 2, the computation continues with $E[VU_2]$. But what should happen in case 3? A signal, $l_1$, has been raised for which there is no handler. The only reasonable thing to do is to throw away the current continuation, i.e. $E$, and report to the top level or the next outer handler, that the exception $l_1$ was raised with the value $U_3$. What is clear from the explanation is that exceptions can cause the computation to discard its current continuation up to a handler.

Consider the evaluation of the term $E[\text{handle} ((\text{raise } l U) \ N) \ V]$. The program will evaluate $(\text{handle} ((\text{raise } l U) \ N) \ V)$ causing the handler $V$ to be installed for the exception $l$, and then proceed with the evaluation of $((\text{raise } l U) \ N)$, which in turn will cause it to evaluate $(\text{raise } l U)$. Evaluating $(\text{raise } l U)$ will result in a packet $[\text{pk } l U]$, which can only be handled by $V$. However, notice that the answer computed by $(\text{raise } l U)$ is passed not to the current continuation at that point, but rather to the continuation at the time when the handler $V$ was installed! Thus exceptions can also change the current continuation.

My point is that a natural operational reading of the evaluation mechanism involves the idea of continuations, a concept which has no direct representation in the syntax of natural semantics. The clever presentation in [HMT89] is a way of overcoming this shortcoming of natural semantics. However, one has to pay a price for extending natural semantics to cope with exceptions — the blow up in the number of rules for $\Lambda_v$ terms from eight to twenty. Of particular concern is the fact that the evaluator has to be redefined for the term constructs of $\Lambda_v$, e.g. one has four rules for application now. In this sense, the specification of exceptions in [HMT89] is not modular. An exception mechanism very similar to that of UML has been specified in [WF91] using term rewriting machines. The specification is indeed modular — specifications for the existing term constructs do not change when the language with exceptions is considered. However, the style in [WF91] introduces two new notions of contexts, one used to maintain the scope of exceptions and the other used to match a raised signal with its handler — the basic intuition about manipulation of continuations is not brought out very clearly.

The specification of $\Lambda_{ve}$ in FORUM follows the same pattern as the specification of $\Lambda_v$. I
\[
\begin{align*}
\text{ex} &: \text{ext} \to \text{vl} \\
\text{exn} &: (\text{ext} \to \text{tm}) \to \text{tm} \\
\text{install} &: \text{tm} \to \text{tm} \to \text{tm} \\
\text{signal} &: \text{tm} \to \text{tm} \\
\text{uncaught} &: \text{ext} \to \text{vl} \to o
\end{align*}
\]

Figure 3.14: Constants for translating \( \Lambda_{\text{ve}} \) terms

\[
\begin{align*}
\phi_{\text{ve}}(l) &= \langle \text{ex } l \rangle \\
\mathcal{H}_{\text{ve}}(\text{exception } l \ M) &= \text{exn } l \ \mathcal{H}_{\text{ve}}(M) \\
\mathcal{H}_{\text{ve}}(\text{handle } M \ N \ P) &= \text{install } \mathcal{H}_{\text{ve}}(M) \ \mathcal{H}_{\text{ve}}(N) \ \mathcal{H}_{\text{ve}}(P) \\
\mathcal{H}_{\text{ve}}(\text{raise } M \ N) &= \text{signal } \mathcal{H}_{\text{ve}}(M) \ \mathcal{H}_{\text{ve}}(N)
\end{align*}
\]

Figure 3.15: Translating the new construct of \( \Lambda_{\text{ve}} \) to FORUM

introduce a new type \text{ext}, the type of exception names in FORUM. First, I translate \( \Lambda_{\text{ve}} \) terms into FORUM syntax using \( \Sigma_{\text{ve}} \) augmented with the constants in figure 3.14. For example, \text{install} : \text{tm} \to \text{tm} \to \text{tm} \to \text{tm} is the target of the translation of \text{handle} terms. [pk \ l \ V] is translated using \text{uncaught} : \text{ext} \to \text{vl} \to o. I define translations \( \phi_{\text{ve}} : \text{Values}_{\text{ve}} \to \text{vl}, \mathcal{H}_{\text{ve}} : \text{ve} \to \text{tm}, \psi_{\text{ve}} : \text{vl} \to \text{Values}_{\text{ve}} \) and \( \mathcal{L}_{\text{ve}} : \text{tm} \to \text{ve} \). I show the translation from \( \Lambda_{\text{ve}} \) to FORUM for the new constructs in figure 3.15. The complete translations are deferred to the appendix. The following lemmas regarding the translations and substitutions are proved in the appendix.

\textbf{Lemma 3.9} Let \( M \in \Lambda_{\text{ve}} \), \( V \in \text{Values}_{\text{ve}} \), \( N \) and \( U \) be a FORUM terms of types \text{tm} and \text{vl} respectively:

\[
\begin{align*}
\cdot \mathcal{H}_{\text{ve}}(M[x := V]) &= \mathcal{H}_{\text{ve}}(M)[x := \phi_{\text{ve}}(V)]. \\
\cdot \mathcal{L}_{\text{ve}}(N[x := U]) &= \mathcal{L}_{\text{ve}}(N)[x := \psi_{\text{ve}}(U)].
\end{align*}
\]

Instead of specifying the semantics of catching raised exceptions in the style of [HMT89], I take the route of using the continuations explicitly. Consequently, I have to manage
explicitly the installation and removal of handlers, and the matching of raised exceptions with handlers. I find this description to be more enlightening because it highlights the manipulation of continuations by exceptions, and explains the maintenance of exception handlers separately. I represent the exceptions via a predicate, \texttt{exnst}, which takes a list of exception handlers as an argument. An exception handler is a term, \((\texttt{pkt} \, l \, V \, K)\), where \(l\) is the exception name, \(V\) is the function installed as the handler, and \(K\) is the continuation which will be invoked if this handler is chosen. To maintain \texttt{exnst} I need predicates \texttt{push} and \texttt{pop} to add and remove handlers from the list of exception handlers. Further, I need a predicate \texttt{lookup} to search the list of handlers.

\((\texttt{lookup} \, l \, V)\) searches the list \(X\) in \((\texttt{exnst} \, X)\) for the first packet whose exception name is \(l\). The search program for \texttt{lookup} would therefore need to check whether two exception names are equal or not, \(i.e.,\) I need an inequality predicate for exception names. Specifying inequality in the presence of \(\forall\) is a rather delicate matter. Suppose I am able to prove \(\forall x, y. \; x \neq y\), how can I then use it? Obviously, instantiating the proof for \(x = c\) and \(y = c\), where \(c\) is a constant, will lead to inconsistencies. [HSH90, SH93, Gir92] analyze such situations. However, the results are not very conclusive as yet.

The solution I adopt is to take the type \texttt{ext} to be \texttt{nat} of section 2.1, and as shown in section 2.1, inequality between numbers can be specified in FORUM. I generate new exception names using the paradigm of \texttt{inc} in section 2.1. The signature for the various constants is given in figure 3.16. In the translation in figure 3.15, \(l\) is translated to \((\texttt{ex} \, l)\). Given the new interpretation of \texttt{ext}, I have to redefine the translation. As \texttt{ExnNames} is a countable set, there is an isomorphism between \texttt{ExnNames} and \texttt{ext}. Hence, by abuse of notation I let \(l\) denote a term of type \texttt{ext}, which denotes the number to which \(l\) is mapped by the isomorphism.

The signature for the additional constants is given in figure 3.16. The signature for specifying \(\Lambda_{ve}, \Sigma_{ve}\), is defined as the set consisting of the elements of \(\Sigma_{ve}\), the constants in figure 3.16, and the constants in figure 3.14.
\begin{verbatim}
 pkt : ext \rightarrow vl \rightarrow (vl \rightarrow o) \rightarrow packet
 push : ext \rightarrow vl \rightarrow (vl \rightarrow o) \rightarrow tm \rightarrow o
 pop : o \rightarrow o
 nil : packet list
 :: : packet \rightarrow packet list \rightarrow packet list
 exnst : packet list \rightarrow o
 isexn : vl \rightarrow (ext \rightarrow o) \rightarrow o
 z : ext
 s : ext \rightarrow ext
 neq : ext \rightarrow ext \rightarrow o
 sigctr : ext \rightarrow o
\end{verbatim}

Figure 3.16: Constants for exception management in FORUM

I consider \texttt{handle} as an instruction to store an exception handler with the current continuation, and \texttt{raise} as an instruction to search for the most recently installed handler for the raised exception. \texttt{(handle M l V)} will cause \texttt{V} to be \texttt{pushed} on the stack of handlers as the handler for \texttt{l} before the execution of \texttt{M} begins, and if the execution of \texttt{M} terminates without raising \texttt{l} to the handler \texttt{V}, then \texttt{V} is removed from the stack of exception handlers. \texttt{(raise l V)} causes the system to look in the exception stack from top to bottom for a handler for \texttt{l}. If a handler is found for \texttt{l}, then control is passed to the handler and its continuation, or else the exception returns to the top-level as an uncaught exception. The specification for the new constructs along with the specification for the auxiliary non-logical constants is presented in figure 3.17. The evaluator, \( E_v \), is defined to be the universal closure of the clauses in figure 3.17, the clauses in \( E_v \).

The clauses in figure 3.17 highlight various aspects of the exception mechanism. The fact that new exceptions can be created on the fly is implicit in the usage of the \texttt{sigctr} to generate a new constants for exception names. The way continuations are handled by exceptions is made explicit by \texttt{handle} storing the current continuation along with the handler in the exception stack. The search for a matching exception explains how exceptions can cause a program to discard its current continuation and reinstate the continuation stored with the handler. The issue of locally installing and removing handlers is a separate concern, managed by the stack like maintenance of \texttt{exnst} via \texttt{push} and \texttt{pop}. The clause for \texttt{lookup}
Figure 3.17: Specification in FORUM for new constructs in $\Lambda_v$.

uses a new constant $\otimes$. Although the use of $\otimes$ makes the intention clear, it is not essential because in FORUM

$$:: A \rightarrow B \rightarrow C \rightarrow (A \otimes B) \rightarrow C$$

is provable.

The key point is that neither do I introduce any new clauses for the constructs in $\Lambda_v$, nor do I modify the existing clauses in $\Lambda_v$. The specification of the exception mechanism is a modular extension to the specification for $\Lambda_v$. The fact that I can specify the stack in the environment using a logical connective, i.e. $\varphi$, is rather important. If I attempt to specify $\Lambda_{ve}$ using non-logical constants in the spirit of abstract machines, then I will not have to add new clauses for $\Lambda_v$, but I will have to redefine the clauses for $\Lambda_v$ to explicitly carry along the exception stack. However, because of the proof rules for $\varphi$, I am able to keep the stack passively in the environment, and interact with it only when I have to deal with exceptions. The rich proof theory of FORUM provides me the tools to specify the maintenance of the environment in a logical fashion. I now state the correspondence theorem.
between the two specifications, namely $\mathcal{E}_{ue}$ and [HMT89]. If $Ex$ is a set of exception names, let $l_{Ex}$ be one plus the largest number to which any element of $Ex$ is mapped. Further, $\mathcal{G}(l, P) = \text{def } P \varphi(\text{exnst nil}) \varphi(\text{sigctr } l)$.

**Definition 3.10** [Translating Answers$_{ue}$ to FORUM] $\mathcal{A}_{ue}$ translates terms in Answers$_{ue}$ to terms in FORUM of type $o$. Let $K$ be a constant of type $vl \rightarrow o$ in FORUM.

\[
\mathcal{A}_{ue}(V, K) = (K \varphi_{ue}(V)) \\
\mathcal{A}_{ue}([pk l V], K) = (\text{uncaught } l \varphi_{ue}(V))
\]

\[\]

**Definition 3.11** [Evaluating $\Lambda_{ue}$ terms in FORUM] Let $M$ be a $\Lambda_{ue}$-term, $A \in$ Answers$_{ue}$, and $Ex_0$ be an exception state such that all the free variables of $M$ are in $Ex_0$. $M$ in exception state $Ex_0$ evaluates to $A$ with the exception state $Ex_1$, written as $\text{eval}_{ue}(\mathcal{H}_{ue}(M), l_{Ex_0}, \mathcal{A}_{ue}(A, K), l_{Ex_1})$, if

\[
\Sigma_{ue} : \mathcal{E}_{ue} ; \quad \forall K : vl \rightarrow o. \mathcal{G}(l_{Ex_1}, \mathcal{A}_{ue}(A, K)) \rightarrow o \mathcal{G}(l_{Ex_0}, (\text{eval } \mathcal{H}_{ue}(M) K))
\]

is provable in FORUM.\[\]

**Theorem 3.12** (Correspondence theorem for $\Lambda_{ue}$) Let $M$ be a $\Lambda_{ue}$ term, $Ex_0$ be the exception state consisting of all the exception names in $M$, $A \in$ Answers$_{ue}$, and $Ex_1$ be the exception state consisting of all the exception names in $A$. All the free variables of $M$ are in $Ex_0$.

\[
\langle M, Ex_0 \rangle \Downarrow \langle A, Ex_1 \rangle \quad \text{if and only if } \text{eval}_{ue}(\mathcal{H}_{ue}(M), l_{Ex_0}, \mathcal{A}_{ue}(A, K), l_{Ex_1})
\]

The proof of theorem 3.12 is deferred to the appendix. The proof is along the same lines as the proof for theorem 3.3. The logical maintainence of environments permits me to use cuts to compose proofs. I illustrate the proof strategy going left to right when $\langle N P, E_0 \rangle \Downarrow \langle [pk l V], E_3 \rangle$. One way this may happen is that

49
\[\{N, E_0\} \downarrow \{\lambda x. Q, E_1\},\]
\[\{P, E_1\} \downarrow \{U, E_2\}\] and
\[\{Q[x:=U], E_2\} \downarrow \{[pk l V], E_3\}.\]

The evaluation trees of \(N, P\) and \(Q[x:=U]\) are smaller than the evaluation tree of \((N P)\). Let \(N_1 = \text{def } \mathcal{H}_{\nu e}(N), P_1 = \text{def } \mathcal{H}_{\nu e}(P), Q_1 = \text{def } \mathcal{H}_{\nu e}(Q), U_1 = \text{def } \phi_{\nu e}(U), V_1 = \text{def } \phi_{\nu e}(V)\) and \(A = \text{def } (\text{uncaught } l V)\). By induction hypothesis I get proofs \(\delta_1, \delta_2\) and \(\delta_3\) in FORUM respectively for

- \(\Sigma_\nu, K_1 : \mathcal{E}_\nu; \mathcal{G}(l_{Ex_1}, (K_1 (\text{abs } \lambda x. Q_1))) \rightarrow \mathcal{G}(l_{Ex_0}, (\text{eval } N_1 K_1))\),
- \(\Sigma_\nu, K_2 : \mathcal{E}_\nu; \mathcal{G}(l_{Ex_2}, (K_2 U_1)) \rightarrow \mathcal{G}(l_{Ex_1}, (\text{eval } P_1 K_2))\), and
- \(\Sigma_\nu, K_3 : \mathcal{E}_\nu; \mathcal{G}(l_{Ex_3}, A) \rightarrow \mathcal{G}(l_{Ex_2}, (\text{eval } Q_1[x:=U_1] K_3))\).

In the proof \(\delta_3\) I use lemma 3.9 to rewrite \(\mathcal{H}_\nu(Q[x:=U])\) as \(Q_1[x:=U_1]\). Using the above proofs I need to construct a proof for the sequent

\[\Sigma_\nu, K : \mathcal{E}_\nu; \mathcal{G}(l_{Ex_0}, A) \rightarrow \mathcal{G}(l_{Ex_0}, (\text{eval } (\text{app } N_1 P_1 K)))\].

I construct the required proof below. To keep the proof readable, I do not write \(\mathcal{E}_{\nu e}\) in the intuitionistic context, and the signature which is \(\Sigma_\nu, K\). In the proof let \(C_1 = \text{def } \lambda \nu. (\text{eval } P_1 \lambda u. (\text{apply } v u K))\), and \(C_2 = \text{def } \lambda u. (\text{apply } (\lambda x. Q_1) u K)\). Note that \((C_1 (\text{abs } \lambda x. Q_1))\) is \(\beta\) equivalent to \((\text{eval } P_1 C_2)\), and that \((C_2 U_1)\) is \(\beta\) equivalent to \((\text{apply } (\text{abs } \lambda x. Q_1) U_1 K)\). The last proof rule is \(\text{CutL}\) in the construction below.

\[
\frac{
\gamma \quad \sigma_1
}{
g(l_{Ex_0}, A) \rightarrow g(l_{Ex_1}, (\text{eval } C_1)) \quad g(l_{Ex_1}, (\text{eval } C_2)) \quad g(l_{Ex_0}, (\text{eval } N_1 C_1))
\rightarrow \frac{g(l_{Ex_2}, (\text{eval } P_1 K))}{\begin{array}{c}
\text{backchain}
\end{array}}
\}
\]

To complete the proof I have to construct \(\gamma\) and \(\sigma_1\). \(\sigma_1\) is constructed from \(\delta_1\), by using \(\text{CutS}\) on \(K_1\) with \(C_1\). I show the construction of \(\gamma\). I use a \(\text{CutL}\) followed by a \(\text{backchain}\) on the left sub-proof of \(\text{CutL}\).
The construction of $\sigma_2$ and $\sigma_3$ is along the lines of $\sigma_1$, using $\delta_2$ and $\delta_3$, respectively. The surprising fact is that the structure of the proof construction given above, and the construction of $\sigma_i, i \in [1, 3]$ is identical to the construction given in the example for theorem 3.3. The fact that I can use Cut rules to compose proofs and cut-elimination to obtain computations in the presence of exceptions is a convincing argument for the claim that $E_{vc}$ is a declarative specification of $\Lambda_{vc}$.

### 3.4 $\Lambda_{vc} —$ Continuations in UML

In this subsection I specify $\Lambda_{vc} - \Lambda_v$ with continuations. A *continuation* is that part of a program which must be evaluated to obtain the final value of the computation. For example, consider the evaluation of $(+ (+ 3 2) 4)$. When I am evaluating $(+ 3 2)$, the remaining part of the computation is $(+ \bullet 4)$. The idea being that the value of $(+ 3 2)$ will replace $\bullet$, and then computation will start once again, or, in other words, $\lambda x.(+ x 4)$ is the continuation.

Informally, one may think of a continuation as a function which takes the value of the term being evaluated currently as its argument, and then continues the computation. If I can represent the continuation of a program in the program itself, then I can change the control flow of the program during its evaluation. Thus, continuations provide a functional means of representing the control flow of the program.

Continuations have been used widely in the semantics, compilation and design of programming languages [Gor79, Sto77, SW74, AJ89, App92, Ste78, RC86, SF92]. Many programming languages provide language constructs which allow the programmer to take control of the continuation of the program during the computation and manipulate it, e.g. callcc in SML of New Jersey, and call-with-current-continuation in scheme. Programming with continuations introduces a rich programming style. [Fri88, DH89, Wan80] implemented
coroutines and process schedulers using scheme-style call-with-current-continuation. [Rep91] implemented CML, a concurrent higher order functional language with concurrency primitives, using \texttt{callcc} in SML of New Jersey.

Continuations are not a part of the core SML [HMT89]. The reasons for not providing language support for manipulating continuations are not presented in [HMT89]. However, given the obvious utility of continuations as a programming paradigm, I consider first-class continuations as an important aspect of a programming language. Moreover, SML of New Jersey does include \texttt{callcc}, a construct that allows the programmer to capture the current continuation. We saw in section 3.1 that exceptions can be specified in natural semantics, and that continuations are manipulated when exceptions are raised. Hence, some level of continuation manipulation can be specified in the natural semantics framework. Thus encouraged, I attempted to specify \texttt{callcc} in the natural semantics framework outlined in [HMT89]. The operational semantics of \texttt{callcc} is informally specified as | first evaluate $M$ to $V$, and then apply $V$ to whatever the current continuation is. Hence, to specify \texttt{callcc} I need to formalize the notion of current continuation within the syntax of the language. Without changing the essential nature of the natural semantics framework it seemed impossible to specify \texttt{callcc}.

In natural semantics, the continuation at any point of the computation is not a part of the syntax. For example, let us look at the evaluation for $(M \ N)$ in figure 3.2.

$$
\frac{M \Downarrow \lambda x. P \quad N \Downarrow U \quad P[x := U] \Downarrow V}{M \ N \Downarrow V}
$$

The continuation of $M$ is a program which will take $\lambda x. P$, compute $N$ to $U$ and then compute $P[x := U]$ to $V$. This information is encoded in the above rule by the three separate hypotheses of the rule, and demanding that the hypotheses be read from left to right. If I want to capture the current continuation in the program, then I must be able to represent the continuation in the evaluation rule — it is not clear how to do this in natural semantics. Exceptions also manipulate continuations — how can exceptions be specified in natural semantics? Exceptions can be specified in natural semantics because of two
key reasons. Firstly, exceptions do not allow the programmer to capture the continuation as a part of the program. Secondly, exceptions do not allow the programmer to use a continuation more than once. Although the path taken in [CF94] is a beginning towards using a concurrent meta-theory with side-effects, it is however restricted by the fact that there are essentially two threads, the program evaluator and the administrator.

The syntax for $\Lambda_{vc}$ is presented in figure 3.18. The definitions of $M$ and $E$ in figure 3.1 are extended with the clauses in figure 3.18 to obtain $\Lambda_{vc}$ and $EvCont_{vc}$ respectively. The definition for $V$ in figure 3.1 remains unchanged for $Values_{vc}$. However, the non-terminal $M$ ranges over the extended definition. I think of (callcc $V$) as creating a process out of the current continuation, called the continuation process — the process identifier for the created process is passed to $V$. Unlike processes created in concurrent languages, the continuation process does not start computing on its own, in fact it lies dormant. The process identified by $l$ is invoked by (throw $lU$). As $\Lambda_{vc}$ is a sequential language, only one process should be active at any given time. Hence, the continuation process is dormant when created, and when a continuation process is invoked using throw, the evaluation thread terminates and the continuation process becomes the evaluation thread. I find this view of callcc to be very useful, especially in light of the fact that one of the most significant uses of callcc has been to implement CML in SML of New Jersey [Rep91]. Adopting the paradigm used in section 2.1 to specify the $HO\pi$-calculus in FORUM, I directly specify the above process-style reading of callcc in FORUM.
catch : tm \rightarrow \text{tm}
jump : \text{tm} \rightarrow \text{tm} \rightarrow \text{tm}
resume : \text{vl} \rightarrow \text{vl} \rightarrow o
cont : \text{vl} \rightarrow (\text{vl} \rightarrow o) \rightarrow o

Figure 3.19: Constants for translating \(\Lambda_{vc}\) terms

\[
\begin{align*}
(\text{eval} (\text{catch} \ M) \ K) & \leftarrow (\text{eval} \ M \ \lambda v. \ \forall l. \ \text{contC}(l, K) \Rightarrow [(\text{apply} \ v \ l \ K) \ \varphi(\text{cont} \ l \ K)]) \\
(\text{eval} (\text{jump} \ M \ N) \ K) & \leftarrow (\text{eval} \ M \ \lambda v. (\text{eval} \ N \ \lambda u. (\text{resume} \ v \ u)))
\end{align*}
\]

where \(\text{contC}(l, K) = \text{def} \ \forall U. [(\text{resume} \ l \ U) \ \varphi(\text{cont} \ l \ K)] \leftarrow [(KU) \ \varphi(\text{cont} \ l \ K)]\)

Figure 3.20: Specification in FORUM for new constructs in \(\Lambda_{vc}\)

The signature for the specification, \(\Sigma_{vc}\), is defined to be \(\Sigma_v\) extended with the constants in figure 3.19, \textit{e.g.} \text{callcc} is translated to \text{catch}. I define translations \(\phi_{vc} : \text{Values}_{vc} \rightarrow \text{vl}\), \(\mathcal{H}_{vc} : \Lambda_{vc} \rightarrow \text{tm}\), \(\psi_{vc} : \text{vl} \rightarrow \text{Values}_{vc}\) and \(\mathcal{L}_{vc} : \text{tm} \rightarrow \Lambda_{vc}\) in the appendix. For example, \(\mathcal{H}_{vc}(\text{callcc} \ M) = \text{catch} \ \mathcal{H}_{vc}(M)\). The following lemmas regarding the translations and substitutions are proved in the appendix.

**Lemma 3.13** Let \(M \in \Lambda_{vc}\), \(V \in \text{Values}_{vc}\), \(N \) and \(U \) be a FORUM terms of types \text{tm} and \text{vl} respectively:

- \(\mathcal{H}_{vc}(M[x := V]) = \mathcal{H}_{vc}(M)[x := \phi_{vc}(V)].\)
- \(\mathcal{L}_{vc}(N[x := U]) = \mathcal{L}_{vc}(N)[x := \psi_{vc}(U)].\)

The evaluator \(\mathcal{E}_{vc}\) for \(\Lambda_{vc}\) is the universal closure of the clauses in figure 3.20 along with the clauses in \(\mathcal{E}_v\). Below, I outline the evaluation of \((\text{eval} \ (\text{catch} \ M) \ K)\) to highlight the novel features of this specification.

- To evaluate \((\text{eval} \ (\text{catch} \ M) \ K)\), I first evaluate \(M\) to a value \(V\), which is passed to the continuation \(\lambda v. \ \forall l. \ \text{contC}(l, K) \Rightarrow [(\text{apply} \ v \ l \ K) \ \varphi(\text{cont} \ l \ K)].\)
- Evaluation of $\forall l. \text{contC}(l, K) \Rightarrow [(\text{apply } V \; l \; K) \varphi (\text{cont } l \; K)]$ creates a new name $l : v$ and makes $\text{contC}(l, K)$ a part of the evaluator.

- Finally, I proceed with the evaluation of $(\text{apply } V \; l \; K) \varphi (\text{cont } l \; K)$.

Placing $(\text{cont } l \; K)$ in the environment of $(\text{apply } V \; l \; K)$ is the creation of the continuation process as a dormant entity — $l$ is the process identifier for the continuation process $(\text{cont } l \; K)$. Every time I create a continuation process, I introduce the $\text{contC}(l, K)$ clause which associates the process identifier $l$ to $(\text{cont } l \; K)$, the continuation processes. When $(\text{resume } l \; U)$ synchronizes with $(\text{cont } l \; K)$, it results in the evaluation of $(K \; V)$ with $(\text{cont } l \; K)$ in the environment. $(\text{jump } M \; N)$ terminates the current evaluation thread and invokes one of the dormant continuation processes. The main novelty lies in the way I can create new entities in the environment, e.g. $(\text{cont } l \; K)$. The created processes lie in the environment dormant, and can be activated by passing control explicitly to them, e.g. the $\text{contC}(l, K)$ clause. This specification provides a view of $\text{catch}$ and $\text{jump}$ as a restricted form of concurrency where it is possible to have more than one process. However, only one of the processes can compute at a given time. Given this basic understanding, it is natural that $\text{callcc}$ in SML of New Jersey was the basis of the CML implementation in SML of New Jersey [Rep91] and $\text{call-with-current-continuation}$ was used to implement co-routines and process schedulers [Fri88, DH89, Wan80].

I would like to define when a term $M$ evaluates to a value $V$. When computation starts, there may be continuation processes defined in the environment in which $M$ is to be evaluated. As the computation of $M$ proceeds, new continuation processes may be created. Hence, the value $V$ will have a continuation state, i.e. the continuation processes associated with it. In this sense continuations behave like a mutable store. Hence, to evaluate $\Lambda_{vc}$ terms one has to carry the continuation processes and the corresponding process identifiers created thus far, just as in evaluating programs with assignable variables one must carry along the mutable store. A continuation state, written as $\Gamma_c$, possibly subscripted, is defined to be a finite function from variables to $\text{EvCont}_{vc}$. $\text{dom}(\Gamma_c)$ denotes the domain of the function $\Gamma_c$. Let $l_{\Gamma_c}$ denote a list of the variables in $\text{dom}(\Gamma_c)$ in any order. A key difference between
the continuation state and the mutable store is that the process identifiers in \( \Gamma c \) cannot be reassigned. Consequently, \( \Gamma c \) is cycle-free, i.e., there is an ordering on \( l_{1 \Gamma c}, \ldots, l_{m \Gamma c} \) where \( \text{dom}(\Gamma c) \) has \( m \) elements, such that for all \( l \in FV(\Gamma c(l_{i j})), l < l_{i j}, j \in [1, m] \). In particular, \( l \notin FV(\Gamma c(l_{ij})) \), for all \( l \in \text{dom}(\Gamma c) \) and \( FV(\Gamma c(l_{i j})) = \emptyset \). By a continuation state I will henceforth mean a cycle-free continuation state. \( \mathcal{C}_{vc} \) in figure 3.21 translates \( E \in EvCont_{vc} \) to the terms in FORUM of type \( \text{vt} \rightarrow o \).

**Definition 3.14** [Translating continuation state into FORUM] The translation of \( \Gamma c \), written as \( \Gamma c^o \), is a FORUM term of type \( o \rightarrow o \).

\[
\Gamma c^o = \text{def } \lambda u : o. \forall l_{\Gamma c} \text{ contC}(l_1, K_1) \Rightarrow \ldots \Rightarrow \text{contC}(l_m, K_m) \Rightarrow [u \varphi (\text{cont } l \ K_1) \varphi \ldots \varphi (\text{cont } l_m \ K_m)]
\]

Where \( m \) is the number of elements of \( \text{dom}(\Gamma c) \) and \( K_i = \text{def } \mathcal{C}_{vc}(\Gamma c(l_i), K), i \in [1, m] \).

Note the substitution of a term \( M \) for \( u \) in \( \Gamma c^o \) is treated like the substitution a term in in the translation of a state. In particular, the free variables \( l_{\Gamma c} \) in \( M \) will get “captured” by the substitution.
(eval (catch M) K) \(\varphi\) (exnst X) \rightarrow
(eval M \lambda v. \forall P, l, \text{contC}(l, K, X) \Rightarrow \left[\left(\text{apply } v l K\right) \varphi \left(\text{cont } l K X\right) \varphi \left(\text{exnst } X\right)\right])

where, \text{contC}(l, K, X) = \text{def} \\forall U, Y. \left[\left(\text{resume } l U\right) \varphi \left(\text{cont } l K X\right) \varphi \left(\text{exnst } Y\right)\right] \rightarrow \left[\left(\text{K } U\right) \varphi \left(\text{cont } l K X\right) \varphi \left(\text{exnst } X\right)\right]

Figure 3.22: Specification for callcc in the presence of exceptions

**Definition 3.15** [Evaluating \(\Lambda_{vc}\) terms in FORUM] Let \(M\) be a \(\Lambda_{vc}\)-term, and \(\Gamma_1\) be a continuation state such that all free variables of \(M\) are in \(\text{dom}(\Gamma_1)\). \(M\) with the continuation state \(\Gamma_1\) evaluates to \(V\) with the continuation state \(\Gamma_2\), written \(\text{eval}_{vc}(M, \Gamma_1, V, \Gamma_2)\) if

\[
\Sigma_{vc} : \mathcal{E}_{vc}; \rightarrow \forall K : \text{vl} \rightarrow o. \Gamma_2^o(K V) \rightarrow o \Gamma_1^o(\text{eval } \mathcal{H}_{vc}(M) K)
\]

is provable in FORUM. ■

A further issue arises when callcc and throw need to be specified in the presence of exceptions, i.e., I am extending \(\Lambda_{vc}\) with first-class continuations. The question is regarding the status of exception handlers when I throw to a continuation process. In SML of New Jersey, the exception handlers are stored along with the continuation. Consequently, if I throw to a continuation, then along with restoring the saved continuation as the current continuation, I install the saved exception handlers as the currently installed exception handlers. Since the definition of the continuation is extended to include exception handlers, the clause for catch is changed to treat exceptions. The achievement here is that only the specifications for catch and contC need to be changed; everything else remains the same. The new specifications for catch and contC are given in figure 3.22. Note that the type of cont is \(\text{vl} \rightarrow (\text{vl} \rightarrow o) \rightarrow \text{packet list} \rightarrow o\). The integration of first-class continuations will be done in detail later, when I put together the specification for UML using all the pieces. As [HMT89] does not specify callcc and throw, I cannot prove equivalence between the two semantics for \(\Lambda_{vc}\).
3.5 UML — Putting it together

In this section I provide the specification for UML — $\Lambda_v$ with state, exceptions and continuations. The syntax for UML is obtained by putting together the syntax in figures 3.1, 3.8, 3.12, and 3.18. A natural semantic specification for $\Lambda_v$ with state and exceptions is obtained in two steps. First, take the rules in figures 3.2, 3.9, and 3.13. Now, apply the state and exception convention to the rules thus obtained. The resulting set of rules is the natural semantic specification for $\Lambda_v$ with state and exceptions. As discussed in section 3.4, I do not have a natural semantics specification of first-class continuations.

In FORUM, it is possible to specify UML as evidenced by the specifications of its parts in sections 3.1, 3.2, 3.3 and 3.4. The signature of the translation, $\Sigma_{ml}$, is the union of $\Sigma_v$, $\Sigma_v$, $\Sigma_{vc}$ and $\Sigma_{vc}$. The translations from UML to FORUM and back, $\phi : \text{Values} \rightarrow \nu l$, $H : \Lambda \rightarrow \text{tm}$, $A : \text{Answers} \rightarrow (\nu l \rightarrow o) \rightarrow o$, $C : \text{EvCont} \rightarrow (\nu l \rightarrow o)$, $\psi : \nu l \rightarrow \text{Values}$, and $L : \text{tm} \rightarrow \Lambda$, are obtained by putting the translations for the fragments together.

The definitions are deferred to the appendix. The evaluator for UML, $E$ is defined as the universal closure of the clauses in figures 3.6, 3.11, 3.17, and 3.20 — I put together the different modules specifying different pieces of FORUM. A configuration is a triplet of continuation state, state and the exception names with respect to which a UML term is evaluated. I now define the translation of a configuration in FORUM, and the evaluation for UML terms in FORUM.

**Definition 3.16** [Translating configuration to FORUM] Let $C =_{def} (\Gamma c, S, Ex)$ be a configuration, where $\Gamma c$ is a continuation state, $S$ is a state and $Ex$ a set of exception names. The translation of $C$, written as $C^o$, is a FORUM term of type $o \rightarrow o$.

\[
C^o =_{def} \lambda u : o.\forall P, l, p_{tr}. \text{getC}(P, l_1) \Rightarrow \text{setC}(P, l_1) \Rightarrow \ldots \Rightarrow \text{getC}(P_n, l_n) \Rightarrow \text{setC}(P_n, l_n) \Rightarrow \\
\text{contC}(p_1, K_1) \Rightarrow \ldots \Rightarrow \text{contC}(p_m, K_m) \Rightarrow \\
[u \circ \Gamma S \circ (\text{cont } p_1 \ K_1) \circ \ldots \circ (\text{cont } p_m \ K_m) \circ (\text{sigctr } l_{\bar{x}}) \circ (\text{exnst } \text{nil})]
\]
Where

- \( n \) is the number of elements in \( \text{dom}(\Gamma e) \), and \( m \) is the number of elements in \( \text{dom}(S) \).

- \( K_i = \text{def} \ C(\Gamma e(p_i), K) \), \( i \in [1, n] \).

The domain of a configuration \( C = \text{def} \ (\Gamma c, S, Ex) \), written as \( \text{dom}(C) \), is the union of \( \text{dom}(\Gamma c) \), \( \text{dom}(S) \) and \( Ex \). These are the variables which have been declared to be values by the configuration \( C \). Note that analogous to the case of \( S^o \) and \( \Gamma e^o \), applying a term \( M \) to \( C^o \) causes the free occurrences of \( l_S \) and \( p_{tr} \) in \( M \), if any, to get captured.

**Definition 3.17** [Evaluating UML terms in FORUM] Let \( M \) be a UML term, \( A \in \text{Answers}_{uc} \), and \( C_1 \) a configuration such that \( FV(M) \subseteq \text{dom}(C_1) \). \( M \) in the configuration \( C_1 \) evaluates to \( A \) in the configuration \( C_2 \), written as \( \text{eval}(M, C_1, A, C_2) \), if

\[
\Sigma_{ml} : \mathcal{E} ; \rightarrow \forall K : \forall l : o. C^o A(A, K) \rightarrow C^o(\text{eval } M, K)
\]

is provable in FORUM.

The part of UML specified by natural semantics, \( \Lambda_{use} \), does not contain first-class continuations. If I restrict the above definitions to UML without \( \text{callcc} \) and \( \text{throw} \), then I can prove a correspondence theorem between the two specifications. For this fragment I consider configurations in which the continuation state is empty. The proof is deferred to the appendix.

**Theorem 3.18 (Correspondence theorem for \( \Lambda_{use} \))** Let \( M \in \Lambda_{use} \), \( A \in \text{Answers}_{use} \), \( C_1 = \text{def} \ \{\emptyset, S_1, Ex_1\} \) and \( C_2 = \text{def} \ \{\emptyset, S_2, Ex_2\} \). Further, \( M \) is closed with respect to \( C_1 \).

\( \{M, C_1\} \downarrow \{A, C_2\} \) if and only if \( \text{eval}(H(M), C_1, A(A, K), C_2) \)
Chapter 4

Program Equivalence for \( \Lambda_{vs} \) in FORUM

In this chapter, I study observational equivalence for \( \Lambda_{vs} \) programs in FORUM. A program is treated like a black-box — the only way to determine the behavior of a program is to give it some inputs and watch the output it generates. In this paradigm, two programs are equivalent if whenever they are given identical inputs they generate identical outputs. Equivalence of programs is relative to what can be observed about their computation. For example, the two sorting routines, binary sort and quicksort [Set89], are equivalent if all I can observe is that the algorithms sort their inputs. However, if I can observe the time taken to sort inputs by binary sort and quicksort, then they will not be equivalent.

Understanding observational equivalence of programs is of fundamental importance for a variety of reasons. For example, the compiler for SML of New Jersey first transforms programs into a continuation-passing-style (CPS) intermediate language, and then performs various transformations on the program in the intermediate language to improve its runtime performance [App92]. Let us assume that SML programs and their translations to CPS language evaluate to the same value. Still, there is a question regarding the transformations done on the CPS programs — what is the relationship between the original CPS program
and the transformed CPS program? Unfortunately, [App92] does not answer this question in a formally precise manner. Suppose I take a program $P$ in the CPS language and transform it to $Q$ following [App92]. What is the relationship between the observations I can make of $P$ and $Q$? If the observable properties of $P$ and $Q$ are identical, then up to our definition of observations the transformation maintains the essential nature of the program. However, if the observable properties of $P$ and $Q$ are not identical, then there will be situations in which the observable results produced by $Q$ will differ from the ones produced by $P$. Such a transformation can only be justified if $Q$ is observationally equivalent to $P$ in the environment in which $Q$ is used. The claim is that observational equivalence provides a framework to study transformations performed by compilers. Along similar lines, one can want to change pieces of existing programs as better algorithms and implementations are developed. Observational equivalence provides a framework to verify whether two programs can be exchanged.

The study of observational equivalence for functional languages with state has been particularly difficult [MS88, OT93, SF92, MT92, Sie93, OT92]. In specifying $\Lambda_{\text{vs}}$, I have placed the evaluator for $\Lambda_{\text{vs}}$ within the rich proof theory of FORUM. I use the meta-theory of FORUM to analyze observational equivalence. I first define observational equivalence for $\Lambda_{\text{vs}}$ programs with respect to the natural semantics and the specification in FORUM, $\mathcal{E}_{\text{vs}}$. Next, using a theorem in [MT92], I prove that the two definitions are equivalent. I prove that reduction preserves observational equivalence. This is the basis for the equational theory for $\Lambda_{\text{vs}}$ in [SF92]. I also prove the observational equivalence for some of the examples in [MS88, OT92]. One of the main focuses of my future work will be to analyze proof-theoretic properties of the transformations required for proving observational equivalences.

4.1 Defining Observational Equivalence

In this section, I define two notions of observational equivalence. Firstly, I define when two $\Lambda_{\text{vs}}$ (defined in figure 3.8) programs are observationally equivalent with respect to their natural semantics specification (defined in section 3.2). Secondly, I define when two
\(\Lambda_{vs}\) programs are observationally equivalent with respect to \(E_{vs}\), the specification of \(\Lambda_{vs}\) in FORUM, defined in section 3.2. Next, I prove that the two notions of observational equivalence are identical. This result lets me study the observational equivalence for \(\Lambda_{vs}\) programs in FORUM — the main aim of this chapter.

The basic idea of observational equivalence is that one places a program in a context, and then observes its behavior. Two programs are observationally equivalent if and only if in all contexts the observable behavior remains the same. There are two key concepts here — context and observable behavior. Suppose I have two programs, \(M\) and \(N\), in \(\Lambda_{vs}\). In what contexts can I place \(M\) and \(N\)? If I think in terms of program transformations, then \(M\) is a part of a larger \(\Lambda_{vs}\) program, and I am replacing it with \(N\) — the contexts must come from the syntax of \(\Lambda_{vs}\). Following the argument, a context would be a \(\Lambda_{vs}\) program with one of its pieces missing. The idea would then be that programs \(M\) and \(N\) have identical observational behavior with respect to contexts of the given language, in this case \(\Lambda_{vs}\).

What are the observable properties of a program? Many observations can be made about the computation of a program in a context. Some of the many observable properties of interest are — whether the evaluation of programs terminate, whether the programs evaluate to the same answer, whether the programs yield the same answer with identical execution time, whether the programs create identical number of new memory locations, etc. The most primitive of the choices listed above is observing whether the programs evaluate to the same answer. As the evaluator for the language is deterministic, observing whether programs evaluate to the same answer yields the same relation as observing whether two programs terminate. The other choices for observations are more refined versions of these two primitive observations. If I transform \(M\) to \(N\), the minimal property that I would want of the resulting programs is that one terminates if and only if the other does — guaranteeing the safety of the transformation. I take termination of the evaluation of programs as the observation I make.

Contexts, ranged over by \(C\), are defined in figure 4.1. I use \(c\) to denote the place in the context where the program will be placed. For example, \(\lambda x.c\) is a context. Note that
placing a term, $M$, in the holes in a context, $C$, written as $C[M]$, may cause binding of free variables in the term. For example, if I place $x$ in the hole in $\lambda x. c$, then I get $\lambda x. x$ — the free variable $x$ gets captured by the context. I now define when a program terminates in a state with respect to natural semantics.

**Definition 4.1** [Termination of programs in $\Lambda_{vs}$ in natural semantics] Let $M \in \Lambda_{vs}$ and $S_0$ be a state such that $\text{close}(S_0, M)$. $M$ in the state $S_0$ terminates, written as $(M, S_0) \#_{ns}$, if there exists a value $V$ and a state $S_1$ such that $(M, S_0) \downarrow (V, S_1)$.

Observational equivalence using program contexts was first defined by [Mor68]. It has been studied extensively for call-by-name and call-by-value $\Lambda$-calculus, and its extensions with state [Abr87, AO89, Abr90, Hoa69, MT92, MS88, OT93, Plo76, PS93, SF92]. The definition of observational equivalence for $\Lambda_{vs}$ is an extension to the definition for $\Lambda_v$ [Mor68] along the lines of [MT92].

**Definition 4.2** [Observational Equivalence with natural semantics] Let $M$ and $N$ be two $\Lambda_{vs}$ terms, and $\emptyset$ be the empty state. $M$ is said to be observationally equivalent to $N$ with respect to the natural semantics, written as $M \equiv_{ns} N$, if
\[
\forall C \in \text{Contexts}_{ns} \text{ such that } C[M] \text{ and } C[N] \text{ are closed terms}, (C[M], \emptyset) \#_{ns} \text{ if and only if } (C[N], \emptyset) \#_{ns}.
\]
To determine whether two programs are equivalent, I have to check for their termination in all contexts in $\text{Context}_{\lambda s}$. This obviously becomes a very difficult problem because one has no control over what the arbitrary contexts might do. [MT92] provides an alternate definition of observational equivalence in which they are able to reduce the contexts to $\text{EvCont}_{\lambda s}$, the evaluation contexts. I use a slight variation of the definition of $\equiv_{ciu}$ in [MT92].

**Definition 4.3** $[\equiv_{ciu}]$ Let $M$ and $N$ be two $\Lambda_{\mu s}$ terms. $M$ is said to be observationally equivalent to $N$ with respect to natural semantics, written as $M \equiv_{ciu} N$, if for all $E \in \text{EvCont}_{\lambda s}$ and for all states $S$, such that $\text{close}(S, E[M])$ and $\text{close}(S, E[N])$, $(E[M], S) \uparrow_{ns}$ if and only if $(E[N], S) \uparrow_{ns}$. ■

Going from arbitrary contexts to evaluation contexts, one loses the capability to bind variables using $\lambda$ and state. Since we are only evaluating programs which are closed with respect to the $\lambda$ bound variables, the main concern is the binding of assignable variables. Consider the two programs $(\text{asg } l \ 5; 1)$ and $(\text{asg } l \ 7; 1)$, where $M; N$ is syntactic sugar for $((\lambda d. N) \ M)$, $d \not\in \text{FV}(N)$. I evaluate the programs in a state which maps $l$ to 0. Clearly, the two programs are not observationally equivalent with respect to the definition for $\equiv_{ns}$. However, if in the definition of $\equiv_{ciu}$ I had insisted on $E$ being closed, then the two programs would have been equivalent! The point is that $E[M]$ may have free variables which are defined in the state in which $E[M]$ will be evaluated. The testing of the program is done by not only placing it in different evaluation contexts, but also by altering the state in which it is evaluated. The following theorem showing the equivalence of $\equiv_{ns}$ and $\equiv_{ciu}$ was proved in [MT92].

**Theorem 4.4** *(Theorem ciu, [MT92])* Let $M$ and $N$ be two $\Lambda_{\mu s}$ terms.

$M \equiv_{ns} N$ if and only if $M \equiv_{ciu} M$.

I translate the definition of $\equiv_{ciu}$ into FORUM. In order to do this, I have to generalize the definition of $\text{eval}_{\lambda s}$, because the initial continuation instead of being $K$, as in the definition
\[ C_{\text{us}}([], K) = K \]
\[ C_{\text{us}}(E[f [] M], K) = \lambda v. (\text{eval} \ H_{\text{us}}(M) \lambda u. K' f v u) \]
\[ C_{\text{us}}(E[f V []], K) = \lambda v. K' f \phi_{\text{us}}(V) v \]
\[ C_{\text{us}}(E[[] M], K) = \lambda v. (\text{eval} \ H_{\text{us}}(M) \lambda u. (\text{apply} v u K')) \]
\[ C_{\text{us}}(E[V []], K) = \lambda v. (\text{apply} \ \phi_{\text{us}}(V) v K') \]
\[ C_{\text{us}}(E[if [] M N], K) = \lambda v. (\text{eval} (\text{ifbr} v H_{\text{us}}(M) H_{\text{us}}(N)) K') \]
\[ C_{\text{us}}(E[\text{let} \ x = [] \text{ in } M], K) = \lambda v. (\text{eval} \ H_{\text{us}}(M)[x := v] K') \]
\[ C_{\text{us}}(E[\text{cell} []], K) = \lambda v. (\text{eval} \ H_{\text{us}}(M)) K' \]
\[ C_{\text{us}}(E[\text{read} []], K) = \lambda v. (\text{get} v K') \]
\[ C_{\text{us}}(E[\text{write} [] M], K) = \lambda v. (\text{eval} \ H_{\text{us}}(M) \lambda u. (\text{set} v u K')) \]
\[ C_{\text{us}}(E[\text{write} V []], K) = \lambda u. (\text{set} \ \phi_{\text{us}}(V) u K') \]

where
\[ K' \overset{def}{=} C_{\text{us}}(E, K : \text{vl} \rightarrow o) \]

Figure 4.2: \( C_{\text{us}} \), translation of \( \text{EvCont}_{\text{us}} \) to FORUM terms of type \( \text{vl} \rightarrow o \)

of \( \text{eval}_{\text{us}} \), is now specified by \( E \). I first define \( C_{\text{us}} \), the translation of \( E \) into FORUM in figure 4.2. The definition of \( \text{eval}_{\text{us}} \) needs to be changed.

**Definition 4.5** [Evaluation in FORUM given an initial continuation \( E \)] Let \( M \in \Lambda_{\text{us}}, E \in \text{EvCont}_{\text{us}} \) and \( S_0 \) be a state such that \( \text{close}(S_0, E[M]) \). \( M \) with \( S_0 \) in the initial continuation \( E \) evaluates to \( V \) with \( S_1 \), written as \( \text{eval}_{\text{s}}(E, M, S_0, V, S_1) \), if

\[
\Sigma_{\text{us}} : C_{\text{us}} ; \rightarrow \forall K : \text{vl} \rightarrow o, S_1 \circ (K V) \rightarrow S_0 \circ (\text{eval} \ H_{\text{us}}(M) C_{\text{us}}(E, K))
\]

is provable in FORUM. ■

It is quite obvious that \( \text{eval}_{\text{us}} \) will be true of \( E[M] \) exactly when \( \text{eval}_{\text{s}} \) is true of \( M \) in the continuation \( E \). The following lemma states the relationship precisely, and follows from an easy induction on the structure of \( E \) and the definition of \( C_{\text{us}} \).

**Lemma 4.6** Let \( M \in \Lambda_{\text{us}}, E \in \text{EvCont}_{\text{us}}, V \in \text{Values}_{\text{us}}, S_1 \) be a state, and \( S_0 \) be a state such that \( \text{close}(S_0, E[M]) \).

\( \text{eval}_{\text{us}}(E[M], S_0, V, S_1) \) if and only if \( \text{eval}_{\text{s}}(E, M, S_0, V, S_1) \)
I now define the termination property, $\downarrow_f$, for programs with respect to $E_{vs}$ using the notion of evaluation defined by $\text{eval}_v$. Next, I define when two $\Lambda_{vs}$ programs are observationally equivalent on the basis of their evaluation in FORUM.

**Definition 4.7** [Termination of programs in $\Lambda_{vs}$ in FORUM] Let $M \in \Lambda_{vs}$, $E \in \text{EvCont}_{vs}$, and $S_0$ be a state such that close$(S_0, E[M])$. $M$ in the state $S_0$ with the initial continuation $E$ terminates, written as $(E, M, S_0) \downarrow_f$, if there exists a value $V$, and a state $S_1$ such that $\text{eval}_v(E, M, S_0, V, S_1)$. ■

**Definition 4.8** [Observational Equivalence with $E_{vs}$] Let $M$ and $N$ be two $\Lambda_{vs}$ terms. $M$ is said to be observationally equivalent to $N$ with respect to $E_{vs}$, written as $M \equiv_f N$, if for all $E \in \text{EvCont}_{vs}$ and for all states $S$, such that close$(S, E[M])$ and close$(S, E[N])$, $(E, M, S) \downarrow_f$ if and only if $(E, N, S) \downarrow_f$. ■

In order to use the translation in FORUM to prove observational equivalence, I first need to prove that $\equiv_f$ defines the same relation as $\equiv_{ciu}$.

**Theorem 4.9** ($\equiv_{ciu}$ and $\equiv_f$ coincide) Let $M$ and $N$ be two $\Lambda_{vs}$ terms.

$M \equiv_{ciu} N$ if and only if $M \equiv_f N$.

**Proof:** To prove the above theorem, it is enough to prove that $(E, M, S_0) \downarrow_f$ if and only if $(E[M], S_0) \downarrow_{vs}$, for any state $S_0$ and $E \in \text{EvCont}_{vs}$, such that close$(S_0, E[M])$ and close$(S_0, E[N])$.

Unraveling definitions of $\downarrow_f$ and $\downarrow_{vs}$, all I need to prove is that if for some $V \in \text{Values}_{vs}$ and state $S_1$, $\text{eval}_v(E, M, S_0, V, S_1)$ then $(E[M], S_0) \Downarrow (V, S_1)$. Conversely, if for some $V \in \text{Values}_{vs}$ and state $S_1$, $(E[M], S_0) \Downarrow (V, S_1)$ then $\text{eval}_v(E, M, S_0, V, S_1)$.

By the Correspondence theorem 3.8 in section 3.2

$\text{eval}_v(E[M], S_0, V, S_1)$ if and only if $(E[M], S_0) \Downarrow (V, S_1)$.
Further note that by lemma 4.6

\[ \text{eval}_v(E[M], S_0, V, S_1) \text{ if and only if } \text{eval}_s(E, M, S_0, V, S_1). \]

Using these facts the proof is completed. ■

Now suppose I want to prove that two programs \( M \) and \( N \) are observationally equivalent. Suppose, for some \( E \in \text{EvCont}_u \), states \( S_0 \) and \( S_1 \), and \( V \in \text{Values}_v \), such that \( \text{FV}(E[M]) \subseteq \text{dom}(S) \), \( \text{eval}_s(E, M, S_0, V, S_1) \) is true. I look at the resulting proof tree in FORUM and transform it to a proof tree for \( \text{eval}_v(E, N, S_0, V, S_1) \). If I can exhibit such transformation to and fro, then I have established that \( M \) is observationally equivalent to \( N \). The problem of determining whether two programs are observationally equivalent has been reduced to specifying proof transformations.

### 4.2 Reduction in \( \text{L}_v \) preserves Observational Equivalence

In this section, I use \( E_v \) to prove that if a program reduces to another program, then the two programs are observationally equivalent — the basis for the equational theory in [FH92]. \( \text{eval}_s \) defines the evaluation of a term \( M \) in the state \( S \), and continuation \( E \) to a value \( V \) and state \( S \). In this sense, \( \text{eval}_s \) is not specifying reductions, rather, it is specifying complete evaluations. However, using lemma 4.6 it can be easily proved that

\[ \text{eval}_v(E, M, S_0, V, S_1) \text{ if and only if } \text{eval}_s([], E[M], S_0, V, S_1). \]

Using this basic intuition, I define when \( M \) reduces to \( N \).

**Definition 4.10** [Reduction in \( \text{L}_v \)] Let \( M, N \in \text{L}_v \), \( S_1 \) be state, and \( S_0 \) be a state such that \( \text{close}(S_0, M) \). \( M \) in state \( S_0 \) evaluates to \( N \) in state \( S_1 \), written as \( \text{red}_s(M, S_0, N, S_1) \), if
I would like to prove that reduction preserves observational equivalence, i.e., if
\( \text{red}_s(M, S_0, S_1, N) \) then \( M \equiv_f N \). Unfortunately, as stated my claim would be false. Following is a counter-example.

**Example 4.11** Let \( M = \text{def} \; \text{ref} \; 0, \; S_0 = \text{def} \; \emptyset, \; N = \text{def} \; \lambda, \; \text{and} \; S_1 = \text{def} \; (\lambda, 0) \). Clearly \( \text{red}_s(M, S_0, S_1, N) \) is true.

However, \( \text{ref} \; 0 \) is not observationally equivalent to \( \lambda \). To distinguish the two terms, take \( E = \text{deref} \; [\; ] \) and \( S = (\lambda, P) \), where \( P \) is a divergent program.

The problem is that evaluation of a program may create new memory cells and change the existing state. However, the statement \( M \equiv_f N \) throws away this information. In the above example \( \lambda \) is defined in the state \( S_1 \), but this information was not used in the attempted proof of \( \text{ref} \; 0 \equiv_f \lambda \). On the other hand, \((\text{ref} \; 0)\) is observationally equivalent to \( \lambda \) with respect to all states which map \( \lambda \) to 0. Suppose we had a way of representing state in the syntax of \( \Lambda_{vs} \), then the situation can be repaired. I would change the succeedent of my claim to
\( M' \equiv_f N' \), where \( M' \) and \( N' \) are \( \Lambda_{vs} \) terms, such that \( M' \) incorporates the state \( S_0 \) and \( M \), while \( N' \) incorporates the state \( S_1 \) and \( N \). I will prove the new statement of my claim below. I define the translation of a state into the syntax of \( \Lambda_{vs} \) [MT92, SF92]. \( (M; N) \) is syntactic sugar for \( (\lambda d. \; N) \; M, \; d \notin \text{FV}(N) \).

**Definition 4.12** Given a state \( S \), I define a \( \Lambda_{vs} \) term \( S^+ \), the encoding of state \( S \) in \( \Lambda_{vs} \).

\[
S^+ = \text{def} \; ((\lambda x_1, \ldots, x_n, y. \text{asg} \; x_1 \; V_1; \ldots; \text{asg} \; x_1 \; V_1; y)(\text{ref} \; 1)\ldots(\text{ref} \; n))
\]

Where

- \( l_S = x_1, \ldots, x_n \).
\[ y \text{ is distinct from all } x_i, \ i \in [1, n]. \]

\[ S(x_i) = V_i, \ i \in [1, n]. \]

The proof of the following lemma is immediate from the construction of \( S^+ \) for a state \( S \).

**Lemma 4.13** Let \( N \in \Lambda_{vs} \), and \( S \) a state such that \( \text{close}(S, N) \).

\[ \text{red}_s((S^+ N), \emptyset, N, S) \text{ is true}. \]

**Theorem 4.14** Let \( M \) be a redex, \( N \in \Lambda_{vs}, S_1 \) be a state, and \( S_0 \) a state such that \( \text{close}(S_0, M) \).

If \( \text{red}_s(M, S_0, N, S_1) \) then \( (S_0^+ M) \equiv_f (S_1^+ N) \).

**Proof:** Assume given \( M, N \in \Lambda_{vs}, S_0 \) a state such that \( \text{close}(S_0, M) \), and \( \text{red}_s(M, S_0, N, S_1) \) is true. Let \( M' = \text{def} \ (S_0^+ M) \) and \( N' = \text{def} \ (S_1^+ N) \). Unraveling the definition of \( M' \equiv_f N' \), I have to prove that for any arbitrary state \( S \) and \( E \in \text{EvCont}_{vs} \), such that \( \text{close}(S, E[M']) \) and \( \text{close}(S, E[N']) \), \((E, M', S) \vdash_f \) if and only if \((E, N', S) \vdash_f \).

**Going from right to left,** I have to prove that if there exists \( V \in \text{Values}_{vs} \), and a state \( S' \) such that \( \text{eval}_s(E_1, N', S, V, S') \), then \( \text{eval}_s(E_1, M', S, V, S') \).

Assume \( \text{eval}_s(E_1, N', S, V, S') \). I have to construct a proof for \( \text{eval}_s(E_1, M', S, V, S') \). \( \Sigma_{vs}, K \) is in the signature of all the sequents shown in all the proofs that I construct. I start by constructing a proof below, called \( \sigma_1 \). In \( \sigma_1 \), \( \delta_1 \) is obtained by unfolding the definition for \( \text{red}_s(M, S_0, N, S_1) \) and lemma 4.13 gives me \( \delta_2 \). Let \( M_1 = \text{def} \ \mathcal{H}_{vs}(M) \), \( N_1 = \text{def} \ \mathcal{H}_{vs}(N) \), \( M'_1 = \text{def} \ \mathcal{H}_{vs}(M') \), \( N'_1 = \text{def} \ \mathcal{H}_{vs}(N') \), \( K' = \text{def} \ \mathcal{C}_{vs}(E, K) \), and \( V_1 = \text{def} \ \phi_{vs}(V) \).
\[ \frac{\delta_1 \cdot \text{eval } N \cdot K \quad \delta_2 \cdot \text{eval } M \cdot K}{\text{CutS}} \]

By definition of \text{eval}, and \( N' \) I must have the following proof in FORUM.

\[ \Sigma_S, \Sigma_{S_i}, K : \text{CL}_S ; S^\omega(KV_1) \quad \gamma_1 \quad \text{(eval } N \cdot K') \varnothing \Gamma_S \varnothing \Gamma_{S_i} \]

\[ \therefore \text{backchain on ref clause, } \forall R, \Rightarrow R \quad \Sigma_S, K : \text{CL}_S ; S^\omega(KV_1) \quad \text{(eval } N_1 \cdot K') \varnothing \Gamma_S \]

\[ \therefore \forall R, \Rightarrow R \quad \Sigma_S, K : \text{CL}_S ; S^\omega(KV_1) \quad \text{(eval } N_1 \cdot K') \varnothing \Gamma_S \]

Using \( \gamma_1 \) in the above proof, I construct \( \sigma_2 \) below.

\[ \Sigma_S, \Sigma_{S_i}, K : \text{CL}_S ; S^\omega(KV_1) \quad \gamma_1 \quad \text{(eval } N \cdot K') \varnothing \Gamma_S \varnothing \Gamma_{S_i} \]

\[ \therefore \text{, } \Rightarrow R, \forall R \quad \Sigma_S, K : \text{CL}_S ; S^\omega(KV_1) \quad \text{(eval } N_1 \cdot K') \varnothing \Gamma_S \]

Using \( \sigma_2 \), I construct the required proof below. To keep the proof readable, I do not write \( \Sigma_S, K \) in the signature, and \( \text{CL}_S \) in the intuitionistic part of some of the sequents.

\[ S^\omega(KV_1) \quad \sigma_2 \quad \text{(eval } N \cdot K') \varnothing \Gamma_S \quad S^\omega(\text{eval } N \cdot K') \quad \sigma_4 \quad \text{(eval } M \cdot K') \quad \text{CutS} \]

\[ \Sigma_S, K : \text{CL}_S ; S^\omega(KV_1) \quad \text{(eval } M_1 \cdot K') \varnothing \Gamma_S \]

\[ \therefore \forall R, \Rightarrow R \quad \Sigma_S, K : \text{CL}_S ; S^\omega(KV_1) \quad \text{(eval } M_1 \cdot K') \]

To complete the construction, I build \( \sigma_3 \) below. \( \sigma_3 \) is obtained from \( \sigma_1 \) by inflating the signature and the intuitionistic parts of the sequents in the proof.

\[ \Sigma_S, K : \text{CL}_S ; S^\omega(\text{eval } N \cdot K_1) \quad \sigma_3 \quad \text{CutS} \]

\[ \Sigma_S, K : \text{CL}_S ; S^\omega(\text{eval } N \cdot K_1) \quad \text{(eval } M_1 \cdot K') \]
Going from left to right, I use the fact the $\mathcal{E}_{vs}$ is deterministic, because there is exactly one clause for every term constructor. Hence, if $\text{red}_1(M, S_0, N, S_1)$, then every evaluation of $M$ that evaluates the term beyond $N$ must pass through the reduction of $M$ to $N$. Assuming $\text{eval}_1(E_1, M', S, V, S')$, I have to construct a proof of $\text{eval}_1(E_1, N', S, V, S')$. The proof in $\text{eval}_1(E_1, M', S, V, S')$ must have the following shape. To keep the proof readable, I do not write the signature and the intuitionistic part of the sequent in all the sequents in the proof.

\[
\Sigma_{S}, \Sigma_{S_1} : \text{CL}_S, \text{CL}_{S_0} ; S^\delta(K \nu_1) \rightarrow (\text{eval}_1 N'_1 K') \varphi \Gamma_S \varphi \Gamma_S_1
\]

: reduction of $M_1$ in $S_0$ to $N_1$ in $S_1$

\[
S^\delta(K \nu_1) \rightarrow (\text{eval}_1 M'_1 K') \varphi \Gamma_S \varphi \Gamma_{S_0}
\]

: backchain on ref clause, $\forall R, \Rightarrow R$

\[
S^\delta(K \nu_1) \rightarrow (\text{eval}_1 M'_1 K') \varphi \Gamma_S
\]

: $\forall R, \Rightarrow R$

\[
S^\delta(K \nu_1) \rightarrow S_0^\delta(\text{eval}_1 M'_1 K')
\]

Using $\delta$ I construct the required proof below.

\[
\Sigma_{S}, \Sigma_{S_1} : \text{CL}_S, \text{CL}_{S_0} ; S^\delta(K \nu_1) \rightarrow (\text{eval}_1 N'_1 K') \varphi \Gamma_S \varphi \Gamma_S_1
\]

: backchain on ref clause, $\forall R, \Rightarrow R$

\[
\Sigma_{S} : \text{CL}_S ; S^\delta(K \nu_1) \rightarrow (\text{eval}_1 N'_1 K') \varphi \Gamma_S
\]

: $\forall R, \Rightarrow R$

\[
S^\delta(K \nu_1) \rightarrow S_0^\delta(\text{eval}_1 N'_1 K')
\]

The above proof highlights some key aspects $\mathcal{E}_{vs}$. If $M$ reduces to $N$, then $N'$, the translation of $N$, entails $M'$, the translation of $M$. Hence, whenever $N'$ evaluates to a value, so will $M'$. The proof going from right to left uses this fact and CutS to construct the required proof. Now, $N'$ entails $M'$. This does not necessarily imply that an evaluation of $M'$ has to have $N'$ as an intermediate state. Since $\mathcal{E}_{vs}$ is deterministic, i.e. for every term construct of type $\text{tm}$ there is exactly one clause, if $M'$ evaluates to a value, and $N'$ entails $M'$, it
necessarily follows that the evaluation has $N'$ in some intermediate state. This observation yields the proof going from left to right in the above theorem.

### 4.3 Observational Equivalence proofs in FORUM

In this section I present the Meyer-Sieber examples [MS88] in the UML, and prove the desired equivalences, interpreting equivalence as $\equiv_f$. I have converted the Algol-like notation of [MS88] into UML syntax following [MT92]. State introduces new nuances into the programming language, and correspondingly into any theory which tries to study the equivalence of programs with state. The idea behind the examples was to highlight some of the novel issues that come up when state is added to a higher order language. Let $\Omega$ be a divergent program in UML.

**Example 4.15** [Example 1] Let $M \in L_{us}$, and $M_1 = \text{def let } x = \text{ref } 0 \text{ in } M$, $x \not\in FV(M)$.

$$M \equiv_f M_1$$

The intuitive justification is simple. As $x \not\in FV(M)$, the creation of $x$ has no effect on the behavior of $M$. ■

The proof for this example follows from the following more general statement of the problem. The lemma is proved by a straightforward induction on the height of the evaluations of $(S^+_3 M)$ and $M$ in FORUM. The essential point is that the evaluation of $(S^+_3 M)$ creates new locations by picking new eigen-variables in FORUM. Hence, if $(S^+_3 M)$ is placed in an evaluation context $E$, then $E$ cannot access the newly created locations. The computations of $(S^+_3 M)$ and $M$ are identical except for this difference.

**Lemma 4.16 (Elimination of inaccessible cells)** Let $M \in L_{us}$, $L = \text{def } l_1, \ldots, l_n$, and $S_1$ be a state such that :
- close($S_1, M$).
- $L \subseteq \text{dom}(S_1)$.
- for all $l \in L$, $l \not\in \text{FV}(M)$.
- if $l \in \text{FV}(M)$, then for all $l' \in L$, $l' \not\in \text{FV}(S_1(l))$.
- $S_3$ is the restriction of $S_1$ to $L$.

$$(S_3^L M) \equiv_f M$$

I first point out some equivalences for which I do not need induction on the height of proofs. The proofs in FORUM for these terms are permuted to each other.

**Remark 4.17** Let $M, N, PA_{\nu s}$, and $V \in \Lambda_{\nu s}$.

1. let val $x = V$ in $M \equiv_f (\lambda x. M)V$.
2. $E[M] \equiv_f \text{let val } x = M \text{ in } E[x], x \not\in \text{FV}(E)$.
3. $E[\text{let val } x = M \text{ in } N] \equiv_f \text{let val } x = M \text{ in } E[N], x \not\in \text{FV}(E)$.
4. let val $x = \text{ref } M$ in $E[\bullet] \equiv_f \text{let val } x = \text{ref } V \text{ in } E[\text{asg} x M], x \not\in \text{FV}(M)$.
5. let val $x = \text{ref } P \text{ in } ((\lambda z. M) N) \equiv_f ((\lambda z. \text{\text{let val } x = \text{ref } P \text{ in } \lambda z. M}) N), x \not\in \text{FV}(N)$
6. let val $x = \text{ref } V \text{ in } ((\lambda z. M) N) \equiv_f ((\lambda z. \text{\text{let val } x = \text{ref } V \text{ in } M}) N), x \not\in \text{FV}(N), z \not\in \text{FV}(M) \cup \text{FV}(V)$

Equivalence 1 is obvious from the interpretation of letval. Equivalence 2 and 3 follow from lemma 4.6. For equivalence 4, observe that $x$ is inaccessible from $M$, thus evaluating $M$ with or without $x$ declared in the environment makes no difference. Equivalences 5 and 6 result from the permutation of proofs because of information about variable occurrences.

**Example 4.18** [Example 2] $M \in \Lambda_{\nu s}$. 73
\begin{itemize}
\item $M_1 = \text{def } \Omega$.
\item $M_2 = \text{def } \text{let val } x = (\text{ref } \bullet) \in ((\text{asg } x \text{ true}) ; (M \bullet) ; \text{if} (\text{deref } x) \Omega \text{ 1}), x \not\in \text{FV}(M)$.
\end{itemize}

\textbf{Lemma 4.19} $M_{2,1} \equiv_f M_{2,2}$

\textbf{Proof:} The strategy is to show that for any $E \in \text{EvCont}_{vs}$, and any state $S$ such that $	ext{close}(S, M_2), (E, M_2, S) \not\vdash_f$ is not true.

On the contrary, suppose that $(E, M_2, S_0) \not\vdash_f$ is true for some $E_0 \in \text{EvCont}_{vs}$, and state $S_0$ such that $	ext{close}(S_0, E_0[M_2])$. Then there exists a value $V$ and a state $S_1$ such that $	ext{eval}_s(E_0, M_2, S_0, V, S_1)$ is true. Now by the equivalence in remark 4.17,

\[ M_2 \equiv_f ((M \bullet); \text{let val } x = (\text{ref } 0) \in (\text{if} (\text{deref } x) \Omega \text{ 1})). \]

\begin{itemize}
\item $V' = \text{def } \phi_{vs}(V)$.
\item $M' = \text{def } \mathcal{H}_{vs}(M)$.
\item $\Omega' = \text{def } \mathcal{H}_{vs}(\Omega)$.
\item $Q = \text{def } (\text{let val } x = (\text{ref } 0) \in (\text{if} (\text{deref } x) \Omega \text{ 1})), Q' = \text{def } \mathcal{H}_{vs}(Q)$.
\item $N = \text{def } ((M \bullet); Q), N' = \text{def } \mathcal{H}_{vs}(N)$.
\item $C_1 = \text{def } \mathcal{C}_{vs}(E K)$.
\item $C_2 = \text{def } \lambda u. (\text{apply } (\text{abs } \lambda d. Q') u C_1), d \not\in \text{FV}(Q)$.
\end{itemize}

By definition of $\equiv_f$ I have a proof $\delta$ of

\[ \Sigma_{vs}, K, \Sigma_{S_0} : \mathcal{E}_{vs}, \mathcal{C}_{L_{S_0}} ; S_1^\prime(K V') \longrightarrow (\text{eval } N' C_1) \not\in \Gamma_S \]

74
in FORUM.

The last rules of $\delta$ must be as shown below.

\[
\begin{array}{c}
\Sigma_{u3}, K, \Sigma_{S_0} : \mathcal{E}_{v3}, \mathcal{C}L_{S_0} ; S_1^0(K V^\prime) \xrightarrow{\delta_1} \text{eval \ (app \ M' \bullet) } C_2) \varnothing \Gamma_{S_0} \\
\vdots \\
\Sigma_{u3}, K, \Sigma_{S_0} : \mathcal{E}_{v3}, \mathcal{C}L_{S_0} ; S_1^0(K V^\prime) \xrightarrow{\delta_2} \text{eval \ N' \ C_1) \varnothing \Gamma_{S_0} }
\end{array}
\]

Suppose $\langle (M \bullet), S_0 \rangle$ diverges then there cannot exist a $\delta_1$ because of theorem 3.8. There is a contradiction, hence I am done.

Suppose $\langle (M \bullet), S_0 \rangle \Downarrow (U, S_2)$. Hence, by theorem 3.8 I would get $\delta_1$ as shown below. Let $U' =_{def} \phi_{u3}(U), S_3 =_{def} S_2[l \mapsto 0], l \not\in \text{dom}(S_2)$.

\[
\begin{array}{c}
\Sigma_{u3}, K, \Sigma_{S_0} : \mathcal{E}_{v3}, \mathcal{C}L_{S_0} ; S_1^0(K V^\prime) \xrightarrow{\delta_2} \text{eval \ N' \ C_1) \varnothing \Gamma_{S_3} \\
\vdots \\
\Sigma_{u3}, K, \Sigma_{S_2} : \mathcal{E}_{v3}, \mathcal{C}L_{S_2} ; S_1^0(K V^\prime) \xrightarrow{\delta_3} \text{eval \ Q' \ C_1) \varnothing \Gamma_{S_3} \\
\Sigma_{u3}, K, \Sigma_{S_2} : \mathcal{E}_{v3}, \mathcal{C}L_{S_2} ; S_1^0(K V^\prime) \xrightarrow{\delta_4} \text{eval \ Q' \ C_1) \varnothing \Gamma_{S_3} }
\end{array}
\]

Using theorem 3.8 and the fact that $\Omega$ is divergent I get that $\delta_2$ cannot exist in FORUM.

\[\text{Example 4.20 [Example 3]} \text{ Let } M \in A_{u3}. \]

- $M_{3,1} =_{def} \text{let \ val } x = (\text{ref } 0) \text{ in (let \ val } y = (\text{ref } 0) \text{ in } M)$. 

- $M_{3,2} =_{def} \text{let \ val } y = (\text{ref } 0) \text{ in (let \ val } x = (\text{ref } 0) \text{ in } M)$. 

\[\text{Lemma 4.21 } M_{3,1} \cong_f M_{3,2} \]

75
**Proof:** The idea is to show that given a proof for the evaluation of \( M_{3.1} \) I can transform it to an evaluation of \( M_{3.2} \), and vice versa.

Let \( E \) be an arbitrary evaluation context and \( S_0 \) any state such that \( \text{close}(S_0, E[M_{3.1}]) \). Suppose \( \langle E[M_{3.1}], S_0 \rangle \Downarrow \langle V, S_1 \rangle \), for some value \( V \) and state \( S_1 \). Let \( V' = \text{def} \ \phi_v s(V) \), \( C_1 = \text{def} \ C_v s(E, K) \), \( N'_1 = \text{def} \ \mathcal{H}_v s(M_{3.1}) \), \( N'_2 = \text{def} \ \mathcal{H}_v s(M_{3.2}) \).

By theorem 3.8, I have a proof in \( \text{FORUM}, \delta_1 \), of

\[
\Sigma_{v3}, \Sigma_{S_0}, K : \mathcal{E}_{v3}, CL_{S_0}; S'_3(K V') \longrightarrow (\text{eval } N'_1 C_1) \varphi \Gamma_{S_0}.
\]

I need to transform \( \delta_1 \), into a proof for the evaluation of \( N'_2 \) in \( \text{FORUM} \). The shape for \( \delta_1 \) is shown below.

\[
\Sigma_{v3}, \Sigma_{S_1}, K : \mathcal{E}_{v3}, CL_{S_1}; S'_3(K V') \longrightarrow (\text{eval } (\text{app } Q' l_1) l_2) C_1) \varphi \Gamma_{S_0} \varphi (P_1 0) \varphi (P_2 0)
\]

Using \( \delta_2 \), I can trivially get the evaluation for \( M_{3.2} \) in \( \text{FORUM} \). The proof in the other direction follows from a similar argument. □

The above proof shows the advantage of using eigen-variables to generate names of new assignable variables. The two terms in this example can essentially be renamed to each other. This fact is made precise by the usage of the eigen-variables.

**Example 4.22** [Example 4] Let \( M \in \Lambda_{v3} \).

- \( M_{4.1} = \text{def} \ \Omega \).
- \( M_{4.2} = \text{def} \ \text{let val } x = (\text{ref } 0) \text{ in}
  \begin{align*}
  &\text{let val } f = \lambda d. (\text{asg } x (+ 2 (\text{deref } x))); (\text{deref } x) \text{ in} \\
  & (M f); (\text{if } (= 0 (\text{mod} (\text{deref } x) 2)) \Omega 1), x, f \notin \text{FV}(M).
  \end{align*} \)
Lemma 4.23  \( M_{4.1} \approx_f M_{4.2} \)

**Proof:** The strategy is to show that \( M_{4.2} \) will have no evaluation in FORUM. I only consider the case when \( (M \ f) \) converges, as otherwise the argument is trivial. I introduce a new idea in this proof.

The essential point of the example is that access to the local variable \( x \) is passed to \( M \) only via \( f \). Hence, the content of \( x \) can only be incremented by 2, and then read out — no other operation is possible on \( x \). In a way, an abstract data type has been created with the only interface function being \( f \). The if statement checks whether this abstraction was maintained by \( M \) or not.

Let

\[
F = \text{def} \ \forall K, V, U. [(\text{apply } c \ V \ K) \phi(P \ U)] \leftarrow [(K \ 2 + (U/2 + 1)) \phi(P \ 2 + (U/2 + 1))].
\]

\( P \) and \( c \) are declared in the signature. \( F \) encapsulates the computational behavior of \( f \). Note how it does away with the name of the cell \( (P \ U) \), and need for \text{get}C and \text{set}C clauses for \( P \).

Suppose I have a proof, \( \delta \), in FORUM of

\[
\Sigma_{\nu_3}, \Sigma_{S_0}, K, c, P, l : \epsilon_{\nu_3}, \text{CL}_{S_0}, F; S_{i}^{(i)}(K, V') \rightarrow (\text{eval}(\text{app}\ M' \ c) \ C_1) \phi \Gamma_{S_0} \phi(P \ 0),
\]

where \( E \in \text{EvCont}_{\nu_3} \),

- \( S_0 \) is a state such that \( \text{close}(S_0, E[M]) \).
- \( V' = \text{def} \ \phi_{\nu_3}(V) \).
- \( M' = \text{def} \ \mathcal{H}_{\nu_3}(M) \).
\[ C_1 = \text{def } \mathcal{C}_{us}(E, K). \]

In \( \delta \) the only interaction for \( P \) is via \( F \) as there are no \texttt{getC} and \texttt{setC} clauses for \( P \) and \( l \) in the sequent above. Thus, it is trivially true that \( S_1(l) \) is a multiple of 2.

I construct \( \delta_1 \) from \( \delta \) using \textit{CutS}. Let \( F' = \text{def } F[c := \text{Add}_2] \), and

\[
\text{Add}_2 = \text{def } \mathcal{H}_{us}(\lambda d. (\text{asg } x (+ 2 (\text{deref } x))); (\text{deref } x)).
\]

\[
\text{Add}_2 \text{ is a } \Sigma_{us}, l \text{ term } \quad \text{c, P, l : } F; S_1(l)(K, V') \quad \delta \quad (\text{eval } (\text{app } M' c) C_1) \varnothing \Gamma_{S_0} \varnothing (P 0) \\
\text{P, l : F}; S_1'(K, V') \quad \delta \quad (\text{eval } (\text{app } M' \text{Add}_2) C_1) \varnothing \Gamma_{S_0} \varnothing (P 0)
\]

I construct \( \delta_2 \) below. Let \( \text{CL}_l = \text{def } \{ \text{getC}(P, l), \text{setC}(P, l) \} \). In this proof I have used the equation \( 2 * (U/2 + 1) = U + 2 \), which is true if the division is for real numbers.

\[
\text{Add}_2 \text{ is the computation of } (\text{app } M' \text{Add}_2) \text{ in FORUM!}
\]

Using \( \delta_1 \) and \( \delta_2 \), I construct \( \delta_3 \) below. \( \delta_3 \) is the computation of \( (\text{app } M' \text{Add}_2) \) in FORUM!

By the observation regarding \( \delta_1 \), \( S_1(l) \) is a multiple of 2. Hence, it is clear that \( (if (= 0 (\text{mod } (\text{deref } l) 2)) \Omega 1) \) will evaluate to the value of \( \Omega \). Using theorem 3.8 and the fact that \( \Omega \) is a divergent program, I have proved that \( M_{4,2} \) does not converge.

\[
\text{To complete the proof I have to convert any evaluation of } (\text{app } M' \text{Add}_2) \text{ into an evaluation of } (\text{app } M' c). \text{ I prove this by a straightforward induction on the height of the evaluation of } (\text{app } M' \text{Add}_2). \]
Example 4.24 [Example 5] Let $M \in \Lambda_v$.

- $M_{5.1} = \text{def } (M \lambda d. *)$
- $M_{5.2} = \text{def } \text{let } \text{val } x = (\text{ref } 0) \text{ in }$
  
  \hspace{1cm} \text{let } \text{val } f = \text{in } \lambda d. \text{asg } x (1 + (\text{deref } x))$
  
  \hspace{1cm} (M f)$

\[\]

Lemma 4.25 $M_{5.1} \approx_f M_{5.2}$

\textbf{Proof:} The strategy is to show that an evaluation of $M_{5.1}$ can be transformed to an evaluation of $M_{5.2}$, and vice versa. The essential point of this example is that incrementing achieved by $f$ is useless for $M$ because it can never read the contents of $x$. Hence, might as well as use $\lambda d. *$ instead. This equivalence has been especially problematic for various denotational semantics [OT92].

The proof strategy is the same as in example 4.

Let $F = \text{def } \forall K, V. (\forall U. [(\text{apply } c V K) \varphi (P U)] \Leftarrow \forall W. [(K *) \varphi (P W)]$. $P$ and $c$ are declared in the signature. $F$ encapsulates the computational behavior of $f$ and $\lambda d. *$. Note how it does away with the name of the cell $(P U)$, and need for $\text{getC}$ and $\text{setC}$ clauses for $P$.

Suppose I have a proof, $\delta$, in FORUM of

$$\Sigma_{us}, \Sigma_S, K, c, P, l: E_{us}, CL_S, F; S_1^o(K, V') \rightarrow (\text{eval } (\text{app } M' c) C_1) \varphi \Gamma_S, \varphi (P 0),$$

where $E \in \text{EvCont}_{us}$.

- $S_0$ is a state such that $\text{close}(S_0, E[M])$.
- $V' = \text{def } \phi_{us}(V)$.

79


- \( M' = \text{def} \ \mathcal{H}_{uv}(M) \).
- \( C_1 = \text{def} \ \mathcal{C}_{uv}(E, K) \).

I construct \( \delta \) from \( \delta \) using \( \text{CutS}, i \in [1, 2] \). Let \( F_i = \text{def} \ F[e := D_i] \), where

\[
D_1 = \text{def} \ \mathcal{H}_{uv}(\lambda d. \text{asg} \ l (+1 \text{ (deref} \ l))) , \ D_2 = \text{def} \ \mathcal{H}_{uv}(\lambda d. \cdot) , \text{ and } i \in [1, 2].
\]

\[
\begin{align*}
D_i \text{ is a } \Sigma_{uv,l} \text{ term } c, P, l : F ; S_i^c(K, V') & \rightarrow (\text{eval (app } M' \text{ c) } C_1) \varphi \Gamma_{S_0} \varphi(P 0) \\
P, l : F_i ; S_i^c(K, V') & \rightarrow (\text{eval (app } M' \text{ } D_i) \ C_1) \varphi \Gamma_{S_0} \varphi(P 0)
\end{align*}
\]

I construct \( \delta_3 \) below. Let \( \text{CL}_i = \text{def} \ \{ \text{getC} (P, l), \text{setC} (P, l) \} \).

\[
\begin{align*}
P, l, K, U, V : \text{CL}_i, \mathcal{E}_{uv} ; (K \bullet) \varphi(P U + 1) & \rightarrow (K \bullet) \varphi(P U + 1) \\
P, l, K, U, V : \text{CL}_i, \mathcal{E}_{uv} ; \forall W. (K \bullet) \varphi(P W) & \rightarrow (K \bullet) \varphi(P U + 1)
\end{align*}
\]

\[
\begin{align*}
\vdash \text{evaluate } D_1 \\
P, l, K, U, V : \text{CL}_i, \mathcal{E}_{uv} ; \forall W. (K \bullet) \varphi(P W) & \rightarrow (\text{apply } D_1 V K) \varphi(P U)
\end{align*}
\]

\[
\begin{align*}
\vdash \forall R, \neg \rightarrow \ R \\
P, l, \Sigma_{uv} : \text{CL}_i, \mathcal{E}_{uv} ; & \rightarrow F'
\end{align*}
\]

I construct \( \delta_4 \) below. Let \( \text{CL}_i = \text{def} \ \{ \text{getC} (P, l), \text{setC} (P, l) \} \).

\[
\begin{align*}
P, l, K, U, V : \text{CL}_i, \mathcal{E}_{uv} ; (K \bullet) \varphi(P U) & \rightarrow (K \bullet) \varphi(P U) \\
P, l, K, U, V : \text{CL}_i, \mathcal{E}_{uv} ; \forall W. (K \bullet) \varphi(P W) & \rightarrow (K \bullet) \varphi(P U)
\end{align*}
\]

\[
\begin{align*}
\vdash \text{evaluate } D_2 \\
P, l, K, U, V : \text{CL}_i, \mathcal{E}_{uv} ; \forall W. (K \bullet) \varphi(P W) & \rightarrow (\text{apply } D_2 V K) \varphi(P U)
\end{align*}
\]

\[
\begin{align*}
\vdash \forall R, \neg \rightarrow \ R \\
P, l, \Sigma_{uv} : \text{CL}_i, \mathcal{E}_{uv} ; & \rightarrow F'
\end{align*}
\]

Composing \( \delta_1 \) and \( \delta_3 \), I construct the evaluation of \( \text{app } M' D_1 \) in FORUM. Composing \( \delta_2 \) and \( \delta_4 \), I construct \( \delta_6 \). \( \delta_6 \) fails to be the evaluation of \( \text{app } M' D_2 \) in FORUM, because of the extra cell \( P \) in the environment. However, I can use lemma 4.16 to get rid of this extra cell. To complete the proof I have to convert proofs of evaluation of \( \text{app } M' D_i \), \( i \in [1, 2] \) to an evaluation of \( \text{app } M' c \). I prove this by a straightforward induction on the height of the evaluation of \( \text{app } M' D_i \). ❄️
Using the meta-theory of FORUM, I have proved many of the Meyer-Sieber examples from [MS88] and an example from [OT92]. This style of reasoning bears a close resemblance to the style in [OT93] using logical relations. The use was most remarkable in the fourth and fifth example where I was able to use $\text{Cut}S$ and $\text{Cut}L$ to get the proofs. One direction in this proof still needed to induct on the height of evaluation proof tree, a weakness of the argument that I would like to get rid of in my future work. The first three proofs were essentially arguments about permutations of evaluations in FORUM. I want to investigate whether I can make these proofs compositional using resolution on the proof rules. The meta-theory gives me distinct advantage in the above proofs. However, I would like to develop the meta-theory so that composition, and permutation of resolution can be studied in finer detail.
Chapter 5

Specifying DLX - a RISC architecture

In this chapter I specify the sequential and pipelined operational semantics for the DLX architecture – a prototypical RISC (Reduced Instruction Set Computer) architecture – in FORUM. DLX is a generic load/store machine representative of the RISC machines which have become very popular since the late 1980’s, e.g. Intel i860, MIPS R2000/R3000, Motorola 88000, SPARC, PowerPC. I will prove that the sequential and pipelined specifications of DLX are identical, and using this equivalence give a simple proof of the correctness of code rescheduling. The main point of this chapter is that FORUM facilitates the declarative specification of the concurrent pipelined operational semantics of DLX with complex synchronizations. Furthermore, the framework allows me to handle structural and data-hazards, and specify optimizations such as call-forwarding and early-branch prediction declaratively.

The key feature of the FORUM specification is that it specifies the computation of the pipeline as compared to the existing specifications in the literature which specify the pipeline’s temporal behavior [AL93]. The specifications in FORUM are executable as logic
programs yielding a prototype implementation of the pipeline which can be used for collecting statistics and experimentation. This seems to be a unique feature of the FORUM specification amongst all the specifications for DLX style pipelines. Furthermore, the FORUM specification is concurrent – different stages of the pipelines can be computed independently of each other. Moreover, the equivalence of sequential and pipeline operational semantics provides me with a tool to prove correct various optimizations done by the back-end optimizer and/or hardware such as code rescheduling. The proofs of program-equivalence are once again achieved by proof transformations very much along the lines of chapter 4.

In this chapter I first introduce the DLX architecture, and specify its sequential operational semantics. I specify only the integer part of DLX and as such, the discussion will be restricted to relevant parts of the architecture, the reader is referred to [HP90] for a detailed description. The specification of the floating-point operations and interrupts in the pipeline do not require any new specification techniques, and thus they have been left out from the present discussion. Next, I specify the operational semantics of DLX pipeline. I prove that the DLX pipeline specification is equivalent to the sequential specification. The program is then extended to call-forwarding and early-branch prediction. Finally, I use the sequential operational semantics to prove the correctness of code scheduling.

5.1 The DLX architecture

The architecture of the DLX machine – the user visible part of the instruction set of DLX – emphasizes design for pipeline efficiency, an easily decoded instruction set and efficiency as a compiler target. In this section, I describe the architecture for the integer part of DLX, for a complete description and discussion of the entire DLX architecture the reader is referred to [HP90]. I have left out the jump instructions because their introduction needs no new ideas – the presence of branch instructions causes all the complications that they may cause.

DLX has thirty-two 32 bit general-purpose registers (GPRs). Memory is word addressable
<table>
<thead>
<tr>
<th>Instruction</th>
<th>Instruction name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>lw R1, I(R2)</td>
<td>Load word</td>
<td>( R1 := M[I + R2] )</td>
</tr>
<tr>
<td>add R1, R2, R6</td>
<td>Add</td>
<td>( R1 := R2 + R6 )</td>
</tr>
<tr>
<td>sll R1, R2, R4</td>
<td>Shift left logical</td>
<td>( R1 := R2 \ll R4 )</td>
</tr>
<tr>
<td>seq R1, R2, R3</td>
<td>Set equal to</td>
<td>if ( R2 = R3 ) ( R1 := 1 ) else ( R1 := 0 )</td>
</tr>
<tr>
<td>beq R4, ;I</td>
<td>Branch equal zero</td>
<td>if ( R4 = 0 ) ( PC := PC + I + 4; ) ( ((PC + 4) - 2^{15}) \leq PC + I \leq ((PC + 4) + 2^{15}) )</td>
</tr>
</tbody>
</table>

Figure 5.1: Semantics of example instructions in DLX.

in Big Endian mode with 32-bit address, and all memory references are through loads or stores between memory and the GPRs. I treat memory as word addressable to avoid unilluminating details regarding byte and halfword addressability. All instructions are 32-bits and all memory accesses must be aligned. Since I am only specifying the integer part of DLX, I use the GPRs for integer multiply and divide instructions.

There are three classes of instructions for the integer part of DLX: loads and stores, ALU operations, and branches. A load instruction is written as lw R1, I(R2) with the intended semantics being that R1 is assigned the contents of the memory array, M, from the address R2 plus the 16-bit integer I. A store operation written as sw I(R1), R2 results in \( M[I + R1] := R2 \). The operands and results for all ALU operations are stored in registers. The operations include simple arithmetic and logical operations: add, subtract, and, or, exclusive or, shifts and compares. One of the arguments in the operations can be the number itself (called the immediate) instead of a register. However, to focus on the central issues in the specification, I do not consider these variants of the instructions. Typical ALU instruction is written as op R1, R2, R3 with the intended semantics being that R1 is assigned the value \( R2 \ op \ R3 \).

The branch instructions can only test for equality with zero and the offsets are limited to 16-bit integers. To present the specification in a more understandable way, I have chosen a representative set of instructions from each class for the DLX [HP90]. I have added instructions end and begin, which cause the computation to halt and start, respectively.
Data transfers

lw, sw

Move data between registers and memory
Load-word, store word (to/from GPR).

Arithmetic/Logical

add, addu, sub, subu
mul, div
mulu, divu
and, or, xor
sll, srl, sra
s-

Operations on integer or logical data in GPRs
Add and subtract; signed and unsigned.
Multiply and divide; signed and unsigned.
And, or, exclusive or.
Shifts: left and right logical, and right arithmetic.
Set conditional: "-" may be lt, le, eq, ne.

Control

beqz, bnez
end, begin

Conditional branches
Branch GPR equal/not equal to zero;
16-bit offset from PC + 1.
Halt, start computation.

Figure 5.2: List of DLX instructions selected for specification.

A sample of instructions from the different classes along with their intended semantics is given in figure 5.1. A list of the selected DLX instructions is given in figure 5.2.

5.2 Sequential specification for DLX architecture

In this section, I specify the sequential semantics of the DLX machine. The DLX instructions are of a very simple nature. In particular, no instruction can both perform an arithmetic operation and a memory operation. Consequently all instructions can be broken into five distinct parts: fetch the instruction to be executed, decode instruction, execute instruction, perform required memory-access, and write-result. The block diagram detailing the connectivity of the various units is shown in figure 5.3. In the figure only MAR (Memory Address Register) can set the address for a memory load/store, and only MDR (Memory Data Register), can send/receive data from memory. IR (Instruction Register), and PC (Program Counter), save the current instruction being executed and the address of the next instruction to be executed, respectively. The latches A, B and AOUT provide storage for the inputs and the outputs of the ALU. The block named CONTROL decodes the instruction in IR, and sets all the switches to generate the appropriate flow of data required
Figure 5.3: Block diagram for the connectivity of functional blocks in the DLX.
to execute the current instruction. The individual parts in the execution of an instruction are further elaborated below.

1. Instruction fetch (IF): \( \text{MAR} := \text{PC}; \text{IR} := \text{M}[	ext{MAR}] \)
   
   \textit{Operation}: Send out the PC and fetch the instruction from memory into the instruction register, \( \text{IR} \). PC is transferred to MAR because PC is connected to memory only via MAR.

2. Instruction decode/register fetch (ID): \( \text{A} := \text{Rs}1; \text{B} := \text{Rs}2; \text{PC} := \text{PC} + 1 \)
   
   \textit{Operation}: Decode the instruction and access the source registers from the register file. The PC is also incremented to point to the next instruction to be executed. Decoding is done in parallel with reading registers to the latches \( \text{A} \) and \( \text{B} \), because of the fixed format of the DLX instructions. Moreover, as the immediate argument occurs in the same bits in all DLX instructions, the sign-extended immediate, if needed, is also calculated in this step.

3. Execution/effective address (EX): The ALU operates on the operands, performing one of the following functions depending upon the DLX instruction type.

   - Memory reference: \( \text{MAR} := \text{A} + (\text{IR}_{16})^{16} || \text{IR}_{16..31}; \text{MDR} := \text{Rd} \)
     
     \textit{Operation}: The immediate is calculated by taking the upper 16-bits of the \( \text{IR} \) and filling the lower 16-bits of the immediate with the 16th bit of the \( \text{IR} \). The immediate is added to the latch \( \text{A} \). The destination register \( \text{Rd} \) is stored in \( \text{MDR} \), because GPRs can store data into memory only via \( \text{MDR} \). \( \text{IR}_{16..31} \) are the lower 16 bits of \( \text{IR} \), and \( (\text{IR}_{16})^{16} \) is the 16th bit of \( \text{IR} \) repeated 16 times.

   - ALU instruction: \( \text{AOUT} := \text{A op B} \)
     
     \textit{Operation}: The ALU performs the operation specified in the opcode on the value in latch \( \text{A} \) and on the value in \( \text{B} \). The result is stored in another latch called AOUT.

   - Branch: \( \text{AOUT} := \text{PC} + (\text{IR}_{16})^{16} || \text{IR}_{16..31}; \text{cond} := (\text{A op 0}) \)
     
     \textit{Operation}: The ALU computes the branch target address by adding the PC to the immediate. It then compares \( \text{A} \) to \( 0; \text{op} \) can be either \( = \) or \( \neq \).
4. Memory access/branch completion (MEM): The only DLX instructions active in this step are loads, stores and branches.

- Memory reference: $\text{MDR} := \text{M}[\text{MAR}]$ or $\text{M}[\text{MAR}] := \text{MDR}$
  
  *Operation*: For a load data comes from memory to MDR, and for a store data from MDR goes to memory.

- Branch: if (cond) $\text{PC} := \text{AOUT}$
  
  *Operation*: For branch instructions PC is updated if cond is 1.

5. Write-back (WB): $\text{Rd} := \text{AOUT or MDR}$

  *Operation*: Write the result into the register file, whether coming from the memory system or from the ALU.

The idea behind specifying DLX is pretty clear, given the above explanation of the operation semantics of the instructions. I look at the various registers and the memory array as entities in some common pool. The evaluation of the program consists of a synchronization between the PC, the registers, and the memory, which contains both the program and the data. Memory is represented by a two place predicate, $(m, n, V)$, the first argument is the address - an integer - of the memory cell and the second argument is the contents of the cell. Similarly, registers are represented by binary predicates $(r, i, V)$ and program memory by $(p, i, x)$. An instruction is represented as $(i, x, s_1, s_2, d, I, \text{op})$, where $X$ denotes the class of the instruction, $s_1$ the first register argument, $s_2$ the second register argument, $d$ the destination register, $I$ is an immediate and op the particular function to calculate. It should be obvious that all instructions can be represented with this representation, however, it is not necessary that all fields will have meaningful data for all the instructions.

The signature, called $\Sigma_x$, specifying the constants used in the sequential specification of DLX is given in figure 5.4. A function for every arithmetic/logical function of DLX is assumed in FORUM – I treat these functions as if they were in-built. Further, the type int is declared with all the integers as terms of type int. The built-in functions are represented by tokens which are members of func. Members of the type class represent the classes of instructions. The predicate cont signals that computation should continue, and num is a
predicate whose argument counts the number of instructions executed. \((\texttt{eq} N)\) is provable if and only if \(N = 1\), and \((\texttt{ne} N)\) is provable if and only if \(N = 0\).

\[
\begin{align*}
\text{cont} & : o \\
\text{num, eq, ne} & : \text{int} \rightarrow o \\
\text{op} & : \text{int} \rightarrow \text{int} \rightarrow \text{int} \\
\text{r, m} & : \text{int} \rightarrow \text{int} \rightarrow o \\
\text{a, b, pc} & : \text{int} \rightarrow o \\
\text{p} & : \text{int} \rightarrow \text{inst} \rightarrow o \\
\text{ix} & : \text{class} \rightarrow \text{int} \rightarrow \text{int} \rightarrow \text{int} \rightarrow \text{int} \rightarrow \text{func} \rightarrow \text{inst} \\
\text{op} & : \text{func} \\
\text{alu, br, ld, st, ht, bg, noop} & : \text{class}
\end{align*}
\]

Figure 5.4: Signature for specification of DLX

The specification for DLX, called \(E_s\), is the set of universal closure of clauses in figure 5.6. Different kinds of parentheses have been used in figure 5.6 to enhance readability. Note that for every instruction type – specified by its class – there is exactly one clause in figure 5.6 with a matching head. The one subtlety in the specification is that I cannot perform a multi-way synchronization amongst the operand registers, destination register, and PC for the ALU operations. The problem is that there is only one \(R_n\), while the instruction may need three copies – a deadlock would occur. The specification thus decouples the reading of the two source registers into individual unsynchronized steps. The idea is that the two source registers are read in any order into the latches \(A\) and \(B\). Once the latches are loaded the operation is performed and result is stored in the destination register. In this approach all of the register arguments may be identical or different – a deadlock will not occur.

The DLX programs in FORUM are defined by the terms parsed by the non-terminal \(P_l\), \(l \in \text{nat}\) in figure 5.5. A program is loaded in program memory, and then evaluated with respect to a state specified by the contents of the special purpose registers, general purpose registers, memory and PC – the data state. The definitions of program state and data state are given below.

**Definition 5.1 [Program State]** Let \(\mathcal{P}\) be an abbreviation for
\[
\lambda m, x_1, \ldots, x_m. (p \ 1 \ x_1) \circ \ldots \circ (p \ m \ x_m) \text{ for } m \in \text{nat}. \text{ For any dlx program } P_l, (\mathcal{P} \ l \ \bar{P}) \text{ is a}
\]
\[
R := n \quad n \in [1, 32]
\]

\[
C := \text{ld} | \text{st} | \text{alu}
\]

\[
I_{l', l, m, n} := q \quad q \in [l', l] \cup [0, m] \cup [-n, 0]
\]

\[
I_{x, l', l, m, n} := (\text{ix bg} R R R)_{l', l, m, n}\op
\quad | (\text{ix C} R R R)_{0, 0, 0}\op
\]

\[
B_{l, l, m, n} := I_{m, l-n-1, 0, n+2}
\]

\[
B_{l', l, m, n} := I_{m+\nu-1, l-n-1, \nu-1, n+2}; B_{l', l, l, m, n+1}
\]

\[
A_{l, l, m, n} := I_{-l+1, -(m+n+2), 0, n+1}
\]

\[
A_{l', l, m, n} := I_{-(l+2-\nu), -(m+n+2), \nu-1, -(n+1)}; A_{l', l, l, m, n+1}
\]

\[
H_{l, l} := I_{0, 0, -1, l}
\]

\[
H_{l', l} := I_{0, 0, \nu-2, l-1}; H_{l', l-1, l}
\]

\[
S := (\text{ix bg} R R R)_{0,0}\op
\]

\[
E := (\text{ix ht} R R R)_{0,0}\op
\]

\[
Q_{l+2} := S; B_{l+2, 0, 0}; E
\]

\[
Q_{l+2} := S; B_{l', l, m, 0}; H_{m}; A_{l', l, m, 0}; E
\]

\[
P_{l} := Q_{l} \ | \ Q_{l'}; P_{l'}
\]

Figure 5.5: Grammar for DLX programs

**Program state.** \(P\) are the instructions in \(P_{l}\), and by definition \(l\) is the number of instructions in \(P\).

**Definition 5.2** [Data State] \((r 1) \ldots (r 32)\) are the DLX registers, \(n \in \text{nat}\) is the number of memory cells, \(\text{pc}\) is the program counter, and \(\text{num}\) stores the number of instructions executed. Let \(S\) be an abbreviation for

\[
\lambda n, \bar{V}, \bar{U}, L. (\text{pc} L) \varphi (r 1 V_{1}) \varphi \ldots (r 32 V_{32}) \varphi (m 1 U_{1}) \varphi \ldots (m n U_{n}) \varphi (\text{num} U_{n+1}).
\]

The lengths of \(V\) and \(U\) are 32 and \(n + 1\) respectively – if implicit, are assumed to be as required.

For any \(n \geq 0, L, V_{1}, \ldots, V_{32}, U_{1}, \ldots, U_{n+1} : \text{int}, (S n \bar{V} \bar{U} L)\) is a data state.

It is worth pointing out that execution cannot start without executing \(\text{bg}\) as clauses for all instructions other than \(\text{bg}\) synchronize with \(\text{cont}\), and that the execution cannot terminate.
\[(\text{pc } L) \varphi (p L (\text{ix } \text{alu } S_1 S_2 D I O)) \varphi \text{cont} \varphi (\text{num } M) \leftarrow \]

\[(\text{pc } L) \varphi (p L (\text{ix } \text{alu } S_1 S_2 D I O)) \varphi (\text{num } (M + 1)) \leftarrow \]

\[\left[\left[ (r S_1 V_1) \leftarrow (r S_1 V_1) \varphi (a V_1) \right] \otimes (r S_2 V_2) \leftarrow (r S_2 V_2) \varphi (b V_2) \right] \otimes (r D V_3) \varphi (a V_1) \varphi (b V_2) \leftarrow (r D (V_1 O V_2)) \varphi \text{cont} \varphi (\text{pc } (L + 1))\right] \]

\[(\text{pc } L) \varphi (p L (\text{ix } \text{ld } S_1 S_2 D I O)) \varphi \text{cont} \varphi (\text{num } M) \leftarrow \]

\[(\text{pc } L) \varphi (p L (\text{ix } \text{ld } S_1 S_2 D I O)) \varphi (\text{num } (M + 1)) \leftarrow \]

\[\left[\left[ (r S_1 V_1) \leftarrow (r S_1 V_1) \varphi (a V_1) \right] \otimes (r D V_2) \varphi (a V_1) \varphi (m (V_1 + I) V_3) \leftarrow (r D V_3) \varphi \text{cont} \varphi (\text{pc } (L + 1)) \varphi (m (V_1 + I) V_3)\right] \]

\[\varphi (p L (\text{ix } \text{st } S_1 S_2 D I O)) \varphi \text{cont} \varphi (\text{num } M) \leftarrow \]

\[\varphi (p L (\text{ix } \text{st } S_1 S_2 D I O)) \varphi (\text{num } (M + 1)) \leftarrow \]

\[\left[\left[ (r S_1 V_1) \leftarrow (r S_1 V_1) \varphi (a V_1) \right] \otimes (r D V_2) \varphi (a V_1) \varphi (m (V_1 + I) V_3) \leftarrow (r D V_2) \varphi \text{cont} \varphi (\text{pc } (L + 1)) \varphi (m (V_1 + I) V_2)\right] \]

\[\varphi (p L (\text{ix } \text{br } S_1 S_2 D I O)) \varphi (r S_1 V_1) \varphi \text{cont} \varphi (\text{num } M) \leftarrow \]

\[\left[\left[ (\text{eq } (V_1 O 0)) \otimes (\text{pc } (L + 1))\right] \oplus [(\text{ne } (V_1 O 0)) \otimes (\text{pc } (L + 1))]\right] \]

\[(\text{pc } L) \varphi (p L (\text{ix } \text{ht } S_1 S_2 D I O)) \varphi \text{cont} \varphi (\text{num } M) \leftarrow \]

\[(\text{pc } L + 1) \varphi (p L (\text{ix } \text{ht } S_1 S_2 D I O)) \varphi (\text{num } (M + 1)) \]

\[(\text{pc } L) \varphi (p L (\text{ix } \text{bg } S_1 S_2 D I O)) \varphi (\text{num } M) \leftarrow \]

\[(\text{pc } L + 1) \varphi (p L (\text{ix } \text{bg } S_1 S_2 D I O)) \varphi \text{cont} \varphi (\text{num } (M + 1)) \]

\[\text{eq } 1 \leftarrow 1 \]

\[\text{ne } 0 \leftarrow 1 \]

Figure 5.6: Sequential specification of DLX
successfully without executing ht because the clause for ht is the only one that consumes cont. Hence, the value of pc in the initial data state must address a bg instruction in the program state. The program state is static—its contents do not change during execution—it instructs the machine how to alter the data state. Thus, the evaluation, given a program state, transforms one data state into another. The definition of evaluation is made precise below.

**Definition 5.3** [Sequential evaluation in DLX, \(E_x\)] Given data states \(S_1\) and \(S_2\), and \((P \ l \ \vec{l})\), a program state. \(P_l\) evaluates in \(S_1\) to \(S_2\) written as \(S_1 \varphi (P \ l \ \vec{l}) \rightarrow S_2 \varphi (P \ l \ \vec{l})\), if

\[
\Sigma : E_x; S_2 \varphi (P \ l \ \vec{l}) \rightarrow S_1 \varphi (P \ l \ \vec{l})
\]

is provable in FORUM. ■

A small example of an evaluation will explain the specification better. I consider a program that adds the contents of second and third register, and places the result in the second register. The program is stored in the memory starting at the first cell. The data state, and the program state are described in figure 5.7. I use \(S_1\) as an abbreviation for the entire expression in the figure. At the end of the computation the resultant state, \(S_2\) will have 9 in the second register, 4 in the pc, 3 in the num and otherwise be identical to \(S_1\). The proof of the computation is detailed below.

\[
(pc \ 0) \varphi (r \ 1 \ 0) \varphi (r \ 2 \ 4) \varphi (r \ 3 \ 5) \varphi \ldots \varphi (r \ 32 \ V_{32}) \varphi (m \ \overrightarrow{U}_n) \varphi (num \ 0) \varphi \\
(p \ 1 (ix \ bg \ 1 \ 2 \ 3 \ 0 +)) \varphi (p \ 2 (ix \ alu \ 2 \ 3 \ 2 \ I +)) \varphi (p \ 3 (ix \ ht \ 1 \ 2 \ 3 \ 0 +)) \varphi (p \ m \ \overrightarrow{W}_m)
\]

Figure 5.7: Example program in DLX

Let \(S'_1\) be identical to \(S_1\), except that it does not contain pc and num, and \(S''_1\) be identical to \(S_1\), except that it does not contain pc, \(r \ 2\) and num. I begin the proof by backchaining on bg clause, which generates cont. Next, I backchain on the alu clause, and then the only clauses that I can use for backchaining are the ones introduced by the alu clause to read
the two source operands. Hence, I read second and third registers into a and b, respectively.
Next the actual computation is performed using the values in the latches, and the value of \((r_2)\) is updated to 9. To complete the proof, I have to construct \(\delta\). \(E_s\) and \(\Sigma_s\) are not written in the proofs for the sake of readability. It is also worth pointing out that neither provability nor the actual answers computed depend upon the order in which the source registers are read.

\[
\begin{align*}
S_2 &\rightarrow S''_2 \varphi(r_29) \varphi\text{cont}\varphi(pc3) \varphi(\text{num}2) \\
S_2, (r_24) \varphi(a4) \varphi(b5) &\leftarrow (r_29) \varphi\text{cont}\varphi(pc3) \rightarrow S'_1 \varphi(b5) \varphi(a4) \varphi(\text{num}2) \\
(r_24) \varphi(a4) \varphi(b5) &\leftarrow (r_29) \varphi\text{cont}\varphi(pc3), \\
S_2, (r_24) &\leftarrow (r_24) \varphi(a4) \rightarrow S'_1 \varphi(b5) \varphi(\text{num}2) \\
(r_24) \varphi(a4) \varphi(b5) &\leftarrow (r_29) \varphi\text{cont}\varphi(pc3), \\
(r_35) &\leftarrow (r_35) \varphi(b5), \\
S_2, (r_24) &\leftarrow (r_24) \varphi(a4) \rightarrow S'_1 \varphi(\text{num}2)
\end{align*}
\]

\[
\begin{align*}
\text{;alu clause, } &-R, \otimes -L \\
S_2 &\rightarrow (pc2) \varphi S'_1 \varphi\text{cont}\varphi(\text{num}1) \\
\Sigma_s : E_s ; S_2 &\rightarrow S_1
\end{align*}
\]

To construct \(\delta\) notice that by definition of \(S_2\), the state \(S''_2 \varphi(r_29) \varphi(pc4) \varphi(\text{num}3)\) is identical to \(S_2\) up to associativity and commutativity of \(\varphi\) — the states are being treated as multisets. Hence, \(\delta\) is constructed by first backchaining on the ht clause, and then using the identity rule. The proof is detailed below. \(E_s\) and \(\Sigma_s\) are not written in the proofs for the sake of readability.

\[
\begin{align*}
\Sigma_s : E_s ; S_2 \delta \rightarrow S_2 \varphi P \rightarrow S_2 \varphi P \\
\Sigma_s : E_s ; S_2 \delta \rightarrow S_2 \varphi P \\
\Sigma_s : E_s ; S_3 \varphi P \rightarrow S_1 \varphi P
\end{align*}
\]

One consequence of definition 5.3 is that evaluation can be composed using cuts. Suppose \(S_1, S_2\) and \(S_3\) are three states, and \(P\) is a program state such that \(S_1 \varphi P \leftarrow S_2 \varphi P\) and \(S_2 \varphi P \leftarrow S_3 \varphi P\). The computation for \(S_1 \varphi P \leftarrow S_3 \varphi P\) is obtained by a linear cut on \(\delta_1\) and \(\delta_2\), which are provided by definition 5.3.

\[
\begin{align*}
\Sigma_s : E_s ; S_2 \delta_2 &\rightarrow S_2 \varphi P \\
\Sigma_s : E_s ; S_3 \varphi P &\rightarrow S_1 \varphi P
\end{align*}
\]
Hence, proofs of evaluations for the DLX machine will be composed using cuts – I use the cut-elimination theorem for FORUM. The main point is that the computation of an entirely imperative program is being represented declaratively and analyzed by proof theoretic tools such as cut-elimination.

5.3 Pipelining DLX - facing the hazards

Pipelining is an implementation technique in which multiple instructions are overlapped in execution. Today, pipelining is a key technique used to make fast CPU’s (Central Processing Unit). The basic idea of pipelining was first implemented in IBM 7030 [Blo59, Buc62]. The CDC 6600 [Tho70] and IBM 360/91 [AST89] introduced many important concepts in pipelining, including scoreboard, use of multiple functional units, simple architecture for efficient pipelining, and tagging of data, dynamic memory hazard resolution, and generalized forwarding. With the advent of RISC [AC87], efficient pipelining and compilation became integral parts of the architecture design. Many new ideas and designs for RISC architectures have been explored in the last decade, resulting in the design of important machines such as the Intel i860, MIPS R2000/R3000, Motorola 88000, SPARC, PowerPC. As we see in figure 5.8, all instructions in DLX have five distinct parts. The intention is to execute the five distinct parts of five instructions at the same time, and complete the execution of an instruction every time unit. Since the time unit for the pipeline will be determined by the execution time of its slowest segment it is important to have similar work loads for the different stages of the pipeline. Furthermore, since the performance gain for the pipeline will be maximized if all stages of the pipeline are kept busy it is important to minimize stalls in the pipeline.

Pipelining exploits the simple nature of the DLX architecture which facilitates well balanced pipeline segments, and early detection and elimination of possible stalls in the pipeline. Although the intentions are concisely stated, the design of pipelines is a tight rope walk, balancing various parameters to improve performance. The DLX instructions are not seen as atomic entities any longer - the five stages, i.e. IF, ID, EX, MEM and WB, are the atomic
activities. The parallel execution of these atomic steps for different instructions requires complex control to ensure that the results computed are identical to the ones computed by \( \varepsilon \).

I view pipelining as an alternate operational semantics for DLX programs—a different evaluator for DLX programs. The declarative specification of the pipelines which can be executed to simulate pipelined computation has not been attempted to the best of my knowledge.

The thrust of the existing work has been either to verify the correctness of a pipelined processor all the way down to the circuit level [TK93], or to automate the production of control circuitry from high level descriptions of pipelines [AL93]. My goal is to provide an executable and declarative specification of pipelines explaining the intricate synchronization required to implement the basic concepts in modern pipelined architectures.

In this section, I begin by introducing the basic pipeline structure for DLX, and structural and data hazards. The pipeline is then specified using external functions to resolve the hazards. The specification in this section does not incorporate any of the standard techniques...
to improve the CPU throughput. The implementation of call-forwarding and early branch resolution will be the subject of the next section. Furthermore, I prove that the pipeline specification is equivalent to the sequential specification, $E$.

**Hazards in Pipeline**

I model IF, ID, EX, MEM and WB and the clock as cyclic processes. The clock generates signals prompting the stages to start their respective computations. When all the stages complete their designated computations they notify the clock, which once again generates the enabling signal for the stages, and the cycle continues. Although there is no synchronization amongst different stages, the computation for every stage is synchronized via the clock signal.

Once the processes for the clock and the pipeline stages start running in parallel, it is possible to run into a variety of problems. Imagine that a machine can only write/read to one register at a time. In such a machine executing the WB stage and ID stage of two alu instructions would cause a contention for register port. The point is that when execution of stages is done in parallel, certain resources need to be duplicated to avoid conflicts over resource usage. Existence of such a conflict can stall the pipeline because one instruction will have to wait while the other uses the resource, causing the entire pipeline to waste one time unit. Such conflicts are called structural hazards. *Informally, if there is a combination of instructions which causes contention for resources, then the machine is said to have structural hazards.*

In the DLX machine that I outline there will be no structural hazards, essentially because all instructions are assumed to complete in one time unit and resources have been duplicated sufficiently. Although eliminating structural hazards from pipelines yields better throughput, designers sometimes allow structural hazards since, either duplicating resources is too expensive or eliminating them results in a larger time unit for the pipeline.
The concurrent execution of parts of DLX instructions alters the relative timing of instructions. *In the sequential specification an instruction is executed only after the preceding instruction is completed*, whereas, *in the pipeline there are up to five instructions whose various parts are being computed at the same time.* So for example (*ix ld 12 2 Iop*) could be in the EX stage while (*ix alu 12 4 I+) is in the ID stage. Now I have a problem, since (*ix alu 12 4 I+) reads the value in R2 before the ld has fetched the contents of memory addressed by R1 and stored them in R2, which happens at the end of the WB stage (figure 5.8). The problem is that there is a data-dependence in the sequence of instructions above. It causes no problems for E+, because alu starts only after the completion of ld. Thus the pipeline must be stalled while alu waits for ld to finish. Such problems are significantly eliminated by code rescheduling and forwarding [HP90].

**Definition 5.4** [Data Hazard] If the order of access to operands by instructions is changed due to overlapping execution in pipeline, then there is a data hazard. Data Hazards are classified by the order in which the read and writes are supposed to occur in the program. Let i be an instruction occurring before j in the execution.

- **RAW** (read after write): j tries to read a source before i writes it, and thus gets the old value.
- **WAR** (write after read): j tries to write a destination before i has read it, and thus i gets the new value.
- **WAW** (write after write): j tries to write an operand before it is written by i, and thus the writes are performed in the wrong order.

Note that RAR (read after read) is not an error. In the DLX pipeline stages all instructions complete in one time unit, and registers are read early at ID as compared to written later at WB. Moreover the memory accesses are kept in order, and hence no WAR hazard is
possible in DLX. Since the WAW hazard is only possible in pipelines that write in more than one stage of the pipeline, this hazard is avoided by DLX which writes registers only in the WB stage. The only kind of data hazard in the DLX pipeline is RAW, as exhibited in the example above regarding the ld and alu instructions.

The concurrent execution of instructions causes another kind of synchronization problem. Suppose \((ix\ br\ 12\ 4\ i =)\) enters the ID stage. The IF stage now needs to fetch the next instruction to be executed, but this is not possible before the branch instruction is resolved and new pc is available at the end of the MEM stage, figure 5.8. Moreover, at the beginning of the IF stage it is not known whether the previous instruction was a br or not. Thus the IF fetches the next instruction, and invalidates the fetched instruction if the preceding instruction is a br instruction. The pipeline is stalled until the end of the MEM stage for the br instruction when the new pc is available. The fact that I fetch and invalidate an instruction introduces wasteful computation and is different in nature from the stall introduced by data hazards. Informally, if the pipeline is stalled for a br instruction to calculate the new pc, then there is a control hazard.

Even in the simple setting of the integer DLX architecture hazards and stalls arise due to the concurrent nature of the pipeline. The presence of instructions which do not complete in one time unit and interrupts further complicates the control circuitry. However, the specification techniques used to resolve the hazards for the simple unoptimized pipeline present the ideas needed to tackle the various synchronization and control issues that arise in the specification. To keep the presentation concise and convey the essential ideas, I chose to specify only the integer DLX with call-forwarding and early-branch resolution, and not deal with floating-point arithmetic and interrupts. Floating-point arithmetic and interrupts are challenging aspects of pipeline design, the comments above are only with respect to the specification techniques used to specify the pipeline design for these features.
Specifying the pipeline

Given the hazards in DLX pipeline, it will not be enough to specify the processes representing the pipeline stages and the clock. The crucial ingredient in the specification is the synchronization which avoids all the hazards that may occur during execution. The resource conscious language of FORUM provides an ideal setting to specify the complex interactions required to specify the pipeline.

I begin by giving the signature, $\Sigma_p$, for the specification of the pipeline. The signature is the union of $\Sigma_s$ and the constants in figure 5.9. $lr$ stores the data to be written back to the register file. The $\text{cond}$ predicate stores conditional value for branch instructions, $\text{mar}$ stores the memory address to be accessed, $\text{lmdr}$ stores the data read from the memory and $\text{smdr}$ stores the data to be written into memory. The $\text{aout}$ predicate stores the ALU result at the end of the EX, and $\text{aout}_1$ stores the ALU result at the end of the MEM. The $\text{pc}_1$ predicate stores the old pc to EX, $ir$ stores the instruction, and $it$ stores the class, destination, immediate and operator of the instruction. The predicates with a 1 before the above predicate names are latches which hold temporary values in between reading and writing to the registers. The beginning of the read and write phase for IF are signalled by $\text{ifrb}$ and $\text{ifwb}$, respectively, while $\text{ifrd}$ and $\text{ifwd}$ signal the completion of the read and write phase, respectively, for IF. Similarly there are predicates for ID, EX, MEM and WB prefixed by their names. The class of an instruction is tested by $\text{alu}^?$, $\text{ld}^?$, $\text{st}^?$, $\text{br}^?$ and $\text{noop}^?$ predicates.

The following definitions make the presentation more concise. I sometime use $\bar{n}$ as an abbreviation for $n_1, n_2, n_3, n_4, n_5$.

$$
\begin{align*}
\text{crb} & = \text{def} \lambda \bar{n}. (\text{ifrb} n_1) \lor (\text{idrb} n_2) \lor (\text{exrb} n_3) \lor (\text{merb} n_4) \lor (\text{wbrb} n_5) \\
\text{crd} & = \text{def} \lambda \bar{n}. (\text{ifrd} n_1) \lor (\text{idrd} n_2) \lor (\text{exrd} n_3) \lor (\text{merd} n_4) \lor (\text{wbrd} n_5) \\
\text{cwd} & = \text{def} \lambda \bar{n}. (\text{ifwd} n_1) \lor (\text{idwd} n_2) \lor (\text{exwd} n_3) \lor (\text{mewd} n_4) \lor (\text{wbwd} n_5) \\
\text{cwb} & = \text{def} \lambda \bar{n}. (\text{ifwb} n_1) \lor (\text{idwb} n_2) \lor (\text{exwb} n_3) \lor (\text{mewb} n_4) \lor (\text{wbwb} n_5)
\end{align*}
$$
\[ lr : \text{int} \to \text{int} \to o \]
\[
\text{cond, mar, lmdr, smdr} : \text{int} \to o \\
\text{lcond, lmar, llmdr, lsmdr} : \text{int} \to o \\
\text{aout, aout} : \text{int} \to o \\
\text{la, lb, lacut, lacut} : \text{int} \to o \\
\text{pc, lpc} : \text{int} \to o \\
\text{ir, lir} : \text{inst} \to o \\
\text{it, lit} : \text{class} \to \text{int} \to \text{func} \to o \] 
\[ i \in \{1, 2\} \]
\[
\text{ifrb, idrb, exrb, merb, wbrb} : \text{int} \to o \\
\text{idrd, idrd, exrd, merd, wbdrd} : \text{int} \to o \\
\text{ifwdb, idwb, exwb, mewb, wbwb} : \text{int} \to o \\
\text{ifwd, idwd, exwd, mewd, wbwd} : \text{int} \to o \\
\text{alu?, ld?, st?, br?, noop?} : \text{class} \to o \]

Figure 5.9: Signature for specification of DLX

I sometimes use \((\text{crb} \, \bar{n})\) as abbreviation for \((\text{crb} \, n_1 \, n_2 \, n_3 \, n_4 \, n_5)\), and similarly for \text{crd}, \text{cwb} and \text{cwd}. Thus \((\text{crb} \, \bar{n})\), is the signal for the read begin phase for the stages of the pipeline, and the numbers \(n_i, i \in [1, \ldots, 5]\) denote different states of the signal in question. \(\tau_1, \tau_2, \tau_3, \tau_4, \tau_5\), specified in figure 5.10, are the transition functions for \text{IF}, \text{ID}, \text{EX}, \text{MEM} and \text{WB} stages, respectively. The states of the five signals in the clock are arguments to each of the transition functions, and the output is the state of the clock signal for its stage. In figure 5.10, \(n_i \in \{0, 1\}, i \in [3, \ldots, 5]\), and for any input not exhibited the functions \(\tau_1, \tau_2, \tau_3, \tau_4, \tau_5\) return 1. For the \text{ID} clock signal, \(i \in \{-1, -2, -3\}\) are the states for a data hazard where the stalled instruction has to wait for \(-i\) cycles, \(i \in \{-4, -5\}\) are the states for a control hazard, and \(i = 0\) is the state when the pipeline will stop within the next five cycles. For example, from figure 5.10, \(\tau_1(-3)100 = 0\), thus \text{IF} will remain idle in the next cycle. I sometimes use \((\text{crb} \, \tau(\bar{n}))\) as an abbreviation for \((\text{crb} \, \tau_1(\bar{n}) \, \tau_2(\bar{n}) \, \tau_3(\bar{n}) \, \tau_4(\bar{n}) \, \tau_5(\bar{n}))\), and similarly for \text{crd}, \text{cwb} and \text{cwd}.

The specification for the pipeline, \(\mathcal{E}_p\), is the set of universal closures of clauses in figures 5.12, 5.13, 5.14 and 5.15. In the \text{ID} stage in figure 5.12, \text{RAW} and control hazards are detected by the function \(\delta_i\), whose return value sets the state for \text{idrd}. \(\delta_i\) is defined below.

**Definition 5.5** \([\delta_i - \text{Hazard detection function for } \mathcal{E}_p]\)
\[
(\delta_i S_1, S_2, D_i, C) = \min((\rho_i S_1, D_i, \bar{C}i, C), (\rho_i S_2, D_i, \bar{C}i, C)) = (\rho_i S_1, D_i, \bar{C}i, C)
\]

\[C \in \{\text{alu, st}\}\]

\[\text{otherwise}\]

Where, \(\rho_i\) is the function in figure 5.11, \(C\) is the instruction type, and \(S_1\) and \(S_2\) are the two source registers of the current instruction. \(C_i\) is the instruction type and \(D_i\) is the destination register of the \(i\)th preceding instruction, for \(i \in \{1, 2, 3\}\).

Note the difference between the control hazard and data hazard. In the case of a data hazard the instruction signaling the hazard and the ones following it are stalled. However, in the case of a control hazard, the \(\text{br}\) instruction which signals the hazard continues execution while the instructions following it are stalled. Thus, if there is both a control and data hazard, the data hazard must be processed first. The table for \(\rho_1\) is given in figure 5.11. In figure 5.11, \(A \in \{\text{alu, br, ld, st}\},\ Y \in \{\text{alu, br, ld, st, \text{noop, bg}\},\ \) and \(Z \in \{\text{alu, ld}\}\). Furthermore, \(=\) in the column for \(D_i\) means that the destination register \(D_i\) is the source
register \( S_j \neq j \) means that the destination register \( D_i \) is different from the source register \( S_j \), and \( X_j \) means either \( \neq j \) or \( \neq j \). If the current instruction is a noop or the destination registers of the three preceding instructions are different from the source registers, then there is no hazard. A data hazard is detected if any of the previous three instructions write either \( S_1 \) or \( S_2 \). A control hazard is signaled if the current instruction is a branch and there is no data hazard. If the current instruction is \( \text{ht} \), then the pipeline is stopped after the execution of \( \text{ht} \) is finished.

<table>
<thead>
<tr>
<th>( C )</th>
<th>( C_1 )</th>
<th>( C_2 )</th>
<th>( C_3 )</th>
<th>( D_1 )</th>
<th>( D_2 )</th>
<th>( D_3 )</th>
<th>( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>noop</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>( X_j )</td>
<td>( X_j )</td>
<td>( X_j )</td>
<td>1</td>
</tr>
<tr>
<td>( A )</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>( \neq j )</td>
<td>( \neq j )</td>
<td>( \neq j )</td>
<td>1</td>
</tr>
<tr>
<td>( A )</td>
<td>Z</td>
<td>Y</td>
<td>Y</td>
<td>( = j )</td>
<td>( X_j )</td>
<td>( X_j )</td>
<td>-3</td>
</tr>
<tr>
<td>( A )</td>
<td>Y</td>
<td>Z</td>
<td>Y</td>
<td>( \neq j )</td>
<td>( = j )</td>
<td>( X_j )</td>
<td>-2</td>
</tr>
<tr>
<td>( A )</td>
<td>Y</td>
<td>Y</td>
<td>Z</td>
<td>( \neq j )</td>
<td>( \neq j )</td>
<td>( = j )</td>
<td>-1</td>
</tr>
<tr>
<td>( \text{br} )</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>( \neq j )</td>
<td>( \neq j )</td>
<td>( \neq j )</td>
<td>-4</td>
</tr>
<tr>
<td>( \text{ht} )</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>( X_j )</td>
<td>( X_j )</td>
<td>( X_j )</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 5.11: \( \rho_1 \) - table for hazard detection in the DLX pipeline

The complex looking clauses in the \( \mathcal{E}_p \) warrant some explanation. The clock is specified by the first two clauses in figure 5.12. It is a cyclic process consuming the completing signals for read(write) to enabling signals for write(read). The clock stops if the state for \( \text{wbwd} \) is \(-1\).

The state information of the clock is calculated by \( \mathcal{\tau} \) defined in figure 5.10. Each of the five stages of the pipeline – \( \text{IF}, \text{ID}, \text{EX}, \text{MEM} \) and \( \text{WB} \) – are implemented by two clauses. One of the clauses synchronizes with the read begin signal, reads data into temporary latches, and produces the read done signal. The other clause synchronizes with the write begin signal, consumes data in the temporary latches, and produces the write done signal. The actions performed by a stage depend on the type of the instruction being processed and the state of the clock. \textit{Note that a register might be updated in one stage and read in another stage during the same cycle.} To ensure availability of proper data, in the read phase required data is stored into latches, and only after all data is read, registers and memory may be updated in the write phase.

The \( \text{bg} \) instruction is handled specially as the clock is started by this instruction. This
instruction also loads several latches in the environment needed by the pipeline. I use $L$ as an abbreviation for the various latches as defined below, and $(L \bar{C} \bar{D} \bar{I} \bar{S} \bar{V})$ is a latch state. $L_0$ is the initial state for these latches. When the lengths of the vectors are not mentioned explicitly, they are of the required length, as assumed in the definition below.

$$L = \text{def} \lambda \bar{C}, \bar{D}, \bar{I}, \bar{S}, \bar{V}. (\varphi_{i \in \{1, 2, 3\}}(\text{it}_i C_i D_i I_i O_i)) \varphi(\text{ix} (\text{ix} C_4 S_1 S_2 D_4 I_4 O_4)) \varphi(a V_1) \varphi(b V_2) \varphi(\text{aout}_1 V_3) \varphi(\text{aout}_1 V_4) \varphi(\text{pc}_1 V_5) \varphi(\text{mar} V_6) \varphi(1\text{mdr} V_7) \varphi(\text{smdr} V_8) \varphi(\text{cond} V_8)$$

$$L_0 = \text{def} (L \text{n\text{\textop}} \bar{0} \bar{0} \bar{0} \bar{0} \bar{0} \bar{0} \bar{0})$$

Now, evaluation of DLX programs can be defined using $E_p$. The idea behind evaluation is the same as that for $E_s$. A program is loaded in the program memory, and then the program state is evaluated in a given data state. The program state remains static, while the data state may possibly change due to the execution of instructions. The definitions of program and data state are taken from definitions 5.1 and 5.2.

**Definition 5.6** [Pipelined evaluation in DLX, $E_p$] Given data states $S_1$ and $S_2$, and $(P \triangleright \bar{P}_l)$, a program state, $P_l$ evaluates in $S_1$ to $S_2$ written as $S_1 \varphi(P \triangleright \bar{P}_l) \longrightarrow_{p} S_2 \varphi(P \triangleright \bar{P}_l)$, if

$$\Sigma_p : E_p; S_2 \varphi(P \triangleright \bar{P}_l) \longrightarrow S_1 \varphi(P \triangleright \bar{P}_l)$$

is provable in FORUM.

I use the example program in figure 5.7 to illustrate the specification $E_p$. I use $S_1$ as an abbreviation for the entire expression in the figure. At the end of the computation, the resultant state, $S_2$ will have 9 in the second register, 4 in the pc, 3 in the num and otherwise be identical to $S_1$. The proof of the computation is detailed below. Let $S_1'$ be identical to $S_1$, except that it does not contain pc and num, and $S_1''$ be identical to $S_1$ except that it does not contain pc, (r 2 ) and num.
(crd \( \vec{N} \)) \( \land \) (cwb \( \vec{N} \))

(cwb \( \vec{N} \)) \( \Rightarrow \)

\[
[(\text{ne}(N_5 = -1)) \otimes (\text{crb} \, \tau(\vec{N}))] \oplus \\
[(\text{eq}(N_5 = -1)) \Rightarrow [1 \, \varphi((L \, \vec{C} \, \vec{D} \, \vec{T} \, \vec{S} \, \vec{V}) \land 1)]]
\]

\((p \, L (i x \, b g \, D \, S_1 \, S_2 \, I \, O)) \varphi(pc \, L) \varphi(num \, M) \Rightarrow \)

\((p \, L (i x \, b g \, D \, S_1 \, S_2 \, I \, O)) \varphi(pc(L + 1)) \varphi(crb \, \vec{I}) \varphi(\mathcal{L}_0 \varphi(num(m + 1))) \)

(ifrb \( N \)) \( \varphi(p \, L \, V) \varphi(pc \, L) \Rightarrow (pc \, L) \varphi(p \, L \, V) \varphi \)

\[
[[[eq(N = 1)] \otimes ([\, \text{lir} \, V) \varphi(\, lpc(L + 1)) \varphi(ifrd \, N)]]] \oplus \\
[[eq(N = -1)] \otimes ([\, \text{lir} \, (i x \, \text{noop} \, 01 \, 20 \, +) \varphi(\, lpc(L - 1)) \varphi(ifrd \, N)]]] \oplus \\
[[eq(N = 0)] \otimes (ifrd \, N)]]
\]

(ifwb \( N \)) \( \Rightarrow \)

\[
[[eq(N = 0)] \otimes (ifwd \, N)] \oplus \\
[[eq(N = 1)] \oplus (eq(N = -1))] \Rightarrow \\
[1 \varphi((ir \, V^\prime) \varphi(pc \, L) \varphi(ir \, V) \varphi(lpc \, L) \Rightarrow (ir \, V) \varphi(pc \, L) \varphi(ifwd \, N))] \]

(idrb \( N \)) \( \varphi(\varphi_i \in \{1, 2, 3\}(i t_1 \, C; \, D; \, I; \, O_i)) \varphi(pc \, L) \varphi(ir \, (i x \, C \, D \, S_1 \, S_2 \, I \, O)) \Rightarrow \)

\[
[(\text{ne}(N = 1)) \otimes (idrd \, N)] \oplus \\
[(eq(N = 1)) \otimes \\
[[[eq(u = 1)] \oplus (eq(u = -4))] \otimes ((\text{lit}_1 \, C \, D \, I \, O) \varphi(\, lpc_1 \, L)) \Rightarrow \\
[[r \, S_1 \, V_1] \Rightarrow (r \, S_1 \, V_1) \varphi(\, la \, V_1) \varphi(S) \otimes \\
[r \, S_2 \, V_2] \Rightarrow (r \, S_2 \, V_2) \varphi(\, lb \, V_2) \varphi(S) \otimes \\
[S \varphi S] \Rightarrow (idrd \, u)]]] \oplus \\
[[eq(u = 1)] \otimes (eq(u = -4)] \otimes (idrd \, u)]])]
\]

Where \( u = \text{df} (\, \vec{b} \, \vec{S}_1 \, \vec{S}_2 \, \vec{D}; \, \vec{C}^2; \, \vec{C}) \)

(idwb \( N \)) \( \varphi(it_1 \, C \, D \, I \, O) \Rightarrow \)

\[
[[eq(N = 1)] \otimes (eq(N = -4)] \otimes ((it_1 \, \text{noop} \, D \, I \, O) \varphi(\, idwd \, N))] \oplus \\
[[eq(N = 1)] \oplus (eq(N = -4))] \Rightarrow \\
[[a \, W_1] \varphi(\, la \, V_1) \varphi(b \, W_2) \varphi(\, lb \, V_2) \varphi(\, pc_1 \, L_1) \varphi(\, lpc_1 \, L_2) \varphi(\, \text{lit}_1 \, C' \, D' \, I' \, O') \Rightarrow \\
(a \, V_1) \varphi(b \, V_2) \varphi(\, pc_1 \, L_2) \varphi(it_1 \, C' \, D' \, I' \, O') \varphi(\, idwd \, N)]]
\]

Figure 5.12: Specification for the DLX pipeline - clock, bg, IF and ID.
(alu? alu) ← 1
(ld? ld) ← 1
(st? st) ← 1
(br? br) ← 1
(noop? noop) ← 1

(exwb N)⊗(it1 C D I O) ← (lit1 C D I O)⊗(lit2 C D I O)⊗
[(ne(N = 1)) ⊗ (exrd N)]⊕
[(eq(N = 1)) ⊗
[(alu? C) ←
((a V1) ⊗ (b V2) ← (a V1) ⊗ (b V2) ⊗ (lmar (V1 O V2)) ⊗ (exrd N)) ⊗ 1]
[(ld? C) ←
((a V1) ← (a V1) ⊗ (lmar (V1 + I)) ⊗ (exrd N)) ⊗ 1]
[(st? C) ←
((a V1) ⊗ (b V2) ←
(a V1) ⊗ (b V2) ⊗ (lmar (V1 + I)) ⊗ (lsmdr V2) ⊗ (exrd N)) ⊗ 1]
[(br? C) ←
((pc L) ⊗ (a V1) ←
(pc1 L) ⊗ (a V1) ⊗ (lmar (L + I)) ⊗ (lcond (V1 O 0)) ⊗ (exrd N)) ⊗ 1]
[(noop? C) ⊗ (exrd N)]]

(exwb N)⊗(it2 C D' I' O')⊗(lit2 C D I O) ← (lit2 C D I O)⊗
[(ne(N = 1)) ⊗ (exwd N)]⊕
[(eq(N = 1)) ⊗
[(alu? C) ←
((aout V1) ⊗ (lmar V2) ← (aout V2) ⊗ (exwd N)) ⊗ 1]
[(ld? C) ←
((mar L1) ⊗ (lmar L2) ← (mar L2) ⊗ (exwd N)) ⊗ 1]
[(st? C) ←
((mar L1) ⊗ (lmar L2) ⊗ (smdr W1) ⊗ (lsmdr W2) ←
(mar L2) ⊗ (smdr W2) ⊗ (exwd N)) ⊗ 1]
[(br? C) ←
((aout V1) ⊗ (lmar V2) ⊗ (cond U1) ⊗ (lcond U2) ←
(aout V2) ⊗ (cond U2) ⊗ (exwd N)) ⊗ 1]
[(noop? C) ⊗ (exwd N)]]

Figure 5.13: Specification for the DLX pipeline – EX.
\[(\text{merb } N) \circ (\text{it} \text{ } 2 \ C \ D \ I \ O) \leftarrow \circ (\text{it} \text{ } 2 \ C \ D \ I \ O) \circ \ (\text{lit} \text{ } 3 \ C \ D \ I \ O) \circ \]
\[[(\text{ne} \ (N = 1)) \land (\text{merd } N)] \oplus \]
\[[(\text{eq} \ (N = 1)) \land \]
\[[(\text{alu} \ ? \ C) \leftarrow \]
\[((\text{aout } V_1) \leftarrow (\text{aout } V_1) \circ (\text{laout}_1 V_1) \circ (\text{merd } N)) \circ \ 1] \oplus \]
\[[(\text{ld} \ ? \ C) \leftarrow \]
\[((\text{mar} \ L) \circ (\text{m } L \ V) \leftarrow \]
\[((\text{mar} \ L) \circ (\text{m } L \ V) \circ (\text{llmdr } V) \circ (\text{merd } N)) \circ \ 1] \oplus \]
\[[(\text{st} \ ? \ C) \leftarrow \]
\[((\text{mar} \ L) \circ (\text{m } L \ V_2) \circ (\text{smdr } V_2) \circ (\text{merd } N)) \circ \ 1] \oplus \]
\[[\text{noop} \ ? \ C) \leftarrow \]
\[((\text{mar} \ L) \circ (\text{m } L \ V_2) \circ (\text{smdr } V_2) \circ (\text{merd } N)) \circ \ 1] \oplus \]
\[[\text{br} \ ? \ C) \leftarrow \]
\[((\text{cond} \ M) \circ (\text{aout } V_1) \circ (\text{pc } L_2) \leftarrow \]
\[[((\text{ne} \ (M = 1)) \land (\text{lpc} \ 2 \ L_2) \circ (\text{merd } N))] \oplus \]
\[[((\text{eq} \ (M = 1)) \land (\text{lpc} \ 2 \ V_1) \circ (\text{merd } N))] \]
\[\circ (\text{cond} \ M) \circ (\text{aout } L_1) \circ (\text{pc } L_2) \circ \ 1] \oplus \]
\[[[\text{noop} \ ? \ C) \leftarrow \]
\[\text{memb } N) \circ (\text{it} \text{ } 3 \ C \ D \ I \ O) \leftarrow (\text{it} \text{ } 3 \ C \ D \ I \ O) \circ \]
\[[(\text{ne} \ (N = 1)) \leftarrow \ (\text{memd } N) \circ \ 1] \oplus \]
\[[(\text{eq} \ (N = 1)) \land \]
\[[(\text{alu} \ ? \ C) \leftarrow \]
\[((\text{aout}_1 V_1) \circ (\text{laout}_1 V_2) \leftarrow (\text{aout}_1 V_2) \circ (\text{memd } N)) \circ \ 1] \oplus \]
\[[(\text{ld} \ ? \ C) \leftarrow \]
\[((\text{llmdr } V_1) \circ (\text{llmdr } V_2) \leftarrow (\text{llmdr } V_2) \circ (\text{memd } N)) \circ \ 1] \oplus \]
\[[\text{st} \ ? \ C) \leftarrow \]
\[[[\text{noop} \ ? \ C) \leftarrow (\text{memd } N)] \oplus \]
\[[\text{br} \ ? \ C) \leftarrow \]
\[((\text{pc } L_1) \circ (\text{lpc} \ 2 \ L_2) \leftarrow (\text{pc } L_2) \circ (\text{memd } N)) \circ \ 1] \]

Figure 5.14: Specification for the DLX pipeline – MEM.
\[(\text{wbrb } N) \land (\text{it}_3 C D I O) \land (\text{num } M) \leftarrow (\text{it}_3 C D I O) \land\]
\[\left[\left(\text{eq}(N = 0)\right) \otimes (\text{wbrd } N) \land (\text{num } M)\right] \oplus\]
\[\left[\left(\text{eq}(N = -1)\right) \otimes (\text{wbrd } N) \land (\text{num } (M + 1))\right] \oplus\]
\[\left[\left(\text{eq}(N = 1)\right) \otimes\right.\]
\[\left.\left(\text{alu? } C\right) \leftarrow\right.\]
\[\left.\left(\text{aout}_1 V_1\right) \leftarrow (\text{aout}_1 V_1) \land (\text{lr}_1 dV_1) \land (\text{wbrd } N) \land (\text{num } (M + 1)) \land 1\right] \oplus\]
\[\left[\left(\text{ld? } C\right) \leftarrow\right.\]
\[\left.\left(\text{lmdr}_1 V_1\right) \leftarrow (\text{lmdr}_1 V_1) \land (\text{lr}_1 dV_1) \land (\text{wbrd } N) \land (\text{num } (M + 1))\right] \oplus\]
\[\left[\left(\text{st? } C\right) \oplus (\text{br? } C)\right) \otimes (\text{lr}_0 dV_1) \land (\text{wbrd } N) \land (\text{num } (M + 1))\]}
\[\left[\left(\text{noop? } C\right) \otimes (\text{lr}_0 dV_1) \land (\text{wbrd } N) \land (\text{num } M)\right]\]

\[(\text{wbrb } N) \leftarrow\]
\[\left[\left(\text{eq}(N = 1)\right) \otimes (\text{wbrd } N)\right] \oplus\]
\[\left[\left(\text{eq}(N = 1)\right) \otimes\right.\]
\[\left.\left((\text{lr}_1 dV_1) \land (\text{x} dV_2) \leftarrow (\text{x} dV_1) \land (\text{wbrd } N)\right) \land\]
\[\left.((\text{lr}_0 dV_1) \leftarrow (\text{wbrd } N) \land 1\right]\]

Figure 5.15: Specification for the DLX pipeline – WB.

\[
\begin{align*}
S_2 \rightarrow (\text{ifrd1}) \land (\text{ir } X_1) \land (\text{pc } 3) \land (\text{pc } 3) \land (\text{pc } 3) \land \sigma_1 \\
S_2 \rightarrow (\text{crb } \tilde{I}) \land (\text{lir } X_1) \land (\text{lpc } 3) \land (\text{lpc } 3) \land \sigma_2 \\
S_2 \rightarrow (\text{crd } \tilde{I}) \land (\text{lir } X_1) \land (\text{lpc } 3) \land (\text{lpc } 3) \land \sigma_2 \\
S_2 \rightarrow (\text{idrb, exrb} \vdash \text{merb, wbrb}) \\
S_2 \rightarrow (\text{idrb}) \land (\text{lir } X_1) \land (\text{lpc } 3) \land \sigma_1 \\
S_2 \rightarrow (\text{pc } 2) \land (\text{pc } 3) \land (\text{pc } 3) \land (\text{pc } 3) \land \sigma_0 \\
\Sigma_p : E_2 : S_2 \rightarrow S_1
\end{align*}
\]

The proof begins by backchaining on the bg clause. Each of the it registers has a noop instruction (by the definition of \(L_0\)). Thus, only the execution of ifrb and ifwub stages are shown. Note that there is no hazard detected because the instruction in it1 was a noop. X1 is the instruction (ix alu 2.3 2 I+), and X2 is the instruction (ix ht 1.2 3 0 +). Now I complete \(\sigma_1\) below.

107
I use $\Gamma_i$ in the proofs above for the part of the context which is inactive in the rule. The construction of $\sigma_2$ is similar to the plan above. First the ID stage is completed for $X_1$, the EX, MEM and WB have noop instructions. In the next cycle, ht is detected and the state of idwb becomes 0, $X_1$ proceeds to EX stage. Now the pipeline goes through the final stages as the clock shuts down. The construction of $\sigma_2$ is not made explicit because it gets very tedious and the idea is clear from the above constructions.

Note that in the read and the write phase I choose a given order for backtracking on the various clauses. It is a key fact that the order in which the various read clauses are chosen for backtracking during the read phase is irrelevant to provability, and similarly for the write phase – a similar point was made in the context of process theories in [Mil93]. The reason for this “permutability” in the read phase is that information is read by the various stages into different latches, and in the write phase the stages use the latches they “created” to update the values of distinct storage locations. In other words, there is no contention for resources in a given phase, read or write, amongst the various stages. It is this permutability of backtracking which underlines the concurrent nature of the pipeline specification in FORUM.

Now I have $\mathcal{E}_s$ and $\mathcal{E}_p$, two operational semantics for DLX. On the one hand, $\mathcal{E}_s$ is much easier to work with and understand, and on the other hand, $\mathcal{E}_p$ specifies a much more interesting algorithm. The natural question to ask is whether $\mathcal{E}_s$ and $\mathcal{E}_p$ are equivalent. Following is the Correspondence theorem that I would like to prove.

**Theorem 5.7 (Correspondence between $\mathcal{E}_s$ and $\mathcal{E}_p$)** For any DLX program $P_l$, and
data states $S_1$ and $S_2$,

$$S_1 \phi (P i P_1) \leftarrow S_2 \phi (P i P_1) \iff S_1 \phi (P i P_1) \leftarrow_p S_2 \phi (P i P_1)$$

The theorem will be proved by induction on the number of instructions executed. Since the internal states of the two specifications are different, the theorem is obtained as a corollary to the lemma 5.9 which establishes relationships between the sequential and pipelined evaluation using the different internal states. Before I present these lemmas, I prove some properties about the evaluation using $E_p$. I define it to be consistent if $(it_1 C_1 D_1 I_1 O_1)$ is the instruction in EX stage, $(it_2 C_2 D_2 I_2 O_2)$ is the instruction in MEM stage, and $(it_3 C_3 D_3 I_3 O_3)$ is the instruction in WB stage. The actual proofs, which are constructions of proofs in FORUM, are not shown in the proofs of the lemmas in this section, essentially because the construction gets very tedious and unilluminating. However, the steps outlined provide the recipe for constructing the required proofs.

**Lemma 5.8 ($E_p$ - it consistency, data hazards, and control hazards)** If

- $L_1$ and $L_2$ are two latch states, and it is consistent in $L_1$,
- $S_1$ and $S_2$ are two data states, such that the pc in $S_1$ does not address a bg or ht instruction in $P$,
- $P$ is a program state, and
- $\Sigma_p : E_p; S_2 \phi L_2 \phi (\text{cwd} N) \phi P \vdash S_1 \phi L_1 \phi (\text{crb} N) \phi P$,

then

1. it is consistent in $L_2$,

2. if the instruction in the ID stage depends upon a preceding instruction in the pipeline for data, then the data hazard is resolved, and
3. if the instruction in the ID stage is a br, then the control hazard is resolved.

**Proof: Part 1:** For this part, notice that in every completion of a read and write phase \( \text{it}_3 \) gets the values of \( \text{it}_2 \), which in turn gets the values from \( \text{it}_1 \), which in turn get the values from \( \text{ir} \) in the ID. This movement of values of \( \text{it} \) exactly matches the definition of the consistency of \( \text{it} \).

**Part 2:** Given the consistency of \( \text{it} \), \( \delta_l \) returns one of -1, -2 or -3, i.e. a data hazard, if and only if the instruction in ID stage depends upon one of the preceding instructions for data. From the definition of the state transition function in figure 5.10, it is clear that the instruction with the dependency and the ones following that instruction are stalled till the preceding instruction completes execution.

**Part 3:** Given the consistency of \( \text{it} \), \( \delta_l \) detects returns -4, i.e. a control hazard, if and only if the current instruction is a br and there are no data hazards. The absence of data hazards enables the execution of br. From figure 5.10, the pipeline after the br is stalled till the br completes the MEM stage, when the new pc is available. Thus the instruction after a br is fetched from the right address. □

The problem in relating the computations of \( \mathcal{E}_s \) and \( \mathcal{E}_p \) are manifold. Firstly, the number of steps taken to compute a program are different - \( \mathcal{E}_p \) takes five steps for each instruction - hence an induction on the height of the proof trees would not work. Secondly, the internal states of the two evaluators are different. *The induction measure that I use is the number of instructions which have completed evaluation. However, I have to treat the case when the last instruction is \( \text{ht} \) separately from other instructions because it cleans up the environment of the latch state.*

**Lemma 5.9 (Lemmas for Correspondence between \( \mathcal{E}_s \) and \( \mathcal{E}_p \))** Given

- \( m \geq 1 \) and \( m \in \text{nat} \).
• \( S_1 \) and \( S_2 \) two data states minus the pc, num has 0 and \( m \) in \( S_1 \) and \( S_2 \), respectively,

• \( P_1 \) is a DLX program, \( P_1 = \text{def } ( P \downarrow \overline{P}_i ) \) is a program state,

• \( L_1 = \text{def } \forall \overline{C}, \overline{D}, \overline{I}, \overline{O}, \overline{S}, \overline{V}. ( L \overline{C} \overline{D} \overline{I} \overline{O} \overline{S} \overline{V} ) \), and

• \( C = \text{def } \forall \overline{N}. ( \text{cwd}\overline{N} ) \).

1. if the \( m \)th instruction is not \( \text{ht} \) then

\[
\Sigma_s : \varepsilon_s ; S_2 \varphi \text{cont} \varphi(\text{pc}\overline{L}) \varphi P_1 + S_1 \varphi(\text{pc}\overline{M}) \varphi P_1 \iff \\
\Sigma_p : \varepsilon_p ; S_2 \varphi(\text{pc}\overline{K}) \varphi L_1 \varphi C \varphi P_1 + S_1 \varphi(\text{pc}\overline{M}) \varphi P_1
\]

2. if the \( m \)th instruction is \( \text{ht} \) then

\[
\Sigma_s : \varepsilon_s ; S_2 \varphi(\text{pc}\overline{L}) \varphi P_1 + S_1 \varphi(\text{pc}\overline{M}) \varphi P_1 \iff \\
\Sigma_p : \varepsilon_p ; S_2 \varphi(\text{pc}\overline{L}) \varphi P_1 + S_1 \varphi(\text{pc}\overline{M}) \varphi P_1
\]

Proof: Proof Part 1 (Left to Right): The \( m_{th} \) instruction is not \( \text{ht} \).

We have

\[
\Sigma_s : \varepsilon_s ; S_2 \varphi(\text{pc}\overline{L}) \varphi P_1 + S_1 \varphi(\text{pc}\overline{M}) \varphi P_1.
\]

Suppose the \( m_{th} \) instruction is a \( \text{bg} \). Then there are two cases:

\( m = 1 \): Then \( \text{bg} \) is the first instruction, \( S_2 \) and \( S_1 \) are identical except for the count in num, and the proof is completed by following the execution of the \( \text{bg} \) instruction.

\( m = n + 1, 0 < n \): Then \( \text{bg} \) is not the first instruction. By the grammar for programs (figure 5.5), the \( n_{th} \) instruction executed must be a \( \text{ht} \). Thus \( S_2' \), the state at the end of
the $n_{th}$ instruction, is the same as $S_2$ except for the count in $num$. Using induction on Part 2, I get

$$\Sigma_p : \mathcal{E}_p ; S'_2 \varphi\left(pc(L - 1)\right) \varphi P_1 \vdash S_1 \varphi\left(pc M\right) \varphi P_1.$$  

The shape of the right-hand side of the sequent at the point when $num$ is incremented to $n$ will be

$$S'_2 \varphi\left(pc(L - 1)\right) \varphi P_1.$$  

Backchaining on the clause for $bg$ increments $num$ by 1 and the right-hand side of the sequent becomes

$$S_2 \varphi\left(pc L\right) \varphi L_0 \varphi\left(crb \bar{0}\right) \varphi P_1.$$  

Now in the linear part of the left-hand-side of the sequent, I replace

$$S'_2 \varphi\left(pc(L - 1)\right) \varphi P_1 \quad \text{by} \quad S_2 \varphi\left(pc L\right) \varphi L_1 \varphi C \varphi P_1.$$  

Note that this replacement does not effect the structure of the proof thus far, and the sequent that I am now constructing the proof of is

$$\Sigma_p : \mathcal{E}_p ; S_2 \varphi\left(pc L\right) \varphi L_1 \varphi C \varphi P_1 \vdash S_1 \varphi\left(pc M\right) \varphi P_1.$$  

After the transformation all that I have left to prove is

$$\Sigma_p : \mathcal{E}_p ; S_2 \varphi\left(pc L\right) \varphi L_1 \varphi C \varphi P_1 \vdash S_2 \varphi\left(pc L\right) \varphi L_0 \varphi\left(crb \bar{0}\right) \varphi P_1.$$  

The proof is easily completed by using $\varphi = L$, $\forall = L$ and $Id$, and hence I am done.
Suppose the $m_{th}$ instruction is not a bg. Then the definition of the DLX programs implies that $m = n + 1$, $0 < n$ and that the $n_{th}$ instruction cannot be ht. By induction hypothesis on Part 1, I get

$$\Sigma_p : \mathcal{E}_p ; S'_2 \varnothing L_1 \varnothing C \varnothing (pc \; K) \varnothing P_1 + S_1 \varnothing (pc \; M) \varnothing P_1.$$  

$S'_2$ differs from $S_2$ in that it has $n$ in num and is missing the side effects, if any, of the $m_{th}$ instruction.

First consider the case when the $n_{th}$ instruction is not a branch instruction. The computation up to the execution of the $n_{th}$ instruction matches for $\mathcal{E}_s$ and $\mathcal{E}_p$. Note that the $n_{th}$ instruction is at $(L - 2)$ address in the program memory. By lemma 5.8 parts 1 and 2, it is consistent, and data-dependency, if any, for the $m_{th}$ instruction is resolved. Thus, the arguments to the $m_{th}$ instruction are identical to the sequential computation, and hence, so are the results. At the point when num is incremented to $n$ the right-hand side of the sequent has the following shape

$$S'_2 \varnothing (pc \; K') \varnothing (L \; V) \varnothing (crb \; W) \varnothing P_1.$$  

Where $K'$ depends upon the data and control hazards encountered by the instructions following the $n_{th}$ instruction. For example, if there are no hazards then $K'$ will be $L + 3$. Backchaining over clauses to complete the execution of the $m_{th}$ instruction results in a right-hand side that looks like

$$S_2 \varnothing (pc \; K) \varnothing (L \; V') \varnothing (crb \; W') \varnothing P_1.$$  

To complete the proof, in the linear part of the left-hand-side of the sequent, replace

$$S'_2 \varnothing L_1 \varnothing C \varnothing (pc \; K) \varnothing P_1 \text{ by } S_2 \varnothing L_1 \varnothing C \varnothing (pc \; K) \varnothing P_1.$$  

The transformation does not effect the structure of the proof, and as a result of it I am now
constructing the proof for the sequent

$$\Sigma_p : E_p ; S_2 \varphi L_1 \varphi C \varphi (pc K) \varphi P_1 \rightarrow S_1 \varphi (pc M) \varphi P_1.$$ 

Given the right-hand side above, the proof can be finished by $\varphi - L, \forall - L$ and $Id$, and hence I am done.

Now consider the case when the $n_{th}$ instruction is a branch instruction. Firstly, the argument to $br$ is identical to the one in the sequential run, and thus by lemma 5.8 part 3, the next $pc$ calculated at the end of the $MEM$ stage will also be $(L - 1)$. Again, using part 2 of lemma 5.8, the data dependency, if any, for the $m_{th}$ instruction is resolved, and thus the result produced by the $m_{th}$ instruction are identical to the ones produced in the sequential run. By the definition of the transition functions in figure 5.10, at the end of the cycle when num is incremented to $n$ the right-hand side of the sequent has the following shape

$$S'_2 \varphi (pc L) \varphi (L \bar{V}) \varphi (crb \bar{W}) \varphi P_1$$

and the $m_{th}$ instruction will have completed its IF stage. The rest of the argument can be completed analogous to the case above when the $n_{th}$ instruction was not $br$.

Proof Part 2 (Left to Right):

We have

$$\Sigma_s : E_s ; S_2 \varphi (pc L) \varphi P_1 \vdash S_1 \varphi (pc M) \varphi P_1.$$ 

The $m_{th}$ instruction to be executed - $ht$ - is at the $(L - 1)$ address in the program memory. By the grammar for DLX programs (figure 5.5) there has to be at least one instruction that was completed before $ht$, and the immediately preceding instruction could not be another
ht. Thus \( m = n + 1, 0 < n \). The state \( S'_{2} \) at the end of the \( n_{th} \) instruction is the same as \( S_{2} \), except that \( \text{num} \) stores \( n \).

By induction hypotheses on part 1, I have

\[
\Sigma_{p} : \mathcal{E}_{p} ; S'_{2} \circ (pc \ K) \circ L_{1} \circ C \circ P_{1} + S_{1} \circ (pc \ M) \circ P_{1}.
\]

First consider the case when the \( n_{th} \) instruction is not a branch instruction.

Note that the \( n_{th} \) instruction is at the \((L - 2)\) address in the program memory. In the proof I look at the point when the \( n_{th} \) instruction completed its ID phase. At this point the right-hand side of the sequent will have the following shape,

\[
S_{2} \circ (C \ V) \circ (pc \ L) \circ (\text{cmd} \ W) \circ P_{1}.
\]

Now, in this proof at the point \( \text{ht} \) completed the ID stage, the state of the \( \text{idrd} \) was set to 0, and \( \text{pc} \) has \((L + 1)\). \( \tau_{1} \) sets the state for \( \text{idrb} \) is set to -1 which causes the \( \text{pc} \) to be decremented by 1 by the IF stage, and thus \( \text{pc} \) has \( L \). Now, from the definition of the state transition functions it is clear that the stages of the pipeline become idle as the \( \text{ht} \) passes through them. \( \text{WB} \) knows that \( \text{ht} \) has arrived when the state of \( \text{wrb} \) is -1, and it increases the count in \( \text{num} \) without looking at the class of the instruction in \( \text{idt} \). Now the clock will detect the -1 as its fifth argument and it will consume the \((C \ V)\) from the environment, and I am left with \( S_{2} \circ (pc \ L) \circ P_{1} \).

In the proof thus far, in the linear part of the left-hand-side of the sequent, I replace

\[
S_{2} \circ (pc \ K) \circ L_{1} \circ C \circ P_{1} \text{ by } S_{2} \circ (pc \ L) \circ P_{1}.
\]

This transformation does not effect the structure of the proof thus far, and now I am proving

\[
\Sigma_{p} ; \mathcal{E}_{p} ; S_{2} \circ (pc \ L) \circ P_{1} + S_{1} \circ (pc \ M) \circ P_{1}.
\]

As a result of the transformation the proof can be completed immediately using an identity axiom.
Now consider the case when the $n_{th}$ instruction is a branch instruction. Firstly, the argument to $br$ is identical to the one in the sequential run, and thus by lemma 5.8 part 3, the next $pc$ calculated at the end of the $MEM$ will also be $(L-1)$. By the definition of the transition functions in figure 5.10, at the end of the cycle when $num$ is incremented to $n$ the right-hand side of the sequent has the following shape

$$S' \psi(pc, L) \psi(L, \bar{V}) \psi(crd, \bar{W}) \psi P_1$$

and the $m_{th}$ instruction will have completed its IF stage. The rest of the argument can be completed analogous to the case above when the $n_{th}$ instruction was not $br$.

The proofs going from right to left can be completed along similar lines - the proofs for $E_s$ are composed using linear cuts. ■

The proof highlights the main fact that the key ingredient in the specification of the pipeline is the complex synchronization and hazard detection. The lemma 5.9 is a inductive argument using the basic properties of the $E_p$ as proved in lemma 5.8. The proof is rather straightforward, given the choice of induction measure, however, the complete construction of FORUM proofs is rather cumbersome. The nice feature is that the two declarative specifications - $E_s$ and $E_p$ - were proved equivalent, and cuts were used in one direction of the proof.

### 5.4 Call-forwarding and early branch resolution

The specification $E_s$ exhibits the basic ideas of the pipeline. However, the pipeline deals naively with the data and control hazards it faces - it stalls whenever it may need to. In this section I look at - call forwarding, early branch resolution, branch prediction - simple and important techniques to reduce the stalls caused by data dependencies and branch instructions.

Suppose an instruction, $X_1$, in the ID stage needs the result of an alu instruction which is in
Figure 5.16: Changes in the DLX pipeline to reduce branch penalty.

the MEM stage. According to the scheme of $E_p$, $X_1$ has to stall for alu to go through MEM and WB before it can get the result of alu – this is clearly inefficient. The rather straightforward idea of short-circuiting the loop, i.e. “forwarding” the result from alu directly to the input of $X_1$, works very well in practise. Sending results from one functional units in the pipeline to another functional unit directly is called call forwarding. Call forwarding reduces dramatically the stalls generated in the pipeline due to data dependencies.

The second inefficiency of $E_p$ is regarding the cost of branch instructions. The idea here is to calculate the outcome of the branch instruction as early as possible in the pipeline – early branch resolution. The basic change to the flow for branch instructions is shown in figure 5.16. By using dedicated adders, extra latches and other circuitry it is possible to compute the new pc by the end of the ID stage. This reduces the penalty for branches from three cycles to one cycle. The other aspect of optimization for branches is to first predict whether a branch will be taken or not, and then continue to fetch and execute from the predicted address till the branch is resolved – branch prediction. In case the branch alters the pc, the earlier instructions are invalidated, and the execution starts at the new pc. In case the branch does not alter the pc, the machine has incurred no penalty by continuing to compute rather than sit idly as $E_p$ does. The number of instructions computed speculatively will never be more than one in the DLX pipeline, because the new pc is ready at the end
of the ID stage of the branch instruction.

Call forwarding and speculative computation become more critical to the performance of the pipeline in the presence of instructions which take more than one cycle to complete. Extensions to, and variants of these ideas are embodied in many of present day RISC machines [HP90]. The specification of these features require more complex synchronizations, new definitions of the state transition functions, and call forwarding functions. Further, the IF and ID stages will have to synchronize directly with each other. Consider the situation when a branch instruction is in the ID stage. The branch instruction might alter the pc at the same time when IF wants to increment the old value in the pc. Thus, there is a race for the pc, and the final value in the pc is unpredictable. I force the write phase of IF to start after the write phase of the ID stage, and thus only IF writes to the pc, avoiding the race.

I begin by presenting the definition for the new hazard resolution function, $\delta_o$, for detecting hazards given that I am implementing call-forwarding and early branch resolution. The table for $\rho_2$ is given in figure 5.17. In figure 5.17, $A \in \{alu, br, ld, st\}$, and $Y \in \{alu, br, ld, st, noop, bg\}$. Furthermore, $=j$ in the column for $D_i$ means that the destination register $D_i$ is the source register $S_j$, $\neq j$ means that the destination register $D_i$ is different from the source register $S_j$, and $X_j$ means either $=j$ or $\neq j$. The first point to note is that $\delta_o$ only depends upon the two immediately preceding instructions in the pipeline, not three as in the case of $\delta_1$. Note, moreover, the difference between the alu and ld instructions. The result of the alu is available at the end of the EX stage, while that of the ld is available only after the MEM stage. Thus, a dependency with an instruction two cycles ahead of the current one is signaled only when the instruction two cycle ahead is a ld. Thus the cost of data hazard with alu instructions is reduced to one cycle and with ld instructions to two cycles at most. Other than this, $\rho_2$ is same as $\rho_1$.

Using the definition of $\rho_2$, I define below $\delta_o$, the hazard detection function, for $E_o$, the new specification of pipeline incorporating call-forwarding and early branch resolution.

**Definition 5.10** [$\delta_o$ - Hazard detection function for $E_o$]
\[
\begin{array}{cccccc}
C & C_1 & C_2 & D_1 & D_2 & \delta_o \\
noop & Y & Y & X & X & 1 \\
A & Y & Y & \neq j & \neq j & 1 \\
A & ld & Y & =j & X & -2 \\
A & Y & ld & \neq j & =j & -1 \\
A & alu & Y & =j & X & -1 \\
br & Y & Y & \neq j & \neq j & -4 \\
ht & Y & Y & X & X & 0 \\
\end{array}
\]

Figure 5.17: \( \rho_2 \) - new table for hazard detection in DLX pipeline

\[
(\delta_o S_1 S_2 \vec{D}_i \vec{C}_i C) = \min\{(\rho_2 S_1 \vec{D}_i \vec{C}_i C), (\rho_2 S_2 \vec{D}_i \vec{C}_i C)\} \quad C \in \{alu, st\}
\]

\[
= (\rho_2 S_1 \vec{D}_i \vec{C}_i C) \quad \text{otherwise}
\]

Where, \( \rho_2 \) is the function in figure 5.17. \( C \) is the instruction type, and \( S_1 \) and \( S_2 \) are the two source registers of the current instruction. \( C_i \) is the instruction type, and \( D_i \) is the destination register of the \( i \)th preceding instruction, for \( i \in \{1, 2\} \).

To make the pipeline work with call-forwarding and early branch resolution, I need to redefine the state transition functions for the clock. The main point is that fewer stall signals are generated when either data or control hazards are detected. \( \sigma_1, \sigma_2, \sigma_3, \sigma_4 \) and \( \sigma_5 \), specified in figure 5.18, are the new transition functions for IF, ID, EX, MEM and WB stages, respectively. The states of the five signals in the clock are arguments to each of the transition functions, and the output is the state of the clock signal for its stage. In figure 5.18, \( n_i \in \{-1, 0, 1\} \), \( i \in \{3, \ldots, 5\} \), and for any input not exhibited the functions \( \sigma_1, \sigma_2, \sigma_3, \sigma_4 \) and \( \sigma_5 \) return 1. For the ID clock signal, \( i \in \{-1, -2\} \) are the states for a data hazard where the stalled instruction has to wait for \(-i\) cycles, \( i \in \{-4, -5\} \) are the states for a control hazard, and \( i = 0 \) is the state when the pipeline will stop within the next five cycles. For example, from figure 5.18, \( \sigma_1(-2)100 = 0 \), thus IF will remain idle in the next cycle.

I sometimes use \( \text{crb} \sigma(\bar{u}) \) as an abbreviation for \( \text{crb} \sigma_1(\bar{u}) \sigma_2(\bar{u}) \sigma_3(\bar{u}) \sigma_4(\bar{u}) \sigma_5(\bar{u}) \), and similarly for \( \text{crd}, \text{cwb} \) and \( \text{cwd} \).

Other than changing the definitions of the transition functions and hazard detection, I need to know from which unit a value has to be forwarded to which unit. In the case of the DLX
architecture that I am studying, all data is read in the ID stage, and thus the destination of the forwarded data will be one of the input latches to the ALU. \( \varphi_1 \) and \( \varphi_2 \), defined in figure 5.19, are the two call-forwarding functions for the two source registers of a given instruction. The result of a \( \text{ld} \) is forwarded only when \( C_3 \) has \( \text{ld} \), because by this time the instruction has completed its MEM stage. In the case of \( \text{alu} \) instructions, results are available after the ID stage, and the result is forwarded when \( \text{alu} \) result is either \( C_2 \) or \( C_3 \). Note that just as \( \rho_2 \) does not need \( \text{it}_3 \), the call-forwarding functions do not need \( \text{it}_1 \). It is interesting to note that the two tables for \( \rho_2 \) and call-forwarding functions put together cover all the cases in the table for \( \rho_1 \). The notation in figure 5.19 is the same as that in figure 5.17. The numeric return values signal the register from where the data will come - 0 stands for the register file, 1 stands for \( \text{imdr} \), 2 stands for \( \text{aout} \) and 3 stands for \( \text{aout}_1 \).

The specification for the pipeline, \( E_o \), is the set of universal closures of clauses in figures 5.20, 5.21 and 5.22. The signature for the specification \( \Sigma_o \) is the union of \( \Sigma_p \) and \( \{ \text{bta} : \text{int} \rightarrow o \} \). Note that the branch is resolved at the end of ID stage, and hence the EX

---

<table>
<thead>
<tr>
<th>IF</th>
<th>ID</th>
<th>EX</th>
<th>MEM</th>
<th>WB</th>
<th>( \sigma_1 )</th>
<th>( \sigma_2 )</th>
<th>( \sigma_3 )</th>
<th>( \sigma_4 )</th>
<th>( \sigma_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-2</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>-2</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>-2</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>n_3</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>n_3</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-4</td>
<td>n_3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>-5</td>
<td>1</td>
<td>n_3</td>
<td>n_4</td>
</tr>
<tr>
<td>0</td>
<td>-4</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>-5</td>
<td>1</td>
<td>0</td>
<td>n_4</td>
</tr>
<tr>
<td>1</td>
<td>-5</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>n_4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>n_3</td>
<td>4</td>
<td>5</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>n_3</td>
<td>n_4</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>n_4</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n_4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
</tbody>
</table>

Figure 5.18: DLX pipeline state transition functions in the presence of call-forwarding and early branch resolution.
and MEM stages treat the branch instruction as if it were a noop. Other than this change, the clauses for EX and MEM are identical to the clauses for these stages in $C_p$. The clause for the clock, bg instruction, eq, ne, WB stage and for matching classes remain unchanged, but I have rewritten them here for the sake of completeness. The critical difference is in the clauses for IF and ID.

The read phase of ID has to take into account call-forwarding and calculation of the jump address for the branch in bta. Note that the calculation of the jump address is done for all instructions, because at this point the class of the instruction has not been decoded. If ID is not idle then first hazards are checked. If there is a data hazard the instruction stalls. Otherwise, if there is either no hazard or a control hazard then the current instruction will continue. The call forwarding functions are used to obtain the arguments of the instruction from the appropriate registers. Note that the call forwarding functions are used only when there are no data hazards. Furthermore, if a control hazard is detected then the state of idrd will be -.4.

If there were no hazards in the ID stage, the state of idwb is 1, bta is consumed away and appropriate registers are loaded. Note that the register pc1 is no longer needed because this was used only by br in the EX stage. The case remaining is when the state of idwb is -.4, i.e. the instruction in ID is a branch instruction. If the condition of the branch instruction is true, then bta is left in the environment and the state of the idwd is -.4. If, however, the condition of the branch instruction is false then bta is consumed and the state of idwd is set to 1.

The write phase of IF stage must synchronize with idwd – this ensures that ID has already
finished its write phase. If the state of idwd is not -4, then the instruction in ID did not alter the pc, and the actions of IF are identical to the write phase of IF in \( \mathcal{E}_p \). In case state of idwd is -4, bta is used to set the value of pc. The read phase of IF is unchanged from \( \mathcal{E}_p \).

The idea behind evaluation remains the same as for \( \mathcal{E}_p \). The definition of data and program state are from definition 5.1 and 5.2. The definition of latch state is different in that the latch state for \( \mathcal{E}_o \) does not have the registers cond and pc. Having said this, I will use the same notation for the latch state for \( \mathcal{E}_o \) also.

**Definition 5.11** [Pipeline evaluation in DLX, \( \mathcal{E}_o \)] Given data states \( S_1 \) and \( S_2 \), and \( (P l \bar{F}_l) \), a program state. \( P_l \) evaluates in \( S_1 \) to \( S_2 \) written as \( S_l \varphi(P l \bar{F}_l) \leftarrow_o S_2 \varphi(P l \bar{F}_l) \), if

\[
\Sigma_o: \mathcal{E}_o ; S_2 \varphi(P l \bar{F}_l) \longrightarrow S_1 \varphi(P l \bar{F}_l)
\]

is provable in FORUM. ■

I have \( \mathcal{E}_s \) and \( \mathcal{E}_o \), two operational semantics for DLX. I will prove the equivalence of these two specifications along the lines of the proof for theorem 5.7.

**Theorem 5.12** (Correspondence between \( \mathcal{E}_s \) and \( \mathcal{E}_o \)) For any DLX program \( P_l \), and data states \( S_1 \) and \( S_2 \),

\[
S_1 \varphi(P l \bar{F}_l) \leftarrow_o S_2 \varphi(P l \bar{F}_l) \iff S_1 \varphi(P l \bar{F}_l) \leftarrow_o S_2 \varphi(P l \bar{F}_l)
\]

Before proving the theorem, the following properties of the \( \mathcal{E}_o \) will be proved. These lemmas prove that \( \mathcal{E}_o \) maintains the consistency of it, and resolves data and control hazards correctly. Analogous facts were proved for \( \mathcal{E}_p \) in lemma 5.8.

**Lemma 5.13** (\( \mathcal{E}_o \) - it consistency, data hazards, and control hazards) If
(ifrb N) \varphi(p LV) \varphi(pc L) \leftarrow (pc L) \varphi(p LV) \varphi

\[(\text{eq}(N = 1)) \neg (\text{eq}(N = 1)) \neg ((\text{lir V}) \varphi(lpc(L + 1)) \varphi(\text{idrd N})) \neg\]

\[(\text{eq}(N = -1)) \neg ((\text{lir V}) \varphi(lpc(L - 1)) \varphi(\text{idrd N})) \neg\]

\[(\text{eq}(N = 0)) \neg (\text{idrd N})\]

(ifwb N) \varphi(idwd M) \leftarrow (idwd M) \varphi

\[(\text{eq}(N = 0)) \neg (\text{eq}(M = -4)) \neg (\text{idwd N}) \neg\]

\[(\text{eq}(N = 1)) \neg (\text{eq}(M = -4)) \leftarrow\]

\[(\text{bta L'}) \varphi(pc L') \leftarrow (pc L') \varphi(\text{idwd N}) \varphi 1\]

\[(\text{eq}(N = 1)) \neg (\text{eq}(M = -4)) \leftarrow\]

\[(\text{ir V'}) \varphi(pc L') \varphi(lir V) \varphi(lpc L) \varphi(bta L') \leftarrow\]

\[(\text{ir V'}) \varphi(pc L') \varphi(lir V) \varphi(lpc L) \leftarrow (\text{ir V'}) \varphi(pc L) \varphi(\text{idwd N}) \varphi 1\]

(idrb N) \varphi(\varphi_{i \in \{1, 2, 3\}}(\text{lit}_1 C; D_i I; O_i)) \varphi(pc L) \varphi(\text{ir}(\text{ix C DS}_1 S_2 I O)) \leftarrow\]

\[(\varphi_{i \in \{1, 2, 3\}}(\text{lit}_1 C; D_i I; O_i)) \varphi(pc L) \varphi(\text{ir}(\text{ix C DS}_1 S_2 I O)) \varphi\]

\[(\text{ne}(N = 1)) \neg (\text{idrd N})\]

\[(\text{eq}(N = 1)) \neg\]

\[(\text{eq}(u = 1)) \neg (\text{eq}(u = -4)) \neg (\text{bta(} L + I)) \neg (\text{lit}_1 C D I O) \leftarrow\]

\[\left[(\text{argr} f_1 S_1 l a V_1) \neg (\text{argr} f_2 S_2 l b V_2) \neg (S \varphi S \leftarrow (\text{idrd u}))\right]\]

\[\left[(\text{ne}(u = 1)) \neg (\text{ne}(u = -4)) \neg (\text{idrd u})\right]\]

\[u = \text{def} (\delta_o S_1 S_2 D_1 D_2 C_1 C_2 C), \quad f_i = \text{def} (\varphi_{i S_1 D_2 D_3 C_2 C_3 C}) \quad i \in \{1, 2\}\]

\[\text{argr} = \text{def} \lambda u, s, l, v.\left[\left((\text{eq}(0 = u)) \varphi(s v) \leftarrow (s v) \varphi(l v)\right) \land\right.\]

\[\left\{\left((\text{eq}(1 = u)) \varphi(l m d r v) \leftarrow (l m d r v) \varphi(l v)\right) \land\right.\]

\[\left\{\left((\text{eq}(2 = u)) \varphi(a o u t v) \leftarrow (a o u t v) \varphi(l v)\right) \land\right.\]

\[\left\{\left((\text{eq}(3 = u)) \varphi(a o u t v) \leftarrow (a o u t v) \varphi(l v)\right)\right.\right.\]

(idwb N) \varphi(it_1 C D I O) \leftarrow\]

\[(\text{ne}(N = 1)) \neg (\text{ne}(N = -4)) \neg (\text{idwd N}) \neg\]

\[(\text{eq}(N = 1)) \neg (\text{bta L}) \leftarrow\]

\[\left[(\text{aW}_1) \varphi(l a V_1) \varphi(b W_2) \varphi(l b V_2) \varphi(l i t_1 C' D' I' O') \leftarrow\right.\]

\[\left(a V_1) \varphi(b V_2) \varphi(it_1 C' D' I' O') \varphi(\text{idwd N})\right.\]

\[\left[\left((\text{eq}N = 1)) \neg (\text{bta L}) \varphi(\text{idwd N})\right) \neg\right.\]

\[\left[\left((\text{eq}N = 0)) \neg (\text{bta L}) \varphi(\text{idwd N})\right) \neg\right.\]

\[\left[\left((\text{eq}V_1 O' 0)) \neg (\text{idwd 0})\right) \left.(\text{idwd 1})\right]\]

Figure 5.20: Specification for the DLX pipeline – IF and ID.
\[(\text{crd} \, \overline{N}) \leftarrow (\text{cwb} \, \overline{N})\]

\[(\text{cwd} \, \overline{N}) \leftarrow \]
\[\left[\text{ne}(N_5 = -1) \otimes (\text{crd} \, \tau(\overline{N}))\right] \oplus \left[\text{eq}(N_5 = -1) \leftarrow \left[1 \land (\overline{L} \overline{C} \overline{D} \overline{I} \overline{S} \overline{V}) \leftarrow \bot\right]\right]\]

\[(p \, \langle \text{bg} \, \overline{D} \overline{S} \, \overline{S}_1 \, \overline{I} \, \overline{O}\rangle) \varphi (p \, \langle \text{cwb} \, \overline{L}\rangle) \leftarrow \]
\[(p \, \langle \text{bg} \, \overline{D} \overline{S} \, \overline{S}_1 \, \overline{I} \, \overline{O}\rangle) \varphi (p \, \langle \text{cwb} + 1 \rangle) \varphi (\text{crd} \, \overline{I}) \varphi \, \overline{L}_0\]

\[(\text{alu? alu}) \leftarrow 1 \quad (\text{ld? ld}) \leftarrow 1\]

\[(\text{noop? noop}) \leftarrow 1\]

\[(\text{exwb} \, \overline{N}) \varphi (\text{it}_1 \, C \, \overline{D} \overline{I} \, \overline{O}) \leftarrow (\text{it}_1 \, C \, \overline{D} \overline{I} \, \overline{O}) \varphi (\text{lit}_2 \, C \, \overline{D} \overline{I} \overline{O})\]
\[\left[\text{ne}(N = 1) \otimes (\text{exrd} \, N)\right] \oplus \]
\[\left[\text{eq}(N = 1) \otimes \right.\]
\[\left[\text{alu?} \, C\right] \leftarrow \]
\[\left((a \, V_1) \varphi (b \, V_2) \leftarrow (a \, V_1) \varphi (b \, V_2) \varphi (\text{laout} \, (V_1 \, O \, V_2)) \varphi (\text{exrd} \, N)) \varphi 1\right]\]
\[\left[\text{ld?} \, C\right] \leftarrow \]
\[\left((a \, V_1) \leftarrow (a \, V_1) \varphi (\text{lmar} \, (V_1 + I)) \varphi (\text{exrd} \, N)) \varphi 1\right]\]
\[\left[\text{st?} \, C\right] \leftarrow \]
\[\left((a \, V_1) \varphi (b \, V_2) \leftarrow \right.\]
\[\left((a \, V_1) \varphi (b \, V_2) \varphi (\text{lmar} \, (V_1 + I)) \varphi (\text{lsmdr} \, V_2) \varphi (\text{exrd} \, N)) \varphi 1\right]\]
\[\left[\text{(noop? C)} \oplus (\text{br? C}) \left. \otimes (\text{exrd} \, N)\right]\right]\]

\[(\text{exwb} \, \overline{N}) \varphi (\text{it}_2 \, C' \, \overline{D}' \, \overline{I}' \, \overline{O}') \varphi (\text{lit}_2 \, C \, \overline{D} \overline{I} \overline{O}) \leftarrow (\text{it}_2 \, C \, \overline{D} \overline{I} \overline{O})\]
\[\left[\text{ne}(N = 1) \otimes (\text{exwd} \, N)\right] \oplus \]
\[\left[\text{eq}(N = 1) \otimes \right.\]
\[\left[\text{alu?} \, C\right] \leftarrow \]
\[\left((a \, \text{out} \, V_1) \varphi (\text{laout} \, V_2) \leftarrow (a \, \text{out} \, V_2) \varphi (\text{exwd} \, N)) \varphi 1\right]\]
\[\left[\text{ld?} \, C\right] \leftarrow \]
\[\left((\text{mar} \, L_1) \varphi (\text{lmar} \, L_2) \leftarrow (\text{mar} \, L_2) \varphi (\text{exwd} \, N)) \varphi 1\right]\]
\[\left[\text{st?} \, C\right] \leftarrow \]
\[\left((\text{mar} \, L_1) \varphi (\text{lmar} \, L_2) \varphi (\text{smdr} \, W_1) \varphi (\text{lsmdr} \, W_2) \leftarrow \right.\]
\[\left((\text{mar} \, L_2) \varphi (\text{smdr} \, W_2) \varphi (\text{exwd} \, N)) \varphi 1\right]\]
\[\left[\text{(noop? C)} \oplus (\text{br? C}) \left. \otimes (\text{exwd} \, N)\right]\right]\]

Figure 5.21: Specification for the DLX pipeline – clock, bg, EX.

124
\( (\text{merb} \, N) \, \varphi (\text{it}_2 \, C \, DB \, I \, O) \, \text{←} \, (\text{it}_2 \, C \, DB \, I \, O) \, \varphi (\text{lit}_3 \, C \, DB \, I \, O) \, \varphi \)
\[ [[\text{eq} \,(N = 1)] \, \otimes \, (\text{merd} \, N)] \]
\[ [[\text{it} \,(N = 1)] \, \circ]
\[ [[\text{alu} \,(C) \, ← \, ((\text{aout} \, V_1) \, ← \, (\text{aout} \, V_1) \, \varphi (\text{laout}_1 \, V_1) \, \varphi (\text{merd} \, N)) \, \varphi \, 1] \]
\[ [[\text{id} \,(C) \, ← \, ((\text{mar} \, L) \, \varphi (m \, L \, V) \, ← \, (\text{mar} \, L) \, \varphi (m \, L \, V) \, \varphi (\text{lmdr} \, V) \, \varphi (\text{merd} \, N)) \, \varphi \, 1] \]
\[ [[\text{st} \,(C) \, ← \, ((\text{mar} \, L) \, \varphi (m \, L \, V) \, \varphi (\text{smdr} \, V_2) \, ← \, (\text{mar} \, L) \, \varphi (m \, L \, V_2) \, \varphi (\text{smdr} \, V_2) \, \varphi (\text{merd} \, N)) \, \varphi \, 1] \]
\[ [[[\text{nop} \,(C) \, \oplus \, (\text{br} \,(C)) \, \otimes \, (\text{merd} \, N)] \]

\( (\text{mewb} \, N) \, \varphi (\text{it}_3 \, C' \, DB' \, IO' \, DB' \, IO') \, \varphi (\text{lit}_3 \, C' \, DB' \, IO' \, DB' \, IO') \, \text{←} \, (\text{it}_3 \, C' \, DB' \, IO' \, DB' \, IO') \, \varphi \)
\[ [[\text{eq} \,(N = 1)] \, \otimes \, (\text{mewd} \, N)] \]
\[ [[\text{eq} \,(N = 1)] \, \circ]
\[ [[\text{alu} \,(C) \, ← \, ((\text{aout}_1 \, V_1) \, ← \, (\text{aout}_1 \, V_1) \, \varphi (\text{laout}_1 \, V_1) \, \varphi (\text{mewd} \, N)) \, \varphi \, 1] \]
\[ [[\text{id} \,(C) \, ← \, ((\text{lmdr} \, W_1) \, \varphi (\text{lmdr} \, W_2) \, ← \, (\text{lmdr} \, W_2) \, \varphi (\text{mewd} \, N)) \, \varphi \, 1] \]
\[ [[[\text{st} \,(C) \, \oplus \, (\text{nop}\, C) \, \oplus \, (\text{br} \,(C)) \, \otimes \, (\text{mewd} \, N)] \]

\( (\text{wbwb} \, N) \, \varphi (\text{it}_3 \, C \, DB \, IO) \, \varphi (\text{num} \, M) \, \text{←} \, (\text{it}_3 \, C \, DB \, IO) \, \varphi \)
\[ [[\text{eq} \,(N = 0)] \, \otimes \, ((\text{wbwb} \, N) \, \varphi (\text{num} \, M)) \]
\[ [[\text{eq} \,(N = -1)] \, \circ \, ((\text{wbwb} \, N) \, \varphi (\text{num} \,(M + 1))) \]
\[ [[\text{eq} \,(N = 1)] \, \circ]
\[ [[\text{alu} \,(C) \, ← \, ((\text{aout}_1 \, V_1) \, ← \, (\text{aout}_1 \, V_1) \, \varphi (\text{lr} \, 1 \, d \, V_1) \, \varphi (\text{wbwb} \, N) \, \varphi (\text{num} \,(M + 1))) \, \varphi \, 1] \]
\[ [[\text{id} \,(C) \, ← \, ((\text{lmdr} \, V_1) \, \varphi (\text{lr} \, 1 \, d \, V_1) \, \varphi (\text{wbwb} \, N) \, \varphi (\text{num} \,(M + 1))) \, \varphi \, 1] \]
\[ [[[\text{st} \,(C) \, \oplus \, (\text{br} \,(C)) \, \otimes \, ((\text{lr} \, 0 \, d \, V_1) \, \varphi (\text{wbwb} \, N) \, \varphi (\text{num} \,(M + 1))) \]
\[ [[[\text{nop} \,(C) \, \otimes \, ((\text{lr} \, 0 \, d \, V_1) \, \varphi (\text{wbwb} \, N) \, \varphi (\text{num} \, M))] \]

\( (\text{wbwb} \, N) \, \text{←} \)
\[ [[\text{eq} \,(N = 1)] \, \otimes \, (\text{wbwb} \, N)] \]
\[ [[\text{eq} \,(N = 1)] \, \circ]
\[ [[[\text{lr} \, 1 \, d \, V_1) \, \varphi (\text{r} \, d \, V_2) \, ← \, (\text{r} \, d \, V_1) \, \varphi (\text{wbwb} \, N)] \&
[[\text{lr} \, 0 \, d \, V_1) \, ← \, (\text{wbwb} \, N) \, \varphi \, 1]] \]

Figure 5.22: Specification for the DLX pipeline – MEM and WB.
\begin{itemize}
\item $L_1$ and $L_2$ are two latch states, and it is consistent in $L_1$,
\item $S_1$ and $S_2$ are two data states, such that the pc in $S_1$ does not address a bg or ht instruction in $P$,
\item $P$ is a program state, and
\item $\Sigma_o : E_o ; S_2 \varphi L_2 \varphi (\text{cwd } \vec{N}) \varphi P \vdash S_1 \varphi L_1 \varphi (\text{crb } \vec{N}) \varphi P$,
\end{itemize}

then

\begin{enumerate}
\item it is consistent in $L_2$,
\item if the instruction in the ID stage depends upon a preceding instruction in the pipeline for data, then the data hazard is resolved, and
\item if the instruction in the ID stage is a br then the control hazard is resolved.
\end{enumerate}

\textbf{Proof:} Identical to proof of lemma 5.8 using the definitions of $\varepsilon_o$ and call forwarding functions. $\blacksquare$

The following lemma is proved by mutual induction on the number of instructions which have completed execution. Since the proof of lemma 5.13 is identical to the proof of lemma 5.9, I state the lemma without proof. The proof of theorem 5.12 is a corollary of this lemma.

\textbf{Lemma 5.14 (Lemmas for Correspondence between $E$ and $E_o$)} Given

\begin{itemize}
\item $m \geq 1$ and $m \in \text{nat}$,
\item $S_1$ and $S_2$ two data states minus the pc, num has 0 and $m$ in $S_1$ and $S_2$, respectively,
\item $P_l$ is a DLX program, $P_1 = \text{def } (P \upharpoonright \vec{P}_l)$ is a program state,
\item $L_1 = \text{def } \forall \vec{C}, \vec{D}, \vec{I}, \vec{O}, \vec{S}, \vec{V}. (L \vec{C} \vec{D} \vec{I} \vec{O} \vec{S} \vec{V})$, and
\end{itemize}
The technique of using external functions for state transition functions, and using the logic
to manage the synchronizations has provided a powerful and flexible tool. The specification
for the optimizations did not change the global structure of the specification. The key
redefinitions were in the external functions, and the clauses for IF and ID stages. The change
from $E_p$ to $E_o$ is not modular; in fact, I believe it cannot be modular because we are changing
the interpreter for DLX programs, not adding new constructs to the programming language.
In spite of the non-modular changes the structure of the specification is maintained, and
the proof strategies for the $E_p$ suffice for $E_o$.

5.5 Program equivalence for DLX — Correctness of code scheduling

In this section I study the observational equivalence for DLX programs. Two programs
are deemed to be observationally equivalent if the observable behavior of the two programs
is identical with respect to a given set of environments. The problem of deciding when
two code fragments are observationally equivalent is of great importance to compiler optimi-
izations. For example, code rescheduling [HP90], i.e. reordering the instructions in the
program, is one of the most important techniques to reduce penalties due to data and con-
trol hazards. However, code can be reordered only if the reordered code is observationally
equivalent to the original code sequence. Similarly, other optimizations done by back end compilers need to be justified by proving appropriate observational equivalence.

The two key words in the informal definition of observational equivalence are environment and observable behavior. The evaluator for DLX programs translates a given data state into another one. Thus, the observable entity at the end of the computation is the data state. Note, however, that if I observe the entire data state then I will be able to count the number of instructions executed by a program. Although the number of instructions executed to compute a result is useful information, this notion of equivalence would be too fine. I am primarily interested in making sure that the results computed by two programs are identical in all environments—how many steps are taken to achieve the results is a question that I am not investigating here. The result of the computation is to alter the contents of the register file and the memory—the observable entities. Hence, two programs will be equivalent if, when placed in identical environment, the contents of the register file and memory at the end of the computations is identical. The observable state of a computation is defined to be the registers, the number of memory cells and the contents of the memory.

**Definition 5.15** [Observable State] $(r_1, \ldots, r_{32})$ are the DLX registers, and $n \in \text{nat}$ is the number of memory cells. Let $O$ be an abbreviation for

$$\lambda n, \bar{V}, \bar{U}. (r_1 V_1) \varphi \cdots (r_{32} V_{32}) \varphi (m_1 U_1) \varphi \cdots (m_n U_n).$$

The lengths of $\bar{V}$ and $\bar{U} - 32$ and $n$ respectively—if implicit—are assumed to be of appropriate length.

For any $n \geq 0, V_1, \ldots, V_{32}, U_1, \ldots, U_n : \text{int}, (O n \bar{V} \bar{U})$ is an observable state. ■

The next problem is to define the notion of environment. An environment, written as $E[\_]_{l,m}$, is defined to be a DLX program, $P_{l+m}$, in which at most one of the lists parsed by the non-terminal $H_{m,m}$ is missing. A *block* is a list of DLX instructions which can be parsed by $H_{m,m}$. Note that given a block $Q$ of length $q$, replacing the hole in the environment $E[\_]_{l,m}$ by $Q$, written as $E[Q]_{l,q}$, results in a DLX program $P_{l+q}$. Thus, blocks of different
lengths may be substituted for the hole in any given environment. I will drop the subscripts on the environments whenever this will not create any confusion. Note that I will use DLX programs themselves to test equivalence of DLX blocks. Using these two concepts, the definition of observational equivalence is made below.

**Definition 5.16** [Observational Equivalence, \(\cong_{dlx}\)] Let \(Q_i, i \in [1, 2]\) be blocks. \(Q_1\) is observationally equivalent to \(Q_2\), written as \(Q_1 \cong_{dlx} Q_2\), if for any DLX environment \(E\) and observable states \(O_1\) and \(O_2\), the following is provable.

\[
\begin{align*}
(pcl_1) \varphi (\text{num} (N_1 + M_1)) \varphi O_2 \varphi (P l_1 E[\tilde{Q}_1]) \rightarrow, O_1 \varphi (pcl 0) \varphi (P l_1 E[\tilde{Q}_1]) \varphi (\text{num} N_1) \\
\text{if and only if} \\
(pcl_2) \varphi (\text{num} (N_2 + M_2)) \varphi O_2 \varphi (P l_2 E[\tilde{Q}_2]) \rightarrow, O_1 \varphi (pcl 0) \varphi (P l_2 E[\tilde{Q}_2]) \varphi (\text{num} N_2)
\end{align*}
\]

for some \(N_i, M_i, L_i\) and \(l_i\) is the length of \(E[Q_i], i \in [1, 2]\).

Note that the definition is with respect to \(E_s\). This is enough due to the theorems 5.12 and 5.7, which establish that all three, \(E_p, E_o\), and \(E_s\), are equivalent. This fact is very helpful because \(E_s\) is the simplest to work with. The definition may appear rather weak because it is only regarding observational equivalence of blocks. This is only apparent because blocks are a basic entity for which it makes sense to formalize the notion of equivalence. The key property of a block as defined is that it can only be entered at the beginning and exited at the end. If either of these two conditions are violated, then it would be next to impossible to find interesting equivalences. Suppose that the definition of a block was such that it allowed one to enter it at some intermediate point. In this case, the standard code rescheduling would be incorrect. To make this point concrete, let's look at the following example. Let their be two instructions \((\text{ix alu} 1 \ 2 \ 3 \ I+)\) and \((\text{ix alu} 4 \ 5 \ 6 \ I+)\). The first one writes the sum of registers 2 and 3 in register 1, and the second one writes the sum of registers 5 and 6 in register 4. It is clear that interchanging the order of these two instructions is harmless, *only if both the instructions are executed*. Suppose I was allowed to place the instructions in an environment which can jump to the second instruction avoiding the execution of
(ix alu 1 2 3 I +), then the reordering of the instructions will produce different results, i.e. code rescheduling in this case would be unsound. The definition of environments and blocks disallows this possibility because the environment cannot jump to an intermediate point in a block. I now prove the observational equivalence for the general statement of code rescheduling.

**Lemma 5.17 (Observational Equivalence of Code rescheduling)** Let

\[ Q_1 =_d e f \ Ix_1; Ix_2, \quad Q_2 =_d e f \ Ix_2; Ix_1, \text{ where } \ Ix_i =_d e f \ (ix \ C; D; S; R; I; O_i), \quad C_i \in \{alu, ld st\}, \text{ and all of } D_i, S_i, R_i \text{ are pair wise distinct for } i \in [1, 2]. \]

\[ Q_1 \equiv_{\text{dx}} Q_2 \]

**Proof:** Note that both \( Q_1 \) and \( Q_2 \) are blocks by definition. let \( E \) be any environment, and \( O_1 \) and \( O_2 \) be two observable states such that

\[ (\text{pc } L_1) \varphi (\text{num } (N_1 + M_1)) \varphi O_2 \varphi (P l_1 E[\tilde{Q}_1]) \leftarrow_s O_1 \varphi (\text{pc } 0) \varphi (P l_1 E[\tilde{Q}_1]) \varphi (\text{num } N_1). \]

If the computation does not reach \( Q_1 \), it will also not reach \( Q_2 \), and thus the proof for

\[ (\text{pc } L_2) \varphi (\text{num } (N_2 + M_2)) \varphi O_2 \varphi (P l_2 E[\tilde{Q}_2]) \leftarrow_s O_1 \varphi (\text{pc } 0) \varphi (P l_2 E[\tilde{Q}_2]) \varphi (\text{num } N_2) \]

is the proof that I assumed.

So suppose that

\[ (\text{pc } L_3) \varphi (\text{num } (N_3 + M_3)) \varphi O_2 \varphi (P l_1 E[\tilde{Q}_1]) \leftarrow_s O_1 \varphi (\text{pc } 0) \varphi (P l_1 E[\tilde{Q}_1]) \varphi (\text{num } N_1) \]

such that \( Ix_1 \) is at the address \( L_3 \) in the program memory. As the block \( Q_1 \) has not been executed as yet, I must have,

\[ (\text{pc } L_4) \varphi (\text{num } (N_4 + M_4)) \varphi O_2 \varphi (P l_2 E[\tilde{Q}_2]) \leftarrow_s O_1 \varphi (\text{pc } 0) \varphi (P l_2 E[\tilde{Q}_2]) \varphi (\text{num } N_2) \]

130
such that \(Ix_2\) is at the address \(L_4\) in the program memory.

Computing the instruction \(Ix_1\) and \(Ix_2\) in both the proofs will yield

\[
\begin{align*}
(pc \ L_3 + 2) \ &\ \varphi (num \ (N_3 + M_3 + 2)) \ \varphi \ O_2 (P \ l_1 \ E[Q_1]) \rightarrow_s \\
O_1 \ &\ \varphi (pc \ 0) \ \varphi (P \ l_1 \ E[Q_1]) \ \varphi (num \ N_1)
\end{align*}
\]

and

\[
\begin{align*}
(pc \ L_4 + 2) \ &\ \varphi (num \ (N_4 + M_4 + 2)) \ \varphi \ O_2 (P \ l_2 \ E[Q_2]) \rightarrow_s \\
O_1 \ &\ \varphi (pc \ 0) \ \varphi (P \ l_2 \ E[Q_2]) \ \varphi (num \ N_2).
\end{align*}
\]

Since there is no dependence between the two instructions, I can permute the order in which the instructions are computed without effecting the observable state at the end of the execution of the two instructions.

The rest of the proof is obtained by inducting on the number of times the blocks \(Q_1\) and \(Q_2\) are computed. The other side of the transformation can be completed similarly. ■

I have provided a formal definition of observational equivalence for DLX programs which can be tackled using program transformations. As an example, I showed how to justify code-rescheduling from the definitions developed. This points out a subtle assumption in the code rescheduling, that the instructions in the two sequences are computed atomically – either one executes the entire sequence, or none of it. Without this assumption, code rescheduling cannot be justified in general. This opens up an interesting line of investigation into back end optimization using this meta-theoretic tool.
Chapter 6

Conclusion and Future Work

6.1 Conclusion

My goal has been to analyze a meta-theory in which various issues regarding programming languages can be discussed. The first, and key requirement of a framework would be that it can specify the operational semantics of the programming language. However, I want the meta-theory to play a much more significant role than the specification of operational semantics alone. In particular, I want to use the meta-theory to study various interesting and challenging properties of programs. One key feature of a meta-theory should be to facilitate discussion of the programming language at various levels of detail — from high-level specifications down to abstract machines. The meta-theory should provide a uniform framework in which diverse properties — subject reduction, compiler optimizations, observational equivalence, and equivalence of different specifications — can be analyzed.

In this thesis I used FORUM as a meta-theory to study programming languages. FORUM provides a rich structure to proofs which was used to specify concurrency, higher-order functions, exceptions, state, and first-class continuations. I specified a fragment of \( HO\pi \)-calculus in FORUM to show how concurrent computations may be represented in FORUM. Next, I defined an untyped higher-order functional language, UML, which provides an exception
mechanism and state and first-class continuations. UML, without first-class continuations, is untyped SML without data-types. UML encapsulates essential programming constructs which have been challenging to understand in more than one way. For example, modular specifications of UML have not been possible. The semantic analysis of the observational equivalence for $\Lambda_{vs}$, the functional core of UML augmented with state, has been very challenging. Next, I analyze the DLX architecture in FORUM. The executable specification of DLX architecture, to the best of my knowledge, has not been attempted. Formalizing low-level optimizations and observational equivalence for DLX programs has been very challenging.

Specifying UML modularly and declaratively has been challenging for formal systems because of the presence the various imperative features. I have provided modular and declarative specifications of the imperative features in UML. The claim that the specifications are modular is justified by the fact that I obtain the specification for UML by literally putting together my specifications for its different parts. The claim that the specifications are declarative is justified by the fact that my proofs regarding the evaluations in FORUM work by composing proofs using cut rules of FORUM.

As a result of my specification, evaluations become proofs in FORUM — formal objects which can be analyzed using the meta-theory of FORUM. I use this fact to study observational equivalence for $\Lambda_{vs}$. Using the proof structure of evaluations in FORUM, I have proved some of the challenging observational equivalences in the literature for $\Lambda_{vs}$-like languages. The nature of these proofs is very interesting. They seem to fall into two main categories. One kind of proofs essentially permute a given evaluation proof, typically using information regarding variable occurrences. The other kind of proofs are based on abstracting away details of function parameter from computations. The structure of proofs, richer logic, and cut rules play a key role in this analysis. However, the story is far from complete. Proofs of observational equivalence exhibit the need for a richer meta-theory for FORUM.

The declarative specification of the DLX pipeline, with its complex synchronizations, hazard resolution, call-forwarding, branch prediction and early-branch resolution, provides ample
evidence of the flexibility of FORUM as a specification language for concurrent and imperative processes. The specifications allow me to prove the correctness of the pipeline with respect to the simple sequential evaluator for DLX programs. Further, I provide a definition of observational equivalence for DLX programs, and justify code rescheduling using the definitions. This effort highlights the key concept of blocks when discussing equivalence of DLX programs.

FORUM seems to provide an appropriate starting point as a meta-theory for present day programming languages. My results regarding the specification of $HO\pi$-calculus, UML and DLX prove that the proofs in FORUM are rich enough to represent a variety of computational paradigms. The analyses of the observational equivalence for $A_{\nu}$, and DLX programs justifies the claim that FORUM can be used to study meta-theoretic properties also.

### 6.2 Future Work

The future work that I want to do in this area has three principal directions. First, I want to develop proof theory required to better specify imperative features and analyze proofs in FORUM. Second, I wish to study the derivation of abstract machines from high-level specifications in FORUM. Finally, I wish to consider other specification tasks. I describe each of these topics below.

**Proof-theoretic challenges**

**Quantifiers for location names**

Although FORUM is able to specify imperative features declaratively, there is one aspect which is not captured entirely by the specification. In the specification for $A_{\nu}$, I used natural numbers to generate unique names for exceptions. The reason for using the `sigctr` was that I had to do inequality checks on the exception names when I searched the exception stack for an appropriate handler. The same problem would come up in the specification of
if the language permitted us to check for the equality of location names. The problem of the mismatch between restriction in \( HO\pi\)-calculus and universal quantifier is also related.

The usage of \( \forall \) to represent creation of new location names is not entirely appropriate. This may be an overkill, because I instantiate the universal quantifier with location names only, not arbitrary values. In some sense, *I need a quantifier for pointers*. When the \( \forall \) is introduced, it discharges a constant from the signature. Along with the discharging of the constant, it might be possible to manage inequality clauses between all the location names. In this sense, the new quantifier then may handle both the “newness” and “uniqueness” of the location names. A solution along these lines would allow for a completely logical specification of the exception mechanism. Further, a proper proof-theoretic understanding might help in the search for semantics for such languages.

**Proof transformations**

In chapter 4 we saw how proofs were manipulated to yield observational equivalences. There were two basic flavors to the proofs. First, the proof would essentially permute a given evaluation proof typically based on information regarding variable occurrences. Second, the proof would attempt to abstract the details of a function parameter from the computation of a program.

Some of the proofs were by induction on the height of proofs in FORUM. This might lead one to believe that there was not much uniformity to the transformations. Fortunately, just the contrary is true. For example, in lemma 4.23, I transform evaluations of \( \text{app} \ M' \ \text{Add}_2 \) to evaluations of \( \text{app} \ M' \ c \) by induction on the height of the evaluation of \( \text{app} \ M' \ \text{Add}_2 \). The shape of the evaluation of \( \text{app} \ M' \ \text{Add}_2 \) will be as shown below. \( \delta \) is the evaluation of \( \text{app} \ \text{Add}_2 \ V \).
\[
\begin{align*}
\text{Evaluation of } (\text{app } \text{Add}_2 \ V) \\
\text{Evaluation of } (\text{app } \text{Add}_2 \ V) \\
\text{Evaluation of } (\text{app } \text{Add}_2 \ V) \\
\text{Sequent for start state of } (\text{app } M' \text{Add}_2)
\end{align*}
\]

The transformation replaces the proof fragment \( \delta \) with the constant \( c \). The structure of the rest of the proof does not matter to the transformation. The question is, how to study these transformations proof theoretically so that the above transformations can be specified compositionally instead of having to induct every time.

**Deriving abstract machines from high-level specifications**

One of the problems in language development is to show the link between the actual abstract machine that is implemented and the high-level semantics that one starts with. The traditional specification techniques are not able to present these different levels of abstractions. Hence, one has to mediate between dissimilar formalisms using some hairy induction arguments. I would like to logically transform the evaluator I have to a CEK style abstract machine for UML. A similar transformation was done for \( \Lambda \nu \) like language in [HM92]. My work would extend it to the richer language UML. The transformation in [HM92] was not carried out entirely within the logic. I want to investigate whether the mixture of FORUM and continuation-passing-style specification will overcome some of the problem encountered in [HM92].

**Specification of other aspects of programming languages**

One of the natural questions that comes after having specified UML, is whether it can be typed in FORUM, i.e., can I specify the static semantics of UML in FORUM? I have some preliminary ideas on this problem. The Subject Reduction theorem would then be a statement about the compositionality between the typing derivation and the evaluation.
The general setting of FORUM would also allow us to analyze other static information about the programs, such as *effects*. The proof theory may provide us insight into the logical nature of *effects*, if any.

**Implementation of FORUM and interpreters for DLX**

An implementation of a fragment of FORUM with first-order unification would suffice to play with interpreters for DLX pipelines. This could be a big step forward in understanding the role of FORUM as a prototyping language for such applications. The critical use of prototypes to experiment with new ideas and negotiate contracts is gaining much recognition recently. As such, this direction might result in some tangible applications of FORUM.
Appendix A

Proofs from chapter 3

In figure A.1 the constants from $\Sigma_{ml}$ used for the translation of UML into FORUM is presented. In figure A.2 the translation from UML to FORUM is presented, in figure A.3 the translation for Answers to FORUM is presented, and in figure A.4 the translation from FORUM terms of type $\text{vl}$ and $\text{tm}$ to UML is presented. I prove lemmas A.1 and A.2 which implies lemmas 3.1, 3.4, 3.9 and 3.13. I use the notation from chapter 3 regarding freely in the appendix.

Lemma A.1 Let $M$ and $N$ be UML terms, $V$ and $U$ be values in UML.

1. $\phi(U[x := V]) = \phi(U)[x := \phi(V)].$
2. $\mathcal{H}(M[x := V]) = \mathcal{H}(M)[x := \phi(V)].$

Proof: The proof of the lemma works by mutual induction on the two claims. The induction is done on the structure of the term.

Proof for claim 1: The claim is vacuously proved in the cases where $U$ is an exception name, or a constant in the language.

Case $U = z, z \neq x : \text{LHS} = z = \text{RHS}$
abs : (vl → tm) → vl
  e : vl
  e ∈ B ∪ Z ∪ {•}

  ⟨.⟩ : vl → tm
  f : tm → tm → tm
  app : tm → tm → tm
  cond : tm → tm → tm → tm
  ifbr : vl → tm → tm → tm
  letval : (vl → tm) → tm → tm
  letfun : (vl → tm) → (vl → vl → tm) → tm
  cell : tm → tm
  read : tm → tm
  write : tm → tm → tm
  ex : ext → vl
  exon : (ext → tm) → tm
  install : tm → tm → tm → tm
  signal : tm → tm → tm
  catch : tm → tm
  jump : tm → tm → tm

get : vl → (vl → o) → o
set : vl → vl → (vl → o) → o
apply : tm → tm → o
uncaught : ext → vl → o
resume : vl → vl → o
cont : vl → (vl → o) → o
eval : tm → (vl → o) → o

Figure A.1: Constants in $\Sigma_{ml}$ used in translating UML to FORUM
\begin{align*}
\phi(x) &= x \\
\phi(\lambda x. M) &= \text{abs} \ x : \text{vl} \ \mathcal{H}(M) \\
\phi(c) &= c \quad c \in \mathbb{Z} \cup \mathbb{B} \cup \{\bullet\} \\
\phi(l) &= (\text{ex } l) \quad l \in \text{ExnNames}
\end{align*}

\begin{align*}
\mathcal{H}(V) &= \{\phi(V)\} \\
\mathcal{H}(f M N) &= f \mathcal{H}(M) \mathcal{H}(N) \\
\mathcal{H}(M N) &= \text{app} \ \mathcal{H}(M) \mathcal{H}(N) \\
\mathcal{H}((f \ M \ N \ P)) &= \text{cond} \ \mathcal{H}(M) \mathcal{H}(N) \mathcal{H}(P) \\
\mathcal{H}((\text{let val } x = M \ \text{in } N)) &= \text{letval} \ (\lambda x. \mathcal{H}(N)) \mathcal{H}(M) \\
\mathcal{H}((\text{let fun } fx = M \ \text{in } N)) &= \text{letfun} \ (\lambda f. \mathcal{H}(N)) \ (\lambda f, x. \mathcal{H}(M)) \\
\mathcal{H}(\text{ref } M) &= \text{cell} \ \mathcal{H}(M) \\
\mathcal{H}(\text{deref } M) &= \text{read} \ \mathcal{H}(M) \\
\mathcal{H}(\text{asg } M \ N) &= \text{write} \ \mathcal{H}(M) \mathcal{H}(N) \\
\mathcal{H}(\text{exception } l \ M) &= \text{exn} \ \lambda l. \mathcal{H}(M) \\
\mathcal{H}(\text{handle } M \ N \ P) &= \text{install} \ \mathcal{H}(M) \mathcal{H}(N) \mathcal{H}(P) \\
\mathcal{H}(\text{raise } M \ N) &= \text{signal} \ \mathcal{H}(M) \mathcal{H}(N) \\
\mathcal{H}(\text{callcc } M) &= \text{catch} \ \mathcal{H}(M) \\
\mathcal{H}(\text{throw } M \ N) &= \text{jump} \ \mathcal{H}(M) \mathcal{H}(N)
\end{align*}

Figure A.2: Translating UML to FORUM

\begin{align*}
\mathcal{A}(V, K) &= (K, \phi(V)) \\
\mathcal{A}([pk \ l, V, ]K) &= (\text{uncaught } l \ \phi_{\text{ex}}(V))
\end{align*}

Figure A.3: Translating answers in UML to FORUM
\( \psi(x) = x \)
\( \psi(\text{abs } \lambda x. M) = \lambda x. \mathcal{L}(M) \)
\( \psi(c) = c \quad \text{for } c \in \mathcal{Z} \cup \mathcal{B} \cup \{\bullet\} \)
\( \psi(\text{ex } l) = l \)

\[ \begin{align*}
\mathcal{L}(\{V\}) &= \psi(V) \\
\mathcal{L}(f M N) &= f \mathcal{L}(M) \mathcal{L}(N) \\
\mathcal{L}(\text{app } M N) &= \mathcal{L}(M) \mathcal{L}(N) \\
\mathcal{L}((\text{ifbr } V N P)) &= \text{if } \psi(V) \mathcal{L}(N) \mathcal{L}(P) \\
\mathcal{L}(\text{cond } M N P) &= \text{if } \mathcal{L}(M) \mathcal{L}(N) \mathcal{L}(P) \\
\mathcal{L}(\text{letval } R N) &= \text{let } \text{val } x = \mathcal{L}(N) \text{ in } \mathcal{L}(R x) \\
\mathcal{L}(\text{letfun } R_1 R_2) &= \text{let } \text{fun } f x = \mathcal{L}(R_2 f x) \text{ in } \mathcal{L}(R_1 f) \\
\mathcal{L}(\text{cell } M) &= \text{ref } \mathcal{L}(M) \\
\mathcal{L}(\text{read } M) &= \text{deref } \mathcal{L}(M) \\
\mathcal{L}(\text{write } M N) &= \text{asg } \mathcal{L}(M) \mathcal{L}(N) \\
\mathcal{L}(\text{exn } R) &= \text{exception } l \mathcal{L}(R l) \\
\mathcal{L}(\text{install } M N P) &= \text{handle } \mathcal{L}(M) \mathcal{L}(N) \mathcal{L}(P) \\
\mathcal{L}(\text{signal } M N) &= \text{raise } \mathcal{L}(M) \mathcal{L}(N) \\
\mathcal{L}(\text{catch } M) &= \text{callcc } \mathcal{L}(M) \\
\mathcal{L}(\text{jump } M N) &= \text{throw } \mathcal{L}(M) \mathcal{L}(N)
\end{align*} \]

Figure A.4: Translating FORUM terms to UML
Case $U = z, z \equiv x$ : LHS $= \phi(V) = \text{RHS}$

Case $U = \lambda z. M, z \not\equiv x$ :
If $z \equiv x$ then the equality is immediate. I write the case when $z \not\equiv x$.
LHS $= \phi((\lambda z. M)[x := V]) = \phi(\lambda z. (M[x := V])) = \text{abs} \lambda z. \mathcal{H}(M[x := V])$
$= \text{abs} \lambda z. (\mathcal{H}(M)[x := \phi(V)]), \text{by induction on claim 2.}$
$= (\text{abs} \lambda z. \mathcal{H}(M))[x := \phi(V)] = \text{RHS}$

This completes the proof of claim 1.

Proof for claim 2:
Case $M = U, U \in \text{Values}$ : By claim 1.

Case $M = (f N P), f \in O$ :
LHS $= f \mathcal{H}(N[x := V]) \mathcal{H}(P[x := V]), \text{by induction on claim 2}$
$= f \mathcal{H}(N)[x := \phi(V)] \mathcal{H}(P)[x := \phi(V)] = \text{RHS}$

Case $M = (N P)$ :
LHS $= \text{app} \mathcal{H}(N[x := V]) \mathcal{H}(P[x := V]), \text{by induction on claim 2}$
$= \text{app} \mathcal{H}(N)[x := \phi(V)] \mathcal{H}(P)[x := \phi(V)] = \text{RHS}$

Case $M = (\text{if } N P L)$ :
LHS $= \text{cond} \mathcal{H}(N[x := V]) \mathcal{H}(P[x := V]) \mathcal{H}(L[x := V]), \text{by induction on claim 2}$
$= \text{cond} \mathcal{H}(N)[x := \phi(V)] \mathcal{H}(P)[x := \phi(V)] \mathcal{H}(L)[x := \phi(V)] = \text{RHS}$
Case $M = (\text{let val } z = N \text{ in } P)$ :

When $z \equiv x$ the result is immediate. I write the proof for $z \not\equiv x$.
LHS = letval $(\lambda z. \mathcal{H}(P[x := V])) \circ \mathcal{H}(N[x := V])$, by induction on claim2
= letval $((\lambda z. \mathcal{H}(P))[x := \phi(V)]) \circ \mathcal{H}(N)[x := V] = \text{RHS}$

Case $M = (\text{let fun } f, z = N \text{ in } P)$ :

When $z \equiv x$ or $z \equiv f$ the result is immediate. I write the proof for $z \not\equiv x$ and $z \not\equiv f$.
LHS = letfun $(\lambda f. \mathcal{H}(P[x := V])) \circ (\lambda f, z. \mathcal{H}(N[x := V]))$, by induction on claim2
= letfun $((\lambda f. \mathcal{H}(P))[x := \phi(V)]) \circ ((\lambda f, z. \mathcal{H}(N))[x := \phi(V)]) = \text{RHS}$

Case $M = (\text{ref } N)$ :

LHS = cell $\mathcal{H}(N[x := V])$, by induction on claim2
= cell $\mathcal{H}(N)[x := \phi(V)] = \text{RHS}$

Case $M = (\text{deref } N)$ :

LHS = read $\mathcal{H}(N[x := V])$, by induction on claim2
= read $\mathcal{H}(N)[x := \phi(V)] = \text{RHS}$

Case $M = (\text{asg } N \ P)$ :

LHS = write $\mathcal{H}(N[x := V]) \circ \mathcal{H}(P[x := V])$, by induction on claim2
= write $\mathcal{H}(N)[x := \phi(V)] \circ \mathcal{H}(P)[x := \phi(V)] = \text{RHS}$

Case $M = (\text{exception } l N)$ :

Note $x \not\equiv l$.
LHS = exn $\mathcal{H}(N[x := V])$, by induction on claim2
= exn $\mathcal{H}(N)[x := \phi(V)] = \text{RHS}$
Case $M = (\text{raise } N \ P)$:
LHS = \text{raise } H(N[x := V]) \ H(P[x := V]), \text{ by induction on claim 2}
= \text{raise } H(N)[x := \phi(V)] \ H(P)[x := \phi(V)] = \text{RHS}

Case $M = (\text{handle } N \ P \ L)$:
LHS = \text{signal } H(N[x := V]) \ H(P[x := V]) \ H(L[x := V]), \text{ by induction on claim 2}
= \text{signal } H(N)[x := \phi(V)] \ H(P)[x := \phi(V)] \ H(L)[x := \phi(V)] = \text{RHS}

Case $M = (\text{callcc } N)$:
LHS = \text{catch } H(N[x := V]), \text{ by induction on claim 2}
= \text{catch } H(N)[x := \phi(V)] = \text{RHS}

Case $M = (\text{throw } N \ P)$:
LHS = \text{jump } H(N[x := V]) \ H(P[x := V]), \text{ by induction on claim 2}
= \text{jump } H(N)[x := \phi(V)] \ H(P)[x := \phi(V)] = \text{RHS}

The proof for lemma A.2 is a mutual induction exactly along the lines of lemma A.1. As the details do not reveal anything new I have not written down the proof.

\textbf{Lemma A.2} Let $N : \text{tm}$, $U : \text{vl}$, and $V : \text{vl}$ be FORUM terms.

1. $\psi(U[x := V]) = \psi(U)[x := \psi(V)]$.

2. $\mathcal{L}(N[x := V]) = \mathcal{L}(N)[x := \psi(V)]$.

I now want to prove theorem 3.18. This theorem states that the natural semantics specification of $\Lambda_{\text{vse}}$ and $\mathcal{E}_{\text{vse}}$, and the FORUM specification of $\Lambda_{\text{vse}}$ are equivalent. I begin with the definition of a \textit{configuration} and the translation of the configuration to FORUM. Configurations are the initial environments in which $\Lambda_{\text{vse}}$ terms need to be evaluated. The
\[ \Sigma : \Phi, C L_S ; Q \varphi \Delta_C \rightarrow Q \varphi \Delta \overset{Id}{\rightarrow} \]

\[ \vdash \forall L, \Rightarrow L \]

\[ \Sigma : \Phi, C L_S ; (C \varphi Q) \rightarrow Q \varphi \Delta \overset{c}{\rightarrow} \]

\[ \vdash \text{Associativity and commutativity of } \varphi \]

\[ \Sigma : \Phi, C L_S ; (C \varphi Q) \rightarrow Q \varphi \Gamma \]

Figure A.5: \( C \) is a configuration

translation of a configuration to FORUM returns a \( \lambda \) which places its arguments in an environment in which all the cells and exceptions have been declared, and appropriately quantified.

**Definition A.3** [Configurations in \( \Lambda_{\text{vsr}} \)] A configuration, \( C \) is a pair of state, \( S \), and set of exception names, \( Ex \), such that for all \( l \in \text{dom}(S) \), \( \text{FV}(S(l)) \subseteq \text{dom}(S) \cup \text{Ex} \).

**Definition A.4** [Translating Configurations to FORUM] \( C = \text{def} \ (S, Ex) \), be a configuration. The translation of \( C \) into FORUM is a term of type \( o \rightarrow o \), written as \( C^o \).

\[
C^o = \text{def} \\
\lambda u : o. \forall P_S, l_S. \text{getC}(P_1, l_1) \Rightarrow \text{setC}(P_1, l_1) \Rightarrow \ldots \Rightarrow \text{getC}(P_n, l_n) \Rightarrow \text{setC}(P_n, l_n) \Rightarrow \\
[u \varphi \Gamma_S \varphi (\text{sigctr } l_{Ex}) \varphi (\text{exnst nil})]
\]

The domain of a configuration, \( C = \text{def} \ (S, Ex) \), written as \( \text{dom}(C) \), is the union of \( \text{dom}(S) \) and \( \text{Ex} \). In figure A.5 I show a proof figure, \( \text{Compl} \), which I will use to end the computations in \( \Lambda_{\text{vsr}} \). Let \( C = (S, Ex) \).

\[
C^o Q = \forall P_S, l_S. \text{getC}(P_1, l_1) \Rightarrow \text{setC}(P_1, l_1) \Rightarrow \ldots \Rightarrow \text{getC}(P_n, l_n) \Rightarrow \text{setC}(P_n, l_n) \Rightarrow \\
[Q \varphi \Gamma_S \varphi (\text{sigctr } l_{Ex}) \varphi (\text{exnst nil})]
\]

145
\[ l_C = \text{def } l_{E_x}, \Gamma_C = \text{def } \Gamma_S, \Delta_C = \text{def } \Gamma_C \varphi(\text{sig} \ l_C) \varphi(\text{exn} \ nil), \text{ and } \Delta \text{ is some permutation of } \Delta_C. \text{ I restate Theorem 3.18.} \]

**Theorem A.5 (Correspondence theorem for } \Lambda_{\text{vsc}} \text{) Let } M \in \Lambda_{\text{vsc}}, A \in \text{Answers}_{\text{vsc}}, \text{ and } C_2 \text{ be a configuration and } C_1 \text{ be a configuration such that } FV(M) \subseteq \text{dom}(C). \]

\[ \langle M, C_1 \rangle \Downarrow \langle A, C_2 \rangle \text{ if and only if eval}(H(M), C_1, A(A, K), C_2) \]

**Proof:** Assume \( M \in \Lambda_{\text{vsc}}, A \in \text{Answers}_{\text{vsc}}, C_2 \) is a configuration and \( C_1 \) is a configuration such that \( FV(M) \subseteq \text{dom}(C). \)

**Left to right direction:** Proof is by induction on the height of the evaluation tree in the natural semantics. I do a case analysis on the structure of \( M \). There are in total thirty-nine rules in natural semantics that I have to analyze. Since many of the cases are repetitive, I will prove the ones which have different features and leave out the proofs of the rest. In the proofs I show, I will not show parts of the signature and intuitionistic context. For example, I do not write \( \Sigma_{\text{vsc}} \) in the signature, and \( \mathcal{E}_{\text{vsc}} \) in the intuitionistic context, since they are present in all the exhibited sequents.

**Case** \( M = V \): There is only one possible evaluation tree. The proof in FORUM follows trivially from *Identity*.

**Case** \( M = (NP) \): There are four possible evaluation trees. Suppose the evaluation is

\[
\begin{align*}
\langle N, C_1 \rangle \Downarrow \langle \lambda x, Q, C_2 \rangle & \quad \langle P, C_2 \rangle \Downarrow \langle U, C_3 \rangle & \quad \langle Q[x := U], C_3 \rangle \Downarrow \langle V, C_4 \rangle \\
\langle (NP), C_1 \rangle \Downarrow \langle V, C_4 \rangle
\end{align*}
\]

The proof is essentially the one given in section 3.1 after theorem 3.3. There are three more natural semantics trees, each for the fact that we have an uncaught exception. The proof for these cases is completed as shown in section 3.3 after theorem 3.12.
**Case** $M = (f N P)$, $f \in \mathcal{O}$: In this case, there are three possible natural semantics which are applicable. The proof is along the lines of the case $M = (N P)$.

**Case** $M = (if N P Q)$: In this case, there are six possible natural semantics which are applicable. I show one of the cases, the others may be completed similarly. Suppose the last natural semantics rule used is:

\[
\langle N, C_1 \rangle \Downarrow \langle \text{true}, C_2 \rangle \quad \langle P, C_2 \rangle \Downarrow \langle V, C_3 \rangle
\]

\[
\langle (if N P Q), C_1 \rangle \Downarrow \langle V, C_3 \rangle
\]

The evaluation trees for $\langle N, C_1 \rangle \Downarrow \langle \text{true}, C_2 \rangle$ and $\langle P, C_2 \rangle \Downarrow \langle V, C_3 \rangle$ are smaller than the evaluation of $M$. Let $N_1 = \text{def } \mathcal{H}(N)$, $P_1 = \text{def } \mathcal{H}(P)$, $Q_1 = \text{def } \mathcal{H}(Q)$, and $V_1 = \text{def } \phi(V)$. Thus, by using induction hypothesis I get proofs $\delta_1$, and $\delta_2$ of

- $K_1 : \mathcal{E}_{\text{w}} ; C_2^0(K_1 \text{true}) \rightarrow C_1^0(\text{eval } N_1 K_1)$, and
- $K_2 : \mathcal{E}_{\text{w}} ; C_3^0(K_2 V_1) \rightarrow C_2^0(\text{eval } P_1 K_2)$, respectively.

The required proof is constructed below. In the proof,

$L_1 = \text{def } \lambda v. (\text{eval } ((\text{if} \ v \ P_1 Q_1)) K)$, and $(L_1 \text{true}) = \beta (\text{eval } ((\text{if} \ v \ P_1 Q_1)) K)$.

\[
\frac{C_2^0(K V_1) \rightarrow \sigma_2 C_2^0(\text{eval } P_1 K)}{\vdash \forall R. R} \quad \frac{C_3^0(K V_1) \rightarrow (\text{eval } ((\text{if} \ v \ P_1 Q_1)) K) \varphi \Delta C_2}{\vdash \forall \Sigma C_1, K : \text{Cl}_{S_1} ; C_3^0(K V_1) \rightarrow (\text{eval } N_1 L_1) \varphi \Delta C_1}
\]

In the above proof, $\sigma_1$ is obtained by $\delta_1$ by a $\text{CutS}$ rule $K_1$ with $L_1$, similarly $\sigma_2$ is obtained from $\delta_2$ by a $\text{CutS}$ rule on $K_2$ with $K$.

Suppose the last natural semantics rule used is:
The evaluation tree for \( \langle N, C_1 \rangle : \langle \text{let } x = N \text{ in } P \rangle, C_2 \) is smaller than the evaluation of \( M \). Let \( N_1 = \text{def } \mathcal{H}(N) \), \( P_1 = \text{def } \mathcal{H}(P) \), \( Q_1 = \text{def } \mathcal{H}(Q) \), and \( V_1 = \text{def } \phi(V) \). Thus, by using induction hypothesis I get proof \( \delta_1 \) of

- \( K_1 : \mathcal{E}_{\text{eval}} ; C_2^\sigma(\text{uncaught } l V_1) \rightarrow C_1^\sigma(\text{eval } N_1 K_1) \).

The required proof is constructed below.

\[
\frac{\Sigma_C, K : \text{CL}_{S_1} ; C_2^\sigma(\text{uncaught } l V_1) \rightarrow \text{eval } N_1 L_1 \Delta C_1}{\Sigma_C, K : \text{CL}_{S_1} ; C_2^\sigma(\text{uncaught } l V_1) \rightarrow \text{eval } \text{if } N_1 P_1 Q_1 K \Delta C_1} \quad \text{backchain}
\]

Where \( \sigma_1 \) is obtained from \( \delta_1 \) using \( \text{CutS} \). The proof, for the other natural semantics rules for if are completed in a similar manner.

**Case** \( M = (\text{let } x = N \text{ in } P) \) : This case is handled exactly like \( (\lambda x. P) \ N \).

**Case** \( M = (\text{let } \text{fun } f, x = N \text{ in } P) \) : The last evaluation rule for letfun will be:

\[
\frac{\langle P[f := \lambda x. \text{let } f x = N \text{ in } P], C_1 \rangle \downarrow \langle A, C_2 \rangle}{\langle \text{let } \text{fun } f x = N \text{ in } P, C_1 \rangle \downarrow \langle A, C_2 \rangle}
\]

The evaluation tree for \( \langle P[f := \lambda x. \text{let } f x = N \text{ in } P], C_1 \rangle : \langle \text{let } \text{fun } f x = N \text{ in } P, C_1 \rangle \) is smaller than the evaluation of \( M \). Let \( N_1 = \text{def } \mathcal{H}(N) \), \( P_1 = \text{def } \mathcal{H}(P) \), and \( V_1 = \text{def } \phi(V) \). Thus, by using induction hypothesis I get the proof \( \delta_1 \) of

- \( K_1 : \mathcal{E}_{\text{eval}} ; C_2^\sigma \mathcal{A}(A, K_1) \rightarrow C_1^\sigma(\text{eval } P_1[f := (\lambda x. \text{let } f x = N_1 \text{ in } N_1)] K_1) \).

The required proof is constructed below. Let \( Q = \text{def } (\lambda x. \text{let } f x = N_1 \text{ in } N_1) \).
\[
\begin{align*}
\Sigma_{\Gamma_1}, K : \text{CL}_{S_1} ; C^\sigma_2 \xrightarrow{\sigma_1} (\text{eval } P_1[f := Q] L_1) \varnothing \Delta_{C_1} \\
\Sigma_{\Gamma_1}, K : \text{CL}_{S_1} ; C_2 \mathcal{A}(A, K) \xrightarrow{\text{eval } \mathcal{H}(\text{let fun } f \ x = N \ in \ P) K} \varnothing \Delta_{C_1}
\end{align*}
\]
backchain

Where \( \sigma_1 \) is obtained from \( \delta_1 \) using \( \text{CutS} \).

Case \( M = (\text{ref } N) \) : I consider the case when the last evaluation rule for \( \text{ref} \) is:

\[
\begin{align*}
\langle N, C_1 \rangle \downarrow \langle V, C_2 \rangle & \quad \langle \text{ref } N, C_1 \rangle \downarrow \langle I, C_2[l \mapsto V] \rangle \\
& \quad l \notin \text{State in } C_2
\end{align*}
\]

The evaluation tree for \( \langle N, C_1 \rangle \downarrow \langle V, C_2 \rangle \) is smaller than the evaluation of \( M \). Let \( N_1 = \text{def} \mathcal{H}(N) \), and \( V_1 = \text{def} \varnothing(V) \). Thus by using induction hypothesis I get proof \( \delta_1 \) of

- \( K_1 : \mathcal{E}_{\text{res}} ; C^\sigma_2(K V) \rightarrow C^\sigma_1(\text{eval } N_1 K_1) \).

The required proof is constructed below. Let \( C_3 = \text{def} \langle S_2[l \mapsto V], E x_2 \rangle \), where \( C_2 = \text{def} \langle S_2, E x_2 \rangle \), and \( L_1 = \text{def} \lambda v. \forall P, l. \text{getC}(P, l) \Rightarrow \text{setC}(P, l) \Rightarrow \ [(K l) \varnothing (P v)] \).

\[
\Sigma_{\Gamma_3}, P, l : \text{CL}_{S_3} ; C^\sigma_3(K l) \rightarrow (K l) \varnothing (P V_1) \varnothing \Delta_{C_2}
\]

\[
\Sigma_{\Gamma_3}, P, l : \text{CL}_{S_3} ; C^\sigma_3(K l) \rightarrow (K l) \varnothing (P V_1)
\]

\[
\Sigma_{\Gamma_3}, P, l : \text{CL}_{S_3} ; C^\sigma_3(K l) \rightarrow (\text{eval } N_1 L_1) \varnothing \Delta_{C_1}
\]

Where \( \sigma_1 \) is obtained from \( \delta_1 \) using \( \text{CutS} \). The other case when the evaluation of \( N \) returns an uncaught exception is treated similarly.

Case \( M = (\text{deref } N) \) : I consider the case when the last evaluation rule for \( \text{deref} \) is:

\[
\begin{align*}
\langle N, C_1 \rangle \downarrow \langle [p k \ l \ V], C_2 \rangle & \quad \langle \text{deref } N, C_1 \rangle \downarrow \langle [p k \ l \ V], C_2 \rangle
\end{align*}
\]

149
The evaluation tree for \( \langle N, C_1 \rangle \downarrow \langle [p_k \ l \ v], C_2 \rangle \) is smaller than the evaluation of \( M \). Let \( N_1 = \text{def} \ \mathcal{H}(N) \), and \( V_1 = \text{def} \ \phi(V) \). Thus, by using induction hypothesis I get proof \( \delta_1 \) of

\[ \bullet \quad K_1 : \mathcal{E}_{\text{ass}} \ ; \ C_2 \langle \text{uncaught} \ l \ V_1 \rangle \rightarrow \ C_1 \langle \text{eval} \ N_1 \ K_1 \rangle. \]

The required proof is constructed below. Let \( L_1 = \text{def} \ \lambda v. (\text{get} \ v \ K) \).

\[
\begin{array}{c}
\Sigma_{C_1}, K : \text{CL}_{S_1} : C_2 \langle \text{uncaught} \ l \ V_1 \rangle \rightarrow (\text{eval} \ N_1 \ L_1) \ \varphi \ \Delta_{C_1} \\
\Sigma_{C_1}, K : \text{CL}_{S_1} : C_2 \langle \text{uncaught} \ l \ V_1 \rangle \rightarrow (\text{eval} \ (\text{read} \ N_1) \ K) \ \varphi \ \Delta_{C_1}
\end{array}
\]

backchain

Where \( \sigma_1 \) is obtained from \( \delta_1 \) using CutS. The other case when evaluation of \( N \) returns a value is treated similarly. Essentially, the computation continues with \( (\text{get} \ l \ K) \), where \( l \) would be the value returned by \( N \).

**Case** \( M = (\text{asg} \ N \ P) : \) There are three natural semantics rules for \( \text{asg} \). The proofs for all of these is completed along the lines of the proof after theorem 3.8.

**Case** \( M = (\text{exception} \ l \ N) : \) This case is rather straightforward given the examples above. It is handled along the lines of \( \text{ref} \).

**Case** \( M = (\text{raise} \ N \ P) : \) This case is rather straightforward given the examples above. It is handled along the lines of \( \text{app} \).

**Case** \( M = (\text{handle} \ N \ P \ Q) : \) Suppose the last rule in the evaluation was

\[
\langle P, C_1 \rangle \downarrow \langle l, C_2 \rangle \ ; \ \langle Q, C_2 \rangle \downarrow \langle W, C_3 \rangle \ ; \ \langle N, C_3 \rangle \downarrow \langle [p_k \ l \ U], C_4 \rangle \ ; \ \langle (W U), C_4 \rangle \downarrow \langle V, C_5 \rangle \\
\langle (\text{handle} \ N \ P \ Q, C_1) \downarrow \langle V, C_5 \rangle \rangle
\]

The evaluation trees for \( \langle P, C_1 \rangle \downarrow \langle l, C_2 \rangle \), \( \langle Q, C_2 \rangle \downarrow \langle W, C_3 \rangle \), \( \langle N, C_3 \rangle \downarrow \langle [p_k \ l \ U], C_4 \rangle \), and \( \langle (W U), C_4 \rangle \downarrow \langle V, C_5 \rangle \) are smaller than the evaluation of \( M \). Let \( N_1 = \text{def} \ \mathcal{H}(N) \),
$P_1 = \text{def } \mathcal{H}(P), Q_1 = \text{def } \mathcal{H}(Q), V_1 = \text{def } \phi(V), U_1 = \text{def } \phi(U), \text{ and } W_1 = \text{def } \phi(W)$. Thus, by using induciton hypothesis I get proofs $\delta_1, \delta_2, \delta_3, \text{ and } \delta_4$ of

- $K_1 : \mathcal{E}_{\text{us}} ; C_2^\alpha(K_1(\text{ex } l)) \rightarrow C_1^\nu(\text{eval } P_1 K_1)$,
- $K_2 : \mathcal{E}_{\text{us}} ; C_3^\nu(K_2 W_1) \rightarrow C_2^\alpha(\text{eval } Q_1 K_2)$,
- $K_3 : \mathcal{E}_{\text{us}} ; C_4^\alpha(\text{uncaught } l U_1) \rightarrow C_3^\nu(\text{eval } N_1 K_3)$, and
- $K_4 : \mathcal{E}_{\text{us}} ; C_5^\nu(K_4 V) \rightarrow C_4^\alpha(\text{app } W_1 U_1 K_4)$, respectively.

The required proof is constructed below.

Let $L_1 = \text{def } \lambda v. (\text{isexn } v \lambda u. (\text{eval } Q_1 \lambda u. (\text{push } w u K N_1)))$, $L_2 = \text{def } \lambda w. (\text{eval } Q_1 \lambda u. (\text{push } w u K N_1))$, and $L_3 = \text{def } \lambda u. (\text{push } l u K N_1)$.

Note $(L_1(\text{ex } l)) = \text{def } (\text{isexn } (\text{ex } l) \lambda w. (\text{eval } Q_1 \lambda u. (\text{push } w u K N_1)))$, $(L_2 l) = \text{def } (\text{eval } Q_1 \lambda u. (\text{push } l u K N_1))$, and $(L_3 W_1) = \text{def } (\text{push } l W_1 K N_1)$.

To complete the proof I am left with constructing $\gamma_1$.

$\Sigma_{C_2} : \mathcal{C}_{\text{CL}_2} ; C_2^\alpha(K V_1) \rightarrow (L_1(\text{ex } l)) \varphi \Delta_{C_2}$

\[
\vdash \forall R, \Rightarrow R \quad C_2^\alpha(K V_1) \rightarrow C_2^\alpha(L_1(\text{ex } l)) \quad C_2^\alpha(L_1(\text{ex } l)) \rightarrow C_1^\nu(\text{eval } N_1 L_1) \quad \sigma_1
\]

$\Sigma_{C_1} : \mathcal{C}_{\text{CL}_1} ; C_1^\nu(K V_1) \rightarrow C_1^\nu(\text{eval } N_1 L_1)$

\[
\Sigma_{C_2} : \mathcal{C}_{\text{CL}_2} ; C_2^\alpha(K V_1) \rightarrow C_2^\alpha(L_2 l) \quad \Sigma_{C_2} : \mathcal{C}_{\text{CL}_2} ; C_2^\alpha(L_2 l) \rightarrow (L_2 l) \varphi \Delta_{C_2}
\]

$\Sigma_{C_2} : \mathcal{C}_{\text{CL}_2} ; C_2^\alpha(K V_1) \rightarrow C_2^\alpha(L_2 l)$

$\Sigma_{C_2} : \mathcal{C}_{\text{CL}_2} ; C_2^\alpha(L_2 l) \rightarrow (L_1(\text{ex } l)) \varphi \Delta_{C_2}$

$\Sigma_{C_2} : \mathcal{C}_{\text{CL}_2} ; C_2^\alpha(K V_1) \rightarrow (L_1(\text{ex } l)) \varphi \Delta_{C_2}$

$\gamma_2$ is constructed below.
\[
\gamma_3 (K V_1) \rightarrow (\text{push } l W_1 K N_1) \varphi \Delta_{C_4}
\]
\[
\therefore \forall R, \exists R
\]
\[
\mu_4 (K V_1) \rightarrow \mu_4 (L_3 W_1) \quad \mu_4 (L_3 W_1) \rightarrow (\text{eval } Q_1 L_3) \varphi \Delta_{C_3}
\]
\[
\mu_4 (K V_1) \rightarrow (\text{eval } Q_1 L_3) \varphi \Delta_{C_3}
\]

\(\gamma_3\) is constructed below.

\[
\begin{align*}
\sigma_3 \quad & \\
\mu_4 (K V_1) & \rightarrow (\text{eval } N_1 \lambda v. (\text{pop } (K v))) \varphi \Gamma_{C_4} \varphi (\text{sigctr } l_{C_4}) \varphi (\text{exnst } (\text{pkt } l W_1 K) :: \text{nil}) \\
\mu_4 (K V_1) & \rightarrow (\text{push } l W_1 K N_1) \varphi \Gamma_{C_4} \varphi (\text{sigctr } l_{C_4}) \varphi (\text{exnst } \text{nil})
\end{align*}
\]

The proof \(\sigma_i\), \(i \in [1, 2]\). \(\sigma_3\) is obtained from \(\delta_3\). \(\delta_3\) results in \(\text{(uncaught } l U_1)\), under the assumption that the initial exception stack is empty. However, the exception stack is \(\text{(exnst } (\text{pkt } l W_1 K) :: \text{nil})\). I change \(\delta_3\) to reflect the initial exception stack, which will catch the exception and evaluate \(\text{(apply } W_1 U_1 K)\). The evaluation of \(\text{(apply } W_1 U_1 K)\) is completed using \(\delta_4\). The other cases for \text{handle} are handled similarly.

**Right to left direction**: In this direction, I induct on height of the sequent proof in FORUM. I then do a case analyses based on the outermost term constructor for the term being evaluate. The proof in this direction is very similar in nature to the one for the other direction. I illustrate the basic strategy using \((\text{asg } N P)\) as an example, and do not write down the other cases. The computation for \((\text{asg } N P)\) in FORUM is as shown below. \(P_1 = \text{def } \mathcal{H}(P), \ N_1 = \text{def } \mathcal{H}(N), \ A \in \text{Answers}_{\text{set}}, \ V_1 = \text{def } \phi(V), \ L_1 = \text{def } \lambda v. (\text{eval } P_1 \lambda u. (\text{set } v u K)), \) and \(L_2 = \text{def } \lambda u. (\text{set } l u K)\).

\[
\begin{align*}
\gamma_1 \quad & \\
\mu_4 A(A, K) & \rightarrow (\text{eval } P_1 L_2) \varphi \Delta_{C_2} \\
\mu_4 A(A, K) & \rightarrow (\text{eval } l L_1) \varphi \Delta_{C_2}
\end{align*}
\]
\[
\therefore \text{Computation of } P_1, \text{ Pt1}
\]
\[
\begin{align*}
\mu_4 A(A, K) & \rightarrow (\text{eval } P_1 L_2) \varphi \Delta_{C_2} \\
\mu_4 A(A, K) & \rightarrow (\text{eval } l L_1) \varphi \Delta_{C_2}
\end{align*}
\]
\[
\therefore \text{Computation of } N_1, \text{ Pt1}
\]
\[
\begin{align*}
\Sigma_{C_1} : \text{CL}_{C_1} \quad & \\
\Sigma_{C_1} : \mu_4 A(A, K) & \rightarrow (\text{eval } (\text{asg } N_1 P_1 K)) \varphi \Delta_{C_1}
\end{align*}
\]

Where \(\gamma_1\) is shown below.
\[
\begin{align*}
C_1^P \Gamma (A, K) & \rightarrow (K \bullet) \varphi \Gamma \varphi (P V_1) \varphi (\text{sigctr } l C_3) \varphi (\text{exnst } \text{nil}) \\
C_1^P \Gamma (A, K) & \rightarrow (m \set V_1 K) \varphi \Gamma \varphi (P V_1) \varphi (\text{sigctr } l C_3) \varphi (\text{exnst } \text{nil}) \\
C_1^P \Gamma (A, K) & \rightarrow \text{eval } P_1 L_2 \varphi \Delta C_2
\end{align*}
\]

From this proof the computations for \( \langle N, C_1 \rangle \downarrow \langle l, C_2 \rangle \) and \( \langle P, C_2 \rangle \downarrow \langle l, C_3 \rangle \) can be extracted easily. Next, the variable conditions in the proof imply that the state can be updated suitable so that the final answer is \( \langle \bullet, C_3[l \rightarrow V] \rangle \). However, if the evaluation of \( N_1 \) raised an uncaught exception, (\text{uncaught } l V_1), then the proof would have ended with the \text{Compl} construction at the point Pt1. In this case, I can get a computation of \( \langle N, C_1 \rangle \downarrow \langle [p k l V], C_2 \rangle \).

The case when \( N_1 \) evaluates to a value, and \( P_1 \) raises an uncaught exception, (\text{uncaught } l V_1), is handled similarly.
Bibliography


