Self and OpenSmalltalk VMs

Jecel Mattos de Assumpção Jr
jecel@merlintec.com

RISC-V J Extension Meeting
September 20, 2018
● Smalltalk history and hardware
● Self virtual machine
● OpenSmalltalk virtual machine
Smalltalk history

Alan Kay

Flex 1968/1969

Seymour Papert

Ivan Sutherland's Sketchpad

B220 tape format

Simula I

Dynabook

LISP 1.5
Smalltalk-72: one page design

Dan Ingalls

DG Nova
Basic =>
Nova
assembly
Xerox Alto

Butler Lampson, Chuck Thacker
Smalltalk-74 and OOZE

Object-Oriented Zoned Environment

When the Smalltalk interpreter needs an object, it is first looked up in the Resident Object Table (ROT). If the object is not found in the ROT, the ROT is searched. If the object is found in the ROT, its memory address is retrieved. If the object is not found in the ROT, a hash function is used to find the object in memory. If the object is found in memory, its fields are accessed. If the object is not found in memory, it is not in core.

Operations that do not move ROT entries:
- Insert an Object Pointer
- Delete an Object Pointer
- Change the memory address of an Object

These operations are safe as they do not move the ROT entries of other objects.
Smalltalk-76

Fixed syntax
Bytecodes
Reflection
Inheritance
SystemTracer
Smalltalk-78 and the Notetaker

Multiple 8086 processors => kernels
Small integers are unboxed
Object Table
Smalltalk-80

- Apple
- Tektronix
- DEC => Berkeley
- HP
**lookupMethodInDictionary**: dictionary

```
<table>
<thead>
<tr>
<th>length index mask wrapAround nextSelector methodArray</th>
</tr>
</thead>
<tbody>
<tr>
<td>length ← memory fetchWordLengthOf: dictionary.</td>
</tr>
<tr>
<td>mask ← length - SelectorStart - 1.</td>
</tr>
<tr>
<td>index ← (mask bitAnd: (self hash: messageSelector)) + SelectorStart.</td>
</tr>
<tr>
<td>wrapAround ← false.</td>
</tr>
</tbody>
</table>

true] whileTrue:

[nextSelector ← memory fetchPointer: index |
  ofObject: dictionary. |

nextSelector = NilPointer ifTrue: [ifFalse]. |

nextSelector = messageSelector |

ifTrue: [methodArray ← memory fetchPointer: MethodArrayIndex |
  ofObject: dictionary. |

newMethod ← memory fetchPointer: index - SelectorStart |
  ofObject: methodArray. |

primitiveIndex ← self primitiveIndexOf: newMethod. |

true]. |

index ← index + 1. |

index = length |

ifTrue: [wrapAround ifTrue: [ifFalse]. |

wrapAround ← true. |

index ← SelectorStart]|
```
Hardware: Xerox PARC


Alto
  TTL

Alto II
  TTL

Dolphin
  TTL

Dandelion
  2900 bitslice

Dorado
  ECL
Hardware: bytecodes

**Sword32 (Katana32)**
- 1984
- ASIC
- University of Tokyo

**AI32**
- 1986
- ASIC
- Hitachi

**Swamp**
- 1986
- Bitslice
- University of Toronto

<table>
<thead>
<tr>
<th>31</th>
<th>Positive 31 bit integer</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>Object pointer</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Stack reference</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Negative 31 bit integer</td>
<td></td>
</tr>
</tbody>
</table>
Hardware: RISC III (SOAR)

Smalltalk On A RISC

1983 - 1985

Joan Pendleton, David Ungar, Shing Kong, Will Brown, Frank Dunlap, and Chris Marino

Professors David Hodges and David Patterson
Hardware: RISC III (SOAR)

Inline caches

L Peter Deutsch and Alan Schiffman
Hardware

COM
1984
Caltech
James T. Kajiya and
William J. Dally

Mushroom
1987
University of Manchester
Ifor Wyn Williams, Mario I. Wolczko, and Trevor P. Hopkins

Rekursive
1986
Linn
David Harland

J-Machine
1988
MIT
William J. Dally and others
Self: The Power of Simplicity

Smalltalk-86 call for projects at PARC

David Ungar and his graduate students at Stanford, later Sun

ARK: Alternate Reality Kit by Randy Smith
Self: like Smalltalk only more so

Smalltalk-80 vs Self

- Blocks
- Messages
- Assignment
- Instance variables
- Instances
- Classes
- Metaclasses
- Temporary variables
- Class variables
- Class instance variables
- Globals
- Pool variables

- Objects
- Named slots (3 kinds)
- Parent slots
- Blocks
- Messages
Self: memory organization

- Integer immediate (or virtual machine address)
  - 30-bit signed integer
  - Top 30 bits of word-aligned address
  - Reference to SELF heap object
  - 30 bits of IEEE floating point number
  - Scavenging fields and hash field
  - Mark header word (begins SELF heap object)

- Plain object
  - Mark
  - Map
  - Slot contents

- Object array
  - Mark
  - Map
  - Array length
  - Slot contents

- Byte array (segregated bytes part)
  - Mark
  - Map
  - Array length
  - Bytes pointer
  - Bytes

- VM memory display
  - Indicators and VM memory display
  - Object memory
  - Code cache
  - Code
  - Deps
  - Debug
Self 1 and 2: type inference

The computer wants to run Fortran (or C)
The programmer wants to write in Self

Polymorphism
Factored code
Generic code
Dynamic types
Language evolution
Compiler technology
PICs
Inlining
Customization
Type inference
Inlining
Customization
Type inference

Craig Chambers

The computer wants to run Fortran (or C)
The programmer wants to write in Self
Self 3 and 4: adaptive compilation

Type feedback instead of type inference

Two compilers: simple and fast generates instrumented code, good compiler for "hot spots"
Squeak: Back to the Future

Boxer

StarLogo

Hypercard

Vivarium

Etoys
Squeak: Slang and Plugins

"Back to the Future":

intermediate + object memory + primitives

hardware + os

simulated image

image

slang

gcc

.C
interpreter
object memory + primitives
Squeak: compilers

Ian Piumarta

Jitter 1: threaded code

Jitter 2: x86 and PowerPC was released with Squeak 2.0

Jitter 3: interesting experiments

Bryce Kampjes

Exupery: bytecode to x86 compiler in the image
Cog => OpenSmalltalk VM

OpenCroquet => Qwaq (later Terf)

Squeak
Pharo
Cuis
Newspeak
Eliot Miranda
OpenSmalltalk VM evolution

Binary tree benchmark

- Interpreter 2005
- Stack 2009
- Cog V1 2010
- Cog V2 2011
- Spur 2014
- Sista future
OpenSmalltalk: object memory

Version 3 ➔ Spur

**Binary representation of object pointers ; x is 1 or 0**

First bits are 00; this is a direct pointer to an object in the heap

First bit is 1; this is a SmallInteger instance (31 bits signed int)

Unused tag

**Spur’s object header**

- s: number of slots
- f: object format
- x: remaining bits
- h: identity hash
- c: class index

**Binary representation of object pointers ; x is 1 or 0**

First bits are 00; this is a direct pointer to an object in the heap

First bit is 1; this is a SmallInteger instance (31 bits signed int)

First bits are 10; this is a Character instance
"The Early History of Smalltalk", Alan Kay, History of programming languages---II, pages 511-598, 1993

"The Design and Implementation of the Self Compiler, an Optimizing Compiler for Object-Oriented Programming Languages", Craig Chambers, PhD Thesis, Stanford University, 1992


Cog Blog, Eliot Miranda