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 Faught

June 2008

AUTHORIZATION FOR REPRODUCTION OF MASTER'S THESIS

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

APPROVAL PAGE

TITLE: Interoperation for Incompatible E	valuation 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 Faught

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 v						
1	Inti	roduction	1			
2	Mo	del of Computation	4			
	2.1	Natural Numbers	7			
	2.2	Lists	8			
	2.3	Functions	8			
	2.4	Types	8			
	2.5	Interoperation	9			
3	Cor	nclusion	26			
$\mathbf{B}_{\mathbf{i}}$	Bibliography					

List of Figures

2.1	Function and list conversions with unforced points	6
2.2	Haskell grammar and evaluation contexts	10
2.3	Haskell typing rules	11
2.4	Haskell operational semantics	12
2.5	Haskell-ML operational semantics	13
2.6	Haskell-Scheme operational semantics	14
2.7	ML grammar and evaluation contexts	15
2.8	ML typing rules	16
2.9	ML operational semantics	17
2.10	ML-Haskell operational semantics	18
2.11	ML-Scheme operational semantics	19
2.12	Scheme grammar and evaluation contexts	20
2.13	Scheme typing rules	21
2.14	Scheme operational semantics	22
2.15	Scheme-Haskell operational semantics	23
2.16	Scheme-ML operational semantics	24
2.17	Unbrand function	25
2.18	Lump equality relation	25

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. Adhoc 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 comprises three dependent models of computation, based on that of Matthews and Findler (((CITATION))). The Haskell and ML models are based on System F, extended with a fixed-point operation. The Scheme model is based on lambda calculus, having a simple type system to ensure no free variables, and extended with type predicates. All models have natural numbers, arithmetic, conditions, lists, and errors. The Haskell model has a call-by-name evaluation strategy, and ML and Scheme have call-by-value evaluation strategies. The models are presented with grammars and operational semantics in the style of (((CITATION))) and typing rules.

The Haskell model has a (((STRICT?))) subset of the strictness points of the ML and Scheme models, and hence forces the reduction of fewer expressions where those expressions are not used. When a function from the Haskell model is converted to a function in the ML or Scheme models, this same subset must be preserved or the meaning of the function will change and parametricity will not hold. Concretely, this means that function arguments and list construction operands must not be reduced. Thus evaluation contexts for the ML and Scheme models are made aware of whether Haskell language boundary guards are in these places, and if so, to make them irreducible.

Since Haskell language boundary guards are forced in some places but not others, they must be considered a value sometimes, but not others. Thus there are two kinds of values: all values, called unforced values, which include imported Haskell expressions, and forced values, which exclude imported Haskell expressions. Unforced values occur in the evaluation contexts and reduction rules where Haskell importations should not be forced, namely function arguments and list construction operands, and forced values occur everywhere else a value is required.

Evaluation contexts are split into two: forced, E, and unforced, U. Only forced evaluation contexts can reduce anything, including Haskell importations, and unforced evaluation contexts restrict where expressions are forced. Where U appears in an evaluation context, any Haskell importation matching that expression is not forced because only E can reduce it.

Since the Haskell and ML models have their own types, type abstractions from one imported into the other cannot be easily converted, because any type the conversion is applied to cannot be substituted into the other language. Instead, and like the importation of parametric polymorphic types into the Scheme model, the lump type is substituted into the imported type abstraction, and a lump equality relation, (((LUMP EQUALITY REL))) asserts that corresponding parts of the inner and outer types of the importation must be equal, or one of them must be a lump.

2.1

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

Figure 2.1: Function and list conversions with unforced points.

$$zeroes = \texttt{fix} \ (\lambda x_H : \{\texttt{N}\}.\texttt{cons} \ \overline{\texttt{0}} \ x_H)$$

$$(\texttt{hs} \ (\{\texttt{N}\} \rightarrow \{\texttt{N}\}) \ (\lambda x_S.x_S)) \ zeroes \qquad \rightarrow \\ (\lambda x_H' : \{\texttt{N}\}.\texttt{hs} \ \{\texttt{N}\} \ ((\lambda x_S.x_S) \ (\texttt{sh} \ \{\texttt{N}\} \ x_H'))) \ zeroes \qquad \rightarrow \\ \texttt{hs} \ \{\texttt{N}\} \ ((\lambda x_S.x_S) \ (\texttt{sh} \ \{\texttt{N}\} \ zeroes)) \qquad \rightarrow \\ \texttt{hs} \ \{\texttt{N}\} \ (\texttt{sh} \ \{\texttt{N}\} \ zeroes)) \qquad \rightarrow \\ \texttt{hs} \ \{\texttt{N}\} \ (\texttt{cons} \ \overline{\texttt{0}} \ zeroes)) \qquad \rightarrow \\ \texttt{hs} \ \{\texttt{N}\} \ (\texttt{cons} \ \overline{\texttt{0}} \ (\texttt{sh} \ \{\texttt{N}\} \ zeroes)) \qquad \rightarrow \\ \texttt{cons} \ (\texttt{hs} \ \mathbb{N} \ \overline{\texttt{0}}) \ (\texttt{hs} \ \{\texttt{N}\} \ zeroes)) \qquad \rightarrow \\ \texttt{cons} \ (\texttt{hs} \ \mathbb{N} \ \overline{\texttt{0}}) \ (\texttt{hs} \ \{\texttt{N}\} \ zeroes))$$

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 2.2,

2.7, and 2.12.

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 \overline{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 if 0 e e e, test whether a value is the natural number $\overline{0}$. 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 $\overline{0}$ if true and $\overline{1}$ 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{array}{lll} e_{H} & = & x_{H} \mid v_{H} \mid e_{H} \mid e_{H} \mid e_{H} \mid t_{H} \mid | \text{ fix } e_{H} \mid o \mid e_{H} \mid e_{H} \mid | \text{ fin } e_{H} \mid
```

Figure 2.2: Haskell grammar and evaluation contexts

Figure 2.3: Haskell typing rules

```
\mathcal{E}[(\lambda x_H : t_H.e_H) \ e'_H]_H \to \mathcal{E}[e_H[e'_H/x_H]]
\mathcal{E}[(\Lambda u_H.e_H)\langle t_H \rangle]_H \to \mathcal{E}[e_H[b \diamond t_H/u_H]]
\mathcal{E}[\text{fix } (\lambda x_H : t_H.e_H)]_H \to \mathcal{E}[e_H[\text{fix } (\lambda x_H : t_H.e_H)/x_H]]
\mathcal{E}[+\overline{n} \ \overline{n}']_H \to \mathcal{E}[\overline{n+n'}]
\mathcal{E}[-\overline{n} \ \overline{n}']_H \to \mathcal{E}[\overline{max(n-n',0)}]
\mathcal{E}[\text{if } 0 \ \overline{0} \ e_H \ e'_H]_H \to \mathcal{E}[e_H]
\mathcal{E}[\text{if } 0 \ \overline{n} \ e_H \ e'_H]_H \to \mathcal{E}[e'_H] \ (n \neq 0)
\mathcal{E}[\text{hd } (\text{nil } t_H)]_H \to \mathcal{E}[\text{wrong } t_H \text{ "Empty list"}]
\mathcal{E}[\text{tl } (\text{nil } t_H)]_H \to \mathcal{E}[\text{wrong } \{t_H\} \text{ "Empty list"}]
\mathcal{E}[\text{hd } (\text{cons } e_H \ e'_H)]_H \to \mathcal{E}[e_H]
\mathcal{E}[\text{tl } (\text{cons } e_H \ e'_H)]_H \to \mathcal{E}[e'_H]
\mathcal{E}[\text{null? } (\text{nil } t_H)]_H \to \mathcal{E}[\overline{0}]
\mathcal{E}[\text{null? } (\text{cons } e_H \ e'_H)]_H \to \mathcal{E}[\overline{1}]
\mathcal{E}[\text{wrong } t_H \ string]_H \to \text{Error: } string
```

Figure 2.4: Haskell operational semantics

```
\begin{split} \mathscr{E}[\operatorname{hm} t_H \operatorname{L} (\operatorname{mh} \operatorname{L} t'_H e_H)]_H &\to \mathscr{E}[e_H] \quad (t_H = t'_H \text{ and } t_H \neq \operatorname{L}) \\ \mathscr{E}[\operatorname{hm} t_H \operatorname{L} (\operatorname{mh} \operatorname{L} t'_H e_H)]_H &\to \mathscr{E}[\operatorname{wrong} t_H \text{ "Type mismatch"}] \\ \quad (t_H \neq t'_H \text{ and } t_H \neq \operatorname{L}) \\ \mathscr{E}[\operatorname{hm} t_H \operatorname{L} (\operatorname{ms} \operatorname{L} w_S)]_H &\to \mathscr{E}[\operatorname{wrong} t_H \text{ "Bad value"}] \quad (t_H \neq \operatorname{L}) \\ \mathscr{E}[\operatorname{hm} \operatorname{N} \operatorname{N} \overline{n}]_H &\to \mathscr{E}[\overline{n}] \\ \mathscr{E}[\operatorname{hm} \{t_H\} \{t_M\} (\operatorname{nil} t'_M)]_H &\to \mathscr{E}[\operatorname{nil} t_H] \\ \mathscr{E}[\operatorname{hm} \{t_H\} \{t_M\} (\operatorname{cons} v_M v'_M)]_H &\to \\ \mathscr{E}[\operatorname{cons} (\operatorname{hm} t_H t_M v_M) (\operatorname{hm} \{t_H\} \{t_M\} v'_M)] \\ \mathscr{E}[\operatorname{hm} (t_H \to t'_H) (t_M \to t'_M) (\lambda x_M : t''_M . e_M)]_H &\to \\ \mathscr{E}[\lambda x_H : t_H . \operatorname{hm} t'_H t'_M ((\lambda x_M : t''_M . e_M) (\operatorname{mh} t_M t_H x_H))] \\ \mathscr{E}[\operatorname{hm} (\forall u_H . t_H) (\forall u_M . t_M) (\Lambda u'_M . e_M)]_H &\to \mathscr{E}[\Lambda u_H . \operatorname{hm} t_H t_M [\operatorname{L}/u_M] e_M [\operatorname{L}/u'_M]] \end{split}
```

Figure 2.5: Haskell-ML operational semantics

```
 \mathscr{E}[\text{hs N} \, \overline{n}]_H \to \mathscr{E}[\overline{n}] 
 \mathscr{E}[\text{hs N} \, w_S]_H \to \mathscr{E}[\text{wrong N "Not a number"}] \, (w_S \neq \overline{n}) 
 \mathscr{E}[\text{hs } \{k_H\} \, \text{nil}]_H \to \mathscr{E}[\text{nil } \lfloor k_H \rfloor] 
 \mathscr{E}[\text{hs } \{k_H\} \, (\text{cons } v_S \, v_S')]_H \to \mathscr{E}[\text{cons (hs } k_H \, v_S) \, (\text{hs } \{k_H\} \, v_S')] 
 \mathscr{E}[\text{hs } \{k_H\} \, w_S]_H \to \mathscr{E}[\text{wrong } \lfloor \{k_H\} \rfloor \, \text{"Not a list"}] 
 (w_S \neq \text{nil and } w_S \neq \text{cons } v_S \, v_S') 
 \mathscr{E}[\text{hs } (b \diamond t_H) \, (\text{sh } (b \diamond t_H) \, e_H)]_H \to \mathscr{E}[e_H] 
 \mathscr{E}[\text{hs } (b \diamond t_H) \, w_S]_H \to \mathscr{E}[\text{wrong } t_H \, \text{"Brand mismatch"}] \, (w_S \neq \text{sh } (b \diamond t_H) \, e_H) 
 \mathscr{E}[\text{hs } (k_H \to k_H') \, (\lambda x_S.e_S)]_H \to \mathscr{E}[\lambda x_H : \lfloor k_H \rfloor.\text{hs } k_H' \, ((\lambda x_S.e_S) \, (\text{sh } k_H \, x_H))] 
 \mathscr{E}[\text{hs } (k_H \to k_H') \, w_S]_H \to \mathscr{E}[\text{wrong } \lfloor k_H \to k_H' \rfloor \, \text{"Not a function"}] 
 (w_S \neq \lambda x_S.e_S) 
 \mathscr{E}[\text{hs } (\forall u_H.k_H) \, w_S]_H \to \mathscr{E}[\Lambda u_H.\text{hs } k_H \, w_S]
```

Figure 2.6: Haskell-Scheme operational semantics

Figure 2.7: ML grammar and evaluation contexts

Figure 2.8: ML typing rules

```
 \mathscr{E}[(\lambda x_M : t_M.e_M) \ v_M]_M \to \mathscr{E}[e_M[v_M/x_M]] 
 \mathscr{E}[(\Lambda u_M.e_M)\langle t_M\rangle]_M \to \mathscr{E}[e_M[b \diamond t_M/u_M]] 
 \mathscr{E}[\operatorname{fix} \ (\lambda x_M : t_M.e_M)]_M \to \mathscr{E}[e_M[\operatorname{fix} \ (\lambda x_M : t_M.e_M)/x_M]] 
 \mathscr{E}[+\overline{n} \ \overline{n}']_M \to \mathscr{E}[\overline{n+n'}] 
 \mathscr{E}[-\overline{n} \ \overline{n}']_M \to \mathscr{E}[\overline{max(n-n',0)}] 
 \mathscr{E}[\operatorname{if0} \ \overline{0} \ e_M \ e_M']_M \to \mathscr{E}[e_M] 
 \mathscr{E}[\operatorname{if0} \ \overline{n} \ e_M \ e_M']_M \to \mathscr{E}[e_M] \ (n \neq 0) 
 \mathscr{E}[\operatorname{hd} \ (\operatorname{nil} \ t_M)]_M \to \mathscr{E}[\operatorname{wrong} \ t_M \ \operatorname{"Empty \ list"}] 
 \mathscr{E}[\operatorname{tl} \ (\operatorname{nil} \ t_M)]_M \to \mathscr{E}[\operatorname{wrong} \ \{t_M\} \ \operatorname{"Empty \ list"}] 
 \mathscr{E}[\operatorname{hd} \ (\operatorname{cons} \ v_M \ v_M')]_M \to \mathscr{E}[v_M] 
 \mathscr{E}[\operatorname{tl} \ (\operatorname{nil} \ t_M)]_M \to \mathscr{E}[v_M] 
 \mathscr{E}[\operatorname{null?} \ (\operatorname{nil} \ t_M)]_M \to \mathscr{E}[\overline{0}] 
 \mathscr{E}[\operatorname{null?} \ (\operatorname{cons} \ v_M \ v_M')]_M \to \mathscr{E}[\overline{1}] 
 \mathscr{E}[\operatorname{wrong} \ t_M \ string]_H \to \operatorname{Error:} \ string
```

Figure 2.9: ML operational semantics

Figure 2.10: ML-Haskell operational semantics

```
\mathcal{E}[\operatorname{ms} \operatorname{N} \overline{n}]_{M} \to \mathcal{E}[\overline{n}]
\mathcal{E}[\operatorname{ms} \operatorname{N} w_{S}]_{M} \to \mathcal{E}[\operatorname{wrong} \operatorname{N} \operatorname{"Not a number"}] (w_{S} \neq \overline{n})
\mathcal{E}[\operatorname{ms} \{k_{M}\} \operatorname{nil}]_{M} \to \mathcal{E}[\operatorname{nil} \lfloor k_{M} \rfloor]
\mathcal{E}[\operatorname{ms} \{k_{M}\} (\operatorname{cons} v_{S} v_{S}')]_{M} \to \mathcal{E}[\operatorname{cons} (\operatorname{ms} k_{M} v_{S}) (\operatorname{ms} \{k_{M}\} v_{S}')]
\mathcal{E}[\operatorname{ms} \{k_{M}\} w_{S}]_{M} \to \mathcal{E}[\operatorname{wrong} \lfloor \{k_{M}\} \rfloor \operatorname{"Not a list"}]
(w_{S} \neq \operatorname{nil} \operatorname{and} w_{S} \neq \operatorname{cons} v_{S} v_{S}')
\mathcal{E}[\operatorname{ms} (b \diamond t_{M}) (\operatorname{sm} (b \diamond t_{M}) v_{M})]_{M} \to \mathcal{E}[v_{M}]
\mathcal{E}[\operatorname{ms} (b \diamond t_{M}) w_{S}]_{M} \to \mathcal{E}[\operatorname{wrong} \lfloor b \diamond t_{M} \rfloor \operatorname{"Brand mismatch"}]
(w_{S} \neq \operatorname{sm} (b \diamond t_{M}) e_{M})
\mathcal{E}[\operatorname{ms} (k_{M} \to k_{M}') (\lambda x_{S}.e_{S})]_{M} \to
\mathcal{E}[\lambda x_{M} : \lfloor k_{M} \rfloor.\operatorname{ms} k_{M}' ((\lambda x_{S}.e_{S}) (\operatorname{sm} k_{M} x_{M}))]
\mathcal{E}[\operatorname{ms} (k_{M} \to k_{M}') w_{S}]_{M} \to \mathcal{E}[\operatorname{wrong} \lfloor k_{M} \to k_{M}' \rfloor \operatorname{"Not a function"}]
(w_{S} \neq \lambda x_{S}.e_{S})
\mathcal{E}[\operatorname{ms} (\forall u_{M}.k_{M}) w_{S}]_{M} \to \mathcal{E}[\Lambda u_{M}.\operatorname{ms} k_{M} w_{S}]
```

Figure 2.11: ML-Scheme operational semantics

```
\begin{array}{lll} e_S & = & x_S \mid v_S \mid e_S \, e_S \mid o \, e_S \, e_S \mid p \, e_S \mid \text{ifO} \, e_S \, e_S \mid cons \, e_S \, e_S \mid f \, e_S \\ & & \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 \overline{n} \mid \text{nil} \mid 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{ifO} \, E_S \, e_S \, e_S \\ & & \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{array}
```

Figure 2.12: Scheme grammar and evaluation contexts

$$\overline{\vdash_S \mathsf{TST}}$$

$$\frac{\Gamma, x_S : \mathsf{TST} \vdash_S e_S : \mathsf{TST}}{\Gamma \vdash_S \lambda x_S.e_S : \mathsf{TST}} \quad \frac{}{\vdash_S \overline{n} : \mathsf{TST}} \quad \overline{\vdash_S \mathsf{nil} : \mathsf{TST}}}{\Gamma \vdash_S \mathsf{nil} : \mathsf{TST}}$$

$$\frac{\Gamma \vdash_S e_S : \mathsf{TST} \quad \Gamma \vdash_S e_S' : \mathsf{TST}}{\Gamma \vdash_S \mathsf{cons} \ e_S \ e_S' : \mathsf{TST}} \quad \overline{\Gamma, x_S : \mathsf{TST} \vdash_S x_S : \mathsf{TST}}}{\Gamma, x_S : \mathsf{TST} \vdash_S x_S : \mathsf{TST}}$$

$$\frac{\Gamma \vdash_S e_S : \mathsf{TST} \quad \Gamma \vdash_S e_S' : \mathsf{TST}}{\Gamma \vdash_B e_S e_S' : \mathsf{TST}} \quad \frac{\Gamma \vdash_S e_S : \mathsf{TST}}{\Gamma \vdash_S f \ e_S : \mathsf{TST}}$$

$$\frac{\Gamma \vdash_S e_S : \mathsf{TST} \quad \Gamma \vdash_S e_S' : \mathsf{TST}}{\Gamma \vdash_S o \ e_S e_S' : \mathsf{TST}} \quad \frac{\Gamma \vdash_S e_S : \mathsf{TST}}{\Gamma \vdash_S p \ e_S : \mathsf{TST}}$$

$$\frac{\Gamma \vdash_S e_S : \mathsf{TST} \quad \Gamma \vdash_S e_S' : \mathsf{TST}}{\Gamma \vdash_S if0 \ e_S e_S' e_S'' : \mathsf{TST}} \quad \frac{\Gamma \vdash_S e_S' : \mathsf{TST}}{\vdash_S \mathsf{wrong} \ string} : \mathsf{TST}$$

$$\frac{\Gamma \vdash_H [k_H] \quad \Gamma \vdash_H e_H : t_H \quad [k_H] = t_H}{\Gamma \vdash_S \mathsf{sh} k_H e_H : \mathsf{TST}} \quad \frac{\Gamma \vdash_M [k_M] \quad \Gamma \vdash_M e_M : t_M \quad [k_M] = t_M}{\Gamma \vdash_S \mathsf{sm} k_M \ e_M : \mathsf{TST}}$$

Figure 2.13: Scheme typing rules

```
\mathscr{E}[(\lambda x_S.e_S) \ v_S]_S \to \mathscr{E}[e_S[v_S/x_S]]
\mathscr{E}[w_S \ v_S]_S \to \mathscr{E}[\text{wrong "Not a function"}] \ (w_S \neq \lambda x_S.e_S)
\mathscr{E}[+\overline{n}\ \overline{n}']_S \to \mathscr{E}[\overline{n+n'}]
\mathscr{E}[-\overline{n}\ \overline{n}']_S \to \mathscr{E}[\overline{max(n-n',0)}]
\mathscr{E}[o\ w_S\ w_S']_S \to \mathscr{E}[\text{wrong "Not a number"}]\ (w_S \neq \overline{n} \ \text{or} \ w_S' \neq \overline{n})
\mathscr{E}[\mathsf{if0}\ \overline{0}\ e_S\ e_S']_S \to \mathscr{E}[e_S]
\mathscr{E}[\mathsf{if0}\ \overline{n}\ e_S\ e_S']_S \to \mathscr{E}[e_S']\ (n \neq 0)
\mathscr{E}[\mathtt{if0}\ w_S\ e_S\ e_S']_S \to \mathscr{E}[\mathtt{wrong}\ \mathrm{``Not\ a\ number''}]\ (w_S \neq \overline{n})
\mathscr{E}[f \text{ nil}]_S \to \mathscr{E}[\text{wrong "Empty list"}]
\mathscr{E}[\operatorname{hd}(\operatorname{cons} v_S v_S')]_S \to \mathscr{E}[v_S]
\mathscr{E}[\mathsf{tl}\;(\mathsf{cons}\;v_S\;v_S')]_S \to \mathscr{E}[v_S']
\mathscr{E}[f \ w_S]_S \to \mathscr{E}[\text{wrong "Not a list"}] \ (w_S \neq \text{nil and } w_S \neq \text{cons } v_S \ v_S')
\mathscr{E}[\operatorname{fun}?(\lambda x_S.e_S)]_S \to \mathscr{E}[\overline{0}]
\mathscr{E}[\mathtt{fun}? w_S]_S \to \mathscr{E}[\overline{1}] \ (w_S \neq \lambda x_S.e_S)
\mathscr{E}[\mathtt{list?\,nil}]_S \to \mathscr{E}[\overline{0}]
\mathscr{E}[\mathtt{list}? (\mathtt{cons} \ v_S \ v_S')]_S \to \mathscr{E}[\overline{0}]
\mathscr{E}[\mathtt{list}? w_S]_S \to \mathscr{E}[\overline{1}] \ (w_S \neq \mathtt{nil} \ \mathtt{and} \ w_S \neq \mathtt{cons} \ v_S \ v_S')
\mathscr{E}[\mathtt{null}?\,\mathtt{nil}]_S \to \mathscr{E}[\overline{0}]
\mathscr{E}[\mathtt{null}?\ w_S]_S \to \mathscr{E}[\overline{1}]\ (w_S \neq \mathtt{nil})
\mathscr{E}[\operatorname{num}? \overline{n}]_S \to \mathscr{E}[\overline{0}]
\mathscr{E}[\text{num}; w_S]_S \to \mathscr{E}[\overline{1}] \ (w_S \neq \overline{n})
\mathscr{E}[\mathsf{wrong}\ string]_S \to \mathbf{Error}: string
```

Figure 2.14: Scheme operational semantics

```
\begin{split} \mathscr{E}[\operatorname{sh} \mathsf{L} \ (\operatorname{hm} \mathsf{L} \ k_M \ w_M)]_S &\to \mathscr{E}[\operatorname{wrong} \ \text{``Bad value''}] \\ \mathscr{E}[\operatorname{sh} \mathsf{L} \ (\operatorname{hs} \mathsf{L} \ w_S)]_S &\to \mathscr{E}[w_S] \\ \mathscr{E}[\operatorname{sh} \mathsf{N} \ \overline{n}]_S &\to \mathscr{E}[\overline{n}] \\ \mathscr{E}[\operatorname{sh} \ \{k_H\} \ (\operatorname{nil} \ t_H)]_S &\to \mathscr{E}[\operatorname{nil}] \\ \mathscr{E}[\operatorname{sh} \ \{k_H\} \ (\operatorname{cons} \ e_H \ e'_H)]_S &\to \mathscr{E}[\operatorname{cons} \ (\operatorname{sh} \ k_H \ e_H) \ (\operatorname{sh} \ \{k_H\} \ e'_H)] \\ \mathscr{E}[\operatorname{sh} \ (k_H \to k'_H) \ (\lambda x_H : t_H.e_H)]_S &\to \\ \mathscr{E}[\lambda x_S.\operatorname{sh} \ k'_H \ ((\lambda x_H : t_H.e_H) \ (\operatorname{hs} \ k_H \ x_S))] \\ \mathscr{E}[\operatorname{sh} \ (\forall u_H.k_H) \ (\Lambda u'_H.e_H)]_S &\to \mathscr{E}[\operatorname{sh} \ k_H[\mathsf{L}/u_H] \ e_H[\mathsf{L}/u'_H]] \end{split}
```

Figure 2.15: Scheme-Haskell operational semantics

```
\begin{split} \mathscr{E}[\operatorname{sm} \mathsf{L} \; (\operatorname{mh} \mathsf{L} \; k_H \; e_H)]_S &\to \mathscr{E}[\operatorname{wrong} \; \text{``Bad value''}] \\ \mathscr{E}[\operatorname{sm} \mathsf{L} \; (\operatorname{ms} \mathsf{L} \; w_S)]_S &\to \mathscr{E}[w_S] \\ \mathscr{E}[\operatorname{sm} \mathsf{N} \; \overline{n}]_S &\to \mathscr{E}[\overline{n}] \\ \mathscr{E}[\operatorname{sm} \; \{k_M\} \; (\operatorname{nil} \; t_M)]_S &\to \mathscr{E}[\operatorname{nil}] \\ \mathscr{E}[\operatorname{sm} \; \{k_M\} \; (\operatorname{cons} \; v_M \; v_M')]_S &\to \mathscr{E}[\operatorname{cons} \; (\operatorname{sm} \; k_M \; v_M) \; (\operatorname{sm} \; \{k_M\} \; v_M')] \\ \mathscr{E}[\operatorname{sm} \; (k_M \to k_M') \; (\lambda x_M : t_M.e_M)]_S &\to \\ \mathscr{E}[\lambda x_S.\operatorname{sm} \; k_M' \; ((\lambda x_M : t_M.e_M) \; (\operatorname{ms} \; k_M \; x_S))] \\ \mathscr{E}[\operatorname{sm} \; (\forall u_M.k_M) \; (\Lambda u_M'.e_M)]_S &\to \mathscr{E}[\operatorname{sm} \; k_M[\mathsf{L}/u_M] \; e_M[\mathsf{L}/u_M']] \end{split}
```

Figure 2.16: Scheme-ML operational semantics

$$\begin{bmatrix} L \end{bmatrix} &= L \\
 \begin{bmatrix} N \end{bmatrix} &= N \\
 \begin{bmatrix} u_H \end{bmatrix} &= u_H \\
 \begin{bmatrix} u_M \end{bmatrix} &= u_M \\
 \begin{bmatrix} \{k_H\} \} \end{bmatrix} &= \{ \lfloor k_H \rfloor \} \\
 \begin{bmatrix} \{k_M\} \rfloor &= \{ \lfloor k_M \rfloor \} \end{bmatrix} \\
 \begin{bmatrix} k_H \to k_H \rfloor &= \lfloor k_H \rfloor \to \lfloor k_H \rfloor \\
 [k_M \to k_M] &= \lfloor k_M \rfloor \to \lfloor k_M \rfloor \\
 [\forall u_H \cdot k_H] &= \forall u_H \cdot \lfloor k_H \rfloor \\
 [\forall u_M \cdot k_M] &= \forall u_M \cdot \lfloor k_M \rfloor \\
 [b \diamond t_H] &= t_H \\
 [b \diamond t_M] &= t_M$$

Figure 2.17: Unbrand function

$$x \doteq x$$

$$x \doteq y \Rightarrow y \doteq x$$

$$x \doteq y \text{ and } y \doteq z \Rightarrow x \doteq z$$

$$t_H \doteq L$$

$$t_M \doteq L$$

$$t_H = t_M \Rightarrow t_H \doteq t_M$$

Figure 2.18: 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. SIG-PLAN 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.