

INTEROPERATION FOR INCOMPATIBLE EVALUATION STRATEGIES

A Thesis

Presented to

the Faculty of California Polytechnic State University

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Computer Science

by

William Faight

June 2008

AUTHORIZATION FOR REPRODUCTION OF MASTER'S THESIS

I reserve the reproduction rights of this thesis for a period of seven years from the date of submission. I waive reproduction rights after the time span has expired.

Signature

Date

APPROVAL PAGE

TITLE: Interoperation for Incompatible Evaluation Strategies

AUTHOR: William Faught

DATE SUBMITTED: June 2008

Dr. John Clements

Advisor or Committee Chair

Signature

Dr. Gene Fisher

Committee Member

Signature

Dr. Aaron Keen

Committee Member

Signature

Abstract

Interoperation for Incompatible Evaluation Strategies

by

William Faight

Software components written in different programming languages can cooperate through interoperation. Differences between languages—incompatibilities—complicate interoperation. This paper explores and resolves incompatible type systems, support for parametricity, and evaluation strategies with a model of computation, gives a thorough proof of its type soundness, and describes an implementation of it. The model uses contracts for higher-order functions and lump types to resolve incompatible type systems, label types to resolve incompatible support for parametricity, and delayed conversions for list constructions to resolve incompatible evaluation strategies. These mechanisms enable the interoperation of Haskell, ML, and Scheme without compromising their semantics.

Acknowledgements

I want to thank my parents, Jerry and Jo Ann, for their encouragement, advice, and support, without which this would not have been possible.

I want to thank my adviser, John Clements, for helping me along the way. I very much appreciate the time he set aside for me and his advice.

Contents

List of Figures	vii
1 Introduction	1
2 Model of Computation	4
2.1 Natural Numbers	5
2.2 Lists	6
2.3 Functions	6
2.4 Types	6
2.5 Interoperation	7
3 Conclusion	24
Bibliography	25

List of Figures

2.1	Haskell grammar and evaluation contexts	8
2.2	Haskell typing rules	9
2.3	Haskell operational semantics	10
2.4	Haskell-ML operational semantics	11
2.5	Haskell-Scheme operational semantics	12
2.6	ML grammar and evaluation contexts	13
2.7	ML typing rules	14
2.8	ML operational semantics	15
2.9	ML-Haskell operational semantics	16
2.10	ML-Scheme operational semantics	17
2.11	Scheme grammar and evaluation contexts	18
2.12	Scheme typing rules	19
2.13	Scheme operational semantics	20
2.14	Scheme-Haskell operational semantics	21
2.15	Scheme-ML operational semantics	22
2.16	Unbrand function	23
2.17	Lump equality relation	23

Chapter 1

Introduction

The complexities of software interoperation in part engender the proverbial reinvention of the wheel. Programmers forgo preexisting solutions to problems where interoperation proves too cumbersome; they reimplement software components, rather than reuse them. Disparate programming language features complicate the conversion of values exchanged between components of different languages. Resolving language incompatibilities transparently at boundaries between component languages facilitates interoperation by unburdening programmers. This paper explores and resolves two such incompatibilities with a model of computation and then proves its type soundness and describes its implementation.

The first incompatibility is type systems. Static type systems calculate and validate the types of expressions before run time, thereby ensuring that well-typed programs do not encounter type errors during run time. Dynamic type systems detect invalid operations on values using value predicates during run time and do not calculate or validate the types of expressions at compile time. Statically-typed languages—languages that use static type systems—that use values from

dynamically-typed languages must verify that the values match their expected types. Languages are assumed to exchange a common set of values that can be checked straightforwardly without coercion. Mismatched values and expected types could cause type errors during run time and violate type soundness. Ad-hoc polymorphism in dynamically-typed languages enables argument types to determine polymorphic function behavior. Since determining function behavior is undecidable [2], actual types for these functions cannot be reliably calculated at language boundaries and compared to expected types. Instead, they are wrapped in contracts [4] that defer the checking of their parameter and result types until they are used during run time. If they are never used, their types cannot be checked, but neither can they cause type errors.

The second incompatibility is parametricity. Parametric polymorphism in statically-typed languages enables function types to be abstracted with type variables and then instantiated into concrete types. Parametricity constrains the behavior of parametric polymorphic functions by ensuring that they behave the same regardless of the types and values of their arguments, and that functions with instantiated result types produce as their results the arguments associated with the same instantiated types. Functions from dynamically-typed languages that use value predicates or conditions on arguments and are used as parametric polymorphic functions by languages that have parametricity can violate their parametricity. Arguments for these functions must be obscured such that value predicates and conditions cannot examine them and annotated to ensure the correct ones are produced as results.

The third incompatibility is evaluation strategies. Evaluation strategies determine the order in which languages evaluate expressions. Eager evaluation evaluates expressions regardless of necessity, and lazy evaluation evaluates expressions

only where necessary. Lazy languages—languages that use lazy evaluation—can construct infinite streams as lists because they do not evaluate list elements when lists are constructed, but eager languages cannot because they do. Since there exist lazy lists—lists in lazy languages—for which no naturally equivalent eager lists exist, lazy lists crossing to eager languages are not converted to eager lists. Instead, elements of lazy lists are converted when accessed by eager languages if they are not lazy lists too.

The languages in the model must be able to express programs in which the aforementioned three incompatibilities arise. Haskell, ML, and Scheme each possess a unique combination of properties that together are sufficient for this purpose: Haskell and ML use static type systems and have parametricity, Scheme uses a dynamic type system, ML and Scheme use eager evaluation, and Haskell uses lazy evaluation.

The rest of the paper is organized as follows: Chapter 2 defines the model of computation, Chapter 3 proves the type soundness of the model, Chapter 4 describes an implementation of the model, Chapter 5 discusses related work, Chapter 6 discusses future work, and Chapter 7 discusses the conclusions.

Chapter 2

Model of Computation

The model of computation is based on that of Matthews and Findler ??.

Their model consists of two simple models, one representing ML and the other Scheme.

The ML model is a simply-typed lambda calculus extended with parametric polymorphism called System F. The substitution semantics by which type abstractions are applied means the ML model has parametricity, which is a property that ensures that programs behave the same regardless of the types applied to by type abstractions. The ML model introduces new expressions, type abstractions and type applications, to express parametric polymorphism, and new types, type variables and forall types, for them. The ML model uses an eager evaluation strategy.

The Scheme model is an extended untyped lambda calculus using an eager evaluation strategy. Value predicates enable ad-hoc polymorphism. Uses a simple type system to check for free variables.

To this mix we introduce a Haskell model identical to the ML model, except

it uses a lazy evaluation strategy.

The definitions of the Haskell, ML, and Scheme models begin in figures 2.1, 2.6, and 2.11.

Expressions are written e , types are written t , forced values are written w , unforced values are written v , forced evaluation contexts are written E , and unforced evaluation contexts are written U . Symbols that represent grammar non-terminals or relations typically have letter subscripts that specify a model.

The Haskell and ML models have static type systems that use typing environments, written Γ , and typing relations, written \vdash . Typing judgements for expressions are written $\Gamma \vdash e : t$, where e is bound in Γ and has the type t . Typing judgements for types are written $\Gamma \vdash t$ and mean t is bound in Γ . Extended typing environments are written $\Gamma, x : t$ for variables and Γ, u for type variables. Typing environments are omitted where empty. The Scheme model uses a simple type system to ensure no free variables. Every well-typed Scheme model expression has the type TST. Type substitution within types is written $x[y/z]$, where the type y is substituted for free occurrences of the type variable z in the type x .

TODO: opsem

Expression and type substitutions within expressions are written like type substitutions within types.

2.1 Natural Numbers

Natural numbers are written \bar{n} , which syntactically represents the natural number n . Natural numbers have the type \mathbb{N} for the Haskell and ML models.

Addition and subtraction are written $+ e e$ and $- e e$. In the Scheme model, if either arithmetic operand is not a natural number, then the result is an error, written **wrong** “Not a number”, which discards the evaluation context and halts the computation. Conditions, written **if0** $e e e$, test whether a value is the natural number $\bar{0}$. If it is, then it reduces to the first alternative; otherwise, it reduces to the second. The Scheme model has predicates that determine whether values are functions, lists, empty lists, and natural numbers, written **fun?** e_S , **list?** e_S , **null?** e_S , and **num?** e_S . Predicates reduce to $\bar{0}$ if true and $\bar{1}$ if false.

2.2 Lists

Empty lists are written **nil** t in the Haskell and ML models and **nil** in the Scheme model. List constructions are written **cons** $e e$ in the Haskell, ML, and Scheme models. Lists of elements that have the type t in the Haskell and ML models have the type $\{t\}$. The Scheme model has predicates that determine whether values are lists and empty lists, written **list?** e and **null?** e .

2.3 Functions

TODO

2.4 Types

TODO

2.5 Interoperation

TODO

$$\begin{aligned}
e_H &= x_H \mid v_H \mid e_H e_H \mid e_H \langle t_H \rangle \mid \mathbf{fix} \, e_H \mid o \, e_H e_H \mid \mathbf{if0} \, e_H e_H e_H \mid f \, e_H \\
&\quad \mathbf{null?} \, e_H \mid \mathbf{wrong} \, t_H \, string \mid \mathbf{hm} \, t_H \, t_M \, e_M \mid \mathbf{hs} \, k_H \, e_S \\
v_H &= \lambda x_H : t_H . e_H \mid \Lambda u_H . e_H \mid \bar{n} \mid \mathbf{nil} \, t_H \mid \mathbf{cons} \, e_H e_H \mid \mathbf{hm} \, L \, t_M \, w_M \\
&\quad \mathbf{hs} \, L \, w_S \\
t_H &= L \mid N \mid u_H \mid \{t_H\} \mid t_H \rightarrow t_H \mid \forall u_H . t_H \\
k_H &= L \mid N \mid u_H \mid \{k_H\} \mid k_H \rightarrow k_H \mid \forall u_H . k_H \mid b \diamond t_H \\
o &= + \mid - \\
f &= \mathbf{hd} \mid \mathbf{tl} \\
E_H &= []_H \mid E_H e_H \mid E_H \langle t_H \rangle \mid \mathbf{fix} \, E_H \mid o \, E_H e_H \mid o \, v_H \, E_H \\
&\quad \mathbf{if0} \, E_H e_H e_H \mid f \, E_H \mid \mathbf{null?} \, E_H \mid \mathbf{hm} \, t_H \, t_M \, E_M \mid \mathbf{hs} \, k_H \, E_S
\end{aligned}$$

Figure 2.1: Haskell grammar and evaluation contexts

$$\begin{array}{c}
\overline{\vdash_H \mathbf{L}} \quad \overline{\vdash_H \mathbf{N}} \quad \overline{\Gamma, u_H \vdash_H u_H} \\
\frac{\Gamma \vdash_H t_H}{\Gamma \vdash_H \{t_H\}} \quad \frac{\Gamma \vdash_H t_H \quad \Gamma \vdash_H t'_H}{\Gamma \vdash_H t_H \rightarrow t'_H} \quad \frac{\Gamma, u_H \vdash_H t_H}{\Gamma \vdash_H \forall u_H. t_H} \\
\\
\frac{\Gamma \vdash_H t_H \quad \Gamma, x_H : t_H \vdash_H e_H : t'_H}{\Gamma \vdash_H (\lambda x_H : t_H. e_H) : t_H \rightarrow t'_H} \quad \frac{\Gamma, u_H \vdash_H e_H : t_H}{\Gamma \vdash_H \Lambda u_H. e_H : \forall u_H. t_H} \quad \overline{\vdash_H \bar{n} : \mathbf{N}} \\
\\
\frac{\Gamma \vdash_H t_H : \{t_H\}}{\Gamma \vdash_H \mathbf{nil} \, t_H : \{t_H\}} \quad \frac{\Gamma \vdash_H e_H : t_H \quad \Gamma \vdash_H e'_H : \{t_H\}}{\Gamma \vdash_H \mathbf{cons} \, e_H \, e'_H : \{t_H\}} \quad \overline{\Gamma, x_H : t_H \vdash_H x_H : t_H} \\
\\
\frac{\Gamma \vdash_H e_H : t_H \rightarrow t'_H \quad \Gamma \vdash_H e'_H : t_H}{\Gamma \vdash_H e_H \, e'_H : t'_H} \quad \frac{\Gamma \vdash_H e_H : t_H \rightarrow t_H}{\Gamma \vdash_H \mathbf{fix} \, e_H : t_H} \\
\\
\frac{\Gamma \vdash_H t_H \quad \Gamma \vdash_H e_H : \forall u_H. t'_H}{\Gamma \vdash_H e_H \langle t_H \rangle : t'_H[t_H/u_H]} \quad \frac{\Gamma \vdash_H e_H : \{t_H\}}{\Gamma \vdash_H \mathbf{hd} \, e_H : t_H} \quad \frac{\Gamma \vdash_H e_H : \{t_H\}}{\Gamma \vdash_H \mathbf{tl} \, e_H : \{t_H\}} \\
\\
\frac{\Gamma \vdash_H e_H : \mathbf{N} \quad \Gamma \vdash_H e'_H : \mathbf{N}}{\Gamma \vdash_H o \, e_H \, e'_H : \mathbf{N}} \quad \frac{\Gamma \vdash_H e_H : \{t_H\}}{\Gamma \vdash_H \mathbf{null?} \, e_H : \mathbf{N}} \quad \frac{\Gamma \vdash_H [k_H] \quad \Gamma \vdash_S e_S : \mathbf{TST}}{\Gamma \vdash_H \mathbf{hs} \, k_H \, e_S : [k_H]} \\
\\
\frac{\Gamma \vdash_H e_H : \mathbf{N} \quad \Gamma \vdash_H e'_H : t_H \quad \Gamma \vdash_H e''_H : t_H}{\Gamma \vdash_H \mathbf{if0} \, e_H \, e'_H \, e''_H : t_H} \quad \frac{\Gamma \vdash_H t_H}{\Gamma \vdash_H \mathbf{wrong} \, t_H \, \mathit{string} : t_H} \\
\\
\frac{\Gamma \vdash_H t_H \quad \Gamma \vdash_M t_M \quad \Gamma \vdash_M e_M : t'_M \quad t_H \doteq t_M \quad t_M = t'_M}{\Gamma \vdash_H \mathbf{hm} \, t_H \, t_M \, e_M : t_H}
\end{array}$$

Figure 2.2: Haskell typing rules

$$\begin{aligned}
& \mathcal{E}[(\lambda x_H : t_H.e_H) e'_H]_H \rightarrow \mathcal{E}[e_H[e'_H/x_H]] \\
& \mathcal{E}[(\Lambda u_H.e_H)\langle t_H \rangle]_H \rightarrow \mathcal{E}[e_H[b \diamond t_H/u_H]] \\
& \mathcal{E}[\mathbf{fix} (\lambda x_H : t_H.e_H)]_H \rightarrow \mathcal{E}[e_H[\mathbf{fix} (\lambda x_H : t_H.e_H)/x_H]] \\
& \mathcal{E}[+ \bar{n} \bar{n'}]_H \rightarrow \mathcal{E}[\overline{n + n'}] \\
& \mathcal{E}[- \bar{n} \bar{n'}]_H \rightarrow \mathcal{E}[\overline{\max(n - n', 0)}] \\
& \mathcal{E}[\mathbf{if0} \bar{0} e_H e'_H]_H \rightarrow \mathcal{E}[e_H] \\
& \mathcal{E}[\mathbf{if0} \bar{n} e_H e'_H]_H \rightarrow \mathcal{E}[e'_H] \ (n \neq 0) \\
& \mathcal{E}[\mathbf{hd} (\mathbf{nil} t_H)]_H \rightarrow \mathcal{E}[\mathbf{wrong} t_H \text{ “Empty list”}] \\
& \mathcal{E}[\mathbf{tl} (\mathbf{nil} t_H)]_H \rightarrow \mathcal{E}[\mathbf{wrong} \{t_H\} \text{ “Empty list”}] \\
& \mathcal{E}[\mathbf{hd} (\mathbf{cons} e_H e'_H)]_H \rightarrow \mathcal{E}[e_H] \\
& \mathcal{E}[\mathbf{tl} (\mathbf{cons} e_H e'_H)]_H \rightarrow \mathcal{E}[e'_H] \\
& \mathcal{E}[\mathbf{null?} (\mathbf{nil} t_H)]_H \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\mathbf{null?} (\mathbf{cons} e_H e'_H)]_H \rightarrow \mathcal{E}[\bar{1}] \\
& \mathcal{E}[\mathbf{wrong} t_H \text{ string}]_H \rightarrow \mathbf{Error: string}
\end{aligned}$$

Figure 2.3: Haskell operational semantics

$$\begin{aligned}
& \mathcal{E}[\mathbf{hm} \ t_H \ \mathbf{L} \ (\mathbf{mh} \ \mathbf{L} \ t'_H \ e_H)]_H \rightarrow \mathcal{E}[e_H] \quad (t_H = t'_H \text{ and } t_H \neq \mathbf{L}) \\
& \mathcal{E}[\mathbf{hm} \ t_H \ \mathbf{L} \ (\mathbf{mh} \ \mathbf{L} \ t'_H \ e_H)]_H \rightarrow \mathcal{E}[\mathbf{wrong} \ t_H \ \text{“Type mismatch”}] \\
& \quad (t_H \neq t'_H \text{ and } t_H \neq \mathbf{L}) \\
& \mathcal{E}[\mathbf{hm} \ t_H \ \mathbf{L} \ (\mathbf{ms} \ \mathbf{L} \ w_S)]_H \rightarrow \mathcal{E}[\mathbf{wrong} \ t_H \ \text{“Bad value”}] \quad (t_H \neq \mathbf{L}) \\
& \mathcal{E}[\mathbf{hm} \ \mathbf{N} \ \mathbf{N} \ \bar{n}]_H \rightarrow \mathcal{E}[\bar{n}] \\
& \mathcal{E}[\mathbf{hm} \ \{t_H\} \ \{t_M\} \ (\mathbf{nil} \ t'_M)]_H \rightarrow \mathcal{E}[\mathbf{nil} \ t_H] \\
& \mathcal{E}[\mathbf{hm} \ \{t_H\} \ \{t_M\} \ (\mathbf{cons} \ v_M \ v'_M)]_H \rightarrow \\
& \quad \mathcal{E}[\mathbf{cons} \ (\mathbf{hm} \ t_H \ t_M \ v_M) \ (\mathbf{hm} \ \{t_H\} \ \{t_M\} \ v'_M)] \\
& \mathcal{E}[\mathbf{hm} \ (t_H \rightarrow t'_H) \ (t_M \rightarrow t'_M) \ (\lambda x_M : t''_M . e_M)]_H \rightarrow \\
& \quad \mathcal{E}[\lambda x_H : t_H . \mathbf{hm} \ t'_H \ t'_M \ ((\lambda x_M : t''_M . e_M) \ (\mathbf{mh} \ t_M \ t_H \ x_H))] \\
& \mathcal{E}[\mathbf{hm} \ (\forall u_H . t_H) \ (\forall u_M . t_M) \ (\Lambda u'_M . e_M)]_H \rightarrow \mathcal{E}[\Lambda u_H . \mathbf{hm} \ t_H \ t_M [\mathbf{L}/u_M] \ e_M [\mathbf{L}/u'_M]]
\end{aligned}$$

Figure 2.4: Haskell-ML operational semantics

$$\begin{aligned}
\mathcal{E}[\mathbf{hs} \ N \ \bar{n}]_H &\rightarrow \mathcal{E}[\bar{n}] \\
\mathcal{E}[\mathbf{hs} \ N \ w_S]_H &\rightarrow \mathcal{E}[\mathbf{wrong} \ N \ \text{“Not a number”}] \ (w_S \neq \bar{n}) \\
\mathcal{E}[\mathbf{hs} \ \{k_H\} \ \mathbf{nil}]_H &\rightarrow \mathcal{E}[\mathbf{nil} \ \lfloor k_H \rfloor] \\
\mathcal{E}[\mathbf{hs} \ \{k_H\} \ (\mathbf{cons} \ v_S \ v'_S)]_H &\rightarrow \mathcal{E}[\mathbf{cons} \ (\mathbf{hs} \ k_H \ v_S) \ (\mathbf{hs} \ \{k_H\} \ v'_S)] \\
\mathcal{E}[\mathbf{hs} \ \{k_H\} \ w_S]_H &\rightarrow \mathcal{E}[\mathbf{wrong} \ \lfloor \{k_H\} \rfloor \ \text{“Not a list”}] \\
&\quad (w_S \neq \mathbf{nil} \text{ and } w_S \neq \mathbf{cons} \ v_S \ v'_S) \\
\mathcal{E}[\mathbf{hs} \ (b \diamond t_H) \ (\mathbf{sh} \ (b \diamond t_H) \ e_H)]_H &\rightarrow \mathcal{E}[e_H] \\
\mathcal{E}[\mathbf{hs} \ (b \diamond t_H) \ w_S]_H &\rightarrow \mathcal{E}[\mathbf{wrong} \ t_H \ \text{“Brand mismatch”}] \ (w_S \neq \mathbf{sh} \ (b \diamond t_H) \ e_H) \\
\mathcal{E}[\mathbf{hs} \ (k_H \rightarrow k'_H) \ (\lambda x_S. e_S)]_H &\rightarrow \mathcal{E}[\lambda x_H : \lfloor k_H \rfloor. \mathbf{hs} \ k'_H \ ((\lambda x_S. e_S) \ (\mathbf{sh} \ k_H \ x_H))] \\
\mathcal{E}[\mathbf{hs} \ (k_H \rightarrow k'_H) \ w_S]_H &\rightarrow \mathcal{E}[\mathbf{wrong} \ \lfloor k_H \rightarrow k'_H \rfloor \ \text{“Not a function”}] \\
&\quad (w_S \neq \lambda x_S. e_S) \\
\mathcal{E}[\mathbf{hs} \ (\forall u_H. k_H) \ w_S]_H &\rightarrow \mathcal{E}[\Lambda u_H. \mathbf{hs} \ k_H \ w_S]
\end{aligned}$$

Figure 2.5: Haskell-Scheme operational semantics

$$\begin{aligned}
e_M &= x_M \mid v_M \mid e_M e_M \mid e_M \langle t_M \rangle \mid \mathbf{fix} \ e_M \mid o \ e_M \ e_M \mid \mathbf{if0} \ e_M \ e_M \ e_M \\
&\quad \mathbf{cons} \ e_M \ e_M \mid f \ e_M \mid \mathbf{null?} \ e_M \mid \mathbf{wrong} \ t_M \ \mathit{string} \mid \mathbf{ms} \ k_M \ e_S \\
v_M &= w_M \mid \mathbf{mh} \ t_M \ t_H \ e_H \\
w_M &= \lambda x_M : t_M . e_M \mid \Lambda u_M . e_M \mid \bar{n} \mid \mathbf{nil} \ t_M \mid \mathbf{cons} \ v_M \ v_M \mid \mathbf{mh} \ L \ t_H \ e_H \\
&\quad \mathbf{ms} \ L \ w_S \\
t_M &= L \mid N \mid u_M \mid \{t_M\} \mid t_M \rightarrow t_M \mid \forall u_M . t_M \\
k_M &= L \mid N \mid u_M \mid \{k_M\} \mid k_M \rightarrow k_M \mid \forall u_M . k_M \mid b \diamond t_M \\
o &= + \mid - \\
f &= \mathbf{hd} \mid \mathbf{tl} \\
E_M &= U_M \mid \mathbf{mh} \ t_M \ t_H \ E_H \\
U_M &= []_M \mid E_M \ e_M \mid w_M \ U_M \mid E_M \langle t_M \rangle \mid \mathbf{fix} \ E_M \mid o \ E_M \ e_M \mid o \ w_M \ E_M \\
&\quad \mathbf{if0} \ E_M \ e_M \ e_M \mid \mathbf{cons} \ U_M \ e_M \mid \mathbf{cons} \ v_M \ U_M \mid f \ E_M \mid \mathbf{null?} \ E_M \\
&\quad \mathbf{ms} \ k_M \ E_S
\end{aligned}$$

Figure 2.6: ML grammar and evaluation contexts

$$\begin{array}{c}
\overline{\vdash_M \mathbf{L}} \quad \overline{\vdash_M \mathbf{N}} \quad \overline{\Gamma, u_M \vdash_M u_M} \\
\frac{\Gamma \vdash_M t_M}{\Gamma \vdash_M \{t_M\}} \quad \frac{\Gamma \vdash_M t_M \quad \Gamma \vdash_M t'_M}{\Gamma \vdash_M t_M \rightarrow t'_M} \quad \frac{\Gamma, u_M \vdash_M t_M}{\Gamma \vdash_M \forall u_M. t_M} \\
\\
\frac{\Gamma \vdash_M t_M \quad \Gamma, x_M : t_M \vdash_M e_M : t'_M}{\Gamma \vdash_M (\lambda x_M : t_M. e_M) : t_M \rightarrow t'_M} \quad \frac{\Gamma, u_M \vdash_M e_M : t_M}{\Gamma \vdash_M \Lambda u_M. e_M : \forall u_M. t_M} \quad \overline{\vdash_M \bar{n} : \mathbf{N}} \\
\\
\frac{\Gamma \vdash_M t_M}{\Gamma \vdash_M \mathbf{nil} \ t_M : \{t_M\}} \quad \frac{\Gamma \vdash_M e_M : t_M \quad \Gamma \vdash_M e'_M : \{t_M\}}{\Gamma \vdash_M \mathbf{cons} \ e_M \ e'_M : \{t_M\}} \quad \frac{}{\Gamma, x_M : t_M \vdash_M x_M : t_M} \\
\\
\frac{\Gamma \vdash_M e_M : t_M \rightarrow t'_M \quad \Gamma \vdash_M e'_M : t_M}{\Gamma \vdash_H e_M \ e'_M : t'_M} \quad \frac{\Gamma \vdash_M e_M : t_M \rightarrow t_M}{\Gamma \vdash_M \mathbf{fix} \ e_M : t_M} \\
\\
\frac{\Gamma \vdash_M t_M \quad \Gamma \vdash_M e_M : \forall u_M. t'_M}{\Gamma \vdash_M e_M \langle t_M \rangle : t'_M[t_M/u_M]} \quad \frac{\Gamma \vdash_M e_M : \{t_M\}}{\Gamma \vdash_M \mathbf{hd} \ e_M : t_M} \quad \frac{\Gamma \vdash_M e_M : \{t_M\}}{\Gamma \vdash_M \mathbf{tl} \ e_M : \{t_M\}} \\
\\
\frac{\Gamma \vdash_M e_M : \mathbf{N} \quad \Gamma \vdash_M e'_M : \mathbf{N}}{\Gamma \vdash_M o \ e_M \ e'_M : \mathbf{N}} \quad \frac{\Gamma \vdash_M e_M : \{t_M\}}{\Gamma \vdash_M \mathbf{null?} \ e_M : \mathbf{N}} \quad \frac{\Gamma \vdash_M [k_M] \quad \Gamma \vdash_S e_S : \mathbf{TST}}{\Gamma \vdash_M \mathbf{ms} \ k_M \ e_S : [k_M]} \\
\\
\frac{\Gamma \vdash_M e_M : \mathbf{N} \quad \Gamma \vdash_M e'_M : t_M \quad \Gamma \vdash_M e''_M : t_M}{\Gamma \vdash_M \mathbf{if0} \ e_M \ e'_M \ e''_M : t_M} \quad \frac{\Gamma \vdash_M t_M}{\Gamma \vdash_M \mathbf{wrong} \ t_M \ string : t_M} \\
\\
\frac{\Gamma \vdash_M t_M \quad \Gamma \vdash_H t_H \quad \Gamma \vdash_H e_H : t'_H \quad t_M \doteq t_H \quad t_H = t'_H}{\Gamma \vdash_M \mathbf{mh} \ t_M \ t_H \ e_H : t_M}
\end{array}$$

Figure 2.7: ML typing rules

$$\begin{aligned}
& \mathcal{E}[(\lambda x_M : t_M.e_M) v_M]_M \rightarrow \mathcal{E}[e_M[v_M/x_M]] \\
& \mathcal{E}[(\Lambda u_M.e_M)\langle t_M \rangle]_M \rightarrow \mathcal{E}[e_M[b \diamond t_M/u_M]] \\
& \mathcal{E}[\mathbf{fix} (\lambda x_M : t_M.e_M)]_M \rightarrow \mathcal{E}[e_M[\mathbf{fix} (\lambda x_M : t_M.e_M)/x_M]] \\
& \mathcal{E}[+ \bar{n} \bar{n}']_M \rightarrow \mathcal{E}[\overline{n + n'}] \\
& \mathcal{E}[- \bar{n} \bar{n}']_M \rightarrow \mathcal{E}[\overline{\max(n - n', 0)}] \\
& \mathcal{E}[\mathbf{if0} \bar{0} e_M e'_M]_M \rightarrow \mathcal{E}[e_M] \\
& \mathcal{E}[\mathbf{if0} \bar{n} e_M e'_M]_M \rightarrow \mathcal{E}[e'_M] \ (n \neq 0) \\
& \mathcal{E}[\mathbf{hd} (\mathbf{nil} t_M)]_M \rightarrow \mathcal{E}[\mathbf{wrong} t_M \text{ “Empty list”}] \\
& \mathcal{E}[\mathbf{tl} (\mathbf{nil} t_M)]_M \rightarrow \mathcal{E}[\mathbf{wrong} \{t_M\} \text{ “Empty list”}] \\
& \mathcal{E}[\mathbf{hd} (\mathbf{cons} v_M v'_M)]_M \rightarrow \mathcal{E}[v_M] \\
& \mathcal{E}[\mathbf{tl} (\mathbf{cons} v_M v'_M)]_M \rightarrow \mathcal{E}[v'_M] \\
& \mathcal{E}[\mathbf{null?} (\mathbf{nil} t_M)]_M \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\mathbf{null?} (\mathbf{cons} v_M v'_M)]_M \rightarrow \mathcal{E}[\bar{1}] \\
& \mathcal{E}[\mathbf{wrong} t_M \text{ string}]_H \rightarrow \mathbf{Error: string}
\end{aligned}$$

Figure 2.8: ML operational semantics

$$\begin{aligned}
& \mathcal{E}[\text{mh } t_M \text{ L } (\text{hm L } t'_M w_M)]_M \rightarrow \mathcal{E}[w_M] \text{ } (t_M = t'_M \text{ and } t_M \neq \text{L}) \\
& \mathcal{E}[\text{mh } t_M \text{ L } (\text{hm L } t'_M w_M)]_M \rightarrow \mathcal{E}[\text{wrong } t_M \text{ "Type mismatch"}] \text{ } (t_M \neq t'_M \text{ and } t_M \neq \text{L}) \\
& \mathcal{E}[\text{mh } t_M \text{ L } (\text{hs L } w_S)]_H \rightarrow \mathcal{E}[\text{wrong } t_M \text{ "Bad value"}] \text{ } (t_M \neq \text{L}) \\
& \mathcal{E}[\text{mh N N } \bar{n}]_M \rightarrow \mathcal{E}[\bar{n}] \\
& \mathcal{E}[\text{mh } \{t_M\} \{t_H\} (\text{nil } t'_H)]_M \rightarrow \mathcal{E}[\text{nil } t_M] \\
& \mathcal{E}[\text{mh } \{t_M\} \{t_H\} (\text{cons } e_H e'_H)]_M \rightarrow \mathcal{E}[\text{cons } (\text{mh } t_M t_H e_H) (\text{mh } \{t_M\} \{t_H\} e'_H)] \\
& \mathcal{E}[\text{mh } (t_M \rightarrow t'_M) (t_H \rightarrow t'_H) (\lambda x_H : t''_H.e_H)]_M \rightarrow \\
& \quad \mathcal{E}[\lambda x_M : t_M.\text{mh } t'_M t'_H ((\lambda x_H : t''_H.e_H) (\text{hm } t_H t_M x_M))] \\
& \mathcal{E}[\text{mh } (\forall u_M.t_M) (\forall u_H.t_H) (\Lambda u'_H.e_H)]_M \rightarrow \mathcal{E}[\Lambda u_M.\text{mh } t_M t_H [\text{L}/u_H] e_H [\text{L}/u'_H]]
\end{aligned}$$

Figure 2.9: ML-Haskell operational semantics

$$\begin{aligned}
& \mathcal{E}[\mathbf{ms} \ N \ \bar{n}]_M \rightarrow \mathcal{E}[\bar{n}] \\
& \mathcal{E}[\mathbf{ms} \ N \ w_S]_M \rightarrow \mathcal{E}[\mathbf{wrong} \ N \ \text{“Not a number”}] \ (w_S \neq \bar{n}) \\
& \mathcal{E}[\mathbf{ms} \ \{k_M\} \ \mathbf{nil}]_M \rightarrow \mathcal{E}[\mathbf{nil} \ [k_M]] \\
& \mathcal{E}[\mathbf{ms} \ \{k_M\} \ (\mathbf{cons} \ v_S \ v'_S)]_M \rightarrow \mathcal{E}[\mathbf{cons} \ (\mathbf{ms} \ k_M \ v_S) \ (\mathbf{ms} \ \{k_M\} \ v'_S)] \\
& \mathcal{E}[\mathbf{ms} \ \{k_M\} \ w_S]_M \rightarrow \mathcal{E}[\mathbf{wrong} \ [\{k_M\}] \ \text{“Not a list”}] \\
& \quad (w_S \neq \mathbf{nil} \text{ and } w_S \neq \mathbf{cons} \ v_S \ v'_S) \\
& \mathcal{E}[\mathbf{ms} \ (b \diamond t_M) \ (\mathbf{sm} \ (b \diamond t_M) \ v_M)]_M \rightarrow \mathcal{E}[v_M] \\
& \mathcal{E}[\mathbf{ms} \ (b \diamond t_M) \ w_S]_M \rightarrow \mathcal{E}[\mathbf{wrong} \ [b \diamond t_M] \ \text{“Brand mismatch”}] \\
& \quad (w_S \neq \mathbf{sm} \ (b \diamond t_M) \ e_M) \\
& \mathcal{E}[\mathbf{ms} \ (k_M \rightarrow k'_M) \ (\lambda x_S. e_S)]_M \rightarrow \\
& \quad \mathcal{E}[\lambda x_M : [k_M]. \mathbf{ms} \ k'_M \ ((\lambda x_S. e_S) \ (\mathbf{sm} \ k_M \ x_M))] \\
& \mathcal{E}[\mathbf{ms} \ (k_M \rightarrow k'_M) \ w_S]_M \rightarrow \mathcal{E}[\mathbf{wrong} \ [k_M \rightarrow k'_M] \ \text{“Not a function”}] \\
& \quad (w_S \neq \lambda x_S. e_S) \\
& \mathcal{E}[\mathbf{ms} \ (\forall u_M. k_M) \ w_S]_M \rightarrow \mathcal{E}[\Lambda u_M. \mathbf{ms} \ k_M \ w_S]
\end{aligned}$$

Figure 2.10: ML-Scheme operational semantics

$$\begin{aligned}
e_S &= x_S \mid v_S \mid e_S e_S \mid o e_S e_S \mid p e_S \mid \text{if0 } e_S e_S e_S \mid \text{cons } e_S e_S \mid f e_S \\
&\quad \text{wrong } string \mid \text{sm } k_M e_M \\
v_S &= w_S \mid \text{sh } k_H e_H \\
w_S &= \lambda x_S. e_S \mid \bar{n} \mid \text{nil} \mid \text{cons } v_S v_S \mid \text{sh } (b \diamond t_H) e_H \mid \text{sm } (b \diamond t_M) w_M \\
o &= + \mid - \\
f &= \text{hd} \mid \text{tl} \\
p &= \text{fun?} \mid \text{list?} \mid \text{null?} \mid \text{num?} \\
E_S &= U_S \mid \text{sh } k_H E_H \\
U_S &= []_S \mid E_S e_S \mid w_S U_S \mid o E_S e_S \mid o w_S E_S \mid p E_S \mid \text{if0 } E_S e_S e_S \\
&\quad \text{cons } U_S e_S \mid \text{cons } v_S U_S \mid f E_S \mid \text{sm } k_M E_M
\end{aligned}$$

Figure 2.11: Scheme grammar and evaluation contexts

$$\begin{array}{c}
\overline{\vdash_S \text{TST}} \\
\\
\frac{\Gamma, x_S : \text{TST} \vdash_S e_S : \text{TST}}{\Gamma \vdash_S \lambda x_S. e_S : \text{TST}} \quad \overline{\vdash_S \bar{n} : \text{TST}} \quad \overline{\vdash_S \text{nil} : \text{TST}} \\
\frac{\Gamma \vdash_S e_S : \text{TST} \quad \Gamma \vdash_S e'_S : \text{TST}}{\Gamma \vdash_S \text{cons } e_S e'_S : \text{TST}} \quad \overline{\Gamma, x_S : \text{TST} \vdash_S x_S : \text{TST}} \\
\frac{\Gamma \vdash_S e_S : \text{TST} \quad \Gamma \vdash_S e'_S : \text{TST}}{\Gamma \vdash_H e_S e'_S : \text{TST}} \quad \overline{\Gamma \vdash_S e_S : \text{TST}} \\
\overline{\Gamma \vdash_S f e_S : \text{TST}} \\
\frac{\Gamma \vdash_S e_S : \text{TST} \quad \Gamma \vdash_S e'_S : \text{TST}}{\Gamma \vdash_S o e_S e'_S : \text{TST}} \quad \overline{\Gamma \vdash_S e_S : \text{TST}} \\
\overline{\Gamma \vdash_S p e_S : \text{TST}} \\
\frac{\Gamma \vdash_S e_S : \text{TST} \quad \Gamma \vdash_S e'_S : \text{TST} \quad \Gamma \vdash_S e''_S : \text{TST}}{\Gamma \vdash_S \text{if0 } e_S e'_S e''_S : \text{TST}} \quad \overline{\vdash_S \text{wrong string} : \text{TST}} \\
\frac{\Gamma \vdash_H [k_H] \quad \Gamma \vdash_H e_H : t_H \quad [k_H] = t_H}{\Gamma \vdash_S \text{sh } k_H e_H : \text{TST}} \quad \frac{\Gamma \vdash_M [k_M] \quad \Gamma \vdash_M e_M : t_M \quad [k_M] = t_M}{\Gamma \vdash_S \text{sm } k_M e_M : \text{TST}}
\end{array}$$

Figure 2.12: Scheme typing rules

$$\begin{aligned}
& \mathcal{E}[(\lambda x_S. e_S) v_S]_S \rightarrow \mathcal{E}[e_S[v_S/x_S]] \\
& \mathcal{E}[w_S v_S]_S \rightarrow \mathcal{E}[\mathbf{wrong} \text{ "Not a function"}] \ (w_S \neq \lambda x_S. e_S) \\
& \mathcal{E}[+ \bar{n} \bar{n}']_S \rightarrow \mathcal{E}[\overline{n + n'}] \\
& \mathcal{E}[- \bar{n} \bar{n}']_S \rightarrow \mathcal{E}[\overline{\max(n - n', 0)}] \\
& \mathcal{E}[o w_S w'_S]_S \rightarrow \mathcal{E}[\mathbf{wrong} \text{ "Not a number"}] \ (w_S \neq \bar{n} \text{ or } w'_S \neq \bar{n}) \\
& \mathcal{E}[\text{if0 } \bar{0} e_S e'_S]_S \rightarrow \mathcal{E}[e_S] \\
& \mathcal{E}[\text{if0 } \bar{n} e_S e'_S]_S \rightarrow \mathcal{E}[e'_S] \ (n \neq 0) \\
& \mathcal{E}[\text{if0 } w_S e_S e'_S]_S \rightarrow \mathcal{E}[\mathbf{wrong} \text{ "Not a number"}] \ (w_S \neq \bar{n}) \\
& \mathcal{E}[f \text{ nil}]_S \rightarrow \mathcal{E}[\mathbf{wrong} \text{ "Empty list"}] \\
& \mathcal{E}[\text{hd} (\text{cons } v_S v'_S)]_S \rightarrow \mathcal{E}[v_S] \\
& \mathcal{E}[\text{tl} (\text{cons } v_S v'_S)]_S \rightarrow \mathcal{E}[v'_S] \\
& \mathcal{E}[f w_S]_S \rightarrow \mathcal{E}[\mathbf{wrong} \text{ "Not a list"}] \ (w_S \neq \text{nil} \text{ and } w_S \neq \text{cons } v_S v'_S) \\
& \mathcal{E}[\text{fun? } (\lambda x_S. e_S)]_S \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\text{fun? } w_S]_S \rightarrow \mathcal{E}[\bar{1}] \ (w_S \neq \lambda x_S. e_S) \\
& \mathcal{E}[\text{list? nil}]_S \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\text{list? } (\text{cons } v_S v'_S)]_S \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\text{list? } w_S]_S \rightarrow \mathcal{E}[\bar{1}] \ (w_S \neq \text{nil} \text{ and } w_S \neq \text{cons } v_S v'_S) \\
& \mathcal{E}[\text{null? nil}]_S \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\text{null? } w_S]_S \rightarrow \mathcal{E}[\bar{1}] \ (w_S \neq \text{nil}) \\
& \mathcal{E}[\text{num? } \bar{n}]_S \rightarrow \mathcal{E}[\bar{0}] \\
& \mathcal{E}[\text{num? } w_S]_S \rightarrow \mathcal{E}[\bar{1}] \ (w_S \neq \bar{n}) \\
& \mathcal{E}[\mathbf{wrong } string]_S \rightarrow \mathbf{Error: } string
\end{aligned}$$

Figure 2.13: Scheme operational semantics

$$\begin{aligned}
\mathcal{E}[\text{sh } L \text{ (hm } L \ k_M \ w_M)]_S &\rightarrow \mathcal{E}[\text{wrong "Bad value"}] \\
\mathcal{E}[\text{sh } L \text{ (hs } L \ w_S)]_S &\rightarrow \mathcal{E}[w_S] \\
\mathcal{E}[\text{sh } N \ \bar{n}]_S &\rightarrow \mathcal{E}[\bar{n}] \\
\mathcal{E}[\text{sh } \{k_H\} \text{ (nil } t_H)]_S &\rightarrow \mathcal{E}[\text{nil}] \\
\mathcal{E}[\text{sh } \{k_H\} \text{ (cons } e_H \ e'_H)]_S &\rightarrow \mathcal{E}[\text{cons } (\text{sh } k_H \ e_H) \ (\text{sh } \{k_H\} \ e'_H)] \\
\mathcal{E}[\text{sh } (k_H \rightarrow k'_H) \ (\lambda x_H : t_H.e_H)]_S &\rightarrow \\
&\quad \mathcal{E}[\lambda x_S.\text{sh } k'_H \ ((\lambda x_H : t_H.e_H) \text{ (hs } k_H \ x_S))] \\
\mathcal{E}[\text{sh } (\forall u_H.k_H) \ (\Lambda u'_H.e_H)]_S &\rightarrow \mathcal{E}[\text{sh } k_H[L/u_H] \ e_H[L/u'_H]]
\end{aligned}$$

Figure 2.14: Scheme-Haskell operational semantics

$$\begin{aligned}
\mathcal{E}[\mathbf{sm} \ L \ (\mathbf{mh} \ L \ k_H \ e_H)]_S &\rightarrow \mathcal{E}[\mathbf{wrong} \ \text{“Bad value”}] \\
\mathcal{E}[\mathbf{sm} \ L \ (\mathbf{ms} \ L \ w_S)]_S &\rightarrow \mathcal{E}[w_S] \\
\mathcal{E}[\mathbf{sm} \ N \ \bar{n}]_S &\rightarrow \mathcal{E}[\bar{n}] \\
\mathcal{E}[\mathbf{sm} \ \{k_M\} \ (\mathbf{nil} \ t_M)]_S &\rightarrow \mathcal{E}[\mathbf{nil}] \\
\mathcal{E}[\mathbf{sm} \ \{k_M\} \ (\mathbf{cons} \ v_M \ v'_M)]_S &\rightarrow \mathcal{E}[\mathbf{cons} \ (\mathbf{sm} \ k_M \ v_M) \ (\mathbf{sm} \ \{k_M\} \ v'_M)] \\
\mathcal{E}[\mathbf{sm} \ (k_M \rightarrow k'_M) \ (\lambda x_M : t_M.e_M)]_S &\rightarrow \\
&\quad \mathcal{E}[\lambda x_S. \mathbf{sm} \ k'_M \ ((\lambda x_M : t_M.e_M) \ (\mathbf{ms} \ k_M \ x_S))] \\
\mathcal{E}[\mathbf{sm} \ (\forall u_M.k_M) \ (\Lambda u'_M.e_M)]_S &\rightarrow \mathcal{E}[\mathbf{sm} \ k_M[L/u_M] \ e_M[L/u'_M]]
\end{aligned}$$

Figure 2.15: Scheme-ML operational semantics

$$\begin{aligned}
\lfloor L \rfloor &= L \\
\lfloor N \rfloor &= N \\
\lfloor u_H \rfloor &= u_H \\
\lfloor u_M \rfloor &= u_M \\
\lfloor \{k_H\} \rfloor &= \{\lfloor k_H \rfloor\} \\
\lfloor \{k_M\} \rfloor &= \{\lfloor k_M \rfloor\} \\
\lfloor k_H \rightarrow k_H \rfloor &= \lfloor k_H \rfloor \rightarrow \lfloor k_H \rfloor \\
\lfloor k_M \rightarrow k_M \rfloor &= \lfloor k_M \rfloor \rightarrow \lfloor k_M \rfloor \\
\lfloor \forall u_H. k_H \rfloor &= \forall u_H. \lfloor k_H \rfloor \\
\lfloor \forall u_M. k_M \rfloor &= \forall u_M. \lfloor k_M \rfloor \\
\lfloor b \diamond t_H \rfloor &= t_H \\
\lfloor b \diamond t_M \rfloor &= t_M
\end{aligned}$$

Figure 2.16: Unbrand function

$$\begin{aligned}
x &\dot{=} x \\
x &\dot{=} y \Rightarrow y \dot{=} x \\
x &\dot{=} y \text{ and } y \dot{=} z \Rightarrow x \dot{=} z \\
t_H &\dot{=} L \\
t_M &\dot{=} L \\
t_H = t_M &\Rightarrow t_H \dot{=} t_M
\end{aligned}$$

Figure 2.17: Lump equality relation

Chapter 3

Conclusion

This work resolved three language incompatibilities in a system of interoperation for three diverse languages. It resolved incompatible type systems with contracts for higher-order functions and lump types. It resolved incompatible support for parametricity with label types. It resolved incompatible evaluation strategies with delayed conversions for list constructions. It defined a model of computation that can express interoperation where the aforementioned incompatibilities arise and resolve them, provided a proof of its type soundness, and described an implementation of it that supported additional language features.

Bibliography

- [1] N. Benton. Embedded interpreters. *J. Funct. Program.*, 15(4):503–542, 2005.
- [2] M. Blume and D. McAllester. A sound (and complete) model of contracts. *SIGPLAN Not.*, 39(9):189–200, 2004.
- [3] ECMA. *Common Language Infrastructure (CLI)*, 4th edition, June 2006.
- [4] R. B. Findler and M. Felleisen. Contracts for higher-order functions. *SIGPLAN Not.*, 37(9):48–59, 2002.
- [5] A. Guha, J. Matthews, R. B. Findler, and S. Krishnamurthi. Relationally-parametric polymorphic contracts. In *DLS '07: Proceedings of the 2007 symposium on Dynamic languages*, pages 29–40, New York, NY, USA, 2007. ACM.
- [6] F. Henglein and J. Rehof. Safe polymorphic type inference for a dynamically typed language: translating scheme to ml. In *FPCA '95: Proceedings of the seventh international conference on Functional programming languages and computer architecture*, pages 192–203, New York, NY, USA, 1995. ACM.
- [7] D. L. Kinghorn. Preserving parametricity while sharing higher-order, polymorphic functions between scheme and ml. Master’s thesis, California Polytechnic State University, San Luis Obispo, June 2007.

- [8] J. Matthews and R. B. Findler. Operational semantics for multi-language programs. *SIGPLAN Not.*, 42(1):3–10, 2007.
- [9] Microsoft. *Microsoft Interface Definition Language*, November 2007.
- [10] Object Management Group. *Common Object Request Broker Architecture (CORBA) Specification*, 3.1 edition, August 2004.
- [11] B. C. Pierce. *Types and programming languages*. MIT Press, Cambridge, MA, USA, 2002.
- [12] S. Tobin-Hochstadt and M. Felleisen. Interlanguage migration: from scripts to programs. In *OOPSLA '06: Companion to the 21st ACM SIGPLAN conference on Object-oriented programming systems, languages, and applications*, pages 964–974, New York, NY, USA, 2006. ACM.