

TOWARDS HYBRID INTENSIONAL PROGRAMMING
WITH JLUCID, OBJECTIVE LUCID, AND GENERAL
IMPERATIVE COMPILER FRAMEWORK IN THE GIPSY

SERGUEI A. MOKHOV

A THESIS
IN
THE DEPARTMENT
OF
COMPUTER SCIENCE AND SOFTWARE ENGINEERING

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF COMPUTER SCIENCE
CONCORDIA UNIVERSITY
MONTRÉAL, QUÉBEC, CANADA

OCTOBER 2005

© SERGUEI A. MOKHOV, 2005

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Serguei A. Mokhov

Entitled: Towards Hybrid Intensional Programming with JLucid, Objective Lucid, and General Imperative Compiler Framework in the GIPSY

and submitted in partial fulfillment of the requirements for the degree of

Master of Computer Science

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. H.F. Li

_____ Examiner
Dr. P. Grogono

_____ Examiner
Dr. P. Chalin

_____ Supervisor
Dr. J. Paquet

Approved _____
Chair of Department or Graduate Program Director

_____ 20 _____

Dr. Nabil Esmail, Dean
Faculty of Engineering and Computer Science

Abstract

Towards Hybrid Intensional Programming with JLucid, Objective Lucid, and General Imperative Compiler Framework in the GIPSY

Serguei A. Mokhov

Pure Lucid programs are concurrent with very fine granularity. Sequential Threads (STs) are functions introduced to enlarge the grain size; they are passed from server to workers by Communication Procedures (CPs) in the General Intensional Programming System (GIPSY). A JLucid program combines Java code for the STs with Lucid code for parallel control. Thus first, in this thesis, we describe the way in which the new JLucid compiler generates STs and CPs. JLucid also introduces array support.

Further exploration goes through the additional transformations that the Lucid family of languages has undergone to enable the use of Java objects and their members, in the Generic Intensional Programming Language (GIPL), and Indexical Lucid: first, in the form of JLucid allowing the use of *pseudo-objects*, and then through the specifically-designed the Objective Lucid language. The syntax and semantic definitions of Objective Lucid and the meaning of Java objects within an intensional program are provided with discussions and examples.

Finally, there are many useful scientific and utility routines written in many imperative programming languages other than Java, for example in C, C++, Fortran, Perl, etc. Therefore, it is wise to provide a framework to facilitate inclusion of these languages into the GIPSY and their use by Lucid programs. A General Imperative Compiler Framework and its concrete implementation is proposed to address this issue.

Acknowledgments

I would like to thank my supervisor Dr. Joey Paquet and Dr. Peter Grogono for ever lasting patience and caring guidance throughout the variety of learning experience and their advices and insightful comments to make these contributions possible. I would also like to thank my friendly team members with whom we together were lifting the complex GIPSY system off the ground. Specifically, I would like to mention Chun Lei Ren, Paula Bo Lu, Ai Hua Wu, Yimin Ding, Lei Tao, Emil Vassev, and Kai Yu Wan for outstanding team work. Thanks to Dr. Patrice Chalin for an in-depth introduction to semantics of programming languages. Thanks to Dr. Sabine Bergler and Dr. Leila Kosseim for the journey through the internals of natural language processing side related to this work. Thanks to my beloved Irina for helping me to carry through.

This work has been sponsored by NSERC and the Faculty of Engineering and Computer Science of Concordia University, Montréal, Québec, Canada. This document was produced in L^AT_EX with the guidance of Dr. Grogono's manual in [Gro01] and Concordia University L^AT_EX thesis styling maintained by Steve Malowany, Stan Swiercz, and Patrice Chalin.

Contents

List of Figures	xii
List of Tables	xv
1 Introduction	1
1.1 Thesis Statement	1
1.2 Contributions	2
1.3 Scope of the Thesis	3
1.4 Structure of the Thesis	4
2 Background	5
2.1 Intensional Programming	5
2.2 The Lucid Programming Language	7
2.2.1 Brief History and The Family	7
2.2.2 Indexical Lucid	9
2.2.2.1 Streams	9
2.2.2.2 Basic Operators	10
2.2.2.3 Sequentiality Problem	12
2.2.2.4 Random Access to Streams	12
2.2.2.5 Definition of Lucid Operators By Means of @ and #	13
2.2.2.6 Abstract Syntax of Lucid	14
2.2.2.7 Concrete GIPL Syntax	14
2.2.2.8 Semantic Rules	15
2.2.2.9 Examples of Lucid Programs	18
2.2.3 Lucid Now	19
2.3 Hybrid Programming	19

2.3.1	ML _{<}	20
2.3.2	FC++	20
2.3.3	GLU	20
2.3.4	GLU#	21
2.4	Compiler Frameworks	21
2.5	General Intensional Programming System	23
2.5.1	Introduction	23
2.5.2	Goals	25
2.5.3	General Intensional Programming Compiler	25
2.5.4	General Education Engine	27
2.5.4.1	Demand Propagation Resources for the GEE	30
2.5.4.2	Synchronization	31
2.5.5	Run-time Interactive Programming Environment	32
2.6	Tools	35
2.6.1	Java as a Programming Language	35
2.6.1.1	Java Reflection	35
2.6.1.2	Java Native Interface (JNI)	36
2.6.1.3	JUnit	37
2.6.2	javacc – Java Compiler Compiler	37
2.6.3	MARF	38
2.6.4	CVS	38
2.6.5	Tomcat	39
2.6.6	Build System	39
2.6.6.1	Makefiles	39
2.6.6.2	Eclipse	41
2.6.6.3	JBuilder	41
2.6.6.4	Ant	41
2.6.6.5	NetBeans	41
2.6.7	readmedir	41
2.7	Summary	42
3	Methodology	43
3.1	JLucid: Lucid with Embedded Java Methods	43
3.1.1	Rationale	43

3.1.1.1	Modeling Non-Determinism	44
3.1.1.2	Loading Existing Java Code with <code>embed()</code>	44
3.1.1.3	The <code>#JAVA</code> and <code>#JLUCID</code> Code Segments	47
3.1.1.4	Is JLucid an Intensional Language?	51
3.1.2	Syntax	51
3.1.3	Semantics	52
3.2	Objective Lucid: JLucid with Java Objects	53
3.2.1	Rationale	53
3.2.1.1	Pseudo-Objectivism in JLucid	53
3.2.1.2	Stream of Objects	54
3.2.1.3	Pure Intensional Object-Oriented Programming	56
3.2.2	Syntax	56
3.2.3	Semantics	56
3.3	General Imperative Compiler Framework	60
3.3.1	Rationale	60
3.3.2	Matching Lucid and Java Data Types	63
3.3.3	Sequential Thread and Communication Procedure Generation	64
3.3.3.1	Java Sequential Threads	64
3.3.3.2	Java Communication Procedures	65
3.3.3.3	C Sequential Threads and Communication Procedures with the JNI	66
3.3.3.4	Worker Aggregator Definition in the Generator-Worker Architecture	67
3.4	Summary	68
3.4.1	Benefits	69
3.4.2	Limitations	69
4	Design and Implementation	70
4.1	Internal Design	70
4.1.1	General Intensional Programming Compiler Framework	70
4.1.1.1	General Imperative Compiler Framework	71
4.1.1.2	Generalization of a Concrete Implementation	71
4.1.1.3	Resolving Generalization Issues and Binary Compat- ibility	74

4.1.1.4	GIPC Preprocessor	79
4.1.1.5	GIPSY Type System	83
4.1.1.6	GICF Design	85
4.1.1.7	Intensional Programming Languages Compiler Framework	85
4.1.1.8	Sequential Thread and Communication Procedure Interfaces	86
4.1.1.9	GIPC Design	88
4.1.1.10	GIPC Class as a Meta Processor	91
4.1.1.11	Calling Sequence	91
4.1.1.12	Compiling and Linking	97
4.1.1.13	Semantic Analyzer	98
4.1.1.14	Interfacing GIPC and GEE and Compiled GIPSY Program	99
4.1.2	JLucid	101
4.1.2.1	Design	101
4.1.2.2	Grammar Generation	102
4.1.2.3	Free Java Functions and Java Compiler	103
4.1.2.4	Arrays	104
4.1.2.5	Implementing <code>embed()</code>	105
4.1.2.6	Abstract Syntax Tree and the Dictionary	105
4.1.3	Objective Lucid	106
4.1.3.1	Design	107
4.1.3.2	Grammar Generation	107
4.1.3.3	Object Instantiation	107
4.1.3.4	The Dot-Notation	108
4.1.3.5	Abstract Syntax Tree and the Dictionary	108
4.1.3.6	Objects as Arrays and Arrays as Objects	109
4.2	External Design	112
4.2.1	User Interface	112
4.2.1.1	WebEditor – A Web Front-End to the GIPSY	112
4.2.1.2	GIPSY Command-Line Interface	114
4.2.1.3	RIPE Command-Line Interface	115

4.2.1.4	GIPC Command-Line Interface	116
4.2.1.5	GEE Command-Line Interface	118
4.2.1.6	Regression Testing Application Command-Line Inter- face	119
4.2.2	External Software Interfaces	120
4.2.2.1	JavaCC API	120
4.2.2.2	MARF Library API	122
4.2.2.3	Servlets API	126
4.2.3	Architectural Design and Unit Integration	126
4.2.3.1	GIPSY	126
4.2.3.2	GIPSY Exceptions Framework	128
4.2.3.3	GEE Design	129
4.2.3.4	RIPE Design	132
4.2.3.5	Data Flow Graphs Integration	133
4.3	Summary	135
5	Testing	136
5.1	Regression Testing	136
5.1.1	Introduction	136
5.1.2	Regression Testing Suite	137
5.1.2.1	Unit Testing with JUnit	137
5.1.2.2	Unit Testing with <code>diff</code>	137
5.1.2.3	Tests	137
5.2	Portability Testing	138
5.3	Solving Problems	139
5.3.1	Prefix Sum	139
5.3.2	Dining Philosophers	144
5.3.3	Fast Fourier Transform	147
5.3.3.1	Fast Fourier Transform in JLucid.	147
5.3.3.2	Fast Fourier Transform code fragment in Java from MARF.	149
5.3.4	Moving Car	152
5.3.5	Game of Life	158
5.4	Summary	160

6	Conclusion	161
6.1	Results	161
6.1.1	Experiments	161
6.1.2	Interpretation of Results	161
6.2	Discussions and Limitations	162
6.2.1	Lack of Hybrid Intensional-Imperative Semantics Proofs . . .	162
6.2.2	Genuine Imperative Compilers	162
6.2.3	Cross-Language Data Type Mapping	162
6.2.4	Dimension Index Overflow	162
6.2.5	Hybrid-DFG Integration	163
6.2.6	Dealing With Side Effects and Abrupt Termination	163
6.2.7	Imperative Function Overloading	163
6.2.8	Cross-Imperative Language Calls	163
6.2.9	Security	164
7	Future Work	165
7.1	Formal Verification of Semantic Rules and the GIPSY Type System .	166
7.2	Dealing with Data Flow Graphs in Hybrid Programming	166
7.3	Security	166
7.4	Implementation of the C Compiler in GICF	167
7.5	Fully Explore Array Properties	167
7.6	Genuine Imperative and Functional Language Compilers	167
7.7	Visualization and Control of Communication Patterns and Load Balancing	167
7.8	Target Host Compilation	168
7.9	The GIPSY Screen Saver	168
7.10	The GIPSY Server	168
	Bibliography	169
	Appendix	178
A	Definitions and Abbreviations	178
A.1	Abbreviations	178

B	Sequential Thread and Communication Procedure Interfaces	181
B.1	Sequential Thread Interface	181
B.2	Communication Procedure Interface	181
B.3	Generated Sequential Thread Wrapper Class	181
B.4	Sample Worker’s Implementation	187
C	Architectural Module Layout	189
C.1	GIPSY Java Packages and Directory Structure	189
C.2	GIPSY Modules Packaging	191
D	Grammar Generation Scripts for JLucid and Objective Lucid	193
D.1	jlucid.sh	193
D.2	JGIPL.sh	194
D.3	JIndexicalLucid.sh	194
D.4	ObjectiveGIPL.sh	195
D.5	ObjectiveIndexicalLucid.sh	195
	Index	196

List of Figures

1	Concrete Indexical Lucid Syntax	14
2	GIPL Expressions	14
3	GIPL where Definitions	15
4	Concrete GIPL Syntax	15
5	Operational Semantics of GIPL	17
6	Natural numbers problem in Indexical Lucid.	18
7	Natural numbers problem in GIPL.	18
8	Indexical Lucid program implementing the merge() function.	18
9	The GIPSY Logo representing the distributed nature of GIPSY.	23
10	Structure of the GIPSY	24
11	Initial Conceptual Design of the GIPC	26
12	Conceptual Design of the GEE	28
13	Conceptual Design of the RIPE	33
14	Tomcat Web Applications Manager	40
15	Indexical Lucid program implementing the merge() function.	45
16	Indexical Lucid program implementing the merge() function as inline Java method.	45
17	Indexical Lucid program implementing the merge() function as embed()	45
18	Illustration of the embed() syntax.	45
19	Generated corresponding ST to that of Figure 18.	46
20	Inline Java function declaration.	47
21	Java method declaration split out from the Lucid part.	47
22	Natural numbers problem in plain GIPL.	48
23	Natural numbers problem with two Java methods calling each other.	48
24	Generated Sequential Thread Class.	50
25	JLucid Extension to GIPL Syntax	52

26	JLucid Extension to Indexical Lucid Syntax	52
27	Additional basic semantic rules to support JLucid	53
28	Pseudo-objectivism in JLucid.	55
29	Using pseudo-free Java functions to access object properties in JLucid. . .	55
30	Objective Lucid example.	57
31	Objective Lucid Syntax	58
32	Additional basic semantic rules to support Objective Lucid	60
33	Hybrid GIPSY Program Compilation Process	65
34	Generator-Worker Architecture	68
35	Example of a hybrid GIPSY program.	72
36	Another example of a hybrid GIPSY program.	73
37	Original Framework for the General Intensional Programming Com- piler in the GIPSY	76
38	Modified Framework for the General Intensional Programming Com- piler in the GIPSY	77
39	The <code>FormatTag</code> API.	78
40	The <code>GIPC Preprocessor</code>	80
41	Preprocessor Grammar for a GIPSY program.	82
42	GIPSY Type System.	84
43	GICF Design.	86
44	IPLCF Design.	87
45	SIPL to GIPL Translator Integration.	88
46	Sequential Thread and Communication Procedure Class Diagram. . .	89
47	All GIPC Compilers.	90
48	Overall GIPC Design.	92
49	Sequence Diagram of GIPSY Program Compilation Process.	94
50	Sequence Diagram of Intensional Compilation Process.	95
51	Sequence Diagram of Imperative Compilation Process.	96
52	Semantic Analyzer.	98
53	Class diagram describing <code>GIPSYProgram</code>	100
54	JLucid Design.	101
55	JLucid Compilation Sequence.	102
56	Java Compilation Sequence.	103

57	Objective Lucid Design.	106
58	Objective Lucid Compilation Sequence.	107
59	GIPSY WebEditor Interface.	113
60	JavaCC- and JJTree-generated Modules Used by Several GIPC Modules.	121
61	MARF Utility Classes used by the GIPSY.	123
62	Dictionary and DictionaryItem API	124
63	Dictionary Usage within the GIPSY	125
64	GIPSY Main Modules.	127
65	GIPSY Exceptions Framework.	128
66	GEE Design.	130
67	The Demand Dispatcher Integrated and Implemented based on Jini. .	131
68	Integration of the Intensional Value Warehouse and Garbage Collection.	132
69	RIPE Design.	133
70	DFG Integration Design.	134
71	Pseudocode of a thread j for the Prefix Sum Problem.	139
72	The Prefix Sum Problem in JLucid in GIPL Style.	139
73	The Prefix Sum Problem in JLucid in Indexical Lucid Style.	143
74	Objective Lucid example of a Car object that changes in time.	153
75	Eduction Tree for the Natural Numbers Problem.	155
76	The Natural Numbers Problem in Objective Lucid.	156
77	Eduction Tree for the Natural Numbers Problem in Objective Lucid.	157
78	The Life in Haskell.	158
79	The Life in Indexical Lucid.	159
80	Sequential Thread Interface.	182
81	Communication Procedure Interface.	183
82	GIPSY Java Packages Hierarchy.	190

List of Tables

1	Matching data types between Lucid and Java.	64
2	Correspondence of the GIPSY .jar files and the modules.	192

Chapter 1

Introduction

1.1 Thesis Statement

In the previous prototype of the General Intensional Programming System (GIPSY) there existed limitations to its potential in distributed computing – lack of sequential threads and communication procedures. Additionally, the capabilities of Indexical Lucid and GIPL, the primary GIPSY’s languages, were limited to only computing aspects without input/output, arrays, and some other essential features (e.g. math, non-determinism, dynamic loading) that exist in imperative (e.g. Java) languages. We discuss an extension to Generic Intensional Programming Language (GIPL) and Indexical Lucid with embedded Java – JLucid. A few problems are solved as an example using the enhanced language.

JLucid brings embedded Java and most of its powers into Indexical Lucid in the GIPSY by allowing intensional languages to manipulate Java methods as first class values¹. However, it is very natural to have objects with Java and manipulate their members in scientific intensional computation, yet JLucid fails to support that Java’s capability. Hence, we design Objective Lucid to address this deficiency. We define the operational semantics of Objective Lucid, and give some examples of its application.

Existence of JLucid, Objective Lucid, and GLU as well as many useful libraries written in other imperative languages, such as C/C++, Perl, Python, Fortran etc. demanded ability to use code written in those languages by intensional programs,

¹The Java methods are not referred to as “functions” as in functional programming – the Java methods can be passed around as values inside the Lucid part, but not to or from Java part of a GIPSY program.

naturally. Thus, we design a first version of the General Imperative Compiler Framework (GICF) as a part of the GIPSY to allow GIPSY programs to use virtually any combination of intensional and imperative languages at the meta level. This is a very ambitious goal; therefore, the proposal is the first iteration of the framework open for later refinements as it matures along with the corresponding changes to the run-time system.

1.2 Contributions

Primary contributions of this thesis are outlined below:

- JLucid
 - Semantics of *pseudo-free* Java methods in Lucid programs
 - Design and implementation of JLucid and its compiler in the GIPSY
- Objective Lucid
 - Semantics of the integration of Java objects in Lucid programs
 - Design and implementation of the Objective Lucid compiler
- General Imperative Compiler Framework
 - Design and Implementation of the GICF
 - Embedding of a Java compiler in the GICF
- WebEditor to edit, compile, and run GIPSY programs online
- System Architecture Issues
 - Rework and refactoring of most existing system design, both at the architectural and detailed design levels
 - Major rework of the architecture and detailed design of GIPC
 - Java sequential threads generation
 - Threaded and RMI communication procedures generation

- GIPSY Type System²
- GIPSY Exceptions Framework
- Regression Testing Infrastructure
- Unit Testing Automation with `JUnit`

The last contributed items touch the rest of the GIPSY, the components and modules done by other team members. The integration performed (outside of the main scope of this thesis) demanded extensive testing. Without the integration and testing work, these other contributions wouldn't be possible. This also includes developing and enforcing Coding Conventions and setting up project's CVS repository [Mok05b, Mok03a, Mok03b] for the entire project as this work is to become a manual for the current and future GIPSY developers and researchers.

1.3 Scope of the Thesis

While the Contributions section outlines the major work done, the below explains what was not done or exhibits some limitations at the time of this writing:

- Integrated imperative compilers aren't native to the GIPSY, instead we call external compilers, such as `javac`, `gcc`, `g++`, `nmake.exe`, `bc.exe`, `perl`, etc. depending on a platform.
- Even though the mechanism was designed and implemented to generate CPs and STs, only two of the concrete implementations of the actual CPs were done: for local execution and distributed execution by extending the RMI implementation done by Bo Lu. The other implementations of CPs for Jini, DCOM+, CORBA and others are being worked on by other team members at the time of this writing.
- Semantic rules to have Java objects in Objective Lucid have been developed, but have not been formally proven to be correct.

²Though the type system may seem not to be related to the architecture, but it impacted the design most of the main modules in it, so it was classified as architectural.

- When presenting GICF and the Preprocessor syntax, no semantic rules are given for any of parts of the hybrid programs, except for JLucid and Objective Lucid, i.e. the semantics of integrated Java itself or C constructs, etc.
- JLucid and Objective Lucid are still in their experimental stage of development and it will take some time before they mature.

1.4 Structure of the Thesis

The next chapter provides the necessary background on the Lucid family of languages, its history, operational semantics, compiler frameworks, and hybrid programming. Then, it gives the context of this thesis, the GIPSY system, and the tools and techniques employed to make the contributions possible. The core of this thesis is based on three publications, namely [MPG05, MP05b, MP05a]. Chapter 3 describes the approach and methodology used to overcome and provide a solution to the problems stated in Section 1.1. Then, the design implementation details are presented in Chapter 4. Chapter 5 introduces the Regression Testing Suite for GIPSY and what kinds of tests were performed and their limitations. Finally, Chapter 6 and Chapter 7 conclude on the work done, discuss the results and limitations of the implementation, and lay down some paths towards enhancing the GIPSY in various areas further. At the end, there is a list of references, Bibliography, and an Appendix with most common abbreviations found in this work, CP and ST interfaces, JLucid and Objective Lucid grammar generation scripts, etc., followed by an overall index.

Chapter 2

Background

While there is a complete and comprehensive set of references in the Bibliography chapter that was a great deal of help to the creation of this work, there are some keynotes that require special mention. The following are some of the related readings that were sources of inspiration and invaluable informational food for thought. These include Joey Paquet’s PhD thesis “Scientific Intensional Programming” [Paq99], related hybrid intensional-imperative programming in various GLU-related work, such as [JD96, JDA97], other recent hybrid programming papers, such as [PK04, MS01, SM02], the PhD thesis of Paula Bo Lu [Lu04] and other theses of the GIPSY group, such as [Ren02, Din04, Tao04, Wu02], and semantics of programming languages in [Gro02a, HJ02, Moe04]. Additionally, since this work also deals with compiler frameworks, a general overview of existing frameworks is presented. An on-line encyclopedia, Wikipedia [WSoafaotw05], was a valuable resource for the background and literature review, some of which is summarized in the sections that follow.

2.1 Intensional Programming

Intensional programming is a generalization of unidimensional contextual (also known as *modal logic* [Car47, Kri59, Kri63]) programming such as temporal programming, but where the context is multidimensional and implicit rather than unidimensional and explicit. Intensional programming is also called *multidimensional* programming

because the expressions involved are allowed to vary in an arbitrary number of dimensions, the context of evaluation is thus a *multidimensional context*. For example, in intensional programming, one can very naturally represent complex physical phenomena such as plasma physics (e.g. in Tensor Lucid in [Paq99]), which are in fact a set of charged particles placed in a space-time continuum that behaves according to a limited set of laws of intensional nature. This space-time continuum becomes the different dimensions of the context of evaluation, and the laws are expressed naturally using intensional definitions [Paq99]. Joey Paquet’s PhD thesis discusses the syntax and semantics of the Lucid language, designs GIPL and Tensor Lucid. While we omit the Tensor Lucid part, the reader is reminded about the basic properties of the Indexical Lucid and GIPL languages in the follow up sections in greater detail to provide the necessary context for the follow up work in Chapter 3 and Chapter 4.

Intensional Logic

Intensional programming (IP) is based on intensional (or multidimensional or modal) logic (where semantics was applied first by [Car47, Kri59, Kri63]), which, in turn, are based on Natural Language Understanding (aspects, such as, *time*, *belief*, *situation*, and *direction* are considered). IP brings in **dimensions** and **context** to programs (e.g. space and time in physics or chemistry). Intensional logic adds dimensions to logical expressions; thus, a non-intensional logic can be seen as a constant or a snapshot in all possible dimensions. *Intensions are dimensions* at which a certain statement is true or false (or has some other than a Boolean value). *Intensional operators* are operators that allow us to navigate within these dimensions.

Temporal Intensional Logic

Temporal intensional logic is an extension of temporal logic that allows to specify the time in the future or in the past.

(1) $E_1 :=$ it is raining **here today**

Context: {**place:here**, **time:today**}

(2) $E_2 :=$ it was raining **here** *before(today) = yesterday*

(3) $E_3 :=$ it is going to rain *at(altitude here + 500 m) after(today) = tomorrow*

Let’s take E_1 from (1) above. Then let us fix **here** to **Montreal** and assume it is

a *constant*. In the month of March, 2004, with granularity of day, for every day, we can evaluate E_1 to either *true* or *false*:

```
Tags:   1 2 3 4 5 6 7 8 9 ...
Values: F F T T T F F F T ...
```

If you start varying the **here** dimension (which could even be broken down into X, Y, Z), you get a two-dimensional evaluation of E_1 :

```
City / Day  1 2 3 4 5 6 7 8 9 ...
Montreal    F F T T T F F F T ...
Quebec      F F F F T T T F F ...
Ottawa      F T T T T T F F F ...
```

The purpose of this example is to remind the reader the basic ideas behind intensions and intensional programming and what dimensionality is by using natural language. What follows is formalization of the above in terms of the Lucid programming language.

2.2 The Lucid Programming Language

Let us begin by introducing the Lucid language history and which features of it came at different stages of its evolution to its present form. This is the necessary step to further illustrate the purpose of this thesis.

2.2.1 Brief History and The Family

From 1974 to Lucid Today:

1. Lucid as a Pipelined Dataflow Language through 1974-1977. Lucid was introduced by Anchroft and Wadge in [AW76, AW77]. Features:
 - A purely declarative language for natural expression of iterative algorithms.
 - Goals: semantics and verification of correctness of programming languages (for details see [AW76, AW77]).

- Operators as pipelined streams: one for initial element, and then all for the successor ones.
2. Intensions, Indexical Lucid, GRanular Lucid (GLU, [JD96, JDA97]), circa 1996. More details on these two dialects are provided further in the chapter as they directly relate to the theme of this thesis. Features:
 - Random access to streams in Indexical Lucid.
 - First working hybrid intensional-imperative paradigm (C/Fortran and Indexical Lucid) in the form of GLU.
 - Education or demand-driven execution (in GLU).
 3. Partial Lucid, Tensor Lucid, 1999 [Paq99].
 - Partial Lucid is an intermediate experimental language used for demonstrative purposes in presenting the semantics of Lucid in [Paq99].
 - Tensor Lucid dialect was developed by Joey Paquet for plasma physics computations to illustrate advantages and expressiveness of Lucid over an equivalent solution written in Fortran.
 4. GIPL, 1999 [Paq99].
 - All Lucid dialects can be translated into this basic form of Lucid, GIPL through a set of translation rules. (GIPL is in the foundation of the execution semantics of GIPSY and its GIPC and GEE because its AST is the only type of AST GEE understands when executing a GIPSY program).
 5. RLucid, 1999, [GP99]
 - A Lucid dialect for reactive real-time intensional programming.
 6. JLucid, Objective Lucid, 2003 - 2005
 - These dialects introduce a notion of hybrid and object-oriented programming in the GIPSY with Java and Indexical Lucid and GIPL, and are discussed great detail in the follow up chapters of this thesis.

7. Lucx [WAP05], 2003 - 2005

- Kaiyu Wan introduces a notion of contexts as first-class values in Lucid, thereby making Lucx the true intensional language.

8. Onyx [Gro04], April 2004.

- Peter Grogono makes an experimental derivative of Lucid – Onyx to investigate on lazy evaluation of arrays.

9. GLU# [PK04], 2004

- GLU# is an evolution of GLU where Lucid is embedded into C++.

2.2.2 Indexical Lucid

When Indexical Lucid came into existence, it allowed accessing context properties in multiple dimensions. Prior Indexical Lucid, the only implied dimension was a set of natural numbers. With Indexical Lucid, we can have more than one dimension, and we can query for a part of the context (any dimensions of it). Thus, the syntactic definition has been amended to include an ability to specify which dimensions exactly we are working on.

2.2.2.1 Streams

Lucid variables and expressions are said to be *streams* of values, through which one can navigate using some sort of navigational operators. In the natural language example given earlier the operators were *before()*, *after()*, and *at()*; here we begin by introducing *first()* and *next()* (very much like in LISP).

If the following equations hold¹:

- **FIRST** $X = 0$
- **NEXT** $X = X + 1$ (like **succ** in LISP)

¹Note, these are initial conditions of a definition to illustrate the ideas behind the streams and not an actual declaration of constructs in the language one would normally write.

where 0 is a stream of 0's: $(0, 0, 0, \dots, 0, \dots)$. Likewise, 1 is a stream of 1's, and the '+' operator performs pair-wise addition of the elements in the streams according to the implied current dimension index. Thus, X is defined as a stream, such that:

- $x_0 = 0, x_{i+1} = x_i + 1$, or
- $X = (x_0, x_1, \dots, x_i, \dots) = (0, 1, \dots, i, \dots)$

Similarly, if:

- **FIRST** $X = X$
- **NEXT** $Y = Y + \text{NEXT } X$

Y here becomes a running sum of X :

- $y_0 = x_0; y_{i+1} = y_i + x_{i+1}$
- $Y = (y_0, y_1, \dots, y_i, \dots) = (0, 1, \dots, i(i+1)/2, \dots)$

2.2.2.2 Basic Operators

This section defines properties of basic Lucid operators, which were proven by Paquet in [Paq99].

Operator FBY. Operator **FBY** stands for “followed by”. **FBY** allows simply to suppress dimension index and switch to another stream. As an example the previously shown streams X and Y can be defined as follows using **FBY**:

- $X = 0 \text{ FBY } X + 1 = (0, 1, 2, \dots, i, \dots)$
- $Y = X \text{ FBY } Y + \text{NEXT } X = (0, 1, \dots, i(i+1)/2, \dots)$

To provide an analogy to lists, we can say that that the following operators are equivalent:

- **FIRST** and **hd**
- **NEXT** and **tl**
- **FBY** and **cons**

Informal Definition of **FIRST**, **NEXT**, **FBY**.

- Definitions:
 - **FIRST** $X = (x_0, x_0, \dots, x_0, \dots)$
 - **NEXT** $X = (x_1, x_2, \dots, x_{i+1}, \dots)$
 - X **FBY** $Y = (x_0, y_0, y_1, \dots, y_{i-1}, \dots)$
- These are the three operators of the original Lucid.
- Indexical Lucid has come into existence with the ability to access an arbitrary element by some index i in the stream.

Operators **wvr, **asa**, and **upon**.** The other three operators that are slightly more complex informally defined below:

- X **wvr** $Y =$
 if **FIRST** $Y \neq 0$
 then X **FBY** (**NEXT** X **wvr** **NEXT** Y)
 else (**NEXT** X **wvr** **NEXT** Y)
- X **asa** $Y =$ **FIRST** (X **wvr** Y)
- X **upon** $Y =$
 X **FBY**
 (**if** **FIRST** $Y \neq 0$ **then** (**NEXT** X **upon** **NEXT** Y) **else** (X **upon** **NEXT** Y))

where **wvr** stands for *whenever*, **asa** stands for *as soon as* and **upon** stands for *advances upon*. **wvr** chooses from its left-hand-side operand only values in the current dimension where the right-hand-side evaluates to *true*. **asa** returns the value of its left-hand-side as a first point in that stream as soon as the right-hand-side evaluates to *true*. Unlike **asa**, **upon** switches context of its left-hand-side operand if the right-hand side is *true*.

2.2.2.3 Sequentiality Problem

With tagged-token dataflows of the original Lucid operators one could only define an algorithm with pipelined, or sequential, data flow:

- It is wasteful use of computing resources (e.g. to compute an element i we need $i - 1$, but $i - 1$ may never be used/needed otherwise).
- Sequential access to the stream of values.

2.2.2.4 Random Access to Streams

New intensional operators are introduced to remedy the sequentiality problem: @ and #. The operators are used as an index # corresponding to the current position that allows querying the current context, and @ is intensional navigation to switch the context. With @ and #:

- the computation is defined according to a context (here a single integer),
- Lucid is no longer a data-flow language and is on the road to intensional programming, and
- the previously introduced intensional operators can be redefined in terms of the operators # and @.

In terms of the three original operators of **FIRST**, **NEXT**, and **FBY** the operators **@** and **#** are defined as follows:

Definition 1

$$\# = 0 \text{ fby } (\# + 1)$$

$$X@Y = \text{if } Y = 0 \text{ then FIRST } X \text{ else (NEXT } X)@(Y - 1)$$

Both X and Y in the above definition are variable streams, and their current values are determined by their current context at the time of evaluation. To redefine the meaning of **@** and **#** Paquet uses the denotational form, with the following proposition:

Proposition 1

$$(1) [\#]_i = i$$

$$(2) [X@Y]_i = [X]_{[Y]_i}$$

where (1) means the value of **#** at the current context i is i itself (i.e. we query the value of our current dimension), and (2) says that evaluate Y at the current context i and then use Y as a new context for X .

2.2.2.5 Definition of Lucid Operators By Means of **@** and **#**

First we present the definition of the operators via **@** and **#** denoted in monospaced font, and then we will provide their equivalence to the original Lucid operators, denoted as **SMALL CAPS**.

Definition 2

$$(1) \text{first } X = X@0$$

$$(2) \text{next } X = X@(\# + 1)$$

$$(3) X \text{ fby } Y = \text{if } \# = 0 \text{ then } X \text{ else } Y@(\# - 1)$$

$$(4) X \text{ wvr } Y = X@T \text{ where}$$

$$T = U \text{ fby } U@(T + 1)$$

$$U = \text{if } Y \text{ then } \# \text{ else next } U$$

end

$$(5) X \text{ asa } Y = \text{first } (X \text{ wvr } Y)$$

$$(6) X \text{ upon } Y = X@W$$

$$\text{where } W = 0 \text{ fby (if } Y \text{ then } (W + 1) \text{ else } W) \text{ end}$$

op	$::=$	$intensional-op$
		$data-op$
$intensional-op$	$::=$	$i-unary-op$
		$i-binary-op$
$i-unary-op$	$::=$	$first \mid next \mid prev$
$i-binary-op$	$::=$	$fby \mid wvr \mid asa \mid upon$
$data-op$	$::=$	$unary-op$
		$binary-op$
$unary-op$	$::=$	$! \mid - \mid iseod$
$binary-op$	$::=$	$arith-op$
		$rel-op$
		$log-op$
$arith-op$	$::=$	$+ \mid - \mid * \mid / \mid \%$
$rel-op$	$::=$	$< \mid > \mid <= \mid >= \mid == \mid !=$
$log-op$	$::=$	$\&\& \mid \parallel$

Figure 1: Concrete Indexical Lucid Syntax

E	$::=$	id
		$E(E_1, \dots, E_n)$
		if E then E' else E''
		$\#E$
		$E@E'E''$
		E where Q

Figure 2: GIPL Expressions

2.2.2.6 Abstract Syntax of Lucid

Abstract and concrete syntaxes of Lucid for expressions, definitions, and operators are presented in Figure 2, Figure 3, and Figure 1 for both Indexical Lucid and GIPL.

2.2.2.7 Concrete GIPL Syntax

The GIPL is the generic programming language of all intensional languages, defined by the means of only two intensional operators – @ and #. It has been proven that other intensional programming languages of the Lucid family can be translated into the GIPL [Paq99]. The concrete syntax of the GIPL is presented in Figure 4. It

```

Q ::= dimension id
   | id = E
   | id(id1, id2, ..., idn) = E
   | QQ

```

Figure 3: GIPL where Definitions

```

E ::= id
   | E(E, ..., E) #LUCX
   | E[E, ..., E] (E, ..., E) #GIPL
   | if E then E else E fi
   | # E
   | E @ [E:E] #GIPL
   | E @ E #LUCX
   | E where Q end;
   | [E:E, ..., E:E] #LUCX
   | iseod E; #INDEXICAL
Q ::= dimension id, ..., id;
   | id = E;
   | id(id, ..., id) = E; #LUCX
   | id[id, ..., id] (id, ..., id) = E; #GIPL
   | QQ

```

Figure 4: Concrete GIPL Syntax

has been amended to support the `iseod` operator of Indexical Lucid for completeness and influenced by the productions from Lucx [WAP05] to allow contexts as first-class values while maintaining backward compatibility to the GIPL language designed by Paquet in [Paq99].

2.2.2.8 Semantic Rules

Paquet’s PhD thesis [Paq99] presents details of the operational semantics of GIPL recited here for the unaware reader with a brief description. Figure 5 provides initial operational semantic rules for Indexical Lucid in Hoare Logic [Moe04, HJ02]. Later on, these rules are extended to support free Java methods and Java objects in JLucid and Objective Lucid respectively in Chapter 3.

Notation

- \mathcal{D} represents the definition environment where all symbols are defined (a dictionary of identifiers).
- $\mathcal{D}, \mathcal{P} \vdash E : a$ represents current context of evaluation (a set of dimensions \mathcal{P}) and the dictionary that yields a specified result a under that context given expression E .
- `const`, `op`, `dim`, `func`, and `var` represent what kind of construct types are put into \mathcal{D} as constants, operators, dimensions, functions, and variables respectively.
- the \mathbf{E}_{Xid} type of rules place different identifier types listed above into the definition environment \mathcal{D} .
- the remaining \mathbf{E}_{xyz} -style rules correspond to the execution (or rather application of) of the operators, functions, and conditionals to their argument expressions given the definition of them in \mathcal{D} and the current context. Thus, \mathbf{E}_{op} specifies application of a defined operator function f in the current context to its arguments (usually one for unary operators and two for binary); \mathbf{E}_{fct} applies the named function to its arguments translating the formal arguments to actual; \mathbf{E}_{cT} and \mathbf{E}_{cF} correspond to conditional evaluation of the `then` and `else` branching clauses; \mathbf{E}_{at} and \mathbf{E}_{tag} correspond to the universal intensional operators `@` and `#` for switching of and querying for the current context; and \mathbf{E}_{w} corresponds to the scope definition marked by the `where` clause.
- the \mathbf{Q} -style rules allow definitions within the scope of the dimension \mathbf{Q}_{dim} and variable identifier \mathbf{Q}_{id} types and their composition.

$$\begin{array}{l}
\mathbf{E}_{\text{cid}} : \frac{\mathcal{D}(id) = (\text{const}, c)}{\mathcal{D}, \mathcal{P} \vdash id : c} \\
\mathbf{E}_{\text{opid}} : \frac{\mathcal{D}(id) = (\text{op}, f)}{\mathcal{D}, \mathcal{P} \vdash id : id} \\
\mathbf{E}_{\text{did}} : \frac{\mathcal{D}(id) = (\text{dim})}{\mathcal{D}, \mathcal{P} \vdash id : id} \\
\mathbf{E}_{\text{fid}} : \frac{\mathcal{D}(id) = (\text{func}, id_i, E)}{\mathcal{D}, \mathcal{P} \vdash id : id} \\
\mathbf{E}_{\text{vid}} : \frac{\mathcal{D}(id) = (\text{var}, E) \quad \mathcal{D}, \mathcal{P} \vdash E : v}{\mathcal{D}, \mathcal{P} \vdash id : v} \\
\mathbf{E}_{\text{op}} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}(id) = (\text{op}, f) \quad \mathcal{D}, \mathcal{P} \vdash E_i : v_i}{\mathcal{D}, \mathcal{P} \vdash E(E_1, \dots, E_n) : f(v_1, \dots, v_n)} \\
\mathbf{E}_{\text{fct}} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}(id) = (\text{func}, id_i, E') \quad \mathcal{D}, \mathcal{P} \vdash E'[id_i \leftarrow E_i] : v}{\mathcal{D}, \mathcal{P} \vdash E(E_1, \dots, E_n) : v} \\
\mathbf{E}_{\text{cT}} : \frac{\mathcal{D}, \mathcal{P} \vdash E : \text{true} \quad \mathcal{D}, \mathcal{P} \vdash E' : v'}{\mathcal{D}, \mathcal{P} \vdash \text{if } E \text{ then } E' \text{ else } E'' : v'} \\
\mathbf{E}_{\text{cF}} : \frac{\mathcal{D}, \mathcal{P} \vdash E : \text{false} \quad \mathcal{D}, \mathcal{P} \vdash E'' : v''}{\mathcal{D}, \mathcal{P} \vdash \text{if } E \text{ then } E' \text{ else } E'' : v''} \\
\mathbf{E}_{\text{tag}} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}(id) = (\text{dim})}{\mathcal{D}, \mathcal{P} \vdash \#E : \mathcal{P}(id)} \\
\mathbf{E}_{\text{at}} : \frac{\mathcal{D}, \mathcal{P} \vdash E' : id \quad \mathcal{D}(id) = (\text{dim}) \quad \mathcal{D}, \mathcal{P} \vdash E'' : v'' \quad \mathcal{D}, \mathcal{P} \dagger[id \mapsto v''] \vdash E : v}{\mathcal{D}, \mathcal{P} \vdash E @E' E'' : v} \\
\mathbf{E}_{\text{w}} : \frac{\mathcal{D}, \mathcal{P} \vdash Q : \mathcal{D}', \mathcal{P}' \quad \mathcal{D}', \mathcal{P}' \vdash E : v}{\mathcal{D}, \mathcal{P} \vdash E \text{ where } Q : v} \\
\mathbf{Q}_{\text{dim}} : \frac{}{\mathcal{D}, \mathcal{P} \vdash \text{dimension } id : \mathcal{D} \dagger[id \mapsto (\text{dim})], \mathcal{P} \dagger[id \mapsto 0]} \\
\mathbf{Q}_{\text{id}} : \frac{}{\mathcal{D}, \mathcal{P} \vdash id = E : \mathcal{D} \dagger[id \mapsto (\text{var}, E)], \mathcal{P}} \\
\mathbf{QQ} : \frac{\mathcal{D}, \mathcal{P} \vdash Q : \mathcal{D}', \mathcal{P}' \quad \mathcal{D}', \mathcal{P}' \vdash Q' : \mathcal{D}'', \mathcal{P}''}{\mathcal{D}, \mathcal{P} \vdash Q Q' : \mathcal{D}'', \mathcal{P}''}
\end{array}$$

Figure 5: Operational Semantics of GIPL

2.2.2.9 Examples of Lucid Programs

Two simple examples of Lucid programs are presented. The examples demonstrate absence of iterative/sequential operation as opposed to the traditional imperative programming languages.

Natural Numbers Problem An example program in Indexical Lucid that yields 44 as the result is in Figure 6. The way the program is expanded using the re-definitions of the Lucid operators, such as `fbby`, employing `@` and `#` in GIPL is shown in Figure 7.

```
N @.d 2
where
  dimension d;
  N = 42 fby.d (N + 1);
end;
```

Figure 6: Natural numbers problem in Indexical Lucid.

```
N @.d 2
where
  dimension d;
  N = if (#.d <= 0) then 42 else (N + 1) @.d (#.d - 1) fi;
end;
```

Figure 7: Natural numbers problem in GIPL.

The Hamming Problem This example (see Figure 8) illustrates the simple use of functions in Lucid.

```
H
where
  H = 1 fby merge(merge(2 * H, 3 * H), 5 * H);

  merge(x, y) = if(xx <= yy) then xx else yy
  where
    xx = x upon(xx <= yy);
    yy = y upon(yy <= xx);
  end;
end;
```

Figure 8: Indexical Lucid program implementing the `merge()` function.

2.2.3 Lucid Now

To summarize, Lucid is a functional programming language where a variable (stream), a function, a dimension, or even entire context can be a *first class value* (i.e. can be viewed and manipulated as data). Lucid provides operators, such as @ and #, to navigate within dimensions and switch contexts. The language also exhibits the educative execution model (demand-driven distributed computation) that augments the semantics with a warehouse (intensional value cache) and its consistency².

2.3 Hybrid Programming

There have been previous approaches to couple intensional or functional and imperative and object-oriented paradigms prior to this work. Some recent related work on the same issue is presented in [BM96, PK04, MS01, SM02] with the [PK04] being the most relevant. The two major approaches of addressing the OO issue are – either (1) to extend Lucid to become object-oriented or objects-aware or (2) make a host imperative language be extended to embed Lucid. The authors of [PK04] chose the latter by extending GLU-with-C to GLU#-with-C++, whereas this work approaches the problem from Lucid to Java. This means a Lucid program is the main one driving the computation. We will briefly consider the following approaches to the hybrid programming:

- ML_{\leq}
- FC++
- GLU
- GLU#

²Paquet defines the augmented operational semantics in [Paq99] and Tao implements its first incarnation in GIPSY [Tao04]. This work has an impact on this aspect by introducing the side effects with the imperative languages, which will be discussed later.

2.3.1 ML_{\leq}

ML_{\leq} [BM96] is a system introduced in 1996 that proposed to marry OOP and functional paradigms using their own language and providing the details of the predicative and decidable typing rules and operational semantics of such a system. Their main goal is to be able to induce implicit polymorphism of functional languages in objects. They do not extend an existing functional language with the OO capabilities, instead they reinterpret all data types as either abstract or concrete classes and use the dynamic dispatch, a typical OO feature, on run-time types.

2.3.2 FC++

FC++ [MS01, SM02] tries to promote the functional paradigm in C++. FC++ is a library add-on to enable higher-order polymorphic functions in a novel use of C++ type inference that is not very complex and is still expressive. FC++ adds support for both parametric and subtype polymorphism policies for functions in order to be able to fit FC++ functions within the C++ object model and pass higher-order functions as parameters. The FC++ functions are kept as objects called *functoids* and use a reference counter machinery for allocation and de-allocation. Closures in FC++ (operation on a some state and the state itself) can automatically be created during functoid object creation, but their “closing” of that state is not automatic and the state values have to be passed explicitly during the creation process. The library also adds a set of functional operators from the Haskell Standard Prelude. FC++ comes more from the OOP-to-functional point of view and conforms with standard software engineering design patterns and is suitable for the common OO tasks.

2.3.3 GLU

GLU was the most general intensional programming tool recently available [JD96]. However, experience has shown that, while being very efficient, the GLU system suffers from a lack of flexibility and adaptability [Paq99]. Given that Lucid is evolving continually, there is an important need for the successor to GLU to be able to stand the heat of evolution [Paq99]. The two major successors of GLU are the GIPSY and GLU# systems.

Eduction

The earlier mentioned notion of eduction was first introduced by the GLU compiler. GLU supports so-called *tagged-token demand-driven dataflow* where data elements (tokens) are computed on demand following a dataflow network defined in Lucid. Data elements flow in the normal flow direction (from producer to consumer) and demands flow in the reverse order, both being tagged with their current context of evaluation. This form of lazy computation is inherited by GIPSY from GLU.

2.3.4 GLU#

GLU# [PK04] is a successor of GLU, which enables Lucid within C++. The authors argue for the embedding small functional/intensional-language pieces of Lucid into C++ programs allowing lazy (demand-driven) evaluation of arrays and functions thereby making Lucid easily accessible within a popular imperative programming language, such as C++. Because GLU# appeared quite recently (2004) to when this work was written, its success compared to GLU is yet to be evaluated; however, it seems to suffer from the same inflexibility GLU did and targets only C++ as a host language.

2.4 Compiler Frameworks

A significant number of compiler frameworks emerged for the past decade. All try to enable compilation of more than one language, either hybrid or not, in an uniform manner. Some frameworks or libraries became “frozen” (i.e. non-extendable) and fixed to a specific set of languages, some other ones were build with the extension in mind, so it is relatively easy to “plug-in” yet another compiler into the system (a collection of compilers and the necessary tools) with minimum integration work required. A brief overview of different compiler frameworks is given next:

- GLU tried to accommodate Fortran, C, and Lucid in one system, but was made so inflexible [Paq99] that it would take a significant effort to extend it and add other languages to the system.
- GLU# merges Lucid and C++; however, makes no provisions for extension to other languages on either intensional or imperative side.

- Microsoft .NET can also be thought of a commercial heterogeneous compiler framework (it is more than a compiler framework, but our focus is on compilers) that allows easy cooperation and application development between different language models, such as C#, C++, Visual Basic, and Assembly in a homogeneous environment. However, none of these languages have natively any of the intensional or functional capabilities, so no native debugging support or other tools exist, even if one starts using FC++ or GLU# in this environment. Despite the fact that all programs can be compiled into the common bytecode, the debugging tools have to be aware of the functional paradigms on a higher level and they are not (at least at this writing).
- The GNU Compiler Collection (GCC) can also be said as a compiler framework from the free software [CP05]. It supports C, C++, Objective-C, Objective-C++, Java, Fortran, and Ada. Again, these languages are more of an imperative nature, but it is far easier to add new language into GCC than to Microsoft .NET due to its openness.
- Finally, the GIPSY presents the GIPC framework that is designed for expansion and integration of the intensional and imperative (and later functional) languages. This is presented through the rest of this thesis.

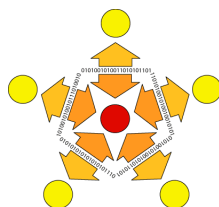


Figure 9: The GIPSY Logo representing the distributed nature of GIPSY.

2.5 General Intensional Programming System

2.5.1 Introduction

GIPSY is broadly presented in [WPG03, Lu04, PW05], and others. Please refer to the online resources [RG05a, PW05, RG05b] to obtain the most current status of the project. GIPSY is primarily implemented in Java. General GIPSY architecture is presented in Figure 10. The essence behind GIPSY is demand-driven computation support for the intensional programming languages, e.g., Indexical Lucid, Tensor Lucid [Paq99], etc.

The GIPSY consists in three modular sub-systems: the General Intensional Programming Language Compiler (GIPC); the General Education Engine (GEE), and the Intensional Run-time Interactive Programming Environment (RIPE). The sub-systems have to be modular so that one implementation of parts of them or the whole can be replaced by another without having major if any impact on the other modules. Although the theoretical basis of the language has been settled, the implementation of an efficient, general and adaptable programming system for this language raises many questions. The following sections outline the theoretical basis and architecture of the different components of the system. All these components are designed in a modular manner to permit the eventual replacement of each of its components – at compile-time or even at run-time – to improve the overall efficiency and productivity of the system [Paq99].

A GIPSY instance sends out little bits of work to others to compute and then gathers the results in distributed fashion. Of course, synchronization, latency tolerance, and maximum utilization of resources are primary goals for the system to be productive. Unlike in most programming language models (see [ST98]) considered for parallel computation, in GIPSY several key concepts are considered:

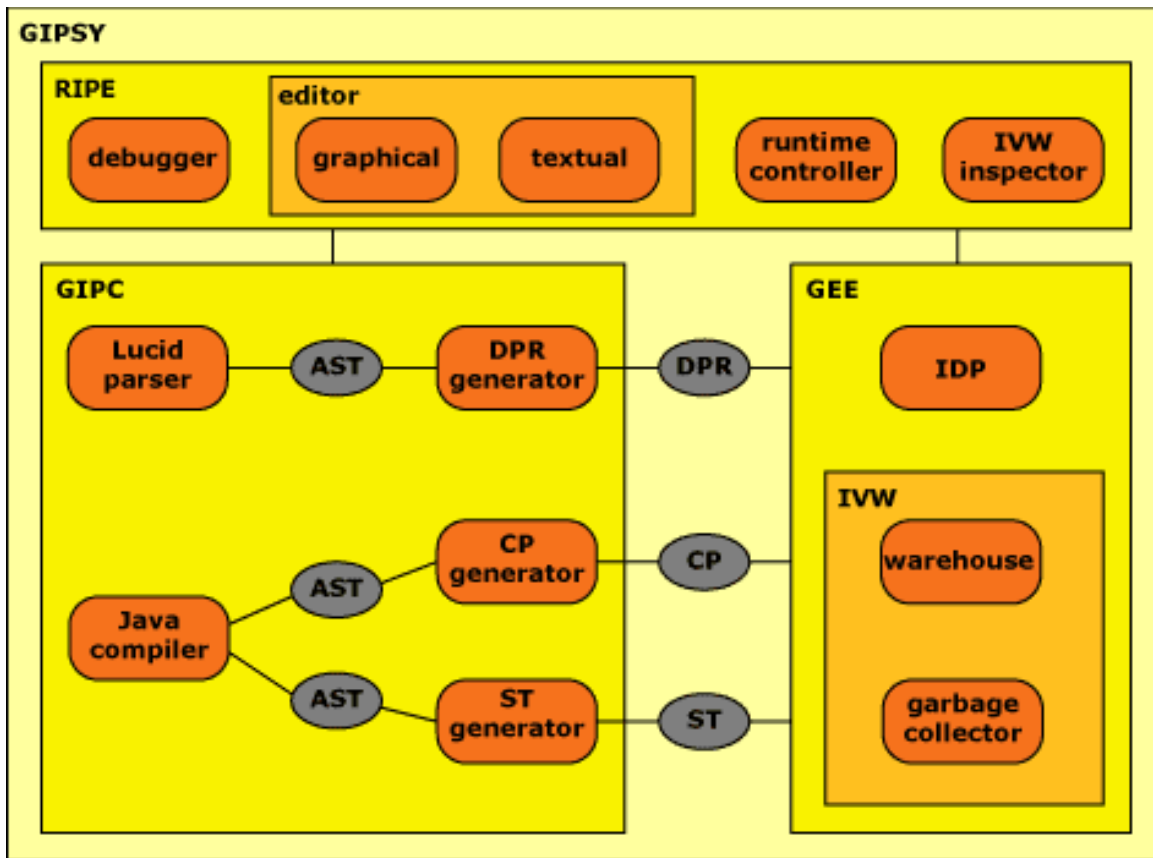


Figure 10: Structure of the GIPSY

- Thread-Level Parallelism (TLP)
- Stream-Level Parallelism (SLP)
- Cluster-Level Parallelism (CLP)

GIPSY's parallelism granularity takes into account the amount of TLP, SLP, and CLP available. TLP determines the maximum number of threads that should or can be created when a Lucid program is being executed. In other words, TLP defines on how many pieces of terminal computational work we can chop a big job into. The goal, as far as programming is concerned, is to program for infinite TLP, and later adjust (load-balance) at run-time to the actual amount of SLP. SLP determines the maximum number of *streams* available to execute the threads. Here, by "streams" we mean processors but, with the invention of multithreaded CPUs for a single processor, there may be several thread streams available in parallel, and hence a more general notion of SLP. The amount of SLP is machine-dependent and has to be discovered at

run-time on remote machines. If a job is to be run on a single machine, GIPSY tries to maximize SLP utilization, providing just enough TLP for the machine in question with the design goal of always assuming infinite TLP. Then load-balancing comes into play. CLP takes GIPSY to another level — distributed computing, involving utilization of SLP of the machines across the network nearby or across the globe over the Internet.

NOTE: the Lucid family of languages has also a notion of streams that refers to Lucid variables that evaluate in multiple contexts. Every Lucid stream (e.g. a variable) can potentially be evaluated on any hardware stream available, but it is important not to confuse the two kinds of streams. The reason for the existence of the two notions is that both terms were used independently in each field. Now that parallel architectures and language models such as Lucid came into proximity, the terms clash.

2.5.2 Goals

The system has to withstand the evolution of the tools, languages, and underlying platforms, thus be flexible and adaptable to the changes. That is one of the most important and stringent requirements put on the development of GIPSY [Pdq99]. Other subordinate requirements in compiler design, run-time system, communication, and user interfaces are presented in detail throughout the follow up sections.

2.5.3 General Intensional Programming Compiler

GIPSY programs are compiled in a two-stage process (see Figure 33, page 65). First, the intensional part of the GIPSY program is translated in Java, then the resulting Java program is compiled in the standard way.

The source code consists of two parts: the Lucid part that defines the intensional data dependencies between variables and the sequential part that defines the granular sequential computation units (usually written in any imperative language, e.g. C or Java). The Lucid part is compiled into an intensional data dependency structure (IDS) describing the dependencies between each variable involved in the Lucid part. This structure is interpreted at run-time by the GEE following the demand propagation mechanism. Data communication procedures used in a distributed evaluation

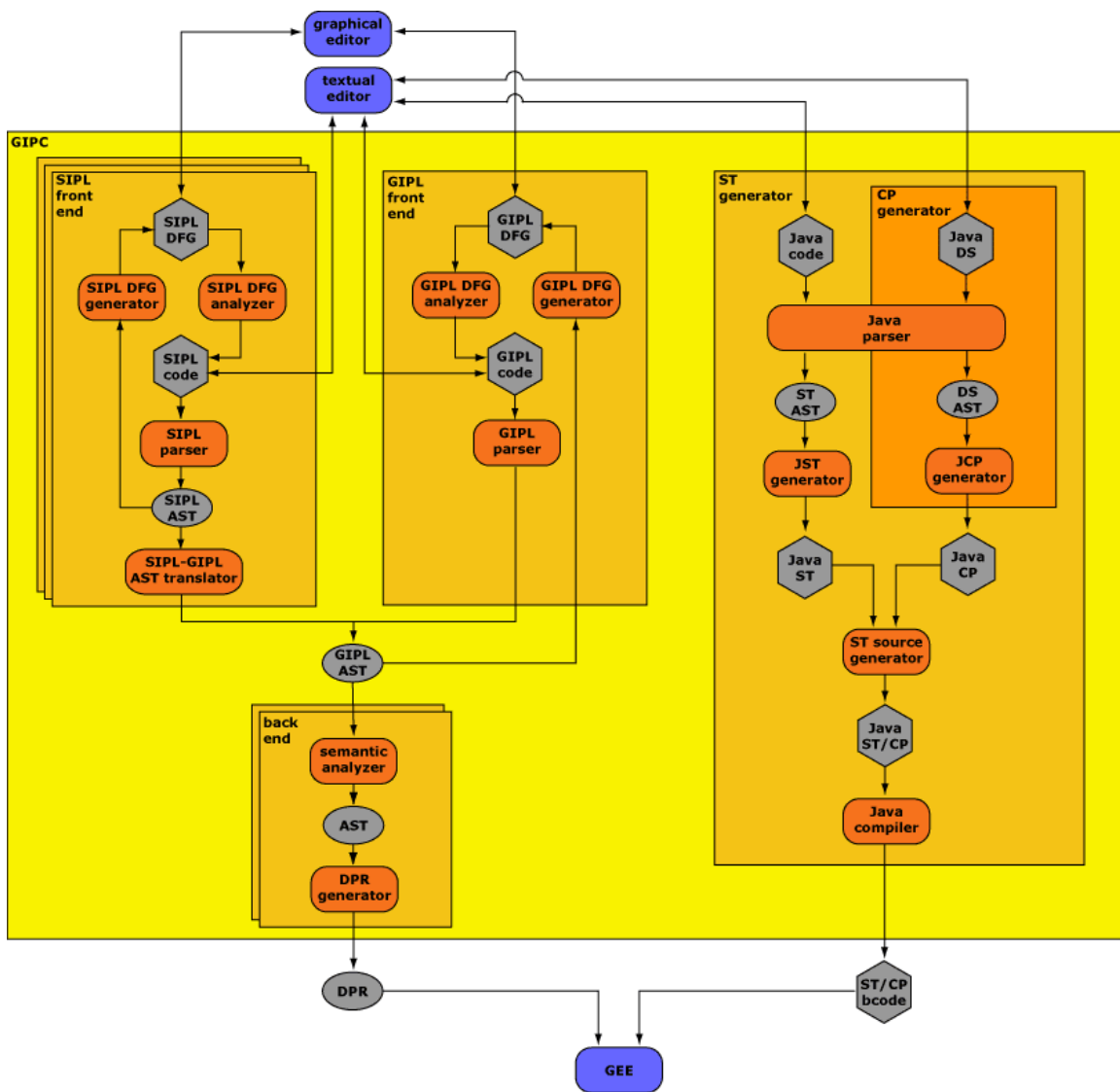


Figure 11: Initial Conceptual Design of the GIPC

of the program are also generated by the GIPC according to the data structures definitions written in the Lucid part, yielding a set of communication procedures (CP). These are generated following a given communication layer definition such as provided by RPC (or rather RMI since GIPSY is implemented in Java), CORBA, Jini, or the WOS [BKU98]. The sequential functions defined in the second part of the GIPSY program are translated into imperative code using the second stage imperative compiler syntax, yielding imperative sequential threads (ST). Intensional function definitions, including higher order functions, will be flattened using a well-known efficient technique [Ron94, Paq99]. The closures in the higher order functions case are still applicable because the function state and the operation on it are correctly passed to the functions by expanding and using function definitions inline. The insignificant limitation here is that self-referential closures for such functions cannot be made. The function elimination in GIPSY pertinent to some of these aspects was implemented by Wu in [Wu02].

The Figure 11 presents the initial conceptual design of the GIPC. Based on this design, the GIPSY module integration and the development of the STs and CPs support has begun. Later on the design was refined in [PGW04, MP05a] and its latest reincarnation is shown in Figure 38 in Chapter 4, page 77; thus, the evolution description is delayed until then.

Prior this work, GIPC supported only two Lucid dialects: GIPL and Indexical Lucid. The initial GIPC compiler was implemented by Chun Lei Ren in [Ren02], and the translation of the Indexical Lucid into GIPL and the semantic analysis was implemented by Aihua Wu in [Wu02]. A large integration and re-engineering effort went into GIPC to approach it to the goals of the GIPSY (see Section 2.5.2) and add more compilers for investigation of the underlying language models. The results of this effort are presented in the Design and Implementation chapter (Chapter 4).

2.5.4 General Education Engine

The GIPSY uses a demand-driven model of computation, which is based on the principle is that certain computation takes effect *only* if there is an explicit demand for it. The GIPSY uses education, which is demand-driven computation in conjunction with an intelligent value cache called a warehouse. Every demand can potentially generate a procedure call, which is either computed locally or remotely, thus eventually in

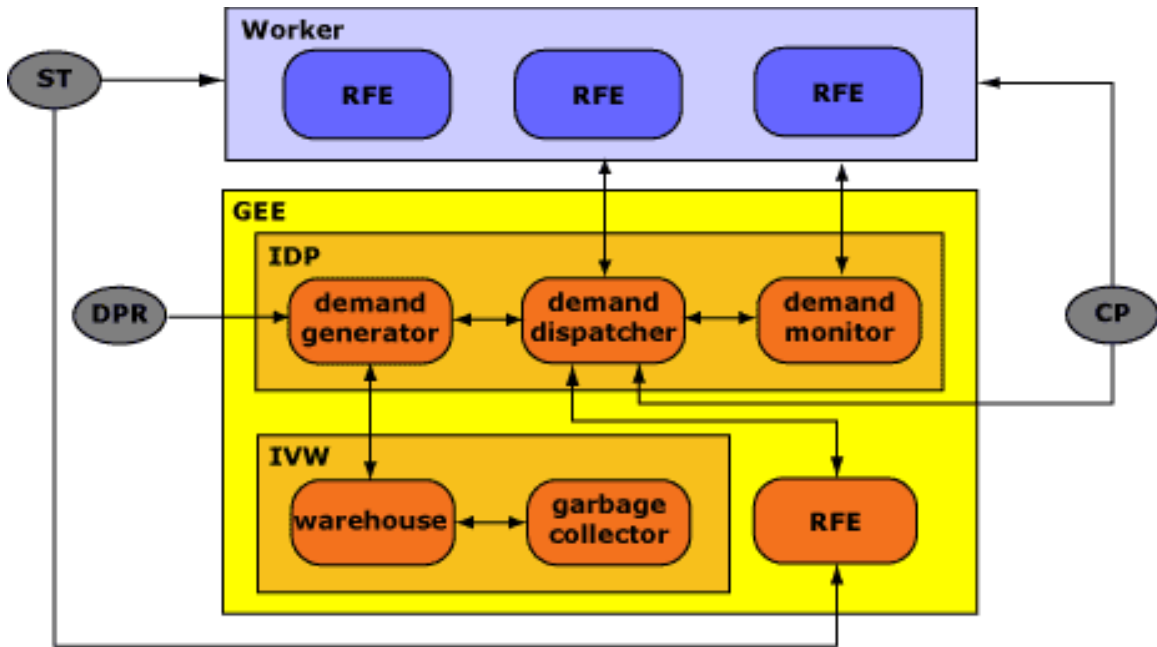


Figure 12: Conceptual Design of the GEE

parallel with other procedure calls. Every computed value is placed in the warehouse, and every demand for an already-computed value is extracted from the warehouse rather than computed again and again (demands that may have side effects, e.g. if we cache results of STs, shall not be cached). Eduction, thus, reduces the overhead induced by the procedure calls needed for the computation of demands sequentially. Figure 12 describes the internal conceptual structure and functioning of the GEE.

The GEE itself is composed of three main modules: the executor, the intensional demand propagator (IDP), and the intensional value warehouse (IVW). First, the intensional data dependency structure (IDS, which represents GEER) is fed to the demand generator (DG) by the compiler (GIPC). This data structure represents the data dependencies between all the variables in the Lucid part of the GIPSY program. This tells us in what order all demands are to be generated to compute values from this program. The demand generator receives the initial demand, that in turn raises the need for other demands to be generated and computed as the execution progresses. For all non-functional demands (i.e. demands not associated with the execution of sequential threads (ST)), the DG makes a request to the warehouse to see if this demand has already been computed. If so, the previously computed value is extracted from the warehouse. If not, the demand is propagated further, until the original

demand resolves to a value and is put in the warehouse for further use. This type of warehousing was introduced by GLU due to its distributed nature to cut down on communication costs, but it can certainly be applicable to any functional language, such as LISP, Scheme, Haskell, ML and others to improve efficiency even on a single machine provided there are no any side effects whatsoever. The garbage collector can run on the background to clean up old function-parameters-values tuples periodically, and given that the large amounts of memory are cheap these days functional languages may gain much more popularity with the increased performance.

For functional demands (i.e. demands associated with the execution of a sequential thread), the demands are sent to the demand dispatcher (DD) that takes care of sending the demand to one of the workers or to resolve it locally (which normally means that a worker instance is running on the processor running the generator process). If the demands are sent to a remote worker, the communication procedures (CP) generated by the compiler are used to communicate the demand to the worker. The demand dispatcher (DD) receives some information about the liveness and efficiency of all workers from the demand monitor (DM), to help it make better decisions in dispatching the demands.

The demand monitor, after some functional demands are sent to workers, starts to gather various types of information about each worker, including, but not limited to:

- liveness status (is it still alive, not responding, or dead)
- network link performance
- response time statistics for all demands sent to it

These data points are accessed by the DD to make better decisions about the load balancing of the workers, and thus achieving better overall run-time efficiency.

Bo Lu was the first one to do the original design of the GEE framework [Lu04] and investigate its performance under threaded and RMI environments. She also introduced the notion of the Identifier Context (IC) classes – demands converted into Java code and using Java Reflection [Gre05] to compile, load, and execute them at run-time. She also contributed the first version of the interpreter-based execution engine. Next, Lei Tao contributed the first incarnation of the intensional

value warehouse and garbage collection mechanisms in [Tao04] based on the popular scientific library called NetCDF. The author of this thesis put an effort to modularize these all and make them easier to extend and customize. He also provided the initial GEE application to start available network services. The GEE was also made aware of the STs and CPs as well as the new type system, described in Section 4.1.1.5. Further, Emil Vassev [VP05] produced a very general and functional framework for demand migration and its implementation, Demand Migration System (DMS) that supports among other things Jini, CORBA, and .NET Remoting for fault-tolerant demand transportation system, a part of the Demand Dispatcher. The DMS is still pending integration as of this writing.

2.5.4.1 Demand Propagation Resources for the GEE

The IDP generates and propagates demands according to the data dependence structure (DPR, now renamed to GEER in [WPG03]) generated by the GIPC. If a demand requires some computation, the result can be calculated either locally or on a remote computing unit. In the latter case, the communication procedures (CP) generated by the GIPC are used by the GEE to send the demand to the worker. When a demand is made, it is placed in a demand queue, to be removed only when the demand has been successfully computed. This way of working provides a highly fault-tolerant system. One of the weaknesses of GLU is its inability to optimize the overhead induced by demand-propagation. The IDP will remedy to this weakness by implementing various optimization techniques:

- Data blocking techniques used to aggregate similar demands at run time, which will also be used at compile-time in the GIPC for automatic granularization of data and functions for data-parallel applications
- The performance-critical parts (IDP and IVW) are designed as replaceable modules to enable run-time replacements by more efficient versions adapted to specific computation-intensive applications
- Certain demand paths identified (at compile-time or run-time) as critical will be compiled to reduce their demand propagation overhead
- Extensive compile-time and run-time rank analysis (analysis of the dimensionality of variables) [Dod96].

2.5.4.2 Synchronization

Distributed vs. Parallel

It is important to make a distinction between parallel and distributed computing. In parallel computing, SLP matters and latency tolerance for memory references with mostly UMA (uniform memory access) characteristics, whereas in distributed computing communication is much more expensive (and perhaps even prohibitive) and CLP matters as well. This setup largely exhibits NUMA (non-UMA) characteristics (see [Pro03b]) and latency tolerance (and so also fault tolerance) has a higher significance. This greatly impacts the way we synchronize in parallel and distributed worlds.

Synchronization in Distributed Environment A distributed environment is a very popular domain these days, so we'll start with it first. Typically, the network is the scarce resource and is the bottleneck for a distributed application because it implies communication (e.g., MPI), which is often unacceptable. Therefore, many distributed applications choose not to communicate at all or communicate very little through message passing. This implies blocking on waiting for the network requests to propagate, i.e. network latency.

Synchronization in Parallel Environment Synchronization in a parallel environment is more fine-grained, often at the hardware level (e.g., a full/empty bit in memory cells). Java does not give us control over such synchronization, so we have to rely on the JVM built for an architecture that has such synchronization. The JVM has to be developed to make use of the full/empty bits that are usually represented as *future variables* [Pro03b, JA03] in the languages specifically designed for parallel computing.

Secure Synchronization

Secure synchronization is especially pertinent in a distributed environment. Like any act of communication within worker-generator architecture (see Section 3.3.3.4) and a warehouse (Figure 33; Section 2.5.4.1), synchronization has to be secure to avoid (a) *over-demanding*, (b) incorrect results sent back, (c) loss of results and demands, and (d) poisoning the warehouse with wrong data. Secure synchronization implies

fault tolerance. In GIPSY, we will rely on Java's RMI and Jini over JSSE for secure communication in a distributed environment, using Java's synchronization primitives (see Section 2.5.4.2) to achieve the goal of secure synchronization. Thus, the reliability and accountability of the results of a GIPSY program are dependent on these properties of underlying Java Runtime Environment (JRE) and the communication protocols used.

Implicit vs. Explicit Synchronization

One of the productivity metrics of a software completing its task on time, is the efficiency of development of (see [Pro03c]) such a software, i.e., the amount of programmer's effort required to create and debug the software. This is essentially a metric, called time-to-solution (TTS) [Pro03c]; from creation until the end result (e.g. completion of some scientific computation). The goal is to minimize TTS. One way to achieve this is ease of programming. As the proportion of the work done by the compiler increases, so does the reliability of the code, but we target scientific researchers, not just programmers. Scientific researchers from math and physics should not care about these issues and, thus, just be concerned mastering the basics of Lucid. Therefore, the programmer has to be freed from taking care of synchronization explicitly, which a source of bugs and inefficiency of programming (e.g., using Java's synchronization primitives, such as `synchronized`, `Object.wait()`, `Object.notify()`, and `Object.notifyAll()`, [Fla97]). The programmer should rather focus on the problem being solved and let the compiler/run-time system deal with the synchronization pain. The GIPSY system, built around the Lucid family, advocates implicit synchronization either by wrapping around the Java's synchronization primitives or through the communication synchronization and data dependencies (although a complete discussion is beyond the scope of this thesis, see [Lu04, VP05]).

2.5.5 Run-time Interactive Programming Environment

The RIPE is a visual programming aid to the run-time environment (GEE) enabling the visualization of a dataflow diagram corresponding to the Lucid part of the GIPSY program, source code editing, launching the compilation and execution of GIPSY programs. The original conceptual design of RIPE [Paq99] is illustrated in Figure 13. The user's points of interaction with the RIPE at run-time vary in the following ways:

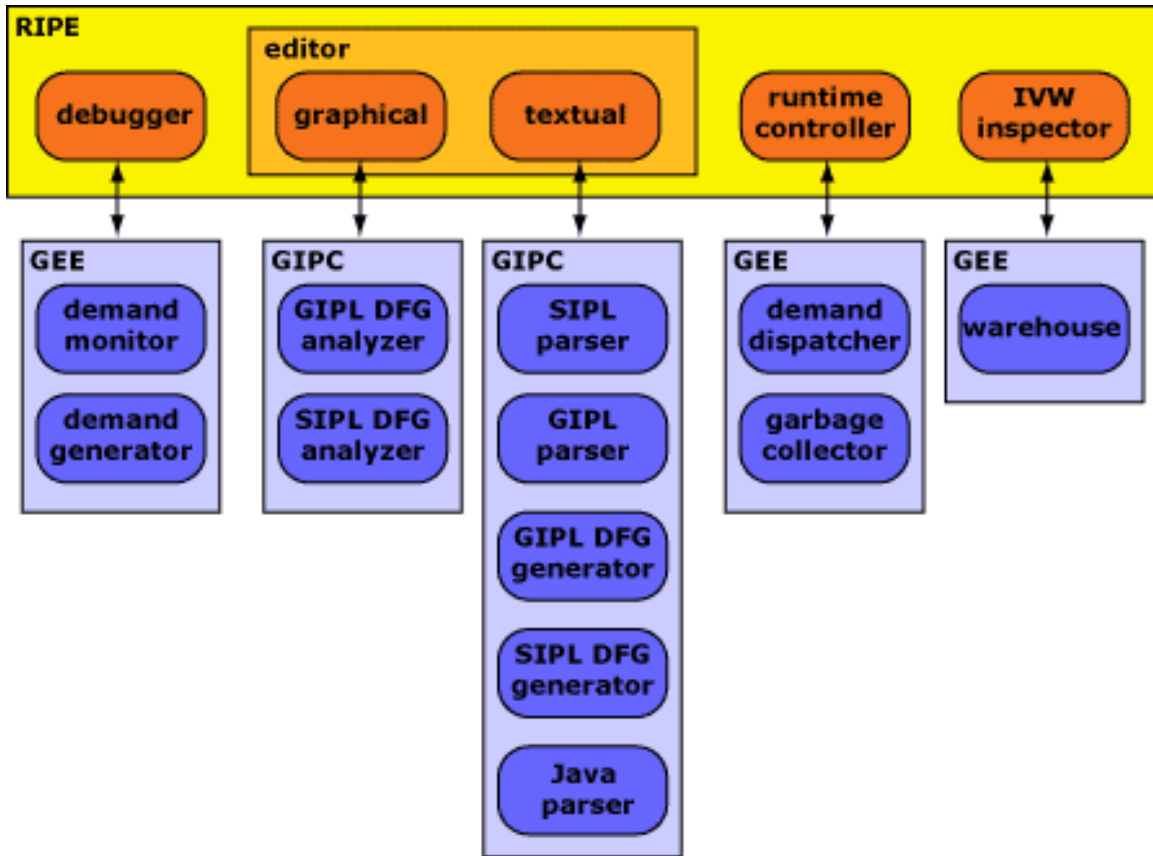


Figure 13: Conceptual Design of the RIPE

- Enable interactive editing of GIPSY programs via a variety of editors (textual, graphical, web).
- Dynamic inspection of the IVW.
- Modification of the input/output channels of the program.
- Recompilation of the GIPSY programs.
- Modification of the communication protocols.
- Swapping of the parts of the GIPSY itself (e.g. garbage collection, optimization, warehouse caching etc. strategies).

Because of the interactive nature of the RIPE, the GIPC is modularly designed to allow the individual on-the-fly compilation of either the IDS (by changing the Lucid code), CP (by changing the communication protocol), or ST (by changing the

sequential code). Such a modular design even allows sequential threads to be programs written in different languages (for now, we are concentrating on Java sequential threads, but a provision is made for easy inclusion of other languages with the GICF, Section 4.1.1.1).

The RIPE even enables the graphic development of Lucid programs, translating the graphic version of the program into a textual version that can then be compiled into an operational version through a DFG generator of Yimin Ding [Din04]. However, the development of this facility for graphical programming posed many problems whose solution is not yet settled, for example representation of the STs and CPs in the DFG nodes. An extensive and general requirements analysis will be undertaken, as this interface will have to be suited to many different types of applications. There is also the possibility to have a kernel run-time interface on top of which we can plug-in different types of interfaces adapted to different applications, such as stand-alone, web-, or server-based.

2.6 Tools

This section presents a brief description of a variety of tools that helped most with the implementation aspects of this work.

2.6.1 Java as a Programming Language

The primary implementation language of GIPSY is Java. This includes using Java's Reflection, JNI, and JUnit frameworks and packages. We have chosen to implement our project using the Java programming language mainly because of the binary portability of the Java applications as well as its facilities, for e.g. memory management and communication tasks, so we can concentrate more on the algorithms instead. Java also provides built-in types and data-structures to manage collections (build, sort, store/retrieve) efficiently [Fla97, Mic05b]. There is also source code written in other languages in the main GIPSY repository. This includes LEFTY code for DFG generation and the code of the test intensional programs in various Lucid dialects. The Java versions supported by GIPSY are 1.4 and 1.5. The GIPSY will no longer build on 1.3 and earlier JDKs.

2.6.1.1 Java Reflection

Java Reflection Framework `java.reflect.*` [Gre05] allows us to load/query/discover a given class for all of its API through enumeration of constructors, fields, methods, etc. at run-time. This is incredibly useful for dynamic loading and execution of our compilers, identifier context classes, and sequential threads on local and remote machines.

The basic API from the reflection framework used in the implementation of GIPSY is the `Class` class that allows getting arrays of declared `Method` objects through the `getDeclaredMethods()` call that will become the STs at the end, then for each `Method` the reflection API allows getting parameter and return types via `getParameterTypes()` and `getReturnType()` calls, which will become the CPs. The `Class.newInstance()` method allows instantiating an object off the newly generated class. Likewise, an enumeration of `Constructor` objects is acquired through the `Class.getConstructors()` call. Constructors in Java are treated differently from methods because they are not inherited and don't have a return type (except that

the type of the object they create). We still need to enumerate them to allow Objective Lucid programs to use the constructors, default or non-default, directly, so we can get a handle on them similarly to STs.

2.6.1.2 Java Native Interface (JNI)

The Java Native Interface (JNI) [Ste05] is very useful for the thread generation component of the GIPC. We rely on JNI to increase the number of popular imperative languages in which the sequential threads could be written. Developers use the JNI to handle some specific situations when an application cannot be written entirely in Java, e.g. when the standard Java classes do not provide some platform-dependent features an application may require, or use a library written in another language be accessible to Java applications, or for performance reasons a small portion of a time-critical code has to be written say in C or assembly, but still be accessible from a Java application [Ste05]. In GIPSY, the second and third of the listed cases are most applicable (e.g. to adopt GLU programs). The JNI will allow us to avoid Lucid-to-C or Lucid-to-C++ type matching as we can do it all through Java and maintain only Lucid-to-Java type mapping table.

The JNI is made so that the native and Java sides of an application can pass back and forth objects, strings, arrays and update their state on either end [Ste05]. The JNI is bi-directional, i.e., allows Java to use the native libraries and applications and provide access to Java libraries from the native applications.

The general methodology of creating a JNI application say that interacts with a C implementation is done in six steps [Ste05]:

1. Write a Java code with a `native` method to be implemented in C, the `main()`, and the dynamic loading statement for a library (to be compiled in the next steps).
2. Compile the Java code with `javac` and produce a `.class` file.
3. Create a C header `.h` file from the compiled `.class` file by calling `javah`. This header file will provide the necessary `#include` directives along with the C-style prototype declaration of the `native` method.
4. Next, write the implementation of the function in regular C in a `.c` file.

5. Then, create a shared library by compiling the `.h` and `.c` files with a C compiler.
6. Run the application regularly with the JVM (`java`).

2.6.1.3 JUnit

JUnit is an open-source Java testing framework used to write and run automated repeatable unit tests in a hassle-free manner [GB04]. The goal is to sustain application correctness over time, especially when undergoing a lot of integration efforts. JUnit is designed with software architecture patterns in mind and follows best software engineering practices. It encourages developers to write tests for their applications that withstand time and bit rot.

The main abstract class is `TestCase` that follows the Command design pattern that implements the `Test` interface. This class maintains the name of the tests (if it fails) and defines the `run()` method that has to be overridden to do the actual testing work. The default Template Method `run()` simply does three things: `setUp()`, `runTest()`, and `tearDown()`. Their default implementation is to do nothing, so a developer can override them as necessary. Then, to collect the test results they apply Collecting Parameter pattern. They use the `TestResult` class for that.

JUnit makes a distinction between *errors* and *failures* in the following way: errors to JUnit are mostly unexpected run-time or regular exceptions, whereas failures are anticipated and are tested for using assertion checks. The errors and failures are collected for further test failure reporting.

To run tests in a general manner from the point of view of the tester, the test classes have with a generic interface using the Adapter pattern. JUnit also offers a pluggable selector capability via the Java Reflection API [Gre05]. The `TestSuite` class represents a collection of tests to run. In the GIPSY, the `Regression` application (see Section 5.1) comprises concrete implementation of such a test suite that tests most of the feasible functionality of the GIPC and GEE modules. See more details of application of JUnit to the GIPSY in Chapter 5.

2.6.2 javacc – Java Compiler Compiler

JavaCC [VC05], accompanied by JJTree, is the tool the GIPSY project is relying on since the first implementation [Ren02] to create Java-language parsers and ASTs off

a source grammar files. The Java Compiler Compiler tool implements the same idea for Java, as do `lex/yacc` [Lou97] (or `flex/bison`) for C – reading a source grammar they produce a parser that complies with this grammar and gives you a handle on the root of the abstract syntax tree. The GIPL, Indexical Lucid, JLucid, Objective Lucid, `PreprocessorParser`, and `DFGGenerator` parsers are generated with the `JavaCC/JJTree` parser generation tools. `JavaCC` is a LL(K) [Lou97] parser generator, so the original GIPL and Indexical Lucid grammars and the new grammars had to be modified to eliminate or avoid the left recursion.

2.6.3 MARF

Modular Audio Recognition Framework (MARF) library [MCSN05] provides a few useful utility and storage classes GIPSY is using to manipulate threads, arrays, option processing, and byte operations. Despite MARF's belonging to a voice/speech/natural language recognition and processing library, it contains a variety of useful utility modules for threading and options processing.

2.6.4 CVS

For managing the source code repository the Concurrent Versions System (CVS) [BddzzP⁺05] is used. The CVS allows multiple developers work on the *up-to-date* source tree in parallel that keeps tracks of the revision history and works in an transactional manner. The author produced a mini-tutorial on the CVS [Mok03a] for the GIPSY Research and Development team, which contains the necessary summary for the team to work with the project repository.

While CVS has a comprehensive set of commands, the basic set includes:

- `init` to initialize the repository
- `checkout` or `co` to checkout the source code tree from the repository to a local directory
- `update` or `up` to make the local tree up-to-date with the one on the server
- `add` to schedule a new file inside the existing local checkout for addition to the repository

- `remove` to schedule a new file inside the existing local checkout for removal from the repository
- `commit` to upload the changes done locally to the server
- `diff` to show the differences between the local and the server versions of the tree

2.6.5 Tomcat

Apache Jakarta Tomcat [Fou05] is an open-source Java application servlet and server pages container project from Apache Foundation to run web Java-based applications written in accordance with the Java Servlet and JavaServer Pages [Mic05a, Mic05c] specifications developed by Sun Microsystems. Tomcat powers up the web front end to GIPSY to test intensional programs online. The web frontend is represented by the `WebEditor` servlet as of this writing a part of RIPE which is discussed later in Chapter 4. Tomcat has an easy interface to deploy Java-based applications and their libraries, e.g. through a manager presented in Figure 14.



Tomcat itself consists from a variety of modules that includes implementation of the JSP (Jasper engine) and Servlet APIs, a webserver called Coyote, the application server called Catalina, and many other things for logging, security, administration, etc.

2.6.6 Build System

The GIPSY's sources can be built using a variety of ways, using different compilers and IDEs on different platforms. This includes Linux Makefiles, IBM's Eclipse, Borland's JBuilder, Apache's Ant, and Sun's NetBeans.

2.6.6.1 Makefiles

UNIX/LINUX Makefiles are targeting all UNIX systems that support GNU `make` (a.k.a `gmake`) [SMSP00, Mok05a]. Often, to compile all of the GIPSY is just enough to type in `make` and the system will be built. All UNIX versions support `make`, and our system has been tested to build on RED HAT LINUX 9, FEDORA CORE 2, MAC OS X, and


The Apache Jakarta Project
<http://jakarta.apache.org/>


Tomcat Web Application Manager

Message: OK - Reloaded application at context path /webeditor

Manager

[List Applications](#) [HTML Manager Help](#) [Manager Help](#) [Server Status](#)

Applications

Path	Display Name	Running	Sessions	Commands
/	Welcome to Tomcat	true	0	Start Stop Reload Undeploy
/balancer	Tomcat Simple Load Balancer Example App	true	0	Start Stop Reload Undeploy
/jsp-examples	JSP 2.0 Examples	true	0	Start Stop Reload Undeploy
/manager	Tomcat Manager Application	true	0	Start Stop Reload Undeploy
/servlets-examples	Servlet 2.4 Examples	true	0	Start Stop Reload Undeploy
/tomcat-docs	Tomcat Documentation	true	0	Start Stop Reload Undeploy
/webdav	Webdav Content Management	true	0	Start Stop Reload Undeploy
<u>/webeditor</u>	GIPSY Web Portal	true	0	Start Stop Reload Undeploy

Deploy

Deploy directory or WAR file located on server

Context Path (optional):

XML Configuration file URL:

WAR or Directory URL:

WAR file to deploy

Select WAR file to upload

Server Information

Tomcat Version	JVM Version	JVM Vendor	OS Name	OS Version	OS Architecture
Apache Tomcat/5.5.9	1.5.0_01-b08	Sun Microsystems Inc.	Windows 2000	5.0	x86

Figure 14: Tomcat Web Applications Manager

SOLARIS 9. There is a test script `make-test.sh` that tests whether we are dealing with the GNU `make` on UNIX systems, as this is the only `make` supported.

2.6.6.2 Eclipse

There are project files `.project` and `.classpath` that belong to this IDE from IBM [c+04]. The GIPSY build with this IDE properly and has its library CLASSPATH set. Eclipse is another open source tool available free of charge and provides extended tools for Java projects development, refactoring, and deployment.

2.6.6.3 JBuilder

There is a project file `GIPSY.jpx` that belongs to this IDE from Borland [Bor03]. The GIPSY build with this IDE properly and has its library CLASSPATH set.

2.6.6.4 Ant

There is a project file `build.xml` that belongs to this build tool from the Apache Foundation [Con05]. The GIPSY build with this tool properly and has its library CLASSPATH set. In this case `build.xml` is a portable way to write a Makefile in XML.

2.6.6.5 NetBeans

There is a project file `nbproject.xml` that belongs to this IDE from Sun [Mic04]. The GIPSY build with this IDE properly and has its library CLASSPATH set.

2.6.7 readmedir

This script generates a human-readable description of a directory structure starting from some directory with file listing and possibly descriptions (for this there should be specially formatted file `README.dir` in every directory traversed. The contents of this file will be a part of the output and is a responsibility of the directory creator/maintainer. The output formats of the script are \LaTeX , HTML, and plain text.

2.7 Summary

In this chapter the reader was introduced to the necessary background on the GIPSY project and how it is being managed starting from the Lucid language origins to its implementations in the GIPSY and the summary of the tools used to aid the advancement of the project. In the GIPSY section the three main modules were introduced, such as GIPC, GEE, and RIPE. While most of the remaining work has gone into the GIPC in this thesis, the author had to perform the necessary integration and adjustments to the GEE and RIPE.

Chapter 3

Methodology

This chapter focuses on the methods and techniques proposed to solve the stated problems (see Section 1.1). The approaches described are based on three publications, namely [MPG05, MP05b, MP05a]. Section 3.1 introduces the JLucid language and all related considerations including the syntax and semantics. Next, Objective Lucid is introduced along with its syntax and semantics. Further, the GICF is introduced by providing the necessary requirements for it to exist and the way to satisfy them. Lastly, the summary is presented outlining the benefits and limitations of the proposed solutions.

3.1 JLucid: Lucid with Embedded Java Methods

3.1.1 Rationale

The name JLucid comes from the GIPC component known as *Java Compiler* within the Sequential Thread (ST) Generator of the GIPSY. It subsumes all of Indexical Lucid and General Intensional Programming Language (GIPL) [Paq99] and syntactically allows embedded Java code. In fact, a JLucid program looks like a partial fusion of the intensional and Java code segments. JLucid gives a great deal of flexibility to Lucid programs by allowing to use existing implementations of certain functions in Java, providing I/O facilities and math routines (that Lucid entirely lacks), and other Java features accessible to Lucid, arrays, and permits to increase the granularity of computations at the operator level by allowing the user to define Java operators, i.e.,

functions manipulating objects, thus allowing *streams of objects*¹ in Lucid. JLucid more or less achieves the same goals and mechanisms as provided by GLU. What we are proposing is a flexible compiler and run-time system that permits the evolution of languages through a framework approach [MP05a, PW05].

3.1.1.1 Modeling Non-Determinism

Lucid, by its nature, is deterministic, so introduction of imperative languages, such as Java, may allow us to model non-determinism in Lucid programs for example by providing access to random number generators available to the imperative languages. Non-determinism can also be introduced as a result of side effects from for example reading a different file each time an ST is invoked, or making a database query against a table where data regularly changes, or say by reading the current time of day value. Of course, a special care should be taken *not* to cache the results of such STs in the warehouse.

3.1.1.2 Loading Existing Java Code with `embed()`

In a nutshell, we want to make the following possible for the Indexical Lucid program in Figure 15 (replicated here from Chapter 2 for convenience) to become something as in Figure 16 or, alternatively as in Figure 17. The latter form would allow us to include objects from any types of URLs, local, HTTP, FTP, etc. The idea behind `embed()` is to include or to import the code written already by someone and not to rewrite it in Lucid (which may not be a trivial task). It is not meant to adjust to URL's existence at run-time as all `embed`-referenced resources are resolved at compile time. We “include” the pointed-to resource and attempt to compile it where the original program-initiator resides. If the URL is invalid at compile time, then there will be a compile error and no computation will be started. `embed()` by itself does not necessarily provoke a remote function call.

Existing Java code, in either `.class` or `.java` form, can be loaded with `embed()`. Intuitively, we would prefer the approach presented in Figure 18. That added flexibility requires syntactical extension of Lucid and is not portable. For the program in

¹A more precise meaning of Java objects within Lucid is explored further in the Objective Lucid language, including the meaning of an object stream and how object members are manipulated (see for example Section 3.2 and Section 4.1.3.6). Additionally, since the actual Java objects are flattened into primitive types, it would be possible to access object members in parallel manner.

```

H
where
  H = 1 fby merge(merge(2 * H, 3 * H), 5 * H);

  merge(x, y) = if(xx <= yy) then xx else yy
  where
    xx = x upon(xx <= yy);
    yy = y upon(yy <= xx);
  end;
end;

```

Figure 15: Indexical Lucid program implementing the `merge()` function.

```

#JAVA

void merge(int x, int y)
{
  // java code here
}

#JLUCID

H
where
  H = 1 fby merge(merge(2 * H, 3 * H), 5 * H);
end;

```

Figure 16: Indexical Lucid program implementing the `merge()` function as inline Java method.

```

H
where
  H = 1 fby merge(merge(2 * H, 3 * H), 5 * H);
  merge(x, y) =
    embed("file://path/to/class/Merge.class", "merge", x, y);
end;

```

Figure 17: Indexical Lucid program implementing the `merge()` function as `embed()`.

```

F
where
  dimension d;
  F = foo(#d);
  where
    foo(i) = embed("file://my/classes/Foo.class", "foo", i);
  end;
end;

```

Figure 18: Illustration of the `embed()` syntax.

Figure 18 to work, `foo()` has to return a Java type of `int`, `byte`, `long`, `char`, `String`, or `boolean`, as per Table 1, page 64. A wrapper class will be created to extend from the `Foo` and implement the `ISequentialThread` interface (see Appendix B.1). General `embed()` syntax would be defined as follows:

```
id(id, id, ...) ::= embed(URI, METHOD, id, id, ...);
```

where `id` is the Lucid function name being defined that is mapped to a Java's method named `METHOD` (which may or may not be of the same name as the first `id`). The `URI` is pointing to either `.class` or `.java` file. Example `URI`'s would be:

```
foo(a,b) = embed("file://files/Foo.java","bar",a,b);
bar(a,b) = embed("http://www.java.com/Foo.class","foo",a,b);
baz(a,b) = embed("ftp://ftp.file.com/pub/Foo.java","zee",a,b);
```

These declarations associate Lucid functions with Java implementations. Name clashes may be avoided, if necessary, by using different function names. Above, for example, Lucid `baz()` is implemented by Java `zee()`.

```
public class <filename>_<machine_name>_<timestamp>
extends my.classes.Foo
implements ISequentialThread
{
    // The definition is provided later in the text
}
```

Figure 19: Generated corresponding ST to that of Figure 18.

There are several ways of making this work. We could extract either a textual or a bytecode definition of `foo()`, wrap it in our own class and, (re)compile it. However, there is an issue here. What about other functions it may use, like shown in Figure 23 with two methods calling each other? That would mean extracting those dependencies as well along with the method of interest. This won't scale very efficiently. Thus, alternate approaches include: to either inherit from the desired class as in Figure 19, encapsulate this class instance, or attempt to wrap the entire class as done for the `JAVA` segment in Section 3.1.1.3 below. The former approach would imply having a class variable instance of the type of that class encapsulated into the wrapper. The latter approach was chosen as more feasible to implement, although it doesn't deal with user-defined classes and subclass and packages the `.class` or `.java` file may require at the moment. Thus, the `embed()` acts in a way similar to `#include` in C/C++ or `import` in Java of a set of Java definitions to be used in a `JLucid` program.

Therefore, `embed()` has to be resolved at compile time. Similar technique may be taken towards other languages than Java at a later time. Lucid's syntax has to be extended to support `embed()`.

3.1.1.3 The #JAVA and #JLUCID Code Segments

This section explores ways of mixing Java and Lucid source code segments in a single text file and ways of dealing with such a merge.

```
F
where
  dimension d;
  F = foo(#d);
  where
    foo(i) = int foo(int i) { return i + 1; }
  end;
end;
```

Figure 20: Inline Java function declaration.

An attempt to use Java's methods inline, such as in Figure 20 would be intuitive, but does not justify the effort spent on syntax analysis. Therefore, we take the inline definition out of the Lucid part, and make it a separate outer definition of the same method. Additionally, we explicitly mark the JLUCID and JAVA code segments to simplify pre-processing of the JLucid code as presented in Figure 21.

```
#JAVA

int foo(int i)
{
  // Some i + PI
  return (int)(java.lang.Math.PI + i);
}

#JLUCID

F
where
  dimension d;
  F = foo(#d);
end;
```

Figure 21: Java method declaration split out from the Lucid part.

Given the Natural Numbers Problem (see [Paq99]) in Figure 22 (replicated here for convenience), one could imagine the function definition for N to be implemented in Java in two functions. To illustrate the point when two separate functions can call

each other in the `JAVA` segment or several `JAVA` segments. This modified `JLucid` code along with line numbers is shown in Figure 23. Since we allow one Java method to call another within, we have to wrap them both into the same class.

```

N @.d 2
where
  dimension d;
  N = if (#.d <= 0) then 42 else (N + 1) @.d (#.d - 1) fi;
end;

```

Figure 22: Natural numbers problem in plain `GIPL`.

The `JLucid` code segments after “`#JAVA`” constructs will be grouped together by the compiler. For all definitions (functions, classes, variables) in these segments, their original location in the `JLucid` source recorded and statically put in the wrapper class. These definitions will end up in that wrapper class as well.

It would be possible to have a class defined within a wrapper class or any other valid Java declaration; even a data member can be included. To summarize, the Java segments in the `JLucid` code are a body of a generated class that implements the `ISequentialThread` interface.

```

1  #JAVA
2
3  int getN(int piDimension)
4  {
5      if(piDimension <= 0)
6          return get42();
7      else
8          return getN(piDimension - 1) + 1;
9  }
10
11 int get42()
12 {
13     return 42;
14 }
15
16 #JLUCID
17
18 N @d 2
19 where
20     dimension d;
21     N = getN(#d);
22 end;

```

Figure 23: Natural numbers problem with two Java methods calling each other.

For the example in Figure 23 the parser would proceed as follows:

- In the preprocessing step the source code is split into two parts: the Java part and the Lucid part. For both parts original source’s line numbers and length of the definitions are recorded.
- Then they both are fed to the respective parsers. Java’s part requires extra handling: the Java methods (one or more) defined in the code, have to be wrapped into a class and then `JavaCompiler` class that takes the Java portion of the source and feeds it to `javac` for syntactic and semantic analyses and byte code generation. They will become parts of a Sequential Thread, ST (see Section 3.3.3.1) definition fed to Workers (see Section 3.3.3.4).
- The Lucid part is processed by the modified Lucid compiler (to include the syntactical modifications for arrays and `embed()`) and comes up with the main AST from that.
- The Java STs are then linked into the main AST in place of nodes where the identifiers of these appear in the Lucid part of the program prior semantic analysis.

Any method or other definition in the `JAVA` segment is wrapped into a class. The generated wrapper class will contain a `Hashtable` that maps method signature strings to their starting line in the original JLucid code plus the length of the definitions in lines of text they occupy statically generated and initialized. This is needed for the error reporting subsystem in case of syntax/semantic errors, report back correctly the line in the original JLucid program and not in the generated class. The class name is created automatically from the original program name, the machine name it’s being compiled on, and a timestamp to guarantee enough uniqueness to the generated class’ name to minimize conflict for multiple such generated classes. Thus, the `JAVA` segment in Figure 23 will transform into the generated class as in Figure 24. This is a short version; for more detailed one please refer to the Section B.3. In fact, after generating this class (and possibly compiling it) this situation can be viewed as a special case for `embed()`, Section 3.1.1.2 or vice versa. Note, since we have no guarantee the Java methods are side-effects free in JLucid, their results are not cached in the warehouse.

In [MPG05] we required `foo()` in the previous examples to be `static`. In fact, any method or other definition in the `JAVA` segment were to be transformed to become `static` while being wrapped into a class. For example, “`int foo() {return 1;}`”

```

public class <filename>_<machine_name>_<timestamp>
implements gipsy.interfaces.ISequentialThread {
    private OriginalSourceCodeInfo oOriginalSourceCodeInfo;

    // Inner class with original source code information
    public class OriginalSourceCodeInfo {
        // For debugging / monitoring; generated statically
        private String strOriginalSource = ...

        // Mapping to original source code position for error reporting
        private Hashtable oLineNumbers = new Hashtable();

        // Body is filled in by the preprocessor statically
        public OriginalSourceCodeInfo() {
            Vector int_getN_int_piDimension = new Vector();
            // Start line and Length in lines
            int_getN_int_piDimension.add(new Integer(3));
            int_getN_int_piDimension.add(new Integer(7));
            oLineNumbers.put("int getN(int piDimension)",
                int_getN_int_piDimension);
            Vector int_get42 = new Vector();
            int_get42.add(new Integer(11));
            int_get42.add(new Integer(4));
            oLineNumbers.put("int get42()", int_get42);
        }
    }

    // Constructor
    public <filename>_<machine_name>_<timestamp>() {
        oOriginalSourceCodeInfo = new OriginalSourceCodeInfo();
    }
    /*
    * Implementation of the SequentialThread interface
    */
    // Body generated by the compiler
    public void run() {
        Payload oPayload = new Payload();
        oPayload.add("d", new Integer(42));

        work(oPayload);
    }

    // Body generated by the compiler statically
    public WorkResult work(Payload poPayload) {
        WorkResult oWorkresult = new WorkResult();
        oWorkresult.add(getN(poPayload.getVaueOf("d")));
        return oWorkResult;
    }

    /*
    * The below are generated off the source file nat2java.ipl
    */
    public static int getN(int piDimension) {
        if(piDimension <= 0) return get42();
        else return getN(piDimension - 1) + 1;
    }

    public static int get42() {
        return 42;
    }
}

```

Figure 24: Generated Sequential Thread Class.

would become “`public static int foo() {...}`”. We insisted on static declarations only because the sequential threads were not instantiated by the workers when executed. This restriction has been lifted during implementation as we instantiate and serialize the sequential thread class as needed.

3.1.1.4 Is JLucid an Intensional Language?

We treat JLucid as a separate specific intensional programming language (SIPL) rather than a part of a GIPSY program within existing Indexical Lucid implementation. Here are some pros and cons of this approach and JLucid as a separate SIPL approach is the winner. Why extend it as a separate SIPL?

- This would serve as an example on how to add other SIPLs.
- This would allow us to keep the original Indexical Lucid clean and working.
- This would allow functions with Java syntax to be used within a Lucid program as well as binary Java function calls of pre-compiled classes.
- It can be extended to other languages as it turns out to be a successful approach.

Why *not* to treat it as a separate SIPL?

- We might want to have embedded Java (or other language) in any intensional language, not just Indexical Lucid. How to make that possible?
- It is not truly an SIPL, but a hybrid.

3.1.2 Syntax

In JLucid, we extend the syntax of both GIPL and Indexical Lucid to support arrays. For example, it is useful to be able to evaluate several array elements under the same context. This is included by the last E rules of $E[E, \dots, E]$ and $[E, \dots, E]$ in both syntaxes. Arrays are useful to manipulate a collection Lucid streams under the same context. JLucid arrays are mapped to Java arrays on the element-by-element basis with the appropriate element type matching and may only correspond to arrays of primitive types in Java. The syntax also includes the `embed()` extension to allow including external Java code. The JLucid syntax extensions to GIPL and Indexical Lucid are presented in Figure 25 and Figure 26.

```

E ::= id
   | E(E,...,E)
   | if E then E else E fi
   | # E
   | E @ E
   | E where Q end;
   | [E:E,...,E:E]
   | embed(URI, METHOD, E, E, ...)
   | E[E,...,E]
   | [E,...,E]
Q ::= dimension id,...,id;
   | id = E;
   | id(id,...,id) = E;
   | QQ

```

Figure 25: JLucid Extension to GIPL Syntax

```

E ::= id
   | E(E,...,E)
   | if E then E else E fi
   | # E
   | E @ E E
   | E where Q end;
   | E bin-op E
   | un-op E
   | embed(URI, METHOD, E, E, ...)
   | E[E,...,E]
   | [E,...,E]
Q ::= dimension id,...,id;
   | id = E;
   | id.id,...,id(id,...,id) = E;
   | QQ
bin-op ::= fby | upon | asa | wvr
un-op  ::= first | next | prev

```

Figure 26: JLucid Extension to Indexical Lucid Syntax

3.1.3 Semantics

The JLucid extension to the operational semantics of Lucid (see Section 2.2.2.8 on page 15) is defined in Figure 27. As in the original Lucid semantics, each type of identifier can only be used in the appropriate situations. Notation:

- **freefun**, **ffid**, **ffdef** mean a type of identifier is a hybrid free (i.e. object-free) function **freefun**, where **ffid** is its identifier and **ffdef** is its definition (body).
- The **E_{ffid}** rule defines JLucid's free functions.
- The JLucid **#JAVA_{ffid}** rule add free function definition to the definition environment.

$$\begin{array}{l}
\mathbf{E}_{\text{ffid}} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}, \mathcal{P} \vdash E_1, \dots, E_n : v_1, \dots, v_n \quad \mathcal{D}(id) = (\text{freefun}, \text{ffid}, \underline{\text{ffdef}}) \quad \mathcal{D}, \mathcal{P} \vdash \langle \text{ffid}(v_1, \dots, v_n) \rangle : v}{\mathcal{D}, \mathcal{P} \vdash E(E_1, \dots, E_n) : v} \\
\#\text{JAVA}_{\text{ffid}} : \frac{\underline{\text{ffdef}} = \text{frttype} \text{ffid}(\text{fargtype}_1 \text{fargid}_1, \dots, \text{fargtype}_n \text{fargid}_n)}{\mathcal{D}, \mathcal{P} \vdash \underline{\text{ffdef}} : \mathcal{D}^\dagger[\text{ffid} \mapsto (\text{freefun}, \text{ffid}, \underline{\text{ffdef}})], \mathcal{P}}
\end{array}$$

Figure 27: Additional basic semantic rules to support JLucid

3.2 Objective Lucid: JLucid with Java Objects

3.2.1 Rationale

Objective Lucid is a direct extension of JLucid. The original syntax of Indexical Lucid (and also for JLucid and GIPL) is augmented to support a so-called *dot-notation*. This allows Lucid to manipulate grouped data by using object’s methods. In fact, the idea is similar to manipulating arrays in JLucid. The difference with the arrays is that they are manipulated as a collection of ordered data of elements of the same type, to be evaluated in the same context. However, an object that varies in some dimension implies that all its members, possibly of different types, also potentially vary along this dimension, but across objects, i.e. the objects themselves are not intensional. An object can be thought of as a heterogeneous collection of different types of members, which you can access individually using their name, whereas arrays can be thought of as a homogeneous collection of members that can be accesses individually using their index.

Just like JLucid [MPG05], Objective Lucid is being developed as a separate specific intensional programming language (SIPL) within the GIPSY for the same reasons: keeping the other implementations undisturbed and working while experimenting on this particular implementation.

3.2.1.1 Pseudo-Objectivism in JLucid

A pseudo-object-oriented approach is already present in JLucid. The program presented in Figure 28 gives an example of a Java function returning an object of type `Integer`. In JLucid we are not able to manipulate this object directly in intensional programming as Java does, though we can provide methods, such as `g()` to access

properties of a particular Java object from within JLucid. However, that reduces legacy Java code reusability by forcing the programmer to add such functions in his code to be able to use it in the GIPSY. Another example in Figure 29 shows how one can make use of objects in JLucid by providing pseudo-free Java accessors similar to `getComputedBar()` in the example. They are pseudo-free because they don't appear as a part of any Java class to a JLucid programmer explicitly, but internally they get wrapped into a class when the code is compiled. In Objective Lucid such explicit workarounds are not necessary anymore, but this gives us some ideas about how to actually implement some features of Objective Lucid in practice, i.e., the compiler can *generate* a number of pseudo-free accessors to object's members and use JLucid's implementation of Java functions internally.

3.2.1.2 Stream of Objects

An interesting question could be to ask: "What is an object stream?" Is it that the members of this object vary in the same dimension(s) or they can have "substreams"? In Objective Lucid we answer this as decomposing public object's data members into primitive types and varying them or in simplified manner we employ object's effectors. Thus, when there is a demand say for the object's state (data members) at some time t , there will have to be generated demands for all of t between $[0, t]$ where at time 0 an instance of the object is created. Therefore, the object state changes in the $[0, t]$ interval represent the object stream in the context of this thesis. There are two possible outcomes of this evaluation: either a portion of object's state is altered by an intensional program or the entire object. In the former case, Lucid only accesses *some* object's members via the dot-notation in the intensional manner, whereas in the latter case *all* the members of an object are altered in the intensional context implicitly. The examples presented in Figure 30, Figure 74, page 153, and Figure 76, page 156 work on portions of an object, whereas the examples in Section 4.1.3.6, page 109 work on all the members of an object at the same time.

```

#JAVA

Integer f()
{
    return new Integer("1234");
}

int g()
{
    return f().intValue();
}

#JLUCID

A
where
    A = g();
end;

```

Figure 28: Pseudo-objectivism in JLucid.

```

#JAVA

class Foo
{
    private int bar;

    public Foo()
    {
        bar = (int)(Math.random() * Integer.MAX_VALUE);
    }

    public int getBar()
    {
        return bar;
    }

    public void computeMod(int piParam)
    {
        bar = bar % piParam;
    }
}

int getComputedBar(int piParam)
{
    Foo oFoo = new Foo();
    oFoo.computeMod(piParam);
    System.out.println("bar = " + bar);
    return oFoo.getBar();
}

#JLUCID

Bar
where
    Bar = getComputedBar(5);
end;

```

Figure 29: Using pseudo-free Java functions to access object properties in JLucid.

3.2.1.3 Pure Intensional Object-Oriented Programming

Objective Lucid has presented a way for Lucid programs to use Java objects. This may seem rather restrictive and may look like a workaround (though practical!). An interesting concept would be to extend the Lucid language itself to create and manipulate pure Lucid objects, not Java objects. This will allow addressing issues like inheritance and polymorphism and other attributes of object-oriented programming and will solve the problem of matching Lucid and Java data types. This is not addressed in this work, but attempted to be solved in [WP05].

3.2.2 Syntax

The parser is extended to support the `<objectref>.<feature>` *dot-notation* for the Lucid part of reference data types. The semantic analysis is augmented to accommodate objects and user-defined data types. In doing so, Lucid is able to manipulate Java objects as well as access public variables and methods of these objects. An example is shown in Figure 30. This example manipulates a simple object `E` by evaluating its state at some time “2”. The program begins with the construction of the object with `f1()` (or one could call the object constructor directly), and then the rest of the expressions access public members `x` and `foo()` of the object during expression evaluation.

The Objective Lucid syntax is in Figure 31. It is a direct extension of the JLucid syntax in Figure 26 to support the dot-notation. Essentially, the extension is the `E.id` productions. Any `E` on the left-hand-side can evaluate to an object type, but the right-hand-side is always an identifier (Java class’ data member or method).

3.2.3 Semantics

To support these extensions to JLucid, the Semantic Analyzer of JLucid requires more non-trivial changes than the syntax analysis and the dot-notation implementation due to arbitrary object data types. In order to perform type checks and apply the semantic rules of Lucid, we place the object data types into the definition environment \mathcal{D} , which is in fact a semantic equivalent to the data dictionary part of the GEER. This is partly solved by using the pseudo-free Java functions, which *de-objectify* the object members, but in order to be able to do so, we need to have the object types

```

#JAVA
class ClassXB
{
    public int x;
    public float b;

    public ClassXB()
    {
        x = 0; b = 1.2;
    }

    public int foo(int a, float c)
    {
        return x = (int)(x * a + b * c);
    }

    ClassXB addx(int b)
    {
        x += b;
        return this;
    }
}

ClassXB f1()
{
    return new ClassXB();
}

#OBJECTIVELUCID

/*
 * The result of this program should be the object E
 * to be evaluated at time dimension 2 with its 'x'
 * member modified accordingly.
 */

E @time 2
where
    dimension time;
    E = f1() fby.time A;
    A = E.addx(B);
    B = E.foo(A @time C, A) + 3;
    C = E.x * 2;
end;

```

Figure 30: Objective Lucid example.


```

E      ::=  id
          |  E(E,...,E)
          |  if E then E else E fi
          |  # E
          |  E @ E E
          |  E where Q end;
          |  E bin-op E
          |  un-op E
          |  embed(URI, METHOD, E, E, ...)
          |  E[E,...,E]
          |  [E,...,E]
          |  E.id
          |  E.id(E,...,E)
Q      ::=  dimension id,...,id;
          |  id = E;
          |  E.id = E;
          |  id.id,...,id(id,...,id) = E;
          |  QQ
bin-op ::=  fby | upon | asa | wvr
un-op  ::=  first | next | prev

```

Figure 31: Objective Lucid Syntax

in the definition environment. The corresponding operational semantic rules from [Paq99] can be extended as follows.

The Objective Lucid extension to the operational semantics of Lucid is defined in Figure 32. As in the original Lucid semantics, each type of identifier can only be used in the appropriate situations. Notation:

- `class`, `cid`, `cdef` means it is a Class type of identifier with name `cid` and a definition `cdef`.
- `classv`, `cid.cvid`, `vdef` means that the variable is a member variable of a class `classv` with identifier `cid.cvid` given the variable definition `vdef` within the class.
- `<cid.cvid>` means object-member reference within an intensional program.
- `classf`, `cid.cfid`, `fdef` means that the function is a member function of a class `classf` with identifier `cid.cfid` given the variable definition `fdef` within the class.
- `<cid.cfid(v_1, \dots, v_n)>` represents a object-function call within an intensional program with actual parameters.
- `freefun`, `ffid`, `ffdef` mean a type of identifier is a hybrid free (i.e. object-free) function `freefun`, where `ffid` is its identifier and `ffdef` is its definition (body).

- By `cdef = Class cid {...}` we declare a class definition. A class can contain member variable `vdef` and member functions definitions `fdef`.

The rules:

- The \mathbf{E}_{c-vid} rule defines an object member variable for an expression for the dot-notation. It is independent from the language in which we define and express our objects. The rule says that under some context given two expressions E and E' that evaluate to a class-type identifier id and a variable type identifier id' respectively and if the two together via a dot-notation represent an object-data-member reference, then the expression $E.E'$ evaluates to a value v .
- Member function calls are resolved by the \mathbf{E}_{c-fct} rule. Similarly to the \mathbf{E}_{c-vid} rule, it defines that given two expressions E and E' under some context that evaluate to a class-type identifier id and a member function type identifier id' and a set of intensional expressions E_1, \dots, E_n evaluates to some values v_1, \dots, v_n and the two identifiers via a dot-notation represent a member function call with parameters v_1, \dots, v_n , then we say the expression $E.E'(E_1, \dots, E_n)$ is a member function call that under the same context evaluates to some value v , i.e. the function *always* returns a value. Here we see why it is necessary for Lucid to map a `void` data type to implicit Boolean `true`. This choice may seem a bit arbitrary (for example, one could pick an integer 1), but aside from practicality aspect the mere choice of `true` may signify a successful termination of a method.
- The \mathbf{E}_{ffd} rule defines JLucid's free functions. The rule is a simpler version of \mathbf{E}_{c-fct} with no class type identifiers present.
- The $\#\mathbf{JAVA}_{objid}$ rule places class definition into the definition environment.
- The $\#\mathbf{JAVA}_{objvid}$ and $\#\mathbf{JAVA}_{objfid}$ rules add `public` Java object member variable and function identifiers along with their definitions to the definition environment.
- The JLucid $\#\mathbf{JAVA}_{ffd}$ rule add free function definition to the definition environment.

$$\begin{array}{l}
\mathbf{E}_{c\text{-}vid} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}, \mathcal{P} \vdash E' : id' \quad \mathcal{D}(id) = (\text{class}, \text{cid}, \underline{\text{cdef}}) \quad \mathcal{D}(id') = (\text{classv}, \text{cid.cvid}, \underline{\text{vdef}}) \quad \mathcal{D}, \mathcal{P} \vdash \langle \text{cid.cvid} \rangle : v}{\mathcal{D}, \mathcal{P} \vdash E.E' : v} \\
\mathbf{E}_{c\text{-}fct} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}, \mathcal{P} \vdash E' : id' \quad \mathcal{D}, \mathcal{P} \vdash E_1, \dots, E_n : v_1, \dots, v_n \quad \mathcal{D}(id) = (\text{class}, \text{cid}, \underline{\text{cdef}}) \quad \mathcal{D}(id') = (\text{classf}, \text{cid.cfoid}, \underline{\text{fdef}}) \quad \mathcal{D}, \mathcal{P} \vdash \langle \text{cid.cfoid}(v_1, \dots, v_n) \rangle : v}{\mathcal{D}, \mathcal{P} \vdash E.E'(E_1, \dots, E_n) : v} \\
\mathbf{E}_{foid} : \frac{\mathcal{D}, \mathcal{P} \vdash E : id \quad \mathcal{D}, \mathcal{P} \vdash E_1, \dots, E_n : v_1, \dots, v_n \quad \mathcal{D}(id) = (\text{freefun}, \underline{\text{ffid}}, \underline{\text{ffdef}}) \quad \mathcal{D}, \mathcal{P} \vdash \langle \text{ffid}(v_1, \dots, v_n) \rangle : v}{\mathcal{D}, \mathcal{P} \vdash E(E_1, \dots, E_n) : v} \\
\# \mathbf{JAVA}_{objid} : \frac{\underline{\text{cdef}} = \text{Class cid} \{ \dots \}}{\mathcal{D}, \mathcal{P} \vdash \underline{\text{cdef}} : \mathcal{D}^\dagger[\text{cid} \mapsto (\text{class}, \text{cid}, \underline{\text{cdef}})], \mathcal{P}} \\
\# \mathbf{JAVA}_{objvid} : \frac{\underline{\text{cdef}} = \text{Class cid} \{ \dots \underline{\text{vdef}} \dots \} \quad \underline{\text{vdef}} = \text{public type vid};}{\mathcal{D}, \mathcal{P} \vdash \underline{\text{cdef}} : \mathcal{D}^\dagger[\text{cid.vid} \mapsto (\text{classv}, \text{cid.vid}, \underline{\text{vdef}})], \mathcal{P}} \\
\# \mathbf{JAVA}_{objfid} : \frac{\underline{\text{cdef}} = \text{Class cid} \{ \dots \underline{\text{fdef}} \dots \} \quad \underline{\text{fdef}} = \text{public frttype fid}(fargtype_1 fargid_1, \dots, fargtype_n fargid_n)}{\mathcal{D}, \mathcal{P} \vdash \underline{\text{cdef}} : \mathcal{D}^\dagger[\text{cid.fid} \mapsto (\text{classf}, \text{cid.fid}, \underline{\text{fdef}})], \mathcal{P}} \\
\# \mathbf{JAVA}_{foid} : \frac{\underline{\text{ffdef}} = \text{frttype foid}(fargtype_1 fargid_1, \dots, fargtype_n fargid_n)}{\mathcal{D}, \mathcal{P} \vdash \underline{\text{ffdef}} : \mathcal{D}^\dagger[\text{ffid} \mapsto (\text{freefun}, \underline{\text{ffid}}, \underline{\text{ffdef}})], \mathcal{P}}
\end{array}$$

Figure 32: Additional basic semantic rules to support Objective Lucid

3.3 General Imperative Compiler Framework

3.3.1 Rationale

Having to deal with JLucid, Objective Lucid, and Java and a future likely possibility to include other than Java imperative languages into intensional ones prompted invention of a general mechanism to handle that and simplify addition of new languages into the GIPSY for research and experiments. This generalization touches several critical aspects exposed by the JLucid and Objective Lucid languages involving such a hybrid programming model. Thus, a core redesign of the GIPC was necessary to enable this feature. The General Imperative Compiler Framework (GICF) addresses the generalization issues (split among this Methodology and Design and Implementation chapters) for the imperative compilers and suggests later development of a similar framework for the intensional languages.

The core areas in the hybrid compilation process affect the way an intensional language program (which now syntactically allows having any number of code segments

written in one or more imperative languages) is compiled. This kind of program has to be preprocessed first to extract the code segments to be compiled by the appropriate language compilers and at the same time maintains syntactic and semantic links between the parts of a hybrid program. This influences the general intensional compiler instrumentation, such as generation of sequential threads and communication procedures, function elimination, GIPL-to-SIPL translation, semantic analysis, and linking (and later interpreting/executing) of a GIPSY program.

Requirements for any such a framework like GICF imply at least the following considerations:

- having a number of compiler interfaces known to the system that any concrete compiler implements,
- ability to pick such compilers at runtime based on a hybrid program being compiled,
- have a generalized AST that is capable of holding intensional and imperative nodes,
- have the semantic analyzer understand possible data types that any language may expose (which is a very challenging goal to do correctly), and deal with function elimination for the imperative parts of the AST,
- preprocess by breaking down a hybrid GIPSY program's source code to be fed to the appropriate compilers gives us flexibility of allowing to include any imperative language we want, but complicates maintenance of semantic links between the intensional and imperative parts for later linking and semantic analysis. This necessitates development of the two other special segments that can declare in a uniform manner for GIPSY providing some meta information about embedded imperative sequential threads, like function and type identifiers, parameter and return types for communication procedures, and user data types. Thus, for the former we need a function prototype declaration segment, that lists all free functions declared within imperative segments to be used by Lucid and the type declaration segment for the user-defined types possibly declared in those same imperative segments. The purpose of this meta-information is two-fold: it will help us maintaining the semantic links via a dictionary and

create so-called “imperative stubs”. The former prompts the development of the GIPSY Type System (see Section 4.1.1.5, page 83) as understood by the Lucid language and its incarnation within the GIPSY to handle types in a more general manner. The latter stubs have to be produced in order for the intensional language compilers (that stay intact with the introduced framework) not to choke on “undefined” symbols that really were defined in the imperative parts, which an existing intensional compiler running in isolation fails to see.

- After all involved compilers are finished doing compilation of their code segments, they all produce a partial AST. For intensional compilers that means the main AST with the intensional and stub nodes. For imperative compilers it is the appropriate imperative AST for each sequential thread. The imperative AST, in fact, need not to be a real tree and may contain a single imperative node that would hold a payload of STs (compiled object or byte code), CPs, type information, and some meta-information (e.g. what language the STs and CPs are in and for which operating system and native compiler environment).
- Then, the imperative stubs have to be replaced by the real imperative nodes at the linking stage before the semantic analysis.
- Once the main tree is formed, the semantic analyzer would use the type system to verify type information of the intensional-imperative calls within taking into consideration imperative nodes when doing function elimination and producing the final “executable” tree, or Demand AST, or DAST, a component of the GEER.

All this work is motivated by the desire to simplify the addition of new compilers into the GIPSY environment with minimal integration hassle. The follow up sections explore some of the issues about primary matching of the Java and GIPSY data types, followed by the definition of sequential threads and communication procedures in the GIPSY, and their **Worker** aggregator. While the below are sections that lay down a concrete example based on JLucid and Java, the discussion addressing the generalization of the design and implementation of these issues are presented in the chapter that follows with the actual sequence diagram showing implementation details of the above hybrid compilation process.

3.3.2 Matching Lucid and Java Data Types

Allowing Lucid to call Java functions brings a new set of issues related to data types. Additional work is required on the semantic analyzer, especially when it comes to type checks between Lucid and Java parts of a JLucid program. This is pertinent when Lucid variables or expressions are used as parameters to Java functions and when a Java function returns a result to be assigned to a Lucid variable or used in an IP expression. The sets of types in both cases are not exactly the same. The basic set of Lucid data types as defined by Grogono [Gro02b] is `int`, `bool`, `double`, `string`, and `dimension`. Lucid's `int` is of the same size as Java's `int`, and so are `double`, `boolean`, and `String`. Lucid `string` and Java `String` are simply mapped to each other since internally we implement the former as the latter; thus, one can think of the Lucid `string` as a reference when evaluated in the intensional program. Based on this fact, the lengths of a Lucid `string` and Java `String` are the same. Java `String` is also an object in Java; however, at this point, a Lucid program has no direct access to any object properties. We also distinguish the `float` data type for single-precision floating point operations. The `dimension` index type is said to be an integer for the time being, but might become a float when higher precision of points in time, for example, will be in demand, or it could even be an enumerated type of unordered values (though `float` dimensions will introduce some very interesting problems). Therefore, we perform data type matching as presented in Table 1. The return and parameter types matching sets are not the same because of the size of the types. Additionally, we allow `void` Java return type which will always be matched to a Boolean expression `true` in Lucid as an expression has to always evaluate to something.

The table does not reflect the fact that JLucid is able to manipulate arrays of values (streams), but these arrays are not Java arrays (Java's arrays are objects). In Objective Lucid (see Section 3.2), we also have Java object data types will also be manipulated by a Lucid program with the Lucid part being able to access object's properties and methods and have them as return types and arguments. As for now our types mapping and restrictions are as per Table 1.

Table 1: Matching data types between Lucid and Java.

Return Types of Java Methods	Types of Lucid Expressions
int, byte, long	int
float	float
double	double
boolean	bool
char, String	string
void	bool::true
Parameter Types Used in Lucid	Corresponding Java Types
string	String
float	float
double	double
int, dimension	int
bool	boolean

3.3.3 Sequential Thread and Communication Procedure Generation

3.3.3.1 Java Sequential Threads

Sequential threads are imperative functions that can be called in the Lucid part of a GIPSY program. The data elements of a Lucid program are integers and the like. Using them as such would result in a very inefficient computation due to the overhead in generation and propagation of demands. STs overcome this problem. The notion of sequential thread and granularization of data was introduced by the GLU (Granular LUCid system [JD96, JDA97]).

Each GIPSY program potentially defines several Java methods that can be called by the Lucid part of the program. Each of these functions are coded in the Java part of the GIPSY program; thus, a sequential thread represents by itself a bit of work to compute split into one or more Java methods. They are compiled (see Figure 33) to Java byte code by the compiler (GIPC, Figure 10) and packed into one executable, along with the Communication Procedures (CP) (see Section 3.3.3.2) needed for the communication between the generator and worker (Section 3.3.3.4, Figure 34). The notion of worker is thus very close to the notion of sequential threads, where a worker is basically the aggregation of the (potentially) several sequential threads that can be executed by a worker, along with the communications procedures needed for the

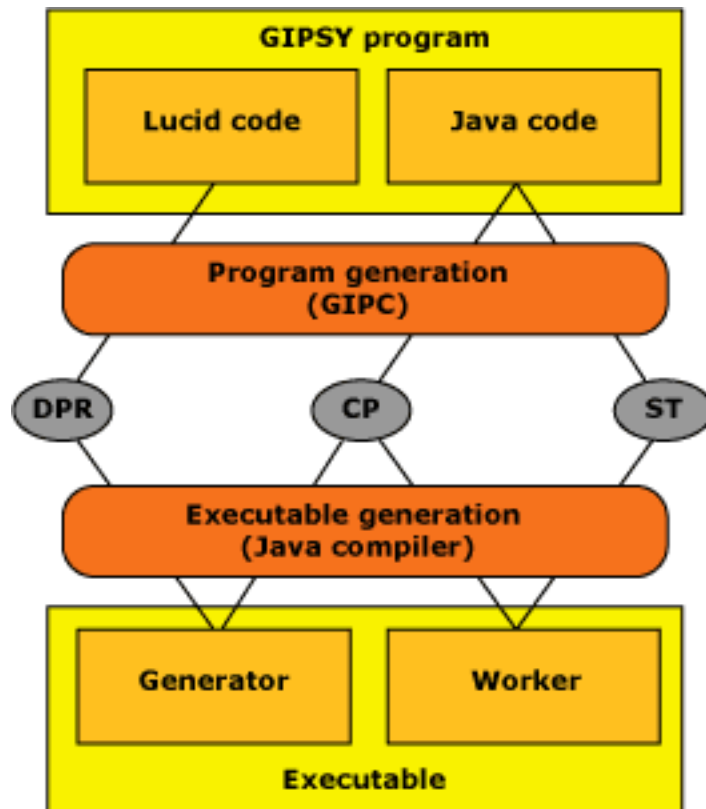


Figure 33: Hybrid GIPSY Program Compilation Process

generator-worker communication.

Notice that the Generator-Worker Architecture may well be extended so that the worker and the generator are fused into one; this is under review and is discussed in [Lu04] and in [VP05]. This gives us distributed generators as outlined in [Gro02b], but as yet is only a topic for discussion.

3.3.3.2 Java Communication Procedures

The functional demands (i.e., demands that raise the need for a Java function call) are potentially computed by remote workers, upon demand by the generator. The demand is sent via the network by the generator to the worker, along with the data representing the parameters of this Java function call. Sending this data through the network requires the breaking of the data structure into packets transmissible via a network. This packing of the demand's input data is done by the Communication Procedures, along with some kind of remote procedure call to the worker using, for example, TCP/IP RPC. Once the function (the sequential thread) resolves, the worker

(Section 3.3.3.4) is responsible for sending back the result to the generator that called for this demand. That is also done by the CPs.

The CPs are generated by the compiler (GIPC) using the first part of the GIPSY program: the definition of the data structures sent over the network (i.e., the parameter and return types of the Java functions). The GIPC parses these Java data structures and translates them into an abstract syntax tree. This tree is then traversed by the CP generator, which generates byte code for the communication procedures, following the communication protocol that was selected. *Serialization* summarizes much of this and Java helps us do it.

The CP generator has to be extremely flexible, as it has to be able to generate code that uses various kinds of communication schemes. In a nutshell, CPs determine the way a ST should be delivered to the computing host's worker depending on the communication environment. For the localhost, it is plain TLP (i.e., we create Java threads on a local machine) so `NullCommunicationProcedure` (Section B.2) is used. For distributed environment CPs wrap transport functions over Jini, DCOM+, CORBA, PVM, and RMI (see [Lu04, VP05]) protocols. Both CP and ST interfaces are presented in Section 4.1.1.8.

3.3.3.3 C Sequential Threads and Communication Procedures with the JNI

This is the methodology of how to extend the Java ST/CP generation concepts to C (and similarly can be done for C++) with the JNI [Ste05] introduced in Section 2.6.1.2, page 36. This approach was designed, but not implemented as of this writing; however, it may serve as a good head start on the implementation of the `CCompiler` in GICF.

Much of the ST wrapper class generation code for C will be similar to that of Java. The main difference is the bodies of the sequential thread functions will not be present in the generated class as-is, but they will be declared as `native` with no Java implementation. The C code chunks will be saved to a `.c` file and the corresponding `.h` file will be generated declaring all the needed prototypes with the `javah` tool provided with the standard distribution of the JDK. After that, we call an external C compiler to compile the C chunks into a shared library. Thus, the other modification to the generated wrapper class the `CCompiler` has to do, is to add a

static initializer with the `System.loadLibrary()` call for the newly compiled library with the C implementation of our ST(s). The generated ST class and the compiled mini-library can be stored together (e.g. the binary library file can be loaded into a byte array of the class and deserialized back when about to be executed) in the imperative node and later be communicated just like Java STs. A more sophisticated alternative is to do the compilation and dynamic loading *after* communication by the engine, but this can be a next step.

As far as type matching concerned, we still can use the same mapping rules defined in Section 3.3.2 (and subsequently the `TypeMap` class of the `JavaCompiler` presented later on) because with the JNI with still work with Java and the JVM can do Java-to-native type translation to C or C++ for us, not only for primitive types, but also for arrays, objects, and strings.

3.3.3.4 Worker Aggregator Definition in the Generator-Worker Architecture

The GIPSY uses a generator-worker execution architecture as shown in Figure 34. The GEER generated by the GIPC is interpreted (or executed) by the generator following the educative model of computation. The low-charge ripe sequential threads are evaluated locally by the generator. The higher-charge ripe sequential threads are evaluated on a remote worker. The generator consists of two systems: the Intensional Demand Propagator (IDP) and the Intensional Value Warehouse (IVW) [Tao04]. The IDP implements the demand generation and propagation mechanisms, and the IVW implements the warehouse. A set of semantic rules that outlines the theoretical aspects of the distributed demand propagation mechanism has been defined in [Paq99]. The worker simply consists of a “Ripe Function Executor” (RFE), responsible for the computation of the ripe sequential threads as demanded by the generator. The sequential threads are compiled and can be either downloaded/uploaded dynamically by/to the remote workers. Better efficiency can be achieved by using a shared network file system.

An example: a GIPSY screen saver would be a sample worker running when the an ordinary PC is going into an idle mode and normally launches ordinary dancing bears screensavers, it can actually run our downloaded worker instead and contribute to computation. When such a worker starts, it has to register it within a system

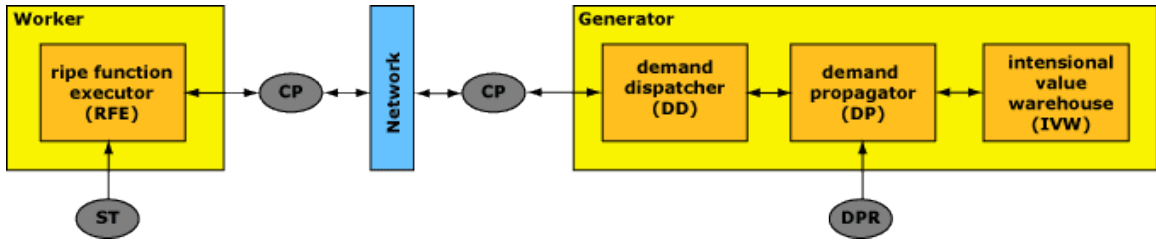


Figure 34: Generator-Worker Architecture

somehow (see [VP05]), so that the generators are aware of its presence and can send demands to it. In the event of merging of semantics of a worker and a generator, such a screensaver would also be able to generate demands and maintain a local warehouse.

3.4 Summary

This chapter presented methodology behind concrete implementations of the first two hybrid languages in the GIPSY – JLucid and Objective Lucid. Semantic rules were presented for free Java functions and Java objects to be included into the Lucid programs and evaluated by the education engine in the hybrid environment. Furthermore, operational semantics of Objective Lucid is clearly defined and is compatible with the semantics of Lucid. The general requirements for the GICF, a tool simplifying imperative compiler management within GIPC, are introduced. The follow up chapter details the architectural and detailed designs and concrete implementation of the languages as well as General Intensional Compiler Framework and overall module integration and their interfaces. Some immediate benefits and limitations are outlined below.

3.4.1 Benefits

- JLucid opens the door for STs and CPs and first hybrid programming paradigm in the GIPSY.
- JLucid provides ability to either write Java code alongside the Lucid code or embed existing one via `embed()`.
- Objective Lucid introduces Java objects and their semantics in the GIPSY.
- GICF generalizes the `embed()` mechanism to all languages in the GIPSY.
- GICF promotes general type handling in the GIPSY.
- GICF promotes general compiler handling in the GIPSY.
- GICF generalizes the notion of the STs and CPs for all compilers.

3.4.2 Limitations

- JLucid is limited only to GIPL-Java and Indexical Lucid-Java hybrids.
- JLucid does not allow Java objects.
- JLucid restricts the `embed()` mechanism only to itself and its derivative – Objective Lucid.
- Objective Lucid is primarily an experimental language to research on Java objects in the intensional environment.
- GICF addresses mostly the imperative compilers, but a similar approach can be applied to the intensional and functional ones.

Chapter 4

Design and Implementation

This chapter combines the architectural and detailed designs and integration of the modules contributed not only by the author of this thesis but also by the other GIPSY team members. Section 4.1 explores the GIPSY architecture and implementation of the major components and frameworks. Then, Section 4.2 focuses on the user interface and external library interfaces. User interfaces, class and sequence diagrams are provided mostly following the top-down approach. For GIPSY Java packages, directory structure with description of each package, and .jar file packaging please refer to Appendix C.

4.1 Internal Design

The GIPC framework redesign along with the realization of the two children frameworks of GICF and IPLCF are presented first followed by the design and implementation of JLucid and Objective Lucid integrated into the new frameworks.

4.1.1 General Intensional Programming Compiler Framework

The GIPC Framework experienced several iterations of refinements as a result of this research. Two new frameworks emerged, namely General Imperative Compiler Framework (GICF) to handle all imperative languages within the GIPSY and, its counterpart Intensional Programming Languages Compiler Framework (IPLCF).

4.1.1.1 General Imperative Compiler Framework

GLU [JDA97, JD96], JLucid [MPG05], and later Objective Lucid [MP05b] prompted the development of a General Imperative Compiler Framework (GICF). The framework targets integration (embedding of) different imperative languages into GIPSY (see [RG05a]) programs for portability and extensibility reasons. GLU promoted C and Fortran functions within; JLucid/Objective Lucid promote embedded Java. Since GIPSY targets to unite all intensional paradigms in one research system, we try to be as general as possible and as compatible as possible and pragmatic at the same time.

For example, if we want to be able to run GLU programs with minimum (if at all) modifications to the code base, GIPSY has to be extended somehow to support C- or Fortran-functions just like it does for Java. What if later on we would need to add C++, Perl, Python, shell scripts, or some other language for example? The need for a general “pluggable” framework arises to add imperative code segments within a GIPSY program. We could go even support multi-segment multi-language (with multiplicity of 3 or more languages) GIPSY programs. Two examples are presented in Figure 35 and in Figure 36.

4.1.1.2 Generalization of a Concrete Implementation

Thus, the `JavaCompiler` component (see Figure 33), part of GIPC, has to be generalized, and the `JavaCompiler` itself be a concrete implementation of this generalization. The generalization would express itself by having an abstract class `ImperativeCompiler`, the generic `Preprocessor` (vs. `JLucidPreprocessor` in Section 4.1.2) should be able to cope with all PLs and know what PLs are supported through enumerating them. Another thing the GICF buys us is an ability to have any supported imperative programming language embedded in any supported intensional programming language. Though this may seem impractical at the first glance, but the framework is designed such that a lot of syntax, semantics, and type mapping work is performed by the individual concrete compiler implementations and not by the generic machinery. The goal here is that as long as any given compiler within the framework conforms to the designed interface specification and produces the required data structures, there should be least possible effort to enable such a compiler in GIPSY. Thus, the compilation process, semantic checks, linking, and execution at

```

#funcdecl

Integer f();
void gee();
void z();

#typedekl

Integer;

#JAVA

Integer f()
{
    return new Integer("123");
}

#CPP

#include <iostream>

void gee()
{
    cout << "gee" << endl;
}

#PERL

sub z()
{
    while(<STDIN>)
    {
        s/\n//;
        print;
    }
}

#OBJECTIVELUCID

A @.d 5
where
    dimension d;
    A = B fby.d (A - 1);
    B = C fby.d (B + f().intValue());
    C = z() && gee();
end;

```

Figure 35: Example of a hybrid GIPSY program.

```

/**
 * Language-mix GIPSY program.
 *
 * $Id: language-mix.ipl,v 1.5 2005/04/25 00:16:30 mokhov Exp $
 * $Revision: 1.5 $
 * $Date: 2005/04/25 00:16:30 $
 *
 * @author Serguei Mokhov
 */
#typedecl

myclass;

#funcdecl

myclass foo(int,double);
float bar(int,int):"ftp://newton.cs.concordia.ca/cool.class":baz;
int f1();

#JAVA
myclass foo(int a, double b)
{
    return new myclass(new Integer((int)(b + a)));
}

class myclass
{
    public myclass(Integer a)
    {
        System.out.println(a);
    }
}

#CPP
#include <iostream>

int f1(void)
{
    cout << "hello";
    return 0;
}

#OBJECTIVELUCID

A + bar(B, C)
where
    A = foo(B, C).intValue();
    B = f1();
    C = 2.0;
end;

/*
 * in theory we could write more than one intensional chunk,
 * then those chunks would evaluate as separate possibly
 * totally independent expressions in parallel that happened
 * to use the same set of imperative functions.
 */

// EOF

```

Figure 36: Another example of a hybrid GIPSY program.

the meta level of implementation of the GIPC and GEE can be reasonably generalized without loss of practicality as we shall see. With this great deal of flexibility, we have several issues:

- Binary portability of compiled languages, such as C/C++ on a different host (this problem theoretically does not exist for Java).
- Though some languages, such as Perl, Python, shell scripts, are interpreted, a version mismatch may happen.
- A compiler for interpreted languages other than Java would be rather simple because should we want to pass the ST code to a remote host, all we need is to pass the source itself. Of course, in both compiled and interpreted variant there is a large potential of security vulnerability exploits (e.g. with malicious code injection), which will have to be dealt with as a part of the future work. As of this writing, there are no embedded checks in GIPSY for that; instead a guide of a sandboxed installation of GIPSY will be provided when the system is released.
- Another important issue is having imperative PL nodes in the AST. The issue is in what such nodes should contain in order for them to be linked back into the main AST, how to perform semantic analysis of the hybrid code based on the contents of such nodes, and GEE should go about executing this code.
- Various languages define their own set of types and typing rules, gluing them all together is a very difficult task for semantic analysis and type inference.

The follow up sections clarify and address most of these issues.

4.1.1.3 Resolving Generalization Issues and Binary Compatibility

In order to fully support GICF, the original GIPC framework in Figure 37 (discussed in detail by Wu and Paquet in [PGW04]) has to be altered in the following way: the Preprocessor has to be added on top of all the front-end modules, and new links drawn between the Preprocessor and the other modules Figure 38. This also changes the data structures flow between the components. For the unaware reader, what follows

is the brief description of the layers, components, and abbreviations of the conceptual design present in Figure 38:

The front-end and back-end layers are the two bottom ones represent the main machinery of the GIPC. The front-end compilers and parsers are responsible for parsing, producing initial syntax trees, STs, and CPs. At this layer, the main abstract syntax tree AST is always compliant to the one of Generic Intensional Programming Language (GIPL). If the source code program was written in some specific intensional programming language (SIPL, e.g. Indexical Lucid or Tensor Lucid), its AST has to be translated first into GIPL. Both, GIPL and SIPL type components may translate a Lucid dialect source code into a data flow (DFG) graph language and back; hence, there is a variety of the DFG translators. Next, the other two types of conceptual components at the front-end layer are the data type (DT) and the sequential thread (ST) front-ends. These correspond to the imperative language compilers and their modules in the implementation. The DT front-end is responsible for analyzing data-type definitions in the ST code and producing *native* (i.e. compiled) representation of communication procedures (NPCs). The ST front-end is responsible for compilation an ST code and producing some equivalent of the native compiled code (NST) as the end result.

The GIPC back-end layer performs finalization of a GIPSY program compilation by doing semantic analysis and eliminating Lucid functions and producing the demand AST (DAST) along with linking in the generated STs and CPs from the imperative side. The GEER generator then produces the final linked version of a GIPSY program as a resource usable by the GEE (GEER).

The first two layers are meta-level layers that prepare information for the front-end and back-end layers. The second layer is the GIPC Preprocessor layer discussed in depth through the rest of this chapter. The top level has to do with some language specification processing and creating corresponding parsers and data structures for the front-end layer. SIPL and GIPL front-end generators have to do with the fact that our SIPL and GIPL parsers are generated out of a source grammar specification by `javacc`. Thus, a GIPL specification corresponds to the GIPL grammar in the `GIPL.jjt` file and the GIPL spec processor is the `javacc` tool. The DT and ST front-end generators exist for the same idea as the GIPL and SIPL ones do. However, in the current implementation they are not present either because they are hand-written

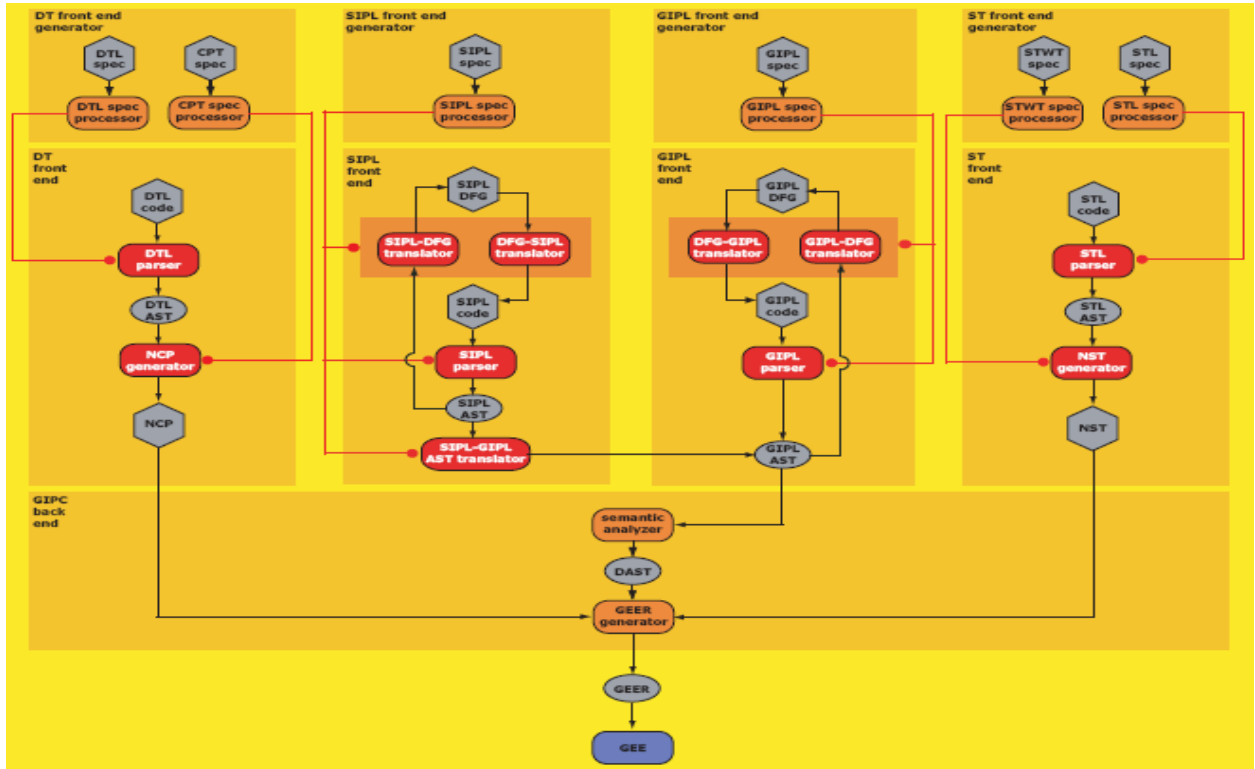


Figure 37: Original Framework for the General Intensional Programming Compiler in the GIPSY

or we rely on the external compiler tools (e.g. `javac` to compile Java STs) to do the processing for us. The design however implies that these components may eventually be converted to the genuine imperative compilers within GIPSY giving greater control and flexibility over the imperative parts than relying on external tools. Therefore, we may acquire a `Java.jjt` one day, for example, and generate a Java parser out of it.

Format Tag To address some binary compatibility issues we invent a notion of a format tag attached to the STs and CPs. The format tag’s purpose is to include meta-information about STs and CPs such that it includes the programming language, the object code format, the operating system, compiler, and their versions. This is important if we are sending platform-dependent compiled code, such as that of C or C++ from one host to another with different architectural platforms. The `FormatTag` API is in Figure 39.

We implement format specifications as a hashtable. We also predefine some common format tags, such as `JAVA`, for conveniences as most frequently used. The class

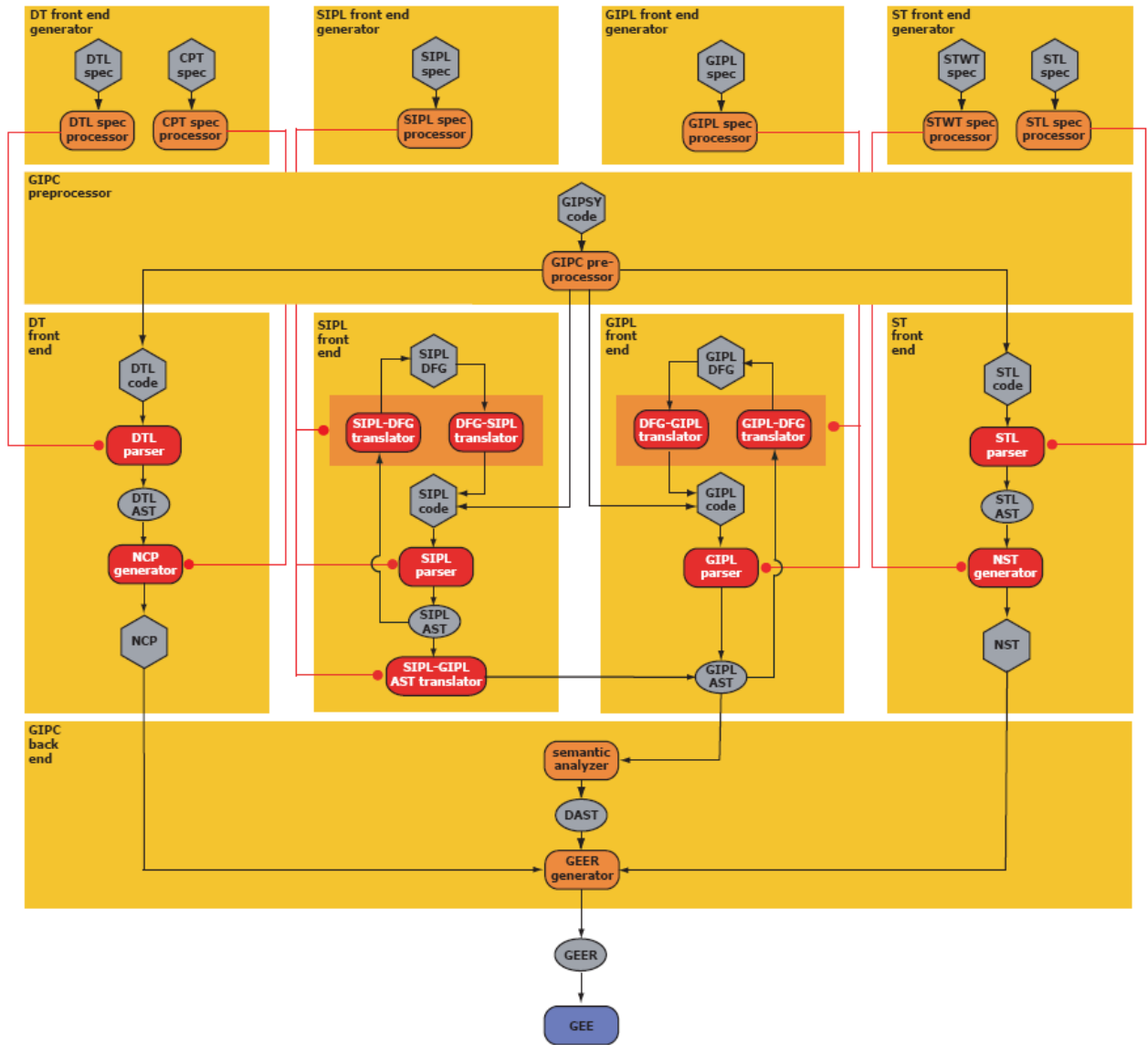


Figure 38: Modified Framework for the General Intensional Programming Compiler in the GIPSY

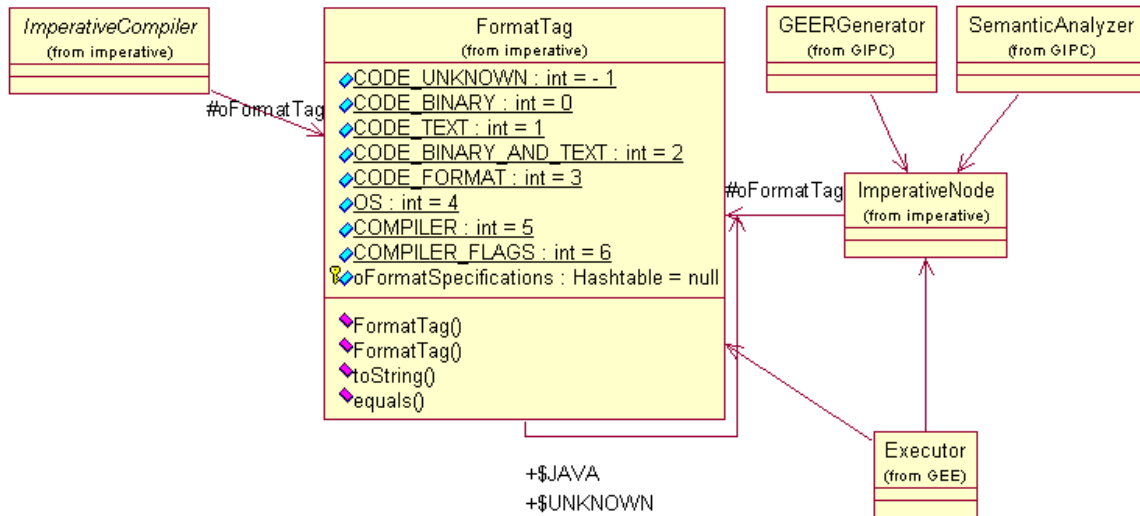


Figure 39: The FormatTag API.

overrides `toString()` and `equals()` of `Object` to define that the two format tags are only equal if the string representation of all their specifications are identical.

Sending Source Code Text Not all non-intensional languages require compilation, e.g. Perl, Python, etc. These can be sent over as plain source code text; thus, the format tag will indicate the fact. We can go even further with this and send *any* language as plain text and compile it on the target host instead prior invocation. For the task of the source code inclusion we reserved the `SequentialThreadSourceGenerator`. Of course, this won't work for embed-included binary code via a URI parameter because that code was already compiled by someone else on some specific platform. As far as current implementation concerned, the generated ST class does always contain the source code of STs from the GIPSY program code segments, but it is unused by the GEE except for debugging as of this writing.

Dictionary The Preprocessor's dictionary will initially be constructed based on the `#funcdecl` and `#typeddecl` program segments. The dictionary will serve as an input to three other components: the NST generator (for error reporting and pointers to the nodes in the AST and the compiled code), to the NCP generator (to analyze the data structures used by STs and generate CPs accordingly), and to the semantic analyzer, to perform data type matching between the intensional and imperative parts. Both NCP and NST generators work under the command of some

imperative language compiler and are referred to as `SequentialThreadGenerator` and `CommunicationProcedureGenerator` in their most general forms, which are subclassed by a concrete language implementation.

4.1.1.4 GIPC Preprocessor

The `Preprocessor` is something that is invoked first by the GIPC on incoming GIPSY program's source code stream. The `Preprocessor`'s job is to do preliminary program analysis, processing, and splitting into chunks. Since a GIPSY program is a hybrid program consisting of different languages in one source file, there ought to be an interface between all these chunks. Thus, the `Preprocessor` after initial parsing and producing the initial parse tree, constructs a preliminary dictionary of symbols used throughout the program. This is important for type matching and semantic analysis later on. The `Preprocessor` then splits the code segments of the GIPSY program into chunks preparing them to be fed to the respective concrete compilers for those chunks. The chunks are represented through the `CodeSegment` class that the GIPC collects. The corresponding class diagram of is in Figure 40.

The `Preprocessor` can also be told to report certain code segments are invalid at the preprocessing stage rather delaying the error until the compiler discovery stage through the `addInvalidSegmentName()` and `addValidSegmentName()` methods and maintaining internal vector of the strings with invalid segment names. This feature is for example used in `Preprocessor`'s extensions of `JLucidPreprocessor` and `ObjectiveLucidPreprocessor` later on that filter out code segments that do not belong to the languages. The filtering logic works like this:

- if no valid and invalid segments are specified, all segments are accepted as valid at the preprocessing stage. This is the default for general GIPC work.
- if some invalid and no valid segments are specified, the `Preprocessor` will error out on the invalid segments
- if only valid segments are specified, everything else will be treated as invalid
- if both valid and invalid segments are present; the invalid set segments are ignored and everything that it is not mentioned in the valid set is said to be invalid.

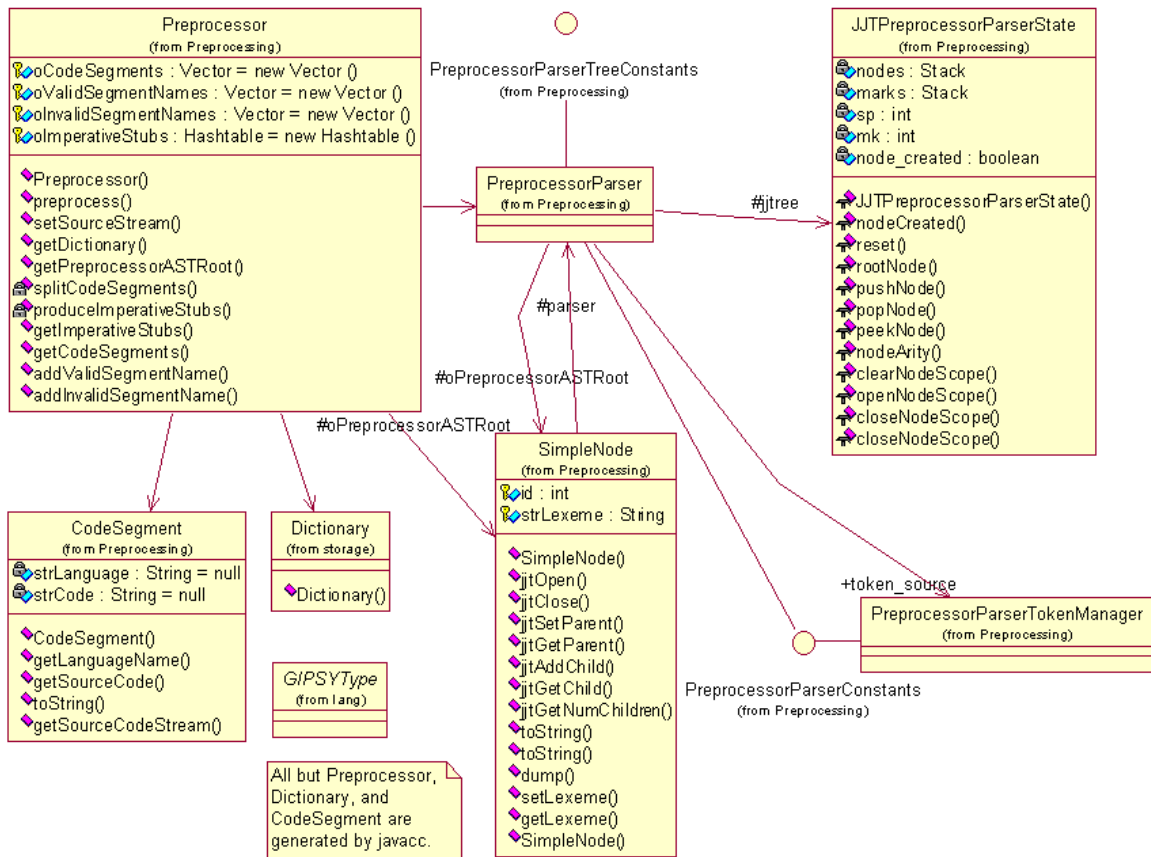


Figure 40: The GIPC Preprocessor.

GIPSY Program Segments Here we define four basic types of segments to be used in a GIPSY program. These are:

- **#funcdecl** program segment declares function prototypes of imperative-language functions defined later or externally from this program to be used by the intensional language part. These prototypes are syntactically universal for all GIPSY programs and need not resemble the actual function definitions they describe in their particular programming language.
- **#typedcl** segment lists all user-defined data types that can potentially be used by the intensional part; usually objects. These are the types that do not appear in the matching table in Table 1.
- **#<IMPERATIVELANG>** segment declares that this is a code segment written in whatever IMPERATIVELANG may be, for example **#JAVA** for Java, **#CPP** for C++, **#PERL** for Perl, **#PYTHON** for Python, etc.

- `#<INTENSIONALLANG>` segment declares that this is a code segment written in whatever `INTENSIONALLANG` may be, for example `#GIPL`, `#INDEXICALLUCID`, `#JLUCID`, `#OBJECTIVELUCID`, `#TENSORLUCID`, `#ONYX`¹, etc. as understood by the GIPSY.

Preprocessor Grammar The initial grammar for the `Preprocessor` to be able to parse a GIPSY program is shown in Figure 41. After having parsed a program, we have a Preprocessor AST (PAST) that will be used further by the compilation process in the GIPC and its submodules. The grammar and the framework were designed in such a way so all the previous neat features of JLucid [MP05b]/Objective Lucid [MP05b] still be present, such as `embed()` and are accessible to other dialects. In the GICF, we generalize our function prototype declaration to be able to include external code of any imperative language.

The lexical elements, such as `LETTER`, `LANGDATA`, `DIGIT`, `CAPLETTER`, and `*LITERALS` are not listed for brevity as they are merely standard and self-explanatory lexical tokens except probably `LANGDATA` – this is character data allowing any character sequence within except `LANGID` that serves as a terminator of a code segment chunk.

Notice, the grammar is not bound to our current set of supported intensional and imperative languages. Rather, the GIPC attempts to look up appropriate compiler for each code segment automatically using `LANGID` for mapping at run-time. The JavaCC version of the grammar can be found the `PreprocessorParser.jjt` file.

The grammar has been amended from what was published in [MP05a] to include `LANGID` in the `EMBED` production, the `immutable` keyword and arrays subscript operator `[]` in the `PSTART` production. `LANGID` in `EMBED` is needed to be able to pick the appropriate compiler for the included code as it may be written in any imperative language. The `immutable` keyword is needed to allow a programmer to assert that certain STs are immutable meaning given the same parameters they always return the same result, and, therefore, their result can be safely cached in the warehouse as such functions are declared side-effects free (e.g. as the `get42()` method in Figure 23, page 48 can be marked as immutable). This marking of methods will allow more efficient caching of the ST results of STs *known* not to have side effects

¹See [Gro04] for details on the Onyx language.

<GIPSY>	::=	<DECLARATIONS> <CODESEGMENTS>
<DECLARATIONS>	::=	<FUNCDECLS> <DECLARATIONS> <TYPEDECLS> <DECLARATIONS> ϵ
<FUNCDECLS>	::=	#funcdecl <PROTOTYPES>
<TYPEDECLS>	::=	#typeddecl <TYPES>
<PROTOTYPES>	::=	<PROTOTYPE> ; <PROTOTYPES> ϵ
<PROTOTYPE>	::=	<PSTART> <EMBED>
<PSTART>	::=	[immutable] <TYPE> [[]] <ID> (<TYPELIST>)
<EMBED>	::=	ϵ : <LANGID> : <URI> : <LANGID> : <URI> : <ID>
<TYPES>	::=	<TYPE> ; <TYPES> ϵ
<TYPELIST>	::=	<TYPE> [[]] <TYPE> [[]] , <TYPELIST> ϵ
<CODESEGMENT>	::=	<LANGDATA> <LANGID> <LANGDATA> <EOF>
<CODESEGMENTS>	::=	<CODESEGMENT> <CODESEGMENTS> ϵ
<URI>	::=	<CHARACTERLITERAL> <STRINGLITERAL>
<ID>	::=	<LETTER> (<LETTER> <DIGIT>)*
<LANGID>	::=	#<CAPLETTER> (<CAPLETTER>)*
<TYPE>	::=	<ID> int double bool float char string void

Figure 41: Preprocessor Grammar for a GIPSY program.

and has to be explicitly set by the programmer. If the programmer by mistake marks a method with side effects as `immutable`, then a program may exhibit erroneous execution at run-time by returning a possibly incorrect value from the warehouse. There is no way to automatically discover immutability of STs in GIPSY at this time (it may only be possible when genuine imperative compilers are implemented). The array subscript operator `[]` has been added to `PSTART` and `TYPELIST` productions to allow GIPSY arrays (as a generalization of JLucid arrays) that are composed of the elements of GIPSY types. The concrete imperative compilers implementing the mapping (if possible) will have to do appropriate conversions from the native arrays to GIPSY arrays.

4.1.1.5 GIPSY Type System

While the main language of GIPSY, Lucid, is polymorphic and does not have explicit types, co-existing with other languages necessitates definition of GIPSY types and their mapping to a particular language being embedded. Figure 42 presents the design aspects of the GIPSY Type System.

Each class is prefixed with `GIPSY` to avoid possible confusion with similar definitions in the `java.lang` package. The `GIPSYVoid` type always evaluates to the Boolean `true`, as described earlier in Section 3.3.2. The other types wrap around the corresponding Java object wrapper classes for the primitive types, such as `Integer`, `Float`, etc. Every class keeps a lexeme (a lexical representation) of the corresponding type in a GIPSY program and overrides `toString()` to show the lexeme and the contained value. These types are extensively used by the `Preprocessor`, imperative and intensional (for constants) compilers, the `SequentialThreadGenerator`, `CommunicationProcedureGenerator`, `SemanticAnalyzer` for the general type of GIPSY program processing, and by the `GEE Executor`.

The other special types that have been created are either experimental or do not correspond to a wrapper of a primitive type. `GIPSYIdentifier` type case corresponds to a declaration of some sort of an identifier in a GIPSY program to be put into the dictionary, be it a variable or a function name with the reference to their definition. This is an experimental type and may be removed in the future. Constants and conditionals may be anonymous and thereby not have a corresponding identifier. `GIPSYEmbed` is another special transitional type that encapsulates embedded code via

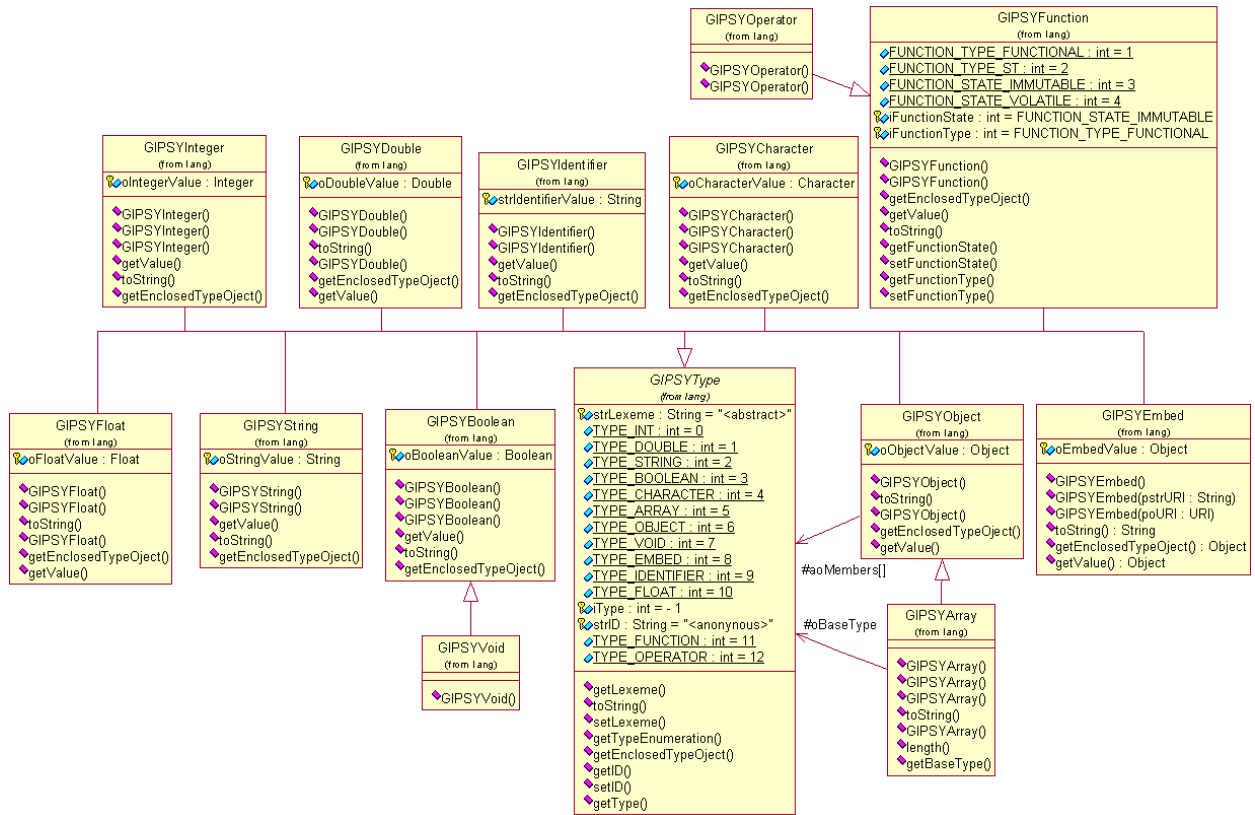


Figure 42: GIPSY Type System.

the URL parameter and later is exploded into multiple types corresponding to STs and their CPs. `GIPSYFunction` and its descendant `GIPSYOperator` correspond to the function types for regular operators and user defined functions. A `GIPSYFunction` can either encapsulate an ordinary Lucid function (as in functional programming an which is immutable) or an ST function (e.g. a Java method), which may easily be volatile (i.e. with side effects). These four types are not directly exposed to a GIPSY programmer and at this point are managed internally. The rest of the type system is exposed to the GIPSY programmer in the preamble of a GIPSY program, i.e., the `#funcdecl` and `#typedec1` segments, which result in the embryo of the dictionary for linking, semantic analysis, and execution. Once ST compilers return, the type data structures (return and parameter types) declared in the preamble are matched against what was discovered by the compilers and if the match is successful, the link is made.

4.1.1.6 GICF Design

The GICF is the first generalization framework of hybrid programming in the GIPSY. Implementation-wise, only Java is implemented as an imperative language with an external compiler. However, provision was made for C/C++, Perl, Fortran and Python with stub compilers. The class diagram describing GICF is shown in Figure 43. On this diagram the interaction between a given imperative compiler and the `SequentialThreadGenerator` and `CommunicationProcedureGenerator` only shown for `JavaCompiler` to keep the clearer picture, but the same kind of association will have to be maintained for all imperative compilers as the `IImperativeCompiler` interface mandates. The `EImperativeLanguages` is a Java interface enumerating all available imperative language compilers. It is used by the GIPC to discover a given compiler for a language dynamically. As of this writing, the enumeration is maintained by hand; however, it is planned to be generated in the near future with a command-line-driven script or a RIPE GUI automagically to facilitate addition of new languages.

4.1.1.7 Intensional Programming Languages Compiler Framework

As a consequence of GICF, a similar approach was applied to the intensional compilers in the form of IPLCF. See the corresponding class diagram in Figure 44. The `IIntensionalCompiler` was designed and implemented by all the intensional compilers we have. An enumeration `EIntensionalLanguages` of all supported intensional languages was created, so the GIPC can pick needed compiler at run-time as determined by the `Preprocessor`.

Translation for all intensional compilers is done through the generic `Translator` implemented by Aihua Wu in [Wu02]. The `Translator` has been integrated into the `GIPC.intensional.GenericTranslator` package and split and renamed as in Figure 45. Thus, every SIPL compiler refers to this translator to acquire a GIPL AST at the end via generic implementation of `IntensionalCompiler.translate()`. The `Translator` was refactored and augmented to understand GIPSY Types (see Section 4.1.1.5) and `ImperativeNode` for imperative languages. The `TranslationParser` and `TranslationLexer` collaborate to compile intensional language translation rules (e.g. `IndexicalLucid.rul`) files provided by each SIPL author.

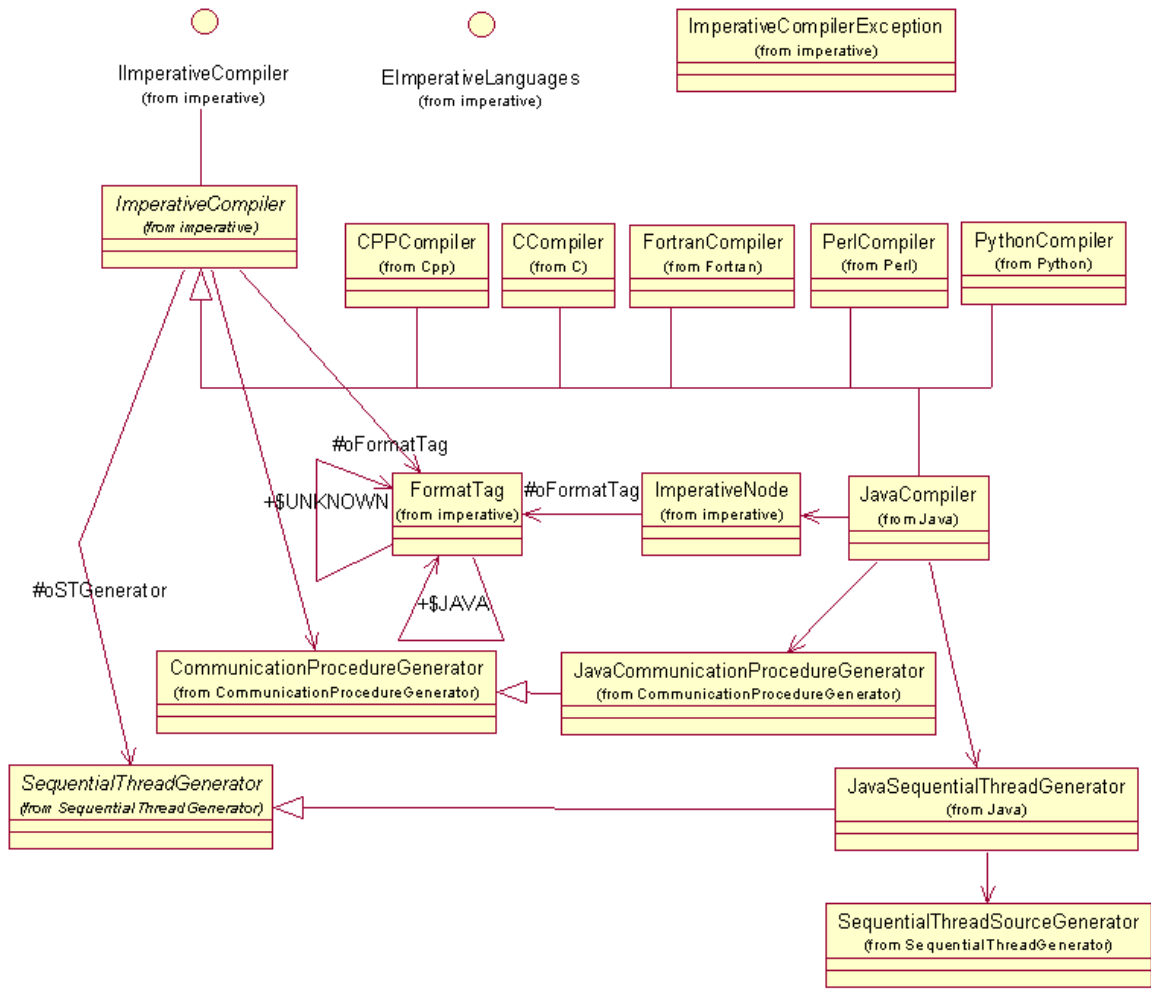


Figure 43: GICF Design.

4.1.1.8 Sequential Thread and Communication Procedure Interfaces

This section details Sequential Thread and Communication Procedure interfaces. The related class diagram is in Figure 46. The `ICommunicationProcedure` and `ISequentialThread` are the core interfaces. Both extend `Serializable` in order for us to be able to dump their concrete implementations to disk or distributed storage using Java’s object serialization machinery. This is needed for the `GIPSYProgram` container to be saved to disk or for an `ST` to be able to reside in `JavaSpaces` [Mam05] implementation of the demand space [VP05]. The `ISequentialThread` also extends `Runnable` to be true thread when materialized, especially for the case of local execution. The `Runnable` interface makes it possible for an implementing class to become a thread in multithreaded environment in Java. The `ICommunicationProceduresEnum`

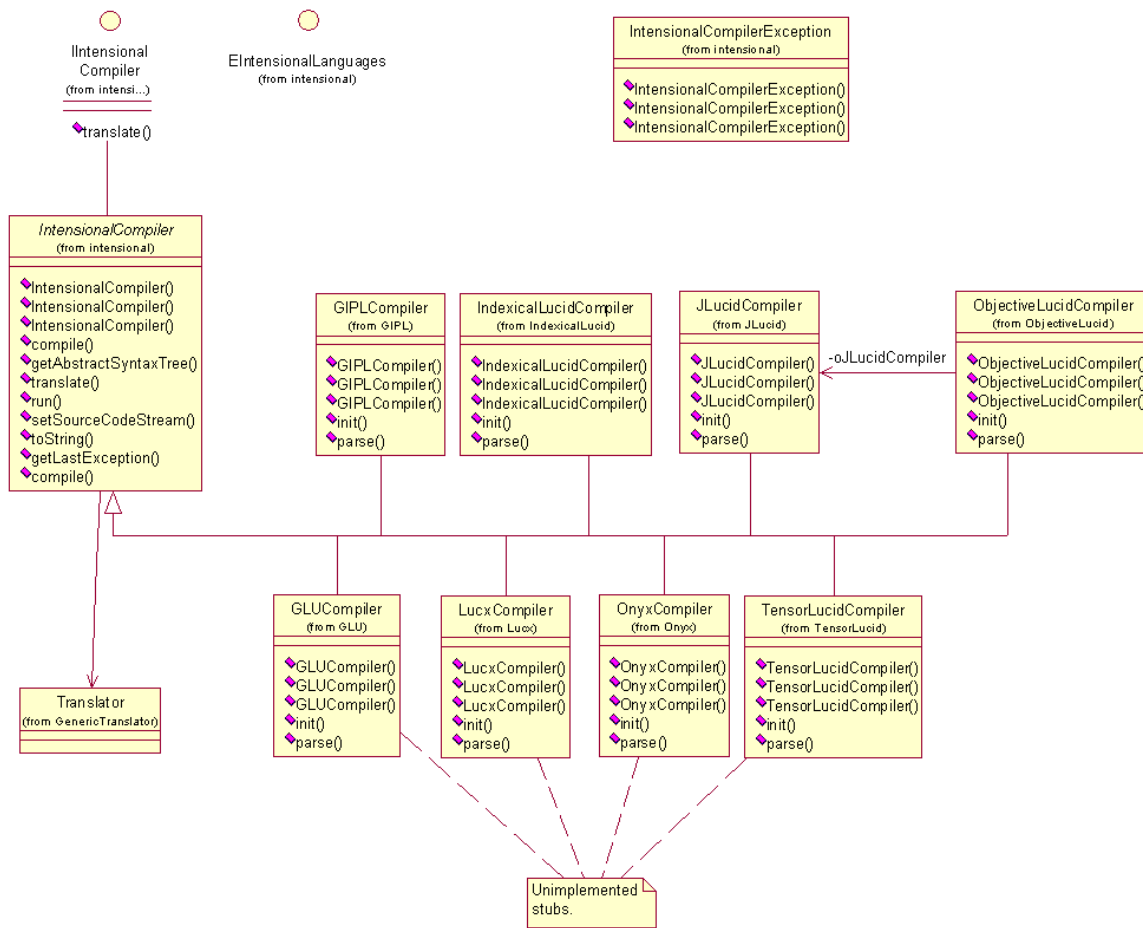


Figure 44: IPLCF Design.

is an enumeration of all known to the GIPSY communication procedure types. The `NullCommunicationProcedure` and `RMICommunicationProcedure` represent concrete implementations for local threaded processing as well as RMI. Therefore, the `SequentialThreadGenerator` is an abstract factory for all sequential threads that has to be overridden by a language-specific sequential thread generator, e.g. such as `JavaSequentialThreadGenerator`. Likewise, `CommunicationProcedureGenerator` is a factory for CPs. The `WorkResult` class represents the result of (computation) work done, which is also has to be `Serializable`. Upon various communication needs the `CommunicationStats` is returned by the `ICommunicationProcedure` API or the `CommunicationException` is thrown indicating an error. The `Worker` class represents a collection of STs and CPs being executed.

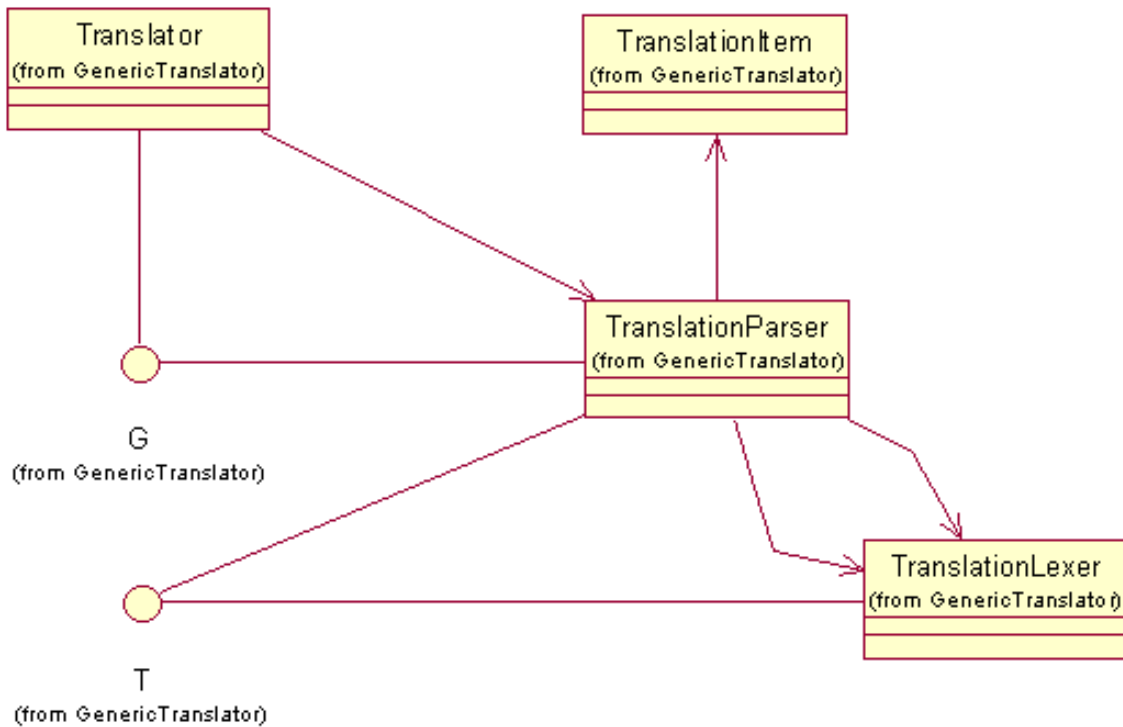


Figure 45: SIPL to GIPL Translator Integration.

4.1.1.9 GIPC Design

In Figure 48 there is a hierarchy that all imperative and intensional compilers should adhere to. The `IImperativeCompiler` interface is something every imperative compiler implements to ease up the job of GIPC. A similar interface has been invented for intensional languages – `IIntensionalCompiler` for consistency.

A set of interfaces has been designed for all the present and future compilers to implement. There are three interfaces so far:

1. `ICompiler` is a superinterface for all compiler interfaces. It is implemented by GIPC itself and by `DFGAnalyzer`, as shown in Figure 47.
2. `IIntensionalCompiler` is a subinterface of `ICompiler` designated to differentiate intensional compilers. It is implemented in part by the `IntensionalCompiler` abstract class that most (for now all) intensional compilers implement.
3. `IImperativeCompiler` is a counterpart of `IIntensionalCompiler`. Its purpose is similar to that of `IIntensionalCompiler` for imperative languages.

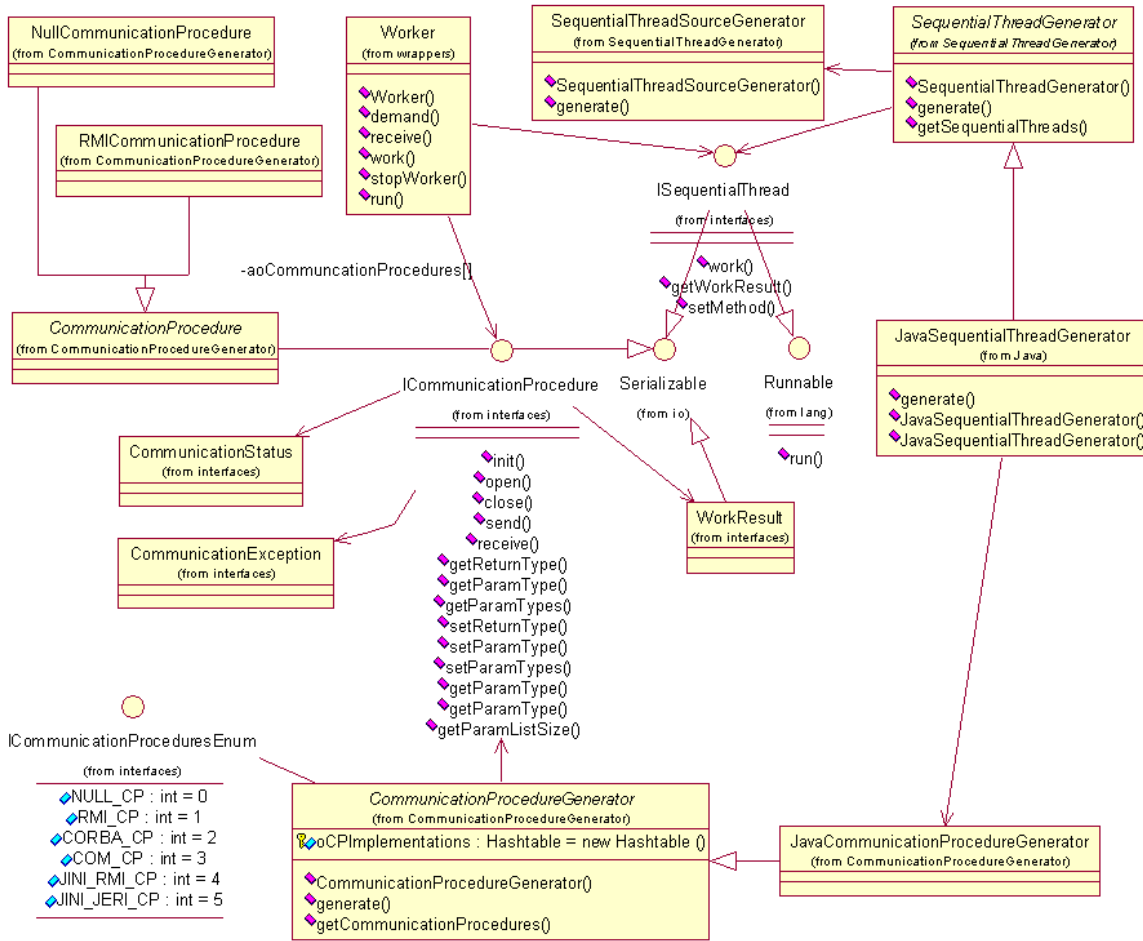


Figure 46: Sequential Thread and Communication Procedure Class Diagram.

The core difference between `IIntensionalCompiler` and `IImperativeCompiler` versus the general `ICompiler` is that most (except for `GIPL`) of the intensional compilers have to perform `SIPL-to-GIPL` translation; hence, the `translate()` method, and all imperative compilers must produce communication procedures and sequential threads as the result of their work; hence, `generateSequentialThreads()` and `generateCommunicationProcedures()` methods are provided. The abstract classes `IntensionalCompiler` and `ImperativeCompiler` provide the most common possible implementation for all intensional and imperative compilers respectively, so the underlying concrete compilers only have to override some parts specific to the language they are to compile. If extension of these classes is not possible for some reason (e.g. when writing external `GIPSY` plugins when a compiler class already inherits from some other class), they must implement their corresponding interface. Out of

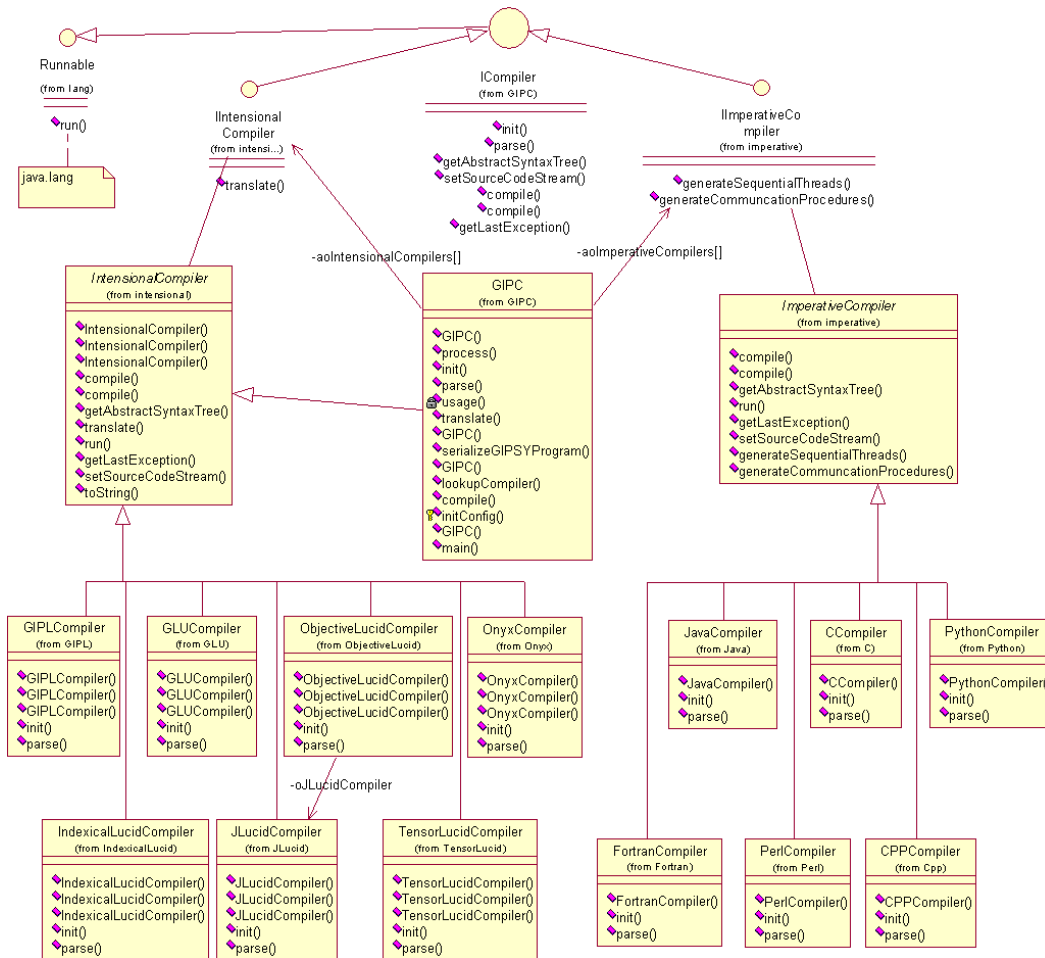


Figure 47: All GIPC Compilers.

the concrete classes on the diagram the author of this thesis fully implemented `GIPC`, `GIPLCompiler`, `IndexicalLucidCompiler`, `JLucidCompiler`, `ObjectiveLucidCompiler`, and `JavaCompiler`. The `DFGAnalyzer` of Yimin Ding was made to implement `ICompiler` as it in fact compiles the “DFG code” out of `GIPL` or `Indexical Lucid`.

The overall design and integration of the `GIPC` participants is illustrated in Figure 48. The `GIPC` class is the main compiler application that drives the compilation process, so in the general case it invokes the `Preprocessor`, intensional and imperative compilers required, the `SemanticAnalyzer`, `IdentifierContextCodeGenerator`, `Translator`, and the `GEERGenerator` linker. It also acts like a facade to other `GIPSY` modules. The major data structures, such as `AbstractSyntaxTree`, `Dictionary`, `CodeSegment`, `FormatTag`, `ImperativeNode`, and `SimpleNode` are created, accessed, or modified throughout the modules during the compilation process. Out of imperative languages only `JavaCompiler` is mentioned as it is the most advanced in this category. The `JLucidCompiler`’s `JLucidParser` underneath invokes both `JGIPLParser` and `JIndexicalLucidParser` as `JLucid` Section 3.1 provides extensions to both of these languages. A number of association links have been removed from the diagram to maintain clarity as these links are intuitive or present in detail diagrams.

4.1.1.10 `GIPC` Class as a Meta Processor

The `GIPC` (a concrete class) acts here as so-called “meta processor” that drives the entire compilation process and invokes appropriate submodules in order to come up with a compiled version of a `GIPSY` program. This involves calling the `Preprocessor`, then feeding its output to whatever concrete compilers for the code segments of the `GIPSY` program, collecting the output of them (various ASTs, dictionaries), performing semantic analysis, and linking all the parts back together in a binary form. This portable binary version of the `GIPSY` program is to either be serialized as an executable file for later execution by the `GEE` or optionally to be fed directly to the `GEE`.

4.1.1.11 Calling Sequence

The sequence diagram in Figure 49 illustrates the entire compilation process and the data structures passed between the modules. This is the roundtrip description of the

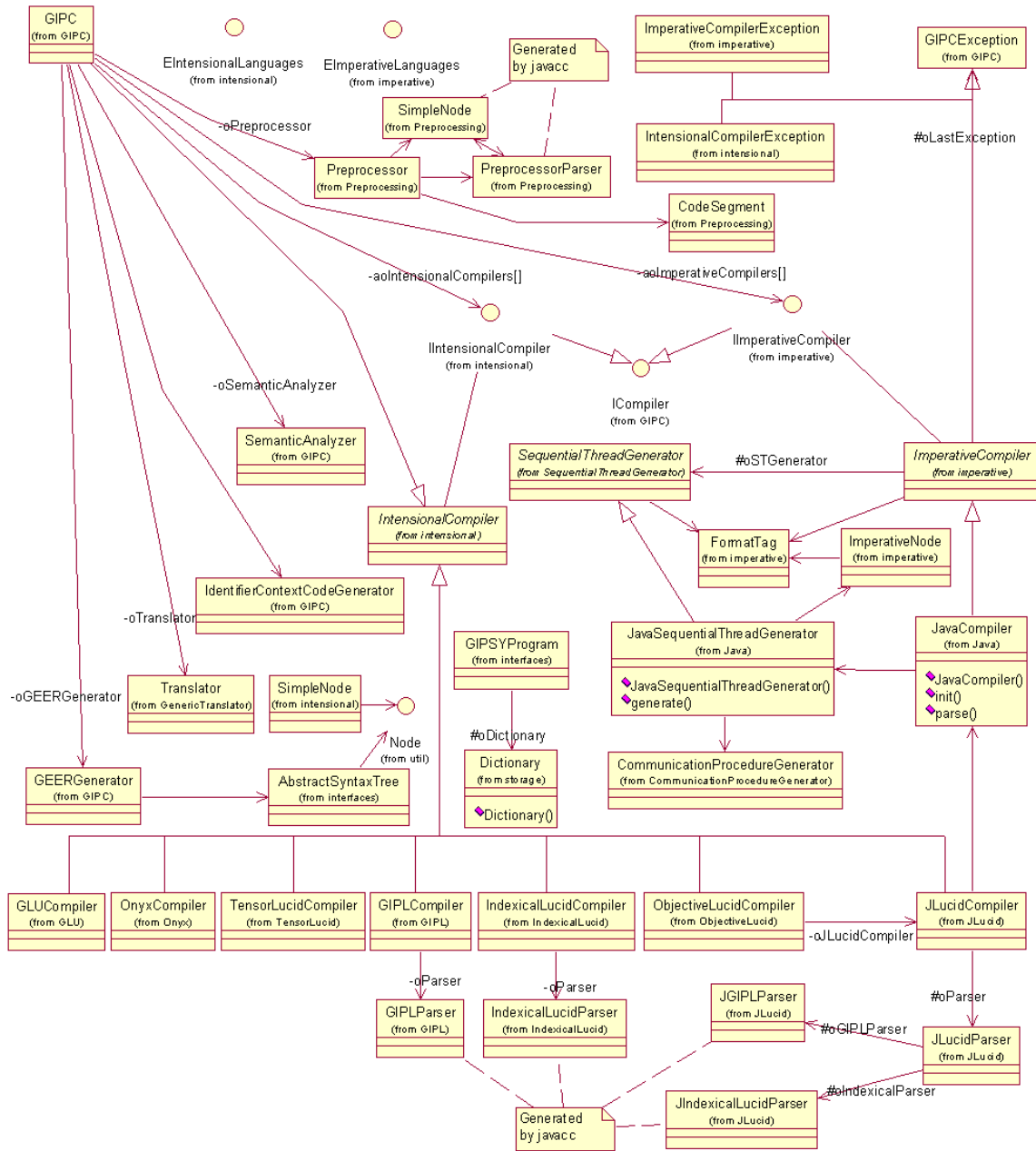


Figure 48: Overall GIPC Design.

implementation efforts. The two followup diagrams detail the differences in the compilation process between the imperative and intensional languages. The general compilation process begins by reading the source GIPSY program and converting it into a meta token stream of types, declarations, and code segments by the **Preprocessor**. The **Preprocessor** takes that input and with its own parser produces a preprocessor AST and an embryo of a dictionary with the identifiers and types declared in the imperative code segments for further semantic linking. The latter is used to produce imperative stubs for cross-segment type checks. The former contains primarily code segments written in various languages. The **GIPC** takes these code segments and creates appropriate compiler threads, one for each code segment. Then, each compiler tries to compile its own chunk and produces a portion of a main AST. Since we treat the IPL part as a main program, its AST is considered to be the main skeleton tree. The ASTs produced by the imperative compilers (which really contain a single **ImperativeNode**) are secondary and should be merged into the main when appropriate. Once all the compiler threads are successfully done, the **GIPC** collects all the ASTs and performs linking via the **GEERGenerator**. The combined AST is now a subject to the semantic analysis and the function elimination. Once semantic analysis is complete, the final post-linking is performed where all the pieces of the **GIPSYProgram** are combined together and its instance is serialized to disk. Optionally, right after compilation the **GEE** may be invoked to start the execution of the just compiled program.

There is no any preference made in **GIPC** on the number and the order of intensional and imperative compilers executed. This may result in several main intensional programs (if the source code contained more than one intensional code segment) or unused imperative nodes (an imperative segment is declared but the code from it is unused). For the former we maintain an array of ASTs in the **GIPSYProgram**, so that when the actual program is executed, the same number of the **GEE Executor** threads are started and all main ASTs are evaluated in parallel providing the result set of a computation instead of a single result. Detailed sequence diagrams of the intensional and imperative compilation processes are in Figure 50 and Figure 51 to illustrate the differences in compiling intensional and imperative code segments.

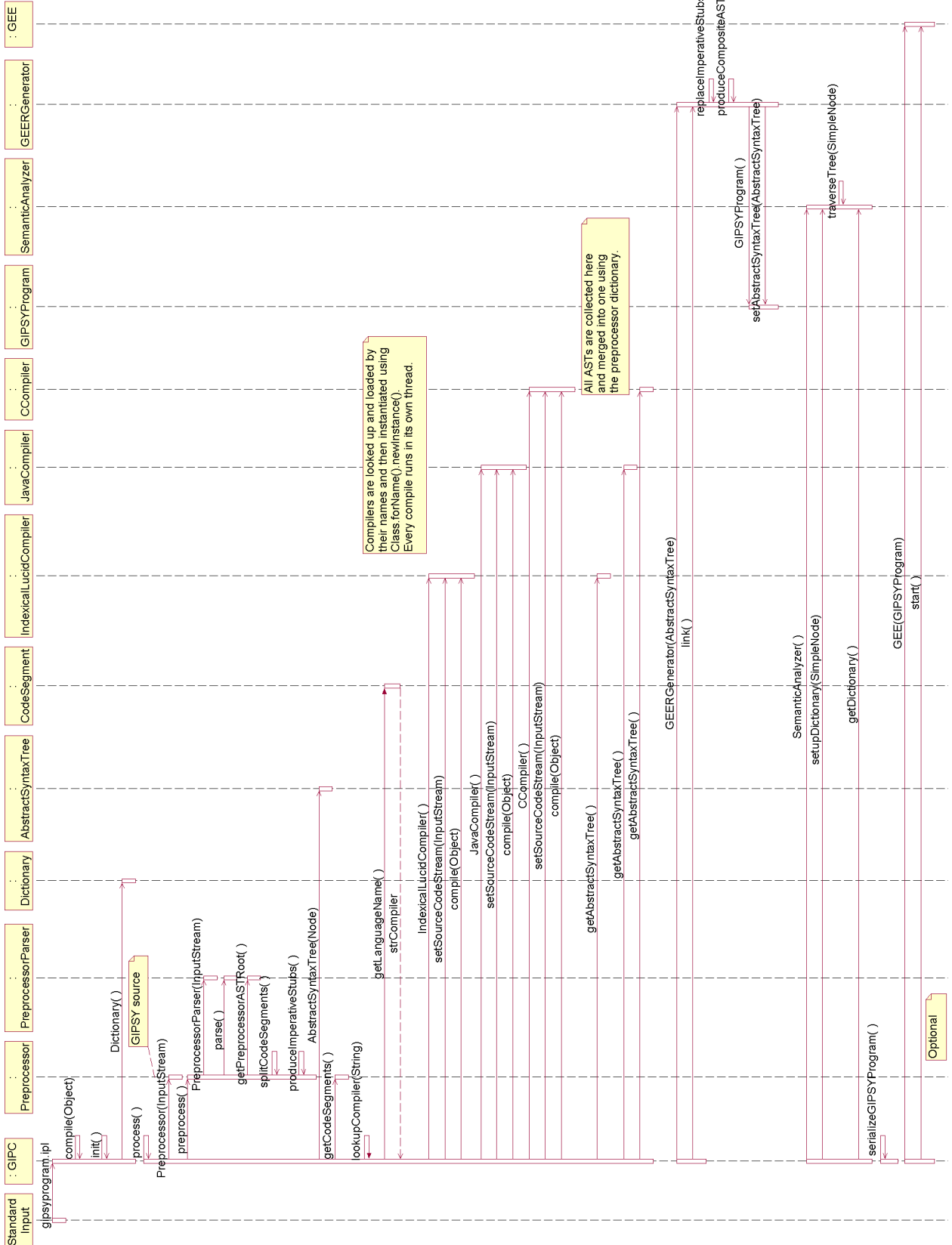


Figure 49: Sequence Diagram of GIPSY Program Compilation Process.

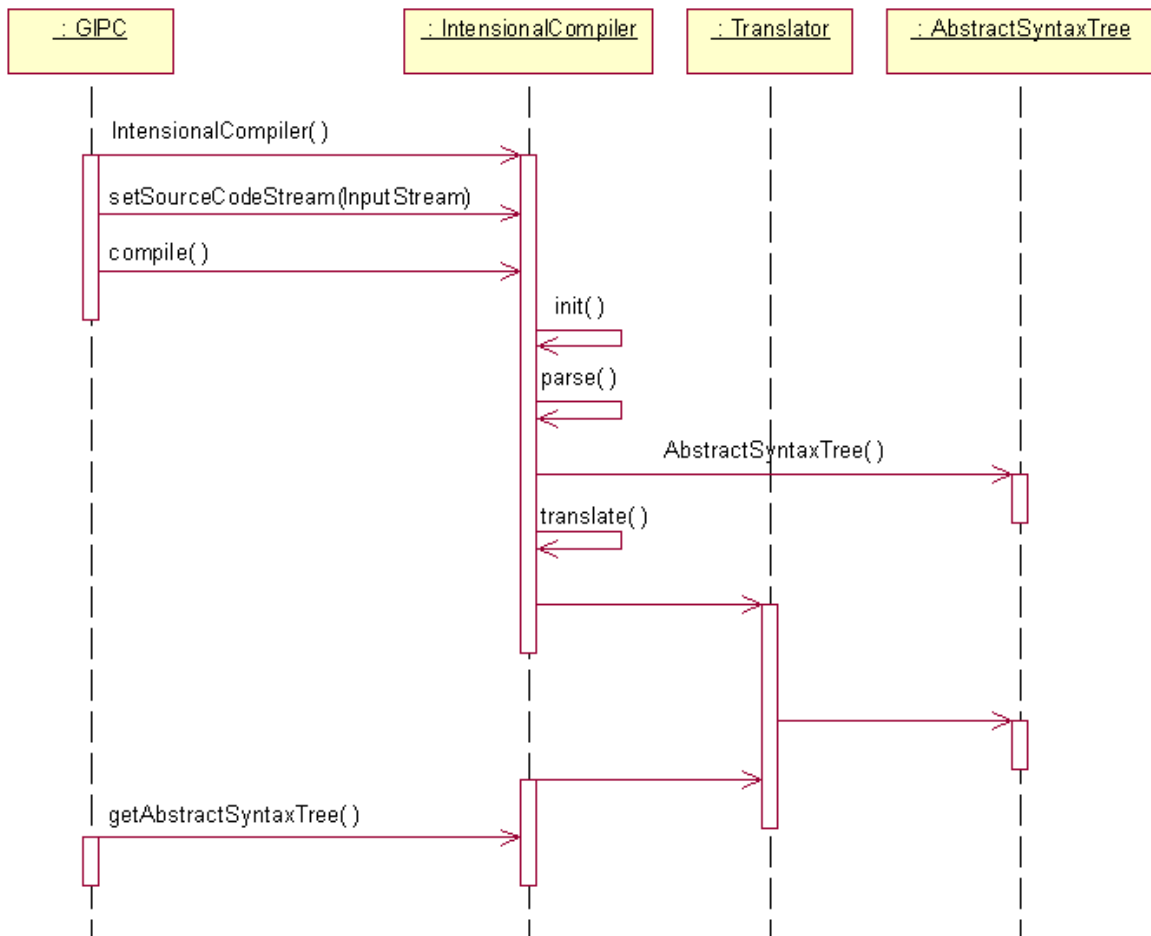


Figure 50: Sequence Diagram of Intensional Compilation Process.

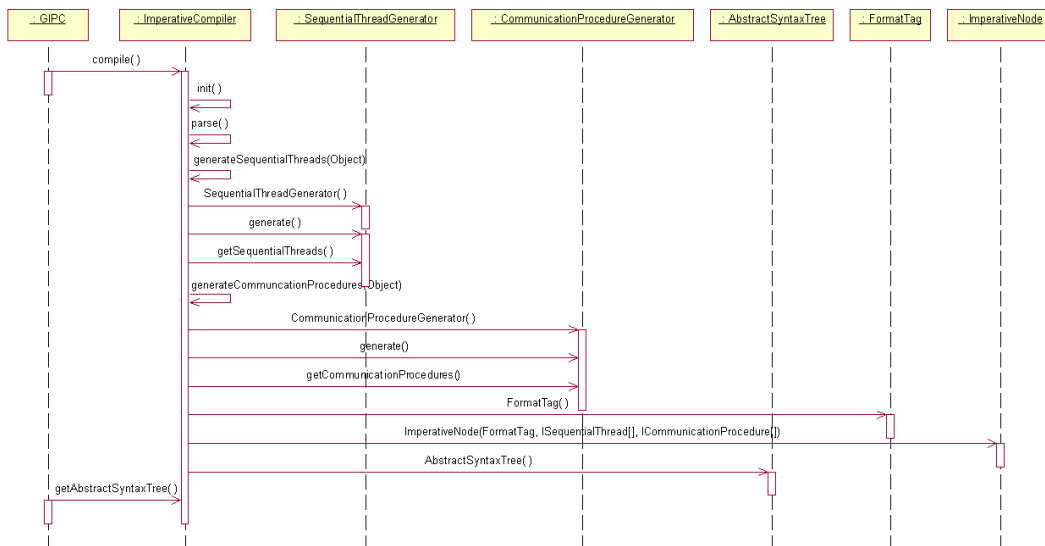


Figure 51: Sequence Diagram of Imperative Compilation Process.

4.1.1.12 Compiling and Linking

Multiple Intensional Parts In a GIPSY program we may possibly have multiple intensional parts. For example, if a GIPSY programmer gave a GIPL expression, an Indexical Lucid expression and a couple of Java procedures in the same source GIPSY program, what is the meaning of that setup would be? In this case, we can say that we evaluate two independent intensional expressions in parallel that happened to share the same imperative part. Thus, for such a GIPSY program there will be two instances of GEE running. The GEE is to extended to accept a forest of ASTs to be processed in parallel.

Imperative Stubs When the **Preprocessor** completes its job, it has to create some stubs in the intensional parts of the program for the symbols declared outside of those parts (e.g. Java functions) so that the appropriate intensional compiler does not complain about undefined symbols when producing the AST because the intensional compilers are not aware of anything outside their work scope. Later on, the corresponding stub nodes in the AST are found and replaced with the real contents at the linking stage.

NCP Generator as a Type Processor The NCP generator will act very much like a type processor and will have to look inside the imperative code segments analyzed/compiled by the ST generator. This kind of type processing is needed to decide on communication procedures (CPs) to be generated for that ST. It issues warnings if the compiled version of the data structures to be sent is not portable. The role of the NCP generators in the GIPSY implementation is played by the imperative compilers, such as `JavaCompiler`.

GEER Generator as a Linker The GEER Generator (see `GEERGenerator` in Figure 53) in the backend acts like a linker of all parts of a GIPSY program. It gathers all the resources from the compiler set, such as ASTs, ICs, CPs, STs, and the dictionary. Then, it replaces the stubs in the intensional part with the nodes from the imperative ASTs (STs accompanied with their respective CPs) forming a complete composite AST ready for consumption by the GEE. All this will be serialized as a `GIPSYProgram` class instance. The `GEERGenerator` is invoked two times – first prior

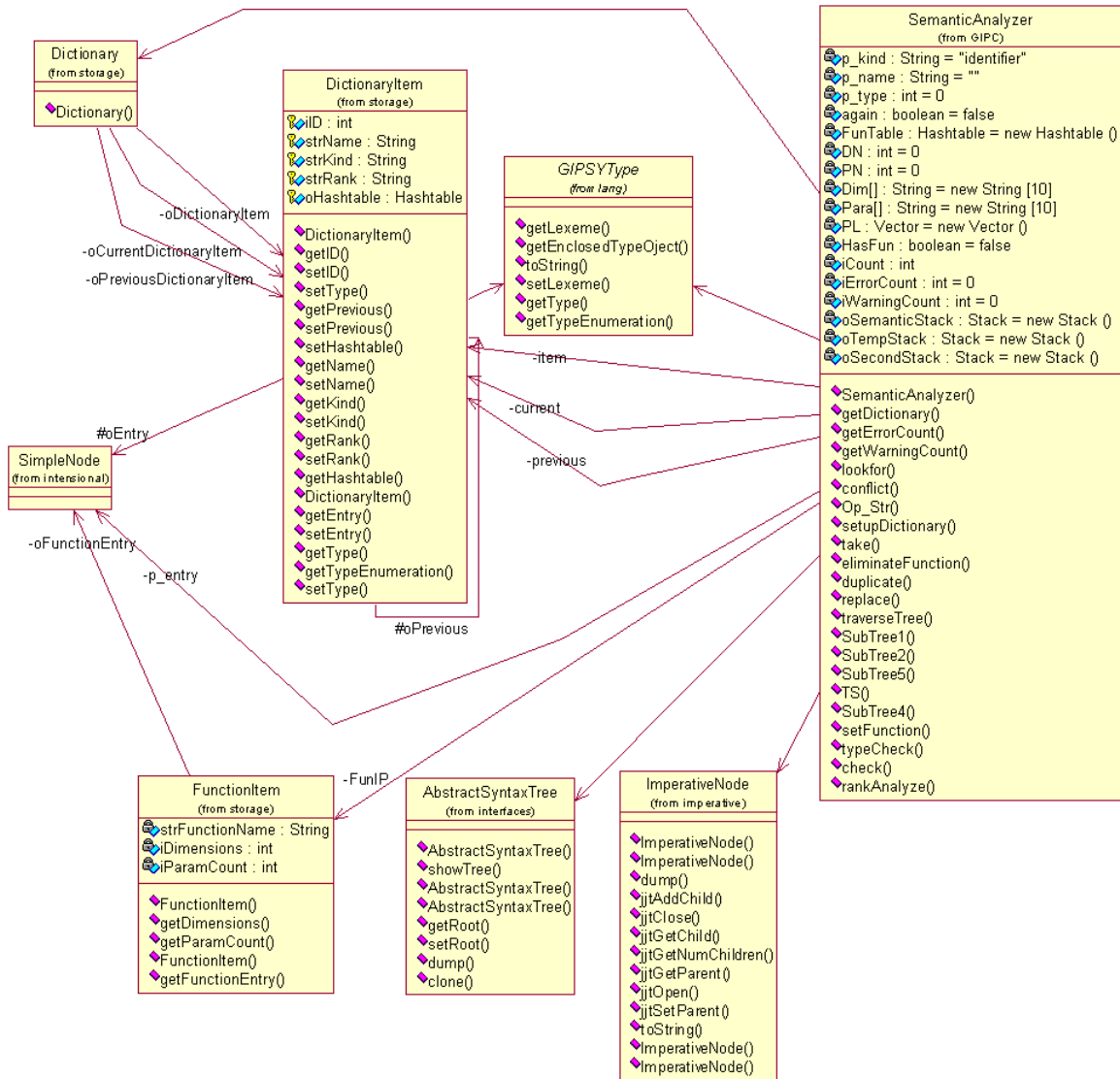


Figure 52: Semantic Analyzer.

SemanticAnalyzer to assemble a complete AST, and then after semantic analysis and function elimination to set up the finalized dictionary and program name.

4.1.1.13 Semantic Analyzer

The semantic analyzer detailed design diagram is shown in Figure 52. Originally implemented by Aihua Wu, the class was renamed from Semantic [Wu02] to a more complete name of SemanticAnalyzer and placed under the GIPC package. Relevant changes include integration of storage.Dictionary (previously was java.util.Vector),

`storage.DictionaryItem` (formerly `Item_in_Dict` [Wu02]), `storage.FunctionItem` (formerly `Fun_Item` [Wu02], serves for function description). The `SemanticAnalyzer` had to be taught to recognize new GIPSY types (see Section 4.1.1.5) with base `GIPSYType` class for object, embed, and array processing, `ImperativeNode` for sequential threads and communication procedures, and a general `AbstractSyntaxTree`.

4.1.1.14 Interfacing GIPC and GEE and Compiled GIPSY Program

Now, let us formally define the notion of a stored compiled GIPSY program, as a GEER or the interface between the two major modules - GIPC and GEE. Until this point, the GEE accepted from GIPC as the input AST of an intensional part and a dictionary of symbols. This suggests having serialized the AST and the dictionary. With the invent of JLucid, communication procedures (CPs) and sequential threads (STs) became relevant and should belong to the GIPC-GEE interface. Thus, a compiled GIPSY program may have several of CPs and STs serialized along. While STs and CPs are present within imperative AST nodes, references to them are recorded here for quicker access and decision making by the GEE. Then, as GEE produces demands (especially over RMI or Jini, [VP05]) for each intensional identifier in the dictionary an Identifier Context (IC) class created [LGP03, Lu04]. This is needed because every such identifier represents a Lucid expression to be evaluated by the engine, and as such should also be part of the compiled GIPSY program. The corresponding class diagram is in Figure 53. It includes the `GIPSYProgram` and all its associations with `GIPC`, `GEE`, `GEERGenerator`, and the storage classes.

To summarize, the GIPC-GEE interface is the `GIPSYProgram` representing encapsulation of the five parts:

1. Linked AST(s)
2. Dictionary
3. A set of STs
4. A set of CPs
5. A set of ICs.

On the diagram in Figure 38 `GIPSYProgram` defines and corresponds to the GEER.

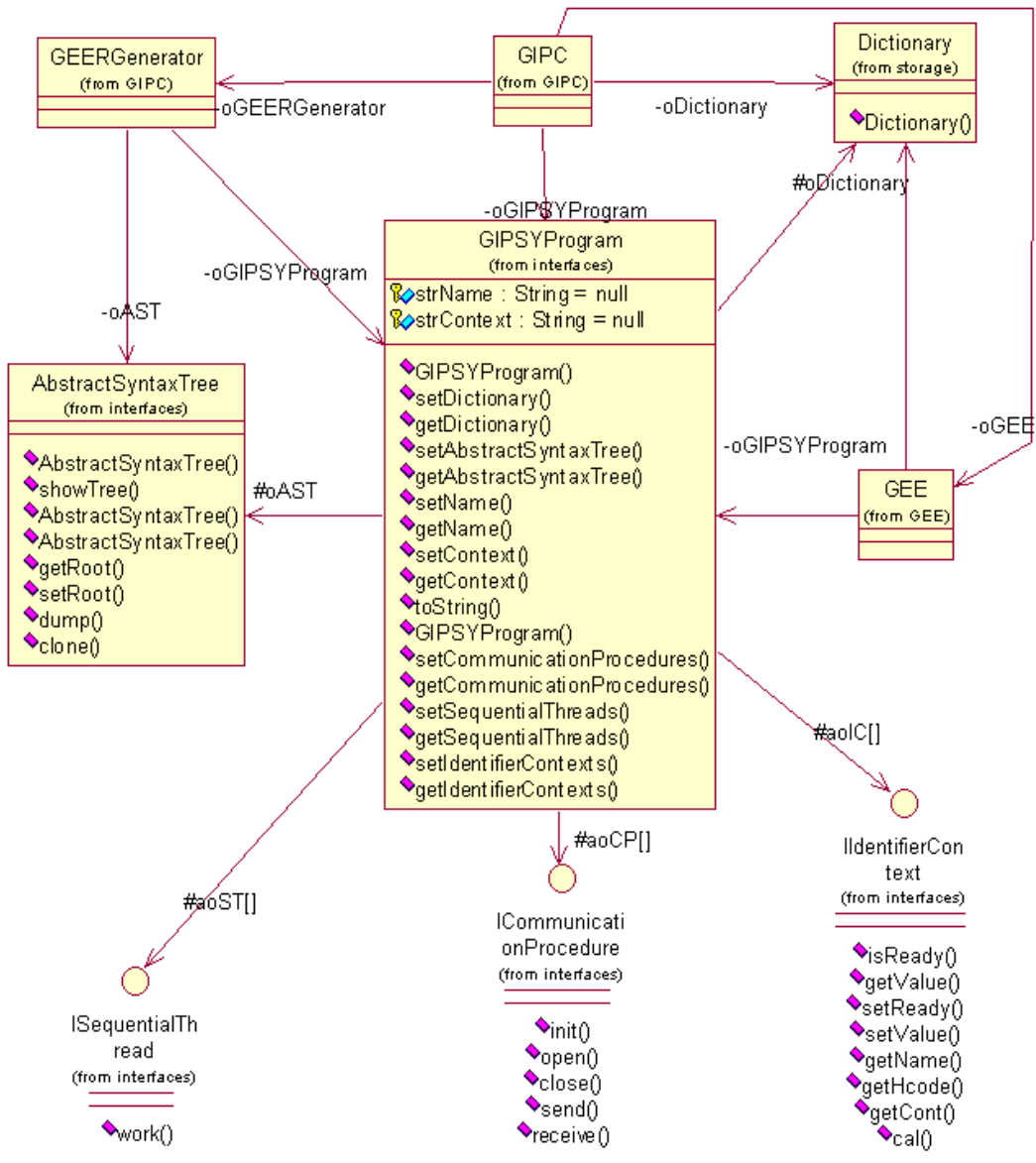


Figure 53: Class diagram describing GIPSYProgram.

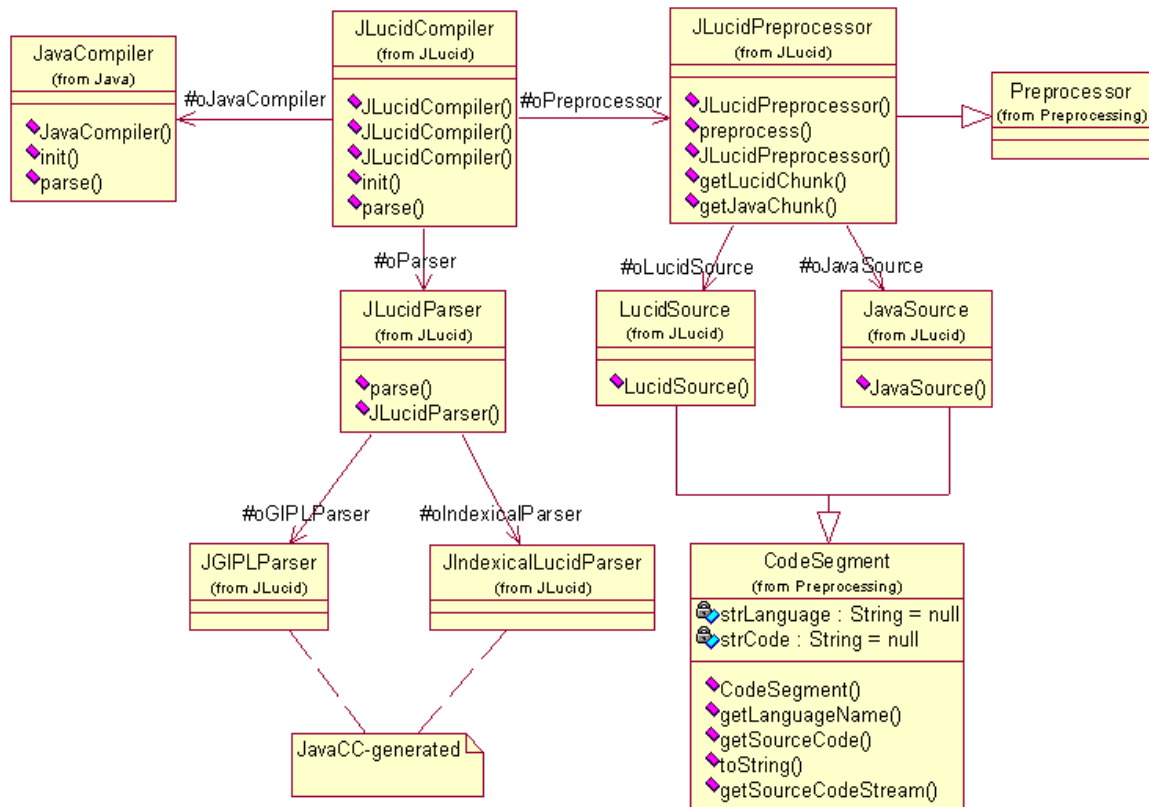


Figure 54: JLucid Design.

4.1.2 JLucid

4.1.2.1 Design

The class diagram describing JLucid is shown in Figure 54. The implementation of JLucid parser-wise is heavily dependent on that of Indexical Lucid as the largest chunk of the IPL work is the same. JLucid adds a preprocessor `JLucidPreprocessor` class that is responsible for parsing initial source JLucid program and extract Java and Lucid parts. The `JLucidParser` class is the one that manipulates `javacc`-generated parsers amended to support `embed()` and arrays. The sequence diagram describing the details of the compilation sequence of JLucid is presented in Figure 55.

JLucid implements generation of Java sequential threads (STs) and their communication procedures (CPs); thus, necessitating `JavaSequentialThreadGenerator` and `JavaCommunicationGenerator`. For uniformity, portability, and testing reasons, we also decided to send the source code over, that can possibly be compiled on the

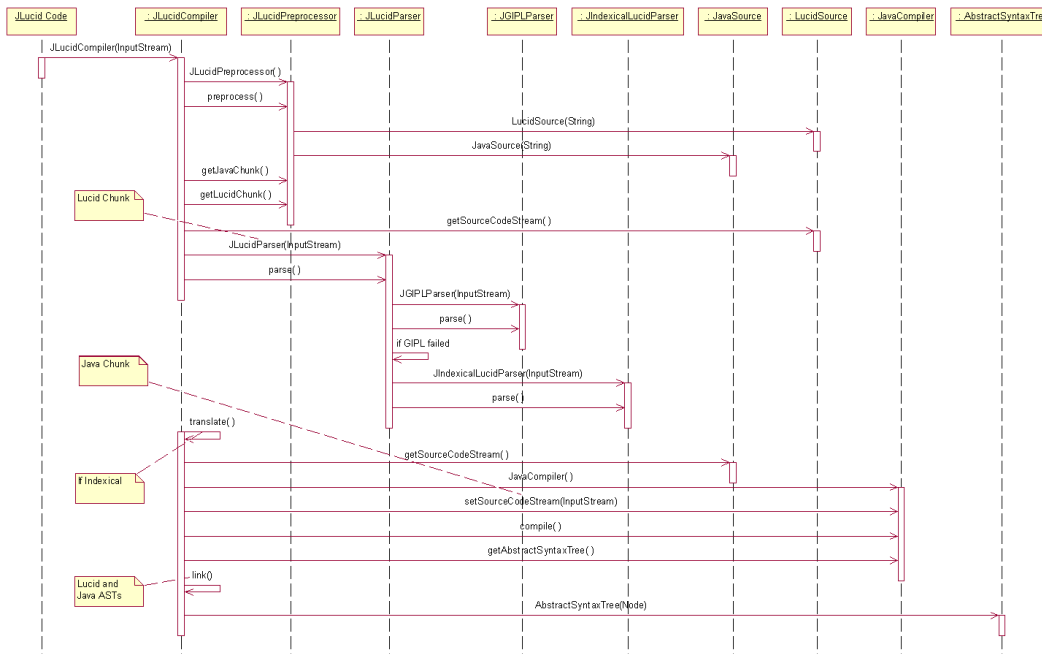


Figure 55: JLucid Compilation Sequence.

remote machine. All this is done by the GICF-integrated `JavaCompiler`, see Section 4.1.2.3.

4.1.2.2 Grammar Generation

As it was shown in Chapter 3, the JLucid syntax extension to GIPL and Indexical Lucid is minimal. The JavaCC grammars we use, are stored in the `.jjt` files for the original two dialects. If we decide to have very similar grammar files for JLucid to support JLucid extensions (arrays and `embed()`), then if the original grammar has a bug, the fix will have to be propagated to all the derived grammars, which will not scale from the maintenance point of view as there will be similar small modifications from Objective Lucid and other dialects. Thus, it was decided to only maintain the original grammars of GIPL and Indexical Lucid and generate the ones for the dialects with the minimal changes, so that each dialect only maintains the part that is relevant to its syntactic extension.

For JLucid three `bash` shell scripts were created to process the original JavaCC grammars of GIPL and Indexical Lucid and generate appropriate extended versions

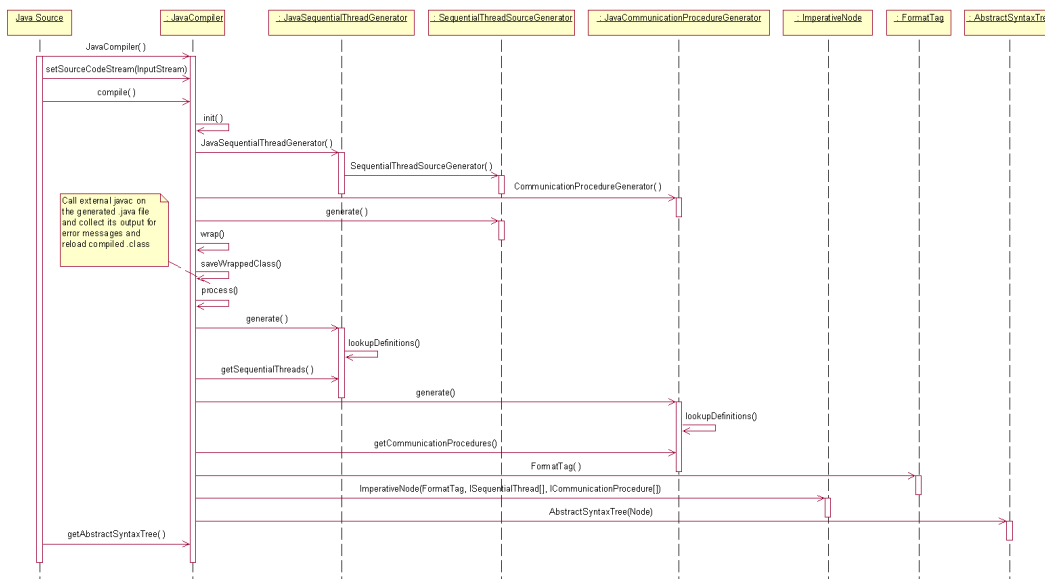


Figure 56: Java Compilation Sequence.

for JLucid. These include `jlucid.sh` that generates JavaCC productions for arrays and `embed()`, `JGIPL.sh` that alters the original `GIPL.jjt` grammar to suit the needs of JLucid mostly in terms of class and package names and the new productions. Similarly, the `JIndexicalLucid.sh` script exists for processing of the `IndexicalLucid.jjt` file. The scripts are rather small and presented in the Appendix D.

4.1.2.3 Free Java Functions and Java Compiler

As defined in Chapter 3, by “free Java functions” we mean is that the corresponding Java STs don’t have an enclosing Java class as far as JLucid source code concerned. However, the enclosing class must exist when compiling a Java program according to Java’s syntax and semantics. Thus, implementation-wise we generate such a class internally that wraps all our sequential threads, as e.g. in Section 4.1.1.8, and we compile that class. This job of wrapping is delegated to the `JavaCompiler`, a member of the imperative compilers framework (see Section 4.1.1.1). The `JLucidCompiler` as shown in Figure 55 at some point invokes the `JavaCompiler`, and what the `JavaCompiler` does internally is illustrated in Figure 56.

Being an imperative compiler, the `JavaCompiler` is obliged to produce the Java

STs and CPs among other things. The core of this process is the `wrap()` method where the actual “wrapping” our pseudo-free Java functions into an internal class occurs. The generated source code `.java` file is saved and is fed to the external `javac` compiler as of this implementation. If there was no compilation errors, a corresponding `.class` or series of `.class` files (for the case of nested classes) is generated. The generated classes are reloaded back by the `JavaCompiler` and their members that are of interest to us retrieved via the Java Reflection Framework [Gre05], thus we obtain an array of references to the ST methods and their parameters and assign them to our own data structures. After this process completes, the corresponding `FormatTag` describing the Java language and the compiler is created and all information is embedded into the `ImperativeNode`, which represents a single and the only node in the imperative `AbstractSyntaxTree`. Later on, this imperative node or its pieces will replace a corresponding stub in the main intensional AST.

4.1.2.4 Arrays

Implementation of arrays in JLucid coincides closely with the implementation of objects in Objective Lucid in Section 4.1.3. As a part of the GIPSY Type System (see Section 4.1.1.5), we employ the `GIPSYArray` (see Figure 42) type to hold the array base type and its members and an overall value. As proposed further, we treat arrays internally as objects (and objects as arrays), so `GIPSYArray` is an extension of `GIPSYObject` that has a base type asserting the data type of the all the elements in the arrays (as our arrays a homogenous collection of elements). Thus, when a syntactic array token is parsed, a corresponding instance of `GIPSYArray` is created to hold the type and value information for later processing. The `SemanticAnalyzer` and the `Executor` are made to understand the array type and apply similar type checking or execution rules to a collection of values instead of a single value.

It might look like this approach will clash with the use of arrays in Java, i.e., when a developer wishes to use Java arrays (or if a library already implements some functionality via Java arrays). This should not be a problem (though will require a more thorough investigation in the future work), when we perform type matching by the base element type, as described in Section 4.1.1.4. The `JavaCompiler` is responsible for the appropriate conversion of the native-to-GIPSY type conversions, by supplying a `TypeMap` such that it can also be used by the GEE at run-time. Similar

comments can be said of the native array types that might exist in other imperative languages that we would be hoping to support.

4.1.2.5 Implementing `embed()`

To implement `embed()` we define a type `GIPSYEmbed` to fetch the file pointed by the URL and hold it in there. In `JLucid`, a `.java` or `.class` file (later also a `.jar` file) is loaded from either local or remote location pointed by the URL as follows: if it is a `.java` file, it's fetched and compiled similarly to the generated class, but the name is static and known; with the `.class` file we skip the compilation process, but extraction of the sequential threads is the same; for the `.jar` its examined with the `JarInputStream` and `JarEntry` Java classes to extract the class information.

4.1.2.6 Abstract Syntax Tree and the Dictionary

When running the `JLucid` compiler in stand-alone mode, all the preprocessing and re-assembling the intensional and imperative pieces into the combined main AST happens in here, not in the `GIPC`, so the `JLucid` compiler returns a complete linked AST with all imperative nodes linked in place and a proper dictionary of identifiers, both intensional and imperative. `JLucid` compiler, however, reused the `Preprocessor` and other parts of the new framework internally instead of re-inventing the wheel.

The `JLucidPreprocessor` uses the general `Preprocessor` class to do the job of chunkanizing the code segments and preparing initial imperative stubs. This necessitated adding the `#funcdecl` segment in the `JLucid` programs that previously did not have one in Chapter 3, to simplify preprocessing and generation of the dictionary. The `JLucidPreprocessor` is set to reject any other code segments than `#JAVA`, `#JLUCID`, or `#funcdecl`.

If the `JLucidCompiler` invoked from the `GIPC` as a part of general compilation process (see Figure 49), the `#JAVA` segment will no longer be really processed internally, and instead, `GIPC` will call `JavaCompiler` externally to the `JLucidCompiler`, so essentially the `JLucidCompiler` will be responsible only for the Lucid part (with arrays and `embed()`).

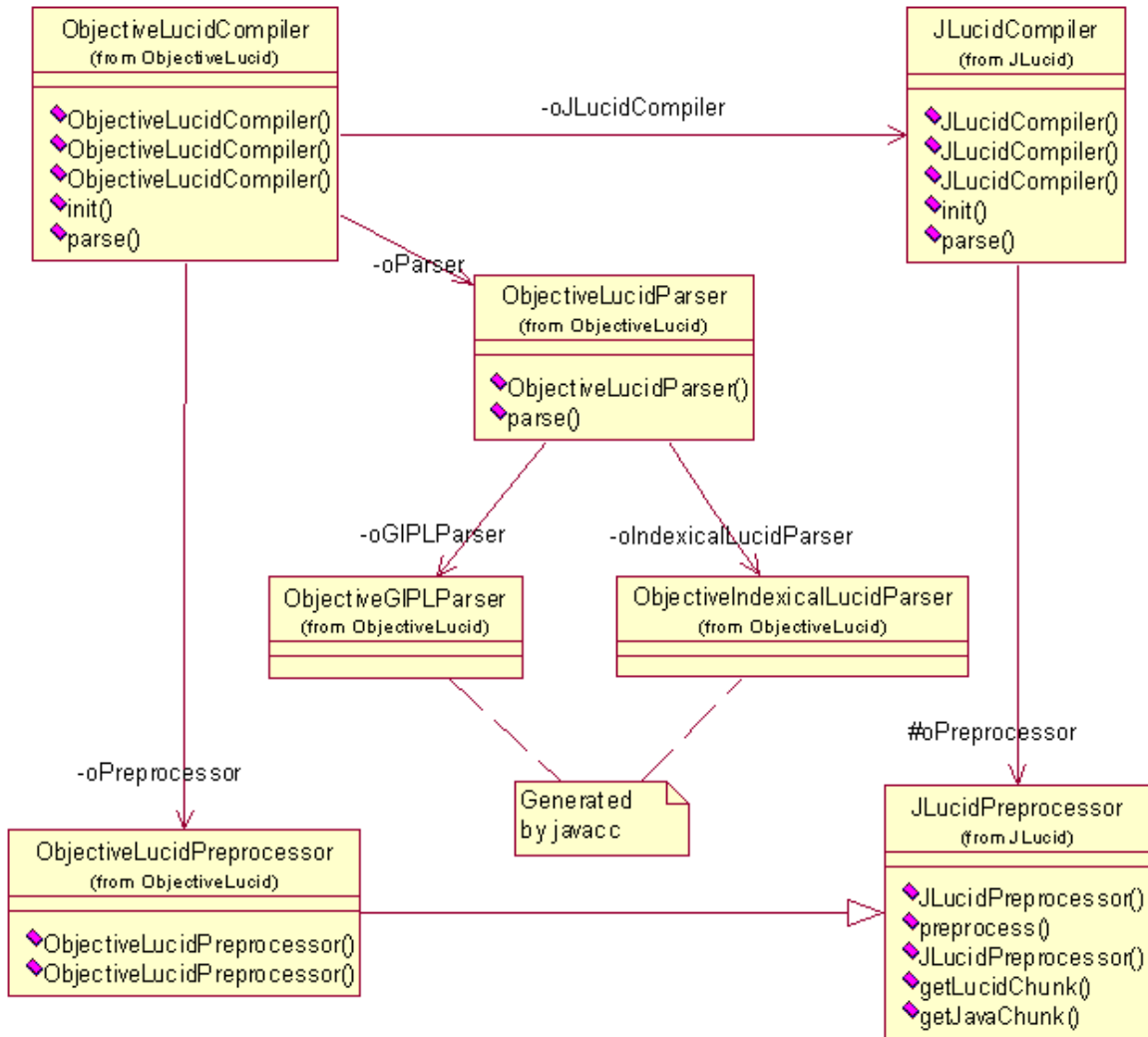


Figure 57: Objective Lucid Design.

4.1.3 Objective Lucid

This section addresses problems that arise when implementing Objective Lucid. These include internal implementation to support the dot-notation, extension to semantic analysis to be able to manipulate object data types (very likely user-defined), and making it all work in the GICF and General Education Engine (GEE) of the GIPSY by correctly forming the abstract syntax tree (AST) that includes object data types.

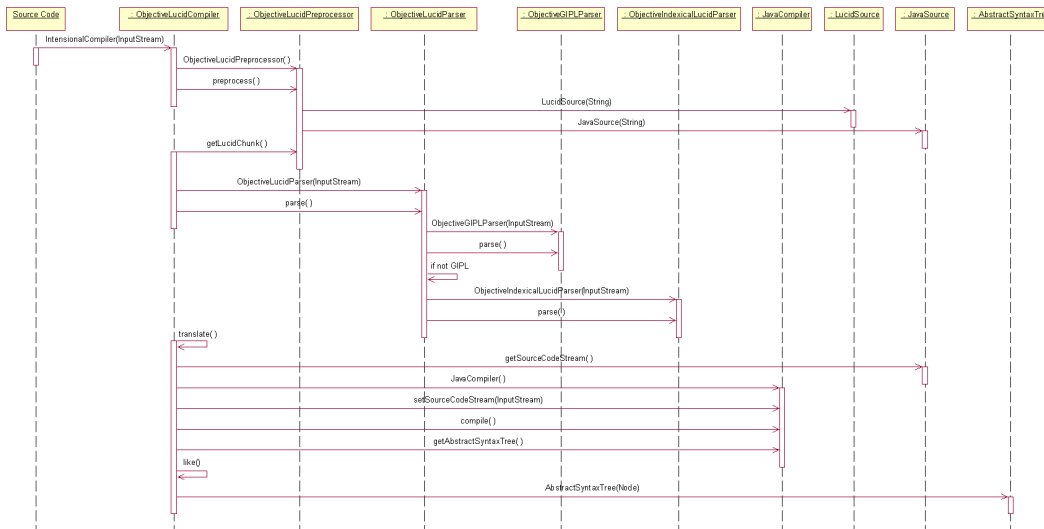


Figure 58: Objective Lucid Compilation Sequence.

4.1.3.1 Design

The class diagram describing Objective Lucid is in Figure 57. Since the JLucid compiler already does most of the legwork, Objective Lucid simply extends it to add the dot-notation and some extra post-processing when unrolling the objects. The corresponding compilation sequence is shown in Figure 58.

4.1.3.2 Grammar Generation

Like with JLucid, the grammar files are generated for Objective Lucid using `bash` shell scripts, `ObjectiveGIPL.sh` and `ObjectiveIndexicalLucid.sh`. These scripts work with the grammars produced by the JLucid scripts (see Section 4.1.2.2) by simply extending them with the dot-notation production and fixing up names of classes and packages. These scripts are presented in the Appendix D.

4.1.3.3 Object Instantiation

Normally, when a Lucid program refers to a Java object, it has to instantiate it first by either calling a pseudo-free Java function that returns an object instance or to call the constructor directly. This instantiation has to be explicit at the beginning of the program to avoid Java's `NullPointerException` at run-time. Internally, the object

instance is created using Java Reflection [Gre05] by first loading and then initializing the needed class with `Class.forName("ClassXB").newInstance()`. Referencing static members do not require a class instance, and can be accessed using the class name, in this case we just keep the `Class.forName("ClassXB")`. We also keep the needed references to the object itself and its members in the `GIPSYObject` type of the GIPSY Type System.

4.1.3.4 The Dot-Notation

Implementing the dot-notation extension of JLucid is the easiest task of the three. In fact, the `E.id` productions are just a syntactic sugar that can be wrapped around already existing mechanisms of JLucid to include Java functions as mentioned in Section 3.2.1.1. The compiler simply generates a set of pseudo-free Java functions for every object member referenced from the intensional program. These will be easy to place into the AST just the way JLucid does it. In other words, this is achieved by automatic generation of implicit accessor Java functions that had to be explicit in JLucid.

4.1.3.5 Abstract Syntax Tree and the Dictionary

The GIPC (General Intensional Programming Compiler) generates abstract syntax trees (AST) of all compiled GIPSY program parts, and constructs the GEER (General Education Engine Resources), which is a data dictionary storing all program identifiers, encapsulated with all ASTs generated at compile time. Simply put, the GEER encapsulates all the meaning of a GIPSY program, and all necessary resources to enable the GEE to execute the programs correctly. The AST and the dictionary contain the generated accessor identifiers that are processed by the JLucid mechanisms, as described previously. This is possible because Java's built-in class `Class` can provide us with all the meta-information about its members through enumeration that we can place in the AST and the dictionary. Little changes from the way JLucid processes that except that the object members are put into the dictionary and acted upon as an array of homogeneous types as described in the follow up section.

The `ObjectiveLucidPreprocessor` also makes use of the general `Preprocessor`, but unlike `JLucidPreprocessor`, it also accepts the `#typeddecl` segment as with objects come user-defined types, so these have to be listed if used by the Lucid part.

4.1.3.6 Objects as Arrays and Arrays as Objects

Implementation-wise, we propose to treat arrays of JLucid as a special case of objects and, the other way around, the objects be a generalization of arrays. An array can be broken into its elements where every element is evaluated as an expression under the same context. Thus, evaluating:

```
A[4] @ [d:4]
where
  dimension d;
  A[#.d] = 42 * #.d fby.d (#.d - 1);
end;
```

is equivalent to evaluating four Indexical Lucid expressions (possibly in parallel). Under this point of view objects can be viewed as arrays where every atomic member is evaluated as if it were an array element. Basically, we denormalize an object into primitives and evaluate them. If an object encapsulates other objects, then these are in turn denormalized and put into the definition environment (dictionary). In other words, if you have an array of four elements `a[4]`, the elements are evaluated as four independent expressions. Likewise, an object that has four data members, each of them is evaluated as an expression under the same context.

Essentially, an array is a collection of atomic elements of the same type. When evaluating say an array of four elements `a[4]` at some context `[d:4]`, we are, in fact, evaluating four ordinary Lucid expressions (possibly in parallel) in the same context. Likewise, an object is a collection of atomic elements of (possibly) different types. In case an object encapsulates another object, that other object can in turn be split into atoms, and so on. All atoms of an object evaluate as independent Lucid expressions, just like array elements.

Thus, from Objective Lucid's point of view, the following are equivalent:

```
(a) int a[4];
```

```
(b) class foo
    {
        int a1;
        int a2;
        int a3;
        int a4;
    }
```

So, internally, we represent (a) in the definition environment as:

```
a_4 // scope identifier
a_4.a1
a_4.a2
a_4.a3
a_4.a4
```

Under the scope of array `a_4` (a generated *id*) there are four members, and `a_4.a*` comprise a denormalized identifier, also generated. And (b) will become:

```
foo // scope identifier
foo.a1
foo.a2
foo.a3
foo.a4
```

where `foo.a*` are generated variable identifiers in the definition environment. Encapsulation will be handled in the following way:

```
class bar
{
    int b1;
    int b2;

    foo oFoo = new foo();
}
```

```
bar
bar.b1
bar.b2
bar.foo
bar.foo.a1
bar.foo.a2
bar.foo.a3
bar.foo.a4
```

To paraphrase and explain in another example, if we have three separate Lucid expressions:

```
// float
a @ [d:2]
where
    dimension d;
    a = 2.5
    fby.d (a + 1);
end;
```

```
// integer
b @ [d:2]
where
    dimension d;
    b = 1
    fby.d (b + 1);
end;
```

```
// ASCII Char
c @ [d:2]
where
    dimension d;
    c = 'a'
    fby.d (c + 1);
end;
```

Now if we group a, b, and c as a class:

```
class foo
{
    float a = 2.5;
    int b = 1;
    char c = 'a';
    public foo() {}
}
```

So when we write:

```
f @ [d:2]
where
    dimension d;
    f = foo() fby.d (f + 1);
end;
```

we mean there start three subexpression evaluations:

```
f.a @ [d:2]
where
    dimension d;
    f.a = foo().a
    fby.d (f.a + 1);
end;
```

```
f.b @ [d:2]
where
    dimension d;
    f.b = foo().b
    fby.d (f.b + 1);
end;
```

```
f.c @ [d:2]
where
    dimension d;
    f.c = foo().c
    fby.d (f.c + 1);
end;
```

We say these are equivalent where the **f** in all expressions refers to the same object's instance (i.e. there are not three objects constructed, only one). Similarly (nearly identically) we implement arrays:

```
a[3] @ [d:2]
where
    dimension d;
    a = [1, 2, 3] fby.d (a + 1);
end;
```

The above means:

```
array a
{
    int a1 = 1;
    int a2 = 2;
    int a3 = 3;
    int length = 3;
}
```

```
a1 @ [d:2]
where
  dimension d;
  a1 = 1 fby.d (a1 + 1);
end;
```

```
a2 @ [d:2]
where
  dimension d;
  a2 = 1 fby.d (a2 + 1);
end;
```

```
a3 @ [d:2]
where
  dimension d;
  a3 = 1 fby.d (a3 + 1);
end;
```

The three subexpressions run in parallel, but refer back to the same array. Should there be a need in one of the three subexpressions to use an array value produced by another subexpression, they generate a demand for that value.

4.2 External Design

The external design encompasses user interface design as well as external software interfaces. In this work, a web interface to the GIPSY as well as command-line interfaces are presented as a part of UI followed by the API of the two external libraries used, JavaCC and MARF.

4.2.1 User Interface

4.2.1.1 WebEditor – A Web Front-End to the GIPSY

The user interface designed for the GIPSY in the scope of this thesis includes a Servlet-driven web interface to the GIPSY daemon server running on our development server for trying out GIPSY programs online. The web interface in a form of a web page allows a connected user to enter, compile, run, and trace GIPSY programs. Users are able to submit their own GIPSY programs (in any supported Lucid dialect) or choose and modify from existing programs from the GIPSY CVS repository (see [RG05a]) and then launch the computation. The GIPSY servlet front-end generates demands through RIPE and returns back results along with an execution trace to a web form. A screenshot of this interface is illustrated in Figure 59.

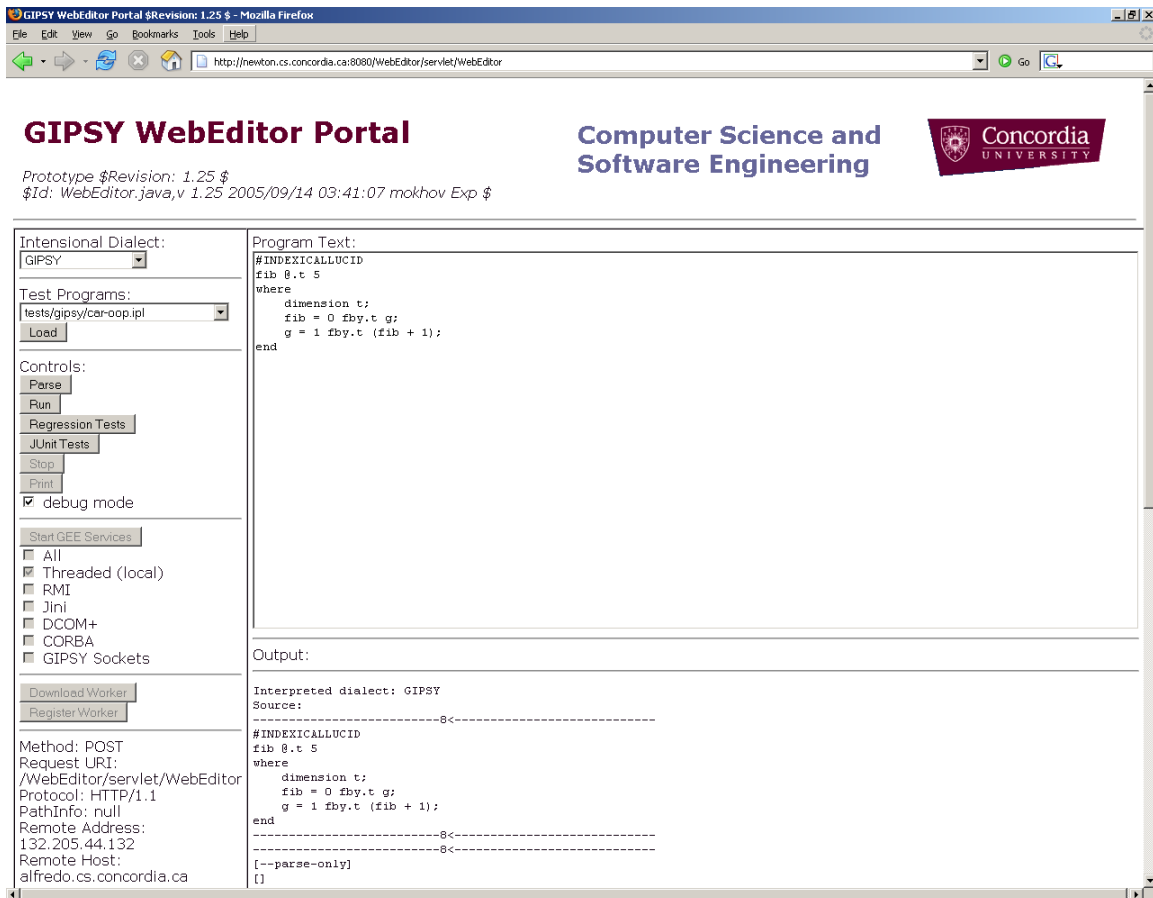


Figure 59: GIPSY WebEditor Interface.

4.2.1.2 GIPSY Command-Line Interface

Synopsis:

```
gipsy [ OPTIONS ]  
gipsy --help | -h
```

This is an all-entry point for all of GIPSY that bundles all the modules. It generally passes all the options to RIPE for further dispatching. When the server part (see Section 7.10) is complete, this will be a GIPSY daemon server. The command line interface includes the following options:

- `--help` or `-h` displays application's usage information.
- `--compile-only` tells to compile a GIPSY program only and return the result of the compilation (error or success messages) and the compiled program itself. This will not invoke the GEE for execution after compilation. The option is primarily for quick tests in development setups.
- `--debug` tells to run in the debug/verbose mode.

It is possible to run the GIPSY by either invoking the `GIPSY.class` directly, by running a corresponding `gipsy.jar` (see Appendix C.2) file, or using a provided wrapper script `gipsy`. The latter is the simplest one to use as it includes all the necessary options for the JVM and searches for the executable `.jar` in several common places. A good idea is to put `gipsy` somewhere under one's `PATH`. (A similar approach applies to the other tools mentioned in the follow up sections, such as `ripe`, `gipc`, `gee`, and `regression`. The tools exist for both Unix and Windows in the form of shell scripts and batch files.)

Example uses of the GIPSY application include:

- `gipsy` or `gipsy --help`
- `gipsy --compile-only`
- `gipsy --compile-only --debug`

Where `--debug` can be combined with any of these, otherwise the options are exclusive.

4.2.1.3 RIPE Command-Line Interface

Synopsis:

```
ripe [ OPTIONS ]
ripe --help | -h
```

The RIPE command-line interface right now acts mostly to activate various own submodules (e.g. textual or DFG editors) or dispatch requests from users to the other main modules, such as GIPC and GEE. The command-line interface includes the following options:

- `--help` or `-h` displays application's usage information.
- `--gipc='<GIPC OPTIONS>'` tells RIPE to invoke GIPC with a set of GIPC options (see Section 4.2.1.4).
- `--gee='<GEE OPTIONS>'` tells RIPE to invoke GEE with a set of GEE options (see Section 4.2.1.5).
- `--regression='<REGRESSION OPTIONS>'` tells RIPE to invoke `Regression` testing with a set of its options (see Section 4.2.1.6).
- `--dfg='<DFG EDITOR OPTIONS>'` tells RIPE to start the DFG editor with its options. Currently, the `DFGEditor` Java class is a stub, and instead, the DFG Editor of Yimin Ding [Din04] is started via a separate program, `lefty`. It is planned the `DFGEditor` class would be a wrapper for the program in the future. Therefore, all DFG editor options are ignored for now, but a provision is made for the future.
- `--txt='<TEXTUAL EDITOR OPTIONS>'` tells RIPE to start the textual editor with its options. Note, at the time of this writing `TextualEditor` is just a stub, and as such does not have any options, but a provision is made when it does.
- `--debug` tells to run in the debug/verbose mode.

Example uses of the RIPE application include:

- `ripe` or `ripe --help`
- `ripe --compile-only`
- `ripe --compile-only --debug`

4.2.1.4 GIPC Command-Line Interface

Synopsis:

```
gipc [ OPTIONS ] [ FILENAME1.ipl [ FILENAME2.ipl ] ... ]
gipc --help | -h
```

The command line interface for GIPC inherited some options from Lucid [Ren02] and includes the following options:

- `--help` or `-h` displays application's usage information.
- `[FILENAME1.ipl [FILENAME2.ipl] ...]` tells GIPC to compile a GIPSY program as indicated by the `FILENAME`. It is possible to have more the one input file for compilation. If this is the case, the same number of instances of GIPC threads will be initially spawned to compile those programs. Notice, however, this does not mean all the files (in case of multiple `.ipl` files) comprise one program and then linked together afterwards as in typical C or C++ compilers. Instead, each `.ipl` file is treated as a stand-alone independent GIPSY program.
- `--stdin` tells GIPC to interpret the standard input as a source GIPSY program. This is the default if no `FILENAME` is supplied.
- `--gipl` or `-G` (came from Lucid [Ren02] for backwards compatibility) tells GIPC to interpret the source program unconditionally as a GIPL program (by default no assumption is made and GIPC attempts to treat the incoming source code as a general GIPSY program). It is primarily used to quickly test the GIPL compiler only, without extra overhead or preprocessing. It is also used by the `Regression` application for that same reason.
- `--indexical` or `-S` (came from Lucid [Ren02]) tells GIPC to interpret the source program unconditionally as an Indexical Lucid program.

- `--jlucid` tells GIPC to interpret the source program unconditionally as a JLucid program.
- `--objective` tells GIPC to interpret the source program unconditionally as an Objective Lucid program.
- `--translate` or `-T` (came from Lucid [Ren02]) enables SIPL-to-GIPL translation. This option is implied by default (as opposed to be optional in Lucid). It tells the GIPC to interpret the input program unconditionally as a non-GIPL program that requires operator and function translation. The option has no effect with `--gipl` as GIPL is the only intensional language that does not require any further translation.
- `--disable-translate` turns off automatic translation (in case the user knows that an incoming non-GIPL program has nothing to translate, which is rarely the case; otherwise, the GIPC will bail out with an error).
- `--warnings-as-errors` tells to treat compilation warnings as errors and stop compilation after displaying them.
- `--gee` tells GIPC to run the compiled program immediately after compilation (if successful) by feeding it directly to the GEE. The default is that the compiled GIPSY program is saved into a file where the original name is suffixed with the `.gipsy` extension.
- `--dfg` tells GIPC to perform DFG code generation as a part of the compilation process.
- `--debug` to run in a debug/verbose mode.

Example uses of the GIPC application include:

- `gipc` or `gipc --help` or `gipc -h`
- `gipc life.ipl`
- `gipc --disable-translate --gee --debug life.ipl`
- `gipc --gipl --debug gipl.ipl`
- `gipc --jlucid --stdin`

4.2.1.5 GEE Command-Line Interface

Synopsis:

```
gee [ OPTIONS ] [ FILENAME1.gipsy [ FILENAME2.gipsy ] ... ]  
gee --help | -h
```

The command line interface includes the following options:

- `--help` or `-h` displays application's usage information.
- `[FILENAME1.gipsy [FILENAME2.gipsy] ...]` tells GEE to run a stored version of a compiled GIPSY program as indicated by the FILENAME. It is possible to have more than one input file for execution. If this is the case, the same number of instances of GEE threads will be initially spawned to run those programs. The programs will run concurrently, but there should not be any interference or communication in their execution except they may share the output resource.
- `--stdin` tells GEE to interpret the standard input as a compiled GIPSY program. This is the default if no FILENAME is supplied.
- `--all` tells GEE to start all implemented services/servers locally (threaded, RMI, Jini, DCOM+, and CORBA).
- `--threaded` tells GEE to start the threaded server only.
- `--rmi` tells GEE to start the RMI service.
- `--jini` tells GEE to start the Jini service.
- `--dcom` tells GEE to start the DCOM+ service.
- `--corba` tells GEE to start the CORBA service.
- `--debug` tells GEE to run in the debug/verbose mode.

Example uses of the GEE application include:

- `gee` or `gee --help` or `gee -h`
- `gee life.gipsy`

- `gee --disable-translate --threaded --debug life.gipsy`
- `gee --all --debug gipl.gipsy`
- `gipc --rmi --jini indexical.gipsy`

4.2.1.6 Regression Testing Application Command-Line Interface

Synopsis:

```
regression [ OPTIONS ]
regression --help | -h
```

The `Regression` application and its test suite are presented in detail in Section 5.1. The application, based on options, invokes either `GIPC` or `GEE` or both directly feeding a pre-selected list of test source programs. The command line interface includes the following options:

- `--help` or `-h` displays application's usage information.
- `--sequential` tells to run sequential tests (default).
- `--parallel` tells to run parallel tests.
- `--gipl` tells to test pure GIPL programs only.
- `--indexical` tells to test pure GIPL and Indexical programs with the Indexical Lucid compiler.
- `--gipsy` tells to test general-style GIPSY programs with code segments.
- `--gee` if specified, tells to run the GEE after compilation (default).
- `--all` tells to do all of the above tests in one run (default).
- `--directory` tells to pick source test files from a specified directory instead of pre-set directories from the GIPSY source tree
- `--debug` tells to run in the debug/verbose mode.

Example uses of the `Regression` application include:

- `regression` or `regression --help` or `regression -h`
- `regression --gipl`
- `regression --parallel --indexical`
- `regression --all --debug`
- `regression --directory=/some/gipsy/misc/tests --all --debug`

4.2.2 External Software Interfaces

4.2.2.1 JavaCC API

JavaCC-generated code contains a number of common classes and interfaces, regardless of the language a parser is generated for. These have to do with AST nodes, tokens, token types, character streams, and alike. The most often used class out of this bundle is `SimpleNode`, which is a concrete node in the AST. These classes have to be periodically refreshed by compiling the source grammar when a newer version of `javacc` comes out.

The below are JavaCC API/modules [VC05] used by the GIPSY and their description. The corresponding class diagram is in Figure 60.

- `Node` is the common interface for all occurrences of `SimpleNode` to implement (see below).
- The `SimpleNode` class represents a concrete node in every AST in the GIPC. Once generated, this class is usually customized according to the needs of the given parser/compiler. All concrete instances, however, implement the same `Node` interface above. At the time of this writing, there are three `SimpleNode` occurrences in the GIPSY source tree: the common one in `gipsy.GIPC.intensional` for all the SIPLs and GIPL, as per original implementation presented in [Ren02]. It is a basis for a GIPL AST aside from the related parsers known to the `SemanticAnalyzer` and GEE's `Executor`. This implementation is wrapped-around by `AbstractSyntaxTree` that the rest of the modules know. Then, a customized subclass of it is in `gipsy.GIPC.DFG.DFGAnalyzer` of Yimin Ding [Din04]. It was made a subclass because a large portion of the code is identical. Finally, the last one is in `gipsy.GIPC.Preprocessing` used by the

Preprocessor. This occurrence of `SimpleNode` was kept as-is due to the significant differences and purpose with the former two.

- The `ImperativeNode` is another implementation of the `Node` interface created manually for all the imperative language compilers. The `ImperativeNode` represents an AST of a single node encapsulating STs, CPs, some meta information that came from a given imperative compiler. The reason for this is to maintain a global AST for a GIPSY program where all nodes implement the same interface.
- `SimpleCharStream` is a common `javacc` utility that treats incoming source code stream as a set of ASCII characters without extra UNICODE processing.
- `ParseException` is a common generated type of exception to indicate a parse error. It was made manually to subclass `GIPCEException` from the GIPSY Exceptions Framework (see Section 4.2.3.2) for uniform exception handling.
- `TokenMgrError` a subclass of `java.lang.Error` primarily to signal lexical errors in the incoming source code or token processing in general by a given parser (e.g. by invoking a static parser twice).

4.2.2.2 MARF Library API

MARF (see Section 2.6.3) has a variety of useful utility and storage-related modules that conveniently found their place in GIPSY. Most of these come from the `marf.util` package as well as `marf.Storage`.² The below are MARF API/modules used by GIPSY and their description:

- `marf.util.FreeVector` is an extension of `java.util.Vector` that allows theoretically vectors of infinite length, so it is possible to set or get an element of the vector beyond its current physical bounds. Getting an element beyond the boundaries returns `null`, as if the object at that index was never set. Setting an element beyond bounds automatically grows the vector to that element. In the GIPSY, `marf.util.FreeVector` is used as a base for our `Dictionary` as shown in Figure 62. Figure 63 shows all the modules that are now using `Dictionary`

²Later some natural language processing (NLP) modules in `marf.nlp` of MARF might also get used in the GIPSY as a part of another research project.

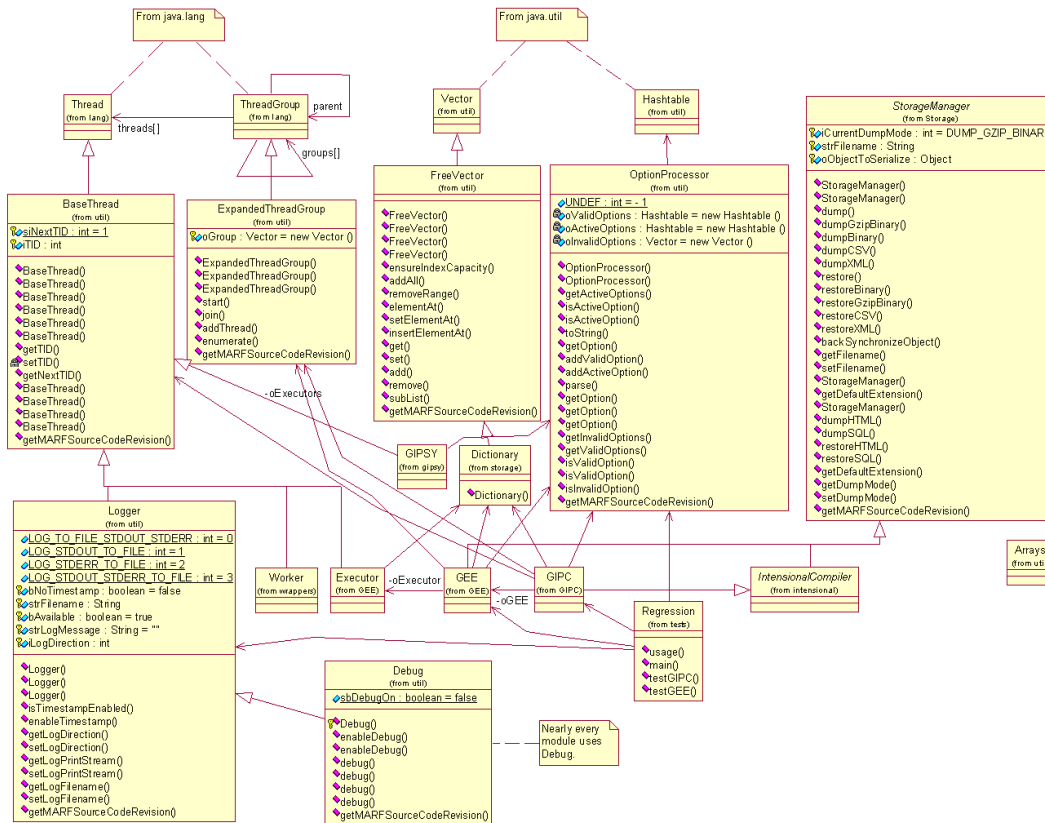


Figure 61: MARF Utility Classes used by the GIPSY.

instead of `java.util.Vector`. The corresponding class diagram of the MARF's util API is shown in Figure 61.

- `marf.util.OptionProcessor` module is extensively used by the command-line user interfaces (see Section 4.2.1) of GIPSY, GIPC, GEE, and Regression. A convenient way of managing command-line options in a hash table and validating them.
- `marf.util.BaseThread` class encapsulates some useful functionality used in threaded versions of GEE and GIPC, which Java's `java.lang.Thread` does not provide:
 - maintaining unique thread ID (TID) among multiple threads and reporting it (for tracing, debugging, and RIPE). A note is added here that Java 1.5.* now also provides a notion of a TID, but `marf.util.BaseThread` was

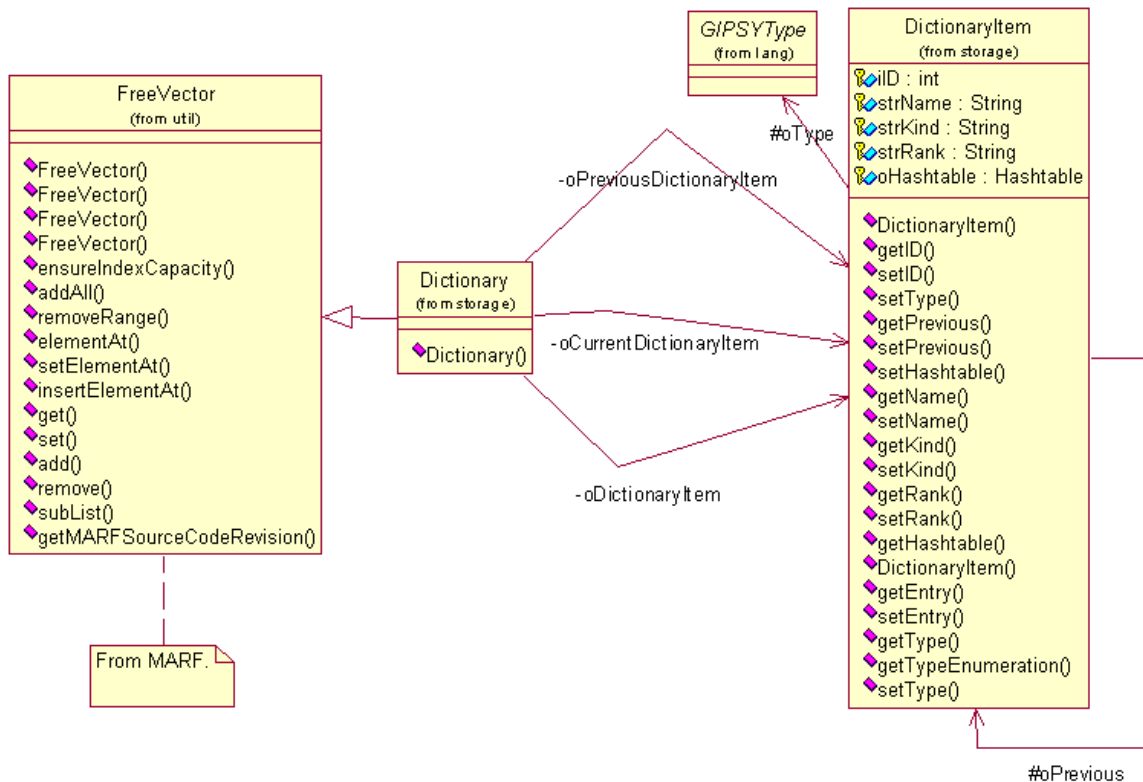


Figure 62: Dictionary and DictionaryItem API

written prior to that and GIPSY remains Java 1.4-compliant still. Plus, MARF’s way of handling this is more flexible.

- adapted human-readable trace information via `toString()`
- access to the `Runnable` target that was specified upon creation.
- integration with `marf.util.ExpandedThreadGroup`, see below.

- `marf.util.ExpandedThreadGroup` allows to start, stop, or other group operations that Java’s `java.lang.ThreadGroup` doesn’t provide. `ExpandedThreadGroup` is, for example, used in GIPC to create a group of compiler threads (in GIPSY every compiler is a thread), one for each language chunk, that will run concurrently. Additionally, a group of GEE, or rather, `Executor` threads would run in the case of a forest of ASTs.
- `marf.util.Arrays` groups more array-related functionality together than the `java.util.Arrays` class does, for example copying (homo- and heterogeneous types) and converting to `java.util.Vector`, and provides some extras.

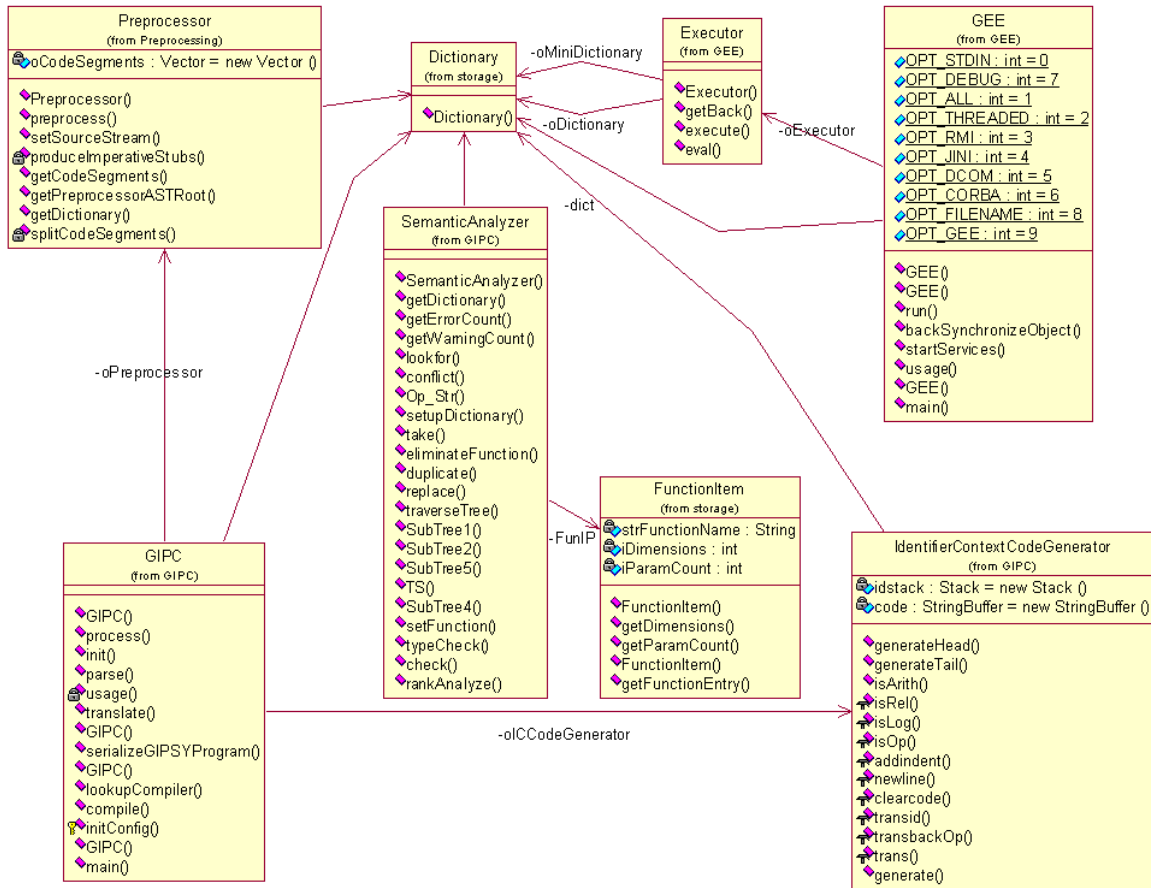


Figure 63: Dictionary Usage within the GIPSY

- `marf.Storage.StorageManager` provides basic implementation of the (possibly compressed) object serialization, and in our case the GIPC and GEE are storage manager with respect to a compiled GIPSY program.
- `marf.util.Logger` is primarily used by the Regression application to log standard output before calling GIPC or GEE to a file, for future comparison with an expected output.
- `marf.util.Debug` is used in many places for debugging convenience allowing to issue debug messages only if the debug mode is globally on, which is also maintained within the class.

4.2.2.3 Servlets API

The Java Servlets technology from Sun [Mic05a] was used to implement the `WebEditor` interface outlined earlier. While the actual API specification of servlets is rather vast, the key used components used here are listed:

- The `HttpServlet` class is the base for all servlets, including `WebEditor`.
- The `doGet()` must be overridden to respond to the GET HTTP requests.
- The `doPost()` must be overridden to respond to the POST HTTP requests. In our implementation, `doPost()` is simply a wrapper around `doGet()`, so both GET and POST requests are handled identically.

4.2.3 Architectural Design and Unit Integration

Unit integration according to the initial design decisions of the GIPSY system and setting up package hierarchy played an important role in the success of this work. A proposed directory structure (see Appendix C.1) and a corresponding breakdown of the Java packages (see Appendix C.1) hierarchy are important to the success of GIPSY, especially for public use. The author of this work inherited the previous GIPSY iteration without any structure or packaging and proposed and restructured the system to what it is now.

4.2.3.1 GIPSY

When integrating several components of a large system and redesigning some of their API, the overall system design has to be considered. In Figure 64 is a high-level view of the main GIPSY modules. These modules can be run as stand-alone Java applications or start each other.

- The `GIPSY` class on the diagram represents a stand-alone server for a client-server type of application, which is capable of spawning `GIPC` and `GEE` upon client's request. The prime goal of it is testing of intensional programs that users can submit online and get the result in case they don't have the development environment set up from where they are working.

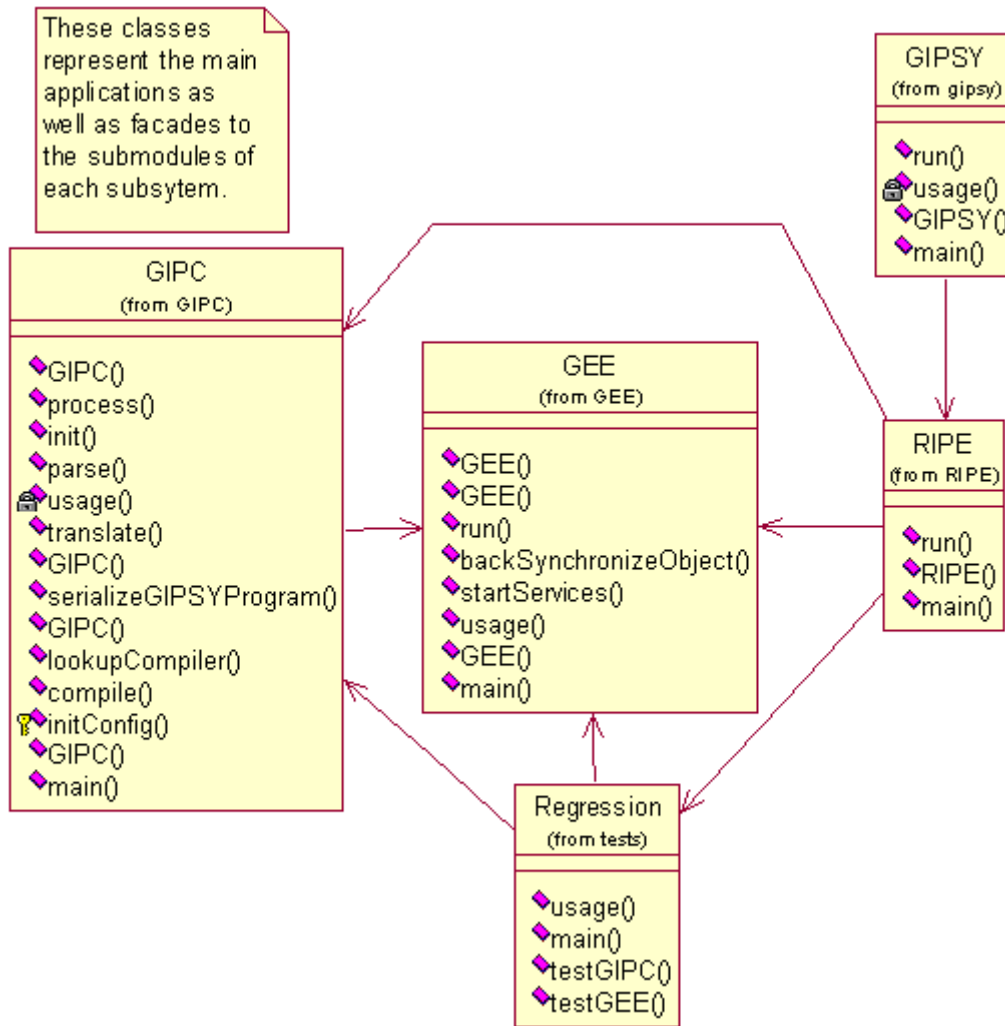


Figure 64: GIPSY Main Modules.

- The **GIPC** class when run as a stand-alone application invokes all the intensional and imperative compilers required and produces a compiled version of a submitted GIPSY program. Optionally, if requested, **GIPC** can pass the compiled program on to **GEE** for execution. The **GIPC** along with **GEE** subsumes what was previously known as **Lucid** and **Facet** defined by Chun Lei Ren in [Ren02].
- The **GEE** when run as a stand-alone application, begins demand-driven execution of a GIPSY program that was either compiled and stored or compiled and passed from **GIPC**.
- The **Regression** class is the main driver for the Regression Testing Suite of

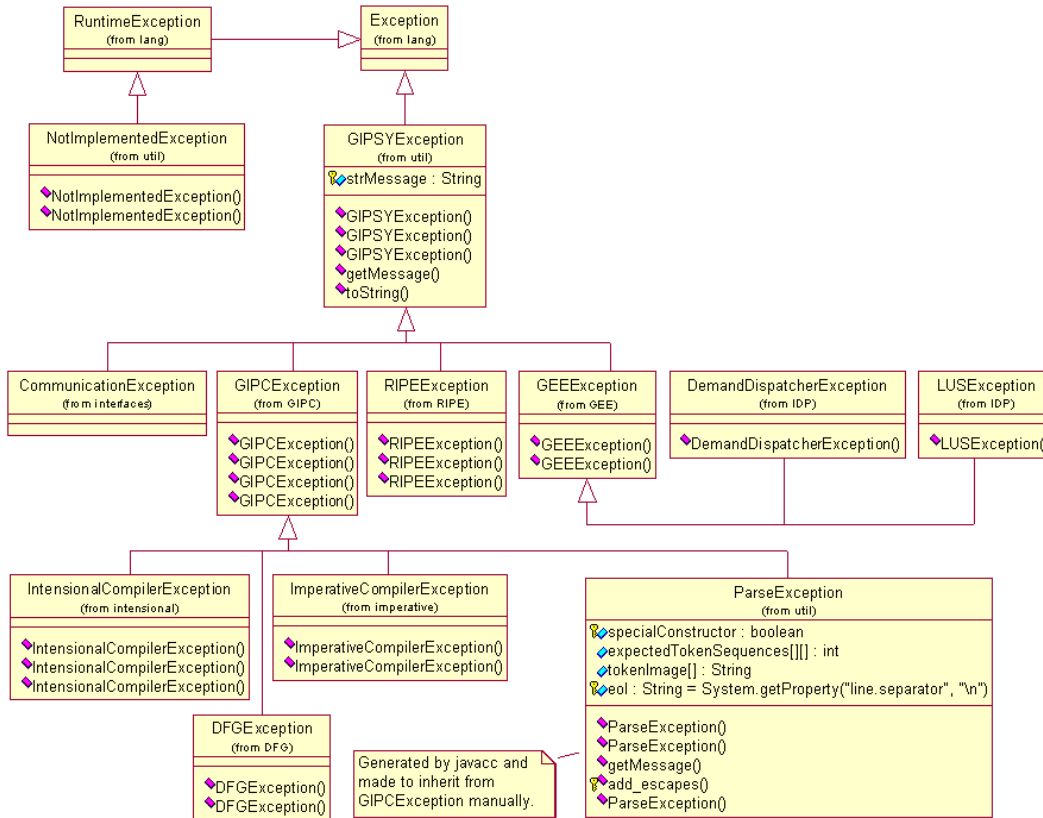


Figure 65: GIPSY Exceptions Framework.

GIPSY, that also calls these modules for regression and unit testing.

4.2.3.2 GIPSY Exceptions Framework

The class diagram describing the GIPSY Exceptions Framework is in Figure 65. The main exception type is `GIPSYException` that provides some machinery encapsulating other exceptions. Every major module, like GIPC, GEE, or RIPE in GIPSY defines its own subclass of `GIPSYException`. By doing this, the applications using the modules can differentiate the exception types and handle them appropriately. The `NotImplementedException` is an easy way to use to indicate some unimplemented but important stubs, if called. It is a subclass of `RuntimeException` because it can happen virtually everywhere and run-time exceptions do not need to be declared to be thrown or caught. The `GIPCEException`, `GEEException`, and `RIPEException` represent base exception objects for the corresponding modules; the rest are primarily

subclasses of these.

4.2.3.3 GEE Design

The general overview of GEE is in Figure 66. The several modules under the `gipsy.GEE` package carry out a complex GIPSY program execution task. The `GEE` is the facade and the main starting point for all of GEE. `GEE` may act as either an application on its own or be invoked by the `GIPC`. For the stand-alone execution a user has to supply a filename of a valid compiled `GIPSYProgram`. This program is loaded and `GEE` starts the `Executor` thread to actually execute it. Before `Executor` begins the `GEE` may optionally start the available demand propagation services, such as local (just threads), RMI, Jini-based and the like. The `Executor` while executing the program generates demands for the identifiers listed in the program and then performs the final calculation based on the results received. The `Executor` was formerly known as `XLucidInterpreter` and the Java version of which was implemented by Bo Lu in [Lu04] and reworked to handle sequential threads, arrays, objects, and other than integer and float data types.

Demand Dispatcher In Figure 67 is a high-level overview of the `DemandGenerator` and related classes. Most of the demand propagation in Jini and JavaSpaces is implemented by Emil Vassev in [VP05]. The integration part included making sure the `IDemandList` interface is consistently used by the `DemandGenerator` along with the `DemandDispatcherAgent` to be compliant to the rest of the GEE. The `IDemandList` interface was originally designed by Bo Lu in [Lu04] and redesigned by the author of this thesis to be implemented by the RMI and threaded versions of GEE and was formerly known as `DemandList`. Next, the temporary class `WorkTask` was made to implement the `ISequentialThread` interface according to the overall GIPSY design for sequential threads. This class is marked as deprecated (and later on will be removed) as every sequential thread class is generated by the `SequentialThreadGenerator` and is different from one GIPSY program to another. Finally, the `LUException` (service look up exception) and `DemandDispatcherException` were made to be a part of the GIPSY Exceptions Framework Section 4.2.3.2 by inheriting from the `GEEException`. For further implementation details of the `DemandDispatcher` please refer to Emil's work [VP05].

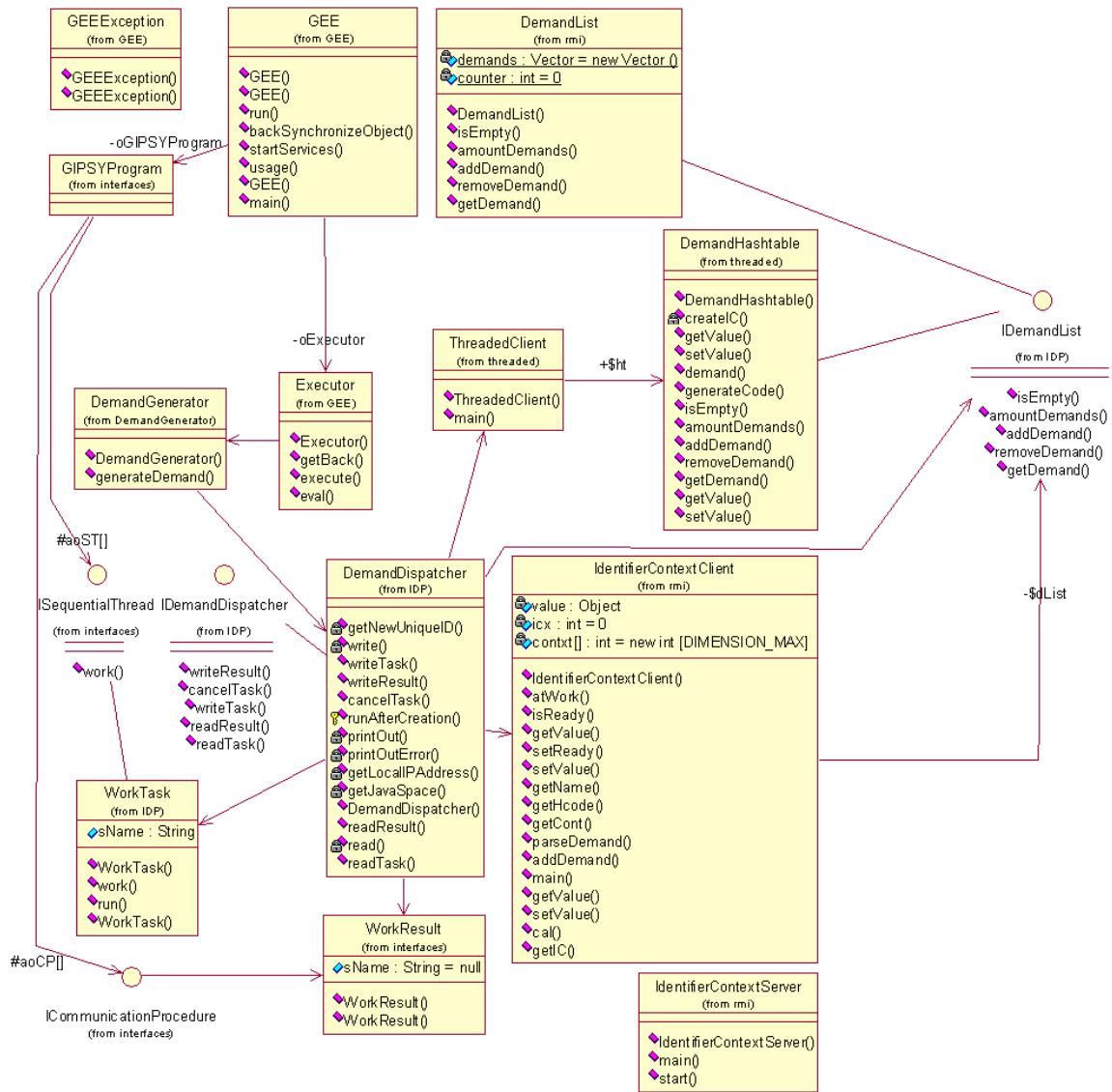


Figure 66: GEE Design.

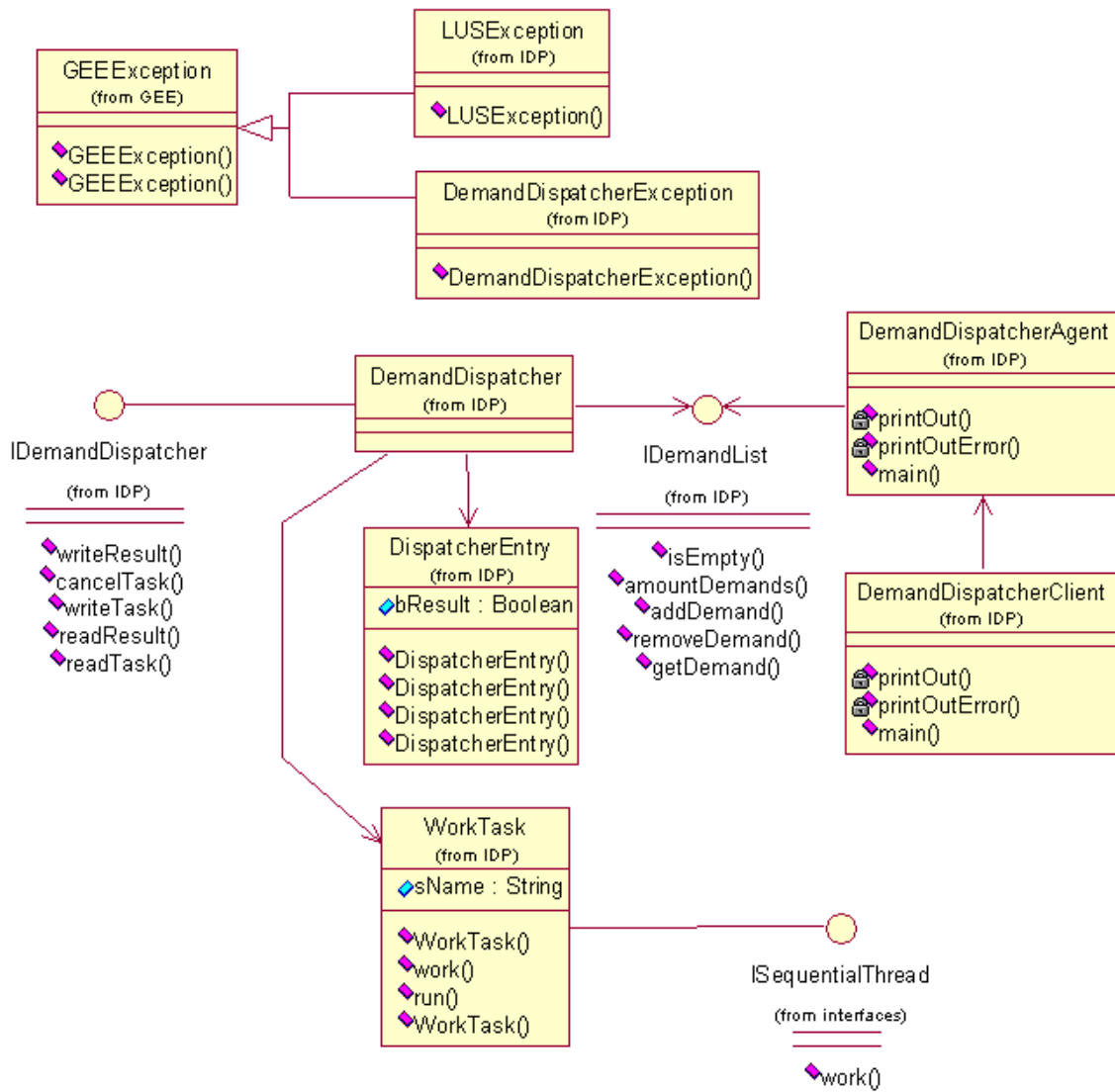


Figure 67: The Demand Dispatcher Integrated and Implemented based on Jini.

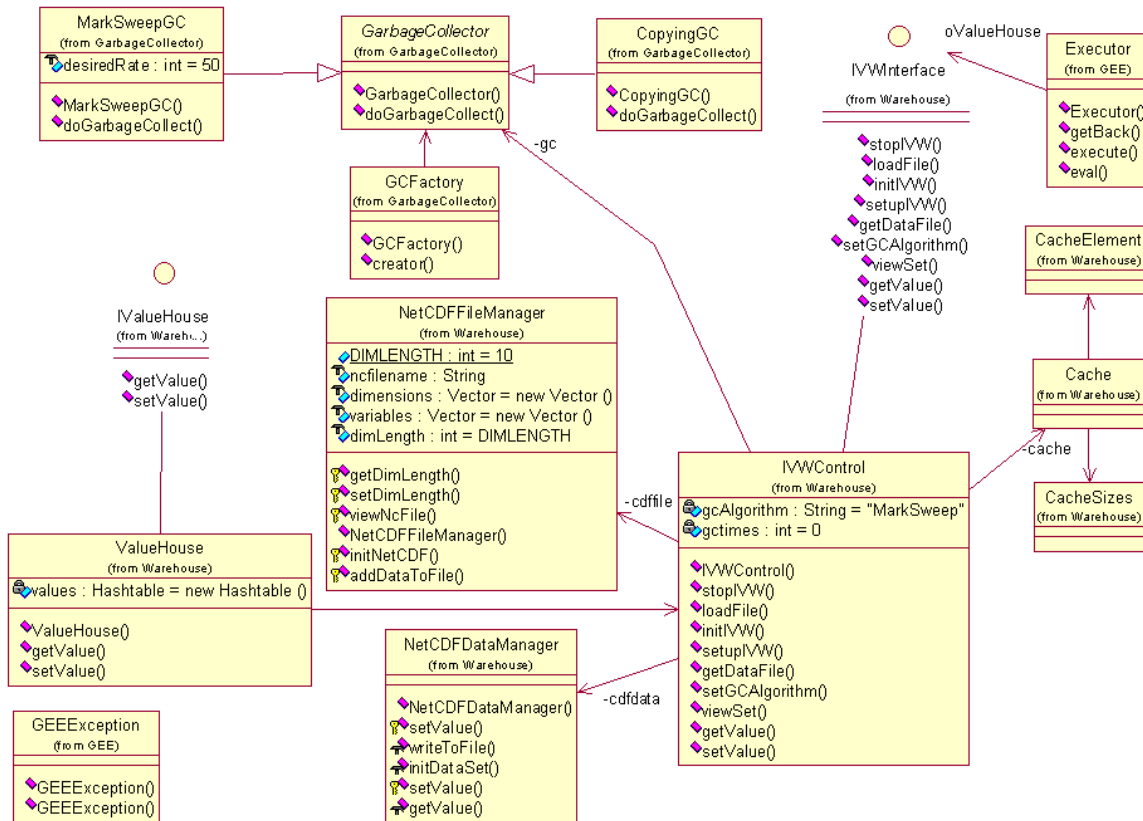


Figure 68: Integration of the Intensional Value Warehouse and Garbage Collection.

Intensional Value Warehouse and Garbage Collection Intensional Value Warehouse and Garbage Collection were implemented by Lei Tao in [Tao04]. After integration, his contributions became to look like as shown in Figure 68. The `IVValueHouse` and its extension `IVWInterface` are the ones used by the `Executor` to communicate to a concrete implementation of a warehouse, allowing adding/changing warehouse implementations easily without affecting the `Executor`. All the exception handling is based on the `GEEException`.

4.2.3.4 RIPE Design

The class diagram describing RIPE is in Figure 69. The `RIPE` class represents a facade to the rest of the RIPE modules. It is semi-implemented, as many things are not clear on this side of the project yet. The only part of RIPE that was advanced well so far by Yimin Ding in [Din04] is the Data-Flow-Graph (DFG) editor, which is not implemented in Java. The `DFGEditor` Java class is meant to be main Java

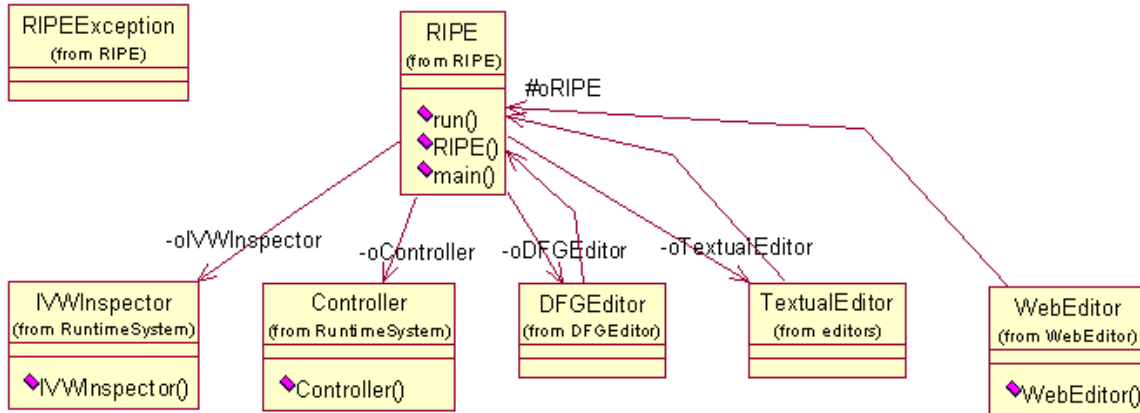


Figure 69: RIPE Design.

program acting like a bridge between Java and the LEFTY language, but did not get implemented yet. The rest of the modules are planned stubs.

4.2.3.5 Data Flow Graphs Integration

The integration of Yimin Ding’s [Din04] DFG-related work is presented in Figure 70. The `DFGAnalyzer` was augmented to implement the `ICompiler` interface as it follows the same structure as the rest of our compilers, which compiles a Lucid code from DFG. The `DFGException` class, a subclass of `GIPCException` has been created to indicate an error situation in the DFG processing. `DFGAnalyzer`’s `SimpleNode` was updated to inherit from `GIPC.intesional.SimpleNode` due to vast functionality overlap. The two analyzer and generator modules have been placed under the `GIPC.DFG.DFGAnalyzer` and `GIPC.DFG.DFGGenerator` packages.

4.3 Summary

This chapter presented most of the development effort went into integration, design, and implementation of JLucid, Objective Lucid, and GICF. User interfaces (both web and command line) has been outlined. Regression Test Suite has been introduced. The follow up chapter presents a variety of testing approaches went into the GIPSY to prove successful integration of the old and implementation of new modules.

To summarize, Objective Lucid, as opposed to GLU [JD96, JDA97] and JLucid, provides access to the object members and is real object-oriented hybrid language. While JLucid may indirectly manipulate objects through pseudo-free functions, the actual objects are still a “black box” to it.

The GICF and IPLCF gave an ability for an easier integration of intensional and imperative languages in the GIPSY. The below are the steps one needs to perform to add a new compiler to the GIPSY:

- create a package where the language compiler will reside (usually under `imperative/LANGUAGE` or `intensional/SIPL/LANGUAGE`).
- add a compiler class that extends either one of `IntensionalCompiler`, `ImperativeCompiler`, or implements one of their superinterfaces
- the code segment and fully qualified class name should be added to either `EImperativeLanguages` or `EIntensionalLanguages`
- optionally implement a custom version of a preprocessor if it is a hybrid language
- implement translation rules to GIPL if it is a SIPL if it is an intensional language
- implement proper ST/CP generation for an imperative language according to that language’s semantics and typing instructions
- implement type mapping table upon the need if it is an imperative language

The above might still sound complex, but it is much more easier and flexible than before. Additionally, some of the steps can be abstracted and simplified, but it is impossible to eliminate manual work altogether.

Chapter 5

Testing

This chapter addresses the testing aspect of this thesis for the following two main reasons: integration and refactoring of a variety of the GIPSY modules including GICF and the development and operation of the two new Lucid dialects developed in this work, namely JLucid and Objective Lucid. Notice, this testing is far from comprehensive and does not include testing of the execution performance of any of the programs and many compilation aspects are still to be resolved as of this writing (and be resolved in the final version). This is, however, a starting point of setting up the GIPSY testing infrastructure for the projects to come to do mandatory systematic tests, which are now a necessity given the size of the system, a centralized source tree, and the number of subprojects developed simultaneously.

5.1 Regression Testing

5.1.1 Introduction

The regression testing is a comprehensive set of tests for the implementation and integration of the GIPSY modules. They test most of the operations and capabilities of the GIPSY. The test cases primarily are various intensional programs (hybrid or not) that exercise the main modules, such as GIPC and GEE as well as their submodules with the major focus on GIPC.

5.1.2 Regression Testing Suite

The regression tests can be run against already pre-compiled `gipsy.jar`, or by using a temporary installation within the source tree using the `Regression` application. Next, there are a “sequential” and “parallel” modes to run the tests. In the sequential mode tests run in strict sequence, whereas in the parallel mode multiple threads are started to run groups of tests in parallel.

5.1.2.1 Unit Testing with JUnit

The core of the `Regression` application is based on the JUnit framework introduced in Section 2.6.1.3. `Regression` represents a `TestSuite`, that contains `ParallelTestCase` and `SequentialTestCase`, a subclasses of `TestCase`. Both types of tests are customizable based on the options supplied to the `Regression` application (see Section 4.2.1.6). JUnit helps to tell us what errors happened and in which modules and the reason of the failures dynamically at run-time.

5.1.2.2 Unit Testing with diff

It becomes cumbersome to use JUnit for all possible cases, in a large system, where often we are generally interested in the output behaviour changes only. Here the UNIX tool `diff` helps us. A collection of hand-checked outputs are said to be “expected”, one ore more file for each test case. Then, when the next time the test is run, a current directory is created with the current outputs, and the current and expected output directories are compared with the `diff` to show the differences in the output produced by the modules. This is all achieved by the `regression` script.

5.1.2.3 Tests

The actual test cases in the form of GIPL, Indexical Lucid, Objective Lucid, JLucid, and GIPSY programs, are located under the corresponding `src/tests/*` directories in the source tree in the form of `*.ipl` files. These comprise most of the examples presented earlier in this work as well as developed in [Paq99], [Ren02], [Wu02], and [Lu04]. The regression tests for the DFG generation ([Din04]), Intensional Value Warehouse and Garbage Collector [Tao04] and Demand Migration System (DMS) [VP05] are not present as of this implementation.

5.2 Portability Testing

GIPSY has been tested and is known as expected (regression tests pass) to run on RED HAT LINUX 9, FEDORA CORE 2, MAC OS X, SOLARIS 9, WINDOWS 98SE/2000/XP systems under JDK 1.4 and 1.5. The corresponding hardware architectures were Intel or Intel-compatible processors (Pentium II, III, and IV with 233 MHz to 1.4 GHz) and G3 and G4 processors from Apple and IBM. For the WebEditor interface, Tomcat 5 on MAC OS X were tested, but it is believed to run on other platforms the Jakarta Tomcat runs on.

5.3 Solving Problems

This section is targeting some common problems of synchronization in parallel and distributed environment and how they are solved using the GIPSY system relieving the programmer from the need of explicitly synchronize the objects. They also illustrate the use of arrays and embedded Java, and Java objects. These programs are among many other test cases from the Regression Tests Suite.

5.3.1 Prefix Sum

```
pseudocode (for thread 'j')

'shared' a 'future' 'int' 'array' [1..logP, 1..P] := undefined;
'private' sum 'int' := j,
        hop 'int' := 1;

'do' level = 1, logP --->
    'if' j <= P - hop ---> a[level, j] := sum      'fi'
    'if' j >      hop ---> sum += a[level, j - hop] 'fi'
        hop := 2 * hop
'od'
```

Figure 71: Pseudocode of a thread j for the Prefix Sum Problem.

```
/*
 * PREFIX SUM in GIPL-style JLucid program.
 * Numbers are from 1 to 8.
 * S[I] will contain prefix sum for number 'i'
 */

#JLUCID

// Array of prefix sums
S[8] @d 8
where
    dimension d;

    S[I] = if(#d = 0)
            then 1
            else (2 * S[I] - 1) @d (#d - 1)
            fi;

// Index the array varies within.
I @i 8
where
    dimension i;
    I = if(#i = 0) 1 else (I - 1) @d (#i - 1);
end;
end;
```

Figure 72: The Prefix Sum Problem in JLucid in GIPL Style.

The pseudocode of for a thread j is in Figure 71 [Pro03a]. The Figure 72 shows the program translated into Lucid. The Figure 73 shows the program translated into Indexical Lucid for numbers from 1 to 8. Below is an equivalent implementation of the problem (targeting only TLP) in Java; compare the program's line count and complexity to that of JLucid:

```
// Modified from Dr. Probst's Cyclic.java
public class PrefixSum
{
    public static final int P      = 8; // number of workers
    public static final int logP  = 3; // number of rows in logP x P matrix

    // For permutation of workers
    private static int[] col = {3, 6, 5, 7, 4, 2, 1, 0};

    // These two mimic a 2D array of future variables
    public static int[][] a      = new int      [logP][P];
    public static Semaphore[][] futures = new Semaphore[logP][P];

    // The resulting sums are to be placed here.
    public static int[] sums    = new int[P];

    public static void main(String[] argv)
    {
        Worker w[] = new Worker[P];

        for(int j = 0; j < futures.length; j++ )
            for(int k = 0; k < futures[j].length; k++)
                futures[j][k] = new Semaphore(0);

        for(int j = 0; j < P; j++)
        {
            w[col[j]] = new Worker(col[j] + 1);
            w[col[j]].start();
        }

        for(int j = 0; j < P; j++)
        {
            try
            {
                w[j].join();
            }
            catch(InterruptedException e)
            {
            }
        }

        for(int j = 0; j < P; j++)
            System.out.println ("Prefix Sum of " + (j + 1) + " = " + sums[j]);
    }
}
```

```

        System.out.println ("System terminates normally.");
    }
}

class Semaphore
{
    private int value;

    Semaphore(int value1)
    {
        value = value1;
    }

    public synchronized void Wait()
    {
        try
        {
            while(value <= 0)
            {
                wait();
            }

            value--;
        }
        catch (InterruptedException e)
        {
        }
    }

    public synchronized void Signal()
    {
        ++value;
        notify();
    }
}

class Worker extends Thread
{
    private int j;
    private int sum;
    private int hop = 1;

    public Worker(int col)
    {
        sum = j = col;
    }

    public void run()
    {
        System.out.println("Worker " + j + " begins execution.");
    }
}

```

```

yield();

for(int level = 0; level < PrefixSum.logP; level++)
{
    if(j <= PrefixSum.P - hop)
    {
        System.out.println
        (
            "Worker " + j +
            " defines a[" + level + "," + (j-1) +"]."
        );

        PrefixSum.a[level][j - 1] = sum;
        PrefixSum.futures[level][j - 1].Signal();
    }

    if(j > hop)
    {
        PrefixSum.futures[level][j - 1 - hop].Wait();

        System.out.println
        (
            "Worker " + j +
            " uses a[" + level + "," + (j - 1 - hop) +"]."
        );

        sum += PrefixSum.a[level][j - 1 - hop];
    }

    hop = 2 * hop;
}

PrefixSum.sums[j - 1] = sum;
System.out.println ("Worker " + j + " terminates.");
}
}

```

```

/*
 * PREFIX SUM in Indexical Lucid-style JLucid
 */

#JLUCID

S[8] @d 8
where
  dimension d;

  S[I] = 1 fby.d (2 * S[I] - 1);

  I @i 8
  where
    dimension i;
    I = 1 fby.i (I - 1);
  end;
end;

```

Figure 73: The Prefix Sum Problem in JLucid in Indexical Lucid Style.

5.3.2 Dining Philosophers

Below is a JLucid implementation of the Dining Philosophers problem [Dij65, Dij71, Gin90]. We have arrays of 8 philosophers and 8 forks, each represented as integers. A philosopher is either thinking (1) or eating (2); likewise for forks, taken or not. A philosopher may eat when they have exactly two forks, not less, if the forks are available. If none available, the philosopher waits (implicit, guaranteed by the GEE). The special variable *I* serves as an intensional index for our arrays.

```
/**
 * Dining Philosophers Problem
 * in JLucid
 *
 * @author Serguei Mokhov, mokhov@cs.concordia.ca
 * @version $Revision: 1.10 $ $Date: 2005/03/02 02:57:31 $
 */

#funcdecl

int getIninitalRandomState();
boolean chew(int);
boolean brainstormIdea(int);

#JLUCID

/*
 * Assume 8 philosophers and two states.
 * States: 2 - eating, 1 - thinking
 * Forks are either available or not; hence, 2 states as well.
 */
PHILOSOPHERS[8] @states 2
where
  dimension states;

  // Initialize all forks
  FORKS[8] @availability 2
  where
    dimension availability;

    FORKS[I] = getIninitalRandomState();

  I @d 8
  where
    dimension d;
    I = 1 fby.d (I - 1);
  end;
end;
```

```

/*
 * Run the actual algorithm.
 * NOTE: in this implementation the computation
 * never terminates (normally). It is an infinite loop.
 */
PHILOSOPHERS[I] =
  if(#states == 1) then
    eat(I) @states 2

    eat(I) =
      getForks(I) && chew(I);

    getForks(I) = g(l, r)
    where
      l = FORK[I] @availability 1;
      r = FORK[I] @availability 1;
    end;
  else
    think(I) @states 1

    think(I) =
      putForks(I) && brainstormIdea(I);

    putForks(I) = p(l, r)
    where
      l = FORK[I] @availability 2;
      r = FORK[I] @availability 2;
    end;
  fi;

I @d 8
where
  dimension d;
  I = 1 fby.d (I - 1);
end;
end;

#JAVA

int getInitalRandomState()
{
  // Either 1 or 2
  return new Random().nextInt(2) + 1;
}

boolean chew(int i)
{
  try
  {
    System.out.println("Philo " + i + " is chewing smth tasty now.");
  }
}

```



```
        sleep(new Random().nextInt(i * 1200));
        System.out.println("Philo " + i + " finished chewing.");
        return true;
    }
    catch(InterruptedException e)
    {
        return false;
    }
}

boolean brainstormIdea(int i)
{
    try
    {
        System.out.println("Philo " + i + " is heavily thinking now.");
        sleep(new Random().nextInt(i * 1200));
        System.out.println("Philo " + i + " finished thinking.");
        return true;
    }
    catch(InterruptedException e)
    {
        return false;
    }
}
```

5.3.3 Fast Fourier Transform

This is an example on how one would compute Fast Fourier Transform (FFT) in the GIPSY for an array of double values. This is straightforward in Lucid because it's deterministic with plenty of parallelism. A JLucid program implementing FFT is in Section 5.3.3.1. The algorithm is based on the Java algorithm implemented in MARF [MCSN05, Pre93, Ber05], a code fragment of which is in Section 5.3.3.2, originally written by Stephen Sinclair. The JLucid version omits the imaginary part of the transform, but it would not be hard to add it.

5.3.3.1 Fast Fourier Transform in JLucid.

```
/*
 * FFT implementation in JLucid.
 * Serguei Mokhov
 * $Id: fft.ipl,v 1.2 2005/08/13 01:37:23 mokhov Exp $
 */

#funcdecl
double sin(double);
double pi();

#JAVA

double sin(double pdValue)
{
    return Math.sin(pdValue);
}

double pi()
{
    return Math.PI;
}

#JLUCID

A
where
    // A is an array of 9 FFT values with a
    // normal FFT applied to the array below.

    A = fft([1, 2, 3, 4, 6, 7, 8, 9], 9, 1);

    fft(inputValues, length, sign) = fftValues
        where
            fftValues = apply(length, reverse(length, inputValues), sign);
```

```

apply(len, coeffs, direction) = coeffs @.s (N - 1)
  where
    dimension s;

    N = 2 * len;
    mmax = (2 fby.s istep) upon(mmax < N);

    coeffs[J / 2] = coeffs[I / 2] - tempr;
    coeffs[I / 2] = coeffs[I / 2] + tempr;

  where
    istep = mmax fby.s (istep) * 2;

    M @.m mmax
  where
    dimension m;

    M = (0 fby.m (M + 2)) upon (M < mmax);

    tempr = wr * coeffs[J / 2] - wi * coeffs[J / 2];

    J = I + mmax;

    wr = 1.0 fby.m ((wtemp = wr) * wpr - wi * wpi + wr);
    wi = 0.0 fby.m (wi * wpr + wtemp * wpi + wi);

  where
    dimension i;
    I = (M fby.i (I + istep)) upon (I < N);
    theta = (direction * 2 * pi()) / mmax;
    wtemp = sin(0.5 * theta);
    wpr = -2.0 * wtemp * wtemp;
    wpi = sin(theta);
  end;
end;
end;
end;

// Binary reversion
reverse(l, vals) = out @.i length
  where
    dimension i;
    out[t] = vals[#.i] @ (#.i + 1) @.bit maxbits(length);
  where
    dimension bit;

    t = 0 fby.bit ((t * 2) | (n & 1));
    n = #i fby.bit (n / 2);
  end;
end;

```

```

// Determine max number of bits
maxbits(len) = (mbits - 1) @.m 16
  where
    dimension m;

    mbits = ( 0 fby.m (mbits + 1) ) upon (mbits < 16 && n != 0);
    n = len fby.m (n / 2);
  end;
end;
end;

// EOF

```

5.3.3.2 Fast Fourier Transform code fragment in Java from MARF.

```

...
/**
 * <p>FFT algorithm, translated from "Numerical Recipes in C++" that
 * implements the Fast Fourier Transform, which performs a discrete Fourier transform
 * in  $O(n \log(n))$ .</p>
 *
 * @param InputReal InputReal is real part of input array
 * @param InputImag InputImag is imaginary part of input array
 * @param OutputReal OutputReal is real part of output array
 * @param OutputImag OutputImag is imaginary part of output array
 * @param direction Direction is 1 for normal FFT, -1 for inverse FFT
 * @throws MathException if the sizes or direction are wrong
 */
public static final void doFFT
(
    final double[] InputReal,
    double[] InputImag,
    double[] OutputReal,
    double[] OutputImag,
    int direction
)
throws MathException
{
    // Ensure input length is a power of two
    int length = InputReal.length;

    if((length < 1) | ((length & (length - 1)) != 0))
        throw new MathException("Length of input (" + length + ") is not a power of 2.");

    if((direction != 1) && (direction != -1))
        throw new MathException("Bad direction specified. Should be 1 or -1.");

    if(OutputReal.length < InputReal.length)
        throw new MathException("Output length (" + OutputReal.length + ") < Input length (" + InputReal.length + ")");
}

```

```

// Determine max number of bits
int maxbits, n = length;

for(maxbits = 0; maxbits < 16; maxbits++)
{
    if(n == 0) break;
    n /= 2;
}

maxbits -= 1;

// Binary reversion & interlace result real/imaginary
int i, t, bit;

for(i = 0; i < length; i++)
{
    t = 0;
    n = i;

    for(bit = 0; bit < maxbits; bit++)
    {
        t = (t * 2) | (n & 1);
        n /= 2;
    }

    OutputReal[t] = InputReal[i];
    OutputImag[t] = InputImag[i];
}

// put it all back together (Danielson-Lanczos butterfly)
int mmax = 2, istep, j, m; // counters
double theta, wtemp, wpr, wr, wpi, wi, tempr, tempi; // trigonometric recurrences

n = length * 2;

while(mmax < n)
{
    istep = mmax * 2;
    theta = (direction * 2 * Math.PI) / mmax;
    wtemp = Math.sin(0.5 * theta);
    wpr = -2.0 * wtemp * wtemp;
    wpi = Math.sin(theta);
    wr = 1.0;
    wi = 0.0;

    for(m = 0; m < mmax; m += 2)
    {
        for(i = m; i < n; i += istep)
        {
            j = i + mmax;
            tempr = wr * OutputReal[j / 2] - wi * OutputImag[j / 2];

```

```
    tempi = wr * OutputImag[j / 2] + wi * OutputReal[j / 2];

    OutputReal[j / 2] = OutputReal[i / 2] - tempr;
    OutputImag[j / 2] = OutputImag[i / 2] - tempi;

    OutputReal[i / 2] += tempr;
    OutputImag[i / 2] += tempi;
}

wr = (wtemp = wr) * wpr - wi * wpi + wr;
wi = wi * wpr + wtemp * wpi + wi;
}

mmax = istep;
}
...

```

5.3.4 Moving Car

A less contrived example of an Objective Lucid program is presented in Figure 74. This is an example where a `Car` object changes with time. Eliminating S , and ignoring the print call, we have have:

```
C @.time 15 where
  C = Car() fby.time C.move(#.time)
```

Using the definition of `fby` gives:

```
C @.time 15
  = (Car() fby.time C.move(#.time)) @.time 15
  = if 15 <= 0 then Car() else (C.move(#.time)) @.time (15 - 1)
  = C.move(14)
```

Our intention is that `fby` will give the sequence:

```
Car() Car.move(1) Car.move(2) ... Car.move(15)
```

This will work as follows. When one generates a demand for `C.move(15)` it's not satisfied until `C.move(14)` is until `C.move(13)` is ... until `C.move(1)` is until `Car()`, so it recurses back and finally the `Car()` object instance gets constructed, and then the demands flow from 1 to 15 and the instance already exists.

The car also does not accelerate indefinitely. It moves until it has enough fuel, else it returns the car object with its members unmodified. The drop of speed is also in place when fuel is depleted.

To further illustrate this idea let's take the existing example of a simpler problem of natural numbers presented in Figure 22 and convert it into Objective Lucid as in Figure 76. First, we will present the eduction tree of the natural numbers problem (see Figure 75, a corrected version of the one produced by Paquet in [Paq99]) and then transmute it into the eduction tree of the execution of the equivalent Objective Lucid program, as shown in Figure 77. The program in Figure 76 exhibits the same properties as the `Car` example, so the eduction tree will be similar but will take more space. The important aspect here is to illustrate the difference between demands for STs and their lazy execution (which is italisized, e.g. `N.inc()`); thus, the actual invocation of a ST method happens at a later time after the demand is made so we avoid not having called constructor prior execution of an instance method. In the

```

#typedecl
Car;

#JAVA
public class Car
{
    public int x = 0;

    public float speed;
    public float speeddrop;
    public float fuel;
    public float fueldrainrate;

    public Car()
    {
        // Assume initially car was already moving.
        speed = 100.0; fuel = 40.5;
        fueldrainrate = 0.018; speeddrop = 0.1;
    }

    // Move by a number of steps assuming constant speed
    // and decelerate when ran out of fuel.
    public Car move(int steps)
    {
        if(fuel > 0)
        {
            fuel -= fueldrainrate * speed * steps;
            x += steps;
        }
        else if(speed > 0)
        {
            x += steps;
            speed -= speeddrop * steps;
        }

        return this;
    }

    public void printCarState()
    {
        System.out.println
        (
            "Speed: " + speed + ", fuel: " + fuel +
            ", drain: " + fueldrainrate + ", x: " + x +
            ", speeddrop: " + speeddrop
        );
    }
}

#OBJECTIVELUCID

(C @.time 15).printCarState()
where
    C = Car() fby.time S;
    S = C.move(#time);
end;

```

Figure 74: Objective Lucid example of a Car object that changes in time.

education trees the normal arrows correspond to demands made for expressions and the bullet arrows correspond to the result of evaluation of the demands, which are also ***bold and italic***. In the Objective Lucid education tree object instance is denoted as *ClassName:MemberName:value* and the $\{d:X\}$ presents the context of evaluation. The result of evaluation of the Objective Lucid variant is said to be *true* because, as previously defined, `void` methods are mapped to return *true* and the last expression bit that is evaluated here is the `print()` method call of the instance of a `Nat32` class, which returns `void`.

```

#typedecl
Nat42;

#JAVA
class Nat42
{
    private int n;

    public Nat42()
    {
        n = 42;
    }

    public Nat42 inc()
    {
        n++;
        return this;
    }

    public void print()
    {
        System.out.println("n = " + n);
    }
}

#OBJECTIVELUCID

(N @.d 2).print[d]()
where
    dimension d;
    N = Nat42[d]() fby.d N.inc[d]();
end

```

Figure 76: The Natural Numbers Problem in Objective Lucid.

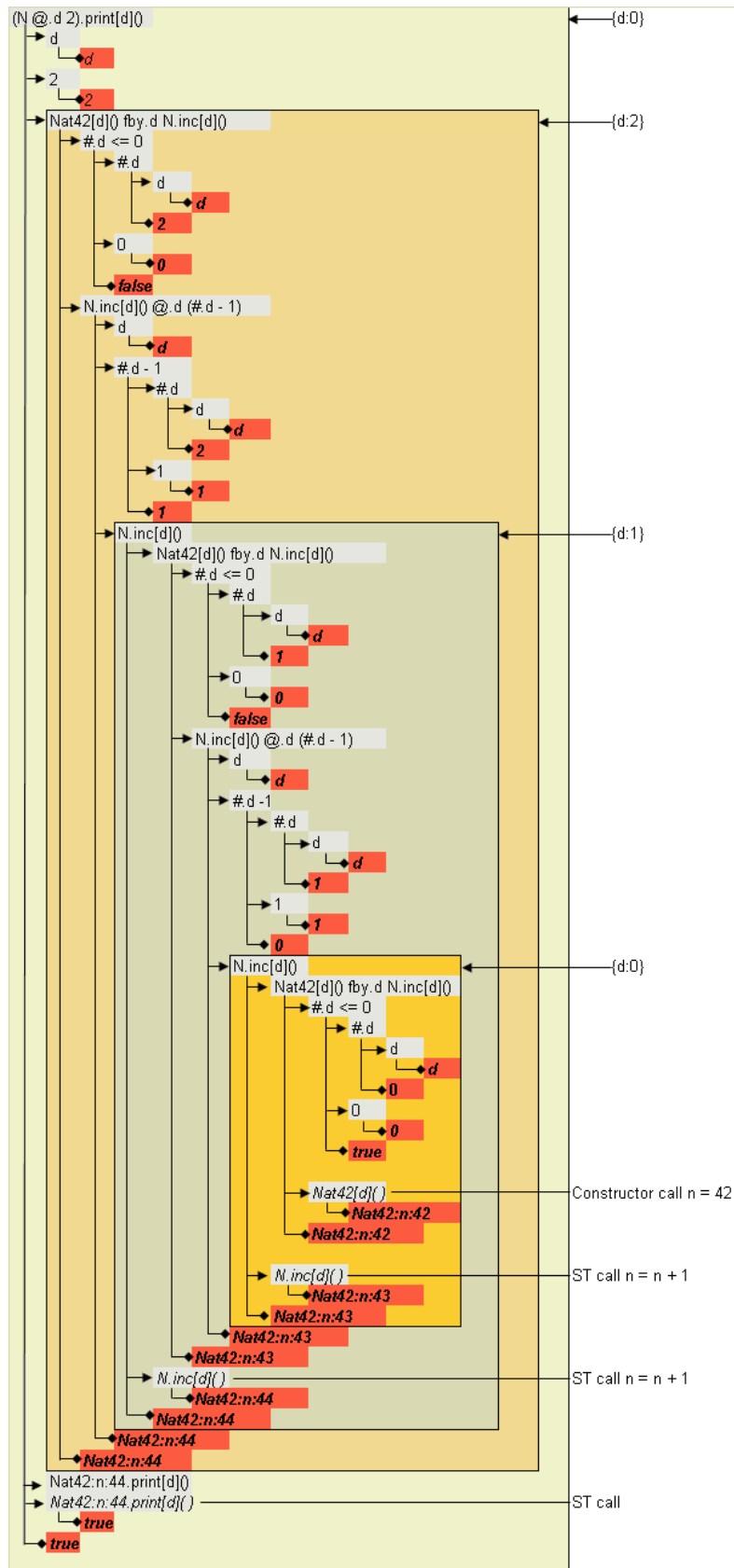


Figure 77: Education Tree for the Natural Numbers Problem in Objective Lucid.

5.3.5 Game of Life

The Game of Life [Gar70] would make a good benchmark for the GIPL. Life takes place on a 2D grid and evolves in time, so it's a 3D problem. The value of a cell at time $T+1$ depends on the value of the cell and its 8 neighbours at time T . Thus, there is a high branching factor and the IVW will get plenty of exercise. Peter Grogono wrote a version in Haskell, which is functional and lazy but is not concurrent and does not have an IVW. The author of this work made a version in Indexical Lucid. In Figure 78 is the top-level function. The Game of Life program is included in the test suite as a good elaborate test case, but this work does not address any of the performance and efficiency issues related to the execution and warehousing, so no measurements have been done to compare the efficiency of the program with and without the warehouse nor with the Haskell program.

```
life = evolve T initial (conway life) where
  initial = F(\i ->
    if val Y i == 0 && 0 <= val X i && val X i < 5 then 1 else 0)

  conway v = F(\i ->
    let neighbours v =
      ev v (n i) + ev v (ne i) + ev v (e i) + ev v (se i) +
      ev v (s i) + ev v (sw i) + ev v (w i) + ev v (nw i) in
      b2i(neighbours v == 3 || ev v i == 1 && neighbours v == 2))

  evolve d s e = F(\i ->
    if val d i == 0 then ev s i else ev e (prev d i))

  b2i b = if b then 1 else 0

  n i = F(...)
```

Figure 78: The Life in Haskell.

Explanations:

- $evolve(d, u, v)$ allows a value to evolve in the dimension d . The first value of the stream is given by u and subsequent values by v .
- $initial(d)$ defines the initial configuration (five ones in the row 0, zeroes everywhere else in the matrix 5-by-5).
- $conway(d, v)$ computes the successor of state v . The functions n , ne , e , se , s , sw , w , and nw are “navigators” that find values of neighbours.
- $b2i(d)$ converts a Boolean to integer to decide the new value of an entity.

```

#INDEXICALLUCID

life = evolve(T, initial(T), conway(life, T))
where
  dimension T;

  evolve(d, u, v) = u fby.d v;

  initial(d) =
    if(Y == 0 && 0 <= X && X < 5) then 1 else 0
  where
    X = 0 fby.d X + 1;
    Y = 0 fby.d Y + 1;
  end;

  conway(d, v) = b2i(neighbours == 3 || (v == 1 && neighbours == 2))
  where
    neighbours = n(d) + ne(d) + e(d) + se(d) + s(d) + sw(d) + w(d) + nw(d);
  where
    n(d) = v @.(d - 5);
    ne(d) = v @.(d - 4);
    e(d) = v @.(d + 1);
    se(d) = v @.(d + 6);
    s(d) = v @.(d + 5);
    sw(d) = v @.(d + 4);
    w(d) = v @.(d - 1);
    nw(d) = v @.(d - 6);
  end;

  b2i(b) = if(b) then 1 else 0;
end;
end;

```

Figure 79: The Life in Indexical Lucid.

5.4 Summary

There were many tests developed and exercised for the GIPSY. This section attempted to show the reader the most representative ones and how the Regression Tests Suite works in the GIPSY for the most modules of GIPC and GEE and how JUnit is applied to make it possible and maintainable. Now, every new module added to the GIPSY system will have to have a corresponding unit and/or regression test (or several tests) exercising most of the features of this module added.

Chapter 6

Conclusion

To conclude, it is believed GIPSY is well off the ground and is steadily getting ready for its first large public release to the research community. It is becoming a lot more usable not only by a small circle of GIPSY developers, but also by scientists and researchers from other research groups. Preliminary testing (see Chapter 5) and results (Section 6.1) give confidence in the success of an important step for the GIPSY in the are of flexible hybrid intensional-imperative programming. To summarize, the newly introduced features for the innovative intensional research platform GIPSY are a valuable asset allowing us to release GIPSY to the masses and a new release will be made at the SourceForge.net at <http://sf.net/projects/sfgipsy> circa the end of December 2005 - January 2006.

6.1 Results

6.1.1 Experiments

The experiments conducted on the GIPSY research platform were primarily design, development, and testing of hybrid programming paradigms by fusing together intensional and imperative languages. For test experiments please refer to Chapter 5.

6.1.2 Interpretation of Results

After extensive testing of the design and implementation of ideas presented in Chapter 3 we can see an enhanced, more flexible GIPSY system taking off the ground.

Most of regression tests pass for the developed sample programs with known errors and failures.

6.2 Discussions and Limitations

6.2.1 Lack of Hybrid Intensional-Imperative Semantics Proofs

The semantics for the GIPSY Type System was not defined and the one of JLucid and Objective Lucid was not formally proven to be correct.

6.2.2 Genuine Imperative Compilers

The most serious limitation of the current implementation of the hybrid paradigm is that there are no genuine imperative GIPSY compilers. The Java wrapper compiler classes merely resort to the external tools from the library of enumerated tools. This makes overall error checking and reporting cumbersome. Additionally, this slows down the compilation process.

6.2.3 Cross-Language Data Type Mapping

When implementing other imperative language compilers than Java, or a genuine compiler for Java, a special mapping has to be explicitly established in the form of `TypeMap`. We can avoid this for C/C++ with the JNI [Ste05], but not for other popular languages.

6.2.4 Dimension Index Overflow

While this limitation is not directly related to the main topics of this thesis, it has to be mentioned. In the current implementation of the dimension type in all Lucid variants is done as a simple Java integer, and as such, is finite. Thus, incorrectly written Lucid programs or programs that may require high dimension values may overflow the dimension index rendering execution of the program incorrect. This limitation is not handled by the GEE nor constrained in the operational semantics of Lucid.

6.2.5 Hybrid-DFG Integration

This thesis does not address placement, rendering, and integration of the hybrid AST nodes into DFGs.

6.2.6 Dealing With Side Effects and Abrupt Termination

As of this implementation, GEE has very limited control over what's happening inside the STs in terms side effects, exceptions, non-termination, etc. in the Java (or other imperative language) code causing it to exit prematurely or to hang. Likewise, we cannot do warehousing of non-immutable STs due to the side effects, i.e. when the same arguments are given to an ST may yield a different result at different invocations. This is serious aspect, which is related to the development of any future semantics of the hybrid programming languages and deserves a separate publication.

6.2.7 Imperative Function Overloading

It is an error to write the following:

```
#funcdecl
int foo(int);
int foo(float);
...
```

but it shouldn't be. This is an error in the sense that only the last declaration is retained due to the way function identifiers are handled, so no function overloading at this moment is officially supported. The issue of dealing with the semantics of a type system in which this is possible, especially if we support multiple imperative PLs, where each may have potentially its own type system or even paradigm is complex. However, this feature is nice to have and some practical aspects can be implemented, which will be a research topic on its own.

6.2.8 Cross-Imperative Language Calls

Normally, an ST written say in #JAVA cannot call another ST in say #C. This limitation is that only the intensional part can make calls to the imperative functions. This

eliminates the need to keep the type mappings between all possible combinations of the imperative languages and semantics associated with this.

However, depending on the language, procedures written in the same language can possibly communicate by calling each other. E.g. in Java, defining free members and passing state between free functions is possible as nothing is done to prevent this.

```
#JAVA
int i;

int foo() {
    return i + 1;
}

int bar() {
    i++;
    return foo();
}
```

This is based on the knowledge about the internal implementation i.e. the “`int i;`” bit will also be wrapped in the class, so it’d be legal to have it from the Java’s point of view; however, is considered to be a kludge and non-portable feature. To be on the safer side, the STs like that should be written assuming no knowledge of internal state for communication is available.

6.2.9 Security

JLucid, Objective Lucid, and GICF opened up doors for very flexible use of external languages and resources as a part of intensional computation. Unfortunately, there are security considerations to deal with when embedding a vulnerable unsigned code from possibly untrusted remote location and then propagate it to all the workers participating in computation can result resulting either gaining some unwanted privileges to the attackers or DDoS.

Chapter 7

Future Work

The future work to take on will focus in the following areas to either address the limitations outlined in Section 6.2 or to introduce new features, not necessarily all related to the topics of this thesis.

- Integration of the Demand Migration System (DMS) [VP05].
- Formal semantic verification from Indexical Lucid through Objective Lucid.
- Placement of hybrid nodes into DFGs.
- Security.
- Trial C compiler with JNI.
- Fully Explore Array Properties.
- Genuine imperative compilers in GICF.
- Introduction functional language compilers.
- Visualization and control of communication patterns and load balancing.
- Target Host Compilation.
- Java wrapper for the DFG Editor of Yimin Ding.

7.1 Formal Verification of Semantic Rules and the GIPSY Type System

One needs to formally conduct verification proofs of the semantic rules from Indexical Lucid to Objective Lucid in PVS or Isabelle, so this project can be undertaken in the near future and the work on it has already begun. Specifically, a relation to the semantic of objects and Java's operational semantics has to be made. Likewise, the semantics of the newly introduced GIPSY type system has to be formally defined.

7.2 Dealing with Data Flow Graphs in Hybrid Programming

This thesis did not deal with the way on how to augment `DFGAnalyzer` and `DFGGenerator` to support hybrid GIPSY programs. This can be addressed by adding an unexpandable imperative DFG node to the graph. To make it more useful, i.e. expandable and so it's possible to generate the GIPSY code off it or reverse it, would require having the genuine compilers as in Section 7.6 for imperative languages, which is far from trivial.

7.3 Security

Security is a substantial concern in distributed computing. The great flexibility provided by embedded Java in JLucid (and later in Objective Lucid) can be misused and be a source of security breaches or DDoS attacks (e.g., due to explicit oversynchronization using Java's synchronization primitives explicitly). Thus, the follow-up work in this direction would include malicious code detection in embedding and distributing as well as explicit synchronization points so that there are no deadlocks and DDoS potential is reduced. This concern touches the compiler (GIPC), the Generator-Worker architecture, the GIPSY Server, and the GIPSY Screen Saver components of the GIPSY system.

7.4 Implementation of the C Compiler in GICF

An methodology of implementing a C compiler, and therefore, C CPs and STs has been devised, but never implemented, so in the future a C compiler will be implemented as a part of GICF with the JNI [Ste05].

7.5 Fully Explore Array Properties

The arrays in JLucid, Objective Lucid, and their generalization in GICF requires further exploration and formalization and mapping of the GIPSY arrays to their native equivalents.

7.6 Genuine Imperative and Functional Language Compilers

Future work in this area is to focus on writing our genuine compilers for the mentioned imperative languages and extending support for more imperative and functional languages (e.g. LISP, Scheme, or Haskell) and make it as much automated as possible.

7.7 Visualization and Control of Communication Patterns and Load Balancing

It is proposed to have a “3D editor” within RIPE’s `DemandMonitor` that will render in 3D space the current communication patterns of a GIPSY program in execution or replay it back and allow the user visually to redistribute demands if they go off balance between workers. A kind of virtual 3D remote control with a mini expert system, an input from which can be used to teach the planning, caching, and load-balancing algorithms to perform efficiently next time a similar GIPSY application is run.

7.8 Target Host Compilation

This has to do with enabling the GEE to deliver the ST source code around and compile it on the target host instead of sending a pre-compiled version of the STs. This is an experimental feature can be useful and dangerous and requires a lot of research.

7.9 The GIPSY Screen Saver

This is a sample implementation of a worker, outlined in Section 3.3.3.4, would represent an application for a PC as a way to contribute to a GIPSY program execution. Three sample implementations of screen saver workers exist one for Windows, one for Linux and one for MacOS X.

7.10 The GIPSY Server

A so-called “GIPSY server” will be implemented to be able to serve intensional or otherwise requests primarily through the HTTP protocol, thus acting like a mini-GIPSY intensional web server. It would accept request from remote clients via HTTP or local clients via command line and be the starting point of computation (an intensional computation resource) available to all those who have no resources to set up GIPSY. This is not duplicate any of the DMS [VP05] nor it is a part of RIPE, as it is primarily non-interactive and runs on the background.

Bibliography

- [AW76] Edward A. Ashcroft and William W. Wadge. Lucid - A Formal System for Writing and Proving Programs. volume 5. SIAM J. Comput. no. 3, 1976.
- [AW77] Edward A. Ashcroft and William W. Wadge. Erratum: Lucid - A Formal System for Writing and Proving Programs. volume 6(1):200. SIAM J. Comput., 1977.
- [BddzzP⁺05] Brian Berliner, david d ‘zoo’ zuhn, Jeff Polk, Larry Jones, Derek R. Price, Mark D. Baushke, and other authors. *Concurrent Versions System (CVS)*. Free Software Foundation, Inc., 1989-2005. <http://www.nongnu.org/cvs/>.
- [Ber05] Stefan M. Bernsee. *The DFT “à pied”: Mastering The Fourier Transform in One Day*. DSPdimension.com, 1999-2005. <http://www.dspdimension.com/data/html/dftapied.html>.
- [BKU98] Gilbert Babin, Peter Kropf, and Herwig Unger. A Two-Level Communication Protocol for a Web Operating System (WOS). In *Proceedings of IEEE Euromicro Workshop on Network Computing, Vasteras, Sweden*, pages 934–944. IEEE, 1998.
- [BM96] Fancois Bourdoncle and Stephan Merz. On the integration of the functional programming, class-based object-oriented programming, and multi-methods. Technical report, Centre de Mathematiques Appliquees, Ecole des Mines de Paris and Institut fur Informatik, Technische Universitat Munchen, October 1996.

- [Bor03] Borland. *Borland JBuilder X*. Borland Software Corporation, 1997-2003. <http://www.borland.com/us/products/jbuilder/>.
- [c⁺04] Eclipse contributors et al. *Eclipse Platform*. IBM, 2000-2004. <http://www.eclipse.org/platform>.
- [Car47] Rudolf Carnap. *Meaning and Necessity: a Study in Semantics and Modal Logic*. University of Chicago Press, Chicago, USA, 1947.
- [Con05] Ant Project Contributors. *Apache Ant*. The Apache Software Foundation, 2000-2005. <http://ant.apache.org/>.
- [CP05] Various Contributors and GNU Project. *GNU Compiler Collection (GCC)*. Free Software Foundation, Inc., 1988-2005. <http://gcc.gnu.org/onlinedocs/gcc/>.
- [Dij65] E.W. Dijkstra. Cooperating Sequential Processes. Technical report, Technological University, Eindhoven, The Netherlands, 1965.
- [Dij71] E.W. Dijkstra. Hierarchical ordering of sequential processes. pages 115–138. *Acta Inf.*, 1971.
- [Din04] Yi Min Ding. Bi-directional Translation Between Data-Flow Graphs and Lucid Programs in the GIPSY Environment. Master’s thesis, Department of Computer Science and Software Engineering, Concordia University, 2004.
- [Dod96] Chris Dodd. *Intensional Programming I, chapter Rank analysis in the GLU compiler, pages 76-82*. World Scientific, Singapore, 1996.
- [Fla97] D. Flanagan. *Java in a Nutshell*. O’Reily & Associates, Inc., second edition, 1997. ISBN 1-56592-262-X.
- [Fou05] Apache Foundation. *Apache Jakarta Tomcat*. apache.org, 1999-2005. <http://jakarta.apache.org/tomcat/index.html>.
- [Gar70] Martin Gardner. The Fantastic Combinations of John Horton Conway’s New Solitaire Game ‘Life’. pages 120–123. *Scientific American*, October 1970. <http://ddi.cs.uni-potsdam>.

de/HyFISCH/Produzieren/lis_projekt/proj_gamelifel/
ConwayScientificAmerican.htm.

- [GB04] Erich Gamma and Kent Beck. *JUnit*. Object Mentor, Inc., 2001-2004. <http://junit.org/>.
- [Gin90] Armando R. Gingras. Dining Philosophers Revisited. In *SIGCSE Bulletin*, volume 22 No. 3, pages 21–28, September 1990.
- [GP99] Jean-Raymond Gagné and John Plaice. Demand-Driven Real-Time Computing. World Scientific, September 1999.
- [Gre05] Dale Green. *Java Reflection API*. Sun Microsystems, Inc., 2001-2005. <http://java.sun.com/docs/books/tutorial/reflect/index.html>.
- [Gro01] Peter Grogono. *A L^AT_EX₂ε Gallimaufry. Techniques, Tips, and Traps*. Department of Computer Science and Software Engineering, Concordia University, March 2001. <http://www.cse.concordia.ca/~grogono/documentation.html>.
- [Gro02a] Peter Grogono. *Semantics of Programming Languages, COMP745 Course Notes*. Department of Computer Science and Software Engineering, Concordia University, 1996-2002.
- [Gro02b] Peter Grogono. GIPC Increments. Technical report, Department of Computer Science and Software Engineering, Concordia University, April 2002.
- [Gro04] Peter Grogono. Intensional programming in Onyx. Technical report, Department of Computer Science and Software Engineering, Concordia University, April 2004.
- [HJ02] Marieke Huisman and Bart Jacobs. Java Program Verification via a Hoare Logic with Abrupt Termination. Technical report, Computing Science Institute, University of Nijmegen, Nijmegen, The Netherlands, 2002. www.cs.ru.nl/B.Jacobs/PAPERS/FASE00.ps.Z.

- [JA03] Harry F. Jordan and Gita Alaghband. *Fundamentals of Parallel Processing*. Pearson Education, Inc., 2003. ISBN 0-13-901158-7.
- [JD96] Raganswamy Jagannathan and Chris Dodd. GLU programmer's guide. Technical report, SRI International, Menlo Park, California, 1996.
- [JDA97] Raganswamy Jagannathan, Chris Dodd, and Iskender Agi. GLU: A high-level system for granular data-parallel programming. In *Concurrency: Practice and Experience*, volume 1, pages 63–83, 1997.
- [Kri59] Saul A. Kripke. *A Completeness Theorem in Modal Logic*. 1959.
- [Kri63] Saul A. Kripke. *Semantical Considerations on Modal Logic*. 1963.
- [LGP03] Bo Lu, Peter Grogono, and Joey Paquet. Distributed execution of multidimensional programming languages. In *Proceedings 15th IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS 2003)*, volume 1, pages 284–289. International Association of Science and Technology for Development, November 2003.
- [Lou97] Kenneth C. Loudon. *Compiler Construction: Principles and Practice*. PWS Publishing Company, 1997. ISBN 0-564-93972-4.
- [Lu04] Bo Lu. *Developing the Distributed Component of a Framework for Processing Intentional Programming Languages*. PhD thesis, Department of Computer Science and Software Engineering, Concordia University, March 2004.
- [Mam05] Qusay H. Mamoud. *Getting Started With JavaSpaces Technology: Beyond Conventional Distributed Programming Paradigms*. Sun Microsystems, Inc., July 2005. <http://java.sun.com/developer/technicalArticles/tools/JavaSpaces/>.
- [MCSN05] Serguei Mokhov, Ian Clement, Stephen Sinclair, and Dimitrios Nicolacopoulos. *Modular Audio Recognition Framework*. Department of

Computer Science and Software Engineering, Concordia University, 2002-2005. <http://marf.sf.net>.

[Mic04] Sun Microsystems. *NetBeans 4.0*. Sun Microsystems, Inc., December 2004. <http://www.netbeans.org>.

[Mic05a] Sun Microsystems. *Java Servlet Technology*. Sun Microsystems, Inc., 1994-2005. <http://java.sun.com/products/servlets>.

[Mic05b] Sun Microsystems. *The Java Website*. Sun Microsystems, Inc., 1994-2005. <<http://java.sun.com>>.

[Mic05c] Sun Microsystems. *JavaServer Pages Technology*. Sun Microsystems, Inc., 2001-2005. <http://java.sun.com/products/jsp/>.

[Moe04] Anders Moeller. Program Verification with Hoare Logic. Technical report, University of Aarhus, 2004. <http://www.brics.dk/~amoeller/talks/hoare.pdf>.

[Mok03a] Serguei Mokhov. *GIPSY: CVS Service on Newton, A Crush Guide*. Department of Computer Science and Software Engineering, Concordia University, June 2003.

[Mok03b] Serguei Mokhov. *Newton: Basic Security*. Department of Computer Science and Software Engineering, Concordia University, June 2003.

[Mok05a] Serguei Mokhov. *Makefile Tutorial*. Department of Computer Science and Software Engineering, Concordia University, 2003-2005.

[Mok05b] Serguei Mokhov. *Naming & Coding Conventions*. Department of Computer Science and Software Engineering, Concordia University, 2003-2005.

[MP05a] Serguei Mokhov and Joey Paquet. General Imperative Compiler Framework within the GIPSY. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 36–42. CSREA Press, June 2005.

- [MP05b] Serguei Mokhov and Joey Paquet. Objective Lucid – First Step in Object-Oriented Intensional Programming in the GIPSY. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 22–28. CSREA Press, June 2005.
- [MPG05] Serguei Mokhov, Joey Paquet, and Peter Grogono. Towards JLucid, Lucid with Embedded Java Functions in the GIPSY. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 15–21. CSREA Press, June 2005.
- [MS01] Brian McNamara and Yannis Smaragdakis. Functional Programming in C++ using the FC++ Library. In *SIGPLAN Notices*, volume 36(4), pages 25–30, 2001.
- [Paq99] Joey Paquet. *Scientific Intensional Programming*. PhD thesis, Department of Computer Science, Laval University, Sainte-Foy, Canada, 1999.
- [PGW04] Joey Paquet, Peter Grogono, and Ai Hua Wu. Towards a Framework for the General Intensional Programming Compiler in the GIPSY. In *19th Annual ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA 2004)*. Vancouver, Canada. ACM, October 2004.
- [PK04] Nikolaos Papaspyrou and Ioannis T. Kassios. GLU# embedded in C++: a marriage between multidimensional and object-oriented programming. In *Softw., Pract. Exper.*, volume 34(7), pages 609–630, 2004.
- [Pre93] William H. Press. *Numerical Recipes in C*. Cambridge University Press. Cambridge, UK., second edition, 1993.
- [Pro03a] David K. Probst. *Cyclic Reduction*. Concordia University, Montreal, Quebec, Canada, February 2003.
- [Pro03b] David K. Probst. *The Principle Of Least Bandwidth*. Commentary from the HEC. Concordia University, Montreal, Quebec, Canada, February 2003.

- [Pro03c] David K. Probst. *The Programmability Metric*. Commentary from the HEC. Concordia University, Montreal, Quebec, Canada, February 2003.
- [PW05] Joey Paquet and Ai Hua Wu. GIPSY – A Platform for the Investigation on Intensional Programming Languages. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 8–14. CSREA Press, June 2005.
- [Ren02] Chun Lei Ren. General Intensional Programming Compiler (GIPC) in the GIPSY. Master’s thesis, Department of Computer Science and Software Engineering, Concordia University, 2002.
- [RG05a] The GIPSY Research and Development Group. *The GIPSY Project*. Department of Computer Science and Software Engineering, Concordia University, 2002-2005. <http://newton.cs.concordia.ca/~gipsy/>.
- [RG05b] The GIPSY Research and Development Group. *The GIPSYwiki: Online GIPSY collaboration platform*. Department of Computer Science and Software Engineering, Concordia University, 2005. <http://newton.cs.concordia.ca/~gipsy/gipsywiki>.
- [Ron94] Panagiotis Rondogiannis. *Higher-Order Functional Languages and Intensional Logic*. PhD thesis, Department of Computer Science, University of Victoria, Victoria, Canada, 1994.
- [SM02] Yannis Smaragdakis and Brian McNamara. FC++: Functional tools for object-oriented tasks. In *Softw., Pract. Exper.*, volume 32(10), pages 1015–1033, 2002.
- [SMSP00] Richard Stallman, Roland McGrath, Paul Smith, and GNU Project. *GNU Make*. Free Software Foundation, Inc., 1997-2000. <http://www.gnu.org/software/make/>.
- [ST98] David B. Skillicorn and Domenico Talia. *Models and Languages for Parallel Computation*, volume 30, No. 2 of *ACM Computing Surveys*.

- ACM, Inc., June 1998. <http://www.ece.cmu.edu/~ece767/papers/skillicorn.pdf>.
- [Ste05] Beth Sterns. *The Java Native Interface (JNI)*. Sun Microsystems, Inc., 2001-2005. <http://java.sun.com/docs/books/tutorial/native1.1/index.html>.
- [Tao04] Lei Tao. Warehouse and Garbage Collection in the GIPSY Environment. Master's thesis, Department of Computer Science and Software Engineering, Concordia University, 2004.
- [VC05] Sreenivasa Viswanadha and Contributors. *Java Compiler Compiler (JavaCC) - The Java Parser Generator*. java.net, 2001-2005. <https://javacc.dev.java.net/>.
- [VP05] Emil Vassev and Joey Paquet. A Generic Framework for Migrating Demands in the GIPSY's Demand-Driven Execution Engine. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 29–35. CSREA Press, June 2005.
- [WAP05] Kaiyu Wan, Vasu Alagar, and Joey Paquet. Lucx: Lucid Enriched with Context. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 48–14. CSREA Press, June 2005.
- [WP05] Ai Hua Wu and Joey Paquet. Object-Oriented Intensional Programming in the GIPSY: Preliminary Investigations. In *Proceedings of PLC2005, Las Vegas, Nevada, USA*, pages 43–47. CSREA Press, June 2005.
- [WPG03] Ai Hua Wu, Joey Paquet, and Peter Grogono. Design of a Compiler Framework in the GIPSY System. In *Proceedings 15th IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS 2003)*, volume 1, pages 320–328. International Association of Science and Technology for Development, November 2003.
- [WSoafaotw05] Jimmy Wales, Larry Sanger, and other authors from all over the world. *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc., 2001-2005. <http://wikipedia.org>.

- [Wu02] Ai Hua Wu. Semantic Checking and Translation in the GIPSY. Master's thesis, Department of Computer Science and Software Engineering, Concordia University, 2002.

Appendix A

Definitions and Abbreviations

A.1 Abbreviations

- AST - Abstract Syntax Tree
- COM - Component Object Model
- CORBA - Common Object Requester Broker Architecture
- CLP - Cluster-Level Parallelism
- CP - Communication Procedure, Section 3.3.3.2
- CVS - Concurrent Versions System
- DCOM - Distributed COM
- DDoS - Distributed Denial of Service (attack).
- FFT - Fast Fourier Transform
- FTP - File Transfer Protocol
- DPR - Demand Propagation Resource, Section 2.5.4.1, [RG05a, PW05]
- GEE - General Education Engine
- GEER - GEE Resources, Section 4.1.1.14
- GIPC - General Intensional Program Compiler, Figure 10, [RG05a, PW05]

- GIPL - General Intensional Programming Language, [Paq99, RG05a, PW05]
- GIPSY - General Intensional Programming System, [RG05a, PW05]
- GLU - Granular Lucid, [JD96, JDA97, Paq99]
- HTTP - Hyper-Text Transfer Protocol
- IDP - Intensional Demand Propagator, Section 3.3.3.4, [RG05a, PW05]
- IDS - Intensional Data-dependency Structure
- IP - Intensional Programming
- IPL - Intensional Programming Language (e.g. GIPL, GLU, Lucid, Indexical Lucid, JLucid, Tensor Lucid, Objective Lucid, Onyx [Gro04])
- IVW - Intensional Value Warehouse, Section 3.3.3.4, [RG05a, PW05]
- JDK - Java Developer's Kit
- JNI - Java Native Interface
- JRE - Java Runtime Environment
- JSSE - Java Secure Socket Extension
- MARF - Modular Audio Recognition Framework [MCSN05]
- MPI - Message Passing Interface
- NCP - Native Communication Procedure
- NST - Native Sequential Thread
- NUMA - Non-Uniform Memory Access
- PVM - Parallel Virtual Memory System
- RFE - Ripe Function Executor, Section 3.3.3.4, [RG05a, PW05]
- RMI - Remote Method Invocation
- RPC - Remote Procedure Call

- SIPL - Specific IPL (e.g. Indexical Lucid, JLucid, Tensor Lucid, Objective Lucid, Onyx)
- SLP - Stream-Level Parallelism
- ST - Sequential Thread, Section 3.3.3.1
- TLP - Thread-Level Parallelism
- TTS - Time To Solution
- UMA - Uniform Memory Access
- URI - Unified Resource Identifier
- URL - Unified Resource Locatior

Appendix B

Sequential Thread and Communication Procedure Interfaces

In this section the actual definitions of the CP and ST interfaces, an example of a generated wrapper class and a *Worker* are presented.

B.1 Sequential Thread Interface

See Figure 80.

B.2 Communication Procedure Interface

See Figure 81.

B.3 Generated Sequential Thread Wrapper Class

This is a more complete version of the generated wrapper class for the code in Figure 23.

```

package gipsy.interfaces;

import java.io.Serializable;
import java.lang.reflect.Method;

/**
 * <p>Sequential Thread represents a piece work to be done.
 * Has to extend Serializable for RMI, CORBA, COM+, Jini to work.
 * Runnable needed to run it in a separate thread.</p>
 *
 * $Id: ISequentialThread.java,v 1.13 2005/09/12 01:24:38 mokhov Exp $
 *
 * @version $Revision: 1.13 $
 * @author Serguei Mokhov, mokhov@cs.concordia.ca
 * @since Inception
 */
public interface ISequentialThread
extends Runnable, Serializable
{
    /**
     * Work-piece to be done.
     * @return WorkResult container
     */
    public WorkResult work();

    public WorkResult getWorkResult();
    public void setMethod(Method poSTMethod);
}

// EOF

```

Figure 80: Sequential Thread Interface.

```

package gipsy.interfaces;
import gipsy.lang.GIPSYType;
import java.io.Serializable;

/**
 * <p>CommunicationProcedure represents the means of delivery of sequential threads.</p>
 * $Id: ICommunicationProcedure.java,v 1.11 2005/10/11 08:34:11 mokhov Exp $
 * @version $Revision: 1.11 $
 * @author Serguei Mokhov, mokhov@cs.concordia.ca
 * @since Inception
 * @see gipsy.interfaces.SequentialThread
 */
public interface ICommunicationProcedure
extends Serializable
{
    public GIPSYType getReturnType();
    public GIPSYType getParamType(final int piParamNumber);
    public GIPSYType[] getParamTypes();
    public void setReturnType(GIPSYType poType);
    public void setParamType(final int piParamNumber, GIPSYType poType);
    public void setParamTypes(GIPSYType[] paoTypes);
    public GIPSYType getParamType(String pstrLexeme);
    public GIPSYType getParamType(String pstrLexeme, String pstrID);
    public int getParamListSize();
    /**
     * Perform any initialization actions required.
     * @return status object of the result of send operation.
     * @throws CommunicationException in case of error
     */
    public CommunicationStatus init()
    throws CommunicationException;
    /**
     * Open a connection; whatever that means for a given protocol.
     * @return status object of the result of send operation.
     * @throws CommunicationException in case of error
     */
    public CommunicationStatus open()
    throws CommunicationException;
    /**
     * Close a connection; whatever that means for a given protocol.
     * @return status object of the result of send operation.
     * @throws CommunicationException in case of error
     */
    public CommunicationStatus close()
    throws CommunicationException;
    /**
     * Defines the means of sending data. Should be overridden by
     * a concrete implementation, such as JINI, COM, CORBA, etc.
     * @return status object of the result of send operation.
     * @throws CommunicationException in case of error
     */
    public CommunicationStatus send()
    throws CommunicationException;
    /**
     * Defines the means of receiving data. Should be overridden by
     * a concrete implementation, such as JINI, COM, CORBA, etc.
     * @return status object of the result of receive operation.
     * @throws CommunicationException in case of error
     */
    public CommunicationStatus receive()
    throws CommunicationException;
}

```

Figure 81: Communication Procedure Interface.

```

import java.util.Hashtable;
import java.util.Vector;

public class <filename>_<machine_name>_<timestamp>
implements gipsy.interfaces.ISequentialThread
{
    private OriginalSourceCodeInfo oOriginalSourceCodeInfo;

    /**
     * Inner class with original source code information
     */
    public class OriginalSourceCodeInfo
    {
        /**
         * For debugging / monitoring; generated statically
         */
        private String strOriginalSource =
            "int getN(int piDimension)" +
            "{" +
            "    if(piDimension <= 0)" +
            "        return get42();" +
            "    else" +
            "        return getN(piDimension - 1) + 1;" +
            "}" +
            "" +
            "int get42()" +
            "{" +
            "    return 42;" +
            "}";

        /**
         * Mapping to original source code position for error reporting
         */
        private Hashtable oLineNumbers = new Hashtable();

        /**
         * Body is filled in by the preprocessor statically
         */
        public OriginalSourceCodeInfo()
        {
            Vector int_getN_int_piDimension = new Vector();

            // Start line and Length in lines
            int_getN_int_piDimension.add(new Integer(3));
            int_getN_int_piDimension.add(new Integer(7));

            this.oLineNumbers.put
            (
                "int getN(int piDimension)",

```

```

        int_getN_int_piDimension
    );

    Vector int_get42 = new Vector();
    int_get42.add(new Integer(11));
    int_get42.add(new Integer(4));

    this.oLineNumbers.put
    (
        "int_get42()",
        int_get42
    );
}

public Hashtable getLineNumbersHash()
{
    return this.oLineNumbers;
}

public int getLineNumberForFunction(String pstrFunctionSignature)
{
}

public int getFunctionSourceLength(String pstrFunctionSignature)
{
}

public String toString()
{
}
}

/**
 * Constructor
 */
public <filename>_<machine_name>_<timestamp>()
{
    this.oOriginalSourceCodeInfo = new OriginalSourceCodeInfo();
}

public String toString()
{
    return this.oOriginalSourceCodeInfo.toString();
}

/*
 * Implementation of the SequentialThread interface
 */

// Body generated by the compiler
public void run()

```



```

{
    Payload oPayload = new Payload();
    oPayload.add("d", new Integer(42));

    work(oPayload);
}

// Body generated by the compiler statically
public WorkResult work(Payload poPayload)
{
    WorkResult oWorkresult = new WorkResult();
    oWorkresult.add(getN(poPayload.getVaueOf("d")));
    return oWorkResult;
}

/*
 * -----
 * The below are generated off the source file nat2java.ipl
 * -----
 */

public static int getN(int piDimension)
{
    if(piDimension <= 0)
        return get42();
    else
        return getN(piDimension - 1) + 1;
}

public static int get42()
{
    return 42;
}
}

```

B.4 Sample Worker's Implementation

```
package gipsy.wrappers;

//import gipsy.interfaces.SequentialThread;
import gipsy.interfaces.ICommunicationProcedure;
import gipsy.util.*;

import marf.util.BaseThread;

/**
 * Worker Class Definition
 *
 * $Revision: 1.11 $ by $Author: mokhov $ on $Date: 2004/11/06 00:50:09 $
 *
 * @version $Revision: 1.11 $
 * @author Serguei Mokhov
 */
public class Worker extends BaseThread
{
    /**
     * Aggregation of sequential threads.
     */
    private Thread[] aoSequentialThreads = null;

    /**
     * Set of available communication procedures for different protocols.
     */
    private ICommunicationProcedure[] aoCommunicationProcedures = null;

    /**
     * Default settings.
     */
    public Worker()
    {
        super();
    }

    /**
     * Generate a demand.
     */
    public void demand()
    {
    }

    /**
     * Receive a result on a demand.
     */
    public void receive()
    {
    }
}
```

```

}

/**
 * Perform computation.
 */
public void work() throws GIPSYException
{
    try
    {
        for(int i = 0; i < this.aoSequentialThreads.length; i++)
            this.aoSequentialThreads[i].start();
    }
    catch(NullPointerException e)
    {
        throw new GIPSYException
        (
            "Worker TID=" + getTID() +
            " did not have any sequential threads to work on."
        );
    }
}

/**
 * Stops worker thread.
 */
public void stopWorker()
{
}

/**
 * From Runnable interface, for TLP
 */
public void run()
{
    try
    {
        work();
    }
    catch(GIPSYException e)
    {
        System.err.println(e);
    }
}
}

// EOF

```

Appendix C

Architectural Module Layout

C.1 GIPSY Java Packages and Directory Structure

Normally, a directory structure of a Java project corresponds to the package naming; thus, the packages are named and declared after the directories. By the means of Java packages, all the classes within the project and external applications “know” how to identify and import the classes they intend to use. A fully-qualified class name includes all the packages starting from the “root” (the top-level directory of the hierarchy) all the way up to the class itself, separated by a dot. The GIPSY Java packages breakdown as of this writing corresponds to the Figure 82.

The logical breakdown was performed in accordance with the original conceptual design primarily produced by Joey Paquet and further by Aihua Wu and Bo Lu, has been the primary source of the hierarchy plus any exceptions and extensions that various team members come up with or have been forced to during implementation were taken into account.

The basic structure is as follows. The top root hierarchy is logically the `gipsy` package. The major non-utility packages under it, which come from the conceptual design, are `GIPC`, `GEE`, and `RIPE`. The major utility packages under `gipsy` that are not present in the conceptual design are: `interfaces` for most intermodule communication; `wrappers` for object wrapping; `storage` for the serializable interface classes; `util` for most common exceptions and utility modules (e.g. fast linked list [Din04]); and `tests` for the Unit and Regression Testing Suites.

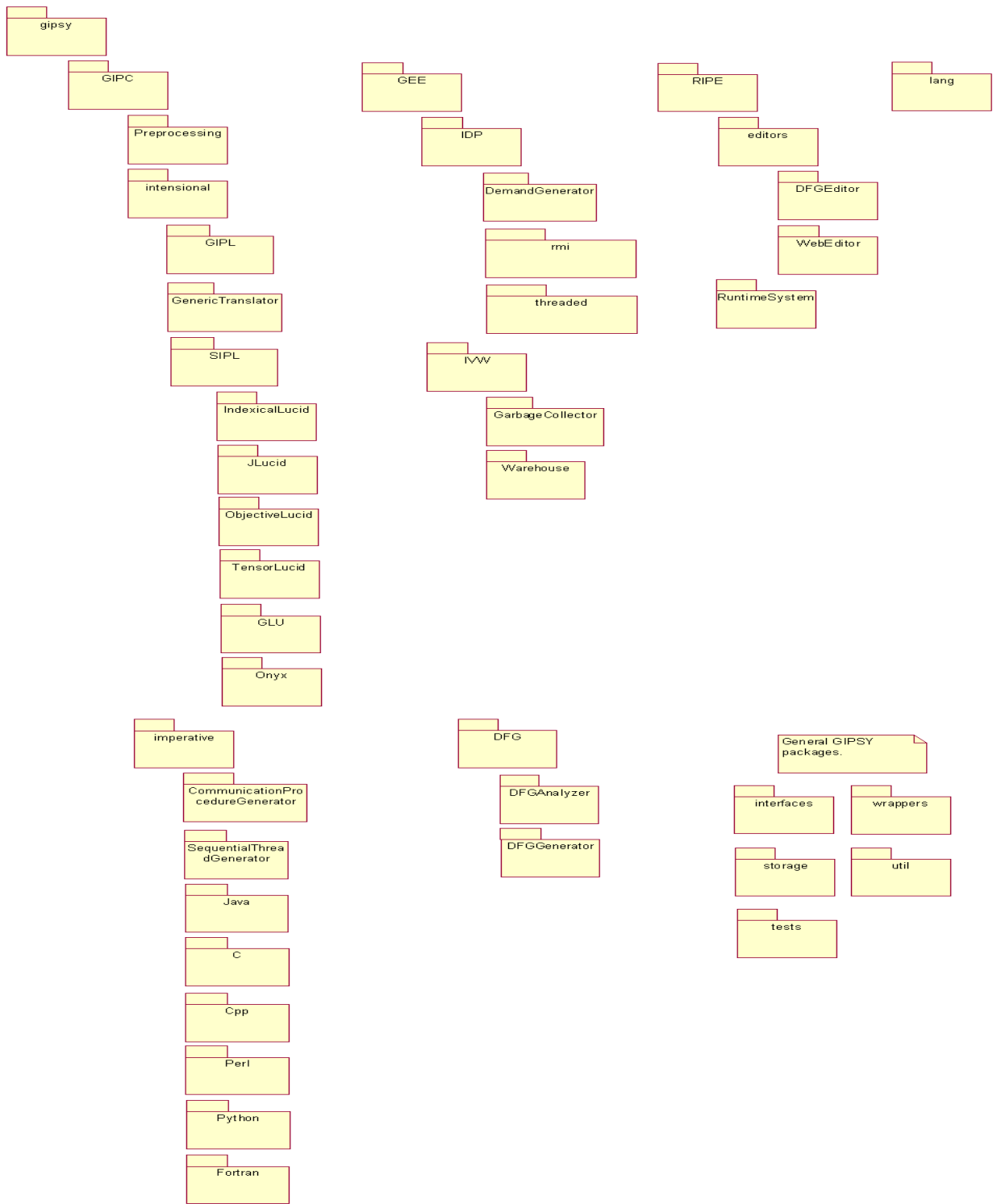


Figure 82: GIPSY Java Packages Hierarchy.

Under the `GIPC` package the major modules (to be discussed later in this chapter) include `Preprocessing` for general GIPSY program preprocessing, `intensional` and `imperative` language compilers and their necessary followers (`GenericTranslator` for the former and `CommunicationProcedureGenerator` and `SequentialThreadGenerator` for the latter). Then the `DFG` package for Lucid-to-data-flow-graph and back generation.

The `GEE`'s main packages includes `IDP` for demand propagation and `IVW` for caching and garbage collection.

Under `RIPE` we have interactive run-time editing and monitoring modules that include textual editor, `DFG` editor, and the web-based editor.

C.2 GIPSY Modules Packaging

GIPSY's major and minor modules are packaged into a set of runnable `.jar` files and distributed with wrapper scripts to be either used as ordinary command line tools as a part of GIPSY Development Kit or the `WebEditor` web application. Different `.jar` files include a subset of all GIPSY modules depending on the need, e.g. `GIPC` includes `GIPC`-related classes plus `GEE` as we allow to optionally invoke `GEE` after successful compilation. `RIPE`, except itself, needs both `GIPC` and `GEE`, whereas `GEE` does not at all require presence of any other module. Thus, the GIPSY binary distribution is broken down into five major `.jar` files (notice, that these files do not include any external libraries GIPSY references):

- `gipsy.jar` simply includes almost all of GIPSY.
- `gipc.jar` should be used/distributed as a part of so-called "GIPSY Development Kit" if someone wants to be able to compile intensional programs and optionally run them.
- `gee.jar` represents GIPSY's non-interactive run-time environment, the `GEE`. This can be distributed alone to the hosts that only wish to run pre-compiled GIPSY programs and have no development environment set up.
- `ripe.jar` includes most of the interactive programming environment of the GIPSY along with `GIPC` and `GEE`.

Table 2: Correspondence of the GIPSY .jar files and the modules.

Module / Jar	<code>gipsy.jar</code>	<code>ripe.jar</code>	<code>gipc.jar</code>	<code>gee.jar</code>	<code>Regression.jar</code>
GIPSY	*				
GIPC	*	*	*		*
RIPE	*	*			
GEE	*	*	*	*	*
DFG/GIPC	*	*	*		*
DFGEditor	*	*			
Regression					*
Interfaces	*	*	*	*	*
WebEditor					
<code>gipsy.lang</code>	*	*	*	*	*
<code>gipsy.wrappers</code>	*	*	*	*	*
<code>gipsy.util</code>	*	*	*	*	*
<code>gipsy.storage</code>	*	*	*	*	*

- `Regression.jar` includes the Regression Testing application plus all of GIPC and GEE as the most exercised modules for testing as of this writing.

The Table 2 shows correspondence between the variety of modules and their containment within a .jar file.

Appendix D

Grammar Generation Scripts for JLucid and Objective Lucid

D.1 jlucid.sh

```
#!/bin/bash

strDate='date'

cat <<GRAMMAR_TAIL
/*
 * Generated by jlucid.sh on $strDate
 */

/**
 * @since $strDate
 */
void embed() #EMBED : {}
{
    //<EMBED> <LPAREN> url() E() ( <COMMA> E() )* <RPAREN> <SEMICOLON>
    <EMBED> <LPAREN> url() <COMMA> <STRING_LITERAL> ( <COMMA> E() )* <RPAREN> <SEMICOLON>
}

/**
 * @since $strDate
 */
void array() #ARRAY : {}
{
    <LBRACKET> E() ( <COMMA> E() )* <RBRACKET>
}

/**
 * URL -> CHARACTER_LITERAL | STRING_LITERAL.
```



```

* @since $strDate
*/
void url() #URL :
{
    Token oToken;
}
{
    (
        oToken = <CHARACTER_LITERAL>
    | oToken = <STRING_LITERAL>
    )
    {
        jjtThis.setImage(oToken.image);
    }
}

// EOF
GRAMMAR_TAIL

# EOF

```

D.2 JGIPL.sh

```

#!/bin/bash

cat ../../GIPL/GIPL.jjt | \
    # Filter out unneeded stuff
    grep -v '// EOF' | \
    #grep -v 'import gipsy.GIPC.intensional.SimpleNode' | \
    # Fix package
    sed 's/intensional\GIPL/intensional\SIPL\JLucid/g' | \
    # JLucid GIPL
    sed 's/GIPL/JGIPL/' | \
    sed 's/\\/{EXTEND-E}/\/{EXTEND-E}\n\t\t| embed()/' | \
    sed 's/\\/{EXTEND-FACTOR}/\/{EXTEND-FACTOR}\n\t| array()/' | \
    sed 's/<WHERE: "where">/<WHERE: "where">\n\t| <EMBED: "embed">/g' \
    > JGIPL.jjt

./jlucid.sh >> JGIPL.jjt

# EOF

```

D.3 JIndexicalLucid.sh

```

#!/bin/bash

cat ../../SIPL/IndexicalLucid/IndexicalLucid.jjt | \
    # Filter out unneeded stuff

```

```

grep -v '// EOF' | \
#grep -v 'import gipsy.GIPC.intensional.SimpleNode' | \
# Fix package
sed 's/intensional\.\SIPL\.\IndexicalLucid/intensional\.\SIPL\.\JLucid/g' | \
# JLucid Indexical
sed 's/IndexicalLucid/JIndexicalLucid/' | \
sed 's/\\/{EXTEND-E}/\\/{EXTEND-E}\n\t\t| embed()/' | \
sed 's/\\/{EXTEND-FACTOR}/\\/{EXTEND-FACTOR}\n\t\t| array()/' | \
sed 's/<WHERE: "where">/<WHERE: "where">\n\t\t| <EMBED: "embed">/g' \
> JIndexicalLucid.jjt

./jllucid.sh >> JIndexicalLucid.jjt

```

EOF

D.4 ObjectiveGIPL.sh

```

#!/bin/bash

cat JGIPL.jjt | \
# Filter out unneeded stuff
grep -v '// EOF' | \
# Fix package
sed 's/intensional\.\SIPL\.\JLucid/intensional\.\SIPL\.\ObjectiveLucid/g' | \
# ObjectiveLucid GIPL
sed 's/JGIPL/ObjectiveGIPL/' | \
sed 's/\\/{EXTEND-E1}/\\/{EXTEND-E1}\n\t\t\t| ( <DOT> ID() ) #OBJREF E1()/' \
> ObjectiveGIPL.jjt

# EOF

```

D.5 ObjectiveIndexicalLucid.sh

```

#!/bin/bash

cat JIndexicalLucid.jjt | \
# Filter out unneeded stuff
grep -v '// EOF' | \
# Fix package
sed 's/intensional\.\SIPL\.\JLucid/intensional\.\SIPL\.\ObjectiveLucid/g' | \
# ObjectiveLucid Indexical
sed 's/JIndexicalLucid/ObjectiveIndexicalLucid/' | \
sed 's/\\/{EXTEND-E1}/\\/{EXTEND-E1}\n\t\t\t| ( <DOT> ID() ) #OBJREF E1()/' \
> ObjectiveIndexicalLucid.jjt

# EOF

```

Index

.NET Remoting, 30

API

AbstractSyntaxTree, 91, 99, 104, 120

addInvalidSegmentName(), 80

addValidSegmentName(), 80

bool, 63

boolean, 63

Car, 152

CCompiler, 66

Class, 35, 108

Class.getConstructors(), 35

Class.newInstance(), 35

CodeSegment, 79, 91

CommunicationException, 87

CommunicationProcedureGenerator, 78,
83, 85, 87, 182

CommunicationStats, 87

Constructor, 35

DemandDispatcher, 129

DemandDispatcherAgent, 129

DemandDispatcherException, 129

DemandGenerator, 129

DemandList, 129

DemandMonitor, 167

DFG, 182

DFGAnalyzer, 88, 91, 133, 166

DFGEditor, 115, 132

DFGException, 133

DFGGenerator, 38, 166

Dictionary, 91, 122–125

DictionaryItem, 124

dimension, 63

doGet(), 126

doPost(), 126

double, 63

EImperativeLanguages, 85, 135

EIntensionalLanguages, 85, 135

embed(), 44–47, 49, 51, 69, 81, 101–
103, 105

equals(), 76

Executor, 83, 93, 104, 120, 124, 129,
132

ExpandedThreadGroup, 124

Facet, 127

Float, 83

float, 63

FormatTag, 76, 78, 91, 104

Fun_Item, 99

GEE, 30, 93, 99, 115, 118, 119, 123–
127, 129, 180, 182

GEEException, 128, 129, 132

GEERGenerator, 91, 93, 97, 99

generateCommunicationProcedures(),
89

generateSequentialThreads(), 89

GenericTranslator, 182
 get42(), 81
 getDeclaredMethods(), 35
 getParameterTypes(), 35
 getReturnType(), 35
 GIPC, 79, 80, 85, 88, 91, 93, 98, 99,
 105, 115–117, 119, 123–127, 129,
 180, 182
 GIPC.DFG.DFGAnalyzer, 133
 GIPC.DFG.DFGGenerator, 133
 GIPC.intensional.GenericTranslator, 85
 GIPC.intensional.SimpleNode, 133
 GIPCException, 122, 128, 133
 GIPLCompiler, 91
 GIPSY, 114, 123, 126
 gipsy, 180
 gipsy.GEE, 129
 GIPSYArray, 104
 GIPSYEmbed, 83, 105
 GIPSYException, 128
 GIPSYFunction, 84
 GIPSYIdentifier, 83
 GIPSYObject, 104, 108
 GIPSYOperator, 84
 GIPSYProgram, 86, 93, 99, 100, 129
 GIPSYType, 99
 GIPSYVoid, 83
 Hashtable, 49
 HttpServlet, 126
 ICommunicationProcedure, 86, 87
 ICommunicationProceduresEnum, 86
 ICompiler, 88, 89, 91, 133
 IDemandList, 129
 IdentifierContextCodeGenerator, 91
 IDP, 182
 IImperativeCompiler, 85, 88, 89
 IIntensionalCompiler, 85, 88, 89
 imperative, 182
 ImperativeCompiler, 71, 89, 135
 ImperativeNode, 85, 91, 93, 99, 104,
 122
 IndexicalLucidCompiler, 91
 int, 63
 Integer, 53, 83
 intensional, 182
 IntensionalCompiler, 88, 135
 IntensionalCompiler.translate(), 85
 interfaces, 180
 IntensionalCompiler, 89
 ISequentialThread, 86, 129
 ISequentialThread, 86
 Item_in_Dict, 99
 IValueHouse, 132
 IVW, 182
 IVWInterface, 132
 JarEntry, 105
 JarInputStream, 105
 JAVA, 76
 java.lang, 83
 java.lang.Error, 122
 java.lang.Thread, 123
 java.lang.ThreadGroup, 124
 java.reflect.*, 35
 java.util.Arrays, 124
 java.util.Vector, 98, 122, 123, 125
 JavaCommunicationGenerator, 101

JavaCompiler, 49, 67, 71, 85, 91, 97,
 102–105
 JavaSequentialThreadGenerator, 87, 101
 JGIPLParser, 91
 JIndexicalLucidParser, 91
 JLucidCompiler, 91, 103, 105
 JLucidParser, 91, 101
 JLucidPreprocessor, 71, 80, 101, 105,
 108
 Lucid, 116, 117, 127
 LUSException, 129
 main(), 36
 marf.nlp, 122
 marf.Storage, 122
 marf.Storage.StorageManager, 125
 marf.util, 122
 marf.util.Arrays, 124
 marf.util.BaseThread, 123, 124
 marf.util.Debug, 125
 marf.util.ExpandedThreadGroup, 124
 marf.util.FreeVector, 122
 marf.util.Logger, 125
 marf.util.OptionProcessor, 123
 Method, 35
 Nat32, 154
 native, 36, 66
 Node, 120, 122
 NotImplementedException, 128
 NullCommunicationProcedure, 87
 Object, 76
 Object.notify(), 32
 Object.notifyAll(), 32
 Object.wait(), 32
 ObjectiveLucidCompiler, 91
 ObjectiveLucidPreprocessor, 80, 108
 ParallelTestCase, 137
 ParseException, 122
 Preprocessing, 182
 Preprocessor, 71, 79–81, 83, 85, 91,
 93, 97, 105, 108, 122
 PreprocessorParser, 38
 Regression, 37, 115, 116, 119, 123, 125,
 127, 137
 RIPE, 112, 114, 115, 132, 180, 182
 RIPEException, 128
 RMICommunicationProcedure, 87
 run(), 37
 Runnable, 86, 124
 runTest(), 37
 RuntimeException, 128
 Semantic, 98
 SemanticAnalyzer, 83, 91, 98, 99, 104,
 120
 SequentialTestCase, 137
 SequentialThreadGenerator, 78, 83, 85,
 87, 129, 182
 SequentialThreadSourceGenerator, 78
 Serializable, 86, 87
 setUp(), 37
 SimpleCharStream, 122
 SimpleNode, 91, 120, 122, 133
 storage, 180
 storage.Dictionary, 98
 storage.DictionaryItem, 99
 storage.FunctionItem, 99
 String, 63

- string, 63
- synchronized, 32
- System.loadLibrary(), 67
- tearDown(), 37
- Test, 37
- TestCase, 37, 137
- TestResult, 37
- tests, 180
- TestSuite, 37, 137
- TextualEditor, 115
- TokenMgrError, 122
- toString(), 76, 83, 124
- translate(), 89
- TranslationLexer, 85
- TranslationParser, 85
- Translator, 85, 91
- true, 59, 63
- TypeMap, 67, 104, 162
- util, 123, 180
- void, 59, 63, 154
- WebEditor, 39, 126, 182
- Worker, 62, 87, 172
- WorkResult, 87
- WorkTask, 129
- wrap(), 104
- wrappers, 180
- XLucidInterpreter, 129
- Architecture
 - Directory Structure, 180
 - GIPSY Java Packages, 180
 - GIPSY Modules Packaging, 182
- Arrays
 - JLucid, 104
- AST, 8, 37, 49, 61, 62, 74, 75, 78, 93, 97–99, 104–106, 108, 120, 122, 163, 169
- Background, 5
- Build System, 39
 - Ant, 41
 - Eclipse, 41
 - JBuilder, 41
 - Makefiles, 39
 - NetBeans, 41
- C, 1, 4, 8, 19, 21, 22, 25, 36, 38, 66, 67, 71, 74, 76, 85, 162
- C++, 1, 9, 19–22, 66, 67, 71, 74, 76, 80, 85, 162
- CLP, 24, 25
- Command-Line Interfaces
 - GEE, 118
 - GIPC, 116
 - GIPSY, 114
 - Regression, 119
 - RIPE, 115
- Communication Procedure, 64, 65
 - Interface, 172
- Compilation Sequence
 - Java, 103
 - JLucid, 102
 - Objective Lucid, 107
- Compiler Frameworks, 21
- context, 6
- CORBA, 3, 27, 30, 66, 118, 169
- CVS, 3, 38, 169
- data types

- matching Lucid and Java, 63
- DCOM+, 3, 66, 118
- Demand Dispatcher
 - Integration, 129
- Design
 - Architectural, 70
 - Detailed, 70
 - External, 112
 - External Software Interfaces, 120
 - GEE, 129
 - GICF, 85
 - GIPC, 88
 - Internal, 70
 - JLucid, 101
 - Objective Lucid, 107
 - Semantic Analyzer, 98
 - User Interface, 112
- DFG, 35, 75, 137, 163
 - Integration, 133
- dimensions, 6
- Dining Philosophers, 144
- DMS, 30, 137, 168
- DPR, 30, 169
 - GIPSY Program, 99
- education, 21
 - GLU, 21
- embed(), 44
 - implementation of, 105
- Examples
 - Dining Philosophers, 144
 - FFT, 147
 - Game of Life, 158
 - Lucid, 18
 - Moving Car, 152
 - Natural Numbers Problem, 18
 - Prefix Sum, 139
 - The Hamming Problem, 18
- Exceptions, 128
- External Software Interfaces, 120
 - JavaCC API, 120
 - MARF Library API, 122
 - Servlets API, 126
- Fast Fourier Transform, 147
- FC++, 19, 20, 22
- FEDORA CORE 2, 39, 138
- FFT, 147, 149, 169
- Files
 - *.ipl, 137
 - .c, 36, 37, 66
 - .class, 36, 44, 46, 104, 105
 - .h, 36, 37, 66
 - .ipl, 116
 - .jar, 70, 105, 114, 182, 183
 - .java, 44, 46, 104, 105
 - .jvt, 102
 - build.xml, 41
 - gee.jar, 182, 183
 - gipc.jar, 182, 183
 - GIPL.jvt, 75, 103
 - GIPSY.class, 114
 - gipsy.jar, 114, 137, 182, 183
 - GIPSY.jpj, 41
 - imperative/LANGUAGE, 135
 - IndexicalLucid.jvt, 103
 - IndexicalLucid.rul, 85
 - intensional/SIPL/LANGUAGE, 135

- Java.jjt, 76
- nbproject.xml, 41
- PreprocessorParser.jjt, 81
- README.dir, 41
- Regression.jar, 183
- ripe.jar, 182, 183
- src/tests/*, 137
- Format Tag, 76
- Fortran, 1, 8, 21, 22, 71, 85
- Frameworks
 - Compiler, 21
 - GICF, 71, 74, 85, 103
 - GIPC, 70, 74, 88
 - GIPSY Exceptions, 128
 - GIPSY Type System, 83
 - IPLCF, 85
 - JUnit, 137
 - MARF, 38
 - RIPE, 132
- Free Java Functions, 103
 - JLucid, 103
- FTP, 44, 169
- GEE, 8, 23, 25, 27–30, 32, 37, 42, 71, 74, 75, 78, 91, 97, 99, 104, 106, 108, 114, 117–120, 123, 128–130, 136, 144, 160, 162, 163, 168, 169, 182, 183
 - Command-Line Interface, 118
 - Conceptual Design, 27
 - Design, 129
 - Integration, 129
 - Introduction, 27
- GEER, 28, 30, 56, 62, 67, 75, 99, 108, 169
 - GIPSY Program, 99
- General Intensional Programming System, 23
- GICF, 2, 4, 34, 43, 60, 61, 66, 68–71, 74, 81, 85, 86, 103, 106, 135, 136, 164, 165, 167
 - Binary Compatibility, 74
 - Design, 85
 - Dictionary, 78
 - Format Tag, 76
 - GEER Generator as a Linker, 97
 - Generalization Issues, 74
 - Imperative Stubs, 97
 - Introduction, 71
 - Multiple Intensional Parts, 97
 - NCP Generator, 97
 - Sending Source Code Text, 78
 - Type Processor, 97
- GIPC, 2, 8, 22, 23, 25–28, 30, 33, 36, 37, 42, 43, 60, 64, 66–68, 71, 75, 81, 88, 90, 92, 99, 108, 116, 120, 121, 123, 128, 136, 160, 166, 169, 182, 183
 - as a Meta Processor, 91
 - Command-Line Interface, 116
 - Initial Conceptual Design, 25
 - Introduction, 25
 - Linker, 97
 - Preprocessor, 79
 - Sequence Diagram, 91
- GIPL, 1, 6, 8, 14, 15, 18, 27, 38, 43, 48, 51, 53, 61, 69, 75, 88, 89, 91, 102, 117, 120, 137, 158, 170

- Syntax, 14
- GIPSY, 1–4, 8, 19–25, 27, 32, 33, 35–39, 41–43, 53, 54, 60, 62, 67–71, 74, 76, 77, 81, 83, 85, 87, 97, 106, 112–114, 120, 122–129, 135–139, 147, 160, 161, 166, 170, 181, 182
 - Command-Line Interface, 114
 - Compilation process, 65
 - GIPC Framework with Preprocessor, 77
 - Goals, 25
 - Introduction, 23
 - Original GIPC Framework, 76
 - Screen Saver, 168
 - Security, 166
 - Server, 168
 - Structure, 24
 - Type System, 62
 - Types, 83
 - Web Front-End, 112
 - Web Portal, 112
 - WebEditor, 112
- GIPSY Exceptions, 128
- GIPSY Program, 99
 - Compiled, 99
 - GEER, 99
 - Intefacing GIPC and GEE, 99
 - Segments, 80
- GIPSY Type System, 83
- GLU, 1, 8, 9, 19–21, 29, 30, 36, 44, 64, 71, 135, 170
 - education, 21
- GLU#, 9, 19–22
- GNU, 39, 41
- Grammar
 - Generation, JLucid, 102
 - Generation, Objective Lucid, 107
 - Preprocessor, 81, 82
- Haskell, 20, 29, 158, 167
- HTTP, 44, 170
- hybrid
 - JLucid, 51
- Hybrid Programming, 19
- immutable, 81
- Implementation, 70
 - Architectural Design, 126
 - Directory Structure, 180
 - GIPSY Java Packages, 180
 - GIPSY Modules Packaging, 182
 - JLucid, 101
 - Objective Lucid, 106
 - Unit Integration, 126
- Indexical Lucid, 1, 6, 8, 9, 11, 14, 15, 18, 23, 27, 38, 43, 45, 51, 53, 69, 75, 91, 101, 102, 137, 158, 159, 165, 166, 170, 171
 - asa, 11
 - fbv, 10, 11
 - first, 11
 - next, 11
 - upon, 11
 - wvr, 11
- Integration
 - Demand Dispatcher, 129
 - DFG, 133
 - Garbage Collection, 132

- GEE, 129
- Intensional Value Warehouse, 132
- Jini, 129
- Semantic Analyzer, 98
- Intensional
 - Programming, 5
- intensional
 - logic, 6
 - operators, 6
- Intensional Programming, 5
- Interfaces
 - Communication Procedure, 172
 - Sequential Thread, 172
- Internal Design
 - GICF, 71
 - GIPC, 70
 - IPLCF, 85
- Introduction, 1
 - Contributions, 2
 - GICF, 71
 - GIPSY, 23
 - JLucid, 43
 - Scope of the Thesis, 3
 - Structure of the Thesis, 4
 - Thesis Statement, 1
- IPLCF, 70, 85, 87, 135
- Isabelle, 166
- Java, 1, 2, 4, 8, 19, 22, 25, 27, 29, 31, 32, 35, 36, 38, 43, 44, 47, 51, 53, 60, 62, 63, 66, 67, 69, 71, 74, 80, 85, 86, 103, 104, 132, 133, 139, 149, 162, 164
 - Reflection, 29, 35
- Java Compiler
 - JLucid, 103
- Jini, 3, 27, 30, 31, 66, 118, 129, 131
 - Integration, 129
- JLucid, 1, 2, 4, 8, 15, 38, 43, 44, 49, 51–54, 56, 59, 60, 62, 68–71, 81, 91, 99, 101–105, 107–109, 135–137, 139, 140, 143, 147, 162, 164, 166, 167, 170, 171, 184
 - Arrays, 104
 - AST, 105
 - Design, 101
 - Dictionary, 105
 - embed(), 105
 - Examples – FFT, 147
 - Free Java Functions, 103
 - Grammar Generation, 102
 - Implementation, 101
 - Introduction, 43
 - Java Compiler, 103
 - Non-Determinism, 44
 - Pseudo-Objectivism in, 53
 - Purpose, 43
 - Rationale, 43
 - Semantics, 52
 - SIPL, 51
 - Syntax, 51
- JNI, 36, 66, 167, 170
- JRE, 32, 170
- JSSE, 31, 170
- JUnit, 37
- Layout
 - Directory Structure, 180

- GIPSY Java Packages, 180
- GIPSY Modules Packaging, 182
- Libraries
 - MARF, 38, 112, 124, 147, 149, 170
- LINUX, 39
- LISP, 9, 29, 167
- logic
 - Hoare, 15
 - intensional, 6
 - non-intensional, 6
 - temporal, 6
- Lucid, 2, 6–9, 11, 12, 14, 18–21, 25, 32, 42–44, 47, 52–54, 56, 58, 59, 61–64, 68, 83, 108, 140, 147, 162, 170
 - Abstract Syntax, 14
 - Arrays as Objects, 109
 - Basic Operators, 10
 - Examples, 18
 - Family, 7
 - GLU, 64
 - History, 7
 - Indexical, 9, 23
 - Introduction, 7
 - JLucid, 43
 - Non-Determinism, 44
 - Objective, 53
 - Objects as Arrays, 109
 - and #, 12, 13
 - Pipelined Dataflows, 7
 - Semantics, 15
 - State of the Art, 19
 - Streams, 9
 - Tensor, 23
- Lucx, 9, 15
- MAC OS X, 39, 138
- MARF
 - FFT, 147, 149
- Methodology, 43
- ML, 29
- ML_≤, 19, 20
- MPI, 31, 170
- NetCDF, 30
- Non-Determinism, 44
- NUMA, 170
- Objective Lucid, 53
 - AST, 108
 - Design, 107
 - Dictionary, 108
 - Examples – Moving Car, 152
 - Grammar Generation, 107
 - Implementation, 106
 - Introduction, 53
 - Object Instantiation, 107
 - Semantic Rules, 60
 - Semantics of, 56
 - Syntax, 56
 - The Dot-Notation, 56, 108
- Onyx, 9, 81, 170, 171
- Options
 - all, 118, 119
 - compile-only, 114
 - corba, 118
 - dcom, 118
 - debug, 114, 115, 117–119
 - dfg, 117

-dfg='<DFG EDITOR OPTIONS>', Perl, 1, 71, 74, 78, 80, 85
 115
 -directory, 119
 -disable-translate, 117
 -gee, 117, 119
 -gee='<GEE OPTIONS>', 115
 -gipc='<GIPC OPTIONS>', 115
 -gipl, 116, 119
 -gipsy, 119
 -help, 114–116, 118, 119
 -indexical, 116, 119
 -jini, 118
 -jlucid, 117
 -objective, 117
 -parallel, 119
 -regression='<REGRESSION OPTIONS>',
 115
 -rmi, 118
 -sequential, 119
 -stdin, 116, 118
 -threaded, 118
 -translate, 117
 -txt='<TEXTUAL EDITOR OPTIONS>',
 115
 -warnings-as-errors, 117
 -G, 116
 -S, 116
 -h, 114–116, 118, 119
 [FILENAME1.gipsy [FILENAME2.gipsy]
 ...], 118
 [FILENAME1.ipl [FILENAME2.ipl] ...],
 116
 Partial Lucid, 8
 Prefix Sum, 139
 Preprocessor, 79
 GIPC, 79
 Grammar, 81
 Problems
 Dining Philosophers, 144
 FFT, 147
 Game of Life, 158
 Moving Car, 152
 Solving, 139
 PVS, 166
 Python, 1, 71, 74, 78, 80, 85
 RED HAT LINUX 9, 39, 138
 Regression
 Introduction, 136
 Testing, 136
 Regression Testing Application
 Command-Line Interface, 119
 Regression Testing Suite, 137
 Results, 161
 RIPE, 23, 32–34, 39, 42, 123, 128, 132,
 133, 167, 168, 182
 Command-Line Interface, 115
 Conceptual Design, 32
 Introduction, 32
 RMI, 2, 3, 27, 29, 31, 66, 87, 99, 118, 129,
 170
 RPC, 27, 65, 170
 Scheme, 29, 167
 Segments
 #<IMPERATIVELANG>, 80
 #<INTENSIONALLANG>, 80

- #C, 163
- #CPP, 80
- #GIPL, 80
- #INDEXICALLUCID, 80
- #JAVA, 80, 105, 163
- #JLUCID, 81, 105
- #OBJECTIVELUCID, 81
- #ONYX, 81
- #PERL, 80
- #PYTHON, 80
- #TENSORLUCID, 81
- #funcdecl, 78, 80, 105
- #typeddecl, 78, 80, 108
- Semantic Analyzer
 - Design, 98
 - Integration, 98
- Sequential Thread, 64
 - Interface, 172
 - Wrapper, 172
- Sequentiality Problem, 12
- side effect
 - immutable, 81
- SIPL, 51, 53, 61, 75, 88, 120, 135, 171
 - JLucid, 51
- SLP, 24, 25, 171
- SOLARIS 9, 39, 138
- Stream, 24
 - hardware, 24
 - Lucid variable, 25
 - of Objects, 54
 - Random access to, 12
- Synchronization, 31
 - Distributed vs. Parallel, 31
 - Implicit vs. Explicit, 32
 - in Distributed Environment, 31
 - in Parallel Environment, 31
 - Secure, 31
- Syntax
 - GIPL, 14
 - JLucid, 51
 - Objective Lucid, 56
- TCP/IP, 65
- Tensor Lucid, 1–4, 6, 8, 15, 23, 38, 43, 44, 53, 54, 58, 60, 63, 68–71, 75, 81, 102, 104, 106, 107, 110, 135–137, 152, 156, 157, 162, 164–167, 170, 171, 184
- Testing, 136
 - Diff, 137
 - Fedora Core 2, 138
 - MacOS X, 138
 - Portability, 138
 - Red Hat Linux 9, 138
 - Regression, 136
 - Solaris 9, 138
 - Unit, 137
 - Windows 98SE/2000/XP, 138
- Thesis
 - Contributions, 2
 - Scope, 3
 - Statement, 1
 - Structure, 4
- TLP, 23–25, 171
- Tools, 35
 - Ant, 41
 - bash, 102, 107

- bc.exe, 3
- bison, 38
- CVS, 38
- diff, 137
- Eclipse, 41
- flex, 38
- g++, 3
- gcc, 3
- gee, 114
- gipc, 114
- gipsy, 114
- gmake, 39
- Java, 35
- java, 37
- Java Reflection, 35
- javac, 3, 36, 49, 75, 104
- JavaCC, 37, 38, 81, 102, 103, 112
- javacc, 75, 101, 120, 122
- javah, 36, 66
- JBuilder, 41
- JGIPL.sh, 103, 185
- JIndexicalLucid.sh, 103, 185
- jlucid.sh, 103, 184
- JNI, 36
- JUnit, 3, 37, 137, 160
- lefty, 115
- lex, 38
- make, 39, 41
- make-test.sh, 41
- Makefiles, 39
- MARF, 38, 112, 122, 124, 147, 149, 170
- NetBeans, 41
- nmake.exe, 3
- ObjectiveGIPL.sh, 107, 186
- ObjectiveIndexicalLucid.sh, 107, 186
- perl, 3
- readmedir, 41
- regression, 114, 137
- ripe, 114
- Tomcat, 39
- yacc, 38
- TTS, 32, 171
- Types, 83
- UMA, 171
- UNIX, 39, 41, 137
- URI, 171
- URL, 171
- WebEditor, 112
- WINDOWS 98SE/2000/XP, 138
- Worker, 67
 - Definition, 67
 - Implementation, 178