An Introduction to the Objeck Programming Language

Randy Hollines objeck@gmail.com

February 27, 2010

Abstract

A brief introduction to the Objeck programming language and it's features. This article is intended to introduce programmers and compiler designers to the unique features and design of the Objeck language. Unless otherwise noted, this article covers functionality that is included in release 0.9.5.

Contents

1	Intr	roduction	4
2	Get	ting Started	4
	2.1	Compiling Source	5
	2.2	1 0	5
3	The	e Basics	5
	3.1	Variable Declarations	6
	3.2	T The state of the	7
		3.2.1 Mathematical and Logical Expressions	7
		3.2.2 Arrays	9
	3.3	Statements	9
		3.3.1 If Statement	9
		3.3.2 Select Statement	J
		3.3.3 While Statement	J
		3.3.4 For Statements	1
4	Use	r Defined Types	1
	4.1	Enums	1
	4.2	Classes	2
		4.2.1 Class Inheritance	2
		4.2.2 Class Casting and Identification	3
		4.2.3 Methods and Functions	4
5	Cla	ss Libraries 14	1
	5.1	Char	5
	5.2	Byte/Int	5
	5.3	Float	6
	5.4	String	6
	5.5	System Libraries and Data Structures	7
		5.5.1 Console	7
		5.5.2 File	3
		5.5.3 FileReader	3
		5.5.4 FileWriter	3
		5.5.5 Directory	9
		5.5.6 LinkedList/IntLinkedList/FloatLinkedList	

	5.5.7 Vector/IntVector/FloatVector	
6	Examples 6.1 Binary Search Tree Example	
7	Appendix A: VM Instructions	23

1 Introduction

The Objeck program language is an object-oriented computer language that is designed to be a general purpose programming system. The Objeck language allows programmers to quickly create solutions by leveraging pre-existing class libraries. The syntax for the language was designed with symmetry in mind and enforces the notion that there should only be one way to do something. Features of this release include:

- Support for object-oriented programming (all data types are treated as objects)
- Cross platform independence (OS X, Linux and Windows)
- Runtime JIT support for Intel based computers
- Automatic memory management
- Support for static shared library
- Basic block compiler optimizations

2 Getting Started

The Objeck computer language consist of a compiler and virtual machine. The compiler program is named obc, while the runtime virtual machine (VM) program is named obr. Here is the world famous "Hello World" program written in the Objeck language:

```
bundle Default {
   class Test {
     function : Main(), Nil {
        "Hello World!"->PrintLine();
     }
   }
}
```

2.1 Compiling Source

The example below compiles the source program hello.obs into the target binary file hello.obe. The two output file types that the compilers supports are executables and shared libraries. Shared libraries are binary files that contain all of the metadata needed by the compiler to relink them into programs. Both executables and shared libraries contain enough metadata to support runtime introspection (a feature that will be added in a future release). As a naming convention, executables must end in *.obe while shared libraries must end in *.obl.

Below is an example of compiling the "Hello World" program

obc -src tests\hello.obs -dest hello.obe

Additional compiler options are listed below:

Option	Description	
-src path to source files, delimited by the ',' character		
-lib	path to library files, delimited by the ',' character	
-tar	target output exe for executable and lib for library; default is exe	
-opt	optimization level s0-s3 with s3 being the most aggressive; default is s0	
-dest	output file name	

2.2 Executing

The command line example below executes the hello.obe executable. Note, for executables all required libraries are statically linked in the target output file. When compiling shared libraries, other shared libraries are <u>not</u> linked into the target output library file.

obr hello.obe

3 The Basics

Now lets introduce you the core features of the Objeck programming language.

In Objeck, all data types are treated as objects. Basic objects provide supports for boolean, character, byte, integer and decimal types. These basic objects can be used to create complex user defined objects. The listing below defines the basic objects that are supported in the language:

Type	Description
Char	1-byte character
Char[]	character array
Bool	boolean value
Bool[]	boolean array
Byte	1-byte integer
Byte[]	byte array
Int	4-byte integer
Int[]	integer array
Float	8-byte decimal
Float[]	decimal array

As mentioned above, basic types are objects and have associated methods for each basic class type. For example:

```
13->Min(3)->PrintLine();
13->Max(3)->PrintLine();
-22->Abs()->PrintLine();
Float->Pi()->PrintLine();
```

3.1 Variable Declarations

Variables can be declared for all of the basic types described above and for user defined objects. Variables can be declared anywhere in a program and are bound to traditional block scoping rules. Variable assignments can be made during a declaration or at any other point in a program. Variables may be declared as local, class instance or class variables. Class level variables are declared using the static keyword. A class that is derived from another class may access it's parents variables if the parent class is declared in one of the source programs. If a class is derived from a class declared in a shared library then that class cannot access it's parents variables, unless an accessor method is provided. Local variables can be declared without specifying their data type, such variables are bound to a type following their first assignment. Three different declaration styles are shown below:

```
a : Int;
b : Int := 13;
c := 7;
```

Types that are not initialized at declaration time are initialized with the following default values:

Type	Initialization
Char	'\0'
Byte	0
Int	0
Float	0.0
Array	Nil
Object	Nil

3.2 Expressions

The Objeck language supports various expression types. Some of these expression types include mathematical, logical, array and method call expressions. The preceding sections describe some of the expressions that are supported in the Objeck language.

3.2.1 Mathematical and Logical Expressions

The following code example demonstrates two ways to printing the number 42. The first way invokes the PrintLine() method for the literal 42. The second prints the product of a variable and a literal.

```
bundle Default {
  class Test {
    function : Main(), Nil {
      42->PrintLine();
      eight := 8;
      (eight * 7)->PrintLine();
    }
  }
}
```

The following mathematical operators are supported in the Objeck language for integers and decimal values:

- addition (+)
- subtraction (-)
- multiplication (*)
- division (/)
- modulus (% for integer values only)

The [*, /, %] operators have a higher precedence than the [+, -] operators. Operators of the same precedence are evaluated from left-to-right. Logical operations are of lower precedence than mathematical operations. All logical operators are of the same precedence and order is determined via left-to-right evaluation. The [&, |] logical operators use short-circuit logic; meaning that some expressions may not be executed if evaluation criteria is not satisfied.

The following logical operators are supported in the Objeck language:

- and (&)
- or (|)
- equal (=)
- not-equal (<>)
- less-than (<)
- greater-than (>)
- less-than-equal (<=)
- greater-than-equal (>=)

3.2.2 Arrays

The Objeck language supports single and multi-dimensional arrays. Arrays are allocated dynamically from the system heap. The memory that is allocated for arrays is managed automatically by the runtime garbage collector. All of the basic types described above (as well as user defined types) can be allocated as arrays. The code example below shows how a two-dimensional array of type Int is allocated and dereferenced.

```
array := Int[,] = Int->New[2,3];
array[0,2] := 13;
array[1,0] := 7;
```

The size of an array can be obtained by calling the array's GetSize() method. The GetSize() method will return the number of elements in a given array. For a multi-dimensional array the size method returns the number of elements in the first dimension. Character array literals are allocated as System.String objects. It should also be noted that language has a System.String class that provides support for advanced string operations.

```
str := "Hello World!";
str->GetSize()->PrintLine();
```

3.3 Statements

Besides providing support for declaration statements the language has support for conditional and control statements. As with other languages, control statements can be nested in order to provide finer grain logical control. General control statements include if and select statements. Basic looping statements include while and for loops. Note, all statements rather decelerations or controls end with a ';'.

3.3.1 If Statement

An if statement is a control statement that executes the associated block of code if it evaluates to true. If the evaluation statement does not evaluate to true than an else if statement may be evaluated (if it exists), otherwise an else statement will be executed (if it exists). The example below demonstrates an if statement.

```
value : Int := System.ReadLine()->ToInt();
if(value <> 3) {
   "Not equal to 3"->PrintLine();
}
else if(value < 13) {
   "Less than 13"->PrintLine();
}
else {
   "Other number"->PrintLine();
};
```

3.3.2 Select Statement

A select statement maps a value to 1 or more labels. Labels are associated to statement blocks. A label may either be a literal or an enum value. Multiple labels can be mapped to the same statement block. Below is an example of a select statement.

```
select(v) {
    label Color->Red: {
        "Red"->PrintLine();
    }

    label 9:
    label 19: {
        v->PrintLine();
    }

    label 27: {
        (3 * 9)->PrintLine();
    }
};
```

3.3.3 While Statement

A while statement is a control statement that will continue to execute its main body as long as it's conditional expression evaluates to true. When its conditional expression evaluates to false than the loop body will cease to execute.

```
i : Int := 10;
while(i > 0) {
   i->PrintLine();
   i := i - 1;
}
```

3.3.4 For Statements

The for statement is another common looping construct. The for loop consists of a pre-condition statement followed by an evaluation expression and an update statement.

```
name : Char[] := "John";
for(i : Int := 0; i < name->GetSize(); i := i + 1;) {
   name[i]->PrintLine();
}
```

4 User Defined Types

4.1 Enums

Enums are user defined enumerated types. The main use of an enum is to group a class of countable values, for example colors, into a distinct class. Once enum values have been defined they may not be assigned or associated to a other enum groups or integer classes. The valid operations for enums are as follows:

- assignment (:=)
- equal (=)
- not-equal (<>)

In addition, enum values may be used in **select** statements as conditional tests or labels.

```
enum Color {
   Red,
   Black,
   Green
}
```

4.2 Classes

Classes are user defined types that allow programmers to create specialized data types. Classes are made up of attributes (data) and operations (methods). Classes are used to encapsulate programming logic and localize information. Operations that are associated to a class may either be at the class level or instance level. Class instances are created by calling an object's New() function. Note, an object instance can only be created if one or more New() functions have been defined.

4.2.1 Class Inheritance

Classes may be derived from other classes using the from keyword. Class inheritance allows classes to share common functionality. The Objeck language supports single class inheritance, meaning that a derived class may only have one parent. The language also supports virtual classes, which assures that derived classes have been defined for all required operations declared in the base class. Virtual classes also allow the programmer to define non-virtual methods that contain program behavior. Virtual classes are dynamically bound to implementation classes at runtime.

```
class Foo {
    @lhs : Int;

New(lhs : Int) {
    @lhs := lhs;
}

method : native : AddTwo(rhs : Int), Int {
    return 2 + rhs;
}

method : virtual : AddThree(int rhs), Int;

method : GetLhs(), Int {
    return lhs;
}
```

```
class Bar origin Foo {
  New(value : Int) {
    Parent(value);
}

method : native : AddThree(rhs : Int), Int {
    return 3 + rhs;
}

function : Main(), Nil {
    bar : b := Bar->New(31);
    b->AddThree(9)->PrintLine();
}
```

4.2.2 Class Casting and Identification

An object that is inherited from another object may be either upcasted or downcasted. Object casting can be performed using the As() operator. The Object language detects upcasting and downcasting at compile time. Upcasting requires a runtime check, while down casting does not. If cross casting is detected then a compile time error will be generated.

```
method : public : Compare(right : System.Base), Int {
  if(right <> Nil) {
   if(GetClassID() = right->GetClassID()) {
      a : A := right->As(A);

  if(@value = a->GetValue()) {
      return 0;
    };
}
```

The class that a given object instance belongs to can found by calling its GetClassID method. This method returns an enum that is associated with that instance's class type. This method is generally used to determine if two object instances are of the same or different classes.

4.2.3 Methods and Functions

The Objeck language support both methods and functions. Functions are public static procedures that may be executed by any class. Methods are operations that may be performed on an object instance. Methods have public and private qualifiers. Methods that are private may only be called from within the same class, while public methods may be called from other classes. Note, methods are private by default. The Objeck language supports polymorphic methods and functions, meaning that there can be multiple methods with the same name within the same class as long as their declaration arguments vary.

Methods and functions can either be executed in an interpreted or JIT compiled mode. Interpreted execution mimics microprocessor functions in a platform independent manner. JIT execution takes the compiled stack code an produces native machine code. Note, that there is initial overhead involved in the JIT compilation process since it occurs at runtime. In addition, some methods can not be compiled into native machine code but this is a rare case. The keyword native is used to JIT compile methods and function at runtime.

A function or method may be defined as virtual meaning that any class that originates from that class must implement all of the class's virtual methods or functions. Virtual methods are a way to ensure that certain operations are available to a family of classes. If a class declares a virtual method then the class become virtual, meaning that it cannot be directly instantiated.

Below is an example of declaring a virtual method:

```
method : virtual : public : Compare(right : System.Base), Int;
```

5 Class Libraries

Objeck includes class libraries that provides access to system resources, such as files and sockets, while also providing support for basic data structures like lists and vectors. As new class libraries are added they will be documented in this section.

5.1 Char

- IsDigit determines if the character is a digit (in the range of 0-9)
- IsChar determines if the character is a alpha (in the range of A-Z or a-z)
- Min returns the smallest of the two numbers; returns the same number if they are equal
- Max returns the largest of the two numbers; returns the same number if they are equal
- Print prints the current value
- PrintLine prints the current value along with a line return
- ToString converts the current value to a System.String object instance

5.2 Byte/Int

- Min returns the smallest of the two numbers; returns the same number if they are equal
- \bullet Max returns the largest of the two numbers; returns the same number if they are equal
- Abs returns the absolute value of the current number
- Print prints the current value
- PrintLine prints the current value along with a line return
- ToString converts the current value to a System.String object instance

5.3 Float

- Min returns the smallest of the two numbers; returns the same number if they are equal
- \bullet Max returns the largest of the two numbers; returns the same number if they are equal
- Abs returns the absolute value of the current number
- Floor returns the floor of the current number
- Ceiling returns the ceiling of the current number
- Pi returns the value of Pi
- Print prints the current value
- PrintLine prints the current value along with a line return
- ToString converts the current value to a System.String object instance

5.4 String

- Append Appends a System.String, Char[], Char, Int or Float to the current String instance
- GetIndex returns the index of the first occurrence of a given Character
- GetSize returns the size of the String
- ToCharArray converts a string to a Char[]
- ToInt converts a string to a Int
- ToFloat converts a string to a Float
- SubString creates a new string that contains a subset of the string's contents
- Equals compares two string returns true if they are equal
- Compare compares two string returns 0 if they are equal

- Print prints the String object's contents
- PrintLine prints the String object's contents along with a line return

5.5 System Libraries and Data Structures

The following data structures are supported:

- Console
- LinkedList
- IntLinkedList
- FloatLinkedList
- Vector
- IntVector
- FloatVector
- BinaryTree

5.5.1 Console

The Console class allows programmers to read and write information to the system console. The class supports the following operations:

- Print prints all basic types including String and Char[] to standard out.
- PrintLine prints all basic types including String and Char[] to standard out followed by a newline.
- ReadLine reads in a line of text as a Char[] from standard in.

5.5.2 File

The File class allows programmers manipulate system files. The class supports the following operations:

- IsOpen returns true if file is open.
- IsEOF returns true if the file pointer is at the EOF.
- Seek seeks to a position in a file.
- Rewind moves the file pointer to the beginning of a file.
- GetSize returns the size of the file.
- Delete deletes a file.
- Exists returns true if the file exists.
- Rename renames a file

5.5.3 FileReader

The FileReader is inherited from the File class and allows programmers read files. The class supports the following operations:

- Close closes a file.
- ReadByte reads a byte from a file.
- ReadBuffer reads n number of bytes from a file.
- ReadString reads a line from a file.

5.5.4 FileWriter

The FileReader is inherited from the File class and allows programmers read files. The class supports the following operations:

- Close closes a file.
- WriteByte writes a byte to a file.
- WriteBuffer writes n number of bytes to a file.
- WriteString writes a string to a file.

5.5.5 Directory

The Directory class allows programmers manipulate filesystem directories. The class supports the following operations:

- Create creates a new directory.
- Exists returns true if the directory exists.
- List returns vector of file and directory names.

5.5.6 LinkedList/IntLinkedList/FloatLinkedList

The LinkedList class allow values inserted and add to the front and back of a list. The class supports the following operations:

- AddBack adds a new value to the back of the list
- AddFront adds a new value to the front of the list
- InsertElement inserts a new value in the position pointed to the cursor
- RemoveBack removes the last element in the list
- RemoveFront removes the first element in the list
- RemoveElement removes the element pointed to the cursor
- Next advances the internal cursor by one element
- Pervious retreats the internal cursor by one element
- GetValue returns the value of the element pointed to by the cursor
- Forward moves the cursor to the end of the list
- Rewind moves the cursor to the start of the list
- GetSize returns the size of the list

5.5.7 Vector/IntVector/FloatVector

The Vector class support the concept of a growing array. The class supports the following operations:

- AddBack adds a new value to the back of the vector
- RemoveBack removes the last element in the vector
- GetValue returns the value of the element pointed to by the cursor
- SetValue replaces the list value based upon the given index
- GetSize returns the size of the list

5.5.8 BinaryTree

The BinaryTree class supports the concept of an associative array with key/value pairs. The class implements a balance binary tree algorithm such that inserts, deletes and searches are $\log_2 n$:

- Insert adds a new value to the tree
- Delete removes a value from the tree
- Find searches for a value based upon a key
- SetValue replaces the list value based upon the given index
- GetKeys returns a vector of keys
- GetValue returns a vector of values

6 Examples

6.1 Binary Search Tree Example

```
use System;
bundle Default {
  class Test {
    function : Main(args : System.String[]), Nil {
```

```
Run();
      "Done!"->PrintLine();
   }
   function : native : Run(), Nil {
      tree : BinaryTree := BinaryTree->New();
      for(i : Int := 0; i < 1000; i := i + 1;) {
        v : IntHolder := IntHolder->New(i);
        s : System.String := System.String->New("Pug-");
        s->Append((i + 2)->ToString()->ToCharArray());
        tree->Insert(v->As(System.Compare), s->As(System.Base));
      };
      v : Vector := tree->GetKeys();
      for(i : Int := 0; i < v->GetSize(); i := i + 1;) {
       h : IntHolder := v->GetValue(i)->As(Structure.IntHolder);
       h->GetValue()->PrintLine();
      };
      v := tree->GetValues();
      for(i : Int := 0; i < v->GetSize(); i := i + 1;) {
        v->GetValue(i)->As(System.String)->PrintLine();
      };
   }
 }
}
```

6.2 Prime Number Example

```
bundle Default {
  class FindPrime {
    function : Main(), Nil {
      Run(1000000);
    }

  function : native : Run(topCandidate : Int), Nil {
      candidate : Int := 2;
      while(candidate <= topCandidate) {</pre>
```

```
trialDivisor : Int := 2;
        prime : Int := 1;
        found : Bool := true;
        while(trialDivisor * trialDivisor <= candidate & found) {</pre>
          if(candidate % trialDivisor = 0) {
            prime := 0;
            found := false;
          }
          else {
            trialDivisor := trialDivisor + 1;
          };
        };
        if(found) {
          candidate->PrintLine();
        };
        candidate := candidate + 1;
      };
   }
 }
}
```

7 Appendix A: VM Instructions

The appendix below lists the types of stack instructions that are executed by the Objeck VM. The VM was designed to be portable and language independent. Early development versions of the VM included an inline assembler, which may be re-added in future releases.

	Stack Opera	tors
Mnemonic	Opcode(s)	Description
LOAD_INT_LIT	4-byte integer	pushes integer onto stack
LOAD_FLOAT_LIT	8-byte float	pushes float onto stack
LOAD_INT_VAR	variable index	pushes integer onto stack
LOAD_FLOAT_VAR	variable index	pushes float onto stack
LOAD_SELF	n/a	pushes self integer on stack
STOR_INT_VAR	variable index	pops integer from stack and saves to index location
STOR_FLOAT_VAR	variable index	pops float from stack and saves to index location
COPY_INT_VAR	variable index	copies an integer from stack and saves to index location
COPY_FLOAT_VAR	variable index	copies a float from stack and saves to index location
LOAD_BYTE_ARY_ELM	array dimension	pushes byte onto stack; assumes array address was pushed prior
LOAD_INT_ARY_ELM	array dimension	pushes integer onto stack; assumes array address was pushed prior
LOAD_FLOAT_ARY_ELM	array dimension	pushes float onto stack; assumes array address was pushed prior
LOAD_ARY_SIZE	n/a	pushes array size as integer onto stack; assumes array address was pushed prior
STOR_BYTE_ARY_ELM	variable index	stores byte at index location; assumes array address was pushed prior
STOR_INT_ARY_ELM	variable index	stores integer at index location; assumes array address was pushed prior
STOR_FLOAT_ARY_ELM	variable index	stores float at index location; assumes array address was pushed prior

	Logical	Operators
Mnemonic $Opcode(s)$		Description
$\mathrm{EQL}_{-}\mathrm{INT}$	n/a	pops top two integer values and pushes result of equal operation
NEQL_INT	n/a	pops top two integer values and pushes result of not-equal operation
LES_INT	n/a	pops top two integer values and pushes result of less-than operation
GTR_INT	n/a	pops top two integer values and pushes result of greater-than oper- ation
LES_EQL_INT	n/a	pops top two integer values and pushes result of less-than-equal op- eration
GTR_EQL_INT	n/a	pops top two integer values and pushes result of greater-than-equal operation
EQL_FLOAT	n/a	pops top two floats values and pushes result of equal operation
NEQL_FLOAT	n/a	pops top two floats values and pushes result of not-equal operation
LES_FLOAT	n/a	pops top two floats values and pushes result of less-than operation
GTR_FLOAT	n/a	pops top two floats values and pushes result of greater-than operation
LES_EQL_FLOAT	n/a	pops top two floats values and pushes result of less-than-equal op- eration
GTR_EQL_FLOAT	n/a	pops top two floats values and pushes result of greater-than-equal operation
AND_INT	n/a	pops top two integer values and pushes result of add operation
OR_INT	n/a	pops top two integer values and pushes result of or operation

Mathematical Operators			
Mnemonic $Opcode($		Description	
ADD_INT	n/a	pops top two integer values and	
		pushes result of add operation	
SUB_INT	n/a	pops top two integer values and	
	,	pushes result of subtract operation	
MUL_INT	n/a	pops top two integer values and	
	,	pushes result of multiply operation	
DIV_INT	n/a	pops top two integer values and	
-		pushes result of divide operation	
SHL_INT	n/a	pops top two floats values and	
		pushes result of shift left operation	
SHR_INT	n/a	pops top two floats values and	
		pushes result of shift right operation	
MOD_INT	n/a	pops top two integer values and	
		pushes result of modulus operation	
ADD_FLOAT	n/a	pops top two floats values and	
		pushes result of greater-than-equal	
		operation	
SUB_FLOAT	n/a	pops top two floats values and	
		pushes result of subtract operation	
MUL_FLOAT	n/a	pops top two floats values and	
		pushes result of multiply operation	
DIV_FLOAT	n/a	pops top two floats values and	
		pushes result of divide operation	
I2F	n/a	pop top integer and pushes result of	
		float cast	
F2I n/a		pop top float and pushes result of	
		integer cast	

${ m Objects/Methods/Traps}$			
Mnemonic	Opcode(s)	Description	
RTRN	n/a	exits existing method returning con-	
		trol to callee	
MTHD_CALL	integer values for class	synchronous call to given method	
	id and method id	releasing control	
ASYNC_MTHD_CALL	integer values for class	asynchronous call to given method	
	id and method id;		
1 07777 0 7 0 7 7 7 7 7 7 7 7 7 7 7 7 7	pushes new thread id		
ASYNC_JOIN	thread id	waits for identified thread to end ex-	
1.01		ecution	
LBL	label id	identifies a jump label	
JMP	label id and conditional	jump to label id	
	context (1=true, 0=un-		
MEM DIVER ADV	conditional, -1=false)		
NEW_BYTE_ARY	array dimension	pushes address of new byte array	
NEW_INT_ARY	array dimension	pushes address of new integer array	
NEW_FLOAT_ARY	array dimension	pushes address of new float array	
NEW_OBJ_INST	integer value for class id	pushes address of new class instance	
OBJ_INST_CAST	integer values for from	performs runtime class cast check	
	class and to class	(note: only required for up casting)	
TRAP	integer value for trap id	calls runtime subroutine releasing control	
TRAP_RTRN	integer value for trap	calls runtime subroutine releasing	
	id and number of argu-	control and then process an integer	
	ments	return value	
LIB_NEW_OBJ_INST	n/a	symbolic library link for a new ob-	
		ject instance	
LIB_MTHD_CALL	n/a	symbolic library link for a method call	
LIB_OBJ_INST_CAST n/a		symbolic library link for an object	
	,	cast	