

# Refinement-Based Game Semantics for Certified Components

Jérémie Koenig

2020

Current practices ensure software reliability through careful testing, but while testing can reveal the presence of bugs, it cannot entirely guarantee their absence. By contrast, *certified systems* come with a formal specification and a computer-checked proof of correctness, providing strong evidence that the system behaves as expected in all possible scenarios. Over the past decade, researchers have been able to build certified systems of significant size and complexity, including compilers, processor designs, operating system kernels and more. Building on these successes, the DeepSpec project proposes to assemble them as certified components to build large-scale heterogeneous certified systems.

However, by necessity, these certified components use a broad variety of semantic models and verification techniques. To connect them, we must first embed them into a common, general-purpose model. The work I present here unifies the foundations of certified abstraction layers, game semantics, algebraic effects and the refinement calculus to build models suitable for this task. We represent certified abstraction layers, interaction trees, and the certified compiler CompCert in a single framework supporting heterogeneous components, stepwise refinement, and data abstraction.

# Refinement-Based Game Semantics for Certified Components

DRAFT updated on August 11, 2020

A Dissertation Presented to the Faculty of the Graduate School of  
Yale University  
in Candidacy for the Degree of  
Doctor of Philosophy

by  
Jérémie Koenig

Dissertation Director:  
Zhong Shao

December 2020

© 2020 Jérémie Koenig. This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.



DRAFT

## **Acknowledgements**

DRAFT

# Contents

<b>Acknowledgements</b>	<b>ii</b>
<b>Contents</b>	<b>iii</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Certified systems at scale . . . . .	1
1.1.1 The DeepSpec project . . . . .	1
1.1.2 CompCert . . . . .	2
1.1.3 CertiKOS . . . . .	3
1.1.4 Horizontal composition . . . . .	5
1.1.5 Challenges . . . . .	5
1.2 General models for system behaviors . . . . .	6
1.2.1 Symmetric monoidal categories . . . . .	6
1.2.2 Game semantics . . . . .	7
1.2.3 Algebraic effects . . . . .	8
1.2.4 The refinement calculus . . . . .	8
1.3 Compiling certified components . . . . .	9
1.3.1 CompCert as a certified system . . . . .	9
1.3.2 CompCert as a tool for building certified systems . . . . .	10
1.3.3 Integrating CompCert into a general-purpose model . . . . .	10

1.4 Contributions . . . . .	11
-----------------------------	----

## **I Preliminaries 13**

### **2 Background 14**

2.1 Building certified systems . . . . .	14
--	----

2.1.1 Semantic domains . . . . .	14
----------------------------------	----

2.1.2 Specifications and refinement . . . . .	15
---	----

2.1.3 Compositionality . . . . .	16
----------------------------------	----

2.1.4 Monotonicity . . . . .	16
------------------------------	----

2.1.5 Types . . . . .	17
-----------------------	----

2.1.6 Abstraction . . . . .	17
-----------------------------	----

2.1.7 Embedding . . . . .	19
---------------------------	----

2.2 The refinement calculus . . . . .	21
---------------------------------------	----

2.2.1 Dual nondeterminism . . . . .	21
-------------------------------------	----

2.2.2 Distributivity . . . . .	22
--------------------------------	----

2.2.3 Program refinement . . . . .	22
------------------------------------	----

2.2.4 Nondeterministic choice in specifications . . . . .	23
---	----

2.2.5 The refinement calculus . . . . .	23
---	----

2.3 Game semantics . . . . .	24
------------------------------	----

2.3.1 Games . . . . .	25
-----------------------	----

2.3.2 Strategies . . . . .	25
----------------------------	----

2.4 Logical relations . . . . .	26
---------------------------------	----

2.4.1 Binary logical relations . . . . .	27
--	----

2.4.2 Relators . . . . .	27
--------------------------	----

2.4.3 Kripke relations . . . . .	28
----------------------------------	----

2.5 Algebraic effects . . . . .	29
---------------------------------	----

2.5.1 Effects theories . . . . .	29
----------------------------------	----

2.5.2 Effect signatures . . . . .	30
-----------------------------------	----

2.5.3 Computations as terms . . . . .	30
---------------------------------------	----

2.5.4	Interpreting effects	31
2.5.5	Monads	32
2.5.6	Terms as strategies	35
<b>3</b>	<b>A refresher on category theory</b>	<b>37</b>
3.1	Motivation	37
3.2	Basic definitions	38
3.2.1	Categories	38
3.2.2	Functors	40
3.2.3	Natural transformations	40
3.2.4	Adjunctions	42
3.3	Cartesian categories	44
3.3.1	Products	44
3.3.2	Generalized elements	45
3.3.3	Cartesian closure	47
3.3.4	Symmetric monoidal categories	49
3.4	Categorical algebra	50
3.4.1	Monads	50
<b>4</b>	<b>Games and dual nondeterminism</b>	<b>53</b>
4.1	Example: function specifications	53
4.1.1	Elementary function specifications	54
4.1.2	Angelic and demonic choices	54
4.1.3	Data abstraction	56
4.2	Refinement in game semantics	57
4.2.1	Strategies and refinement	57
4.2.2	Strategy specifications	59
4.3	Free completely distributive completions	60
4.3.1	Construction	61
4.3.2	Categorical characterization	61
4.3.3	Computational interpretation	62

4.4	Dually nondeterministic strategies . . . . .	63
<b>II</b>	<b>Refinement-based game semantics</b>	<b>66</b>
<b>5</b>	<b>Certified abstraction layers</b>	<b>67</b>
5.1	Introduction . . . . .	67
5.1.1	Abstraction layers . . . . .	67
5.1.2	Certified abstraction layers in CertiKOS . . . . .	68
5.1.3	Contributions . . . . .	69
5.2	Layer model . . . . .	70
5.2.1	Specification monad . . . . .	70
5.2.2	Layer interfaces . . . . .	71
5.2.3	Client code . . . . .	72
5.2.4	Layer implementations . . . . .	73
5.2.5	Correctness . . . . .	73
5.2.6	Composing certified abstraction layers . . . . .	76
5.3	Horizontal composition . . . . .	77
5.3.1	Layer implementations . . . . .	77
5.3.2	Layer interfaces sharing state . . . . .	78
5.3.3	Monoidal structure . . . . .	79
5.4	CompCertX . . . . .	80
5.4.1	CompCertX . . . . .	81
5.4.2	Extending layer interfaces with memory states . . . . .	81
5.5	Conclusion . . . . .	82
<b>6</b>	<b>The interaction specification monad</b>	<b>83</b>
6.1	Overview . . . . .	83
6.2	Plays . . . . .	84
6.3	Interaction specifications . . . . .	85
6.4	Interaction primitives . . . . .	86



6.5	Categorical structure . . . . .	87
6.6	Products . . . . .	88
6.7	Certified abstraction layers . . . . .	89
6.7.1	Signatures . . . . .	90
6.7.2	Interfaces . . . . .	90
6.7.3	Keeping state . . . . .	90
6.7.4	Simulation relations . . . . .	91
<b>7</b>	<b>Stateful and reentrant strategies</b>	<b>93</b>
7.1	Overview . . . . .	93
7.2	Games . . . . .	94
7.3	Plays and strategies . . . . .	95
7.4	Operations on strategies . . . . .	96
7.4.1	Composition . . . . .	97
7.4.2	Identity . . . . .	98
7.4.3	Tensor . . . . .	99
7.5	Category . . . . .	99
7.6	Embedding $\mathcal{G}_{\sqsubseteq}^{ib}$ . . . . .	100
7.7	Hiding state . . . . .	101
<b>III</b>	<b>Compiling certified open components</b>	<b>102</b>
<b>8</b>	<b>Semantics in CompCert</b>	<b>103</b>
8.1	Whole-program semantics in CompCert . . . . .	103
8.1.1	Transition systems . . . . .	104
8.1.2	Trace semantics . . . . .	105
8.1.3	Nondeterminism . . . . .	107
8.1.4	Forward simulations . . . . .	108
8.1.5	Backward simulations . . . . .	108
8.1.6	Memory model . . . . .	108

8.2	Compositional extensions . . . . .	109
8.2.1	CompCert and SepCompCert . . . . .	110
8.2.2	Contextual compilation . . . . .	110
8.2.3	Compositional CompCert . . . . .	111
8.2.4	CompCertM . . . . .	111
8.3	Limitations . . . . .	112
8.3.1	Decomposing heterogenous systems . . . . .	112
8.3.2	Requirements . . . . .	113
<b>9</b>	<b>CompCertO</b>	<b>115</b>
9.1	Overview . . . . .	117
9.1.1	Semantic model . . . . .	117
9.1.2	Simulations . . . . .	117
9.1.3	Simulation conventions . . . . .	118
9.1.4	Simulation convention algebra . . . . .	120
9.2	Operational semantics . . . . .	120
9.2.1	Open semantics in CompCertO . . . . .	120
9.2.2	Open simulations . . . . .	123
9.3	Simulation convention algebra . . . . .	126
9.3.1	Refinement of simulation conventions . . . . .	126
9.3.2	Kleene algebra . . . . .	127
9.3.3	Compiler correctness . . . . .	129
9.3.4	Compositional compilation and verification . . . . .	130
9.4	Evaluation . . . . .	131
<b>10</b>	<b>Passes of CompCertO</b>	<b>133</b>
10.1	Logical relations for the CompCert memory model . . . . .	134
10.1.1	Memory extensions . . . . .	134
10.1.2	Memory injections . . . . .	134
10.1.3	Modal Kripke relators . . . . .	135
10.1.4	CompCert Kripke logical relations . . . . .	136

10.1.5	From CKLRs to simulation conventions	137
10.1.6	External calls in injection passes	138
10.1.7	Discussion: world transitions and compositionality	139
10.1.8	Properties	140
10.2	Invariants	141
10.2.1	Invariants and language interfaces	142
10.2.2	Typing invariants	142
10.2.3	Value analysis	143
10.2.4	Simulations modulo invariants	144
10.3	Specialized simulation conventions	144
10.3.1	The Allocation pass	145
10.3.2	The Stacking pass	145
10.3.3	The Asmgcn pass	146
<b>IV</b>	<b>Conclusion</b>	<b>148</b>
<b>11</b>	<b>Conclusion and future work</b>	<b>149</b>
11.1	Categorical characterization of $\mathcal{I}_E$	149
11.2	Generalizations of $\mathcal{I}_E$	150
11.3	Richer game models	150
11.4	Linear logic	150
11.5	Parametric language for CompCert memory model	150

# List of Figures

1.1	Structure of the certified compiler CompCert . . . . .	3
1.2	Structure of the certified OS kernel CertiKOS . . . . .	4
2.1	A selection of relators . . . . .	28
5.1	Certified abstraction layer implementing a bounded queue . . . . .	72
8.1	CompCert’s approximation of the C toolchain . . . . .	103
8.2	Outline of the CompCert memory model . . . . .	109
9.1	Two simple C compilation units and their assembly code . . . . .	116
9.2	Simulation identity and vertical composition . . . . .	119
9.3	Horizontal composition of open semantics . . . . .	122
9.4	Forward simulation properties . . . . .	124
9.5	Forward simulation properties (cont.) . . . . .	124
10.1	Defining properties of CKLRs . . . . .	136
10.2	External calls and memory injections . . . . .	139
10.3	Simulation with invariants . . . . .	143
10.4	Separation of arguments in the stacking simulation convention. . . . .	146

# List of Tables

3.1	Some categories relevant to my work . . . . .	39
3.2	Informal correspondence between category theory and programming concepts . .	41
5.1	Correspondance between certified compilation and certified abstraction layers . .	68
8.1	Taxonomy of CompCert extensions . . . . .	114
9.1	Summary of notations . . . . .	119
9.2	Language interfaces used in CompCertO . . . . .	121
9.3	Languages and passes of CompCertO . . . . .	128
9.4	Significant lines of code in CompCertO . . . . .	131

# Chapter 1

## Introduction

### 1.1 Certified systems at scale

Certified software [Shao, 2010] is software accompanied by a mechanized, machine-checkable proof of correctness. To construct a certified program, we must not only write its code in a given programming language, but also formally specify its intended behavior and use specialized tools to construct evidence that the program indeed conforms to the specification.

The past decade has seen an explosion in the scope and scale of practical software verification. Researchers have been able to produce certified compilers [Leroy, 2009], program logics [Appel, 2011], operating system kernels [Gu et al., 2015; Klein et al., 2009], file systems [Chen et al., 2015], processor designs [Azevedo de Amorim et al., 2014; Choi et al., 2017] and more, often introducing new techniques and mathematical models. In this context, there has been increasing interest in making these components interoperable and combining them—and their proofs of correctness—into larger certified systems.

#### 1.1.1 The DeepSpec project

For example, the DeepSpec project [Appel et al., 2017] seeks to connect various components specified and verified in the Coq proof assistant. The key idea behind DeepSpec is to use specifications as *interfaces* between components. When a component providing a certain interface has been verified, client components can rely on this for their own proofs of correctness. Standardizing this process would make it possible to construct large-scale certified systems by assembling off-the-

shelf certified components.

This approach promises benefits beyond the potential increase in the scale of certified systems. As it stands, a certified system is only as trustworthy as its specification. Indeed, it is possible to prove a buggy system correct with respect to a buggy specification. If the specification is only validated by a human expert subjecting it to careful examination, these bugs could go unnoticed and persist in the final system. By contrast, if the same specification is used as a premise in the correctness proof of a client component, its deficiencies will become apparent and prevent us from carrying out this second proof.

Moreover, internal specifications used as intermediate steps in the verification of a complex system disappear from its external characterization and no longer need to be trusted. This reduces the ratio between the size of the system and the size of the trusted specification.

### 1.1.2 CompCert

To an extent, these principles are already demonstrated in the structure of the certified C compiler CompCert [Leroy, 2009]. CompCert can emit assembly code for PowerPC, ARM, RISC-V and x86 processors, and supports an extensive subset of ISO C99 as its source language. However, in the context of verification, CompCert is usually used as compiler for the Clight language, which is a simplified dialect of C more amenable to formal reasoning. The architecture-dependent assembly language which CompCert compiles to is known as Asm.

As a *certified* compiler, CompCert comes with a proof of correctness. Like all components used in the DeepSpec project, CompCert was specified and verified in the Coq proof assistant. The proof shows that if the compiler successfully transforms a source program  $p$  into a target program  $p'$ , then the behavior of the target program *refines* that of the source program:

$$\frac{\text{CompCert}(p) = p'}{\text{Clight}[p] \supseteq \text{Asm}[p']}$$

In the statement above, the semantics of the source and target programs are expressed as the sets of traces  $\text{Clight}[p]$  and  $\text{Asm}[p']$ . Each trace records a possible execution of the corresponding program as a sequence of events. The trace containment property  $\text{Clight}[p] \supseteq \text{Asm}[p']$  expresses that any possible behavior of the target program  $p'$  is a possible behavior of the source program  $p$ .

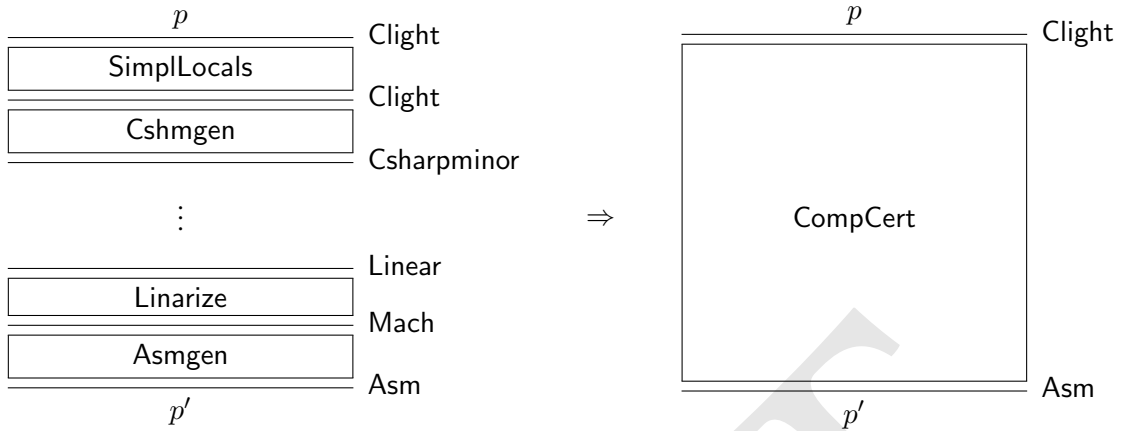


Figure 1.1: Structure of the certified compiler CompCert. The source program  $p$  is progressively transformed into the target program  $p'$  by successive compilation passes, depicted here as rectangular boxes. The horizontal line above each compilation pass depicts its source language, and the one below depicts its target language. Two passes can be composed when the target language of the first one corresponds to the source language of the second one. See also Table 9.3.

The key to this achievement was the decomposition of CompCert into compilation passes, which were then verified individually. When passes are composed to obtain the overall C-to-assembly compiler (Fig. 1.1), their correctness proofs can be composed as well to establish the overall correctness of the compiler. The final theorem does not mention the intermediate programs or language semantics, so that a user only needs to trust the accuracy of the Clight and Asm semantics, and the soundness of the proof assistant.

The introduction of CompCert in 2008 represented a breakthrough in the scale and feasibility of certified software. Since then, CompCert itself has been used as a platform other projects have built upon. For example, verification tools have been created with soundness proofs connecting to CompCert [Appel, 2011; Jourdan, 2016], and the composition techniques used to verify CompCert have been extended in various directions [Kang et al., 2016; Song et al., 2019; Stewart et al., 2015].

### 1.1.3 CertiKOS

The techniques used in CompCert also provided a blueprint for the verification of the operating system kernel CertiKOS [Gu et al., 2015, 2016, 2018]. CertiKOS is divided into several dozen *certified abstraction layers*, which we specified and verified individually.

*Layer interfaces* provide an abstract view of a layer’s functionality, hiding the procedural details



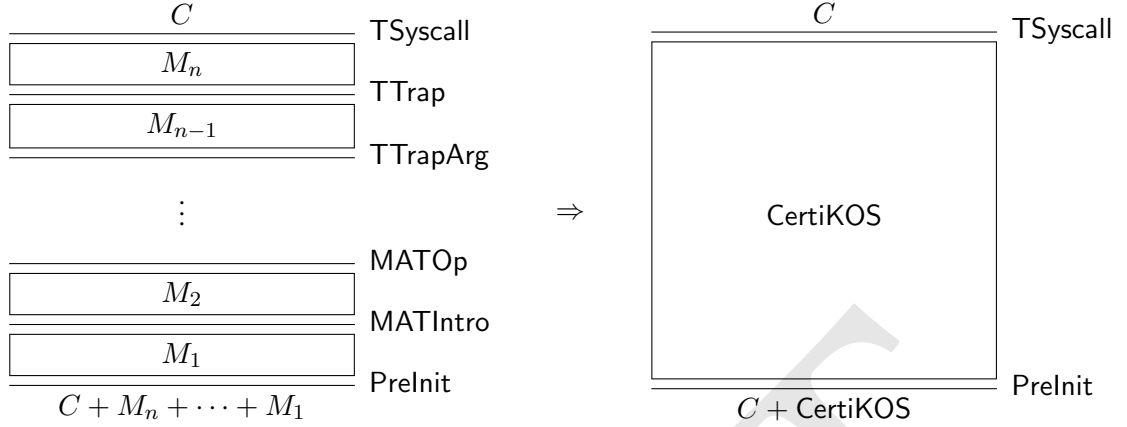


Figure 1.2: Structure of the certified OS kernel CertiKOS. Here, boxes represent certified abstraction layers and horizontal lines represent layer interfaces. The high-level client program  $C$  is transformed into the complete low-level program  $C + \text{CertiKOS}$  by linking it with the successive certified abstraction layers of CertiKOS. As before, the correctness properties of layers can be composed to derive a correctness property for the whole system.

and low-level data representations involved in its implementation. A layer interface enumerates the primitives implemented by an abstraction layer and specifies their expected behavior. Client code can then be verified in terms of this abstract view in order to build higher-level layers.

To make this possible, we extended the CompCert semantics to be parametrized by a layer interface: the expression  $\text{Asm}_L[C]$  denotes the set of traces generated by the client program  $C$  running on top of the layer interface  $L$ . Then a layer  $M$  implements an *overlay* interface  $L_2$  on top of an *underlay* interface  $L_1$  when the following *contextual refinement* property holds:

$$\forall C \cdot \text{Asm}_{L_2}[C] \supseteq \text{Asm}_{L_1}[C + M].$$

Here, the execution of the client code  $C$  on top of the overlay  $L_2$  serves as the specification. The property shows that this specification can be implemented by running  $C$  together with the layer implementation  $M$  on top of the underlay interface  $L_1$ .

Certified abstraction layers with compatible interfaces can then be chained together, in the same way passes of a compiler can be composed when the target language of one corresponds to the source language of another, allowing us to derive a contextual refinement property for the whole kernel (Fig. 1.2). The system call interface of CertiKOS is specified by the topmost layer

interface `TSyscall`, and the basic facilities offered by the hardware are formalized as the base layer interface `Prelnit`. Then, by decomposing  $\text{CertiKOS} = M_n + \dots + M_1$  into 37 certified abstraction layers, we were able to derive the overall theorem:

$$\forall C \cdot \text{Asm}_{\text{TSyscall}}[C] \supseteq \text{Asm}_{\text{Prelnit}}[C + \text{CertiKOS}].$$

#### 1.1.4 Horizontal composition

In addition to the *vertical* composition principles outlined above, `CompCert` and `CertiKOS` provide limited forms of *horizontal* composition, which allow individual programs and layer implementations to be decomposed further.

In `CompCert`, this is used to model *separate compilation*. The original correctness theorem of `CompCert` only characterized the compilation of a whole program, but in practice C programs usually consist of several `.c` files, known as *compilation units*. These components are compiled independently, and the results are combined through *linking* to build an executable image. To reflect this, [Kang et al. \[2016\]](#) generalized the correctness theorem of `CompCert` to the separate compilation property:

$$\frac{\forall 1 \leq i \leq k \cdot \text{CompCert}(p_i) = p'_i}{\text{Clight}[p_1 + \dots + p_k] \supseteq \text{Asm}[p'_1 + \dots + p'_k]}$$

Likewise, in `CertiKOS` the verification of a given layer can be decomposed into the verification of the individual functions which constitute its implementation. This is achieved through the *layer calculus* presented in Chapter 5.

#### 1.1.5 Challenges

The compositional approach to verification presented above is compelling, but using it to build large-scale certified systems out of disparate certified components is difficult. A key aspect enabling composition in `CompCert` and `CertiKOS` is the uniformity of the models underlying their language semantics and correctness proofs. By contrast, across certified software projects there exists a great diversity of semantic models and verification techniques. This makes it difficult to formulate specifications for interfacing specific projects, let alone devise a general method for

connecting certified components.

Worse yet, this diversity is not simply a historical accident. The semantic models used in the context of individual verification projects are often carefully chosen to make the verification task tractable. The semantic model used in CompCert alone has changed multiple times, addressing new requirements and techniques which were introduced alongside new compiler features and optimizations [Leroy, 2012].

Given the difficulty of verification, preserving this flexibility is essential. Therefore, to make it possible to link components verified using a variety of paradigms, we must identify a model which is expressive enough to embed the semantics, specifications and correctness proofs of all these paradigms. The model should provide high-level composition and reasoning principles, allowing us to construct large-scale certified systems.

## 1.2 General models for system behaviors

Fortunately, there is a wealth of semantics research to draw from as we attempt to design general models for certified components. I outline some of it below.

### 1.2.1 Symmetric monoidal categories

As a whole, category theory provides a general study and taxonomy for compositional structures found across mathematics. In a category, two components (morphisms) can be chained together when the interface which the first one provides (its target object) matches the interface which the second one relies on (its source object). For example, as described above, compilation passes and certified abstraction layers both constitute categories.

Symmetric monoidal categories also capture various kinds of *product* constructions, allowing components to be connected not only in series ( $\circ$ ) but also in parallel ( $\otimes$ ):

$$\frac{f : A \rightarrow B \quad g : B \rightarrow C}{g \circ f : A \rightarrow C} \quad \frac{f_1 : A_1 \rightarrow B_1 \quad f_2 : A_2 \rightarrow B_2}{f_1 \otimes f_2 : A_1 \otimes A_2 \rightarrow B_1 \otimes B_2}$$

As such, symmetric monoidal categories can be constructed to model a wide range of systems and processes [Baez and Stay, 2010]. The basic setup can be refined by introducing additional

structures, for instance modeling feedback loops (traced monoidal categories) or allowing wires to run in both directions (closed monoidal categories).

Symmetric monoidal categories appear in different forms in many approaches to logic and programming language semantics. For example, *cartesian closed categories* correspond to models of the simply typed lambda calculus, and are a special case of symmetric monoidal categories. However, symmetric monoidal categories in general do not require multiple components to be able to connect to the same interface ( $\Delta : A \rightarrow A \otimes A$ ), giving us more flexibility when modeling resources. In the same way the simply-typed lambda calculus provides an internal language for cartesian closed categories, various fragments of *linear logic* provide internal languages for various kinds of symmetric monoidal categories.

For our purposes, the theory of symmetric monoidal categories constitutes a repository of algebraic structures formalizing the compositional aspects of all kinds of systems, and can be used as a blueprint as we attempt to design general models for certified heterogeneous systems.

### 1.2.2 Game semantics

A specific way to construct symmetric monoidal categories is *game semantics* [Abramsky, 2010], which uses two-player games to describe the form of the interaction between a program component and its environment, and represents the externally observable behavior of the component as a strategy in this game.

Game semantics represents a synthesis of various approaches to programming language semantics. Game semantics is a form of *denotational semantics*, defining the behavior of complex programs in terms of the behavior of their components. However, since it models components by describing their external interactions across time, game semantics also feature a strong *operational* flavor. Finally, the usual construction of the strategies used in game semantics is a variation on the *trace semantics* used in the context of process algebras and concurrent semantics.

An early success of game semantics was the formulation of the first satisfactory fully complete models for the programming language PCF [Abramsky et al., 2000; Hyland and Ong, 2000]. Since then, game semantics have provided fully complete models for a wide variety of other languages, incorporating features such as state [Abramsky and McCusker, 1997], control [Laird, 1997], general references [Abramsky et al., 1998], concurrency [Ghica and Murawski, 2004] and more.

There are deep connections between game semantics and linear logic [Blass, 1993], and hence symmetric monoidal categories, hinting to its promise as a general approach for modeling the behavior of various kinds of systems and processes. In particular, the typed aspect of many game models makes them well-suited for describing the behavior of heterogeneous systems. However, the generality of game models often translates to a fair amount of complexity, which imposes a high barrier to entry for practitioners and makes them difficult to formalize in a proof assistant.

### 1.2.3 Algebraic effects

While more restricted, the framework of *algebraic effects* [Plotkin and Power, 2001] is sufficient for many modeling tasks, fits within the well-known monadic approach to effectful and interactive computations, and can be seen as a particularly simple version of game semantics.

In this framework, computations are modeled as the *terms* of an algebra whose operations correspond to the available effects. The computation proceeds inward, with each function symbol representing an occurrence of an effect, and each argument representing a possible way for the computation to be continued after the effect has been triggered.

One major advantage of this approach is that the methods and results of universal algebra become available to reason about effectful computations. Equational theories can be used to characterize the behavior of effects, and *interpretations* of one effect algebra into another can be used to model *effect handlers* [Plotkin and Pretnar, 2009]. The *free monad* associated to an algebra can be used to recover the more familiar monadic model of computational effects [Moggi, 1991]. Along these lines, *interaction trees* [Xia et al., 2019] have been developed for use in and across several DeepSpec projects.

Algebraic effects can also be understood as a limited form of game semantics: signatures can be interpreted as simple games, and the abstract syntax trees of terms can be interpreted as strategies. However, their restriction to first-order computation make them easier to formalize and reason about than more general notions of games.

### 1.2.4 The refinement calculus

While game semantics and effect systems have been proposed for a wide variety of programming languages, there has been comparatively less focus on specifications and correctness properties

for game semantics and algebraic effects. By contrast, the general approach of *stepwise refinement*, which suggests a uniform treatment of programs and specifications, has been studied extensively in the context of Dijkstra’s predicate transformer semantics [Dijkstra, 1975] and in the framework known as the *refinement calculus* [Back and Wright, 1998].

In refinement-based approaches, programs and specifications are expressed in a common language, and a certified program is constructed in an incremental manner, by applying a series of correctness-preserving transformation to the (abstract and declarative) specification until we obtain a (concrete and executable) program. Correctness preservation is expressed by a reflexive and transitive *refinement relation*. Language constructions are monotonic with respect to this relation, so that elementary refinement rules can be applied congruently within any program context.

To make it possible to express specifications, the language is extended with non-executable constructions, including in particular infinitary versions of both *angelic* and *demonic* nondeterministic choice operators. In its modern presentation, the refinement calculus is formulated in a lattice-theoretic framework where joins ( $\sqcup$ ) and meets ( $\sqcap$ ) correspond respectively to angelic and demonic choices. The resulting language is remarkably expressive and requires very few additional primitive constructions. The duality inherent in this approach also lends itself to game-theoretic interpretations, and indeed the semantics of the refinement calculus can be expressed as a two-player game between the angel and the demon.

### 1.3 Compiling certified components

There are two ways to look at CompCert in the context of software verification: as a certified system with an interesting structure, or as a tool we can use to build certified programs.

#### 1.3.1 CompCert as a certified system

Until now, I have mostly discussed CompCert as a specific example of a certified system, describing the compositional structures used in its construction as a precedent for the development of certified systems of significant size. From this point of view, the language semantics and correctness theorem formalized alongside the compiler’s code are components of its specification, and are used to formalize what it means for a C compiler to be correct.

The long-standing effort to refine and generalize these semantics and proofs can be understood as an attempt to model real-world compiler usage more accurately, preventing bugs in the compiler from making it through the verification process due to simplifications in its specification.

Connecting CompCert with other components could mean combining the proof of CompCert with that of a certified assembler and certified linker, perhaps even a certified processor design, to verify a larger portion of the toolchain end-to-end, and guarantee that programs compiled by the user will be executed in a way that conforms to the C standard.

### 1.3.2 CompCert as a tool for building certified systems

CompCert is also used as a tool for constructing other certified systems. For example, CertiKOS uses CompCert's assembly semantics to model the execution of the kernel's code and formulate the correctness property which it must satisfy. In addition, CompCert is used to compile the portions of the kernel which are written in C, and the compiler's correctness theorem allows us to use code proofs carried out with respect to the source code to prove the compiled assembly code correct.

CompCert serves in this case as a rudimentary *framework* for the construction of certified C and assembly programs in the Coq proof assistant. Efforts to increase the precision and flexibility of the compiler's correctness theorem have made this framework more general and compositional, even though given the specialized nature of this work it is difficult to imagine it could be extended to reach the level of generality which I have discussed.

### 1.3.3 Integrating CompCert into a general-purpose model

Still, given the importance of compilers in the construction of present-day computer systems, and of CompCert in the formal methods landscape, its integration into any framework attempting to tackle end-to-end verification should be a litmus test, and would provide immediate benefits:

- Embedding the transition systems used to define CompCert semantics into a general model would immediately augment it with a semantics for C and assembly programs. For versions of CompCert semantics which model individual translation units, this would give us the ability to interface program components at a granular level with components of different kinds, for instance file systems, network services, hardware devices, ...

- Then, in conjunction with a sufficiently precise correctness theorem, CompCert would not simply provide a certified compiler, but in fact a *compiler of certified program components*, showing that the interaction of a C translation unit with its environment is reflected by the corresponding compiled assembly code, as specified by the C calling convention in use.

Unfortunately, despite the abundance of research extending CompCert in this direction, none is flexible enough to be used in this way. As noted by [Patterson and Ahmed \[2019\]](#), the problem of certified compositional compilation is extremely complex and challenging, lacking even a single commonly accepted definition. In fact, as illustrated by the connections between certified abstraction layers and certified compilation, the problem is almost as general as the full-blown programme which I am pursuing.

On the flip side, this means that the techniques and conceptual framework developed in this thesis are a perfect match for understanding and tackling the very extensive form of compositional certified compilation described above, and indeed the culmination of my work is a version of CompCert which solves this challenge.

## 1.4 Contributions

The central claim of this dissertation is that a synthesis of game semantics, algebraic effects, and the refinement calculus can be used to construct a hierarchy of semantic models suitable for constructing large-scale, heterogeneous certified systems:

- Part I introduces relevant background and the general approach of *refinement-based game semantics*. A reexamination of the foundations of game semantics under the lens of dual nondeterminism invites us to regard plays as *elementary specifications* for the behaviors of interactive systems. The completion of plays with arbitrary angelic and demonic choices yields a notion of *strategy specification*, which provides support for stepwise refinement and data abstraction in the context of game semantics.
- Part II demonstrates the use of this approach for certified abstraction layers. Starting from the layer calculus of CertiKOS, I construct increasingly expressive models where layer interfaces, layer implementations and simulation relations can be treated in a uniform and



compositional way.

- Part III presents my work on the certified compiler CompCert. I examine existing models and techniques under the lens of refinement-based game semantics, and articulate why none of the existing CompCert extensions can be integrated within the framework in a completely satisfactory way. I then introduce CompCertO, the first extension of CompCert to fit the bill, which nevertheless avoids much of the complexity found in previous approaches.

Most of this work has been formalized in the Coq proof assistant. At the time of submission, the beginning of a Coq library for refinement-based game semantics is available at:

<https://github.com/CertiKOS/rbgs/>

A complete implementation of CompCertO can also be found at the following address:

<https://github.com/CertiKOS/compcert/tree/compcerto>

## **Part I**

# **Preliminaries**

## Chapter 2

# Background

Refinement-based game semantics combines several lines of research which so far have largely been understood as separate. This chapter aims to provide a short introduction to each one, and give the reader a sense of how they relate to one another.

I begin by outlining in §2.1 general principles which are used in the construction of certified systems. In the following sections, I present the *refinement calculus* (§2.2), *game semantics* (§2.3), *logical relations* (§2.4) and *algebraic effects* (§2.5).

### 2.1 Building certified systems

#### 2.1.1 Semantic domains

The goal of certified system design is to create a formal description of the system to be constructed (the program), while ensuring through careful analysis that the system will behave properly. To this end, we assign to every system  $p \in P$  a mathematical object  $\llbracket p \rrbracket \in \mathbb{D}$  representing its behavior. We will call the set  $\mathbb{D}$  a *semantic domain*.

**Example 2.1** (Trace semantics). *In CompCert, the behaviors of programs are ultimately expressed as sets of traces. Each trace records a possible execution of the program as a sequence of externally observable events taken from a set  $\mathbb{E}$ . The program may then terminate with a result taken in a set  $R$ , exhibit undefined behaviors ( $\downarrow$ ) or silently diverge ( $\uparrow$ ). The corresponding semantic domain can be*

defined as:

$$\mathbb{D}_{\text{CompCert}} := \mathcal{P}(\mathbb{E}^* R + \mathbb{E}^* \downarrow + \mathbb{E}^* \uparrow + \mathbb{E}^\omega).$$

In the remainder of this section I elucidate the structure and properties of  $\mathbb{D}$  necessary to the process of building large-scale certified systems.

### 2.1.2 Specifications and refinement

System design starts with a set of requirements on the behavior of the system to be constructed (the specification). These requirements do not capture every detail of the eventual system, but delineate a range of acceptable behaviors.

In refinement-based approaches, programs and specifications are interpreted in the same semantic domain  $\mathbb{D}$ , which is equipped with a *refinement* preorder  $\sqsubseteq \subseteq \mathbb{D} \times \mathbb{D}$ . The proposition  $\sigma_1 \sqsubseteq \sigma_2$  asserts that  $\sigma_1$  is refined by  $\sigma_2$ . The correctness of a system description  $p \in P$  with respect to a specification  $\sigma \in \mathbb{D}$  can then be formulated as  $\sigma \sqsubseteq \llbracket p \rrbracket$ .

Refinement is a correctness-preserving relation, in the sense that for a set  $\mathbb{P} \subseteq \mathcal{P}(\mathbb{D})$  of properties of interest and for two semantic objects  $\sigma_1, \sigma_2 \in \mathbb{D}$ :

$$\sigma_1 \sqsubseteq \sigma_2 \Rightarrow \forall P \in \mathbb{P}. P(\sigma_1) \Rightarrow P(\sigma_2).$$

**Example 2.2** (Trace containment). *In the trace semantics of CompCert, the set of traces associated with the source program is understood as a set of permissible behaviors for the target program. As a first approximation, the corresponding notion of refinement is trace containment, expressed by the superset relation  $\supseteq$ , so that for example  $\text{Clight}[p] \in \mathbb{D}_{\text{CompCert}}$  is refined by  $\text{Asm}[p'] \in \mathbb{D}_{\text{CompCert}}$  when  $\text{Clight}[p] \supseteq \text{Asm}[p']$ . Any property  $P(\sigma)$  asserting that all traces  $t \in \sigma \in \mathbb{D}_{\text{CompCert}}$  satisfy a given condition will be preserved by trace containment.*

CompCert also allows undefined behaviors to be refined by anything, which makes the refinement relation slightly more sophisticated. See Chapter 8 for details.

### 2.1.3 Compositionality

Complex systems are built by assembling components whose behavior is understood, such that their interaction achieves a desired effect. The syntactic constructions of the language used to describe systems correspond to the ways in which they can be composed.

To enable compositional reasoning, a suitable model must provide an account of the behavior of a system in terms of the behavior of its parts. For instance, if the language contains a binary operator  $+$  :  $P \times P \rightarrow P$ , then the semantic domain should be equipped with a corresponding operation  $\oplus$  :  $\mathbb{D} \times \mathbb{D} \rightarrow \mathbb{D}$ . Ideally, the operation  $\oplus$  will characterize  $+$  exactly:

$$\llbracket p_1 \rrbracket \oplus \llbracket p_2 \rrbracket = \llbracket p_1 + p_2 \rrbracket. \quad (2.1)$$

**Example 2.3.** *Denotational semantics are compositional by nature. The behavior of program components is defined recursively on their structure, so that (2.1) holds by definition.*

In the context of refinement-based verification, we can regard  $\oplus$  as a *specification* for  $+$ , and it is sufficient to establish that:

$$\llbracket p_1 \rrbracket \oplus \llbracket p_2 \rrbracket \sqsubseteq \llbracket p_1 + p_2 \rrbracket. \quad (2.2)$$

**Example 2.4** (Linking). *CompCert approximates linking as an operator  $+$  which merges programs. In CompCertO, the semantic model is equipped with a horizontal composition operation  $\oplus$ , which provides a specification for the linking operator. In particular, the correctness of assembly linking is established in Thm. 9.9 as  $\text{Asm}(p_1) \oplus \text{Asm}(p_2) \sqsubseteq \text{Asm}(p_1 + p_2)$ .*

### 2.1.4 Monotonicity

Once a component has been shown to conform to a given specification, we want to abstract it as a “black box” so that further reasoning can be done in terms of the component’s specification rather than its implementation details. To support this, we must establish that semantic composition operators are compatible with refinement:

$$\frac{\sigma_1 \sqsubseteq \sigma'_1 \quad \sigma_2 \sqsubseteq \sigma'_2}{\sigma_1 \oplus \sigma_2 \sqsubseteq \sigma'_1 \oplus \sigma'_2}$$

For example, suppose we have two components  $p_1$  and  $p_2$ , where  $p_2$  relies on  $p_1$  for its operation, and we want to verify that their combination  $p_1 + p_2$  satisfies a specification  $\sigma$ . Once we verify  $p_1$  against its own specification  $\sigma_1 \sqsubseteq \llbracket p_1 \rrbracket$ , by the monotonicity of  $\oplus$  it is sufficient to show that  $\sigma \sqsubseteq \sigma_1 \oplus \llbracket p_2 \rrbracket$ :

$$\sigma \sqsubseteq \sigma_1 \oplus \llbracket p_2 \rrbracket \sqsubseteq \llbracket p_1 \rrbracket \oplus \llbracket p_2 \rrbracket \sqsubseteq \llbracket p_1 + p_2 \rrbracket.$$

### 2.1.5 Types

In many cases, systems are built out of components of various types which can only be composed when their types are compatible. Different semantic domains can then be used to interpret components of different types.

As discussed in Chapter 3, categories offer a systematic framework to realize this separation. Instead of a single set  $\mathbb{D}$ , we use a category  $\mathbf{D}$ . The objects  $A \in \mathbf{D}$  correspond to the possible types or interfaces alongside which components can be connected. The behavior of components which use an interface  $A$  to implement an interface  $B$  can then be interpreted as morphisms of type  $A \rightarrow B$ . Category theory can then be a useful resource to formulate and characterize typed composition principles, and the theory of *enriched categories* can be used to explore how these composition principles interact with the structure of the semantic domains  $\mathbf{D}(A, B)$ .

**Example 2.5** (Signatures). *The models introduced in this thesis use signatures, discussed in §2.5, to enumerate the operations used and provided by program components. Consequently, the semantics of components are often expressed in categories whose objects are signatures. Morphisms of type  $E \rightarrow F$  describe the behavior of components which use the operations of  $E$  to implement the operations of  $F$ . Categorical structures can then be used to describe various composition principles.*

### 2.1.6 Abstraction

Large-scale systems operate across multiple levels of abstraction. Each abstraction level brings to its own understanding of the interaction between a component and its environment. To reason across abstraction layers we need to give an explicit account of how these different views of a component's interactions correspond to one another.

This is studied extensively in the context of abstract interpretation [Cousot and Cousot, 1992].

Suppose  $\mathbb{D}_1$  is the concrete domain and  $\mathbb{D}_2$  is the abstract one. To establish a correspondence between them, the most general approach is to define a *soundness relation*  $\rho \subseteq \mathbb{D}_2 \times \mathbb{D}_1$ . We expect  $\rho$  to be compatible with refinement:

$$\frac{\sigma'_2 \sqsubseteq_2 \sigma_2 \quad \sigma_2 \rho \sigma_1 \quad \sigma_1 \sqsubseteq_1 \sigma'_1}{\sigma'_2 \rho \sigma'_1}$$

In other words, if  $\sigma_2$  is an abstraction which soundly approximates a concrete semantic object  $\sigma_1$ , we can make  $\sigma_2$  less precise or  $\sigma_1$  more precise and preserve this relationship.

Often there will be a more elementary description of the correspondence, based on the construction of the semantic domains, as illustrated by the following example.

**Example 2.6.** *In the model of certified abstraction layers presented in Chapter 5, the semantic domain used to represent layer interfaces is parametrized by a set of abstract states  $S$ . To compare a layer interface formulated in terms of a more concrete set of states  $S_1$  with a layer interface formulated in terms of a more abstract set of states  $S_2$ , we will use a relation  $R \subseteq S_2 \times S_1$ . This relation can then be extended to the level of layer interfaces as a soundness relation  $\leq_R \subseteq \mathbb{D}[S_2] \times \mathbb{D}[S_1]$ , asserting that  $R$  is a simulation relation between a layer interface in  $\mathbb{D}[S_2]$  and a layer interface in  $\mathbb{D}[S_1]$ .*

In the case of categorical semantics where  $\mathbb{D}_1 = \mathbf{D}(A_1, B_1)$  and  $\mathbb{D}_2 = \mathbf{D}(A_2, B_2)$  are homsets, the constituents of the soundness relation may themselves be organized as a category with the same objects as  $\mathbf{D}$ . This induces a double category structure, where horizontal morphisms are the semantic objects, vertical morphisms are abstraction relationships, and 2-cells correspond to the soundness relation.

**Example 2.7** (Simulation conventions in CompCertO). *Interactions between C compilation units are understood very differently from interactions between assembly components. At the level of C, cross-module interaction is defined in terms of function calls; invoking a function means assigning values to the function's parameters, initializing a new stack frame, and finally executing the function's body. At the assembly level, cross-module interactions simply consist in branching to an address outside of the current module with a certain register state. The calling convention used by a compiler specifies the correspondence between the two.*

*In the model used in CompCertO, language interfaces describe the form of cross-component inter-*

actions for a class of languages. They play the role of objects, and semantic domains are assigned a type  $A \rightarrow B$  where  $A$  and  $B$  are language interfaces. In addition, a simulation convention  $\mathbb{R} : A_1 \Leftrightarrow A_2$  is used to formulate the correspondence between the high-level language interface  $A_1$  and the low-level language interface  $A_2$ . A simulation property can then be depicted as the 2-cell:

$$\begin{array}{ccc} A_1 & \xrightarrow{L_1} & B_1 \\ \mathbb{R}_A \Uparrow & & \Uparrow \mathbb{R}_B \\ A_2 & \xrightarrow{L_2} & B_2 \end{array}$$

When semantic domains are rich enough, there may be a most precise abstraction  $\sigma_2 \in \mathbb{D}_1$  for each concrete semantic object  $\sigma_1 \in \mathbb{D}_1$ , defining an abstraction function  $\alpha : \mathbb{D}_1 \rightarrow \mathbb{D}_2$  such that:

$$\sigma_2 \rho \sigma_1 \Leftrightarrow \sigma_2 \sqsubseteq_2 \alpha(\sigma_1)$$

Conversely, there may be a least precise  $\sigma_1$  corresponding to each  $\sigma_2$ , defining a concretization function  $\gamma : \mathbb{D}_2 \rightarrow \mathbb{D}_1$  such that:

$$\gamma(\sigma_2) \sqsubseteq_1 \sigma_1 \Leftrightarrow \sigma_2 \rho \sigma_1$$

When both an abstraction and concretization function exist, they are related by the property:

$$\gamma(\sigma_2) \sqsubseteq_1 \sigma_1 \Leftrightarrow \sigma_2 \sqsubseteq_2 \alpha(\sigma_1)$$

This corresponds to the original formulation of abstract interpretation [Cousot and Cousot, 1977] in terms of Galois connections [Erné et al., 1993].

### 2.1.7 Embedding

It is often useful to first interpret the semantics of a component in a more restricted domain, then embed this domain into a more general one:

$$p \in P \xrightarrow{\llbracket - \rrbracket} \sigma \in \mathbb{D} \xrightarrow{|\cdot|} |\sigma| \in \mathbb{D}'$$



The elements of the restricted domain  $\mathbb{D}$  will often have stronger properties, or be expressed in a way which makes certain proof methods more tractable. The more general domain  $\mathbb{D}'$  can then provide more expressivity and additional algebraic structures, and may embed the behaviors of different kinds of components, but may be less amenable to domain-specific reasoning.

In this situation, we will want the embedding  $|-| : \mathbb{D} \hookrightarrow \mathbb{D}'$  to preserve any relevant structure present in both  $\mathbb{D}$  and  $\mathbb{D}'$ . For example, in the case of categorical semantics, this embedding will be a (faithful) functor of the appropriate kind.

**Example 2.8** (Transition systems and trace semantics in CompCert). *As mentioned in Example 2.1, CompCert semantics ultimately characterize the behavior of programs in terms of traces. However, they are first defined as labelled transition systems, and the correctness properties of compilation passes are first proved as simulations. An embedding then assigns to each transition system  $L$  the set of traces  $|L|$  characterizing its observable behavior, and a soundness proof shows that simulation properties  $L_1 \leq L_2$  imply trace containment  $|L_1| \supseteq |L_2|$ .*

**Example 2.9** (Certified abstraction layers). *The theory of certified abstraction layers developed in Chapter 5 describes deterministic layer interfaces and implementations. Abstraction is expressed using simulation relations. Layer interfaces, layer implementations and simulation relations are represented as three different kinds of objects.*

*In Chapter 6, I show how this theory of certified abstraction layers can be embedded into a more general framework which supports dual nondeterminism, allowing layer interfaces, layer implementations and simulation relations are all represented as morphisms in a uniform way, at the cost of a more sophisticated the notion of refinement.*

*Finally in Chapter 7, I outline a framework where the layers' states can be encapsulated and hidden from the descriptions of the layers' behaviors, and again provide an embedding of the previous theory into this more general one. This means simulation relations are no longer necessary at all, but prevents us from expressing the cartesian products of layers described in §5.3.2.*

*Each of these embeddings preserves categorical composition, refinement, and tensor products.*

## 2.2 The refinement calculus

Correctness properties for imperative programs are often stated as triples of the form  $P\{C\}Q$  asserting that when the program  $C$  is started in a state which satisfies the predicate  $P$  (the *precondition*), then the state in which  $C$  terminates will satisfy the predicate  $Q$  (the *postcondition*). For example:

$$x \text{ is odd } \{x := x * 2\} x \text{ is even}$$

In the *axiomatic* approach to programming language semantics [Hoare, 1969], inference rules corresponding to the different constructions of the language determine which triples are valid, and the meaning of a program is identified with the set of properties  $P\{-\}Q$  which the program satisfies.

### 2.2.1 Dual nondeterminism

Axiomatic semantics can accommodate nondeterminism in two different ways.

In the program  $C_1 \sqcap C_2$ , a *demon* will choose which of  $C_1$  or  $C_2$  is executed. For example, the program  $x := 2 * x \sqcap x := 0$  may be executed arbitrarily as  $x := 2 * x$  or  $x := 0$ , with no guarantee as to which branch will be chosen. The demon works against us, so that if we want  $C_1 \sqcap C_2$  to satisfy a given correctness property, we need to make sure we can deal with either choice:

$$\frac{P\{C_1\}Q \quad P\{C_2\}Q}{P\{C_1 \sqcap C_2\}Q}$$

Conversely, in the program  $C_1 \sqcup C_2$ , an *angel* will decide whether  $C_1$  or  $C_2$  is executed. The statement  $x := x * 2 \sqcup x := 0$  is more difficult to interpret than its demonic counterpart, but roughly speaking magically behaves as  $x := x * 2$  or  $x := 0$  depending on the needs of its user. If possible, the angel will make choices which validate the correctness property. This implies:

$$\frac{P\{C_1\}Q}{P\{C_1 \sqcup C_2\}Q} \quad \frac{P\{C_2\}Q}{P\{C_1 \sqcup C_2\}Q}$$

More generally, a triple  $P\{C\}Q$  can be interpreted as a *game* between the angel and the demon [Back and Wright, 1998, Chapter 14]. The angel resolves the  $\sqcup$  choices, whereas the demon resolves  $\sqcap$ . The triple is valid if there is a winning strategy for the angel.

### 2.2.2 Distributivity

Note that in this setup,  $\sqcap$  and  $\sqcup$  distribute over each other. More precisely, for all  $P, C_1, C_2, C_3, Q$ :

$$\begin{aligned} P\{C_1 \sqcap (C_2 \sqcup C_3)\}Q &\Leftrightarrow P\{(C_1 \sqcap C_2) \sqcup (C_1 \sqcap C_3)\}Q \\ P\{C_1 \sqcup (C_2 \sqcap C_3)\}Q &\Leftrightarrow P\{(C_1 \sqcup C_2) \sqcap (C_1 \sqcup C_3)\}Q \end{aligned}$$

Consider the first equivalence above. If the angel has a winning strategy for the left-hand side triple, they can win both  $P\{C_1\}Q$  and either  $P\{C_2\}Q$  or  $P\{C_3\}Q$ . Although the right-hand side triple reverses the angel and demon's choices, the angel can preemptively choose the left or right disjunct depending on whether they can win  $P\{C_2\}Q$  or  $P\{C_3\}Q$ . Likewise, if the angel can win the right-hand side, then they have a winning strategy for the left-hand side as well. The second equivalence corresponds to a similar situation where the angel and demon have been exchanged.

### 2.2.3 Program refinement

Instead of proving program correctness in one go, *stepwise refinement* techniques use a more incremental approach centered on the notion of program refinement. A refinement  $C_1 \sqsubseteq C_2$  means that any correctness property satisfied by  $C_1$  will also be satisfied by  $C_2$ :

$$C_1 \sqsubseteq C_2 := \forall PQ. P\{C_1\}Q \Rightarrow P\{C_2\}Q$$

We say that  $C_2$  refines  $C_1$  or that  $C_1$  is refined by  $C_2$ .

Typically, under such approaches, the language will be extended with constructions allowing the user to describe abstract specifications as well as concrete programs. Then the goal is to establish a sequence of refinements  $C_1 \sqsubseteq \dots \sqsubseteq C_n$  to show that a program  $C_n$  involving only executable constructions correctly implements a specification  $C_1$ , which may be stated in much more abstract terms.

If the language is sufficiently expressive, then a correctness property  $P\{-\}Q$  can itself be encoded [Morgan, 1988] as a specification statement  $\langle P, Q \rangle$  such that:

$$P\{C\}Q \Leftrightarrow \langle P, Q \rangle \sqsubseteq C.$$

In the context of refinement, the properties associated with demonic and angelic choice become:

$$\frac{C \sqsubseteq C_1 \quad C \sqsubseteq C_2}{C \sqsubseteq C_1 \sqcap C_2} \quad \frac{C \sqsubseteq C_1}{C \sqsubseteq C_1 \sqcup C_2} \quad \frac{C \sqsubseteq C_2}{C \sqsubseteq C_1 \sqcup C_2}$$

Given the symmetry between the demon and angel, it is then natural to interpret demonic and angelic choices respectively as meets and joins of the refinement ordering.

#### 2.2.4 Nondeterministic choice in specifications

Until this point, we have discussed demonic ( $\sqcap$ ) and angelic ( $\sqcup$ ) choices as implementation constructs (appearing to the right of  $\sqsubseteq$ ), taking the point of view of a client seeking to use the program to achieve a certain goal. However, in this work they are used primarily as *specification* constructs (appearing to the left of  $\sqsubseteq$ ), and we are interested in what it means for a system to implement them. As a specification,  $C_1 \sqcap C_2$  allows the system to refine either one of  $C_1$  or  $C_2$ , while  $C_1 \sqcup C_2$  requires it to refine *both* of them:

$$\frac{C_1 \sqsubseteq C}{C_1 \sqcap C_2 \sqsubseteq C} \quad \frac{C_2 \sqsubseteq C}{C_1 \sqcap C_2 \sqsubseteq C} \quad \frac{C_1 \sqsubseteq C \quad C_2 \sqsubseteq C}{C_1 \sqcup C_2 \sqsubseteq C}$$

In other words, demonic choices give us more implementation freedom, whereas angelic choices make a specification stronger and more difficult to implement. Therefore we can think of demonic choices as choices of the *system*, and think of angelic choices as choices of the *environment*.

#### 2.2.5 The refinement calculus

These basic ingredients have been studied systematically in the *refinement calculus*, dating back to Ralph-Johan Back's 1978 PhD thesis [Back, 1978]. In its modern incarnation, the refinement calculus subsumes programs and specifications with *contracts* featuring unbounded angelic and demonic choices [Back and Wright, 1998]. These choice operators constitute a *completely distributive lattice* with respect to the refinement ordering.

Dijkstra's *predicate transformer* semantics [Dijkstra, 1975] are a natural fit for the refinement calculus, but other approaches are possible. For instance, as mentioned above, the understanding of contracts as a game between the angel and the demon can be formalized to provide a form of

game semantics for the refinement calculus.

More generally, The refinement calculus can be presented as a *hierarchy* where simpler models (state transformer functions, relations) can be embedded in more general ones (predicate transformers) in various structure-preserving ways. This makes it possible to reason about simpler components in a limited, stronger version of the framework, while retaining the possibility of embedding them in a setting where more general constructions are available, and where they can be composed with components developed and analyzed in a different setting. This expressivity is highly desirable as a scalable approach to verifying heterogeneous systems.

However, the refinement calculus only applies to imperative programs with no side-effects beyond changes to the program state. Recent research has attempted to extend the paradigm to a broader setting, and the present work can be understood as a step in this direction as well.

## 2.3 Game semantics

Game semantics [Abramsky and McCusker, 1999; Blass, 1992] is a form of denotational semantics which incorporates some operational aspects. An early success of this approach was the formulation of the first fully abstract models of the programming language PCF [Abramsky et al., 2000; Hyland and Ong, 2000]. Typically, game semantics interpret types as two-player games and terms as strategies for these games. Games describe the form of the interaction between a program component (the *system*) and its execution context (the *environment*). Strategies specify which move the system plays for all relevant positions in the game.

Positions are usually identified with sequences of moves, and strategies with the set of positions a component can reach. This representation makes game semantics similar to trace semantics of process algebras, but it is distinguished by a strong polarization between actions of the system and the environment, and between outputs and inputs. This confers an inherent “rely-guarantee” flavor to games which facilitates compositional reasoning [Abramsky, 2010].

For example, in a simple game semantics resembling that of Idealized Algol [Abramsky and McCusker, 1997], sequences of moves corresponding to the execution of  $x := 2 * x$  have the form:

$$\text{run} \cdot \underline{\text{read}}_x \cdot n \cdot \underline{\text{write}}_x[2n] \cdot \text{ok} \cdot \underline{\text{done}} \quad (n \in \mathbb{N})$$

The moves of the system have been underlined. The environment initiates the execution with the move `run`. The system move `readx` requests the value of the variable `x`, communicated in response by the environment move `n`. The system move `writex[2n]` requests storing the value `2n` into the variable `x`, and is acknowledged by the environment move `ok`. Finally, the system move `done` expresses termination.

### 2.3.1 Games

A game is defined by a set of moves players will choose from, as well as a stipulation of which sequences of moves are valid. We focus on two-player, alternating games where the environment plays first and where the players each contribute every other move. As above, when typesetting examples, we underline the moves of the system.

**Example 2.10.** *In the game of chess, moves are taken in the set  $\{a1 \dots h8\} \times \{a1 \dots h8\}$ . From the perspective of the player with black pieces, a valid sequence of moves may look like:*

$$e2e4 \cdot \underline{c7c5} \cdot c2c3 \cdot \underline{d7d5} \dots$$

Most game semantics include additional structure in the description of games. The set of moves is usually partitioned into system and environment moves ( $M = M^O \uplus M^P$ ), and into questions and answers ( $M = M^\circ \uplus M^\bullet$ ). Game models for high-order languages are often more complex, and include *justification pointers* encoding the causal structure of the interaction.

The compositionality of game semantics comes from the ways in which complex games can be derived from simple ones, and used to interpret compound types. For example, in the game  $A \times B$  the environment initially chooses whether to play an instance of  $A$  or an instance of  $B$ . The game  $A \rightarrow B$  usually consists of an instance of  $B$  played together with instances of  $A$  started at the discretion of the system, where the roles of the players are reversed.

### 2.3.2 Strategies

The *plays* of a game are sequences of moves; they both identify a position in the game and describe the succession of actions that led to it. Most game models of sequential computation use *alternating* plays, in which the system and environment each contribute every other move. It is also common

to require the environment to play first and to restrict plays to even lengths, so that they specify which action the system took in response to the latest environment move. We write  $P_G$  for the set of plays of the game  $G$ , partially ordered by the prefix relation  $\sqsubseteq_p$ .

Traditionally [Abramsky and McCusker, 1999], strategies are defined as prefix-closed sets of plays, so that strategies  $\sigma \in S_G$  for the game  $G$  are downsets of  $P_G$  satisfying certain requirements:

$$S_G \subseteq \mathcal{D}(P_G, \sqsubseteq_p)$$

A common constraint is that a strategy  $\sigma \in S_G$  should not contain two plays  $sm_1, sm_2 \in \sigma$  where  $m_1$  and  $m_2$  are distinct moves of the system.

This is usually understood as enforcing *system determinism*: given a set of environment choices, there is only one possible behavior for the system. Therefore, relaxing this constraint has usually been understood as the first step toward modeling nondeterministic systems [Harmer and McCusker, 1999]. We will see in Chapter 4 that by approaching the question from the point of view of dual nondeterminism, we are led to a different interpretation and a different approach.

## 2.4 Logical relations

Logical relations are structure-preserving relations in the way homomorphisms are structure-preserving maps. However, logical relations are more compositional than homomorphisms, because they do not suffer from the same problems in the presence of mixed-variance constructions like the function arrow [Hermida et al., 2014]. In the context of typed languages, this means that type-indexed logical relations can be defined by recursion over the structure of types.

Logical relations have found widespread use in programming language theory. Unary logical relations can be used to establish various properties of type systems: a type-indexed predicate expressing a property of interest is shown to be compatible with the language's reduction, and to contain all of the well-typed terms of the language. Binary logical relations can be used to capture contextual equivalence between terms, as well as notions such as non-interference or compiler correctness. Relational models of type quantification yield Reynold's well-known theory of relational parametricity, and can be used to prove *free theorems* that all terms of a given generic

type must satisfy.

### 2.4.1 Binary logical relations

Logical relations can be of any arity, but we restrict our attention to binary logical relations. Given an algebraic structure  $\mathcal{S}$ , a *logical relation* between two instances  $S_1, S_2$  of  $\mathcal{S}$  is a relation  $R$  between their carrier sets, such that the corresponding operations of  $S_1$  and  $S_2$  take related arguments to related results. We write  $R \in \mathcal{R}(S_1, S_2)$ .

**Example 2.11.** A monoid is a set with an associative operation  $\cdot$  and an identity element  $\epsilon$ . A logical relation of monoids between  $\langle A, \cdot_A, \epsilon_A \rangle$  and  $\langle B, \cdot_B, \epsilon_B \rangle$  is a relation  $R \subseteq A \times B$  such that:

$$(u R u' \wedge v R v' \Rightarrow u \cdot_A v R u' \cdot_B v') \wedge \epsilon_A R \epsilon_B. \quad (2.3)$$

### 2.4.2 Relators

Logical relations between multisorted structures consist of one relation for each sort, between the corresponding carrier sets. In the case of structures which include type operators, we can associate to each base type  $A$  a relation over its carrier set  $\llbracket A \rrbracket$ , and to each type operator  $T(A_1, \dots, A_n)$  a corresponding *relator*: given relations  $R_1, \dots, R_n$  over the carrier sets  $\llbracket A_1 \rrbracket, \dots, \llbracket A_n \rrbracket$ , the relator for  $T$  will construct a relation  $T(R_1, \dots, R_n)$  over  $\llbracket T(A_1, \dots, A_n) \rrbracket$ . Relators for some common constructions are shown in Fig. 2.1. Using them, the proposition (2.3) can be reformulated as:

$$\cdot_A [R \times R \rightarrow R] \cdot_B \wedge \epsilon_A R \epsilon_B.$$

**Example 2.12.** Simulation relations are logical relations of transition systems. Consider the transition systems  $\alpha : A \rightarrow \mathcal{P}(A)$  and  $\beta : B \rightarrow \mathcal{P}(B)$ . A simulation relation  $R \in \mathcal{R}(A, B)$  satisfies:

$$\begin{array}{ccc} s_1 & \xrightarrow{\alpha} & s'_1 \\ R \downarrow & & \downarrow R \\ s_2 & \xrightarrow{\beta} & s'_2 \end{array} \quad \begin{array}{l} \forall s_1 s_2 s'_1. \alpha(s_1) \ni s'_1 \wedge s_1 R s_2 \Rightarrow \\ \exists s'_2. \beta(s_2) \ni s'_2 \wedge s'_1 R s'_2 \end{array}$$



$$\begin{aligned}
x [R_1 \times R_2] y &\Leftrightarrow \pi_1(x) [R_1] \pi_1(y) \wedge \pi_2(x) [R_2] \pi_2(y) \\
x [R_1 + R_2] y &\Leftrightarrow (\exists x_1 y_1 . x_1 [R_1] y_1 \wedge x = i_1(x_1) \wedge y = i_1(y_1)) \\
&\quad \vee (\exists x_2 y_2 . x_2 [R_2] y_2 \wedge x = i_2(x_2) \wedge y = i_2(y_2)) \\
f [R_1 \rightarrow R_2] g &\Leftrightarrow \forall x y . x [R_1] y \Rightarrow f(x) [R_2] g(y) \\
A [\mathcal{P}^\leq(R)] B &\Leftrightarrow \forall x \in A . \exists y \in B . x [R] y
\end{aligned}$$

Figure 2.1: A selection of relators. The relators  $\times$ ,  $+$ ,  $\rightarrow$  are standard.  $\mathcal{P}^\leq$  is an asymmetric version of the relator for the powerset type operator  $\mathcal{P}$ , which can be used to formulate simulations.

Using the relators in Fig. 2.1, we can express the same property concisely and compositionally as:

$$\alpha [R \rightarrow \mathcal{P}^\leq(R)] \beta.$$

### 2.4.3 Kripke relations

Since relations for stateful languages often depend on the current state, Kripke logical relations are parametrized over a set of state-dependent *worlds*. Components related at the same world are guaranteed to be related in compatible ways. We use the following notations.

**Definition 2.13.** A *Kripke relation* is a family of relations  $(R_w)_{w \in W}$ . We write  $R \in \mathcal{R}_W(A, B)$  for a Kripke relation between the sets  $A$  and  $B$ . For  $w \in W$  I will write:

$$[w \Vdash R] := R_w \quad [\Vdash R] := \bigcap_w R_w$$

A simple relation  $R \in \mathcal{R}(A, B)$  can be promoted to a Kripke relation  $[R] \in \mathcal{R}_W(A, B)$  by defining  $[w \Vdash [R]] := R$  for all  $w \in W$ . More generally, for an  $n$ -ary relator  $F$  we have:

$$\frac{F : \mathcal{R}(A_1, B_1) \times \cdots \times \mathcal{R}(A_n, B_n) \rightarrow \mathcal{R}(A, B)}{[F] : \mathcal{R}_W(A_1, B_1) \times \cdots \times \mathcal{R}_W(A_n, B_n) \rightarrow \mathcal{R}_W(A, B)}$$

where for the Kripke relations  $R_i \in \mathcal{R}_W(A_i, B_i)$ :

$$[w \Vdash [F](R_1, \dots, R_n)] := F(w \Vdash R_1, \dots, w \Vdash R_n)$$

In the following, I will use  $\lceil - \rceil$  implicitly when a relator appears in a context where a Kripke logical relation is expected. Since reasoning with logical relations often involves self-relatedness, I use the notation  $x :: R$  to denote  $x R x$ . For legibility, I will also write  $w \Vdash x R y$  for  $x \lceil w \Vdash R \rceil y$  and  $\Vdash x R y$  for  $x \lceil \Vdash R \rceil y$ .

## 2.5 Algebraic effects

The framework of *algebraic effects* [Plotkin and Power, 2001] models computations as terms in an algebra whose operations represent effects: a term  $m(x_1, \dots, x_n)$  represents a computation which first triggers an effect  $m$ , then continues as a computation derived from the subcomputations  $x_1, \dots, x_n$ . For example, the term:

$$w := \text{readbit}(\text{print}[\text{“Hello”}](\text{done}), \text{print}[\text{“World”}](\text{done}))$$

could denote a computation which first reads one bit of information, then depending on the result causes the words “Hello” or “World” to be output, and finally terminates.

Note that somewhat surprisingly, the *arguments* of operations correspond to the possible *outcomes* of the associated effect. For instance the `readbit` operation takes two arguments. Moreover, effects such as `print` which take parameters are represented by *families* of operations indexed by the parameters’ values, so that there is a `print[s]` operation for every  $s \in \text{string}$ .

### 2.5.1 Effects theories

Under this approach, effects can be described as algebraic theories: a signature describes the set of operations together with their arities, and a set of equations describes their behaviors by specifying which computations are equivalent. The example above uses a signature with the operations `done` of arity 0, `readbit` of arity 2, and a family of operations  $(\text{print}[s])_{s \in \text{string}}$  of arity 1. An equation for this signature is:

$$\text{print}[s](\text{print}[t](x)) = \text{print}[st](x),$$

which indicates that printing the string  $s$  followed by printing the string  $t$  is equivalent to printing the string  $st$  in one go.

In this work, I use effect signatures to represent the possible external interactions of a computation, but I will not use equational theories. It will however be possible to interpret effects into another signature, modeling a limited form of *effect handlers* [Plotkin and Pretnar, 2009].

### 2.5.2 Effect signatures

**Definition 2.14.** An *effect signature* is a set  $E$  of operations together with a mapping  $\text{ar}$ , which assigns to each  $e \in E$  a set  $\text{ar}(e)$  called the *arity* of  $e$ . We will use the notation  $E = \{e_1 : \text{ar}(e_1), e_2 : \text{ar}(e_2), \dots\}$  to describe effect signatures.

Note that in this definition, arities are *sets* rather than natural numbers. This allows the representation of effects with a potentially infinite number of outcomes. The examples above use effects from the following signature:

$$E_{\text{io}} := \{\text{readbit} : \mathbb{2}, \text{print}[s] : \mathbb{1}, \text{done} : \emptyset \mid s \in \text{string}\},$$

where  $\mathbb{1} = \{*\}$  and  $\mathbb{2} := \{\text{tt}, \text{ff}\}$  are finite sets of the expected size.

Since we will rely on the construction  $\{m[x] : B \mid x \in A\}$  extensively, I will use the syntactic sugar  $\{m : A \rightarrow B\}$  so that the signature above can be described as:

$$E_{\text{io}} = \{\text{readbit} : \mathbb{2}, \text{print} : \text{string} \rightarrow \mathbb{1}, \text{done} : \emptyset\}$$

### 2.5.3 Computations as terms

The most direct way to interpret an effect signature is the algebraic point of view, in which it induces a set of terms built out of the signature's operations. A term represents a computation which proceeds inward from the top-level operation towards the leaves of the term, namely the constants  $(c : \emptyset) \in E$ . Terms representing partial computations may also involve variables. In that case, the variables may be thought of as placeholders, each one representing a possible intermediate outcome, which may later be replaced by another computation to be sequenced with the initial one. Terms are defined as follows.

**Definition 2.15** (Terms over a signature). The set of *terms* associated with a signature  $E$  and a

set of variables  $X$  is defined as:

$$t \in T_E(X) ::= \underline{m}(t_n)_{n \in N} \mid \underline{v} \quad \text{where } (m : N) \in E, v \in X$$

Consider an operation  $m : N$  in the signature  $E$ . A term of the form  $\underline{m}(t_i)_{i \in N}$  first triggers the effect  $m$ . The effect's outcome  $n \in N$  then resumes the computation as prescribed by the subterm  $t_n \in T_E(X)$ . On the other hand, terms of the form  $\underline{v} \in X$  correspond to a computation which immediately terminates with the corresponding outcome. A substitution  $\sigma : X \rightarrow T_E(Y)$  can then specify how the computation is to be continued.

**Definition 2.16** (Variable substitution). A substitution  $f : X \rightarrow T_E(Y)$  can be *applied* to a term  $t \in T_E(X)$ , yielding the term  $tf \in T_E(Y)$  defined by:

$$\begin{aligned} \underline{v}f &:= f(v) & v &\in X \\ \underline{m}(t_n)_{n \in N}f &:= \underline{m}(t_nf)_{n \in N} & (m : N) &\in E \end{aligned}$$

**Example 2.17.** The computation  $w$  given above can be decomposed into the term  $t \in T_{E_{\text{io}}}(\{x, y\})$  and the substitution  $f : \{x, y\} \rightarrow \emptyset$  defined as:

$$\begin{aligned} t &:= \text{readbit}(x, y) \\ f &:= \{x \mapsto \text{print}[\text{"Hello"}](\text{done}), y \mapsto \text{print}[\text{"World"}](\text{done})\}. \end{aligned}$$

Then the term  $tf \in T_{E_{\text{io}}}(\emptyset)$  obtained applying  $f$  to  $t$  is again:

$$tf = \text{readbit}(\text{print}[\text{"Hello"}](\text{done}), \text{print}[\text{"World"}](\text{done})) = w.$$

## 2.5.4 Interpreting effects

To assign a semantics to the effects of  $E$ , we can interpret the operations of the signature in a semantic domain  $A$  by defining for each  $(m : N) \in E$  a corresponding function  $\alpha_m : A^N \rightarrow A$ . When equational theories are used, we need to make sure that the corresponding equations hold. In our limited setting we can use the simple definition below.

**Definition 2.18.** An *algebra* for the effect signature  $E$  is a *carrier set*  $A$  together with a function  $\alpha_m : A^N \rightarrow A$  for each operation  $(m : N) \in E$ . We can then interpret a term  $t \in T_E(A)$  as an element  $t^\alpha \in A$  of the carrier set, defined recursively by:

$$\begin{aligned} \underline{v}^\alpha &:= v & v &\in A \\ \underline{m}(t_n)_{n \in N}^\alpha &:= \alpha_m(t_n^\alpha)_{n \in N} & (m : N) &\in E \end{aligned}$$

**Example 2.19** (Trace semantics for  $E_{\text{io}}$ ). We can interpret the operations of  $E_{\text{io}}$  using traces in the language  $P = (\mathbb{2} + \text{str})^*$ . A trace  $s \in P$  records a possible execution, and sets of traces are used to characterize computations. Hence the algebra  $\pi$  uses the carrier  $\mathcal{P}(P)$  and operations:

$$\begin{aligned} \pi_{\text{readbit}}(\sigma_0, \sigma_1) &:= \{0 \cdot s_0, 1 \cdot s_1 \mid s_0 \in \sigma_0, s_1 \in \sigma_1\} \\ \pi_{\text{print}[u]}(\sigma) &:= \{u \cdot s \mid s \in \sigma\} \\ \pi_{\text{done}} &:= \{\epsilon\} \end{aligned}$$

For instance, the recurring example  $w$  yields the set:

$$w^\pi = \{0 \cdot \text{“Hello”}, 1 \cdot \text{“Worlds”}\}.$$

Note that as presented, the algebra does not respect the equation:

$$\text{print}[s](\text{print}[t](x)) = \text{print}[st](x),$$

since the left-hand side will add two events  $s \cdot t$  at the beginning of the traces, while the right-hand side will add a single event  $st$ .

### 2.5.5 Monads

The semantics of effectful computations is often expressed using *monads*. There are deep connections between monadic and algebraic effects, which mirror the way monads have long been used in the context of categorical algebra. This is explained in more detail §3.4.1. Below, I give a brief introduction free of categorical jargon.

The notion of algebra given in Def. 2.18 does not provide a very sophisticated account of terms with variables (partial computations). We did allow “variables” to be taken from the algebra’s underlying set  $A$  when interpreting terms. It is possible to interpret more general terms with variables by assigning values in  $A$  to each one: given a term  $t \in T_E(X)$ , and a variable assignment  $\rho : X \rightarrow A$ , we can easily construct the substitution  $f(x) := \rho(x)$  and interpret  $t$  as  $(tf)^\alpha \in A$ .

However, to be able to interpret open terms independently of any assignment, we need a notion of algebra which is *parametric* in the set of variables we use. Instead of using a single carrier set, we will specify for every possible set  $X$  of variables:

- A carrier set  $T(X)$  for the corresponding algebra.
- A function  $\eta_X : X \rightarrow T(X)$  providing the interpretations of variables;
- For every “semantic substitution”  $f : X \rightarrow T(Y)$ , a map  $f^\dagger : T(X) \rightarrow T(Y)$  applying the substitution to the elements of  $T(X)$ .

These data can be interpreted computationally in the following way:

- A semantic object  $\tau \in T(X)$  is a partial computation with an intermediate result in  $X$ .
- The computation  $\eta_X(v) \in T(X)$  immediately yields the intermediate result  $v \in X$ .
- A *continuation*  $\sigma : X \rightarrow T(Y)$  is activated with a value  $v \in X$  to produce a computation with outcomes in  $Y$ . Its extension  $\sigma^\dagger(\tau)$  is the sequential composition of  $\tau \in T(X)$  with the continuation  $\sigma$ , sometimes written  $x \leftarrow \tau ; \sigma(x)$ .

When they behave consistently with each other, these data define a monad.

**Definition 2.20.** A *monad*  $\langle T, \eta, -^\dagger \rangle$  is given as above, and must satisfy for all  $f : X \rightarrow T(Y)$  and  $g : Y \rightarrow T(Z)$  the following properties:

$$\eta_Y^\dagger \circ f = f^\dagger \circ \eta_X = f \qquad (g^\dagger \circ f)^\dagger = g^\dagger \circ f^\dagger.$$

Moreover, if we specify a computation  $\sigma^m \in T(N)$  for each operation  $m : N$  in a signature  $E$ , we can derive an  $E$ -algebra  $\sigma^X$  with  $T(X)$  as its carrier set for any set  $X$  of variables. For every

operation  $(m : N) \in E$  and every family of computations  $(\tau_n)_{n \in N}$  with  $\tau_n \in T(X)$ , we define:

$$\sigma_m^X(\tau_n)_{n \in N} := n \leftarrow \sigma^m; \tau_n.$$

This algebra is compatible with the monad structure, in the sense that for all  $f : X \rightarrow T(Y)$ , the following property holds:

$$f^\dagger(\sigma_m^X(\tau_n)_{n \in N}) = \sigma_m^Y(f^\dagger(\tau_n))_{n \in N}$$

In fact, a “parametric interpretation” of  $m$  satisfying this property uniquely determines  $\sigma_m$  as:

$$\sigma_m = \sigma_m^N(\eta(n))_{n \in N}.$$

**Definition 2.21** (Free monad for a signature). The terms  $T_E(X)$  for a signature  $E$  and a set of variables  $X$  can be used to define a monad which can interpret the operations of  $E$ . For  $v \in X$ ,  $t \in T_E(X)$  and  $f : X \rightarrow T_E(Y)$ , we define:

$$\eta_X(v) := \underline{v} \qquad f^\dagger(t) := tf$$

For each operation  $m : N$  in the signature  $E$ , the corresponding operation on  $T_E$  is given by:

$$\mathbf{I}^m := \underline{m}(\underline{n})_{n \in N}.$$

This monad is known as the *free monad* for the signature  $E$ .

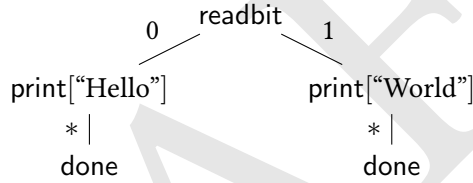
**Remark 2.22.** *This presentation of the connections between monads and algebraic structures seems to be somewhat unusual from the point of view of categorical algebra. In particular, in that context monads correspond to the algebraic theories themselves rather than their models, so that for instance the signature  $E$  would simply be captured by the monad  $T_E$ . It would be unusual to give an interpretation of a theory into a monad in the way presented above, and this would perhaps be formulated instead as a monad homomorphism out of  $T_E$ .*

*However, for my purposes it is useful to retain signatures as the starting point for constructing models, in particular in view of their interpretations as games outlined in the next section. For ex-*

ample, Chapter 6 describes a monad which adds a lattice structure to signatures in a systematic way, constructed by borrowing ideas from game semantics. The presentation above also reveals more clearly the connections between signatures and the practical use of monads in functional programming, where each monad is complemented by a number of monadic operations, given in the style that I have used.

### 2.5.6 Terms as strategies

An effect signature can be seen as a particularly simple *game*, in which the proponent chooses a question  $(m : N) \in E$  and the opponent responds with an answer  $n \in N$ . Then the terms induced by the signature are *strategies* for an iterated version of this game. Indeed, the abstract syntax tree of our example term  $w \in T_{E_{\text{io}}}(\emptyset)$  can directly be read as the strategy:



In this tree, the nodes are labeled with operations of the signatures and can be interpreted as moves of the systems. The edges are labeled with elements of the arity sets and can be interpreted as moves of the environment.

Going one step further, a form of two-sided strategy for the game  $E \rightarrow F$  can be defined as an interpretation of the signature  $F$  in the monad  $T_E$ . The strategy  $\sigma : E \rightarrow F$  responds to an environment move  $(q : R) \in F$  with a term:

$$q^\sigma \in T_E(R).$$

which is to say an interaction over the signature  $E$  which eventually produces an answer  $r \in R$  to the initial question  $(q : R) \in F$ . The identity strategy  $\text{id}_E : E \rightarrow E$  defines for each  $(m : N) \in E$ :

$$m^{\text{id}_E} := \underline{m}(\eta(n))_{n \in N}$$



The composition of  $\sigma : E \rightarrow F$  and  $\tau : F \rightarrow G$  defines for each  $(u : V) \in G$ :

$$u^{(\tau \circ \sigma)} := (u^\tau)^\sigma$$

DRAFT

## Chapter 3

# A refresher on category theory

In general, to keep my work as accessible as possible, I will avoid relying on category theory to describe the constructions and methods it uses. A reader unfamiliar with category theory should be able to skip this chapter and understand the work presented in the remainder of this thesis.

Nevertheless, category theory provides an important taxonomical framework which allows us to understand in a common language the high-level structures exhibited by various theories. Given the unifying ambition which animates the project of refinement-based game semantics, this is an invaluable resource.

Therefore, whenever possible I will mention the categorical structures underlying the constructions described in this thesis. This chapter is a brief summary of the concepts and definitions of category theory I will use for this purpose, but is not a self-contained exposition. For a more complete and careful treatment, the reader may refer to the following: a short introduction to category theory for computer scientists is given in [Pierce \[1991\]](#), a modern textbook covering the basics is provided by [Awodey \[2010\]](#), and a standard reference is [Mac Lane \[1978\]](#).

### 3.1 Motivation

Ultimately, refinement-based game semantics seeks to provide general methods for constructing heterogeneous certified systems, by integrating a vast array of semantic models and verification techniques. Category theory allows us to understand the commonalities and differences between these models in a unified and systematic way. By describing the compositional structure of a

model in categorical terms, we can step back from the details of its construction and focus instead on the abstract, high-level properties which the model exhibits. Because these properties are formulated in the universal language of categories, they can be readily compared with those of similar models, or reveal connections between seemingly unrelated phenomena across distant fields of mathematics.

In many cases, a model's high-level categorical properties will characterize it completely (up to isomorphism), without reference to the details of its construction. This provides evidence that the model is in fact the most general one exhibiting a certain structure, and demonstrate that it is free of arbitrary choices which may prove unadvisable at a later point. Then, category theory can also be used as a powerful *proof engineering* device. Formalizing category theory in a proof assistant like Coq can be useful [Spitters and van der Weegen, 2011], but it is typically a challenging undertaking involving sophisticated techniques. A more mundane approach is to simply spell out the categorical characterization of a given structure. This will give a compact specification which we can nonetheless trust to be complete, with *universal properties* providing complete but representation-independent reasoning principles for the structure of interest.

Regardless, category theory allows us to develop intuition about abstract compositional structures *in and of themselves*, independently of the context in which they may show up, making it possible to transfer intuition between a wide range of mathematical domains and apply general, abstract forms of reasoning across a variety of settings.

## 3.2 Basic definitions

### 3.2.1 Categories

**Definition 3.1.** A *category*  $\mathbf{C}$  is a collection of *objects*  $A \in \mathbf{C}$  together with a set of *morphisms*  $\mathbf{C}(A, B)$  between any two objects  $A, B \in \mathbf{C}$ . We write  $f : A \rightarrow B$  whenever  $f \in \mathbf{C}(A, B)$  is a morphism from  $A$  to  $B$ . There is an *identity* morphism  $\text{id}_A : A \rightarrow A$  for every object  $A \in \mathbf{C}$ , and a *composite* morphism  $g \circ f : A \rightarrow C$  whenever  $f : A \rightarrow B$  and  $g : B \rightarrow C$ .

Composition is associative and admits identities as units. In other words, for all morphisms

Category	Objects	Morphisms	Mon.	See
<b>Set</b>	Sets	Functions	$[\times]$	Ex. 3.2
<b>Pos</b>	Partially ordered sets	Monotonic functions	$[\times]$	
<b>Mon</b>	Monoids	Monoid homomorphisms	$\times$ $[\otimes]$	Ex. 3.3
<b>Sup</b>	Complete lattices	Preserve all sups	$\times$ $[\otimes]$	
<b>CDLat</b>	Completely distributive lattices	Complete homomorphisms	$\times$ $\otimes$	
<b>CAL</b>	Layer interfaces	Certified abstraction layers	$\otimes$	§5
$\mathcal{G}_{\sqsubseteq}^{ib}$	Effect signatures	Innocent strategies	$[\times]$	§??
$\mathcal{G}_{\sqsubseteq}^b$	Effect signatures	Reentrant strategies	$[\otimes]$	

Table 3.1: Some categories relevant to my work. XXX: Check known monoidal structures and provide references; describe the column in caption

$f : A \rightarrow B$ ,  $g : B \rightarrow C$ ,  $h : C \rightarrow D$ , the following properties hold:

$$\text{id}_B \circ f = f \circ \text{id}_A = f \quad (h \circ g) \circ f = h \circ (g \circ f).$$

Table 3.1 lists some of the categories we will use. Categories are usually named after their objects and morphisms, and assigned a short name based on their objects, as in “the category **Foo** of foos and bars”. The prototypical example is the category **Set** of sets and functions.

**Example 3.2.** *Set is the category whose objects are the sets, and morphisms are functions between two sets. The identity for  $A \in \mathbf{Set}$  is the function  $f : A \rightarrow A$  defined by  $f(a) := a$ . The composite of  $f : A \rightarrow B$  and  $g : B \rightarrow C$  is the function  $g \circ f : A \rightarrow C$  defined by  $(g \circ f)(a) := g(f(a))$ .*

Many categories use sets equipped with a given structure as objects, and structure-preserving functions as morphisms. Categories of this form are known as *concrete categories*.

**Example 3.3.** *In the category **Mon** of monoids and monoid homomorphisms:*

- *The objects are monoids, in other words tuples  $\langle A, \cdot, \epsilon \rangle$  where  $A$  is a set and where the binary operation  $\cdot : A \times A \rightarrow A$  is associative and admits  $\epsilon \in A$  as a unit.*
- *The morphisms are monoid homomorphisms. A monoid homomorphism from  $\langle A, \cdot_A, \epsilon_A \rangle$  to  $\langle B, \cdot_B, \epsilon_B \rangle$  is a function  $f : A \rightarrow B$  such that  $f(x \cdot_A y) = f(x) \cdot_B f(y)$  and  $f(\epsilon_A) = \epsilon_B$ .*

*It is easy to verify that the identity function is a monoid homomorphism and that the composition of two monoid homomorphisms is again a monoid homomorphism.*

### 3.2.2 Functors

Like monoids, categories are mathematical structures, and come with a natural notion of structure-preserving maps. These “homomorphisms of categories” are known as *functors*.

**Definition 3.4** (Functor). For two categories  $\mathbf{C}$  and  $\mathbf{D}$ , a *functor*  $F$  from  $\mathbf{C}$  to  $\mathbf{D}$  associates:

- to each object  $X \in \mathbf{C}$  an object  $FX \in \mathbf{D}$ ;
- to each morphism  $f : A \rightarrow B$  in  $\mathbf{C}$ , a morphism  $Ff : FA \rightarrow FB$  in  $\mathbf{D}$ .

Functors must preserve identity and composition, so that for  $f : A \rightarrow B$  and  $g : B \rightarrow C$  in  $\mathbf{C}$ :

$$F(\text{id}_A) = \text{id}_{FA} \quad F(g \circ f) = Fg \circ Ff.$$

I will write  $F : \mathbf{C} \rightarrow \mathbf{D}$  when  $F$  is a functor from  $\mathbf{C}$  to  $\mathbf{D}$ .

We can in fact go one step further, and define a notion of *natural transformation* between two functors of the same type.

### 3.2.3 Natural transformations

One way to think about functors is to consider the source category as a “shape”, and to think of the functor as selecting an instance of this shape in the target category. This point of view is used for example to formalize commutative diagrams as functors, and in the context of other constructions such as limits and colimits.

Under this interpretation, two functors  $F, G : \mathbf{C} \rightarrow \mathbf{D}$  give instances of the same shape  $\mathbf{C}$  in the target category  $\mathbf{D}$ . A natural transformation  $\eta : F \rightarrow G$  defines edges between corresponding target objects, connecting the two shapes as the two bases of a “prism”, and requiring the diagrams formed by each face to commute. For example, suppose the category  $\mathbf{C}$  has three objects, and the following morphisms in addition to identities:

$$\begin{array}{ccc} A & \xrightarrow{g \circ f} & C \\ & \searrow f & \nearrow g \\ & B & \end{array}$$

Category theory	Programming
Object	Type
Morphism	Function
Functor	Generic type
Natural transformation	Parametric function
Adjunction	Introduction and elimination rules

Table 3.2: Informal correspondence between category theory and programming concepts

Then a natural transformation  $\eta : F \rightarrow G$  will induce the following diagram in the category  $\mathbf{D}$ :

$$\begin{array}{ccccc}
 FA & \xrightarrow{F(g \circ f)} & FC & & \\
 \eta_A \downarrow & Ff \searrow & \nearrow Fg & \eta_B \downarrow & \\
 & FB & & & \\
 & \eta_C \downarrow & & & \\
 GA & \xrightarrow{\quad} & GC & & \\
 Gf \searrow & & \nearrow Gg & & \\
 & GB & & & 
 \end{array}$$

**Definition 3.5.** A natural transformation  $\eta : F \rightarrow G$  between the functors  $F, G : \mathbf{C} \rightarrow \mathbf{D}$ , is a family of morphisms  $\eta_X : FX \rightarrow GX$  in  $\mathbf{D}$ , indexed by  $X \in \mathbf{C}$ , such that for all morphisms  $f : X \rightarrow Y$  in  $\mathbf{C}$ , the *naturality* condition  $\eta_Y \circ Ff = Gf \circ \eta_X$  holds:

$$\begin{array}{ccc}
 FX & \xrightarrow{Ff} & FY \\
 \eta_X \downarrow & & \downarrow \eta_Y \\
 GX & \xrightarrow{Gf} & GY
 \end{array}$$

Another way to gain intuition about natural transformations, especially relevant in the context of programming languages, is to think of them as parametric functions. This relies on the correspondence summarized in Table 3.2. If we think of  $\mathbf{D}$  as a category of types and functions, then a functor  $F : \mathbf{C} \rightarrow \mathbf{D}$ , can be thought of as a *generic type* parametrized by the objects of  $\mathbf{C}$ . A natural transformation  $\eta : F \rightarrow G$  corresponds to a generic *function* of type:

$$\eta : \forall X. FX \rightarrow GX,$$

where the naturality condition enforces a form of *parametricity*.

Consider for example the case where  $\mathbf{C} = \mathbf{D} = \mathbf{Set}$ . The functor  $-^* : \mathbf{Set} \rightarrow \mathbf{Set}$  maps a set  $A$  to the corresponding set of lists  $A^*$ . Its action on morphisms corresponds to the usual map operator; in other words, for  $f : A \rightarrow B$  the function  $f^* : A^* \rightarrow B^*$  applies  $f$  independently to every element in a list; in other words:

$$f^*(\vec{a}) := \text{map}(f, \vec{a}).$$

The generic function  $\text{length} : \forall A \cdot A^* \rightarrow \mathbb{N}$  corresponds to a natural transformation  $\eta : -^* \rightarrow \mathbb{N}$ , where  $\mathbb{N}$  is understood as the constant functor mapping all objects to  $\mathbb{N}$  and all functions to  $\text{id}_{\mathbb{N}}$ . Then the naturality condition:

$$\begin{array}{ccc} A^* & \xrightarrow{f^*} & B^* \\ \eta_A \downarrow & & \downarrow \eta_B \\ \mathbb{N} & \xrightarrow{\text{id}} & \mathbb{N} \end{array}$$

shows that  $\text{length}$ , and all generic functions of the same type, satisfy the property:

$$\text{length}(\text{map}(f, \vec{x})) = \text{length}(\vec{x}).$$

### 3.2.4 Adjunctions

The functor  $-^*$  mentioned above satisfies a number of interesting properties. The set of lists  $A^*$  is equipped with a monoid structure, where  $\epsilon$  is given by the empty list and  $\cdot$  corresponds to concatenation. This means that the functor  $-^* : \mathbf{Set} \rightarrow \mathbf{Set}$  can be decomposed into:

$$\begin{array}{ccc} \mathbf{Set} & \xrightarrow{F} & \mathbf{Mon} \\ & \searrow -^* & \downarrow U \\ & & \mathbf{Set} \end{array}$$

where  $F$  first maps the set  $X \in \mathbf{Set}$  to the monoid object associated with  $X^*$ , and  $U$  then “forgets” the structure of any monoid, retaining only the underlying set:

$$FX := \langle X^*, ++, \text{nil} \rangle \quad U \langle Y, \cdot, \epsilon \rangle := Y.$$

Note that  $\text{map}$  preserves  $++$  and  $\text{nil}$ , so given  $f : A \rightarrow B$ , the function  $f^*$  is indeed a monoid homomorphism and can be used to define  $Ff : \langle A^*, ++, \text{nil} \rangle \rightarrow \langle B^*, ++, \text{nil} \rangle$ . Conversely, if  $g : \langle A, \cdot_A, \epsilon_A \rangle \rightarrow \langle B, \cdot_B, \epsilon_B \rangle$  is a monoid homomorphism,  $U$  simply forgets the properties of  $g$ , retaining only the underlying function  $Ug : A \rightarrow B$ .

The monoid  $FA$  is in fact the *free monoid* generated by the elements of  $A$ . This means there is a natural transformation:  $\eta : \text{id}_{\text{Set}} \rightarrow F$  which maps every element of a set  $X$  to the corresponding element of  $FX$ , namely:

$$\eta_X(x) = x :: \text{nil} \quad (x \in X),$$

and that *any* mapping  $f : X \rightarrow UY$  from the set  $X$  into the monoid  $Y$  can be decomposed into  $Ug \circ \eta_A$ .

which is to say the monoid obtained by taking all possible monoid expressions with variables in  $A$ , then quotienting the resulting set by the associativity and identity laws which monoids must obey. This makes  $FA$  unique: it is in a sense the most general, or the *universal* monoid containing  $A$ .

Given another monoid  $\langle M, \cdot, \epsilon \rangle$  and a function  $f : A \rightarrow UM$  assigning elements of  $M$  to the elements of  $A$ , there is a *unique* monoid homomorphism  $g : FA \rightarrow M$ , defined in this case by:

$$g(\text{nil}) := \epsilon$$

$$g(a :: x) := f(a) \cdot g(x),$$

such that the property  $Ug \circ \eta_A = f$  holds in **Set**:

$$\begin{array}{ccc} A & \xrightarrow{\eta_A} & UFA \\ & \searrow f & \downarrow Ug \\ & & UM \end{array} \quad \begin{array}{ccc} FA & & \\ \downarrow g & & \\ M & & \end{array} \quad \begin{array}{l} (\text{Set}) \\ (\text{Mon}) \end{array}$$

This situation generalizes as follows.

**Definition 3.6** (Adjoint functors). Given two categories  $\mathbf{C}, \mathbf{D}$  and two functors  $F : \mathbf{D} \rightarrow \mathbf{C}$



### 3.3 Cartesian categories

#### 3.3.1 Products

**Definition 3.7.** A *product* of a collection  $(A_i)_{i \in I}$  of objects  $A_i \in \mathbf{C}$  is an object  $A \in \mathbf{C}$  together with a collection of morphisms  $(\pi_i : A \rightarrow A_i)_{i \in I}$ , such that for all  $X \in \mathbf{C}$  and collection of morphisms  $(f_i : X \rightarrow A_i)_{i \in I}$ , there exists a unique morphism  $\langle f_j \rangle_{j \in I} : X \rightarrow A$  such that  $f_i = \pi_i \circ \langle f_j \rangle_{j \in I}$  for all  $i \in I$ . A category with all finite products is called *cartesian*.

A nullary product is called a *terminal object* and written  $1$ . The definition boils down to the existence of a unique morphism  $\langle \rangle : X \rightarrow 1$  for each object  $X \in \mathbf{C}$ :

$$\begin{array}{c} X \\ \vdots \langle \rangle \\ 1 \end{array}$$

The unary product of an object  $A$  is just itself, with the morphism  $\text{id}_A$  as the sole “projection”:

$$\begin{array}{ccc} X & & \\ \vdots & \searrow f & \\ A & \xrightarrow{\text{id}} & A \end{array}$$

Binary products are written  $A \times B$ , and satisfy the following property:

$$\begin{array}{ccccc} & & X & & \\ & \swarrow f & \vdots \langle f, g \rangle & \searrow g & \\ A & \xleftarrow{\pi_1} & A \times B & \xrightarrow{\pi_2} & B \end{array}$$

When they exist, products are unique up to isomorphism, so we can talk about *the* product of  $(A_i)_{i \in I}$ , refer to the corresponding object as  $\prod_{i \in I} A_i$ , or in the finitary case used use the notations mentioned above. Note that:

$$1 \times X \cong X \times 1 \cong X \quad X \times Y \cong Y \times X \quad (X \times Y) \times Z \cong X \times (Y \times Z)$$

In other words,  $\langle \mathbf{C}/\cong, \times, 1 \rangle$  behaves like a commutative monoid, so we do not need to be too

careful about parentheses.

**Example 3.8** (Products in **Set** and **Mon**). *In **Set**, singletons like the unit set  $\mathbb{1} = \{*\}$  are terminal objects, and the morphism  $\langle \rangle : A \rightarrow \mathbb{1}$  is the only possible function into  $\mathbb{1}$ , defined by  $\langle \rangle(a) := *$ . Binary products are given by sets of pairs:*

$$A \times B := \{(a, b) \mid a \in A \wedge b \in B\} \quad \pi_1(a, b) := a \quad \pi_2(a, b) := b$$

*Products in **Mon** extend the definitions above to provide a monoid structure on the underlying sets. The terminal monoid uses  $\mathbb{1}$  as its underlying set with the monoid structure defined by  $* \cdot_1 * := *$  and  $\epsilon_1 := *$ . Similarly, the monoid structure on  $A \times B$  is defined componentwise as:*

$$(x_1, x_2) \cdot_{A \times B} (y_1, y_2) := (x_1 \cdot_A y_1, x_2 \cdot_B y_2) \quad \epsilon_{A \times B} := (\epsilon_A, \epsilon_B)$$

*A similar approach can be used for products of arbitrary arities.*

The categories listed in Tbl. 3.1 are all cartesian. The construction of their products is similar to the one used for **Mon**. We start with the product of the underlying sets, and define the appropriate structure componentwise.

### 3.3.2 Generalized elements

Terminal objects are useful for representing elements as morphisms. For example in **Set**, there is a one-to-one correspondence between the morphism  $e : 1 \rightarrow A$  and the element  $e(*) \in A$ . Combining this with the terminal morphism  $\langle \rangle : X \rightarrow 1$ , we can define the constant function  $e \circ \langle \rangle : X \rightarrow A$  mapping all elements of a domain set  $X$  to the element  $e(*)$  of  $A$ . More generally, products can be used to represent functions of  $n$  arguments as morphisms of type  $f : A_1 \times \cdots \times A_n \rightarrow B$ . We can then compose  $f$  with a family  $(f_i : X \rightarrow A_i)_{1 \leq i \leq n}$  of functions producing arguments of  $f$  as  $f \circ \langle f_1, \dots, f_n \rangle : X \rightarrow B$ .

In a category with a terminal object  $1$ , the morphisms of type  $1 \rightarrow A$  are known as the *global elements* of  $A$ , but in fact we can extend this intuition to *any* morphism  $e : X \rightarrow A$  by viewing it as a *generalized element* of  $A$  in the context of  $X$ . This point of view is used quite literally in categorical models of type theories.

**Example 3.9** (Categorical semantics). Consider a simple type theory, for example the simply-typed lambda calculus, built around a typing judgement of the form:

$$x_1 : A_1, \dots, x_n : A_n \vdash M : B,$$

which asserts that the term  $M$  has type  $B$  in a context where each variable  $x_i$  has type  $A_i$ .

To interpret a type theory of this kind in a cartesian category  $\mathbf{C}$ , we can assign to each type  $A$  an object  $\llbracket A \rrbracket \in \mathbf{C}$ , and to each well-typed term a morphism of type:

$$\llbracket x_1 : A_1, \dots, x_n : A_n \vdash M : B \rrbracket : \llbracket A_1 \rrbracket \times \dots \times \llbracket A_n \rrbracket \rightarrow \llbracket B \rrbracket.$$

The syntactic constructions of the language can then be interpreted in terms of (semantic) categorical constructions. In particular, given the well-typed terms:

$$\Gamma, x : A \vdash M : B \qquad \Gamma \vdash N : A,$$

the substitution  $M[x/N]$  of  $x$  by  $N$  in  $M$  will be interpreted as:

$$\llbracket \Gamma \vdash M[x/N] : B \rrbracket : \llbracket \Gamma \rrbracket \rightarrow \llbracket B \rrbracket := \llbracket \Gamma, x : A \vdash M : B \rrbracket \circ \langle \text{id}_{\llbracket \Gamma \rrbracket}, \llbracket \Gamma \vdash N : A \rrbracket \rangle.$$

As another example, the language may contain product types, usually defined by the typing rules:

$$\frac{\Gamma \vdash M : A \quad \Gamma \vdash N : B}{\Gamma \vdash (M, N) : A \times B} \qquad \frac{\Gamma \vdash M : A \times B}{\Gamma \vdash \text{fst}(M) : A} \qquad \frac{\Gamma \vdash M : A \times B}{\Gamma \vdash \text{snd}(M) : B}$$

and the reduction rules:

$$\text{fst}(x, y) \rightsquigarrow x \qquad \text{snd}(x, y) \rightsquigarrow y \qquad (\text{fst}(z), \text{snd}(z)) \rightsquigarrow z$$

Since these rules correspond to the categorical notion of product, we can use the interpretation:

$$\begin{aligned} \llbracket A \times B \rrbracket &:= \llbracket A \rrbracket \times \llbracket B \rrbracket \\ \llbracket \Gamma \vdash (M, N) : A \times B \rrbracket &:= \langle \llbracket \Gamma \vdash M : A \rrbracket, \llbracket \Gamma \vdash N : B \rrbracket \rangle \\ \llbracket \Gamma \vdash \text{fst}(M) : A \rrbracket &:= \pi_1 \circ \llbracket \Gamma \vdash M : A \times B \rrbracket \\ \llbracket \Gamma \vdash \text{snd}(M) : B \rrbracket &:= \pi_2 \circ \llbracket \Gamma \vdash M : A \times B \rrbracket. \end{aligned}$$

Defining categorical semantics for programming languages in this way establishes general principles for interpreting a language in any category with the required structure. For example, the  $\kappa$ -calculus [Hasegawa, 1995] can be interpreted along the lines of Ex. 3.9 in any cartesian category. As such it gives a computational, variable-based syntax which we can use to define morphisms in **Set**, **Mon**, and many other categories.

### 3.3.3 Cartesian closure

A category **C** is *closed* when its collections of morphisms  $\mathbf{C}(X, Y)$  can be represented internally as an object of **C** itself. In *cartesian closed* categories, this correspondance is expressed using products. Once again **Set** is the prototypical example.

**Example 3.10** (Cartesian closure in **Set**). *Given two sets  $A, A \in \mathbf{Set}$ , the function space  $\mathbf{Set}(A, A)$  is itself a set and therefore gives an object  $A^A \in \mathbf{Set}$ . This means there exists a one-to-one correspondence between the morphisms of type  $A \rightarrow A$  and the global elements of  $A^A$ :*

$$\mathbf{Set}(A, A) \cong \mathbf{Set}(1, A^A)$$

*In fact, more generally, there is a correspondence known as currying between the morphisms of type  $\Gamma \times A \rightarrow A$  and the generalized elements  $\Gamma \rightarrow A^A$ :*

$$\mathbf{Set}(\Gamma \times A, A) \cong \mathbf{Set}(\Gamma, A^A)$$

*This correspondence is in fact an adjunction  $- \times A \dashv -^A$  between:*

- the functor  $- \times A$ , which maps an object  $X$  to the product  $X \times A$ , and  $f : X \rightarrow Y$  to the

morphism  $f \times \text{id}_A : X \times A \rightarrow Y \times A$ ;

- the functor  $-^A$ , which maps an object  $X$  to the function space  $X^A$  and  $f : X \rightarrow Y$  to the morphism  $f^A : X^A \rightarrow Y^A$  defined by  $f^A(\phi) := f \circ \phi$ .

This adjunction is spelled out in full in the following diagram, where the morphisms  $f : \Gamma \times A \rightarrow A$  and  $g : \Gamma \rightarrow A^A$  determine each other:

$$\begin{array}{ccc}
 \Gamma \times A & \xrightarrow{\eta_\Gamma} & (\Gamma \times A)^A \\
 \downarrow g \times \text{id}_A & \searrow f & \downarrow f^A \\
 B^A \times A & \xrightarrow{\varepsilon_B} & B
 \end{array}
 \quad
 \begin{array}{ccc}
 \Gamma & \xrightarrow{\eta_\Gamma} & (\Gamma \times A)^A \\
 & \searrow g & \downarrow f^A \\
 & & B^A
 \end{array}$$

$f = \varepsilon_B \circ (g \times \text{id}_A)$   
 $g = f^A \circ \eta_\Gamma$

The adjunction's counit  $\varepsilon_B : B^A \times A \rightarrow B$  is an evaluation map, which takes a function  $\phi \in B^A$  and an argument  $a \in A$  and returns the value  $\phi(a)$ . The unit  $\eta_\Gamma : \Gamma \rightarrow (\Gamma \times A)^A$  takes a value  $\rho \in \Gamma$  and returns a function which constructs a pair  $(\rho, a)$  from its argument  $a \in A$ .

This structure provides a categorical notion of high-order functions we can be introduced by currying and eliminated through uncurrying. The same structure can be obtained in any cartesian category where functors  $- \times A$  have right adjoints.

**Definition 3.11** (Cartesian closed category). An object  $A$  in a cartesian category  $\mathbf{C}$  is called *exponentiable* if there is a functor  $-^A$  and a natural transformation  $\varepsilon^A : -^A \times A \rightarrow \text{id}_{\mathbf{C}}$  such that for all  $f : \Gamma \times A \rightarrow B$ , there exists a unique morphism  $f^\dagger : \Gamma \rightarrow B^A$  such that:

$$\begin{array}{ccc}
 \Gamma & & \Gamma \times A \\
 \downarrow f^\dagger & & \downarrow f^\dagger \times \text{id}_A \\
 B^A & & B^A \times A \\
 & & \downarrow \varepsilon_B^A \\
 & & B
 \end{array}
 \quad
 \begin{array}{ccc}
 \Gamma \times A & \xrightarrow{f} & B \\
 \downarrow f^\dagger \times \text{id}_A & \searrow & \downarrow \varepsilon_B^A \\
 B^A \times A & \xrightarrow{\varepsilon_B^A} & B
 \end{array}$$

$f = \varepsilon_B^A \circ (f^\dagger \times \text{id}_A)$ .

We call  $\varepsilon_B^A$  an *evaluation map* and the object  $B^A$  an *exponential object*. A cartesian category where all objects are exponentiable is called *cartesian closed*.

The cartesian closed categories are the models of the simply typed  $\lambda$ -calculus.

**Example 3.12** (Simply-typed  $\lambda$ -calculus). *In addition to the structures mentioned in Example 3.9, the simply typed  $\lambda$ -calculus has types of the form  $A \rightarrow B$  with the associated rules:*

$$\frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x \cdot M : A \rightarrow B} \quad \frac{\Gamma \vdash f : A \rightarrow B \quad \Gamma \vdash x : A}{\Gamma \vdash f x : B}$$

and reductions:

$$(\lambda x \cdot M) N \rightsquigarrow M[x/N] \quad (\lambda x \cdot f x) \rightsquigarrow f$$

In a cartesian closed category, we can give these constructions the following interpretation:

$$\begin{aligned} \llbracket A \rightarrow B \rrbracket &:= B^A \\ \llbracket \Gamma \vdash \lambda x \cdot M : A \rightarrow B \rrbracket &:= \llbracket \Gamma, x : A \vdash M : B \rrbracket^\dagger \\ \llbracket \Gamma \vdash f x : B \rrbracket &:= \varepsilon_B \circ \langle \llbracket \Gamma \vdash f : A \rightarrow B \rrbracket, \llbracket \Gamma \vdash x : A \rrbracket \rangle. \end{aligned}$$

### 3.3.4 Symmetric monoidal categories

**Example 3.13** (Cartesian structure in **Sup**). *The category **Sup** has complete join-semilattices as objects, and join-preserving functions as morphisms. Since by definition, any  $L \in \mathbf{Sup}$  has all joins, in particular it has an element  $\perp \in L$  defined by  $\perp = \bigvee \emptyset$ , which is preserved by morphisms.*

*Products in **Sup** are defined as expected: the underlying set  $|\prod_{i \in I} L_i|$  is the cartesian product of sets  $\prod_{i \in I} |L_i|$ , so that its elements have the form  $(x_i)_{i \in I}$  with  $x_i \in L_i$ . Their ordering is defined componentwise as:*

$$(x_i)_{i \in I} \leq (y_i)_{i \in I} \Leftrightarrow \forall i \in I \cdot x_i \leq_i y_i,$$

and joins can likewise be computed as:

$$\bigvee_{j \in J} (x_i^j)_{i \in I} = \left( \bigvee_{j \in J} x_i^j \right)_{i \in I}.$$

*In particular, the terminal object is the trivial suplattice  $\{\perp\}$  as expected. However, a global element  $e : 1 \rightarrow L$  in **Sup** is meaningless, because the sup-morphism  $e$  can only map  $\perp$  to  $\perp_L$ .*

*We will see in §??*

**Example 3.14.** *The category **Sup** is catesian. The homset  $\mathbf{Sup}(L, M)$  can be made into a semilattice*

$[L, M]$ , with joins defined pointwise:

$$\left( \bigvee_{i \in I} f_i \right) (x) = \bigvee_{i \in I} f_i(x),$$

so we would like **Sup** to be closed.

However, the object  $0 = 1 = \{\perp\}$  is both initial and terminal. Global elements  $e : 1 \rightarrow L$  are initial morphisms and carry no information, since they can only map  $\perp \in 1$  to  $\perp \in L$ . This means in particular that in general:

$$\mathbf{Sup}(1 \times L, M) \not\cong \mathbf{Sup}(1, [L, M]),$$

since the right-hand side is a unit set.

There is however a semilattice  $I = \{\perp, *\}$  which can take over the role of  $1$ , in the sense that generalized elements  $e : I \rightarrow L$  are in one-to-one correspondence with the elements  $e(*) \in L$ . Moreover, the functor  $[L, -]$  has a left adjoint  $- \otimes L$ ,

**Definition 3.15** (Monoidal category).

**Definition 3.16** (Symmetric monoidal category).

## 3.4 Categorical algebra

**Definition 3.17** (Monad).

**Definition 3.18** ( $F$ -algebra).

### 3.4.1 Monads

[comparing with Def. 2.20:] The usual categorical definition can be recovered by taking  $\mu_X := \text{id}_{T(X)}^\dagger$  and  $T(f) := (\eta_Y \circ f)^\dagger$ .

The concepts discussed above admit an elegant categorical formulation [Trnková et al., 1975]. This allows us to generalize them to categories other than **Set**, and to establish a connection with the monadic approach to modeling effectful computations.

**Definition 3.19.** The *functor* associated to an effect signature  $E$  is  $\hat{E} : \mathbf{Set} \rightarrow \mathbf{Set}$ , defined as:

$$\hat{E}(X) := \sum_{(m:N) \in E} X^N.$$

Then an interpretation of  $E$  is also an algebra for the endofunctor  $\hat{E}$ , that is an object  $A \in \mathbf{Set}$  together with a morphism  $\phi : \hat{E}(A) \rightarrow A$ .

The interpretations of  $E$  form a category  $\mathbf{Int}_E$ . Homomorphisms of interpretations are function  $f : A \rightarrow B$  between the carrier sets of the interpretations  $(A, \varphi^A)$  and  $(B, \varphi^B)$  such that:

$$\begin{array}{ccc} \hat{E}A & \xrightarrow{\varphi^A} & A \\ \hat{E}f \downarrow & & \downarrow f \\ \hat{E}B & \xrightarrow{\varphi^B} & B \end{array}$$

In other words, for each operation  $(m : N) \in E$  and argument tuple  $x \in A^N$ :

$$f(\varphi_m^A(x_n)_{n \in N}) = \varphi_m^B(f(x_n))_{n \in N}.$$

The forgetful functor  $U_E : \mathbf{Int}_E \rightarrow \mathbf{Set}$  retains the underlying set  $A$  of an interpretation  $(A, \varphi^A)$ , and the underlying function  $f$  of homomorphisms. Then  $\mathcal{F}_E$  can be characterized as the left adjoint  $\mathcal{F}_E \dashv U_E$  where  $\mathcal{F}_E(A) \in \mathbf{Int}_E$  is equipped with the interpretation:

$$\varphi_m(x_n)_{n \in N} := m(x_n)_{n \in N} \quad \text{for all } (m : N) \in E,$$

where the adjunction's unit  $\eta : \mathbf{1} \rightarrow U_E \mathcal{F}_E$  is defined as:

$$\eta_A(v) := v \quad v \in A$$

and the counit  $\epsilon : \mathcal{F}_E U_E \rightarrow \mathbf{1}$  is defined as:

$$\begin{aligned} \epsilon_{(A, \varphi^A)}(v) &:= v & v \in A \\ \epsilon_{(A, \varphi^A)}(m(x_n)_{n \in N}) &:= \varphi_m^A(x_n)_{n \in N} & (m : N) \in E \end{aligned}$$



In other words, a defining property of  $\mathcal{F}_E$  is that for a set of variables  $A$ , an interpretation  $(B, \varphi^B)$ , and a valuation  $\rho : A \rightarrow B$ , there exists a unique homomorphism  $\rho^* : \mathcal{F}_E(A) \rightarrow (B, \varphi^B)$ :

$$\begin{array}{ccccc}
 A & \xrightarrow{\eta_E} & \mathcal{F}_E(A) & \xleftarrow{\varphi} & \hat{E} \mathcal{F}_E(A) \\
 & \searrow \rho & \downarrow \rho^* & & \downarrow \hat{E} \rho^* \\
 & & B & \xleftarrow{\varphi^B} & \hat{E} B
 \end{array}$$

This homomorphism uses  $\rho$  to interpret variables and  $\varphi^B$  to interpret operations:

$$\begin{aligned}
 \rho^*(v) &:= \rho(v) & v &\in A \\
 \rho^*(m(x_n)_{n \in N}) &:= \varphi^B(\rho^*(x_n))_{n \in N} & (m : N) &\in E
 \end{aligned}$$

## Chapter 4

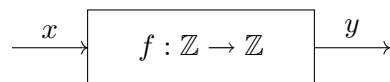
# Games and dual nondeterminism

[XXX reuse if useful: In this chapter, I adapt the work of Morris and Tyrrell [Morris, 2004; Morris and Tyrrell, 2008], which extends the refinement calculus to the level of terms using *free completely distributive completions* of posets, to investigate *dual nondeterminism* in the context of game semantics and construct completely distributive lattices of *strategy specifications*, partially ordered under a form of alternating refinement [Alur et al., 1998].]

Using the lens of dual nondeterminism, §?? outlined briefly some of the challenges associated with nondeterminism and refinement in game semantics. This section presents a new approach to address them, leading to the construction of *strategy specifications* featuring dual nondeterminism and a form of alternating refinement. This approach decouples the nature of nondeterminism from the structure of plays, leading to more uniform and tractable theory.

### 4.1 Example: function specifications

To illustrate the use of dual nondeterminism in the context of specifications and refinement, I will use functions on integers. A function  $f : \mathbb{Z} \rightarrow \mathbb{Z}$  can be seen as a simple system, which accepts a single input and produces a single output:



### 4.1.1 Elementary function specifications

To constrain the behavior of such a system, we can use an elementary specification of the form:

$$x \mapsto y$$

The specification above asserts that whenever the function's input is  $x \in \mathbb{Z}$ , it should produce the output  $y \in \mathbb{Z}$ ; in other words,  $f(x) = y$ . For example, the function defined by  $f(x) := 2x$  satisfies the following specifications:

$$0 \mapsto 0 \quad 1 \mapsto 2 \quad -1 \mapsto -2 \quad 2 \mapsto 4.$$

On their own, these specifications carry very limited information about the function they attempt to characterize. Like the plays of a game, they only give a snapshot of a single interaction between the system and its environment. However, by incorporating angelic and demonic nondeterminism, the model can be made much more expressive.

### 4.1.2 Angelic and demonic choices

Angelic choice allows us to range over various possible choices of the environment, placing more constraint on the function and making the specification stronger. For example, the specification:

$$-1 \mapsto 1 \sqcap 0 \mapsto 0 \sqcap 1 \mapsto 1$$

is satisfied by both of the functions  $g(x) := x^2$  and  $h(x) := |x|$ , but not by  $f$  although it satisfies the middle component. Infinite choice allows us to formulate richer specifications:

$$\hat{f} := \bigsqcup_{x \in \mathbb{Z}} x \mapsto 2x$$

The specification above characterizes the function  $f$  exactly. This allows  $f$  to be represented as a specification, and the fact that  $f$  satisfies  $1 \mapsto 2$  to be expressed as:

$$1 \mapsto 2 \sqsubseteq \bigsqcup_{x \in \mathbb{Z}} x \mapsto 2x = \hat{f}.$$

This also illustrates that a specification can be made stronger and more precise by adding more angelic choices, although if we go too far it may not be possible to implement it; for example, no function satisfies:

$$0 \mapsto -1 \sqcup 0 \mapsto 1.$$

Conversely, demonic choice allows us to express implementation freedom and weaken specifications. For example:

$$1 \mapsto 1 \sqcap 1 \mapsto 2$$

is satisfied the each one of the functions  $f, g, h$ . Using infinite choice, the specification:

$$\bigsqcap_{x \in \mathbb{Z}} x \mapsto x + 1$$

expresses that there must be *at least one*  $x$  mapped by the function to its successor. It can be refined by narrowing down the range of demonic choices:

$$\bigsqcap_{x \in \mathbb{Z}} x \mapsto x + 1 \sqsubseteq 0 \mapsto 1 \sqcap 1 \mapsto 2.$$

This specification is satisfied by  $f$ , but not by  $g$  or  $h$ .

Angelic and demonic choices become even more powerful when they are used together. For example:

$$\bigsqcup_{x \text{ odd}} \bigsqcap_{y \text{ even}} (x \mapsto y)$$

expresses the constraint that *all* odd inputs must be mapped to *some* even output. Again  $f$  satisfies this specification but  $g$  and  $h$  do not.

To summarize, adding arbitrary angelic and demonic choices to a specification framework allows us to gain a significant amount of expressivity. When used in specifications, angelic choice

corresponds to logical conjunction and the *for all* quantifier, and demonic choice corresponds to disjunction and *there exists*.

### 4.1.3 Data abstraction

To illustrate the expressivity of the resulting model, consider the usual construction of integers as pairs of natural numbers.

An integer is represented by a pair  $n = (n_1, n_2) \in \mathbb{N} \times \mathbb{N}$ , where  $n_1$  is understood as the positive component,  $n_2$  is understood as the negative component, and two integers  $n = (n_1, n_2)$  and  $m = (m_1, m_2)$  in this form are considered equal when  $n_1 + m_2 = n_2 + m_1$ . The canonical embedding of natural numbers into integers can be realized as  $n \mapsto (n, 0)$ , and the standard arithmetic operations can be defined as:

$$n + m := (n_1 + m_1, n_2 + m_2)$$

$$n - m := (n_1 + m_2, n_2 + m_1)$$

$$n \times m := (n_1 \times m_1 + n_2 \times m_2, n_1 \times m_2 + n_2 \times m_1).$$

This “low-level” representation can be connected to the more usual and abstract notion of integer by the relation  $\rho \subseteq \mathbb{Z} \times (\mathbb{N} \times \mathbb{N})$  defined by:

$$x \rho (n_1, n_2) \Leftrightarrow x = n_1 - n_2.$$

Then a natural question is: given a function  $g : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$  on the concrete representation, and a function  $f : \mathbb{Z} \rightarrow \mathbb{Z}$  expressed in the more abstract representation, what does it mean for  $g$  to be a correct implementation of  $f$ ? This situation is depicted by the simulation diagram below. Given related inputs  $x \rho n$ , the functions  $f$  and  $g$  are expected to produce related outputs  $f(x) \rho g(n)$ :

$$\begin{array}{ccc}
 x & \xrightarrow{f} & f(x) \\
 \rho \downarrow & & \downarrow \rho \\
 n & \xrightarrow{g} & g(n)
 \end{array}
 \quad
 \begin{array}{c}
 (\forall) \qquad \qquad \qquad (\exists)
 \end{array}$$

$$f [\rho \rightarrow \rho] g$$

However, since this involves objects of two different types, it is not obvious how this property can be formulated in the context of refinement. Dual nondeterminism makes it possible to define concretization and abstraction functions:

$$\begin{aligned} \gamma(f) &:= \bigsqcup_{n \in \mathbb{N} \times \mathbb{N}} \bigsqcup_{x \in \mathbb{Z} \mid x \rho n} \bigsqcap_{m \in \mathbb{N} \times \mathbb{N} \mid f(x) \rho m} & n \mapsto m \\ \alpha(g) &:= \bigsqcup_{x \in \mathbb{Z}} \bigsqcap_{n \in \mathbb{N} \times \mathbb{N} \mid x \rho n} \bigsqcup_{y \in \mathbb{Z} \mid y \rho g(n)} & x \mapsto y \end{aligned}$$

such that the correspondence of  $f$  and  $g$  to be stated as:

$$\gamma(f) \sqsubseteq \hat{g} \quad \Leftrightarrow \quad \hat{f} \sqsubseteq \alpha(g).$$

## 4.2 Refinement in game semantics

Game semantics represents strategies as sets of plays. Each play records a possible interaction between the system and the environment. For instance, if we interpret the elementary function specification  $x \mapsto y$  as a simple play containing one move of the environment ( $x$ ) followed by one move of the system ( $y$ ), the strategy associated to  $f : \mathbb{Z} \rightarrow \mathbb{Z}$  can be represented as:

$$\sigma := \{x \mapsto f(x) \mid x \in \mathbb{Z}\}.$$

### 4.2.1 Strategies and refinement

In most game models, there are restrictions on the contents of  $\sigma$ . One common restriction is the exclusion of nondeterministic choice involving moves of the system:

$$x \mapsto y_1 \in \sigma \wedge x \mapsto y_2 \in \sigma \quad \Rightarrow \quad y_1 = y_2. \quad (4.1)$$

Another common requirement is for  $\sigma$  to range over all possible behaviors of the environment:

$$\forall x \in \mathbb{Z} \cdot \exists y \in \mathbb{Z} \cdot x \mapsto y \in \sigma. \quad (4.2)$$

In our setting, these constraints restrict strategies to precisely describe functions.

To obtain a richer model, we can relax these restrictions, allowing strategies such as:

$$\sigma' := \{0 \mapsto 0, 1 \mapsto 1, 1 \mapsto -1\}.$$

On one hand, the relaxation of (4.1) permits some form of nondeterminism, for example between the elementary specifications  $1 \mapsto 1$  and  $1 \mapsto -1$  in the example above. Indeed this was the approach taken in early work on nondeterminism in game semantics [Harmer and McCusker, 1999]. On the other hand, the relaxation of (4.2) allows strategies which do not completely characterize functions.

Yet even after these constraints have been relaxed, there is a choice to be made as to how this nondeterminism should be interpreted, which will in particular determine how the refinement ordering of strategies is to be constructed:

- We can use an *angelic* interpretation and define refinement as set inclusion ( $\subseteq$ ). This is the original interpretation and allows us to characterize a function completely by ranging over the domain. However, this prevents us from expressing a looser specification permitting a range of outputs for a given input. For example, the component  $1 \mapsto 1 \sqcup 1 \mapsto -1$  of  $\sigma'$  results in an unimplementable specification, since it requires both  $f(1) = 1$  and  $f(1) = -1$ .
- If we use a *demonic* interpretation and define refinement as set containment ( $\supseteq$ ), we cannot formulate any specification stronger than a single  $x \mapsto y$ . In particular functions themselves cannot be represented as sets of elementary specifications in this way. A function  $f$  which satisfies *any one* of the conditions  $f(0) = 0$ ,  $f(1) = 1$  or  $f(1) = -1$  will match  $\sigma'$ .
- To address these issues, we can use an *alternating* interpretation of nondeterminism, which takes into account the polarity of moves, so that  $\sigma'$  above is interpreted as:

$$0 \mapsto 0 \sqcup (1 \mapsto 1 \sqcap 1 \mapsto -1),$$

requiring that the function map 0 to 0 and map 1 to either  $-1$  or  $1$ .

The alternating interpretation of strategy nondeterminism is the one most in line with our goals. In the context of games, choices of the system and the environment must be distinguished,

and it is natural to use the polarity of moves to make this distinction. However, the resulting refinement ordering is complex and unintuitive. In the case of function specifications, we get for example:

$$\{0 \mapsto 0\} \sqsubseteq \{0 \mapsto 0, 1 \mapsto 1, 1 \mapsto -1\} \sqsubseteq \{0 \mapsto 0, 1 \mapsto 1\}$$

In the more general case where plays consist of an unbounded alternation of input and outputs the refinement ordering becomes even more complex to describe.

Another issue is incompleteness. For example, the nondeterministic choices:

$$\{0 \mapsto 1\} \sqcup \{0 \mapsto -1\} \quad \{1 \mapsto 0\} \sqcap \{-1 \mapsto 0\}$$

cannot be represented in the model proposed above. This restriction may at first seem harmless. After all, the first specification above is unimplementable, and the second one is sufficiently weak that it is unclear whether it could ever find any practical use. However, from an algebraic point of view the loss of completeness is a significant shortcoming. It breaks the uniformity of the model and imposes a proliferation of side-conditions which the user must establish.

Consider for example the case of data abstraction as discussed in §4.1.3. The completely distributive lattice structure associated with dual nondeterminism allows us to use arbitrary abstraction relations. If the relation makes an abstract specifications impossible to implement, for instance because it involves abstract data with no concrete representation, or concrete data with an ambiguous high-level interpretation, the associated concretization function  $\gamma$  will simply generate an unimplementable low-level specification.

If this is not possible because this low-level specification cannot be represented, we need instead to formulate, prove and track various restrictions on the abstraction relation or the specifications being manipulated. This is a particular concern in the context of formal proofs mechanized in a proof assistant, where even simple and intuitive side-conditions may interfere with usability.

## 4.2.2 Strategy specifications

Instead of retrofitting nondeterminism and refinement onto the usual notion of strategy, we can reexamine their construction altogether, taking the following approach:



- Individual plays will be understood as elementary specifications. Like the elementary function specifications of §4.1, they indicate how the system may react to a specific behavior of the environment. The prefix relation on plays can then be understood as a notion of refinement, with longer plays placing more precise constraints on the system.
- Sets of plays (strategies) allow us to range over all possible choices of the environment, understood as angelic nondeterminism. They are ordered by inclusion ( $\subseteq$ ). By requiring them to be prefix-closed, we can make sure they incorporate and preserve the refinement ordering of plays.
- Sets of strategies (strategy *specifications*) additionally permits choices of the system, understood as demonic nondeterminism. They are ordered by containment ( $\supseteq$ ). By requiring them to be upward closed with respect to strategy inclusion, we can once again incorporate strategy refinement into this ordering.

Under this approach, dual nondeterminism is built into the construction of strategy specifications, allowing us to avoid the pitfalls of previous approaches. The model outlined above is complete and general in a precise way: this construction corresponds to the free completely distributive lattice generated by the underlying poset of plays.

### 4.3 Free completely distributive completions

Historically, dual nondeterminism in the refinement calculus only operates at the level of statements, in the context of imperative programming. More recently, Morris and Tyrrell were able to extend the lattice-theoretic approach used in the refinement calculus to functional programming [Morris, 2004; Morris and Tyrrell, 2008; Tyrrell et al., 2006].

In essence, their approach works by capturing dual nondeterminism as a *monad* constructing free completely distributive lattices. Importantly, this monad operates on *partially ordered sets*, allowing it to incorporate any preexisting notion of refinement. This allows dual nondeterminism to be used in a variety of new contexts.

**Definition 4.1.** A completely distributive lattice  $L$  is a *free completely distributive completion* of a poset  $C$  if there is a monotonic function  $\phi : C \rightarrow L$  such that for any completely distributive

lattice  $M$  and monotonic function  $f : C \rightarrow M$ , there exists a unique complete homomorphism  $f_\phi^* : L \rightarrow M$  such that  $f_\phi^* \circ \phi = f$ :

$$\begin{array}{ccc} C & \xrightarrow{\phi} & L \\ & \searrow f & \downarrow f_\phi^* \\ & & M \end{array}$$

A free completely distributive completion of a poset always exists and it is unique up to isomorphism. We write  $\mathbf{FCD}(C)$  for the free completely distributive completion of  $C$ .

### 4.3.1 Construction

Morris [2004] gives the following constructions for the free completely distributive completion of a partially ordered set  $(A, \leq)$ :

$$\begin{aligned} \mathbf{FCD}(A, \leq) &:= \mathcal{DU}(A, \leq) & \mathbf{FCD}(A, \leq) &:= \mathcal{UD}(A, \leq) \\ \phi(a) &:= \downarrow \uparrow a & \phi(a) &:= \uparrow \downarrow a. \end{aligned}$$

In the expressions above,  $\mathcal{D}$  and  $\mathcal{U}$  are themselves completions. A *downset* of a poset  $(A, \leq)$  is a subset  $x \subseteq A$  satisfying the *downward closure* property:

$$\forall a, b \in A \cdot a \leq b \wedge b \in x \Rightarrow a \in x.$$

Unions and intersections preserve downward closure, giving rise to the *downset lattice*  $\mathcal{D}(A, \leq)$ . Its elements are the downsets of  $(A, \leq)$ , ordered by set inclusion ( $\subseteq$ ) with unions as joins and intersections as meets. The dual *upset lattice*  $\mathcal{U}(A, \leq)$  is ordered by set containment ( $\supseteq$ ) with intersections as joins and unions as meets.

### 4.3.2 Categorical characterization

Categorically speaking,  $\mathbf{FCD} : \mathbf{Poset} \rightarrow \mathbf{CDLat}$  is the left adjoint to the forgetful functor  $U : \mathbf{CDLat} \rightarrow \mathbf{Poset}$  from the category  $\mathbf{CDLat}$  of completely distributive lattices and complete homomorphisms to the category  $\mathbf{Poset}$  of partially ordered sets and monotonic functions. As

such,  $U \circ \mathbf{FCD}$  is a monad over  $\mathbf{Poset}$ , and can be used to model dual nondeterminism as an effect. In the remainder of this thesis, I will identify  $\mathbf{FCD}$  with the monad  $U \circ \mathbf{FCD}$ , and refer to the function  $U(f_\phi^*)$  as the  $\mathbf{FCD}$  extension of  $f$ .

### 4.3.3 Computational interpretation

Computationally, the  $\mathbf{FCD}$  monad can be used to interpret dual nondeterminism as an effect. As usual,  $\phi(a) \in \mathbf{FCD}(a)$  corresponds to a computation which terminates immediately with the outcome  $a \in A$ . For a computation  $x \in \mathbf{FCD}(A)$  and for  $f : A \rightarrow \mathbf{FCD}(B)$ , the computation  $f_\phi^*(x) \in \mathbf{FCD}(B)$  replaces any outcome  $a$  of  $x$  with the computation  $f(a)$ . We will use the notation  $a \leftarrow x; f(a)$  for  $f_\phi^*(x)$ , or simply  $x; y$  when  $f$  is constant with  $f(a) = y$ .

A computation  $x \in \mathbf{FCD}(a)$  can be understood as a *structured* collection of possible outcomes. More precisely, each element  $x \in \mathbf{FCD}(A)$  can be written as  $x = \bigsqcap_{i \in I} \bigsqcup_{j \in J_i} \phi(a_{ij})$  where the index  $i \in I$  ranges over possible demonic choices, the index  $j \in J_i$  ranges over possible angelic choices, and  $a_{ij} \in A$  is the corresponding outcome of the computation. Note that  $f_\phi^*(x) = \bigsqcap_{i \in I} \bigsqcup_{j \in J_i} f(a_{ij})$ .

The algebraic properties of lattices underlie the model's insensitivity to *branching*. Complete distributivity:

$$\bigsqcap_{i \in I} \bigsqcup_{j \in J_i} x_{i,j} = \bigsqcup_{f \in (\prod_i J_i)} \bigsqcap_{i \in I} x_{i,f_i}$$

further allows angelic and demonic choices to commute, and the status of  $f_\phi^*$  as a complete homomorphism enables the following properties:

$$\begin{aligned} a \leftarrow \left( \bigsqcup_{i \in I} x_i \right); M[a] &= \bigsqcup_{i \in I} (a \leftarrow x_i; M[a]) \\ a \leftarrow \left( \bigsqcap_{i \in I} x_i \right); M[a] &= \bigsqcap_{i \in I} (a \leftarrow x_i; M[a]) \\ a \leftarrow x; b \leftarrow y; M[a, b] &= b \leftarrow y; a \leftarrow x; M[a, b] \end{aligned}$$

Finally, the least element  $\perp := \bigsqcup \emptyset$ , traditionally called abort, merits some discussion. As a specification construct, it places no constraint on the implementation (it is refined by every element). As an implementation construct, we use it indiscriminately to interpret failure, silent

divergence, and any other behavior which we want to exclude (it refines only itself). Writing again  $\mathbb{1} = \{*\}$  for the unit set, the *assertion*  $\{P\} \in \mathbf{FCD}(\mathbb{1})$  of a proposition  $P$  evaluates to the unit value  $\phi(*)$  when the proposition is true and to  $\perp$  otherwise. We will use it to formulate guards blocking a subset of angelic choices.

#### 4.4 Dually nondeterministic strategies

[XXX adapt –] A play  $s \in P_G$  can be promoted to a strategy  $\downarrow s \in \mathcal{D}(P_G, \sqsubseteq_p)$ :

$$\downarrow s := \{t \in P_G \mid t \sqsubseteq_p s\}$$

Set inclusion corresponds to strategy refinement, and the downset completion augments  $P_G$  with angelic choices.

A common constraint is that a strategy  $\sigma \in S_G$  should not contain two plays  $sm_1, sm_2 \in \sigma$  where  $m_1$  and  $m_2$  are distinct moves of the system. This is usually understood as enforcing *system determinism*: given a set of environment choices, there is only one possible behavior for the system. Therefore, relaxing this constraint has usually been understood as the first step toward modeling nondeterministic systems [Harmer and McCusker, 1999]. However, looking at the situation through the lens of refinement and dual nondeterminism yields to a different conclusion, where this requirement corresponds to *environment determinacy*.

Angelic nondeterminism allows us to range over all possible choices of the environment and record the resulting plays. For instance, expressing the strategy for  $x := 2 * x$  as a prefix-closed set of even-length plays, we get:

$$\begin{aligned} \sigma &:= \bigcup_{n \in \mathbb{N}} \downarrow (\text{run} \cdot \underline{\text{rd}}_x \cdot n \cdot \underline{\text{wr}}_x[2n] \cdot \text{ok} \cdot \underline{\text{done}}) \\ &= \{\epsilon, \quad \text{run} \cdot \underline{\text{rd}}_x, \quad \text{run} \cdot \underline{\text{rd}}_x \cdot n \cdot \underline{\text{wr}}_x[2n], \quad \text{run} \cdot \underline{\text{rd}}_x \cdot n \cdot \underline{\text{wr}}_x[2n] \cdot \text{ok} \cdot \underline{\text{done}} \mid n \in \mathbb{N}\} \end{aligned}$$

Note that this strategy admits refinements containing much more angelic nondeterminism, includ-

ing with respect to moves of the system. For instance:

$$\sigma \subseteq \bigcup_{n \in \mathbb{N}} \bigcup_{-1 \leq \delta \leq 1} \downarrow(\text{run} \cdot \underline{\text{rd}}_x \cdot n \cdot \underline{\text{wr}}_x[2n + \delta] \cdot \text{ok} \cdot \underline{\text{done}})$$

These refinements do not correspond to interpretations of concrete programs, and in game models which seek to achieve definability they are usually excluded. In our context, retaining them is algebraically important, and they can in fact appear as intermediate terms in some applications. In the construction above, although  $\delta$  appears in a system move, it is still associated with an *angelic* choice. This can be interpreted as a choice of the environment which is not directly observed (perhaps as a result of abstraction), but which nonetheless influences the behavior of the system.

This is different from allowing the *system* to choose an answer in the interval  $[2n - 1, 2n + 1]$ . The model and refinement lattice which we have presented so far are insufficient to express such a specification, because downsets do not add enough meets, forcing the would-be specification to become much coarser:

$$\sigma' := \bigcup_{n \in \mathbb{N}} \bigcap_{-1 \leq \delta \leq 1} \downarrow(\text{run} \cdot \underline{\text{rd}}_x \cdot n \cdot \underline{\text{wr}}_x[2n + \delta] \cdot \text{ok} \cdot \underline{\text{done}}) = \{\epsilon, \text{run} \cdot \underline{\text{rd}}_x\}.$$

We will remedy this by using a richer completion, as explained in §4.

[– XXX]

Our approach to demonic nondeterminism in game semantics will be to substitute **FCD** for  $\mathcal{D}$  in the construction of strategies presented earlier. The more permissive *strategy specification*  $\sigma'$  which we attempted to construct above can then be expressed as:

$$\sigma' := \bigsqcup_{n \in \mathbb{N}} \prod_{-1 \leq \delta \leq 1} \phi(\text{run} \cdot \underline{\text{rd}}_x \cdot n \cdot \underline{\text{wr}}_x[2n + \delta] \cdot \text{ok} \cdot \underline{\text{done}})$$

Because of the properties of **FCD**, the strategy specification  $\sigma'$  will retain not only angelic choices, but demonic choices as well, expressing possible behaviors of the system.

For the construction **FCD** :=  $\mathcal{UD}$ , strategy specifications correspond to sets of traditional strategies, ordered by containment ( $\supseteq$ ). This outer set ranges over demonic choices. Writing

$s_{n,\delta} := \text{run} \cdot \underline{\text{rd}}_{\text{x}} \cdot n \cdot \underline{\text{wr}}_{\text{x}}[2n + \delta] \cdot \text{ok} \cdot \underline{\text{done}}$ , the strategy specification  $\sigma'$  will be encoded as:

$$\sigma' = \{\sigma \in \mathcal{D}(P_G) \mid \forall n \in \mathbb{N} \cdot \exists \delta \in [-1, 1] \cdot s_{n,\delta} \in \sigma\}.$$

Upward closure ensures that a strategy specification which contains a strategy  $\sigma$  contains all of its refinements as well. For instance, the only strategy specification containing the completely undefined strategy  $\emptyset$  is the maximally permissive strategy specification  $\perp = \mathcal{D}(P_G)$ .

## **Part II**

# **Refinement-based game semantics**

## Chapter 5

# Certified abstraction layers

### 5.1 Introduction

This chapter describes the theory of certified abstraction layers used in the verification of the certified operating system kernel CertiKOS.

#### 5.1.1 Abstraction layers

Software is constructed in layers. The basic facilities provided by the programming environment are used to implement more abstract data structures and operations. The programmer can then forget the implementation details of these data structures and operations, and instead think of them as *primitive* constructions when building the next layer of code.

This core principle is especially relevant in the context of system code, where abstraction layers may transform the programming model significantly. For example, at the level of the bare metal, the memory address space must be manipulated explicitly. The operating system's memory management layers abstract away the hardware's low-level details and provide a much more convenient view of the memory for higher-level code.

This illustrates a conceptual similarity between abstraction layers and compilers. As is the case for abstraction layers, the goal of a compiler is to use a lower-level programming model (like assembly code) to provide a more abstract, higher-level one (for example the C programming language). For abstraction layers, the transformation of client code is less expansive, consisting only in adding a layer's code. Nevertheless they can be understood as a specific instance of the



CompCert	CertiKOS
Language semantics	Layer interface
Source language	Overlay interface
Compilation pass	Abstraction layer
Target language	Underlay interface
Pass correctness	Layer correctness

Table 5.1: Correspondance between certified compilation and certified abstraction layers

more general phenomenon.

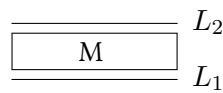
### 5.1.2 Certified abstraction layers in CertiKOS

The analogy between compilers and abstraction layers can be extended to cover correctness and verification techniques. This is illustrated by the *certified abstraction layers* of CertiKOS and their connection to CompCert. The verification of CertiKOS builds on CompCert in two different ways:

- In terms of functionality, CompCert provides mechanized semantics for C and assembly in the Coq proof assistant, and can be used to transport proofs about the C code of the kernel to the level of the compiled assembly code.
- Additionally, many of the patterns and techniques used to establish the correctness theorem of CompCert serve as a starting point for the verification of CertiKOS (Tbl. 5.1).

The kernel is divided into several dozen abstraction layers, which are specified and verified individually. Specifications of abstraction layers are called *layer interfaces*. They extend CompCert semantics to provide a set of *primitives*, which can be invoked as external functions. The behavior of primitives is described in terms of an *abstract state*, maintained alongside CompCert’s memory state, but only updated by primitive invocations.

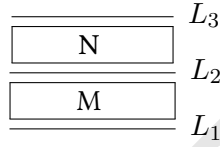
A layer implementation consists in client code for an *underlay* interface, which implements the primitives described by an *overlay* interface. A layer  $M$  implementing the overlay interface  $L_2$  on top of the underlay interface  $L_1$  can be depicted as follows:



The correctness of  $M$  is formulated as the contextual refinement property:

$$\forall C. \llbracket C \rrbracket_{L_2} \sqsubseteq \llbracket C + M \rrbracket_{L_1}$$

expressing that the behavior of the client code  $C$  running alongside  $M$  on top of the underlay interface  $L_1$  will refine the behavior of  $C$  evaluated on top of the overlay interface  $L_2$ . Then, when the underlay interface of one layer corresponds to the overlay interface of another:



the contextual refinement properties can be composed to obtain the composite certified layer:



in the same way the passes of CompCert and their correctness properties can be composed when the target language of one corresponds to the source language of another.

### 5.1.3 Contributions

In our original design [Gu et al., 2015], the formalization of certified abstraction layers was closely tied with the semantic infrastructure of CompCert. Indeed, we introduced the variant CompCertX discussed in §8.2.2 to make it possible for language semantics to use arbitrary underlay interfaces and express the behavior of layer implementations.

As we proceeded to verify the 37 layers of CertiKOS in the framework outlined above, it quickly became apparent that the common structures found in the verification of each layer were a source of excessive redundancy in our proofs. This prompted my work on the *layer calculus* of CertiKOS, which became the starting point for the research presented in this thesis.

This chapter presents a modernized version of the layer calculus decoupled from CompCert. The new version avoids the complications of the original model but captures its essential features. Through a series of embeddings, the new model can be interfaced with CompCertO (§9) to provide the capabilities of the original model and more.

## 5.2 Layer model

A layer interface  $L$  has three components. First, a *signature* enumerates primitive operations together with their types (see Def. 2.14). Second, the set  $S$  contains the *abstract states* of the layer interface. Finally, for each operation  $m : N$ , a *specification* is given as a function:

$$L.m : S \rightarrow \mathcal{P}^1(N \times S).$$

Throughout this thesis, the notation  $v@k \in V \times S$  is used for a pair containing the value  $v \in V$  and the state  $k \in S$ .

### 5.2.1 Specification monad

In the type of  $L.m$  above,  $\mathcal{P}^1$  corresponds to the *maybe* monad:

$$\mathcal{P}^1(A) := \{x \subseteq A : |x| \leq 1\},$$

where a terminating computation producing the value  $v \in A$  is represented as the singleton:

$$\eta_{\mathcal{P}}(v) := \{v\},$$

and where the empty set  $\emptyset \in \mathcal{P}^1(A)$  specifies an undefined computation, which is free to silently diverge, crash, or produce any possible outcome. When two computations are sequentially composed, both must be successful for the result to be defined. For  $x \in \mathcal{P}(\mathcal{P}(A))$ , the multiplication  $\mu_{\mathcal{P}}(x) \in \mathcal{P}(A)$  is defined as:

$$\mu_{\mathcal{P}}(x) := \{v \in A \mid \exists y. x \ni y \ni v\}.$$

The notion of refinement associated with this monad is set inclusion, which allows an unspecified computation to be refined by a computation with a definite outcome:

$$\emptyset \subseteq \{v\}.$$

To account for state, we can combine  $\mathcal{P}^1$  with the state monad transformer. For a set of states  $S$ , we obtain the monad:

$$\mathcal{L}_S(A) := S \rightarrow \mathcal{P}^1(A \times S),$$

equipped with the structure:

$$\begin{aligned} \eta_{\mathcal{L}}(v) &:= \lambda k. \{v @ k\} \\ \mu_{\mathcal{L}}(x) &:= \lambda k. \{v @ k'' \mid \exists y k'. x(k) \ni y @ k' \wedge y(k') \ni v @ k''\}. \end{aligned}$$

This monad is used to specify the primitives in a layer interface using abstract states in the set  $S$ .

## 5.2.2 Layer interfaces

The monad  $\mathcal{L}_S$  is used to define layer interfaces.

**Definition 5.1.** A layer interface is a tuple  $L = \langle E, S, \sigma \rangle$ , where  $E$  is an effect signature,  $S$  is a set of states, and  $\sigma$  assigns to each operation  $(m : N) \in E$  a specification  $\sigma^m \in \mathcal{L}_S(N)$ . We will sometimes use the notation  $L.m$  to refer to  $\sigma^m$ .

As a running example, we will use the certified layer described in Fig. 5.1, which implements a bounded queue with at most  $N$  elements using a circular buffer. We outline the construction of the corresponding layer interfaces below.

**Example 5.2** (Bounded queue and ring buffer interfaces). The layer interface  $L_{\text{rb}} = \langle E_{\text{rb}}, S_{\text{rb}}, \sigma_{\text{rb}} \rangle$  describes a bounded queue. Its states are sequences of values, expected to contain at most  $N$  elements. The two primitives `enq` and `deq` respectively add a new element to the queue and remove the oldest element. If we attempt to add an element which would overflow the queue's capacity  $N$ , or remove an element from an empty queue, the result is  $\emptyset$  (i.e., the operation aborts).

I will demonstrate how the bounded queue interface can be implemented in terms of a circular buffer described by the layer interface  $L_{\text{bq}} = \langle E_{\text{bq}}, S_{\text{bq}}, \sigma_{\text{bq}} \rangle$ . The states of  $L_{\text{rb}}$  contain an array  $f \in V^N$  storing  $N$  values of type  $V$ , and two counters taking values in the interval  $0 \leq c_1, c_2 < N$ . The array can be accessed through the primitives `get` and `set`; the primitives `inc1` and `inc2` increment the corresponding counter and return the counter's old value.

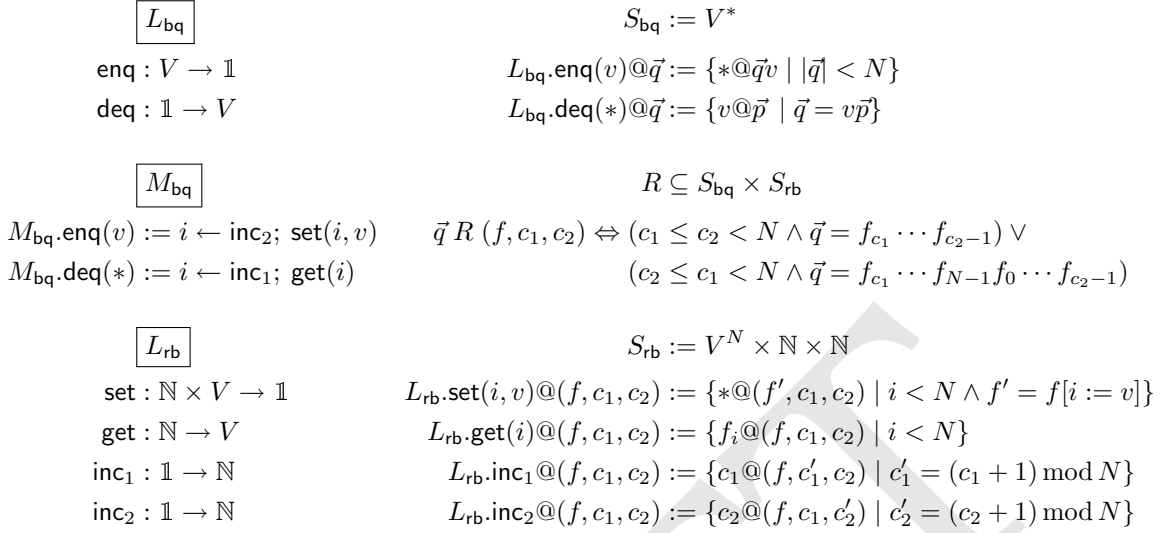


Figure 5.1: A certified abstraction layer  $L_{\text{rb}} \vdash_R M_{\text{bq}} : L_{\text{bq}}$  implementing a bounded queue of size  $N$  using a ring buffer. The left-hand side of the figure shows the signatures of the overlay and underlay interfaces, and the code associated with the layer. The right-hand side shows primitive specifications and the simulation relation used by the correctness proof.

### 5.2.3 Client code

Given an underlay interface  $L = \langle E, S, \sigma \rangle$ , we will use the *free monad*  $\mathcal{F}_E$  on the signature  $E$  as a general representation for client code.

As discussed in §3.4.1, a computation  $x \in \mathcal{F}_E(A)$  in the free monad is a term built out of the operations of  $E$  and the elements of  $A$  regarded as variables. These terms can be interpreted as a computation which proceeds inward from the top-level operation towards the leaves of the term (constants or variables).

Using the underlay interface, we can define an interpretation  $(\varphi_m^L)_{(m:N) \in E}$  of the signature  $E$  into the monad  $\mathcal{L}$  where for each operation  $(m : N) \in E$ :

$$\begin{aligned} \varphi_m^L : \mathcal{L}(A)^N &\rightarrow \mathcal{L}(A) \\ \varphi_m^L(x_n)_{n \in N} &:= n \leftarrow \sigma_m ; x_n \end{aligned}$$

Then  $\varphi^L$  can be extended to interpret any client code  $x \in \mathcal{F}_E(A)$  in terms of the underlay state as  $x[L] \in \mathcal{L}_E(A) := \varphi^L(x)$ .

### 5.2.4 Layer implementations

A layer implementation provides underlay client code for each operation of the overlay interface.

**Definition 5.3.** A layer implementation for an underlay signature  $E$  and an overlay signature  $F$  is a family  $M \in \prod_{(m:N) \in F} \mathcal{F}_E(N)$ . Then, for an underlay interface  $L = \langle E, S, \sigma \rangle$ , the evaluation of  $M$  over  $L$  is the layer interface:

$$M[L] := \langle F, S, M[\sigma] \rangle,$$

where  $M[\sigma]^m := M^m[\sigma]$ .

**Example 5.4.** The layer implementation  $M_{\text{bq}}$  stores the queue's elements into the array, between the indices given by the counters' values. This is expressed by the simulation relation  $R$  given in Fig. 5.1, which explains how overlay states are realized by  $M_{\text{bq}}$  in terms of underlay states.

The code of  $M_{\text{bq}}$  can be interpreted in the monad  $S_{\text{rb}} \rightarrow \mathcal{P}^1(- \times S_{\text{rb}})$ , with calls to primitives of  $L_{\text{rb}}$  replaced by their specifications, to obtain the layer interface  $M_{\text{bq}}[L_{\text{rb}}]$ . Concretely, the behaviors of  $M_{\text{bq}}.\text{enq}(v)$  and  $M_{\text{bq}}.\text{deq}$  running on the underlay interface  $L_{\text{rb}}$  evaluate to:

$$\begin{aligned} M_{\text{bq}}.\text{enq}(v)[L_{\text{rb}}](f, c_1, c_2) &= (i \leftarrow L_{\text{rb}}.\text{inc}_2 ; L_{\text{rb}}.\text{set}(i, v))(f, c_1, c_2) \\ &= L_{\text{rb}}.\text{set}(c_2, v)(f, c_1, (c_2 + 1) \bmod N) \\ &= \{*\text{@}(f[c_2 := v], c_1, (c_2 + 1) \bmod N) \mid c_2 < N\} \\ \\ M_{\text{bq}}.\text{deq}[L_{\text{rb}}](f, c_1, c_2) &= (i \leftarrow L_{\text{rb}}.\text{inc}_1 ; L_{\text{rb}}.\text{get}(i))(f, c_1, c_2) \\ &= L_{\text{rb}}.\text{get}(c_1)(f, (c_1 + 1) \bmod N, c_2) \\ &= \{f_{c_1}\text{@}(f, (c_1 + 1) \bmod N, c_2) \mid c_1 < N\} \end{aligned}$$

### 5.2.5 Correctness

Consider an underlay interface  $L_1$ , an overlay interface  $L_2$ , and a corresponding layer implementation  $M$ . The layer interface  $M[L_1]$  shares a signature with  $L_2$ , but its behavior is described in terms of the states of  $L_1$ . To establish correctness, we need to provide a simulation relation  $R \subseteq S_2 \times S_1$  explaining the correspondence between the more abstract states of  $L_2$  and the more

concrete states of  $L_1$ . Using the relators  $\rightarrow, \times, \mathcal{P}^\leq$  presented in §2.4.2, the associated simulation property can be defined as follows.

**Definition 5.5** (Simulation of layer interfaces). Given a signature  $E$ , two layer interfaces  $L_1 = \langle E, S_1, \sigma_1 \rangle$  and  $L_2 = \langle E, S_2, \sigma_2 \rangle$ , and a relation  $R \subseteq S_2 \times S_1$ , we say that  $L_2$  is *simulated by*  $L_1$  and write  $L_2 \leq_R L_1$  when the following simulation property holds for each operation  $(m : N) \in E$ :

$$L_2.m \ [R \rightarrow \mathcal{P}^\leq (= \times R)] \ L_1.m. \quad (5.1)$$

Then, layer correctness is established by showing a simulation between the overlay interface and the layer implementation.

**Definition 5.6** (Certified abstraction layer). For an underlay interface  $L_1 = \langle E_1, S_1, \sigma_1 \rangle$ , an overlay interface  $L_2 = \langle E_2, S_2, \sigma_2 \rangle$  and a layer implementation  $M : E_1 \rightarrow E_2$ , we say that  $M$  is a *correct implementation of  $L_2$  on top of  $L_1$*  when there is a simulation relation  $R \subseteq S_2 \times S_1$  such that the following property holds:

$$L_1 \vdash_R M : L_2 \quad :\Leftrightarrow \quad L_2 \leq_R M[L_1].$$

We say that  $L_1, M, L_2$  together constitute a *certified abstraction layer*.

**Example 5.7.** We can prove the correctness of the layer implementation  $M_{\text{bq}}$  with respect to the relation  $R$  shown in Fig. 5.1. Writing:

$$[f]_{c_1}^{c_2} := \{f_{c_1} \dots f_{c_2-1} \mid c_1 \leq c_2\} \cup \{f_{c_1} \dots f_{N-1} f_0 \dots f_{c_2} \mid c_2 \leq c_1\},$$

the relation can be formulated as:

$$\vec{q} \ R \ (f, c_1, c_2) \Leftrightarrow c_1, c_2 < N \ \wedge \ \vec{q} \in [f]_{c_1}^{c_2}.$$

To show the correctness property  $L_{\text{rb}} \vdash_R M_{\text{bq}} : L_{\text{bq}}$ , we then need to prove that:

$$L_{\text{bq}} \leq_R M[L_{\text{rb}}].$$

To this end, consider the related states  $\vec{q} R (f, c_1, c_2)$ . We know in particular that  $c_1, c_2 < N$ . Hence, from the calculation carried out in Example 5.4:

$$M_{\text{bq}}.\text{enq}(v)[L_{\text{rb}}](f, c_1, c_2) = \{*\text{@}(f[c_2 := v], c_1, (c_2 + 1) \bmod N)\}$$

$$M_{\text{bq}}.\text{deq}[L_{\text{rb}}](f, c_1, c_2) = \{f_{c_1}\text{@}(f, (c_1 + 1) \bmod N, c_2)\}$$

Since the operations are specified as:

$$L_{\text{bq}}.\text{enq}(v)\text{@}\vec{q} = \{*\text{@}\vec{q}v \mid |\vec{q}| < N\}$$

$$L_{\text{bq}}.\text{deq}\text{@}\vec{q} = \{v\text{@}\vec{p} \mid \vec{q} = v\vec{p}\},$$

we need to show that whenever  $|\vec{q}| < N$ , the following relations hold:

$$* = * \wedge \vec{q}v R (f[c_2 := v], c_1, (c_2 + 1) \bmod N),$$

and that whenever  $\vec{q} = v\vec{p}$ , the following relations hold:

$$v = f_{c_1} \wedge \vec{p} R (f, (c_1 + 1) \bmod N, c_2).$$

This is straightforward to check, using the constraints on the length of the queue to verify that:

$$\begin{aligned} \vec{q} \in [f]_{c_1}^{c_2} &\Rightarrow \vec{q}v \in [f[c_2 := v]]_{c_1}^{(c_2+1) \bmod N} \\ v\vec{p} \in [f]_{c_1}^{c_2} &\Rightarrow \vec{p} \in [f]_{(c_1+1) \bmod N}^{c_2}. \end{aligned}$$

**Remark 5.8.** Note that we place no restriction on the relations we use beyond the simulation property. In particular, in the example above  $R$  is neither one-to-many nor many-to-one.

On one hand, there are many low-level ring buffer states corresponding to a given high-level queue state: the implementation is free to choose the position  $c_1$  where the queue's first element is stored, and arbitrary values for the  $N - |\vec{q}|$  array locations which are not in use.

On the other hand, a low-level state where  $c_1 = c_2$  can be interpreted in two different ways, as the empty queue  $\epsilon$  or as the full queue  $f_{c_1} \dots f_{N-1} f_0 \dots f_{c_2-1}$ . This would make it impossible



to implement an operation specified as  $L_{\text{bq}}.\text{len}@q := \{|\vec{q}|@q\}$ : the implementation would have to simultaneously return 0 and  $N$  when  $c_1 = c_2$ . However, this is not an issue for  $\text{enq}$  and  $\text{deq}$  because both of their specifications are undefined for one of the interpretations:

$$L_{\text{bq}}.\text{enq}(v)@f_{c_1} \dots f_{N-1}f_0 \dots f_{c_2-1} = \emptyset \quad L_{\text{bq}}.\text{deq}@e = \emptyset.$$

In both cases, the implementation can satisfy the specification under both interpretations, because under one of them the specification does not actually place any constraints on the implementation. This allows  $M_{\text{bq}}.\text{enq}$  to be correct by using the “empty” interpretation, and  $M_{\text{bq}}.\text{deq}$  to be correct by using the “full” interpretation.

Once we have shown an abstraction layer correct, we will be able to reason about any client code in terms of the layer’s overlay interface rather than its low-level implementation. This is enabled by the following property.

**Theorem 5.9** (Soundness). *For a certified abstraction layer  $L_1 \vdash_R M : L_2$ , the following contextual refinement property holds:*

$$\forall C \in \mathcal{F}_{E_2}(A) \cdot C[L_2] \leq_R (C \circ M)[L_1]$$

*Proof.* By the monotonicity of substitution and the fact that  $(C \circ M)[L_1] = C[M \circ L_1]$ . □

### 5.2.6 Composing certified abstraction layers

Certified abstraction layers presented in this way can be vertically composed as expected:

$$\frac{L_1 \vdash_R N : L_2 \quad L_2 \vdash_S M : L_3}{L_1 \vdash_{R \circ S} M[N] : L_3}$$

This allows us to decompose a large system into multiple abstraction layers, verify their correctness individually, then derive a correctness property for the whole system.

Note that for every layer interface  $L = \langle E, S, \sigma \rangle$ , we can also derive a trivial certified abstraction layer which leaves everything unchanged:

$$L \vdash_{\mathbf{I}_E} L$$

In other words, we can define a category **CAL** of layer interface and certified layer implementation.

### 5.3 Horizontal composition

While vertical composition is quite useful, individual abstraction layers can themselves be quite complex and involve a significant number of independent operations. To address this, we seek to decompose layers further by introducing *horizontal* composition principles.

#### 5.3.1 Layer implementations

As a starting point, we note that the category of effect signatures and layer implementations is equipped with a cartesian structure. For a simple signature  $F = \{m : A \rightarrow B\}$ , we write:

$$(m[a] \mapsto x) : E \rightarrow F,$$

with  $a \in A$  bound in  $x \in \mathcal{F}_E(B)$ , for a correspondingly simple layer implementation of the signature  $F$  in terms of  $E$ . Cartesian tupling can be used to construct more complex layer implementation from the simple ones.

**Example 5.10** (Decomposition of  $M_{\text{bq}}$ ). *The signature  $E_{\text{bq}}$  and layer implementation  $M_{\text{bq}}$  can be decomposed into the elementary components:*

$$\begin{aligned} E_{\text{enq}} &:= \{\text{enq} : V \rightarrow \mathbb{1}\} & M_{\text{enq}} : E_{\text{rb}} \rightarrow E_{\text{enq}} &:= \text{enq}[v] \mapsto i \leftarrow \text{inc}_2 ; \text{set}[i, v] \\ E_{\text{deq}} &:= \{\text{deq} : \mathbb{1} \rightarrow V\} & M_{\text{deq}} : E_{\text{rb}} \rightarrow E_{\text{deq}} &:= \text{deq}[*] \mapsto i \leftarrow \text{inc}_1 ; \text{get}[i] \end{aligned}$$

*Then:*

$$\begin{aligned} E_{\text{bq}} &= E_{\text{enq}} \otimes E_{\text{deq}} \\ M_{\text{bq}} &= \langle M_{\text{enq}}, M_{\text{deq}} \rangle. \end{aligned}$$

### 5.3.2 Layer interfaces sharing state

When two layer interfaces  $L_1 = \langle E_1, S, \sigma_1 \rangle$  and  $L_2 = \langle E_2, S, \sigma_2 \rangle$  share a common set of abstract states, the specifications  $\sigma_1$  and  $\sigma_2$  will have compatible types, and the product of layer interfaces can be defined as:

$$L_1 \times L_2 := \langle E_1 \otimes E_2, S, (\sigma_1, \sigma_2) \rangle, \quad \text{where } (\sigma_1, \sigma_2)^{i(m)} := \sigma_i^m.$$

Layer correctness properties can be decomposed along the same lines:

$$\frac{L \vdash_R M_1 : L_1 \quad L \vdash_R M_2 : L_2}{L \vdash_R \langle M_1, M_2 \rangle : L_1 \times L_2}$$

Conversely, the cartesian projections:

$$L_1 \times L_2 \vdash = \pi_1 : L_1 \quad L_1 \times L_2 \vdash = \pi_2 : L_2$$

allow us retain only part of a layer's interface.

**Example 5.11** (Decomposition of  $L_{\text{bq}}$ ). *The specification and verification of our bounded queue implementation can be decomposed in the following way:*

$$\frac{L_{\text{rb}} \vdash_{R_{\text{bq}}} M_{\text{enq}} : L_{\text{enq}} \quad L_{\text{rb}} \vdash_{R_{\text{bq}}} M_{\text{deq}} : L_{\text{deq}}}{L_{\text{rb}} \vdash M_{\text{bq}} : L_{\text{bq}}}$$

where:

$$\begin{aligned} L_{\text{enq}} &:= \langle E_{\text{enq}}, S_{\text{bq}}, (\text{enq}[v] \mapsto L_{\text{bq}}.\text{enq}[v]) \rangle \\ L_{\text{deq}} &:= \langle E_{\text{deq}}, S_{\text{bq}}, (\text{deq}[v] \mapsto L_{\text{bq}}.\text{deq}[v]) \rangle \end{aligned}$$

The various forms of composition I have introduced so far are summarized in Fig. [XXX: building blocks figure from slides]. Together, they describe the principles of the layer calculus we used to verify CertiKOS. However, the form of horizontal composition we have introduced so far is largely syntactic in nature. Crucially, they rely on the fact that within an abstraction layer, components share the same low- and high-level state, and correctness proof share the same correspondence

between them. In particular, note that the diagonal morphism:

$$L \vdash_{=} \Delta_L : L \times L$$

does not duplicate any of the resources implemented by  $L$ . That is, in the layer:

$$\frac{L_{rb} \vdash_{R_{bq}} M_{bq} : L_{bq} \quad L_{rb} \vdash_{R_{bq}} M_{bq} : L_{bq}}{L_{rb} \vdash_{R_{bq}} \langle M_{bq}, M_{bq} \rangle : L_{bq} \times L_{bq}}$$

we have simply duplicated the code for each of the queue operations. But each copy uses the same underlying ring buffer, and the layer still implements a single queue.

### 5.3.3 Monoidal structure

[XXX motivation: text+figure from prospectus]

In this section, we present a more general form of horizontal composition which lets two layers operate independently of one another. The layers do not need to share state.

**Definition 5.12.** The *tensor product* of  $L_1 = \langle E_1, S_1, \sigma_1 \rangle$  and  $L_2 = \langle E_2, S_2, \sigma_2 \rangle$  is defined as:

$$L_1 \otimes L_2 := \langle E_1 \otimes E_2, S_1 \otimes S_2, (\hat{\sigma}_1, \hat{\sigma}_2) \rangle,$$

where  $\hat{\sigma}_1$  and  $\hat{\sigma}_2$  act on their respective halves of the state:

$$\begin{aligned} \hat{\sigma}_1^{m@ (k_1, k_2)} &:= \{v@ (k'_1, k_2) \mid \sigma_1^{m@ k_1} \ni v@ k'_1\} \\ \hat{\sigma}_2^{m@ (k_1, k_2)} &:= \{v@ (k_1, k'_2) \mid \sigma_2^{m@ k_2} \ni v@ k'_2\} \end{aligned}$$

Note that the corresponding structure is not cartesian. In particular, two certified abstraction layers  $L \vdash_{R_1} M_1 : L_1$  and  $L \vdash_{R_2} M_2 : L_2$  cannot be composed into a single one in this way:

$$L \vdash_{\langle R_1, R_2 \rangle} \langle M_1, M_2 \rangle : L_1 \otimes L_2$$

because the operations of one layer will update the underlay state in a way that may break the simulation relation for the other layer. However, layers can be composed side-by-side when the

act on two independent underlay interfaces.

**Theorem 5.13** (Tensor product of certified abstraction layers). *Two certified abstraction layers can be composed as follows:*

$$\frac{L_1 \vdash_{R_1} M_1 : L'_1 \quad L_2 \vdash_{R_2} M_2 : L'_2}{L_1 \otimes L_2 \vdash_{R_1 \times R_2} M_1 \times M_2 : L'_1 \otimes L'_2}$$

Then, when doing so makes sense, we can model a situation where two independent layers access a common underlay  $L$ , by showing that the resources of  $L$  can be multiplexed. This can be formalized by introducing a third certified abstraction layer  $L \vdash_R M : L \otimes L$ .

For example, suppose the layer interface  $L^n$  describes a memory allocator maintaining a pool of  $n$  pages. To share this interface between multiple clients, it may be possible to split this pool, so that:

$$L^{n_1+n_2} \vdash_R \Delta : L^{n_1} \otimes L^{n_2}.$$

The allocator is used by

[XXX example of a memory allocator maintaining a pool of  $n$  pages, we we can split as Problem of modeling the nondeterminism of arbitrary page numbers being returned in the context of the limited monad  $\mathcal{L}$ ]

## 5.4 CompCertX

Our original definition of certified abstraction layers was tied up with the semantic model of CompCert [Gu et al., 2015]. In particular, layer specifications were defined to operate on a CompCert memory state component as well as the layer-specific abstract state. The specifications in more abstract layer interfaces would in fact leave this component unchanged, acting instead on the abstract state, but intermediate layer interfaces which could modify both components of the state would be used to define the semantics of layer implementations and define corresponding low-level specifications.

The streamlined version presented in this chapter can however be used to recover and present the original theory.

### 5.4.1 CompCertX

CompCertX is a version of CompCert designed to support certified abstraction layers.

In CompCert, language semantics use a parameter  $\chi$  common to all languages, which describes the behavior of external functions. Schematically,

$$\text{Clight}_\chi[p]$$

describes the semantics of the Clight program  $p$ , where  $\chi^{f[sg](\vec{v})} : \text{mem} \times \mathcal{P}^1(\text{val} \times \text{mem})$  describes the behavior of the external function  $f$  with signature  $sg$  called by  $p$  with arguments  $\vec{v}$ . We extend this facility to full-blown layer interfaces. Given a layer interface  $L = \langle E, S \times \text{mem}, \sigma \rangle$ , where  $E \subseteq \mathcal{C}$ , the semantics:

$$\text{Clight}_L[p]$$

will maintain and abstract state  $s \in S$  alongside the program's memory state, and use the specification  $\sigma$  to interpret external calls as before.

To make it possible to evaluate a CompCert modules as layer implementations (§??), CompCertX also extends the *incoming* interface of language semantics. Instead of only modelling the initial invocation of main, we make it possible to evaluate the component for any incoming call. Then:

$$\text{Clight}_L[M]^{f[sg](\vec{v})} : S \times \text{mem} \rightarrow N \times S \times \text{mem}$$

models the execution of the function  $f$  in the module  $M$ , so that the evaluation of  $M$  on the underlay  $L$  can be defined as:

$$M[L] := \langle E_M, S \times \text{mem}, \text{Clight}_L[M] \rangle.$$

### 5.4.2 Extending layer interfaces with memory states

The layer interfaces we have seen so far have been expressed in purely abstract terms. To use them as specifications for CompCert modules, we can adjoin a memory state to them in the following way.

**Definition 5.14** (Layer interface extension). A layer interface  $L = \langle E, S, \sigma \rangle$  can be promoted to

the layer interface:

$$L@T := \langle E, S \times T, \hat{\sigma} \rangle,$$

where the state has an additional component of type  $S'$  which is left unchanged by the specifications:

$$\hat{\sigma}^m(s, t) := \{(s', t) \mid \sigma^m(s) \ni s'\}.$$

[XXX give an overview of how this translates when we mix with CompCert, and the implementation of the layer library, which was also the inspiration behind CompCert Kripke logical relations.]

## 5.5 Conclusion

The new presentation of certified abstraction layers laid out in this chapter eliminates some of the complexity found in the original code by decoupling the model from CompCert semantics. In doing so, the new exposition makes precise the underlying mathematical structures and some of the challenges involved in our approach to layered verification.

In particular, the categorical structures behind our layer calculus are made explicit. The distinctions between our restricted form of product and the more general monoidal structure hints at the connections between our framework and linear logic, where cartesian and tensor products cohabitate and work together.

## Chapter 6

# The interaction specification monad

In this chapter, I define the *interaction specification monad*, a variant of the free monad on an effect signature which incorporates dual nondeterminism and refinement. The result can be used to formulate a theory of certified abstraction layers in which layer interfaces, layer implementations, and simulation relations are treated uniformly and compositionally.

### 6.1 Overview

Given an effect signature  $E$ , we construct a prefix-ordered set of plays  $\bar{P}_E(A)$  corresponding to the possible interactions between a computation with effects in  $E$  and its environment, including the computation's ultimate outcome in  $A$ . The interaction specification monad  $\mathcal{I}_E(A)$  is then obtained as the free completely distributive completion of the poset  $\bar{P}_E(A)$ .

For each effect  $e \in E$ , the interaction specification monad has an operation  $\mathbf{I}_E^e \in \mathcal{I}_E(\text{ar}(e))$  which triggers an instance of  $e$  and returns its outcome. Given a second effect signature  $F$ , a family  $(f^m)_{m \in F}$  of computations  $f^m \in \mathcal{I}_E(\text{ar}(m))$  can be used to interpret the effects of  $F$  into the signature  $E$ . This is achieved by a *substitution* operator  $\bullet[f]$ , which transforms a computation  $x \in \mathcal{I}_F(A)$  into the computation  $x[f] \in \mathcal{I}_E(A)$ , where each occurrence of an effect  $m \in F$  in  $x$  is replaced by the corresponding computation  $f^m$ .

Effect signatures are used as simple games, and a family  $(f^m)_{m \in F}$  as described above can be interpreted as a certain kind of strategy for the game  $!E \multimap F$ . We use this approach to define a first category of games and strategy specifications  $\mathcal{G}_{\sqsubseteq}^{ib}$ , where  $(\mathbf{I}_E^e)_{e \in E}$  is the identity morphism



for  $E$  and the substitution operator is used to define composition.

Note that rather than providing denotational semantics for specific programming languages, our models are intended as a coarse-grained composition “glue” between components developed and verified in their own languages, each equipped with their own internal semantics. In this context, the models’ restriction to first-order computation applies only to cross-component interactions, but individual components can still make use of high-order languages and reasoning techniques [Mansky et al., 2020].

## 6.2 Plays

We first introduce the partially ordered sets of plays which we use to construct the interaction specification monad. Since we intend to describe *active* computations, we use *odd*-length plays which start with *system* moves, by contrast with the more common approach presented in §??.

**Definition 6.1.** The set  $\bar{P}_E(A)$  of *interactions* for an effect signature  $E$  and a set of values  $A$  is defined inductively:

$$s \in \bar{P}_E(A) ::= \underline{v} \mid \underline{m} \mid \underline{mns},$$

where  $v \in A$ ,  $m \in E$  and  $n \in \text{ar}(m)$ . The set  $\bar{P}_E(A)$  is ordered by the prefix relation  $\sqsubseteq \subseteq \bar{P}_E(A) \times \bar{P}_E(A)$ , defined as the smallest relation satisfying:

$$\underline{v} \sqsubseteq \underline{v}, \quad \underline{m} \sqsubseteq \underline{m}, \quad \underline{m} \sqsubseteq \underline{mnt}, \quad \frac{s \sqsubseteq t}{\underline{mns} \sqsubseteq \underline{mnt}}.$$

A play corresponds to a finite observation of an interaction between the system and the environment. At any point in such an interaction, the system can terminate the interaction with a given value ( $v$ ), or it can trigger an effect  $m \in E$  and wait to be resumed by an answer  $n \in \text{ar}(m)$  of the environment ( $\underline{mns}$ ). A play which concludes before the environment answers a query from the system ( $\underline{m}$ ) denotes that no information has been observed after that point. It can be refined by a longer observation of an interaction which begin with the same sequence of questions and answers.

### 6.3 Interaction specifications

We define our monad as the free completely distributive completion of the corresponding poset of plays.

For the sake of conciseness and clarity, we will use the order embedding associated with **FCD** implicitly, so that an element of a poset  $s \in P$  can also be regarded as an element of its completion  $s \in \mathbf{FCD}(P)$ . Likewise, for a completely distributive lattice  $M$ , we can implicitly promote a monotonic function  $f : P \rightarrow M$  to its extension  $f : \mathbf{FCD}(P) \rightarrow M$ . These conventions are at work in the following definition.

**Definition 6.2.** The *interaction specification monad* for an effect signature  $E$  maps a set  $A$  to the free completely distributive completion of the corresponding poset of plays:

$$\mathcal{I}_E(A) := \mathbf{FCD}(\bar{P}_E(A))$$

An element  $x \in \mathcal{I}_E(A)$  is called an *interaction specification*.

The monad's action on a function  $f : A \rightarrow B$  replaces the values in an interaction specification with their image by  $f$ :

$$\mathcal{I}_E(f)(\underline{v}) := \underline{f(v)}$$

$$\mathcal{I}_E(f)(\underline{m}) := \underline{m}$$

$$\mathcal{I}_E(f)(\underline{m}ns) := \underline{m}n\mathcal{I}_E(f)(s).$$

The monad's unit  $\eta_A^E : A \rightarrow \mathcal{I}_E(A)$  is the embedding of a single play consisting only of the given value:

$$\eta_A^E(v) := \underline{v}$$

Finally, the multiplication  $\mu_A^E : \mathcal{I}_E(\mathcal{I}_E(A)) \rightarrow \mathcal{I}_E(A)$  carries out the outer computation and

sequences it with any computation it evaluates to:

$$\begin{aligned}\mu_A^E(\underline{x}) &:= x \\ \mu_A^E(\underline{m}) &:= \underline{m} \\ \mu_A^E(\underline{mns}) &:= \underline{m} \sqcup \underline{mn}\mu_A^E(s).\end{aligned}$$

The most subtle aspect of Def. 6.2 is the case for  $\mu_A^E(\underline{mns})$ , which includes  $\underline{m}$  as well as  $\underline{mn}\mu_A^E(s)$ . This is both to ensure that the effects of the first computation are preserved when the second computation is  $\perp$ , and to ensure the monotonicity of the underlying function used to define  $\mu_A^E$ . Consider for example  $\underline{m} \sqsubseteq \underline{mn}\perp \in \bar{P}_E(\mathcal{I}_E(A))$ . Since  $\mu_A^E(\perp) = \perp$  and the **FCD** extension of the function  $s \mapsto \underline{mns}$  preserves  $\perp$ , it is not the case that  $\underline{m} \sqsubseteq \underline{mn}\mu_A^E(\perp)$ .

As usual, the Kleisli extension of a function  $f : A \rightarrow \mathcal{I}_E(B)$  is the function  $f^* = \mu_B^E \circ \mathcal{I}_E(f)$ . We extend the notations used for **FCD** to the monad  $\mathcal{I}_E$ .

## 6.4 Interaction primitives

The operations of an effect signature  $E$  can be promoted to interaction specifications of  $\mathcal{I}_E$  as follows.

**Definition 6.3** (Interaction primitive). For an effect signature  $E$  and an operation  $m \in E$ , the interaction specification  $\mathbf{I}_E^m \in \mathcal{I}_E(\text{ar}(m))$  is defined as:

$$\mathbf{I}_E^m := \bigsqcup_{n \in \text{ar}(m)} \underline{mnn}$$

Note that in the play  $\underline{mnn}$ , the first occurrence of  $n$  is the environment's answer, whereas the second occurrence is the value returned by  $\mathbf{I}_E^m$ .

To model effect handling for a signature  $F$ , we use a family of interaction specifications  $(f^m)_{m \in F}$  to provide an interpretation  $f^m \in \mathcal{I}_E(\text{ar}(m))$  of each effect  $m \in F$  in terms of another effect signature  $E$ . This allows us to transform an interaction specification  $x \in \mathcal{I}_F(A)$  into an interaction specification  $x[f] \in \mathcal{I}_E(A)$ , defined as follows. The constructions  $\perp$  and  $\{P\}$  were discussed in §2.2; they carry similar meanings in the context of the interaction specification

monad.

**Definition 6.4** (Interaction substitution). Given the effect signatures  $E, F$  and the set  $A$ , for an interaction specification  $x \in \mathcal{I}_F(A)$  and a family  $(f^m)_{m \in F}$  with  $f^m \in \mathcal{I}_E(\text{ar}(m))$ , the *interaction substitution*  $x[f] \in \mathcal{I}_E(A)$  is defined by:

$$\begin{aligned} \underline{v}[f] &:= \underline{v} \\ \underline{m}[f] &:= r \leftarrow f^m; \perp \\ \underline{mns}[f] &:= r \leftarrow f^m; \{r = n\}; s[f]. \end{aligned}$$

The outcome of the interaction specification is left unchanged, but effects are replaced by their interpretation. Whenever that interpretation produces an outcome  $r$ , the substitution process resumes with the remainder of any matching plays of the original computation.

## 6.5 Categorical structure

As presented so far, the interaction specification monad can be seen as an extension of the refinement calculus able to model effectful computations for a given signature. We now shift our point of view to game semantics and show how interaction substitutions can be used to define a simple category of games and strategies featuring dual nondeterminism and alternating refinement.

**Definition 6.5** (Morphisms). Consider the effect signatures  $E, F$  and  $G$ . We will write  $f : E \rightarrow F$  whenever  $(f^m)_{m \in F}$  is a family of interactive computations such that  $f^m \in \mathcal{I}_E(\text{ar}(m))$ . For  $f : E \rightarrow F$  and  $g : F \rightarrow G$ , we define  $g \circ f : E \rightarrow G$  as:

$$(g \circ f)^m = g^m[f].$$

The completely distributive lattice structure of  $\mathcal{I}_F(-)$  can be extended pointwise to morphisms, so that for a family  $(f_i)_{i \in I}$  with  $f_i : E \rightarrow F$ , we can define  $\bigsqcup_{i \in I} f_i : E \rightarrow F$  and  $\prod_{i \in I} f_i : E \rightarrow F$ . For  $f, g : E \rightarrow F$  we define refinement as:

$$f \sqsubseteq g \Leftrightarrow \forall m \in F. f^m \sqsubseteq g^m.$$

A morphism  $f : E \rightarrow F$  can be interpreted as a *well-bracketed* strategy for the game  $!E \multimap F$ . In this game, the environment first plays a move  $m \in F$ . The system can then ask a series of questions  $q_1, \dots, q_k \in E$  to which the environment will reply with answers  $r_i \in \text{ar}(q_i)$ , and finally produce an answer  $n \in \text{ar}(m)$  to the environment's initial question  $m$ . The plays of  $!E \multimap F$  are restricted to a single top-level question  $m$ . In addition, the well-bracketing requirement imposes that at any point, only the most recent pending question may be answered.

Compared with the usual notion of strategy, our model introduces arbitrary demonic choices and relaxes constraints over angelic choices. The definition of  $g \circ f$  given above otherwise corresponds to the traditional definition of strategy composition. The identity strategy is given by  $\mathbf{I}_E : E \rightarrow E$ .

**Lemma 6.6.** *Consider the effect signatures  $E, F, G, H$  and the morphisms  $f : E \rightarrow F, g : F \rightarrow G$  and  $h : G \rightarrow H$ . The following properties hold:*

$$\begin{aligned}\mathbf{I}_F \circ f &= f \circ \mathbf{I}_E = f \\ h \circ (g \circ f) &= (h \circ g) \circ f\end{aligned}$$

*Composition preserves all extrema on the left, and all non-empty extrema on the right.*

*Proof.* Using properties of **FCD** and inductions on plays. □

Having established the relevant properties, we can now define our first category of games and strategies.

**Definition 6.7.** The category  $\mathcal{G}_{\perp}^{ib}$  has effect signatures as objects. Morphisms, identities and composition have been defined above. The hom-sets  $\mathcal{G}_{\perp}^{ib}(E, F)$  are completely distributive lattices, with composition preserving all extrema on the left, and all non-empty extrema on the right.

## 6.6 Products

Effect signatures can be combined in the following way.

**Definition 6.8.** We define the effect signature  $1 := \emptyset$ . For a family of effect signatures  $(E_i)_{i \in I}$ ,

we define:

$$\bigotimes_i E_i := \{(i, e) : \text{ar}(e) \mid i \in I, e \in E_i\}$$

For example, the signature  $E_{\text{io}}$  above is equivalent to the following composite one:

$$\{\text{readbit} : 2\} \otimes \{\text{print}[s] : 1 \mid s \in \text{string}\} \otimes \{\text{done} : \emptyset\}$$

The construction  $\otimes$  gives products in the category  $\mathcal{G}_{\sqsubseteq}^{ib}$ , as demonstrated below.

**Theorem 6.9.** *The category  $\mathcal{G}_{\sqsubseteq}^{ib}$  has all products. Objects are given by  $\bigotimes_{i \in I} E_i$  and projection arrows are given for each  $i \in I$  by the morphism:*

$$\pi_i : \bigotimes_{j \in I} E_j \rightarrow E_i \quad \pi_i^m := (i, m).$$

*Proof.* We need to show that for an effect signature  $X$  and a collections of morphisms  $(f_i)_{i \in I}$  with  $f_i : X \rightarrow E_i$ , there is a unique  $\langle f_i \rangle_{i \in I} : X \rightarrow \bigotimes_{i \in I} E_i$  such that for all  $i \in I$ :

$$f_i = \pi_i \circ \langle f_j \rangle_{j \in I}.$$

Note that for  $x : X \rightarrow \bigotimes_{i \in I} E_i$ ,  $i \in I$  and  $m \in E_i$ , we have:

$$(\pi_i \circ x)^m = \pi_i^m[x] = (i, m)[x] = x^{(i, m)}$$

Hence,  $\langle f_i \rangle_{i \in I}$  is uniquely defined as:

$$\langle f_i \rangle_{i \in I}^{(j, m)} := f_j^m.$$

□

## 6.7 Certified abstraction layers

Certified abstraction layers can be embedded into the category  $\mathcal{G}_{\sqsubseteq}^{ib}$  as follows.

### 6.7.1 Signatures

The signature of a layer interface or implementation can be encoded as an effect signature. For example:

$$E_{\text{bq}} := \{\text{enq}[v] : \mathbb{1}, \text{deq} : V \mid v \in V\}$$

$$E_{\text{rb}} := \{\text{set}[i, v] : \mathbb{1}, \text{get}[i] : V, \text{inc}_1 : \mathbb{N}, \text{inc}_2 : \mathbb{N} \mid i \in \mathbb{N}, v \in V\}$$

A layer implementation  $M$  with an underlay signature  $E$  and an overlay signature  $F$  can then be interpreted as a morphism  $\llbracket M \rrbracket : E \rightarrow F$  in a straightforward manner, by replacing underlay operations used in the definition of  $M$  with the corresponding interaction primitives:

$$\llbracket M \rrbracket^m := (M.m)[e := \mathbf{I}_E^e]_{e \in E}$$

### 6.7.2 Interfaces

In order to handle the layer's abstract data, we can extend signatures with state in the following way:

$$E @ S := \{m @ k : \text{ar}(m) \times S \mid m \in E, k \in S\}$$

A layer interface  $L$  with a signature  $E$  and states in  $S$  can be interpreted as a morphism  $\llbracket L \rrbracket : 1 \rightarrow E @ S$  almost directly, mapping  $\emptyset$  to  $\perp$  in outcomes of primitive specifications:

$$\llbracket L \rrbracket^{m @ k} := \bigsqcup L.m @ k$$

### 6.7.3 Keeping state

For a morphism  $f : E \rightarrow F$ , we construct  $f @ S : E @ S \rightarrow F @ S$  which keeps updating a state  $k \in S$  as it performs effects in  $E @ S$ , then adjoins the final state to any answer returned by  $f$ . For

a set  $A$ , we first define  $- \# - : \bar{P}_E(A) \times S \rightarrow \mathcal{I}_{E@S}(A \times S)$ :

$$\begin{aligned} \underline{v} \# k &:= \underline{v@k} \\ \underline{m} \# k &:= \underline{m@k} \\ \underline{mns} \# k &:= \bigsqcup_{k' \in S} \underline{m@k} \ n@k' \ s \# k', \end{aligned}$$

and extend it to morphisms as  $(f@S)^{m@k} := f^m \# k$ . Then in particular, running a layer implementation  $\llbracket M \rrbracket : E \rightarrow F$  on top of a layer interface  $\llbracket L \rrbracket : 1 \rightarrow E@S$  yields the morphism  $\llbracket M \rrbracket @S \circ \llbracket L \rrbracket : 1 \rightarrow F@S$ .

#### 6.7.4 Simulation relations

The most interesting aspect of our embedding is the representation of simulation relations. We will see that dual nondeterminism allows us to represent them as regular morphisms.

Recall the definition of the judgment  $L_1 \vdash_R M : L_2$ , which means that a layer implementation  $M$  correctly implements  $L_2$  on top of  $L_1$  through a simulation relation  $R \subseteq S_2 \times S_1$ . If we write  $L'_1 := \llbracket M \rrbracket @S_1 \circ \llbracket L_1 \rrbracket$  for the layer interface obtained by interpreting  $M$  on top of  $L_1$ , then:

$$L_1 \vdash_R M : L_2 \Leftrightarrow \forall m \in E_2 \cdot L_2^m [R \rightarrow \mathcal{P}^{\leq}(= \times R)] L_1'^m$$

We will use the families of morphisms  $R_E^* : E@S_2 \rightarrow E@S_1$  and  $R_*^E : E@S_1 \rightarrow E@S_2$  to encode this judgment:

$$\begin{aligned} (R_E^*)^{m@k_1} &:= \bigsqcup_{k_2 \in R^{-1}(k_1)} n@k_2' \leftarrow \mathbf{I}_{E@S_2}^{m@k_2}; \quad \prod_{k_1' \in R(k_2')} n@k_1' \\ (R_*^E)^{m@k_2} &:= \prod_{k_1 \in R(k_2)} n@k_1' \leftarrow \mathbf{I}_{E@S_1}^{m@k_1}; \quad \bigsqcup_{k_2' \in R^{-1}(k_1')} n@k_2' \end{aligned}$$

They yield two equivalent ways to encode layer correctness as refinement properties.

In the first case,  $R_E^*$  is intended to translate a high-level specification  $\sigma$  which uses overlay states  $k_2, k_2' \in S_2$  into a low-level specification  $R_E^* \circ \sigma$  which uses underlay states  $k_1, k_1' \in S_1$ . The client calls  $R_E^*$  with an underlay state  $k_1$ , with the expectation that if there is *any* corresponding



overlay state, then  $R_E^* \circ \sigma$  will behave accordingly (it is angelic with respect to its choice of  $k_2$ ). On the other hand,  $R_E^* \circ \sigma$  is free to choose any underlay representation  $k'_1$  for the outcome  $k'_2$  produced by  $\sigma$ , and the client must be ready to accept it (it is demonic with respect to its choice of  $k'_1$ ).

In the second case,  $R_*^E \circ \tau$  is the strongest high-level specification which a low-level component  $\tau$  implements with respect to  $R$ . For an overlay state  $k_2 \in S_2$ ,  $\tau$  may behave in various ways depending on the corresponding underlay state  $k_1 \in S_1$  it is invoked with, and so the specification must allow them using demonic choice. On the other hand, when  $\tau$  returns with a new underlay state  $k'_1$ , the environment is free to choose how to interpret it as an overlay state  $k'_2$ .

**Theorem 6.10.** For  $\sigma := \llbracket L_2 \rrbracket$  and  $\tau := \llbracket M \rrbracket @ S_1 \circ \llbracket L_1 \rrbracket$ :

$$R_{E_2}^* \circ \sigma \sqsubseteq \tau \Leftrightarrow L_1 \vdash_R M : L_2 \Leftrightarrow \sigma \sqsubseteq R_*^{E_2} \circ \tau.$$

*Proof.* The proof is straightforward but requires Thm. 5.3 from [Morris and Tyrrell \[2008\]](#).  $\square$

## Chapter 7

# Stateful and reentrant strategies

The model presented in Chapter 6 is designed to be simple but general enough to embed CompCert semantics, certified abstraction layers, and interaction trees. I now sketch a more general model which allows strategies to retain state across successive activations. We explain how the new model  $\mathcal{G}_{\sqsubseteq}^b$  can embed the morphisms of  $\mathcal{G}_{\sqsubseteq}^{ib}$ , and how it can be used to characterize certified abstraction layers independently of the states used in their description.

### 7.1 Overview

As discussed in §6.5, the morphisms of  $\mathcal{G}_{\sqsubseteq}^{ib}(E, F)$  correspond to the well-bracketed strategies for the game  $!E \multimap F$ . As such, they can be promoted to well-bracketed strategies for the more general game  $!E \multimap !F$ , which allows the environment to ask multiple question of  $F$  in a row, and to ask nested questions whenever it is in control.

More precisely, strategies promoted in this way correspond to the *innocent* well-bracketed strategies for  $!E \multimap !F$ , meaning that they will behave in the same way in response to the same question, regardless of the history of the computation. The model we introduce in this section relaxes this constraint, allowing strategies to maintain internal state.

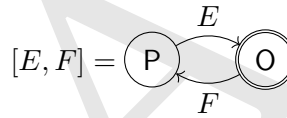
After outlining the construction of a new category  $\mathcal{G}_{\sqsubseteq}^b$  of games and strategies (§7.2–§7.5), we define an embedding of  $\mathcal{G}_{\sqsubseteq}^{ib}$  into  $\mathcal{G}_{\sqsubseteq}^b$  (§7.6), and show how the states used by a strategy  $\sigma : E @ S \rightarrow F @ S$  can be internalized and hidden from its interactions (§7.7).

## 7.2 Games

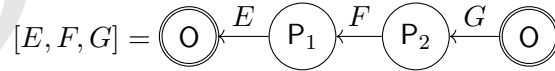
To facilitate reasoning, and make it easier to describe operators on strategies in a systematic way, we describe games as a specific kind of graph where vertices represent players and edges determine which questions can be asked by one player to another. Generalizing from effect signatures, questions are assigned an arity which gives the type of the answer.

**Definition 7.1.** A *game signature*  $\Gamma$  is a set of players with a distinguished element  $O$ , together with an effect signature  $\Gamma(u, v)$  for all  $u, v \in \Gamma$ . The operations  $m \in \Gamma(u, v)$  are called the *questions* of  $u$  to  $v$ , and the elements  $n \in \text{ar}(m)$  are called *answers* to the question  $m$ .

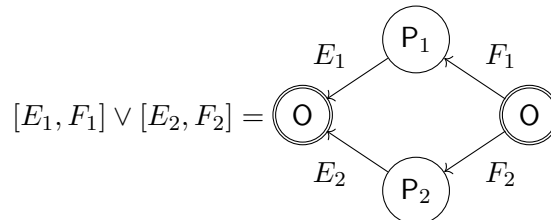
We depict game signatures as directed graphs whose vertices are the players and whose edges are labeled by the corresponding effect signature. Missing edges correspond to the empty signature  $\emptyset$ . For example, the game  $!E \multimap !F$  is generated by the game signature:



When we consider the ways in which questions propagate through a game signature, the distinguished player  $O$  serves the role of both a source and sink. As such, it is visually useful to depict  $O$  as two nodes, one capturing the incoming edges of  $O$ , and one capturing its outgoing edges. For example, the following game signature generates the interaction sequences used in the definition of strategy composition:



As another example of a game signature, a situation where  $\sigma_1 : E_1 \rightarrow F_1$  and  $\sigma_2 : E_2 \rightarrow F_2$  interact with the environment independently of one another can be described as:



The signature above will be used to compute tensor products of strategies. These constructions generalize as follows.

**Definition 7.2** (Constructions on game signatures). For a collection of effect signature  $(E_i)_{1 \leq i \leq n}$  and an effect signature  $F$ , the game signature  $[E_1, \dots, E_n, F]$  has the players  $O, P_1, \dots, P_n$  and the following edges:

$$[E_1, \dots, E_n, F] := \textcircled{O} \xleftarrow{E_1} P_1 \xleftarrow{E_2} \dots \xleftarrow{E_n} P_n \xleftarrow{F} \textcircled{O}$$

For a collection of game signatures  $(\Gamma_i)_{i \in I}$ , the *wedge sum*  $\bigvee_{i \in I} \Gamma_i$  has the players:

$$\{O\} \cup \{(i, p) \mid i \in I \wedge p \in \Gamma_i \setminus \{O\}\}$$

For  $i \in I$  and  $p \in \Gamma_i$ , the corresponding player in  $\bigvee_{j \in I} \Gamma_j$  is:

$$\iota_i(p) := \begin{cases} O & \text{if } p = O \\ (i, p) & \text{otherwise.} \end{cases}$$

Then for each question  $m : u \rightarrow v$  in  $\Gamma_i$ , the wedge sum has a corresponding question  $\iota_i(m) : \iota_i(u) \rightarrow \iota_i(v)$ .

### 7.3 Plays and strategies

The well-bracketing requirement enforces a kind of *stack discipline* on the succession of questions and answers. A well-bracketed play can be interpreted as an activation tree, where questions are understood as function calls and answers are understood as the corresponding calls returning. At any point in a play over a signature  $\Gamma$ , its possible evolutions are characterized by the stack of pending questions.

**Definition 7.3.** For a game signature  $\Gamma$  and a player  $p \in \Gamma$ , a *p-stack* over  $\Gamma$  is a path:

$$O = p_0 \xrightarrow{m_1} p_1 \xrightarrow{m_2} \dots \xrightarrow{m_n} p_n = p$$

where  $p_i \in \Gamma$  and  $m_i \in \Gamma(p_{i-1}, p_i)$ . We will write this path as  $\kappa = m'_1 \cdots m'_n : \mathbf{O} \twoheadrightarrow p \in \Gamma$ .

Such stacks can in turn be arranged in a graph  $\hat{\Gamma}$  over which the game associated with  $\Gamma$  will be played.

**Definition 7.4** (Strategy specifications). For a signature  $\Gamma$ , the graph  $\hat{\Gamma}$  is defined as follows. The vertices of  $\hat{\Gamma}$  are pairs  $(u, \kappa)$  in which  $u \in \Gamma$  and  $\kappa$  is a  $u$ -stack. For each question  $m \in \Gamma(u, v)$  and stack  $\kappa : \mathbf{O} \twoheadrightarrow u$ , there is an edge:

$$m : (u, \kappa) \rightarrow (v, \kappa m) \in \hat{\Gamma}.$$

In addition, for each answer  $n \in \text{ar}(m)$ , there is an edge:

$$n : (v, \kappa m) \rightarrow (u, \kappa) \in \hat{\Gamma}.$$

The *plays* over  $\Gamma$  are paths of type  $(\mathbf{O}, \epsilon) \twoheadrightarrow (\mathbf{O}, \kappa) \in \hat{\Gamma}$ , where  $\kappa : \mathbf{O} \twoheadrightarrow \mathbf{O} \in \Gamma$  is a stack. We will write  $P\Gamma$  for the poset of plays over  $\Gamma$  under the prefix ordering. The *strategy specifications* for  $\Gamma$  are given by the completion:

$$S\Gamma := \mathbf{FCD}(P\Gamma).$$

## 7.4 Operations on strategies

**Definition 7.5.** A *transformation* from the game signature  $\Gamma_1$  to the game signature  $\Gamma_2$  associates to each player  $p \in \Gamma_1$  a player  $f(p) \in \Gamma_2$  with  $f(\mathbf{O}) = \mathbf{O}$ , and to each question  $m \in \Gamma_1(u, v)$  a *path* of questions in  $\Gamma_2$ :

$$f(u) = p_0 \xrightarrow{m'_1} p_1 \xrightarrow{m'_2} \cdots \xrightarrow{m'_n} p_n = f(v),$$

written as  $f(m) = m'_1 m'_2 \cdots m'_n : f(u) \twoheadrightarrow f(v)$ , and such that  $\text{ar}(m'_1) = \cdots = \text{ar}(m'_n) = \text{ar}(m)$ .

We extend  $f$  itself to the paths in  $\Gamma_1$  by taking the image of  $m_1 \cdots m_n : u \twoheadrightarrow v$  to be:

$$f(m_1 \cdots m_n) := f(m_1) \cdots f(m_n) : f(u) \twoheadrightarrow f(v).$$

Game signatures and transformations form a category.

In other words, a transformation is a structure-preserving map on paths. Transformations can be extended to plays.

**Definition 7.6** (Action on plays). A transformation  $f : \Gamma_1 \rightarrow \Gamma_2$  induces a monotonic function  $Pf : P\Gamma_1 \rightarrow P\Gamma_2$  as follows. For  $m \in \Gamma(u, v)$  and  $\kappa : O \rightarrow u$ , the image of the move  $m : (u, \kappa) \rightarrow (v, \kappa m)$  is the path:

$$f(m) : (f(u), f(\kappa)) \rightarrow (f(v), f(\kappa)f(m)).$$

For  $n \in \text{ar}(m)$ , the image of  $n : (v, \kappa m) \rightarrow (u, \kappa)$  is the path:

$$n^{|f(m)|} : (f(v), f(\kappa)f(m)) \rightarrow (f(u), f(\kappa)),$$

where  $n^{|f(m)|}$  denotes a sequence  $nn \cdots n$  of copies of the answer  $n \in \text{ar}(m)$  of same length as the path  $f(m)$ .

Operators on strategies will generally be defined by a game signature of global *interaction sequences*, and will use transformations to project out the corresponding plays of the arguments and the result.

### 7.4.1 Composition

When composing the strategy specifications  $\sigma \in S[E, F]$  and  $\tau \in S[F, G]$  to obtain  $\tau \circ \sigma \in S[E, G]$ , we will use the transformation  $\psi_X^c : [E, F, G] \rightarrow [E, G]$  to describe the externally observable behavior of interaction sequences in  $[E, F, G]$ :

$$\begin{aligned} \psi_X^c(P_1) &= \psi_X^c(P_2) = P & \psi_X^c(O) &= O \\ \psi_X^c(m) &= \begin{cases} \epsilon & \text{if } m \in F \\ m & \text{otherwise} \end{cases} \end{aligned}$$

This can be described concisely as  $\psi_X^c = [1, 0, 1]$ , with:

$$\psi_1^c := [1, 1, 0] : [E, F, G] \rightarrow [E, F]$$

$$\psi_2^c := [0, 1, 1] : [E, F, G] \rightarrow [F, G]$$

defined similarly.

We can now formulate the composition of strategy specifications as follows. The “footprint” of the plays  $s_1 \in P[E, F]$  and  $s_2 \in P[F, G]$  can be defined as:

$$\psi^c(s_1, s_2) := \bigsqcup_{s \in P[E, F, G]} \{ \psi_1^c(s) \sqsubseteq s_1 \wedge \psi_2^c(s) \sqsubseteq s_2 \}; \psi_X^c(s).$$

In other words, the angel chooses a global play  $s$  matching  $s_1$  and  $s_2$  and produces its external view. By extending  $\psi^c$  to strategy specifications in the expected way, we obtain:

$$\tau \circ \sigma = s \leftarrow \sigma; t \leftarrow \tau; \psi^c(s, t).$$

### 7.4.2 Identity

The strategy  $\text{id}_E \in S[E, E]$  uses the signature:

$$[E] = E \hookrightarrow \textcircled{\text{O}}$$

and the transformation:

$$\psi_X^{\text{id}} := [2] : [E] \rightarrow [E, E]$$

$$\psi_X^{\text{id}}(\text{O}) := \text{O} \quad \psi_X^{\text{id}}(m) := mm$$

Then  $\text{id}_E$  is defined as:

$$\text{id}_E := \bigsqcup_{s \in P[E]} \psi_X^{\text{id}}(s).$$

### 7.4.3 Tensor

The tensor product of the strategies  $\sigma_1 \in S[E_1, F_1]$  and  $\sigma_2 \in S[E_2, F_2]$  is a strategy  $\sigma_1 \otimes \sigma_2 \in S[E_1 \otimes E_2, F_1 \otimes F_2]$  defined using interaction sequences in  $\Gamma = [E_1, F_1] \vee [E_2, F_2]$ . The external projection  $\psi_X^\otimes : \Gamma \rightarrow [E_1 \otimes E_2, F_1 \otimes F_2]$  is:

$$\begin{aligned}\psi_X^\otimes(O) &= O & \psi_X^\otimes(P_1) &= \psi_X^\otimes(P_2) = P \\ \psi_X^\otimes(\iota_i(m)) &= \iota_i(m)\end{aligned}$$

The internal projections  $\psi_i^\otimes : \Gamma \rightarrow [E_i, F_i]$  are given by:

$$\begin{aligned}\psi_i^\otimes(p) &= \begin{cases} P & \text{if } p = P_i \\ O & \text{otherwise} \end{cases} \\ \psi_i^\otimes(\iota_j(m)) &= \begin{cases} m & \text{if } i = j \\ \epsilon & \text{otherwise} \end{cases}\end{aligned}$$

The footprint of the plays  $s_1 \in P[E_1, F_1]$  and  $s_2 \in P[E_2, F_2]$  is:

$$\psi^\otimes(s_1, s_2) := \bigsqcup_{s \in P\Gamma} \{\psi_1^\otimes(s) \sqsubseteq s_1 \wedge \psi_2^\otimes(s) \sqsubseteq s_2\}; \psi_X^\otimes(s)$$

The tensor product can then be defined as:

$$\sigma_1 \otimes \sigma_2 := s_1 \leftarrow \sigma_1; s_2 \leftarrow \sigma_2; \psi^\otimes(s_1, s_2)$$

## 7.5 Category

The category  $\mathcal{G}_{\sqsubseteq}^b$  has effect signatures as objects, and has the elements of  $S[E, F]$  as morphisms  $\sigma : E \rightarrow F$ . The categorical structure is defined in the previous section.

The associator, unitor and braiding associated with  $\otimes$  can be obtained by embedding the corresponding morphisms of  $\mathcal{G}_{\sqsubseteq}^{ib}$  using the process outlined in §7.6 below. Note however that unlike that of  $\mathcal{G}_{\sqsubseteq}^{ib}$ , the symmetric monoidal structure of  $\mathcal{G}_{\sqsubseteq}^{ib}$  is not cartesian, because the interactions of



a strategy  $\sigma : E \rightarrow F_1 \otimes F_2$  which involve only one of the games  $F_1$  and  $F_2$  are not sufficient to characterize the behavior of  $\sigma$  in interactions that involve both of them.

## 7.6 Embedding $\mathcal{G}_{\sqsubseteq}^{ib}$

Since a morphism  $f \in \mathcal{G}_{\sqsubseteq}^{ib}(E, F)$  defined using the interaction specification monad only describes the behavior of a component for a single opponent question, to construct a corresponding strategy  $Wf \in \mathcal{G}_{\sqsubseteq}^b(E, F)$  we must duplicate the component's behavior, compounding the angelic and demonic choices of each copy.

We proceed as follows. For a stack  $\kappa : \mathbf{O} \rightarrow \mathbf{O}$ , the set  $P_{\Gamma}^{\kappa}$  contains partial plays of type  $(\mathbf{O}, \kappa) \rightarrow (\mathbf{O}, \kappa') \in \hat{\Gamma}$ , and for a question  $q \in F$ , the set  $\bar{P}_{\Gamma}^{\kappa q}$  contains partial plays of type  $(\mathbf{P}, \kappa q) \rightarrow (\mathbf{O}, \kappa') \in \hat{\Gamma}$ . We will define an operator:

$$\omega^{\kappa} : P_{\Gamma}^{\kappa} \rightarrow \mathbf{FCD}(P_{\Gamma}^{\kappa})$$

which *prepends* an arbitrary number of copies of  $f$  to a play of  $P_{\Gamma}^{\kappa}$ . Starting with  $\omega_0^{\kappa}(t) := t$ , we construct a series of approximations:

$$\omega_{i+1}^{\kappa}(t) := t \sqcup \bigsqcup_{q \in F} q \bar{\omega}_i^{\kappa q}(f^q, \omega_i^{\kappa}(t))$$

The auxiliary construction:

$$\bar{\omega}^{\kappa q} : \bar{P}_E(\text{ar}(q)) \times P_{\Gamma}^{\kappa} \rightarrow \mathbf{FCD}(\bar{P}_{\Gamma}^{\kappa q})$$

embeds an interaction  $s \in \bar{P}_E(\text{ar}(q))$ , inserting reentrant calls as appropriate, and continues with the play  $t$  if  $s$  terminates:

$$\bar{\omega}_i^{\kappa q}(\underline{v}, t) = vt$$

$$\bar{\omega}_i^{\kappa q}(\underline{m}, t) = m \omega_i^{\kappa q m}(\epsilon)$$

$$\bar{\omega}_i^{\kappa q}(\underline{m}ns, t) = m \omega_i^{\kappa q m}(n \bar{\omega}_i^{\kappa q}(s, t))$$

The index  $i$  limits both the number of sequential and reentrant copies of  $f$  which are instantiated.

The strategy specification associated to  $f$  in  $\mathcal{G}_{\sqsubseteq}^b$  is:

$$Wf := \bigsqcup_{i \in \mathbb{N}} \omega_i(\epsilon).$$

## 7.7 Hiding state

The functor  $W : \mathcal{G}_{\sqsubseteq}^{ib} \rightarrow \mathcal{G}_{\sqsubseteq}^b$  can be used to embed the layer theory defined in §6.7 as-is. In addition, the *state* of layer interfaces can be propagated across consecutive calls and eliminated from the representation.

**Definition 7.7.** The *state-free observation* at  $k_0 \in S$  of a partial play  $s : (O, \kappa) \twoheadrightarrow (O, \kappa')$  over the signature  $[E@S, F@S]$  is written  $s/k_0$  and defined recursively as:

$$\begin{aligned} \epsilon/k_0 &:= \epsilon \\ (m@k_1 \ n@k_2 \ s)/k_0 &:= \{k_0 = k_1\}; mn(s/k_2) \end{aligned}$$

For  $\sigma : E@S \rightarrow F@S$ , the strategy  $\sigma/k_0 : E \rightarrow F$  is obtained using the **FCD** extension of the operator above.

When the strategy  $\sigma/k_0$  is first activated,  $\sigma$  is passed the initial state  $k_0$ . Then, whenever  $\sigma$  makes a move  $m@k$ ,  $\sigma/k_0$  removes  $k$  from the visible interaction, but remember it in order to adjoin it to the next incoming move.

## **Part III**

# **Compiling certified open components**

## Chapter 8

# Semantics in CompCert

This chapter describes the semantic model used in CompCert [Leroy, 2009], as well as extensions which have been proposed to make it more compositional. There is no novelty in the technical material presented, however game semantics and dual nondeterminism provide a powerful lens through which the existing design and extensions can be examined and their more sophisticated aspects understood.

Going beyond the idea of a *certified compiler*, I then identify a set of requirements for using CompCert as a *compiler of certified components*.

### 8.1 Whole-program semantics in CompCert

The semantics of CompCert languages are given in terms of a simple notion of process behavior. By *process*, we mean a self-contained computation which can be characterized by the sequence of system calls it performs. For a C program to be executed as a process, its translation units must be compiled to object files, then linked together into an executable binary loaded by the system.

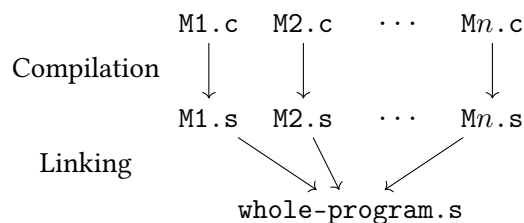


Figure 8.1: CompCert’s approximation of the C toolchain

The model used for verifying CompCert accounts for this in the way depicted in Fig. 8.1. Linking is approximated by merging programs, seen as sets of global definitions. The execution of a program composed of the translation units  $M1.c \dots Mn.c$  which compile to  $M1.s \dots Mn.s$  is modeled as:

$$L_{\text{tgt}} := \text{Asm}[M1.s + \dots + Mn.s].$$

Here,  $+$  denotes CompCert's linking operator and  $\text{Asm}$  maps an assembly program to its semantics. Note that the loading process is encoded as part of the definition of  $\text{Asm}$ , which constructs a global environment laying out the program's code and static data into the runtime address space, and models the conventional invocation of `main`. To formulate compiler correctness, we must also specify the behavior of the source program. To this end, CompCert defines a linking operator and semantics for the language Clight,<sup>1</sup> allowing the desired behavior to be specified as:

$$L_{\text{src}} := \text{Clight}[M1.c + \dots + Mn.c].$$

Given an appropriate notion of refinement, the correctness of CompCert can then be stated as:

$$L_{\text{src}} \sqsubseteq L_{\text{tgt}}.$$

### 8.1.1 Transition systems

Language semantics are given as labelled transition systems (LTS), which characterize a program's behavior in terms of sequences of observable events taken from a fixed set  $\mathbb{E}$ .

Schematically, a CompCert LTS is a tuple  $L = \langle S, I, \rightarrow, F \rangle$  consisting of a set of states  $S$ , a subset  $I \subseteq S$  of initial states, a labelled transition relation  $\rightarrow \subseteq S \times \mathbb{E}^* \times S$ , and a set  $F \subseteq S \times \text{int}$  of final states associated with exit statuses. The relation  $s \xrightarrow{t} s'$  indicates that the state  $s$  may transition to the state  $s'$  with the externally observable interaction  $t \in \mathbb{E}^*$ .

The execution of  $L$  starts by selecting an initial state  $s_0$ , the uses the transition relation to

---

<sup>1</sup>Although CompCert features a frontend for a richer version of the C language, the simplified intermediate dialect Clight is usually used as the source language when using CompCert to build certified artifacts.

repeatedly update the state, until a final state is reached:

$$I \ni s_0 \xrightarrow{t_1} s_1 \xrightarrow{\epsilon} s_2 \xrightarrow{t_2} s_3 \ F \ r$$

Roughly speaking, the externally observable behavior of the program consists of the sequence of events generated by transitions ( $t_1 t_2$  in the execution above), together with the exit status of the process ( $r$  in the execution above). See the next section for a more detailed account.

Note that nondeterministic choices are potentially involved at every step of the execution: the selection of an initial state from the set  $I$ , the selection of a possible transition  $\{(t, s') \mid s \xrightarrow{t} s'\}$ , and the selection of an outcome  $\{r \mid s \ F \ r\}$ . If there a possible transition out of a final state  $s$ , there will also be a choice as to whether the execution should terminate or continue.

The angelic, demonic, or mixed nature of these choices is key to understanding CompCert's notions of *nondeterminacy* and *receptiveness* of transition systems, and the distinction and correspondence between CompCert's notions of *forward* and *backward* simulations.

### 8.1.2 Trace semantics

I outlined above the execution of a labelled transition system in terms of terminating traces. The model used in CompCert also considers other kinds of behaviors:

- An execution reaching a final state is said to *terminate*. For example, the following execution generates the event trace  $t_1 t_2 \cdots t_{n-1}$  and terminates with status  $r$ :

$$I \ni s_1 \xrightarrow{t_1} s_2 \xrightarrow{t_2} \cdots \xrightarrow{t_{n-1}} s_n \ F \ r.$$

I will write this behavior as:

$$t_1 t_2 \cdots t_{n-1} \Downarrow r.$$

- An execution reaching an infinite sequence of  $\epsilon$  transitions is said to *silently diverge*. The following execution diverges after generating the trace  $t_1 t_2 \cdots t_{n-1}$ :

$$I \ni s_1 \xrightarrow{t_1} s_2 \xrightarrow{t_2} \cdots \xrightarrow{t_{n-1}} s_n \xrightarrow{\epsilon} s_{n+1} \xrightarrow{\epsilon} \cdots$$

I will write silently diverging behaviors as:

$$t_1 t_2 \cdots t_{n-1} \uparrow .$$

- By contrast, infinite executions which keep interacting are said to exhibit *reactive* behavior.

The following execution is reactive if and only if  $\forall i \cdot \exists j \geq i \cdot t_j \neq \epsilon$ :

$$I \ni s_1 \xrightarrow{t_1} s_2 \xrightarrow{t_2} s_3 \xrightarrow{t_3} \cdots$$

Then the behavior of the transition system is represented by the infinite sequence:

$$t_1 t_2 \cdots$$

- Finally, an execution which reaches a stuck state is said to *go wrong*. It will have the shape:

$$I \ni s_1 \xrightarrow{t_1} s_2 \xrightarrow{t_2} \cdots \xrightarrow{t_{n-1}} s_n ,$$

with no  $t, s'$  such that  $s_n \xrightarrow{t} s'$ . The corresponding, partially undefined behavior is written:

$$t_1 t_2 \cdots t_{n-1} \not\downarrow .$$

It can be refined by any behavior admitting  $t_1 t_2 \cdots t_{n-1}$  as a prefix.

A transition system with no initial state ( $I = \emptyset$ ) immediately goes wrong. In this case, the transition system will admit the single undefined behavior  $\not\downarrow$ .

CompCert then characterizes the externally observable behavior of a transition system  $L$  as a set  $\text{Beh}(L) \subseteq \mathbb{B}$  of individual behaviors  $b \in B$  as defined above. I will often simply write  $b \in L$  to mean that the transition system  $L$  admits the behavior  $b$ , overlooking the distinction between  $L$  and the set of behaviors it admits.

Refinement is defined mostly as trace containment, but also takes into account undefined behaviors. Writing the prefix relation between individual behaviors as  $\sqsubseteq_{\text{pref}} \subseteq \mathbb{B} \times \mathbb{B}$ , this notion of

refinement can be defined as:

$$L_1 \sqsubseteq L_2 \Leftrightarrow \forall b_2 \in L_2 \cdot \exists b_1 \in L_1 \cdot b_1 \subseteq_{\text{pref}} b_2$$

That is to say, any behavior  $b_2 \in L_2$  of the target program refines a behavior  $b_1 \in L_1$  of the source program. The behaviors of the target program must largely be contained within the behaviors of the source program, but when the source program goes wrong this restriction is lifted.

### 8.1.3 Nondeterminism

While this is not stated explicitly in the literature surrounding CompCert, the interpretation of transition systems presented above uses both angelic and demonic nondeterminism.

For the most part, the interpretation is demonic, as underscored by the trace containment aspect of the definition of  $\sqsubseteq$ . However, undefined behaviors introduce a discontinuity in this model. To illustrate this phenomenon, consider the following sequence of transition systems, where the initial state  $i$  may admit zero, one, or more transitions to the final state  $f$ :

$S := \{i, f\}$	$L_0 := \{S, I, \rightarrow_0, F\}$	$\rightarrow_0 := \emptyset$
$I := \{i\}$	$L_1 := \{S, I, \rightarrow_1, F\}$	$\rightarrow_1 := \rightarrow_0 \cup \{(i, e_1, f)\}$
$F := \{(f, 42)\}$	$L_2 := \{S, I, \rightarrow_2, F\}$	$\rightarrow_2 := \rightarrow_1 \cup \{(i, e_2, f)\}$
	$\vdots$	$\vdots$

The corresponding sets of behaviors are as follows:

$$\begin{aligned}
 \text{Beh}(L_0) &= \{\zeta\} \\
 \text{Beh}(L_1) &= \{e_1 \Downarrow 42\} \\
 \text{Beh}(L_2) &= \{e_1 \Downarrow 42, e_2 \Downarrow 42\} \\
 &\vdots
 \end{aligned}$$



As a consequence of the discontinuity introduced by  $\downarrow$ , we obtain the ordering:

$$L_0 \sqsubseteq \cdots \sqsubseteq L_3 \sqsubseteq L_2 \sqsubseteq L_1 .$$

One way to think about this phenomenon is to state it in more general terms, considering the way CompCert interprets nondeterminism in a set of possible actions  $A$ , depending on the set's cardinality. When  $|A| \leq 1$ , nondeterminism is interpreted as angelic:

$$\bigsqcup A$$

When  $|A| \geq 1$ , nondeterminism is interpreted as demonic:

$$\bigsqcap A$$

Therefore CompCert uses a notion of *discontinuous choice*:

$$\oplus A := \begin{cases} \perp & \text{if } A = \emptyset \\ \bigsqcap A & \text{otherwise} \end{cases}$$

As we will see, CompCert's forward simulations correspond to the angelic interpretation  $\bigsqcup A$ , whereas backward simulations correspond to the discontinuous interpretation  $\oplus A$ . This discontinuity is at the root of the *safety* conditions found in CompCert's definition of backward simulations (§??). For deterministic transition systems,  $A$  is always of size at most 1, so that the two interpretations coincide.

#### 8.1.4 Forward simulations

#### 8.1.5 Backward simulations

#### 8.1.6 Memory model

The construction of states in CompCert language semantics follows common patterns. In particular, all languages start with the same notion of *memory state*.

$$\begin{aligned}
v \in \text{val} &::= \text{undef} \mid \text{int}(n) \mid \text{long}(n) \mid \text{float}(x) \mid \text{single}(x) \mid \text{vptr}(b, o) \\
(b, o) \in \text{ptr} &= \text{block} \times \mathbb{Z} \quad (b, l, h) \in \text{ptrrange} = \text{block} \times \mathbb{Z} \times \mathbb{Z} \\
\\
\text{alloc} &: \text{mem} \rightarrow \mathbb{Z} \rightarrow \mathbb{Z} \rightarrow \text{mem} \times \text{block} \\
\text{free} &: \text{mem} \rightarrow \text{ptrrange} \rightarrow \text{option}(\text{mem}) \\
\text{load} &: \text{mem} \rightarrow \text{ptr} \rightarrow \text{option}(\text{val}) \\
\text{store} &: \text{mem} \rightarrow \text{ptr} \rightarrow \text{val} \rightarrow \text{option}(\text{mem})
\end{aligned}$$

Figure 8.2: Outline of the CompCert memory model

The CompCert memory model [Leroy et al., 2012; Leroy and Blazy, 2008] is the core algebraic structure underlying the semantics of CompCert’s languages. Some of its operations are shown in Fig. 8.2. The idealized version presented here involves the type of memory states  $\text{mem}$ , the type of runtime values  $\text{val}$ , and the types of pointers  $\text{ptr}$  and address ranges  $\text{ptrrange}$ . To keep our exposition concise and clear, we gloss over the technical details associated with modular arithmetic and overflow constraints.

The memory is organized into a finite number of *blocks*. Each memory block has a unique identifier  $b \in \text{block}$  and is equipped with its own linear address space. Block identifiers and offsets are often manipulated together as pointers  $p = (b, o) \in \text{ptr} = \text{block} \times \mathbb{Z}$ . New blocks are created with prescribed boundaries using the primitive  $\text{alloc}$ . A runtime value  $v \in \text{val}$  can be stored at a given address using the primitive  $\text{store}$ , and retrieved using the primitive  $\text{load}$ . Values can be integers ( $\text{int}$ ,  $\text{long}$ ) and floating point numbers ( $\text{float}$ ,  $\text{single}$ ) of different sizes, as well as pointers ( $\text{vptr}$ ). The special value  $\text{undef}$  represents an undefined value. Simulation relations often allow  $\text{undef}$  to be refined into a more concrete value; we write value refinement as  $\leq_v := \{(\text{undef}, v), (v, v) \mid v \in \text{val}\}$ .

The memory model is shared by all of the languages in CompCert. States always consist of a memory component  $m \in \text{mem}$ , alongside language-specific components which may contain additional values ( $\text{val}$ ).

## 8.2 Compositional extensions

The model outlined above makes no attempt to model interactions across components, and is only ever used to describe the behavior of whole programs, and the compiler’s correctness property as

described in §?? only considers uses which follow the pattern approximated in Fig. 8.1. This makes it challenging to reason about the behavior of individual compilation units.

To formulate a more fine-grained and flexible version of the correctness theorem of CompCert, we need an account of the behavior of individual translation units. The model I introduce in §9.2.1 achieves this by recording control transfers to and from the modeled system explicitly.

However, there already exists multiple extensions to CompCert’s semantic model and correctness theorem which address some of limitations of CompCert’s original approach. I present some of them in this section. The conceptual framework of games can be used to classify these extensions. By reinterpreting these models in terms of games and strategies, we can establish the taxonomy presented in Table 8.1.

We focus here on CompCert, but a more general survey, discussion and synthesis of various compositional compiler correctness results is provided by [Patterson and Ahmed \[2019\]](#).

### 8.2.1 CompCert and SepCompCert

As noted in §9.1, the whole-program semantics used by CompCert express strategies for the game  $1 \rightarrow \mathcal{W}$ . CompCert’s original correctness theorem stated the refinement property  $C_{wp}(p) \sqsubseteq \text{Asm}_{wp}(p')$ , where  $C_{wp}$  and  $\text{Asm}_{wp}$  denote the source and target whole-program semantics. SepCompCert [[Kang et al., 2016](#)] later introduced the linking operator  $+$  and generalized the correctness theorem to the form discussed in §8.1.

Since external calls are not accounted for explicitly in this semantic model, they are interpreted using a common global parameter  $\chi$  specifying their behavior. The correctness proof assumes that  $\chi$  is deterministic and that it satisfies a number of healthiness requirements with respect to the memory transformations used in CompCert’s correctness proof, corresponding roughly to self-simulation under the CKLRs `ext` and `injp`.

### 8.2.2 Contextual compilation

CompCertX [[Gu et al., 2015](#)] and Stack-Aware CompCert [[Wang et al., 2019](#)] generalize the incoming interface of programs from  $\mathcal{W}$  to  $\mathcal{C}$ , and as such characterizes the behavior not only of `main` but of any function of the program, called with any argument values. This allows CompCertX and its correctness theorem to be used in the layer-based verification of the CertiKOS kernel: once the

code of a given abstraction layer has been verified and compiled using CompCertX, that layer's specification can be used as the new  $\chi$  when the next layer is verified. However, this approach does not support mutually recursive components, and requires the healthiness conditions on  $\chi$  to be proved before the next layer is added.

### 8.2.3 Compositional CompCert

The *interaction semantics* of Compositional CompCert [Stewart et al., 2015] are closer to our own model but are limited to the language interface  $\mathcal{C}$ . Likewise, the simulations used in Compositional CompCert correspond to our notion of forward simulation for a single convention called *structured injections*, which we will write  $\mathbb{SI}$ . Simulation proofs are updated to follow this model, and the *transitivity* of  $\mathbb{SI}$  is established ( $\mathbb{SI} \cdot \mathbb{SI} \equiv \mathbb{SI}$ ), so that passes can be composed to obtain a simulation for the whole compiler.

Compositional CompCert also introduced a notion of *semantic linking* similar to our horizontal composition (§9.2.1). As in our case, semantic linking is shown to preserve simulations (Thm. 9.8), however semantic linking is not related to syntactic linking of assembly programs, and this was later shown to be problematic [Song et al., 2019].

Another limitation of Compositional CompCert is the complexity of the theory and the proof effort required. Because of the use of a single simulation convention, many assumptions naturally expressed as relational invariants in the simulation relations of CompCert must be either captured by  $\mathbb{SI}$  or handled at the level of language semantics, and simulation proofs essentially had to be rewritten and adapted to structured injections.

### 8.2.4 CompCertM

The most recent extension of CompCert is CompCertM [Song et al., 2019], which shares common themes and was developed concurrently with our work. While its correctness is ultimately stated in terms of closed semantics, CompCertM uses a notion of open semantics as an intermediate construction to enable compositional compilation and verification.

The open semantics used in CompCertM builds on interaction semantics by incorporating an assembly language interface. The resulting semantic model can be characterized as  $\mathcal{C} \times \mathcal{A} \rightarrow \mathcal{C} \times \mathcal{A}$ . Simulations are parametrized by Kripke relations similar to CKLRs (§10.1) and predicates

similar to our invariants (§10.2). While simulations do not directly compose, a new technique called *refinement under self-related context* (RUSC) can nonetheless be used to derive a contextual refinement theorem for the whole compiler with minimal overhead.

This approach has many advantages. CompCertM avoids much of the complexity of Compositional CompCert when it comes to composing passes, and the flexibility of the simulations used makes updating the correctness proofs of passes much easier. CompCertM also charts new ground in several directions. The RUSC relation used to state the final theorem is shown to be adequate with respect to the trace semantics of closed programs. CompCertM has improved support for static variables and the verification of the assembly runtime function `utod` is demonstrated.

In other aspects, CompCertM inherits the limitation of previous approaches whereas CompCertO goes further. Because the compiler correctness theorem is not itself expressed as a simulation, it fails requirement #2 laid out in §8.3.2. While it may be possible to accomodate #3, the parametrization of simulations does not offer the same flexibility as our full-blown notion of simulation convention. As a consequence, a cascade of techniques (*repaired* interaction semantics, *enriched* memory injections, the mixed simulations of [Neis et al., 2015]) need to be deployed to enforce invariants which find a natural relational expression under our approach.

Therefore, an interesting question for future investigation will be to determine to what extent the techniques used by CompCertM and CompCertO could be integrated to combine the strengths of both developments.

## 8.3 Limitations

### 8.3.1 Decomposing heterogeneous systems

Existing work turns CompCert into a platform enabling compositional verification (§8.2). However, in most cases, the horizon is a completed assembly program to be run as a user-level process. This becomes a limitation in the context of heterogeneous systems.

For example, consider the problem of verifying a network interface card (NIC) driver. The NIC and its driver are closely coupled, but the details of their interaction are irrelevant to the rest of the system and should not leak into our reasoning at larger scales. Instead, we wish to treat them as a unit and establish a direct relationship between calls into the driver's C interface and network

communication. Together, the NIC and driver implement a specification  $\sigma : \text{Net} \rightarrow \mathcal{C}$  (see §7). The driver code would be specified ( $\sigma_{\text{drv}}$ ) and verified at the level of CompCert semantics, whereas device I/O primitives ( $\sigma_{\text{io}}$ ) and the NIC ( $\sigma_{\text{NIC}}$ ) would be specified as additional components:

$$\sigma_{\text{NIC}} : \text{Net} \rightarrow \text{IO} \quad \sigma_{\text{io}} : \text{IO} \rightarrow \mathcal{C} \quad \sigma_{\text{drv}} : \mathcal{C} \rightarrow \mathcal{C}$$

By reasoning about their interaction, it would be possible to establish a relationship between the overall specification  $\sigma$  and the composition  $\sigma_{\text{drv}} \circ \sigma_{\text{io}} \circ \sigma_{\text{NIC}}$ . Then a *compiler of certified components* would help us transport specifications and proofs obtained with respect to the driver's C code to the compiled code operating at the level of assembly ( $\sigma' : \text{Net} \rightarrow \mathcal{A}$ ).

In this paper, we show how to adapt the semantic model and correctness proofs of CompCert so that they can be used in this way. Our model is not intended to reach the level of generality required to handle all aspects of the problem above; indeed we want it to remain tailored to CompCert's verification as much as possible. Instead, we envision a hierarchy where components could be verified in suitable models, then soundly embedded in more general ones to be made interoperable. To make it possible for CompCert to be used in this context, the ability to treat the driver code as an independent component is crucial. This excludes approaches to compositional compiler correctness which are formulated in terms of completed programs.

### 8.3.2 Requirements

To handle use cases like the one we have presented above, the compiler's correctness proof should satisfy the following requirements:

1. The semantics of the source and target languages should characterize the behavior of *open* components in terms of their interactions with the rest of the program.
2. The correctness theorem should go beyond refinement under a fixed notion of program context, and relate the interactions of the source and target modules directly.
3. The abstraction gap between C and assembly-level interactions should be made explicit.
4. Some form of certified *linking* should be provided as well as certified compilation.

Variant	Semantic model	(1)	(2)	(3)	(4)	(5)
(Sep)CompCert [Kang et al., 2016; Leroy, 2009]	$\chi : \mathbf{1} \rightarrow \mathcal{C} \vdash \mathbf{1} \rightarrow \mathcal{W}$				✓	✓
CompCertX [Gu et al., 2015]	$\chi : \mathbf{1} \rightarrow \mathcal{C} \times \mathcal{A} \vdash \mathbf{1} \rightarrow \mathcal{C} \times \mathcal{A}$				✓	✓
Compositional CompCert [Stewart et al., 2015]	$\mathcal{C} \rightarrow \mathcal{C}$	✓	✓			
CompCertM [Song et al., 2019]	$\mathcal{C} \times \mathcal{A} \rightarrow \mathcal{C} \times \mathcal{A}$	✓		✓	✓	✓
CompCertO	$A \rightarrow A \quad (A \in \mathbb{L})$	✓	✓	✓	✓	✓

Table 8.1: Taxonomy of CompCert extensions in terms of the corresponding game models (§7) and the requirements they satisfy (§8.3.2). The parameter  $\chi : \mathbf{1} \rightarrow \mathcal{C}$  pre-specifies the behavior of external functions, whereas games on the left of arrows correspond to dynamic interactions. As a distinguishing feature, CompCertO’s model is parametrized by a generic notion of *language interface*  $A \in \mathbb{L} \supseteq \{\mathcal{C}, \mathcal{A}\}$ . See §8.2 for details.

5. To facilitate integration into the official release, changes to the existing proofs of CompCert should be minimal.

As outlined in Table 8.1, each of these requirements is fulfilled by some existing CompCert extension, however none satisfies them all.

This paper introduces CompCertO, the first extension of CompCert to address all of these requirements simultaneously. The key to this achievement is the expressivity of our model.

## Chapter 9

# CompCertO

This chapter presents the high-level design of CompCertO. CompCertO is the first extension of CompCert which satisfies all of the requirements I have identified in §8.3.2.

CompCertO generalizes the semantic model of CompCert to express interactions between compilation units, using *language interfaces* to describe the form of these interactions and *simulation conventions* to describe the correspondence between the interfaces of source and target languages. The behavior of composite programs is specified by a *horizontal composition* operator which is shown to be correctly implemented by the existing linking operator for assembly programs. To combine and reason about simulation proofs, we introduce a rich *simulation convention algebra* and use it to derive our main compiler correctness statement.

[XXX restructure]

The games we start from have a particularly simple structure. We call each one a *language interface*. Their moves are partitionned into questions and answers, where questions correspond to function invocations and answers return control to the caller.

**Definition 9.1.** A *language interface* is a tuple  $A = \langle A^\circ, A^\bullet \rangle$ , where  $A^\circ$  is a set of *questions* and  $A^\bullet$  is a set of *answers*.

We focus on games of the form  $A \rightarrow B$ , where  $A$  and  $B$  are language interfaces. In this setting, the valid positions of  $A \rightarrow B$  are sequences of the form:

$$q \cdot \underline{m_1} \cdot n_1 \cdots \underline{m_k} \cdot n_k \cdot \underline{r} \in B^\circ (A^\circ A^\bullet)^* B^\bullet$$



<u>A.c</u>	int mult(int n, int p) { return n * p; }	<u>A.s</u>	mult: %eax := %ebx %eax *= %ecx ret
<u>B.c</u>	int sqr(int n) { return mult(n, n); }	<u>B.s</u>	sqr: %ecx := %ebx call mult L1: ret

Figure 9.1: Two simple C compilation units and corresponding assembly code. For this example, the calling convention stores arguments in the registers %ebx and %ecx and return values in the register %eax.

and all their prefixes. In our context, this describes a program component responding to an incoming call  $q$ . The component performs a series of external calls  $m_1 \dots m_k$  yielding the results  $n_1 \dots n_k$ , and finally returns from the top-level call with the result  $r$ . The arrows show the correspondence between questions and answers but are not part of the model.

**Example 9.2.** *We use a simplified version of C and assembly to illustrate some of the principles behind our model. Consider the program components in Fig. 9.1. The behavior of B.c as it interacts with A.c is described by plays of the form:*

$$\text{sqr}(3) \cdot \underline{\text{mult}(3, 3)} \cdot 9 \cdot \underline{\phantom{0}} \quad (9.1)$$

This corresponds to the game  $\tilde{C} \rightarrow \tilde{C}$  for a language interface  $\tilde{C} := \langle \text{ident} \times \text{val}^*, \text{val} \rangle$ . Questions specify the function to invoke and its arguments; answers carry the value returned by the function.

To describe the behavior of A.s and B.s, we use a set of registers  $R := \{\text{pc}, \text{eax}, \text{ebx}, \text{ecx}\}$  together with a stack of pending return addresses (pc is the program counter). The corresponding language interface can be defined as  $\tilde{\mathcal{A}} := \langle \text{val}^R \times \text{val}^*, \text{val}^R \times \text{val}^* \rangle$ . A possible execution of B.s is:

$$\begin{bmatrix} \text{pc} \mapsto \text{sqr} \\ \text{eax} \mapsto 42 \\ \text{ebx} \mapsto 3 \\ \text{ecx} \mapsto 7 \\ \text{stack}: x \cdot \vec{k} \end{bmatrix} \cdot \begin{bmatrix} \text{pc} \mapsto \text{mult} \\ \text{eax} \mapsto 42 \\ \text{ebx} \mapsto 3 \\ \text{ecx} \mapsto 3 \\ \text{stack}: \text{L1} \cdot x \cdot \vec{k} \end{bmatrix} \cdot \begin{bmatrix} \text{pc} \mapsto \text{L1} \\ \text{eax} \mapsto 9 \\ \text{ebx} \mapsto 3 \\ \text{ecx} \mapsto 3 \\ \text{stack}: x \cdot \vec{k} \end{bmatrix} \cdot \begin{bmatrix} \text{pc} \mapsto x \\ \text{eax} \mapsto 9 \\ \text{ebx} \mapsto 3 \\ \text{ecx} \mapsto 3 \\ \text{stack}: \vec{k} \end{bmatrix} \quad (9.2)$$

The correspondence between (9.1) and (9.2) is determined by the C calling convention in use. We will discuss this point in more detail in §9.1.3.

## 9.1 Overview

The semantic model of CompCert corresponds to a game  $\mathcal{E} \rightarrow \mathcal{W}$ . Programs are run without any parameters and produce a single integer denoting their exit status. This is described by the language interface  $\mathcal{W} := \langle \mathbb{1}, \text{int} \rangle$ , where  $\mathbb{1} = \{*\}$  is the unit set and  $\text{int}$  is the set of machine integers. Interaction with the environment is captured as a trace of events from a predefined set, each with an output and input component. These events, described by the language interface  $\mathcal{E}$ , correspond to system calls and accesses to volatile variables.

### 9.1.1 Semantic model

In order to model open components and cross-component interactions, we generalize CompCert's labelled transition systems to describe strategies for games of the form:

$$A \twoheadrightarrow B := A \times \mathcal{E} \rightarrow B.$$

The language interface  $B$  describes how a component can be activated, and the ways in which it can return control to the caller. The language interface  $A$  describes the external calls that the component may perform in the course of its execution.

This flexibility allows us to treat interactions at a level of abstraction adapted to each language. For example, CompCertO's source language Clight uses the game  $\mathcal{C} \twoheadrightarrow \mathcal{C}$ . The questions of  $\mathcal{C}$  specify a function to call, argument values, and the state of the memory at the time of invocation; the answers specify a return value and an updated memory state. On the other hand, the target language Asm uses  $\mathcal{A} \twoheadrightarrow \mathcal{A}$ , where  $\mathcal{A}$  describes control transfers in terms of processor registers rather than function calls. The language interfaces used in CompCertO are described in §9.2.1.

### 9.1.2 Simulations

CompCert uses simulation proofs to establish a correspondence between the externally observable behaviors of the source and target programs of each compilation pass. The internal details of simulation relations have no bearing on this correspondence, so these details can remain hidden to fit a uniform and transitive notion of pass correctness. This makes it easy to derive the correctness

of the whole compiler from the correctness of each pass.

To achieve compositionality across compilation units, our model must reveal details about component interactions which were previously internal. Since many passes transform these interactions in specialized ways, this breaks the uniformity of pass correctness properties.

Existing work attempts to recover this uniformity by developing a general notion of correctness covering all passes or by delaying pass composition so that it operates on closed semantics only. Unfortunately, these techniques either conflict with our requirement #2, make proofs more complex, or cascade into subtle “impedence mismatch” problems requiring their own solutions (see §8.2).

By contrast, we capture the particularities of each simulation proof by introducing a notion of *simulation convention* expressing the correspondence between source- and target-level interactions. To describe simulation conventions and reason about them, we use logical relations.

### 9.1.3 Simulation conventions

The framework of Kripke relations allows us to define simulation conventions as follows. The worlds ensure that corresponding pairs of questions and answers are related consistently.

**Definition 9.3.** A *simulation convention* between the language interfaces  $A_1 = \langle A_1^\circ, A_1^\bullet \rangle$  and  $A_2 = \langle A_2^\circ, A_2^\bullet \rangle$  is a tuple  $\mathbb{R} = \langle W, \mathbb{R}^\circ, \mathbb{R}^\bullet \rangle$  with  $\mathbb{R}^\circ \in \mathcal{R}_W(A_1^\circ, A_2^\circ)$  and  $\mathbb{R}^\bullet \in \mathcal{R}_W(A_1^\bullet, A_2^\bullet)$ . We will write  $\mathbb{R} : A_1 \Leftrightarrow A_2$ . In particular, for a language interface  $A$ , the *identity* simulation convention is defined as  $\text{id}_A := \langle \mathbb{1}, =, = \rangle : A \Leftrightarrow A$ . We will usually omit the subscript  $A$ .

A simulation between the transition systems  $L_1 : A_1 \twoheadrightarrow B_1$  and  $L_2 : A_2 \twoheadrightarrow B_2$  is then assigned a type  $\mathbb{R}_A \rightarrow \mathbb{R}_B$ , where  $\mathbb{R}_A : A_1 \Leftrightarrow A_2$  and  $\mathbb{R}_B : B_1 \Leftrightarrow B_2$  are simulation conventions relating the corresponding language interfaces. We write  $L_1 \leq_{\mathbb{R}_A \rightarrow \mathbb{R}_B} L_2$ .

Table 9.1 presents a summary of notations. Simulation conventions will often be derived from more elementary relations, following the internal structure of questions and answers (see §10.1).

**Example 9.4.** The calling convention we used in Example 9.2 can be formalized as a simulation convention  $\tilde{\mathbb{C}} := \langle \text{val}^*, \tilde{\mathbb{C}}^\circ, \tilde{\mathbb{C}}^\bullet \rangle : \tilde{\mathcal{C}} \Leftrightarrow \tilde{\mathcal{A}}$ . We use the set of worlds  $\text{val}^*$  to relate the stack of assembly

Notation	Examples	Description
$R \in \mathcal{R}(S_1, S_2)$	$\leq_v$	Simple relation
$R \in \mathcal{R}_W(S_1, S_2)$	$\hookrightarrow_m$	Kripke relation (Def. 2.13)
$w \Vdash R$		Kripke relation at world $w$
$w \Vdash x R y$		$x$ and $y$ related at world $w$
$A, B, C$	$\mathcal{C}, \mathcal{A}, \mathbf{1}$	Language interface (Def. 9.1)
$\mathbb{R} : A_1 \Leftrightarrow A_2$	alloc	Simulation convention (Def. 9.3)
$L : A \rightarrow B$	Clight( $p$ )	LTS for $A \rightarrow B$ (Def. 9.5)
$L_1 \oplus L_2$		Horizontal composition (Def. 9.6)
$L_1 \leq_{\mathbb{R} \rightarrow \mathbb{S}} L_2$	Thm. 9.16	Simulation property (Def. 9.7)

Table 9.1: Summary of notations

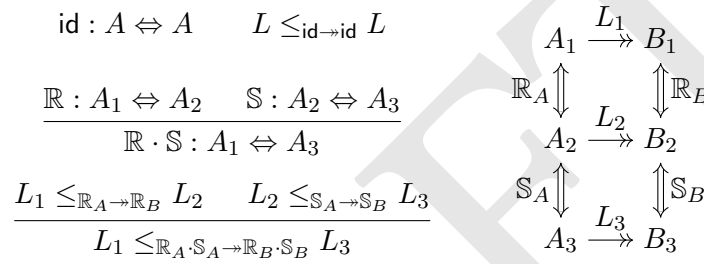


Figure 9.2: Simulation identity and vertical composition

questions to that of the corresponding answers. The relations  $\tilde{\mathcal{C}}^\circ, \tilde{\mathcal{C}}^\bullet$  are defined by:

$$\frac{rs[\text{pc}] = f \quad \vec{v} \text{ is contained in } rs[\text{ebx}, \text{ecx}]}{x\vec{k} \Vdash f(\vec{v}) \quad \tilde{\mathcal{C}}^\circ \quad rs@\vec{x}\vec{k}} \quad \frac{rs[\text{eax}] = v' \quad rs[\text{pc}] = x}{x\vec{k} \Vdash v' \quad \tilde{\mathcal{C}}^\bullet \quad rs@\vec{k}}$$

For a C-level function invocation  $f(\vec{v})$ , we expect the register  $\text{pc}$  to point to the beginning of the function  $f$ , and the registers  $\text{ebx}$  and  $\text{ecx}$  to contain the first and second arguments (if applicable). Other registers may contain arbitrary values. The stack  $x\vec{k}$  has no relationship to the C question, however the assembly answer is expected to pop the return address and branch to it, setting the program counter  $\text{pc}$  accordingly. In addition, the return value  $v'$  must be stored in the register  $\text{eax}$ .

The expressive power of simulation conventions makes the adaptation of existing correctness proofs for the various passes of CompCert straightforward. Instead of forcing all passes into a single one-size-fits-all mold, we can choose conventions matching the simulation relation and invariants used in each pass. Simulations for each pass can then be composed in the way shown in Fig. 9.2.

### 9.1.4 Simulation convention algebra

CompCert’s *injection* passes (see §10.1) pose a particular challenge, already encountered in the work on Compositional CompCert [Stewart et al., 2015]: injection passes make stronger assumptions on external calls than they guarantee for incoming calls. Our framework can express this situation by using different simulation conventions for external and incoming calls ( $\text{injp} \rightarrow \text{inj}$ ). However, since the external calls of one component and the incoming calls of another will not be related in compatible ways, this asymmetry breaks horizontal compositionality (Thm. 9.8).

In CompCertO, we rectify this imbalance *outside* of the simulation proof itself. Simulations for individual passes are not always horizontally compositional, but we are able to derive a symmetric simulation convention for the compiler as a whole. Properties of the Clight and RTL languages allow us to strengthen the correctness proof. These properties are encoded as self-simulations and inserted as pseudo-passes. We can then perform algebraic manipulations on simulation statements to rewrite the overall simulation convention used by the compiler into a symmetric one.

These algebraic manipulations are based on a notion of simulation convention refinement ( $\sqsubseteq$ ) allowing a simulation convention to replace another in all simulation statements. We construct a typed Kleene algebra [Kozen, 1998] based on this order, and use it to ensure that our compiler correctness statement is both compositional and insensitive to the inclusion of optional passes.

## 9.2 Operational semantics

### 9.2.1 Open semantics in CompCertO

The memory model also plays a central role when describing interactions between program components. In our approach, the memory state is passed alongside all control transfers.

**Language interfaces** Our models of cross-component interactions in CompCert languages are shown in Table 9.2. At the source level ( $\mathcal{C}$ ), questions consist of the address of the function being invoked ( $vf \in \text{val}$ ), its signature ( $sg \in \text{signature}$ ), the values of its arguments ( $\vec{v} \in \text{val}^*$ ), and the state of the memory at the point of entry ( $m \in \text{mem}$ ); answers consist of the function’s return value and the state of the memory at the point of exit. This language interface is used for Clight and for the majority of CompCert’s intermediate languages.

Name	Question	Answer	Description
$\mathcal{C}$	$vf[sg](\vec{v})@m$	$v'@m'$	C calls
$\mathcal{L}$	$vf[sg](ls)@m$	$ls'@m'$	Abstract locations
$\mathcal{M}$	$vf(sp, ra, rs)@m$	$rs'@m'$	Machine registers
$\mathcal{A}$	$rs@m$	$rs'@m'$	Arch-specific
$\mathbf{1}$	n/a	n/a	Empty interface
$\mathcal{W}$	*	$r$	Whole-program

Table 9.2: Language interfaces used in CompCertO

As we move towards lower-level languages, this is reflected in the language interfaces we use: function arguments are first mapped into abstract locations alongside local temporary variables ( $\mathcal{L}$ , used by LTL and Linear). These locations are eventually split between in-memory stack slots and a fixed number of machine registers ( $\mathcal{M}$ , used by Mach). Finally, the target assembly language Asm stores the program counter, stack pointer, and return address into their own machine registers, which is reflected in its interface  $\mathcal{A}$ .

The interface of whole-program execution can also be described in this setting: the language interface  $\mathbf{1}$  contains no move; per §9.1, the interface  $\mathcal{W}$  has a single trivial question  $*$ , and the answers  $r \in \text{int}$  give the exit status of a process. Hence the original CompCert semantics described in §8.1 can be seen to define strategies for  $\mathbf{1} \rightarrow \mathcal{W}$ : the process can only be started in a single way, cannot perform any external calls, and indicates an exit status upon termination.

**Transition systems** To account for the cross-component interactions described by language interfaces, CompCertO extends the transition systems described in §8.1 as follows.

**Definition 9.5.** Given an *incoming* language interface  $B$  and an *outgoing* language interface  $A$ , a *labelled transition system for the game  $A \rightarrow B$*  is a tuple  $L = \langle S, \rightarrow, D, I, X, Y, F \rangle$ . The relation  $\rightarrow \subseteq S \times \mathbb{E}^* \times S$  is a *transition relation* on the set of states  $S$ . The set  $D \subseteq B^\circ$  specifies which questions the component accepts;  $I \subseteq D \times S$  then assigns to each one a set of *initial states*.  $F \subseteq S \times B^\bullet$  designates *final states* together with corresponding answers. External calls are specified by  $X \subseteq S \times A^\circ$ , which designates *external states* together with a question of  $A$ , and  $Y \subseteq S \times A^\bullet \times S$ , which is used to select a *resumption state* to follow an external state based on the answer provided by the environment. We write  $L : A \rightarrow B$  when  $L$  is a labelled transition system for  $A \rightarrow B$ .

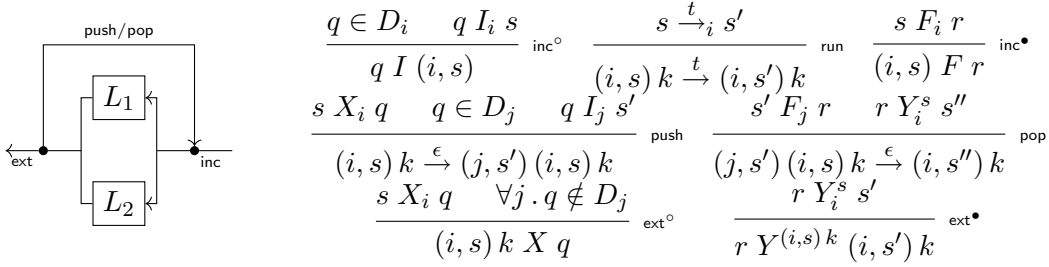


Figure 9.3: Horizontal composition of open semantics. The state is a stack of alternating activations of the two components, initialized as a singleton by an incoming question ( $\text{inc}^\circ$ ). During normal execution ( $\text{run}$ ), the top-level state is updated. Calls into the other component push a new state onto the stack ( $\text{push}$ ), initialized to handle the call in question. If a final state is reached while there are suspended activations ( $\text{pop}$ ), the result is used to resume the most recent one. External calls which are provided by neither component ( $\text{ext}^\circ$ ,  $\text{ext}^\bullet$ ), and final states encountered at the top level ( $\text{inc}^\bullet$ ), are simply passed along to the environment.

We use infix notation for the various transition relations  $I, X, Y, F$ . In particular we write  $n Y^s s'$  to denote that  $n \in A^\bullet$  resumes the suspended external state  $s$  to continue with state  $s'$ . The main reason for treating events  $e \in \mathbb{E}$  and external calls  $mn \in A^\circ A^\bullet$  differently is that while events are expected to be the same between the source and target programs, the form of external calls varies significantly across languages and the simulation convention they follow must be defined explicitly. In addition, while events and event traces bundle together the output and input components of the interaction, our representation of external calls separates them, which simplifies the formulation of horizontal composition and open simulations.

**Horizontal composition** To model linking, we need to express the external behavior of a collection of components in terms of the behaviors of individual components.

Consider the components  $L_1, L_2 : A \multimap A$ . When  $L_1$  is running and performs an external call to one of the functions implemented by  $L_2$ , the execution of  $L_1$  is suspended. The question of  $L_1$  to  $L_2$  is used to initialize a new state for  $L_2$ , and  $L_2$  becomes the active component. Once  $L_2$  reaches a final state, the corresponding answer is used to resume the execution of  $L_1$ . In the process  $L_2$  may itself perform cross-component calls, instantiating *new* executions of  $L_1$ . Therefore, in addition to the state of the active component, we need to maintain a *stack* of suspended states for component instances awaiting resumption. The corresponding transition system is described in Fig. 9.3.

**Definition 9.6** (Horizontal composition). For two transition systems  $L_1, L_2 : A \multimap A$  with  $L_i =$

$\langle S_i, \rightarrow_i, D_i, I_i, X_i, Y_i, F_i \rangle$ , the *horizontal composition* of  $L_1$  and  $L_2$  is defined as:

$$L_1 \oplus L_2 := \langle (S_1 + S_2)^*, \rightarrow, D_1 \cup D_2, I, X, Y, F \rangle$$

where the components  $\rightarrow, I, X, Y, F$  are defined by the rules shown in Fig. 9.3.

### 9.2.2 Open simulations

CompCert is proved correct using a simulation between the transition semantics of the source and target programs. This *forward*<sup>1</sup> simulation is used to establish a *backward* simulation. Backward simulations are in turn proved to be sound with respect to trace containment. We have updated forward and backward simulations to work with CompCertO's semantic model. In this section we present forward simulations, which are used as our primary notion of refinement.

A forward simulation asserts that any transition in the source program has a corresponding transition sequence in the target. The sequence may be empty, but to ensure the preservation of silent divergence this can only happen for finitely many consecutive source transitions. This is enforced by indexing the simulation relation over a well-founded order, and requiring the index to decrease whenever an empty transition sequence is used. This mechanism is unchanged in CompCertO and is largely orthogonal to the techniques we introduce, so we omit this aspect of forward simulations in our exposition below.

**Forward simulations** Our transition systems introduce various forms of external communication, which must be taken into account by our notions of simulation. In CompCertO, a forward simulation between the small-step semantics  $L_1 : A_1 \rightarrow B_1$  and  $L_2 : A_2 \rightarrow B_2$  operates in the context of the simulation conventions  $\mathbb{R}_A : A_1 \Leftrightarrow A_2$  and  $\mathbb{R}_B : B_1 \Leftrightarrow B_2$ .

As depicted in Fig. 9.4, if questions of  $B_1$  and  $B_2$  respectively used to activate  $L_1$  and  $L_2$  are related by the simulation convention  $\mathbb{R}_B$  at a world  $w_B$ , simulations guarantee that the corresponding answers will be related by  $\mathbb{R}_B$  as well: the diagrams can be pasted together horizontally to follow the executions of  $L_1$  and  $L_2$ . Note that the simulation relation  $R \in \mathcal{R}_{w_B}(S_1, S_2)$  is itself indexed by  $w_B$  to ensure that answers are related consistently with the corresponding questions.

<sup>1</sup>In this usage, *forward* pertains to the compilation process, rather than the execution of programs.



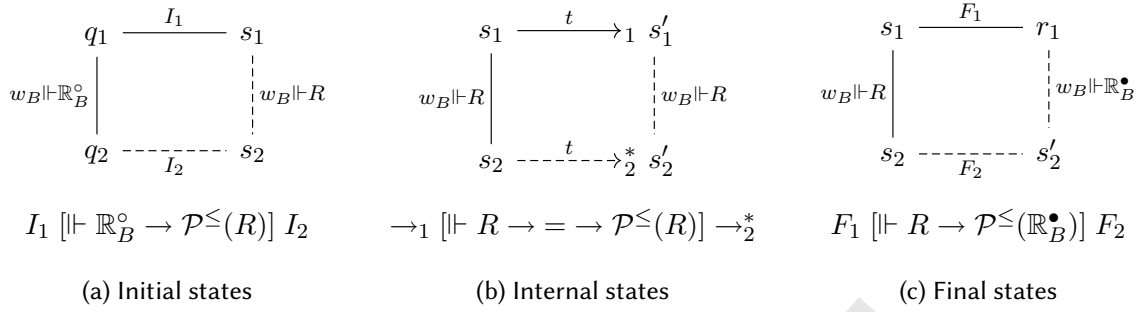


Figure 9.4: Forward simulation properties for initial, internal and final states. (a) When the execution is initiated by two related questions, any initial state of  $L_1$  is matched by a related initial state of  $L_2$ . (b) Every internal transition of  $L_1$  is then matched by a sequence of transitions of  $L_2$ , preserving the simulation relation. (c) When a final state is eventually reached by  $L_1$ , any related state in  $L_2$  is final as well and produces a related answer. The indexing of the relations  $\mathbb{R}_B^\circ$ ,  $R$ ,  $\mathbb{R}_B^\bullet$  by the Kripke world  $w_B$  guarantees that the original questions and their eventual answers are related in a consistent way.

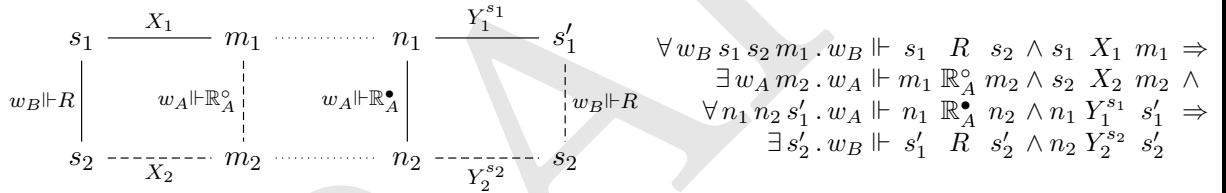


Figure 9.5: Forward simulation property for external states. When  $L_1$  assigns a question  $m_1$  to an external state  $s_1$ ,  $L_2$  assigns a corresponding question  $m_2$  to any related state  $s_2$ . The questions are related according to the simulation convention  $\mathbb{R}_A$ , at a world  $w_A$  chosen by the simulation. When  $L_1$  is resumed by an answer  $n_1$ , then a related answer  $n_2$  also resumes  $L_2$ , and reestablishes the simulation relation between resulting states.

The simulation convention  $\mathbb{R}_A$  determines the correspondence between outgoing questions triggered by the transition systems' external states. The corresponding simulation properties are shown in Fig. 9.5. Compared with the treatment of incoming questions, the roles of the system and environment are reversed: the simulation proof can choose  $w_A$  to relate the outgoing questions, and the environment guarantees that any corresponding answers will be related at that world.

**Definition 9.7** (Forward simulation). Given two simulation conventions  $\mathbb{R}_A : A_1 \Leftrightarrow A_2$  and  $\mathbb{R}_B : B_1 \Leftrightarrow B_2$ , and given the transition systems  $L_1 : A_1 \rightarrow B_1 = \langle S_1, \rightarrow_1, D_1, I_1, X_1, Y_1, F_1 \rangle$  and  $L_2 : A_2 \rightarrow B_2 = \langle S_2, \rightarrow_2, D_2, I_2, X_2, Y_2, F_2 \rangle$ , a *forward simulation* between  $L_1$  and  $L_2$  consists of a relation  $R \in \mathcal{R}_{W_B}(S_1, S_2)$  satisfying the properties shown in Fig. 9.4 and Fig. 9.5. In

addition, the domains of  $L_1$  and  $L_2$  must satisfy:

$$(\lambda q_1 . (q_1 \in D_1)) \llbracket \vdash \mathbb{R}_B^\circ \rightarrow \Leftrightarrow \rrbracket (\lambda q_2 . (q_2 \in D_2))$$

We will write  $L_1 \leq_{\mathbb{R}_A \rightarrow \mathbb{R}_B} L_2$ .

**Horizontal composition** The horizontal composition operator described by Def. 9.6 preserves simulations. Roughly speaking, whenever new component instances are created by cross-component calls, the simulation property for the new components can be stitched in-between the two halves of the callers' simulation property described in Fig. 9.5.

**Theorem 9.8** (Horizontal composition of simulations). *For a simulation convention  $\mathbb{R} : A_1 \Leftrightarrow A_2$  and transition systems  $L_1, L'_1 : A_1 \rightarrow A_1$  and  $L_2, L'_2 : A_2 \rightarrow A_2$ , the following property holds:*

$$\frac{L_1 \leq_{\mathbb{R} \rightarrow \mathbb{R}} L_2 \quad L'_1 \leq_{\mathbb{R} \rightarrow \mathbb{R}} L'_2}{L_1 \oplus L'_1 \leq_{\mathbb{R} \rightarrow \mathbb{R}} L_2 \oplus L'_2}$$

*Proof.* See `common/SmallstepLinking.v` in the Coq development.  $\square$

One interesting and novel aspect of the proof is the way worlds are managed. Externally, only the worlds corresponding to incoming and outgoing questions and answers are observed. Internally, the proof of Thm. 9.8 maintains a *stack* of worlds to relate the corresponding stack of activations in the source and target composite semantics. See also §10.1.7.

Horizontal composition of simulations allows us to decompose the verification of a complex program into the verification of its parts. To establish the correctness of the linked assembly program, we can then use the following result.

**Theorem 9.9.** *Linking Asm programs yields a correct implementation of horizontal composition:*

$$\forall p_1 p_2 . \text{Asm}(p_1) \oplus \text{Asm}(p_2) \leq_{\text{id} \rightarrow \text{id}} \text{Asm}(p_1 + p_2)$$

*Proof.* See `x86/AsmLinking.v` in the Coq development.  $\square$

**Vertical composition** Simulations also compose *vertically*, combining the simulation properties for successive compilation passes into a single one. The convention used by the resulting

simulation can be described as follows.

**Definition 9.10** (Composition of simulation conventions). The composition of two Kripke relations  $R \in \mathcal{R}_{W_R}(X, Y)$  and  $S \in \mathcal{R}_{W_S}(Y, Z)$  is the Kripke relation  $R \cdot S \in \mathcal{R}_{W_R \times W_S}(A, C)$  defined by:

$$(w_R, w_S) \Vdash x [R \cdot S] z \Leftrightarrow \exists y \in Y . w_R \Vdash x R y \wedge w_S \Vdash y R z .$$

Then for the simulation conventions  $\mathbb{R} : A \Leftrightarrow B$  and  $\mathbb{S} : B \Leftrightarrow C$ , we define  $\mathbb{R} \cdot \mathbb{S} : A \Leftrightarrow C$  as:

$$\mathbb{R} \cdot \mathbb{S} := \langle W_{\mathbb{R}} \times W_{\mathbb{S}}, \mathbb{R}^\circ \cdot \mathbb{S}^\circ, \mathbb{R}^\bullet \cdot \mathbb{S}^\bullet \rangle$$

**Theorem 9.11.** *Simulations compose vertically in the way depicted in Fig. 9.2.*

*Proof.* Visually speaking, the diagrams shown in Figs. 9.4 and 9.5 can be pasted vertically. For details, see the Coq proofs `identity_forward_simulation` and `compose_forward_simulations` in the file `common/Smallstep.v`.  $\square$

### 9.3 Simulation convention algebra

Now that we have described the structures we use to model and relate the execution of program components, we shift our attention to the techniques we use to derive a correctness statement for the whole compiler from the correctness properties of each pass.

#### 9.3.1 Refinement of simulation conventions

As discussed in §9.1.4, the composite simulation conventions obtained when we vertically compose the passes of CompCertO are not immediately satisfactory. In the remainder of this section, we introduce the algebraic infrastructure we use to rewrite them into an acceptable form, centered around a notion of refinement for simulation conventions.

A refinement  $\mathbb{R} \sqsubseteq \mathbb{S}$  captures the idea that the convention  $\mathbb{S}$  is more general than  $\mathbb{R}$ , so that any simulation accepting  $\mathbb{S}$  as its incoming convention can accept  $\mathbb{R}$  as well. The shape of the symbol illustrates its meaning: related questions of  $\mathbb{R}$  can be transported to related question of  $\mathbb{S}$ ; when we get a response, the related answers of  $\mathbb{S}$  can be transported back to related answers of  $\mathbb{R}$ .

**Definition 9.12** (Simulation convention refinement). Given two simulation conventions  $\mathbb{R}, \mathbb{S} : A_1 \Leftrightarrow A_2$ , we say that  $\mathbb{R}$  refines  $\mathbb{S}$  and write  $\mathbb{R} \sqsubseteq \mathbb{S}$  when the following holds:

$$\begin{aligned} \forall w m_1 m_2 . w \Vdash m_1 R^\circ m_2 &\Rightarrow \exists v . (v \Vdash m_1 S^\circ m_2 \wedge \\ \forall n_1 n_2 . v \Vdash n_1 S^\bullet n_2 &\Rightarrow w \Vdash n_1 R^\bullet n_2) . \end{aligned}$$

We write  $\mathbb{R} \equiv \mathbb{S}$  when both  $\mathbb{R} \sqsubseteq \mathbb{S}$  and  $\mathbb{S} \sqsubseteq \mathbb{R}$ .

**Theorem 9.13.** For  $\mathbb{R} : A_1 \Leftrightarrow A_2, \mathbb{S} : A_2 \Leftrightarrow A_3, \mathbb{T} : A_3 \Leftrightarrow A_4$ , the following properties hold:

$$(\cdot) :: \sqsubseteq \times \sqsubseteq \rightarrow \sqsubseteq \quad (\mathbb{R} \cdot \mathbb{S}) \cdot \mathbb{T} \equiv \mathbb{R} \cdot (\mathbb{S} \cdot \mathbb{T}) \quad \mathbb{R} \cdot \text{id} \equiv \text{id} \cdot \mathbb{R} \equiv \mathbb{R}$$

In addition, when  $\mathbb{R} \sqsubseteq \mathbb{R}' : A_1 \Leftrightarrow A_2$  and  $\mathbb{S}' \sqsubseteq \mathbb{S} : B_1 \Leftrightarrow B_2$ , for all  $L_1 : A_1 \rightarrow B_1$  and  $L_2 : A_2 \rightarrow B_2$ :

$$L_1 \leq_{\mathbb{R} \rightarrow \mathbb{S}} L_2 \Rightarrow L_1 \leq_{\mathbb{R}' \rightarrow \mathbb{S}'} L_2 .$$

*Proof.* See `open_fsim_ccref` in `common/CallconvAlgebra.v`. □

In CompCert, the passes `Cshngen`, `Renumber`, `Linearize`, `CleanupLabels` and `Debugvar` restrict the source and target memory states and values to be identical. Their simulation proofs require very few changes and can be assigned the convention  $\text{id} \rightarrow \text{id}$  (see Table 9.3). The properties above ensure that these passes have no impact on the overall simulation convention of CompCertO.

### 9.3.2 Kleene algebra

Given a collection of simulation conventions, it is possible to combine them by allowing questions to be related by any one of them. This is the additive operation of our Kleene algebra. The Kleene star combines all possible finite iterations of a simulation convention.

This construction is key in our treatment of asymmetric injection passes ( $\text{injp} \rightarrow \text{inj}$ ). The details of `inj` and `injp` are discussed in §10.1, but are not an essential part of the technique. Schematically, by pre- and post-composing injection passes with self-simulations of types  $\text{injp}^* \rightarrow \text{injp}^*$

Language/Pass	Outgoing $\rightarrow$ Incoming	SLOC	See also
<b>Clight</b>	$\mathcal{C} \rightarrow \mathcal{C}$	+17 (+3%)	
<i>Thm. 10.6</i>	$(\text{ext} + \text{injp})^* \rightarrow (\text{ext} + \text{injp})^*$		§10.1.8
SimplLocals	$\text{injp} \rightarrow \text{inj}$	-4 (-1%)	§10.1.6
Cshmgen	$\text{id} \rightarrow \text{id}$	+0 (+0%)	§9.3.1
<b>Csharpminor</b>	$\mathcal{C} \rightarrow \mathcal{C}$	+15 (+4%)	
Cminorgen	$\text{injp} \rightarrow \text{inj}$	-21 (-2%)	§10.1.6
<b>Cminor</b>	$\mathcal{C} \rightarrow \mathcal{C}$	+27 (+3%)	
Selection	$\text{wt} \cdot \text{ext} \rightarrow \text{wt} \cdot \text{ext}$	+43 (+0%)	§10.2.2
<b>CminorSel</b>	$\mathcal{C} \rightarrow \mathcal{C}$	+15 (+3%)	
RTLgen	$\text{ext} \rightarrow \text{ext}$	+8 (+0%)	§10.1.5
<b>RTL</b>	$\mathcal{C} \rightarrow \mathcal{C}$	+11 (+2%)	
Tailcall <sup>†</sup>	$\text{ext} \rightarrow \text{ext}$	-1 (-1%)	§10.1.5
Inlining	$\text{injp} \rightarrow \text{inj}$	+58 (+3%)	§10.1.6
Renumber	$\text{id} \rightarrow \text{id}$	-14 (-7%)	§9.3.1
<i>Thm. 10.6</i>	$\text{inj} \rightarrow \text{inj}$		§10.1.8
Constprop <sup>†</sup>	$\text{va} \cdot \text{ext} \rightarrow \text{va} \cdot \text{ext}$	-17 (-2%)	§10.2.3
CSE <sup>†</sup>	$\text{va} \cdot \text{ext} \rightarrow \text{va} \cdot \text{ext}$	+3 (+0%)	§10.2.3
Deadcode <sup>†</sup>	$\text{va} \cdot \text{ext} \rightarrow \text{va} \cdot \text{ext}$	-7 (-1%)	§10.2.3
Allocation	$\text{wt} \cdot \text{alloc} \rightarrow \text{wt} \cdot \text{alloc}$	+13 (+0%)	§10.3.1
<b>LTL</b>	$\mathcal{L} \rightarrow \mathcal{L}$	+15 (+6%)	
Tunneling	$\text{ext} \rightarrow \text{ext}$	+2 (+0%)	§10.1.5
Linearize	$\text{id} \rightarrow \text{id}$	-15 (-3%)	§9.3.1
<b>Linear</b>	$\mathcal{L} \rightarrow \mathcal{L}$	+16 (+7%)	
CleanupLabels	$\text{id} \rightarrow \text{id}$	-10 (-4%)	§9.3.1
Debugvar	$\text{id} \rightarrow \text{id}$	-12 (-3%)	§9.3.1
Stacking	$\text{stacking} \rightarrow \text{stacking}$	+291 (+11%)	§10.3.2
<b>Mach</b>	$\mathcal{M} \rightarrow \mathcal{M}$	+100 (+26%)	
Asmggen	$\text{asmgen} \rightarrow \text{asmgen}$	+179 (+6%)	§10.3.3
<b>Asm</b>	$\mathcal{A} \rightarrow \mathcal{A}$	+53 (+5%)	
Total:		+765 (+2%)	

Table 9.3: Languages and passes of CompCertO. Passes are grouped by their source language. <sup>†</sup> indicates optional optimization passes. The simulation conventions  $\text{ext}$ ,  $\text{inj}$ ,  $\text{injp}$  are explained in §10.1. The invariants  $\text{wt}$ ,  $\text{va}$  are explained in §10.2. The conventions  $\text{alloc}$ ,  $\text{stacking}$ ,  $\text{asmgen}$  are explained in §10.3. To handle the asymmetry of injection passes ( $\text{injp} \rightarrow \text{inj}$ ), self-simulation properties are inserted as pseudo-passes (*Thm. 10.6*). Significant lines of code (SLOC) measured by *coqwc*, compared to CompCert v3.6.

and  $\text{inj} \rightarrow \text{inj}$ , we obtain the simulation conventions:

$$\text{injp}^* \cdot \text{injp} \cdot \text{inj} \rightarrow \text{injp}^* \cdot \text{inj} \cdot \text{inj}$$

The property  $\text{injp}^* \cdot \text{injp} \sqsubseteq \text{injp}^*$  on the left-hand side, and the idempotency of  $\text{inj}$  on the right-hand side (Thm. 10.6) allow us to rewrite the above into a symmetric simulation convention.

**Definition 9.14.** Consider  $(\mathbb{R}_i)_{i \in I}$  a family of conventions with  $\mathbb{R}_i = \langle W_i, R_i^\circ, R_i^\bullet \rangle : A_1 \Leftrightarrow A_2$  for all  $i \in I$ . We define the simulation convention  $\sum_{i \in I} \mathbb{R}_i := \langle W, R^\circ, R^\bullet \rangle$ , where:

$$W := \sum_{i \in I} W_i \quad \begin{array}{l} (i, w) \Vdash R^\circ := w \Vdash R_i^\circ \\ (i, w) \Vdash R^\bullet := w \Vdash R_i^\bullet. \end{array}$$

We will write  $\mathbb{R}_1 + \dots + \mathbb{R}_n$  for the finitary case  $\sum_{i=1}^n \mathbb{R}_i$ . Then for  $\mathbb{R} : A_1 \Leftrightarrow A_2$ , we define  $\mathbb{R}^* := \sum_{n \in \mathbb{N}} \mathbb{R}^n$ , where  $\mathbb{R}^0 := \text{id}$  and  $\mathbb{R}^{n+1} := \mathbb{R}^n \cdot \mathbb{R}$ .

**Theorem 9.15** (Kleene iteration of simulations). *The constructions  $\sqsubseteq, \cdot, +, *$  work together as a typed Kleene algebra. Moreover, the following properties hold:*

$$\frac{\forall i. L_1 \leq_{\mathbb{R} \rightarrow \mathbb{S}_i} L_2}{L_1 \leq_{\mathbb{R} \rightarrow \sum_i \mathbb{S}_i} L_2} \quad \frac{L \leq_{\mathbb{R} \rightarrow \mathbb{S}} L}{L \leq_{\mathbb{R}^* \rightarrow \mathbb{S}^*} L}$$

*Proof.* See `cc_star_fsim` and the preceding definitions in `common/CallconvAlgebra.v`.  $\square$

### 9.3.3 Compiler correctness

We can now present our central result. The passes of `CompCertO` are shown in Table 9.3. The techniques outlined above allow us to formulate a simulation convention  $\mathbb{C} : \mathcal{C} \Leftrightarrow \mathcal{A}$  for the compiler, and to establish the following correctness property.

**Theorem 9.16** (Compositional Correctness of `CompCertO`). *For a Clight program  $p$  and an Asm program  $p'$  such that  $\text{CompCert}(p) = p'$ , the following simulation holds:*

$$\text{Clight}(p) \leq_{\mathbb{C} \rightarrow \mathbb{C}} \text{Asm}(p'),$$

where the simulation convention  $\mathbb{C}$  is defined as:

$$\mathbb{C} := (\text{ext} + \text{injp})^* \cdot \text{inj} \cdot (\text{va} \cdot \text{ext})^3 \cdot \text{wt} \cdot \text{alloc} \cdot \text{ext} \cdot \text{stacking} \cdot \text{asmgen}.$$

*Proof.* Use Thm. 9.11 to compose the correctness proofs of the compiler passes and self-simulations shown in Table 9.3. By properties of the Kleene star, the outgoing simulation convention of each of the passes `SimplLocals–Renumber` can be folded into  $(\text{ext} + \text{injp})^*$  to obtain the overall outgoing convention  $\mathbb{C}$ . Likewise, by properties of `inj` and `ext` their incoming simulation conventions can be folded into `inj` to obtain the overall incoming convention  $\mathbb{C}$ . For details, see `driver/Compiler.v`.  $\square$

### 9.3.4 Compositional compilation and verification

Consider the translation units  $M1.c, \dots, Mn.c$  compiled and linked to  $M1.s + \dots + Mn.s = M.s$ .

We can use Thms. 9.8, 9.9 and 9.16 to establish the following separate compilation property:

$$\text{Clight}(M1.c) \oplus \dots \oplus \text{Clight}(Mn.c) \leq_{\mathbb{C} \rightarrow \mathbb{C}} \text{Asm}(M.s) \quad (9.3)$$

That is, the linked `Asm` program  $M.s$  faithfully implements the horizontal composition of the source modules' behaviors, following the simulation convention  $\mathbb{C}$ .

Additionally, suppose we wish to verify that the overall program satisfies a specification  $\Sigma$ , also expressed as a transition system for  $\mathcal{C} \rightarrow \mathcal{C}$ . We must first decompose  $\Sigma$  into:

$$\Sigma \leq_{\text{id} \rightarrow \text{id}} \Sigma_1 \oplus \dots \oplus \Sigma_n.$$

Then for each module, we prove that  $\Sigma_i \leq_{\text{id} \rightarrow \text{id}} \text{Clight}(Mi.c)$ . This can be combined with Thm. 9.8 and Eqn. 9.3 to establish the correctness property  $\Sigma \leq_{\mathbb{C} \rightarrow \mathbb{C}} \text{Asm}(M.s)$ .

Note that Eqn. 9.3 can be established as long as the property  $\text{Clight}(Mi.c) \leq_{\mathbb{C} \rightarrow \mathbb{C}} \text{Asm}(Mi.s)$  holds independently for each module. It is possible for the different  $Mi.s$  to be obtained by different compilers, as long as each one satisfies a correctness property following the simulation convention in Thm. 9.16. Indeed this is the case for versions of `CompCertO` obtained by enabling different optimization passes. Moreover, if some of the  $Mi.s$  are hand-written assembly modules satisfying

Component	SLOC
Semantic framework (§9.2)	+782 (+14%)
Horizontal composition (§9.2.1)	676
Simulation convention algebra (§9.3)	1,052
CKLR theory and instances (§10.1)	1,807
Clight and RTL parametricity (§10.1.5)	2,741
Invariant preservation proofs (§10.2)	+549 (+7%)
Pass correctness proofs (Tbl. 9.3)	+765 (+2%)
<b>Total</b>	<b>8372</b>

Table 9.4: Significant lines of code in CompCertO relative to CompCert v3.6. See Table 9.3 for a per-pass breakdown of the increase in size of pass correctness proofs, and `overhead.py` in the Coq development for the list of files included in each group.

$\mathcal{C}$ -style specifications, then we can prove on a case-by-case basis that  $\Sigma_i \leq_{\mathcal{C} \rightarrow \mathcal{C}} \text{Asm}(\text{Mi.s})$  and proceed with the rest of the proof as before.

Using C functions from arbitrary assembly contexts is also possible, because the compiler’s simulation convention  $\mathbb{C}$  captures all of the guarantees provided by CompCertO and directly specifies the behavior of the compiled assembly code. A proof involving an arbitrary assembly context which invokes a function compiled by CompCertO must establish that the call is performed according to the C calling convention. Then Thm. 9.16 can be used to reason about the behavior of the call in terms of the semantics of the source code or a  $\mathcal{C}$  specification that it satisfies.

## 9.4 Evaluation

To give a sense of the overall complexity of CompCertO, we list in Table 9.4 the increase in significant lines of code it introduces compared to CompCert v3.6. As shown in Table 9.3, our methodology comes with a negligible increase in the complexity of most simulation proofs. Although SLOC is an imperfect measure, and a 1:1 comparison between developments which prove different things is difficult, our numbers represent a drastic improvement over Compositional CompCert, and compare favorably or are on par with the corresponding sections of CompCertM.

Our use of the simulation conventions `injp`, `alloc`, `stacking` and `asmgen` in particular underscores the benefits of our approach. The corresponding passes are the root of much complexity in Compositional CompCert, CompCertX and CompCertM. For instance, to express the requirement on the areas protected by `injp`, both Compositional CompCert and CompCertM introduce



general mechanisms for tracking ownership of different regions of memory as part of an extended notion of memory injection. The approach taken here demonstrates that the requirements placed on external functions by the original CompCert are already good enough for the job! Because our framework is expressive enough to capture them, the corresponding passes barely need any modifications, and the associated issues are resolved before they even show up.

Likewise, the preservation of callee-save registers ensured by the Allocation pass, and the subtle issues associated with argument-passing in the Stacking pass have been the cause of much pain in previous CompCert extensions. The ease with which they are addressed here demonstrates the power of an explicit treatment of abstraction, made possible by our notions of language interface and simulation convention.

## Chapter 10

# Passes of CompCertO

Having described the overall design of CompCertO in Chapter 9, I turn to the details of the techniques I used to update each compilation pass:

- Compositional relational reasoning within CompCertO is explained in §10.1: *CompCert Kripke logical relations* unify CompCert’s memory transformations as structure-preserving relations over the memory model. They can be used to define simulation conventions for most of the compiler’s passes, and to derive parametricity theorems which capture important properties of CompCert languages.
- Several passes of CompCert use typing and soundness invariants. §10.2 explains how these invariants fit into the simulation framework of CompCertO and how the techniques used to verify these passes can be reified into a notion of *simulation modulo invariants*.
- Finally, §10.3 discusses the more specialized simulation conventions used for the passes of CompCert which significantly transform the shape of interactions across compilation units.

The simulation proofs for most of CompCert’s passes can be updated with minimal effort. In the case of more complex passes, simulation conventions can be defined which capture the internal invariants used by existing proofs, avoiding many sources of complexity found in previous work.

## 10.1 Logical relations for the CompCert memory model

In the next few chapters, we examine the correctness properties and simulation conventions at the level of individual passes. In this chapter, we focus on simulations which use *ext*, *inj* and *injp*.

The questions, answers and states used to describe the semantics of CompCert languages all contain a memory state and surrounding runtime values. Likewise, simulation conventions and relations are constructed around *memory transformations* and relate the surrounding components in ways that are compatible with the chosen memory transformation.

### 10.1.1 Memory extensions

For passes where strict equality is too restrictive, but the source and target programs use similar memory layouts, CompCert uses the *memory extension* relation, which allows the values stored in the target memory state to refine the values stored in the source memory at the same location.

By analogy with the value refinement relation  $v_1 \leq_v v_2$  introduced in §8.1.6, we write  $m_1 \leq_m m_2$  to signify that the source memory  $m_1$  is extended by the target memory  $m_2$ . Together, the relations  $\leq_v$  and  $\leq_m$  constitute a logical relation for the memory model, in the sense that loads from memory states related by extension yield values related by refinement, writing values related by refinement preserves memory extension, and similarly for the remaining memory operations.

### 10.1.2 Memory injections

The most complex simulation relations of CompCert allow memory blocks to be dropped, added, or mapped at a given offset within a larger block. These transformations of the memory structure are specified by partial functions of type:

$$\text{meminj} := \text{block} \rightarrow \text{block} \times \mathbb{Z}$$

We will call  $f \in \text{meminj}$  an *injection mapping*. An entry  $f(b) = (b', o)$  means that the source memory block with identifier  $b$  is mapped into the target block  $b'$  at offset  $o$ .

As with refinement and extension, an injection mapping determines both a relation on values and a relation on memory states, which work together as a logical relation for the CompCert

memory model. The relation  $f \Vdash v_1 \hookrightarrow_v v_2$  allows  $v_2$  to refine  $v_1$ , but also requires any pointer present in  $v_1$  to be transformed according to  $f$ . The relation  $f \Vdash m_1 \hookrightarrow_m m_2$  requires that the corresponding addresses of  $m_1$  and  $m_2$  hold values that are related by  $f \Vdash \hookrightarrow_v$ .

Note that corresponding memory allocations in the source and target states cause  $f$  to evolve into a more defined injection mapping  $f \subseteq f'$  relating the two newly allocated blocks. To handle this, we introduce the following constructions.

### 10.1.3 Modal Kripke relators

We defined general Kripke relations in §???. We add structure to sets of Kripke worlds, specifying how they can evolve.

**Definition 10.1.** A Kripke frame is a tuple  $\langle W, \rightsquigarrow \rangle$ , where  $W$  is a set of *possible worlds* and  $\rightsquigarrow$  is a binary *accessibility relation* over  $W$ . Then the Kripke relator  $\Diamond$  is defined by:

$$w \Vdash x [\Diamond R] y \Leftrightarrow \exists w'. w \rightsquigarrow w' \wedge w' \Vdash x R y$$

**Example 10.2** (Injection simulation diagrams). Building on Ex. 2.12, consider once again the simple transition systems  $\alpha : A \rightarrow \mathcal{P}(A)$  and  $\beta : B \rightarrow \mathcal{P}(B)$ . An injection-based simulation relation between them will be a Kripke relation  $R \in \mathcal{R}_{\text{meminj}}(A, B)$  satisfying the property:

$$\begin{array}{ccc}
 s_1 & \xrightarrow{\alpha} & s'_1 \\
 f \Vdash R \downarrow & & \downarrow f' \Vdash R \\
 s_2 & \xrightarrow{\beta} & s'_2
 \end{array}
 \quad (f \subseteq f')
 \quad \forall f \, s_1 \, s_2 \, s'_1. f \Vdash s_1 R s_2 \wedge \alpha(s_1) \ni s'_1 \Rightarrow$$

$$\exists f' \, s'_2. f \subseteq f' \wedge \beta(s_2) \ni s'_2 \wedge f' \Vdash s'_1 R s'_2$$

(10.1)

The new states may be related according to a new injection mapping  $f'$ , but in order to preserve existing relationships between any surrounding source and target pointers, the new mapping must include the original one ( $f \subseteq f'$ ). This pattern is very common in CompCerto and appears in a variety of contexts. By using  $\langle \text{meminj}, \subseteq \rangle$  as a Kripke frame, we can express (10.1) as:

$$\alpha [\Vdash R \rightarrow \mathcal{P}^\subseteq(\Diamond R)] \beta.$$

$$\begin{aligned}
& \rightsquigarrow \text{ is reflexive and transitive} \\
& w \rightsquigarrow w' \Rightarrow f(w) \subseteq f(w') \\
\\
& \text{alloc} :: \Vdash R^{\text{mem}} \rightarrow = \rightarrow = \rightarrow \Diamond(R^{\text{mem}} \times R^{\text{block}}) \\
& \text{free} :: \Vdash R^{\text{mem}} \rightarrow R^{\text{ptrrange}} \rightarrow \text{option}^{\leq}(\Diamond R^{\text{mem}}) \\
& \text{load} :: \Vdash R^{\text{mem}} \rightarrow R^{\text{ptr}} \rightarrow \text{option}^{\leq}(R^{\text{val}}) \\
& \text{store} :: \Vdash R^{\text{mem}} \rightarrow R^{\text{ptr}} \rightarrow R^{\text{val}} \rightarrow \text{option}^{\leq}(\Diamond R^{\text{mem}})
\end{aligned}$$

Figure 10.1: Defining properties of CKLRs. Note the correspondence with the types of operations in Fig. 8.2.

#### 10.1.4 CompCert Kripke logical relations

We can now formalize the idea that extensions and injections constitute logical relations for the CompCert memory model.

**Definition 10.3** (CompCert Kripke Logical Relation). For a tuple  $R = (W, \rightsquigarrow, f, R^{\text{mem}})$ , where  $\langle W, \rightsquigarrow \rangle$  is a Kripke frame,  $f : W \rightarrow \text{meminj}$  associates an injection mapping to each world, and where  $R^{\text{mem}} \in \mathcal{R}_W(\text{mem})$  is a Kripke relation on memory states, the Kripke relations  $R^{\text{ptr}} \in \mathcal{R}_W(\text{ptr})$  and  $R^{\text{ptrrange}} \in \mathcal{R}_W(\text{ptrrange})$  are defined by the rules:

$$\frac{f_w(b) = (b', \delta)}{w \Vdash (b, o) R^{\text{ptr}} (b', o + \delta)} \quad \frac{w \Vdash (b_1, l_1) R^{\text{ptr}} (b_2, l_2) \quad h_1 - l_1 = h_2 - l_2}{w \Vdash (b_1, l_1, h_1) R^{\text{ptrrange}} (b_2, l_2, h_2)}$$

and  $R^{\text{val}} \in \mathcal{R}_W(\text{val})$  is the smallest Kripke relation satisfying:

$$\begin{aligned}
& \forall v \in \text{val} . \Vdash \text{undef } R^{\text{val}} v \quad \text{vptr} :: \Vdash R^{\text{ptr}} \rightarrow R^{\text{val}} \\
& \text{int, long, float, single} :: \Vdash = \rightarrow R^{\text{val}} .
\end{aligned}$$

We say that  $R$  is a *CompCert Kripke logical relation* if the properties shown in Fig. 10.1 are satisfied.

**Rationale** The relation  $R^{\text{mem}}$  is given. We expect  $R^{\text{ptr}}$  to map each source pointer to at most one target pointer and to be shift-invariant in the following sense:

$$\frac{w \Vdash (b_1, o_1) R^{\text{ptr}} (b_2, o_2)}{w \Vdash (b_1, o_1 + \delta) R^{\text{ptr}} (b_2, o_2 + \delta)}$$

Any such relation can be uniquely specified by the injection mapping  $f$ . We expect the other relations to be consistent with  $R^{\text{ptr}}$  and  $\text{undef}$  to act as a least element for  $R^{\text{val}}$ , which determines them completely.

Note that  $R^{\text{mem}}$  is the central component driving world transitions, as witnessed by the uses of  $\Diamond$  in Fig. 10.1. The surrounding relations are monotonic in  $w$ , so that any extra state constructed from pointers and runtime values will be able to “follow along” when world transitions occur.

**Theorem 10.4.** *Extensions and injections correspond to the CompCert Kripke logical relations:*

$$\text{ext} := \langle \{*\}, \{(*, *)\}, * \mapsto (b \mapsto (b, 0)), \leq_m \rangle \quad \text{inj} := \langle \text{meminj}, \subseteq, f \mapsto f, \hookrightarrow_m \rangle$$

*Proof.* The correspondence between  $R_{\text{inj}}^{\text{val}}$  and  $\hookrightarrow_v$  is easily verified, as is the correspondence between  $* \Vdash R_{\text{ext}}^{\text{val}}$  and  $\leq_v$ . The properties of Fig. 10.1 reduce to well-known properties of the memory model already proven in CompCert. See `cklr/Extends.v` and `cklr/Inject.v` for details.  $\square$

### 10.1.5 From CKLRs to simulation conventions

In simulations, the accessibility relation allows us to update the world after each step in the program’s execution. Transitivity allows us to combine multiple steps in one:

$$q \cdot \overset{\curvearrowright}{s_1} \cdot \overset{\curvearrowright}{s_2} \cdots \overset{\curvearrowright}{s_k} \cdot r \quad \Rightarrow \quad q \cdot \overset{\curvearrowright}{s_1 \cdot s_2 \cdots s_k} \cdot r$$

In our approach to simulation conventions, the accessibility relation is not part of the interface of simulations. Instead, a single world is used to formulate the 4-way relationship between pairs of questions and answers. As shown below, in the case of simulation conventions based on CKLRs, this relation does involve the accessibility relations which CKLRs introduce.

Given a specific language interface  $\mathcal{X}$ , the components of any CKLR  $R = \langle W, \rightsquigarrow, f, R^{\text{mem}} \rangle$  can be used to construct a simulation relation  $R_{\mathcal{X}} : \mathcal{X} \Leftrightarrow \mathcal{X}$ . For instance:

$$R_{\mathcal{C}} := \langle W, (R^{\text{val}} \times = \times \vec{R}^{\text{val}} \times R^{\text{mem}}), \Diamond(R^{\text{val}} \times R^{\text{mem}}) \rangle.$$

We will often implicitly promote  $R$  to  $R_{\mathcal{C}}$ . Furthermore, since the semantics of CompCert languages are built out of the operations of the memory model, they are well-behaved with respect

to CKLRs and we can prove the following parametricity theorems.

**Theorem 10.5** (Relational parametricity of Clight and RTL). *For all programs  $p$  and CKLR  $R$ :*

$$\text{Clight}(p) \leq_{R \Rightarrow R} \text{Clight}(p) \quad \text{RTL}(p) \leq_{R \Rightarrow R} \text{RTL}(p)$$

*Proof.* See `cklr/*rel.v` in the Coq development, in particular `Clightrel.v` and `RTLrel.v`.  $\square$

The passes of CompCert which use memory extensions do not feature complex invariants which must be preserved at call sites; it is enough for external calls to preserve the memory extension. Consequently, they are not much more difficult to update than identity passes, and can be assigned the type  $\text{ext} \rightarrow \text{ext}$ . By contrast, injection passes are trickier to handle.

### 10.1.6 External calls in injection passes

Passes which alter the block structure of the memory use *memory injections* (§10.1.2). The convention `inj` can be used for incoming calls, but it is insufficient for outgoing calls.

Consider the `SimplLocals` pass, which removes some local variables from the memory. The corresponding values are instead stored as temporaries in the target function’s local environment, and the correspondence between the two is enforced by the simulation relation. To maintain it, we need to know that external calls do not modify the corresponding source memory blocks.

More generally, as depicted in Fig. 10.2, injection passes expect external calls to leave regions outside of the injection’s footprint untouched. This expectation is reasonable because external calls should behave uniformly between the source and target executions. These requirements can be formalized as the following CompCert Kripke logical relation:

$$\text{injp} := \langle \text{meminj} \times \text{mem} \times \text{mem}, \rightsquigarrow_{\text{injp}}, \pi_1, R_{\text{injp}}^{\text{mem}} \rangle \quad \frac{f \Vdash m_1 \hookrightarrow_m m_2}{(f, m_1, m_2) \Vdash m_1 R_{\text{injp}}^{\text{mem}} m_2}$$

where  $(f, m_1, m_2) \rightsquigarrow_{\text{injp}} (f', m'_1, m'_2)$  ensures that  $f \subseteq f'$  and that the memory states satisfy the constraints in Fig. 10.2 (for details, see `cc_injp` in `common/LanguageInterface.v`).

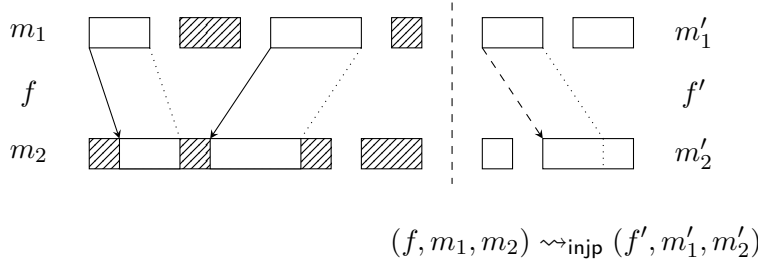


Figure 10.2: External calls and memory injections. The source and target memory states are depicted at the top and bottom of the figure. Arrows describe the injection mapping. The memory block on the left of the dashed line are present at the beginning of the call. Memory blocks on the right are allocated during the call, adding a new entry to the injection mapping. The shaded areas must not be modified by the call.

### 10.1.7 Discussion: world transitions and compositionality

The *injp* convention illustrates a key novelty in the granularity at which we deploy Kripke worlds. In previous work, Kripke worlds are usually assumed to evolve linearly with the execution. Writing  $s_i$  for internal states, this can be depicted as:

$$q \cdot \overset{\curvearrowright}{s_1} \cdot \overset{\curvearrowright}{s_2} \cdot \overset{\curvearrowright}{m} \cdot \overset{\curvearrowright}{s'_1} \cdot \overset{\curvearrowright}{s'_2} \cdot \overset{\curvearrowright}{n} \cdot \overset{\curvearrowright}{s_3} \cdots \overset{\curvearrowright}{s_k} \cdot r$$

To enable horizontal compositionality, the challenge is then to construct worlds, accessibility relations, and simulation relations which are sophisticated enough to express ownership constraints like the ones discussed in §10.1.6, which evolve and shift as the execution switches between components.

In our open simulations, worlds can be deployed independently for incoming and outgoing calls, in a way which follows the structure of plays, as depicted here:

$$q \cdot \overset{\curvearrowright}{m_1} \cdot \overset{\curvearrowright}{n_1} \cdots \overset{\curvearrowright}{m_k} \cdot \overset{\curvearrowright}{n_k} \cdot r$$

Internal steps are not part of a component's observable behavior, and individual simulation proofs are free to use any simulation relation establishing the simulation convention at interaction sites.

Two examples illustrate this flexibility. First, as explained in §9.2.2, to handle nested cross-component calls, composite simulations use an internal stack of worlds. A situation where  $m_1$



and  $m_2$  are cross-component calls and  $m_3$  is an external call can be described as:

$$q \cdot m_1 \cdot m_2 \cdot m_3 \cdot n_3 \cdot n_2 \cdot n_1 \cdot r$$

Second, in simulations which use CKLRs, the simulation relation can simply be qualified as  $w_B \Vdash \Diamond R$  to allow the world to evolve as the execution progresses. The corresponding shape is:

$$q \cdot s_1 \cdot s_2 \cdot s_3 \cdots s_k \cdot r$$

Since  $\Diamond\Diamond R = \Diamond R$ , per-step world transitions can easily be folded into the overall constraint. Moreover, this allows steps which *do not* individually conform to world transitions ( $\Vdash R \rightarrow \Diamond R$ ), but *do* maintain  $\Diamond R$  with respect to the initial world ( $w_B \Vdash \Diamond R \rightarrow \Diamond R$ ).

For instance, the simulation convention of the Stacking pass is based on  $\text{injp} \rightarrow \text{injp}$ . Stacking stores the contents of some temporaries used by the source program into *spilling locations* of the target in-memory stack frames. To prove correctness, we must ensure that spilling locations are only accessed as intended, by enforcing their separation from the injected source memory. This property is maintained by  $\rightsquigarrow_{\text{injp}}$ , which most internal steps and external calls conform to. On the other hand, internal steps which *do* access spilling locations in the expected way do not conform to  $\rightsquigarrow_{\text{injp}}$  at a granular level. However, since the stack frame is a *new* memory block allocated after a function is called, these steps do maintain  $\rightsquigarrow_{\text{injp}}$  with respect to the initial world. This allows us to encode CompCert's original and "instantaneous" assumptions about external calls directly, and existing simulation proofs relying on them can be updated with almost no changes.

Combined together, the two examples above are sufficient to express ownership constraints which require sophisticated permission maps in other approaches, by using conditions already present in CompCert.

### 10.1.8 Properties

Finally, we state some properties which are used to derive Thm. 9.16.

**Theorem 10.6.** *For all Clight and RTL programs:*

$$\forall p. \text{Clight}(p) \leq_{(\text{ext}+\text{inj})^* \multimap (\text{ext}+\text{inj})^*} \text{Clight}(p)$$

$$\forall p. \text{RTL}(p) \leq_{\text{inj} \multimap \text{inj}} \text{RTL}(p)$$

*In addition, the simulation relations derived from ext and inj compose in the following way:*

$$\text{ext} \cdot \text{inj} \equiv \text{inj} \cdot \text{ext} \equiv \text{inj} \cdot \text{inj} \equiv \text{inj}.$$

*Proof.* The first statement is derived from Thms. 9.15 and 10.5; see `driver/Compiler.v`. For the second statement, see `ext_inj`, `inj_ext`, `inj_inj` found under `cklr/`.  $\square$

## 10.2 Invariants

Several passes of CompCert rely on the preservation of invariants by their source program (this situation is illustrated in more detail in the supplementary appendix): when the semantics of a language preserves an invariant, the preservation properties can assist in proving forward simulations which use the language as their source. This allows us to decompose the simulation proof, and in the case of RTL the preservation proofs can be reused for multiple passes.

In CompCert, this technique is deployed in an ad-hoc manner: for each pass using an invariant, the simulation relation is strengthened to assert that the invariant holds on the source state, and the preservation properties for the source language are used explicitly in the simulation proof to maintain this invariant. In CompCertO, this becomes more involved, because the simulation convention must be altered to ensure that invariants are preserved by external calls.

On the other hand, our simulation infrastructure offers the opportunity to capture and reason about invariants explicitly, and to further decouple preservation and simulation proofs. In this section, we give an overview of our treatment of invariants. For details, see the supplementary appendix and `common/Invariant.v` in the Coq development.

### 10.2.1 Invariants and language interfaces

First, we define a sort of “invariant convention”, which describes how a given invariant impacts the questions and answers of the language under consideration.

**Definition 10.7.** An invariant for a language interface  $A$  is a tuple  $\mathbb{P} = \langle W, \mathbb{P}^\circ, \mathbb{P}^\bullet \rangle$ , where  $W$  is a set of worlds and  $\mathbb{P}^\circ, \mathbb{P}^\bullet$  are families of predicates on  $A^\circ, A^\bullet$  indexed by  $W$ .

**Example 10.8.** Typing constraints for the language interface  $\mathcal{C}$  can be expressed as the invariant:

$$\text{wt} := \langle \text{sig}, \mathbb{P}_{\text{wt}}^\circ, \mathbb{P}_{\text{wt}}^\bullet \rangle \quad \frac{\vec{v} <: \text{sg.args}}{\text{sg} \Vdash \text{vf}[\text{sg}](\vec{v}) @ m \in \mathbb{P}_{\text{wt}}^\circ} \quad \frac{v' <: \text{sg.res}}{\text{sg} \Vdash v' @ m' \in \mathbb{P}_{\text{wt}}^\bullet}$$

The proposition  $\vec{v} <: \text{sg.args}$  asserts that the types of the arguments  $\vec{v}$  match those specified by the signature  $\text{sg}$ . The proposition  $v' <: \text{sg.res}$  asserts a similar property for the return value  $v'$ .

Invariants can be seen as a special case of simulation convention which constrain the source and target questions and answers to be equal. This can be formalized as follows.

**Definition 10.9** (Simulation conventions for invariants). A  $W$ -indexed predicate  $P$  on a set  $X$  can be promoted to a Kripke relation  $\hat{P} \in \mathcal{R}_W(X, X)$  defined by the rule:

$$\frac{w \Vdash x \in P}{w \Vdash x \hat{P} x}$$

Then an invariant  $\mathbb{P} = \langle W, \mathbb{P}^\circ, \mathbb{P}^\bullet \rangle$  can be promoted to a simulation convention:  $\hat{\mathbb{P}} := \langle W, \hat{\mathbb{P}}^\circ, \hat{\mathbb{P}}^\bullet \rangle$ .

### 10.2.2 Typing invariants

The typing invariant described in Ex. 10.8 is used by the Selection and Allocation passes. We have updated their correctness proofs as well as the preservation proofs in `Cminortyping.v` and `RTLtyping.v` to use our framework.

The invariant `wt` satisfies one key property: when a simulation convention  $\mathbb{R}$  consists of a sequence of CKLRs and other invariants, the following property holds:

$$\text{wt} \cdot \mathbb{R} \cdot \text{wt} \equiv \mathbb{R} \cdot \text{wt}$$

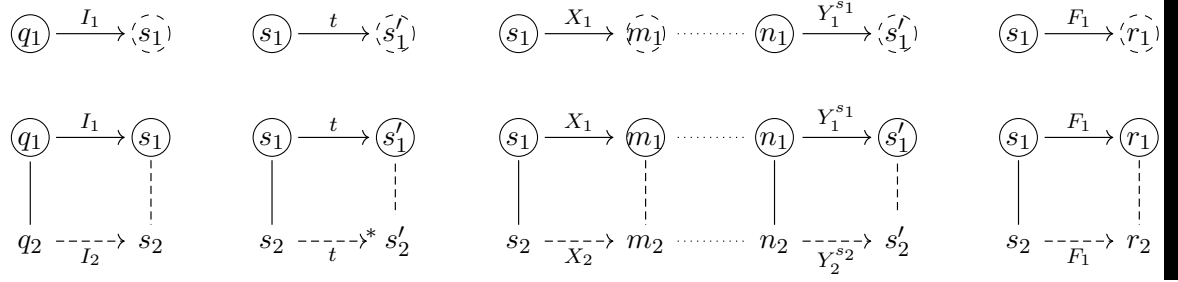


Figure 10.3: Simulation with invariants. Circles indicate questions, answers and states which satisfy the appropriate invariants. When the transition system  $L_1$  preserves the invariants in the way shown in the top row, a simulation of  $L_1$  by  $L_2$  can be established through the weakened diagrams shown in the bottom row. The resulting simulation uses the convention  $\mathbb{P}_A \cdot \mathbb{R}_A \rightarrow \mathbb{P}_B \cdot \mathbb{R}_B$ , ensuring that the environment establishes and preserves the appropriate invariants on questions and answers. The simulation relation  $P \cdot R$  then ensures that the strengthened assumptions used by the weakened simulation diagrams can be satisfied.

This means CompCertO’s overall simulation convention can eliminate the typing invariant for the Selection pass, retaining only that used for Allocation. In turn, this facilitates the simplification of the convention for the passes from Clight to Inlining.

### 10.2.3 Value analysis

The passes Constprop, CSE and Deadcode use CompCert’s value analysis framework. Abstract interpretation is performed on their source program, and the resulting information is used to carry out the corresponding optimizations. The correctness proofs for these passes then rely on the invariant  $va$ , which asserts that the concrete runtime states satisfy the constraints encoded in the corresponding abstract states.

We have updated the value analysis framework and associated pass correctness proofs to fit the invariant infrastructure described in this section. Value analysis passes use the convention  $va \cdot \text{ext} \rightarrow va \cdot \text{ext}$ . Unfortunately, because it combines constraints with mixed variance, the invariant  $va$  does not propagate in the same way as  $wt$ , so the compiler’s simulation convention must retain the component  $(va \cdot \text{ext})^3$  as-is. When some of the corresponding optimization passes are disabled, we use self-simulations of the RTL language to match this convention nonetheless.

### 10.2.4 Simulations modulo invariants

The top row in Fig. 10.3 illustrates the preservation of invariants by transition systems. In the context of a transition system  $L_1 : A_1 \rightarrow B_1$ , we consider three invariants working together:

- an invariant  $\mathbb{P}_A$  for the language interface  $A$ ;
- an invariant  $\mathbb{P}_B$  for the language interface  $B$ ;
- a  $W_B$ -indexed predicate  $P$  on the states of  $L_1$ .

The preservation of  $\langle \mathbb{P}_A, \mathbb{P}_B, P \rangle$  is then analogous to a unary simulation property, where  $\mathbb{P}_A \rightarrow \mathbb{P}_B$  play the roles of the simulation conventions, and  $P$  plays the role of the simulation relation. In fact, when  $L_1$  preserves these invariants, the following property holds:

$$L_1 \leq_{\hat{\mathbb{P}}_A \rightarrow \hat{\mathbb{P}}_B} L_1$$

Once we have established that the source language preserves the invariants, we wish to use this fact to help prove the forward simulation for a given pass. To this end, we define a *strengthened* transition system  $L_1^{\mathbb{P}} : A_1 \rightarrow B_1$ , with the property that  $L_1 \leq_{\hat{\mathbb{P}}_A \rightarrow \hat{\mathbb{P}}_B} L_1^{\mathbb{P}}$ . For a target transition system  $L_2 : A_2 \rightarrow B_2$ , it then suffices to show that  $L_1^{\mathbb{P}} \leq_{\mathbb{R}_A \rightarrow \mathbb{R}_B} L_2$  to establish:

$$L_1 \leq_{\hat{\mathbb{P}}_A \cdot \mathbb{R}_A \rightarrow \hat{\mathbb{P}}_B \cdot \mathbb{R}_B} L_2 .$$

Simulations from  $L_1^{\mathbb{P}}$  are easier to prove, because  $L_1^{\mathbb{P}}$  provides assumption that the invariants hold on all source questions, answers and states. The simulation diagrams reduce to those shown in the bottom row of Fig. 10.3. However, since they are formulated in terms of Def. 9.7, the standard forward simulation techniques defined by CompCert in `Smallstep.v` remain available.

## 10.3 Specialized simulation conventions

For the Alloc, Stacking and Asmgcn passes, we construct more specific simulation conventions which express the correspondence between the higher-level and lower-level representations of function calls and returns. These passes use identical conventions for incoming and external calls.

### 10.3.1 The Allocation pass

The Allocation pass from RTL to LTL is the first pass to modify the interface of function calls. LTL uses *abstract locations* which represent the stack slots and machine registers eventually used in the target assembly program. Abstract locations contain arguments, temporaries and return values. The contents are stored in a *location map*, passed across components by the interface  $\mathcal{L}$  alongside memory states. The compiler also expects the values of abstract locations designated as *callee-save* to be preserved by function calls.

To express the simulation convention used by the Allocation pass, we will use the following notations. For a signature  $sg$  and a location map  $ls$ , we write  $\text{args}(sg, ls)$  to represent the argument values stored in  $ls$ . Likewise,  $\text{retval}(sg, ls)$  extracts the contents of locations used to store the return value. The relation  $\equiv_{CS}$  asserts that two location maps agree on callee-save locations.

The simulation convention  $\text{alloc} : \mathcal{C} \Leftrightarrow \mathcal{L}$  uses its worlds to remember the initial location map and the signature associated with a call, and can then be defined by:

$$\begin{aligned} \text{alloc} &:= \langle \text{signature} \times \text{locmap}, R_{\text{alloc}}^{\circ}, R_{\text{alloc}}^{\bullet} \rangle \\ &\frac{\vec{v} \leq_v \text{args}(sg, ls) \quad m_1 \leq_m m_2}{(sg, ls) \Vdash vf[sg](\vec{v})@m_1 \quad R_{\text{alloc}}^{\circ} \quad vf[sg](ls)@m_2} \\ &\frac{v' \leq_v \text{retval}(sg, ls') \quad ls \equiv_{CS} ls' \quad m'_1 \leq_m m'_2}{(sg, ls) \Vdash v'@m'_1 \quad R_{\text{alloc}}^{\bullet} \quad ls'@m'_2} \end{aligned}$$

### 10.3.2 The Stacking pass

The Stacking pass consolidates the information which Linear stores in abstract stack locations into in-memory stack frames. The simulation proof uses a memory injection, and involves maintaining separation properties ensuring that the source memory and the regions of stack frames introduced by Stacking occupy disjoint areas of the target memory.

With regards to the memory state, the stacking simulation convention is essentially identical to  $\text{injp}$ . Since the new regions of stack frames are outside the image of source memory, and most of them are local to function activations, the properties of  $\text{injp}$  are largely sufficient (see also §10.1.7).

The exception to the rule pertains to argument passing. Loads from a function's argument locations access the caller's stack frame. If the area used to store arguments overlaps with the

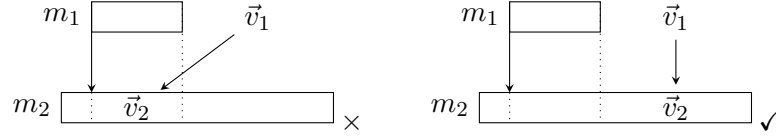


Figure 10.4: Separation of arguments in the stacking simulation convention.

injected source memory state, then the source program and external calls may alter them in unexpected ways. In previous CompCert extensions, sophisticated techniques were required to prevent this from happening.

In our model, we simply encode the required separation condition in the simulation convention  $\text{stacking} : \mathcal{L} \Leftrightarrow \mathcal{M}$ . The convention asserts that the contents of argument locations are stored into corresponding stack slots within the target memory. Additionally, it requires that the region of the target memory used to store arguments within the caller’s stack must be disjoint from the injection image of the source memory (Fig. 10.4). For details, see `cc_stacking` in `backend/Mach.v`.

### 10.3.3 The Asmgén pass

The Asmgén pass from Mach to Asm uses a memory extension. Asm introduces explicit registers for the program counter, stack pointer and return address. The corresponding simulation convention  $\text{asmgen} : \mathcal{M} \Leftrightarrow \mathcal{A}$  ensures that the appropriate components of Mach-level queries are mapped to the new registers. In addition, we must ensure that the call returns the stack pointer to its original value, and set the program counter to the return address specified by the caller.

[XXX: we don’t talk about nextblock in our simplified exposition of the memory model] A more complex challenge is that the Asm language does not have a control stack, but instead executes instruction after instruction in a “flat” manner, making it difficult to distinguish final states. To address this, we keep track of the initial value  $nb_0$  of the memory’s nextblock counter, and use it to distinguish between inner and outer stack pointers. When  $rs[\text{sp}] \geq nb_0$ , we interpret the `ret` instruction as an *internal* return and simply jump to the location pointed to by  $rs[\text{ra}]$ . However, when  $rs[\text{sp}] < nb_0$ , we interpret `ret` as a top-level return to the environment, and model its behavior as a final state rather than an internal step. In addition, the simulation convention `asmgen` must ensure that the initial value of the stack pointer points to a valid block of the initial memory.

This is expressed as:

$$\text{asmgen} := \langle \text{val} \times \text{val}, R_{\text{asmgen}}^{\circ}, R_{\text{asmgen}}^{\bullet} \rangle$$

$$\frac{rs_1 \uplus [\text{sp} := sp, \text{ra} := ra, \text{pc} := vf] \leq_v rs_2 \quad m_1 \leq_m m_2}{(sp, ra) \Vdash vf(sp, ra, rs_1) @ m_1 \quad R_{\text{asmgen}}^{\circ} \quad rs_2 @ m_2}$$

$$\frac{rs'_1 \uplus [\text{sp} := sp, \text{pc} := ra] \leq_v rs'_2 \quad m'_1 \leq_m m'_2}{(sp, ra) \Vdash rs'_1 @ m'_1 \quad R_{\text{asmgen}}^{\bullet} \quad rs'_2 @ m'_2}$$



## **Part IV**

### **Conclusion**

## Chapter 11

# Conclusion and future work

The goal of this thesis is to demonstrate that the concepts and formalisms which underpin game semantics, the refinement calculus, algebraic effects and other areas of programming language theory can be connected and unified in a fruitful manner. Their synthesis, which I call *refinement-based game semantics*, constitutes a plausible approach to the end-to-end verification of large-scale, heterogenous computer systems.

Taking first steps in this direction, I have shown how a simple theory of certified abstraction layers can be embedded into increasingly expressive categorical models, and how the conceptual framework of refinement-based game semantics can be used to integrate CompCert into this hierarchy with minimal effort. A preliminary formalization in the Coq proof assistant of the game models presented in Part II demonstrates the advantages of the techniques I have used [XXX enumerate]. The work on CompCertO presented in Part III has been fully formalized in the Coq proof assistant.

### 11.1 Categorical characterization of $\mathcal{I}_E$

Although I have informally described the interaction specification monad  $\mathcal{I}_E$  as a variant of the free monad on  $E$ , I have not given an explicit characterization of it in terms of categories and adjunctions.

$$\chi : \sum_{m:N \in E} X^N \rightarrow X$$

$X$  completely distributive lattice,  $\chi$  preserves infs and non-empty sups.

Difficulties to formalize it like that:

- Coproducts of lattices, known as *free products* in the literature [?] are difficult to describe in general. Note however that  $\text{FCD}(A) + \text{FCD}(B) \simeq \text{FCD}(A + B)$ .
- $\chi$  does not necessarily preserve  $\perp$ , so not a homomorphism in **CDLat**. If we want to express  $\chi$  as a complete homomorphism, we can reformulate as:

$$\chi : \sum_{m:N \in E} X_{\perp}^N \rightarrow X$$

or come up with a new category **CDLat'**. Equivalent through an adjunction which adds bottom. The new bottom can be thought of as “global failure”, for example a module that can’t be constructed, which propagates to the whole system and all time.

Also relates to typical game semantics requirement that  $\epsilon \in \sigma$ , which completes a pointed poset instead of general poset.

## 11.2 Generalizations of $\mathcal{I}_E$

## 11.3 Richer game models

## 11.4 Linear logic

The **Sup** model of linear logic. Product is as usual. Terminal object  $1 = \{\perp = \top\}$ , initial object  $0 = \{\perp \sqsubseteq \top\}$ , tensor unit  $I = \{\perp \sqsubseteq * \sqsubseteq \top\}$ . Tensor is contravariant galois connection, or a version of product that identifies the  $\perp$ s. Homomorphism  $x : I \rightarrow L$  picks out an element of  $L$ .

## 11.5 Parametric language for CompCert memory model

# Bibliography

- Samson Abramsky. 2010. From CSP to Game Semantics. In *Reflections on the Work of C.A.R. Hoare*. Springer, London, 33–45. [https://doi.org/10.1007/978-1-84882-912-1\\_2](https://doi.org/10.1007/978-1-84882-912-1_2)
- Samson Abramsky, Kohei Honda, and Guy Mccusker. 1998. A Fully Abstract Game Semantics for General References. In *Proceedings of the Thirteenth Annual IEEE Symposium on Logic in Computer Science*. Society Press, 334–344.
- Samson Abramsky, Radha Jagadeesan, and Pasquale Malacaria. 2000. Full Abstraction for PCF. *Inf. Comput.* 163, 2 (2000), 409–470. <https://doi.org/10.1006/inco.2000.2930>
- Samson Abramsky and Guy McCusker. 1997. Linearity, Sharing and State: A Fully Abstract Game Semantics for Idealized Algol with Active Expressions. In *Algol-like Languages*. Birkhäuser, Boston, MA, 297–329. [https://doi.org/10.1007/978-1-4757-3851-3\\_10](https://doi.org/10.1007/978-1-4757-3851-3_10)
- Samson Abramsky and Guy McCusker. 1999. Game semantics. In *Computational logic: Proceedings of the 1997 Marktoberdorf Summer School*. Springer, Berlin, Heidelberg, 1–55. [https://doi.org/10.1007/978-3-642-58622-4\\_1](https://doi.org/10.1007/978-3-642-58622-4_1)
- Rajeev Alur, Thomas A. Henzinger, Orna Kupferman, and Moshe Y. Vardi. 1998. Alternating refinement relations. In *Proceedings of the 9th International Conference on Concurrency Theory (CONCUR '98)*. Springer, Berlin, Heidelberg, 163–178. <https://doi.org/10.1007/BFb0055622>
- Andrew W. Appel. 2011. Verified Software Toolchain. In *Proceedings of the 20th European Symposium on Programming (ESOP 2011)*. Springer, Berlin, Heidelberg, 1–17. [https://doi.org/10.1007/978-3-642-19718-5\\_1](https://doi.org/10.1007/978-3-642-19718-5_1)

- Andrew W Appel, Lennart Beringer, Adam Chlipala, Benjamin C Pierce, Zhong Shao, Stephanie Weirich, and Steve Zdancewic. 2017. Position paper: the science of deep specification. *Phil. Trans. R. Soc. A* 375, 2104 (2017), 20160331. <https://doi.org/10.1098/rsta.2016.0331>
- Steve Awodey. 2010. *Category theory*. Oxford university press.
- Arthur Azevedo de Amorim, Nathan Collins, André DeHon, Delphine Demange, Cătălin Hrițcu, David Pichardie, Benjamin C Pierce, Randy Pollack, and Andrew Tolmach. 2014. A verified information-flow architecture. In *ACM SIGPLAN Notices*, Vol. 49. ACM, 165–178.
- Ralph-Johan Back. 1978. *On the correctness of refinement steps in program development*. Technical Report A-1978-4. Department of Computer Science, University of Helsinki, Helsinki, Finland.
- Ralph-Johan Back and Joakim von Wright. 1998. *Refinement Calculus: A Systematic Introduction*. Springer, New York. <https://doi.org/10.1007/978-1-4612-1674-2>
- John Baez and Mike Stay. 2010. Physics, topology, logic and computation: a Rosetta Stone. In *New structures for physics*. Springer, Berlin, Heidelberg, 95–172. [https://doi.org/10.1007/978-3-642-12821-9\\_2](https://doi.org/10.1007/978-3-642-12821-9_2)
- Andreas Blass. 1992. A game semantics for linear logic. *Ann. Pure Appl. Log.* 56, 1–3 (1992), 183–220. [https://doi.org/10.1016/0168-0072\(92\)90073-9](https://doi.org/10.1016/0168-0072(92)90073-9)
- Andreas Blass. 1993. Is game semantics necessary?. In *International Workshop on Computer Science Logic*. Springer, 66–77.
- Haogang Chen, Daniel Ziegler, Tej Chajed, Adam Chlipala, M. Frans Kaashoek, and Nickolai Zeldovich. 2015. Using Crash Hoare logic for certifying the FSCQ file system. In *Proceedings of the 25th Symposium on Operating Systems Principles (SOSP '15)*. ACM, New York, NY, USA, 18–37. <https://doi.org/10.1145/2815400.2815402>
- Joonwon Choi, Muralidaran Vijayaraghavan, Benjamin Sherman, Adam Chlipala, and Arvind. 2017. Kami: A Platform for High-Level Parametric Hardware Specification and Its Modular Verification. *Proc. ACM Program. Lang.* 1, ICFP, Article 24 (Aug. 2017), 30 pages. <https://doi.org/10.1145/3110268>

- Patrick Cousot and Radhia Cousot. 1977. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *Proceedings of the 4th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*. ACM, 238–252.
- Patrick Cousot and Radhia Cousot. 1992. Abstract interpretation frameworks. *J. Log. Comput.* 2, 4 (1992), 511–547.
- Edsger W Dijkstra. 1975. Guarded commands, nondeterminacy, and formal derivation of programs. *Commun. ACM* 18, 8 (Aug. 1975), 453–457. <https://doi.org/10.1145/360933.360975>
- M. Ern , J. Koslowski, A. Melton, and G. E. Strecker. 1993. A Primer on Galois Connections. *Ann. N. Y. Acad. Sci.* 704, 1 (1993), 103–125. <https://doi.org/10.1111/j.1749-6632.1993.tb52513.x>
- Dan R Ghica and Andrzej S Murawski. 2004. Angelic semantics of fine-grained concurrency. In *International Conference on Foundations of Software Science and Computation Structures*. Springer, 211–225.
- Ronghui Gu, J r mie Koenig, Tahina Ramananandro, Zhong Shao, Xiongnan (Newman) Wu, Shu-Chun Weng, Haozhong Zhang, and Yu Guo. 2015. Deep Specifications and Certified Abstraction Layers. In *Proceedings of the 42nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL ’15)*. ACM, New York, NY, USA, 595–608. <https://doi.org/10.1145/2676726.2676975>
- Ronghui Gu, Zhong Shao, Hao Chen, Xiongnan Wu, Jieung Kim, Vilhelm Sj berg, and David Costanzo. 2016. CertiKOS: An Extensible Architecture for Building Certified Concurrent OS Kernels. In *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation (OSDI’16)*. USENIX Association, Berkeley, CA, USA, 653–669. <https://dl.acm.org/doi/10.5555/3026877.3026928>
- Ronghui Gu, Zhong Shao, Jieung Kim, Xiongnan Newman Wu, J r mie Koenig, Vilhelm Sj berg, Hao Chen, David Costanzo, and Tahina Ramananandro. 2018. Certified concurrent abstraction layers. In *Proceedings of the 39th ACM SIGPLAN Conference on Programming Language Design*

- and Implementation (PLDI 2018)*. ACM, New York, NY, USA, 646–661. <https://doi.org/10.1145/3192366.3192381>
- Russell Harmer and Guy McCusker. 1999. A fully abstract game semantics for finite nondeterminism. In *Proceedings of the 14th Annual IEEE Symposium on Logic in Computer Science (LICS '99)*. IEEE Computer Society, USA, 422–430. <https://doi.org/10.1109/LICS.1999.782637>
- Masahito Hasegawa. 1995. Decomposing typed lambda calculus into a couple of categorical programming languages. In *Proceedings of the 6th International Conference on Category Theory and Computer Science (CTCS '95)*. Springer, Berlin, Heidelberg, 200–219. [https://doi.org/10.1007/3-540-60164-3\\_28](https://doi.org/10.1007/3-540-60164-3_28)
- Claudio Hermida, Uday S. Reddy, and Edmund P. Robinson. 2014. Logical Relations and Parametricity - A Reynolds Programme for Category Theory and Programming Languages. *Electron. Notes Theor. Comput. Sci.* 303 (March 2014), 149–180. <https://doi.org/10.1016/j.entcs.2014.02.008>
- C. A. R. Hoare. 1969. An Axiomatic Basis for Computer Programming. *Commun. ACM* 12, 10 (Oct. 1969), 576–580. <https://doi.org/10.1145/363235.363259>
- J. M. E. Hyland and C.-H. L. Ong. 2000. On Full Abstraction for PCF: I, II, and III. *Inf. Comput.* 163, 2 (2000), 285–408. <https://doi.org/10.1006/inco.2000.2917>
- Jacques-Henri Jourdan. 2016. *Verasco: a formally verified C static analyzer*. Ph.D. Dissertation.
- Jecheon Kang, Yoonseung Kim, Chung-Kil Hur, Derek Dreyer, and Viktor Vafeiadis. 2016. Lightweight Verification of Separate Compilation. In *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages* (St. Petersburg, FL, USA) (*POPL '16*). ACM, New York, NY, USA, 178–190. <https://doi.org/10.1145/2837614.2837642>
- Gerwin Klein, Kevin Elphinstone, Gernot Heiser, June Andronick, David Cock, Philip Derrin, Dhammika Elkaduwe, Kai Engelhardt, Rafal Kolanski, Michael Norrish, et al. 2009. seL4: formal verification of an OS kernel. In *Proceedings of the ACM SIGOPS 22nd Symposium on Operating*

- Systems Principles (SOSP '09)*. ACM, New York, NY, USA, 207–220. <https://doi.org/10.1145/1629575.1629596>
- Dexter Kozen. 1998. *Typed Kleene algebra*. Technical Report TR98-1669. Cornell University. <https://hdl.handle.net/1813/7323>
- James Laird. 1997. Full abstraction for functional languages with control. In *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science (LICS '97)*. IEEE Computer Society, USA, 58–67. <https://doi.org/10.1109/LICS.1997.614931>
- Xavier Leroy. 2009. Formal Verification of a Realistic Compiler. *Commun. ACM* 52, 7 (July 2009), 107–115. <https://doi.org/10.1145/1538788.1538814>
- Xavier Leroy. 2012. Mechanized semantics for compiler verification. In *Proceedings of the 10th Asian Symposium on Programming Languages and Systems (APLAS 2012)*. Springer, Berlin, Heidelberg, 386–388. [https://doi.org/10.1007/978-3-642-35182-2\\_27](https://doi.org/10.1007/978-3-642-35182-2_27)
- Xavier Leroy, Andrew W. Appel, Sandrine Blazy, and Gordon Stewart. 2012. *The CompCert Memory Model, Version 2*. Research report RR-7987. INRIA. <http://hal.inria.fr/hal-00703441>
- Xavier Leroy and Sandrine Blazy. 2008. Formal verification of a C-like memory model and its uses for verifying program transformations. *J. Autom. Reason.* 41, 1 (2008), 1–31.
- Saunders Mac Lane. 1978. *Categories for the working mathematician*. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4757-4721-8>
- William Mansky, Wolf Honoré, and Andrew W Appel. 2020. Connecting Higher-Order Separation Logic to a First-Order Outside World. In *Proceedings of the 29th European Symposium on Programming (ESOP '20)*. Springer, Cham, 428–455.
- Eugenio Moggi. 1991. Notions of computation and monads. *Information and computation* 93, 1 (1991), 55–92.
- Carroll Morgan. 1988. The specification statement. *ACM Trans. Program. Lang. Syst.* 10, 3 (July 1988), 403–419. <https://doi.org/10.1145/44501.44503>



- Joseph M. Morris. 2004. Augmenting Types with Unbounded Demonic and Angelic Nondeterminacy. In *Proceedings of the 7th International Conference on Mathematics of Program Construction (MPC 2004)*. Springer, Berlin, Heidelberg, 274–288. [https://doi.org/10.1007/978-3-540-27764-4\\_15](https://doi.org/10.1007/978-3-540-27764-4_15)
- Joseph M Morris and Malcolm Tyrrell. 2008. Dually nondeterministic functions. *ACM Trans. Program. Lang. Syst.* 30, 6 (Oct. 2008), 34. <https://doi.org/10.1145/1391956.1391961>
- Georg Neis, Chung-Kil Hur, Jan-Oliver Kaiser, Craig McLaughlin, Derek Dreyer, and Viktor Vafeiadis. 2015. Pilsner: A Compositionally Verified Compiler for a Higher-order Imperative Language. In *Proceedings of the 20th ACM SIGPLAN International Conference on Functional Programming (Vancouver, BC, Canada) (ICFP 2015)*. ACM, New York, NY, USA, 166–178. <https://doi.org/10.1145/2784731.2784764>
- Daniel Patterson and Amal Ahmed. 2019. The Next 700 Compiler Correctness Theorems (Functional Pearl). *Proc. ACM Program. Lang.* 3, ICFP, Article 85 (July 2019), 29 pages. <https://doi.org/10.1145/3341689>
- Benjamin C Pierce. 1991. *Basic category theory for computer scientists*. MIT press.
- Gordon Plotkin and John Power. 2001. Adequacy for Algebraic Effects. In *Proceedings of the 4th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS 2001)*. Springer, Berlin, Heidelberg, 1–24. [https://doi.org/10.1007/3-540-45315-6\\_1](https://doi.org/10.1007/3-540-45315-6_1)
- Gordon Plotkin and Matija Pretnar. 2009. Handlers of algebraic effects. In *Proceedings of the 18th European Symposium on Programming (ESOP 2009)*. Springer, Berlin, Heidelberg, 80–94. [https://doi.org/10.1007/978-3-642-00590-9\\_7](https://doi.org/10.1007/978-3-642-00590-9_7)
- Zhong Shao. 2010. Certified Software. *Commun. ACM* 53, 12 (Dec. 2010), 56–66. <https://doi.org/10.1145/1859204.1859226>
- Youngju Song, Minki Cho, Dongjoo Kim, Yonghyun Kim, Jeehoon Kang, and Chung-Kil Hur. 2019. CompCertM: CompCert with C-Assembly Linking and Lightweight Modular Verification. *Proc. ACM Program. Lang.* 4, POPL, Article 23 (Dec. 2019), 31 pages. <https://doi.org/10.1145/3371091>

- Bas Spitters and Eelis van der Weegen. 2011. Type Classes for Mathematics in Type Theory. arXiv:1102.1323 [cs.LO]
- Gordon Stewart, Lennart Beringer, Santiago Cuellar, and Andrew W. Appel. 2015. Compositional CompCert. In *Proceedings of the 42nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages* (Mumbai, India) (POPL '15). ACM, New York, NY, USA, 275–287. <https://doi.org/10.1145/2676726.2676985>
- Věra Trnková, Jiří Adámek, Václav Koubek, and Jan Reiterman. 1975. Free algebras, input processes and free monads. *Commentationes Mathematicae Universitatis Carolinae* 16, 2 (1975), 339–351.
- Malcolm Tyrrell, Joseph M. Morris, Andrew Butterfield, and Arthur Hughes. 2006. A Lattice-Theoretic Model for an Algebra of Communicating Sequential Processes. In *Proceedings of the Third International Colloquium on Theoretical Aspects of Computing (ICTAC 2006)*. Springer, Berlin, Heidelberg, 123–137. [https://doi.org/10.1007/11921240\\_9](https://doi.org/10.1007/11921240_9)
- Yuting Wang, Pierre Wilke, and Zhong Shao. 2019. An Abstract Stack Based Approach to Verified Compositional Compilation to Machine Code. *Proc. ACM Program. Lang.* 3, POPL, Article 62 (Jan. 2019), 30 pages. <https://doi.org/10.1145/3290375>
- Li-yao Xia, Yannick Zakowski, Paul He, Chung-Kil Hur, Gregory Malecha, Benjamin C Pierce, and Steve Zdancewic. 2019. Interaction trees: representing recursive and impure programs in Coq. *Proc. ACM Program. Lang.* 4, POPL, Article 51 (Dec. 2019), 32 pages. <https://doi.org/10.1145/3371119>