

# Zhe Yang

## Research Interests

Design, implementation, theory, and practice of programming languages and systems: program and data structure as data objects; semantic-based program manipulation, in particular program analysis, partial evaluation, and multi-level languages; interaction between programming languages and algorithms; effect type systems; continuation-passing style.

## Education

### Fall 1995–September 2001

New York University

Ph.D. in Computer Science, September 2001.

M.S. in Computer Science, July 1997.

Thesis title: *Language support for program generation: reasoning, implementation, and applications.*

Advisors: Prof. Olivier Danvy and Prof. Benjamin Goldberg.

Other thesis committee members: Richard Cole, Fritz Henglein, Amir Pnueli, Alan Siegel.

GPA: 4.0/4.0

### Summer 1994

Shanghai Jiao-Tong University, China

B.S. in Computer Science and Engineering. *Honor class.*

## Thesis Overview

I develop techniques for formally reasoning about higher-order code-generation algorithms and for implementing them efficiently. My approach is to specify such algorithms in two-level languages (i.e., ones with meta level for code-generation level as well as an object level for generated code) that accommodate common proof obligations about code generation, and embed natively into conventional languages. In the dissertation I (1) develop two such two-level languages, one for call-by-name object languages and one for call-by-value object languages with effects (e.g., ML), (2) prove their formal properties to support reasoning and their native implementation in existing programming languages, (3) develop a type encoding method for writing type-indexed programs (which comprise a large class of code-generation algorithms) in ML-like languages, (4) apply these techniques to non-trivial examples: one-pass CPS transformation, type-directed partial evaluation for call-by-name and for call-by-value, and (5) formally derive self application for type-directed partial evaluation, thus achieving automatic compiler generation.

## Publications

### Refereed conference papers

The second Futamura projection for type-directed partial evaluation.

Bernd Grobauer and Zhe Yang. In *Proceedings of the 2000 ACM SIGPLAN Workshop on Partial Evaluation and Semantics-Based Program Manipulation*, pages 22–32, January 2000, Boston, MA.

An operational investigation of the CPS hierarchy.

Olivier Danvy and Zhe Yang. In *Proceedings of the 1999 European Symposium on Programming*, LNCS 1576, pages 224–242, March 1999, Amsterdam, the Netherlands.

Encoding types in ML-like languages.

In *Proceedings of the 1998 ACM SIGPLAN International Conference on Functional Programming*, pages 289–300, September 1998, Baltimore, MD.

High level reading and data structure compilation.

Robert Paige and Zhe Yang. In *Proceedings of the 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 456–469, January 1997, Paris, France.

### Refereed journal papers

The second Futamura projection for type-directed partial evaluation.

Bernd Grobauer and Zhe Yang. *Higher-Order and Symbolic Computation*, 14(2/3), 2001, Also available as Technical Report BRICS RS-00-44, Department of Computer Science, University of Aarhus, December 2000.

Encoding types in ML-like languages.

*Theoretical Computer Science*, to appear. Preliminary version available as Technical Report BRICS RS-98-9, Department of Computer Science, University of Aarhus, April 1998..

## Others

Reasoning about code generation in two-level languages.

Technical Report BRICS RS-00-46, Department of Computer Science, University of Aarhus, December 2000.

(<http://cs.nyu.edu/zheyang/Reason2Level-long.ps.gz>)

Reasoning about effects.

Andrzej Filinski and Zhe Yang. March 2000. Unpublished.

(<http://cs.nyu.edu/zheyang/TheoryOfEffects.ps.gz>)

A Survey on Multi-Stage Programming Languages.

January 1999, as part of the PhD requirements.

(<http://cs.nyu.edu/zheyang/papers/survey.ps.gz>)

A native ML implementation of type-directed partial evaluation (abstract).

In *Proceedings of the 1998 APPSEM Workshop on Normalization by Evaluation*, number NS-98-1 in BRICS Notes Series, May 1998, Göteborg, Sweden.

## Professional Experience

### Summer 2002–Present

**Aleri, Inc.**

Database architect, under the direction of Dr. Jon Riecke

Our team developed a functional programming language used for implementing vectorized database systems. My work, in particular, involves the design and development of several algorithms for fast searching, sorting, and partitioning operations on vectors. The algorithms fit well into modern memory hierarchy—where cache locality was the most emphasized part—and scale smoothly to parallel versions that run on symmetric multiprocessor systems. The implementation is kept generic, but specializes, through the use of C++ templates, to different element types for maximum efficiency.

### Fall 2001–Summer 2002

**University of Pennsylvania**

Post-doctoral research fellow with Prof. Benjamin C. Pierce.

My research work included developing and maintaining the open-source, multi-platform file synchronization software, Unison, as well as designing the next-generation general purpose synchronizer, Harmony, that takes into account document-specific structural and semantic information. With Prof. Benjamin Pierce, I co-taught the graduate-level seminar course “Distributed Data” (CIS 700/1) in the Spring 2002 semester.

### Spring 1999–Spring 2001

**University of Aarhus (Denmark)**

Visiting Ph.D. student and research assistant, BRICS.

My research focused on semantic-based program manipulation. With Bernd Grobauer, I reformulated and implemented the second Futamura projection (i.e., compiler generation via self application of partial evaluators) for type-directed partial evaluation. With Andrzej Filinski, I developed type theories for reasoning about languages with a hierarchy of effects. I also developed two-level languages that support formal reasoning about higher-order code-generation algorithms and their efficient implementation.

### Spring 1996–Spring 1999

**New York University**

Research assistant with Prof. Robert Paige.

Our research focused on the interaction between languages and algorithms. With a type system that reflects the complexity of compound operations (e.g., associative operations on finite sets and maps), we developed a worst-case linear time algorithm to solve the reading problem, i.e., how to convert external text input to internal, non-linear representation that supports efficient operations. We also developed a prototypical offline partial evaluator for SETL.

**Summer 1998** **NEC Research Institute (Princeton, NJ)**  
Summer internship, working with Dr. Suresh Jagannathan.

I developed a flow analysis for an intermediate language for Java, engineered the analysis for supporting separate compilation, and exploited the flow analysis for optimizations (e.g., flow-directed inlining).

**Fall 1997** **University of Aarhus and University of Copenhagen (Denmark)**  
Visiting PhD student, working with Prof. Olivier Danvy and Prof. Neil D. Jones.

With Olivier Danvy, I developed an operational notion of continuation semantics, used this notion to study the CPS hierarchy of control operators, and derived a native and modular implementation. I also developed a method to write apparently dependently typed programs in languages with Hindley-Milner type systems, using type encodings.

**Fall 1995 and Fall 1998** **New York University**  
Teaching assistant for graduate courses *Programming Languages* and *Fundamental Algorithms*, assisting Prof. Bud Mishra and Prof. Ravi Boppana.

My tasks included grading assignments, discussing with students, and answering their questions.

**August 1994 to July 1995** **Shanghai Metals and Minerals Import and Export Corporation**  
Team Leader of the software development group for an integrated exporting goods and related documentation Management Information System. Designer of the object and dataflow models for the system, and member of the programming group.

**September 1994 to December 1994** **Faculty Club of Shanghai Jiao-Tong University**  
Designer and programmer of the accounting part of JTHIS, a general-purpose Hotel Information System which also includes other components such as room reservation and monitoring, telephone billing, and dining management.

**September 1992 to July 1994** **Shanghai Jiao-Tong University**  
Undergraduate research student, independent study with the Software and Theory Group.

I studied specification languages, program analysis, and partial evaluation. For my undergraduate thesis, I designed HOPFL, a higher-order functional language with simple communication primitives, and developed a self-applicable partial evaluator for HOPFL.

## Professional Activities

**Conference committee member of:** *ACM SIGPLAN 2003 Workshop on Partial Evaluation and Semantics-Based Program Manipulation (PEPM)* (<http://www.ecs.soton.ac.uk/~mal/PEPM03/>).  
*2002 ACM SIGPLAN ASIAN Symposium on Partial Evaluation and Semantics-Based Program Manipulation (ASIA-PEPM)* (<http://www.comp.nus.edu.sg/asia-pepm02/>).

**Referee for (conferences):** *Computer Science Logic (CSL)* (2002), *Logic in Computer Science (LICS)* (2002), *Foundation of Software Science and Computation Structures (FoSSaCS)* (2002), *Symposium on Programs as Data Objects (PADO)* (2001), *International Colloquium on Automata, Languages, and Programming (ICALP)* (2000), *Programming Languages Design and Implementation (PLDI)* (2000), *International Conference on Functional Programming (ICFP)* (1999, 2000), *Static Analysis Symposium (SAS)* (1998), and *Algorithmic Languages and Calculi* (1997).

**Referee for (journal):** *Higher-Order and Symbolic Computation* (1998, 2000).

**Given Invited talks on:** *High Level Reading*, at DIKU TOPPS seminar, University of Copenhagen (1997, hosted by Prof. Neil D. Jones); *Encoding Types*, at BRICS, University of Aarhus (1998, hosted by Prof. Olivier Danvy), and at DIKU TOPPS seminar, University of Copenhagen (1998, hosted by Prof. Fritz Henglein); *the CPS Hierarchy*, at the Programming Language Seminar, Northeastern University (2000, hosted by Prof. Mitchell Wand), and at the Bell Labs, Lucent Technology (2000, hosted by Dr. Jon Riecke); *Language Support for Program Generation*, at Georgia Tech (2001, hosted by Prof. Olin Shivers) and at Princeton University (2001, hosted by Prof. Andrew Appel).

Co-organizer (with Prof. Robert Paige) of the SPLASH (Special Programming Language And Semantics Hour) seminar at NYU, Spring 1998 and Fall 1998.

Served as the vice president of the Courant Student Organization, NYU, 1996–1997.

Served as a student representative for the Computer Science Department, NYU, Spring 1998.

## Honors and Awards

ACM PAC traveling fund (3 times) and NYU GSAS student conference travel award (twice) for attending POPL 1997, ICFP 1998, and PEPM 2000 as a student author.

Awarded Harold Grad Memorial Prize, for outstanding performance and promise as a graduate student, Courant Institute of Mathematical Sciences, NYU, April 1999.

BRICS research fellowship, University of Aarhus, Fall 1997 and Spring 1999–Summer 2000.

Teaching/research assistantship, New York University, Fall 1995–Summer 2001.

Honorable mention, 1994 American Mathematical Competition in Modeling.

First prize winner, 1993 Chinese Mathematical Competition in Modeling.

First prize winner, 1992 Chinese Mathematical Competition in Modeling.

## References

### Dr. Jon Riecke

Aleri, Inc.  
41 East 11th Street, 10th floor  
New York, NY 10003  
jon.riecke@aleri.com

### Prof. Benjamin C. Pierce

Department of Computer and Information Science  
University of Pennsylvania  
200 South 33rd Street  
Philadelphia, PA 19104-6389  
bcpierce@cis.upenn.edu

### Prof. Olivier Danvy

BRICS, Department of Computer Science  
University of Aarhus  
Ny Munkegade, Building 540  
DK-8000, Aarhus C, Denmark  
danvy@brics.dk

### Prof. Fritz Henglein

University of Copenhagen  
Universitetsparken 1  
2100 Copenhagen East, Denmark  
henglein@diku.dk

### Prof. Andrzej Filinski

DIKU, University of Copenhagen  
Universitetsparken 1  
DK-2100, Copenhagen East, Denmark  
andrzej@diku.dk

## Contact Information and Personal Data

1453A Bay Ridge Pkwy, 3rd fl  
Brooklyn, NY 11228  
USA  
(1) 718 236-6428 (Home)  
(1) 646 206-2676 (Mobile)  
zheyang@cis.upenn.edu  
<http://cs.nyu.edu/~zheyang>

Born December 1975 in Shanghai, China.  
Native Chinese, fluent English, and basic Italian.