Understanding Software

Paul Klint

#UnderstandingSoftware
Where we are coming from (circa 1800)
Jacquard Looms at Family Heirloom Weavers in Red Lion, PA.

http://youtu.be/NSjmFD6Q7hw
Using this development environment
Preparing input for Jacquard’s loom
For the rest of this story ...
For the rest of this story ...

... visit the Turing’s Legacy exhibition during the breaks or reception

Wednesday, June 20, 2012
Where you can also see the Lego Turing Machine

http://www.legoturingmachine.org/

Credits:
Construction: Jeroen van den Bos, Davy Landman
Film: Andre Theelen
Fast forward to today ...
Integrated Development Environments for All

IDE lovers

IDE haters

Emacs haters

Vim haters
We have created a lot of software
~900,000,000,000 LOC

Source: Booch, 2005
Also: Capers Jones
~900,000,000,000 LOC

Cumulative lines of code

Source: Booch, 2005
Also: Capers Jones
~900.000.000.000 LOC

Cumulative lines of code

Source: Booch, 2005
Also: Capers Jones

Nearly One (European) Billion Lines!
How much is that?
How much is that?

Source: US Geological Survey
How much is that?

Source: US Geological Survey

Wednesday, June 20, 2012
How much is that?

Source: US Geological Survey

Wednesday, June 20, 2012
How much is that?

All water

All drinkable water

All drinkable and accessible water

Source: US Geological Survey

Wednesday, June 20, 2012
How much is that?

Source: US Geological Survey

Wednesday, June 20, 2012
How much is that?

All water

All drinkable water

All drinkable and accessible water

Requires 100L per LOC to fill the smallest sphere

Nice picture, but ...
How much is that?
How much is that?
Print 1 billion lines in point size 10
How much is that?

Print 1 billion lines in point size 10
How much is that?

Print 1 billion lines in point size 10

0.8 times distance to the moon
How much is that?
Print 1 billion lines in point size 10

0.8 times distance to the moon

Wrap Planet Earth 8 times
How can we understand and manage all that software?
How can we understand and manage all that software?
Word cloud of all papers at International Conference on Software Engineering (ICSE) 2012

Credit: Adrian Kuhn
Software Engineering

Different Perspectives
Software Engineering

Traditional Academic View

How can we build new software?
Software Engineering

Realistic View
Software Engineering

Realistic View

Construction costs are just 1-10% of costs of total lifecycle

Wednesday, June 20, 2012
Software Engineering

Realistic View

How can we understand and improve existing software?

Construction costs are just 1-10% of costs of total lifecycle
Software Engineering

Evolution of Research at CWI/SWAT

Application

Foundation

Construction → Evolution
Software Engineering

Evolution of Research at CWI/SWAT

AF+SDF

Wednesday, June 20, 2012
Software Engineering

Evolution of Research at CWI/SWAT

Application

DSL

Foundation

ASF+SDF

Construction

Evolution

Wednesday, June 20, 2012
Software Engineering

Evolution of Research at CWI/SWAT
Software Engineering

Evolution of Research at CWI/SWAT

- DSL
- ASF+SDF
- COBOL

- Algebraic Specification
- Term Rewriting
- Generalized Parsing
- ASF+SDF Meta-Environment
Software Engineering

Evolution of Research at CWI/SWAT

Principles of DSL design
RISLA: Interest-based products

Algebraic Specification
Term Rewriting
Generalized Parsing
ASF+SDF Meta-Environment
Software Engineering

Evolution of Research at CWI/SWAT

- DocGen
- Cluster analysis
- COBOL transformation
- Software Improvement Group

- Principles of DSL design
- RISLA: Interest-based products

- Algebraic Specification
- Term Rewriting
- Generalized Parsing
- ASF+SDF Meta-Environment

Wednesday, June 20, 2012
## Triggers for Change

<table>
<thead>
<tr>
<th>Technology</th>
<th>Usability</th>
<th>Application domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>New ideas did not fit in old approach</td>
<td>Students without formal background</td>
<td>Analysis &amp; transformation of real languages</td>
</tr>
<tr>
<td>Enable many experiments and easy extension</td>
<td>Usable for real programmers</td>
<td>Repository mining</td>
</tr>
<tr>
<td>Connect to external libraries</td>
<td>IDE features for DSls</td>
<td>New DSLs</td>
</tr>
<tr>
<td>Efficient built-in datatypes</td>
<td>Error reporting &amp; recovery</td>
<td>Software Visualization</td>
</tr>
</tbody>
</table>

Wednesday, June 20, 2012
Current Goals

• “One-stop-shop” for
  • Meta-programming
  • Meta-data analysis
  • Visualization
• Lab infrastructure
• Transfer medium

http://www.rascal-mpl.org
Software Engineering

Evolution of Research at CWI/SWAT
Software Engineering

Evolution of Research at CWI/SWAT
Software Engineering

Evolution of Research at CWI/SWAT

- DSL
- Rascal
- Java, PHP, Lua, ...

Wednesday, June 20, 2012
Software Engineering

Evolution of Research at CWI/SWAT

Generalized Parsing
Pattern Matching
Tree traversal
Relational Calculus
Java Bridge
Eclipse Plugin
Software Engineering

Evolution of Research at CWI/SWAT

Application

DSL

Forensics (NFI)
Financial Auditing (PwC)
Gaming (IC3D)
GPUs (VU)
HPC (Bergen)

Rascal

Generalized Parsing
Pattern Matching
Tree traversal
Relational Calculus
Java Bridge
Eclipse Plugin

Java, PHP, Lua, ...

Construction → Evolution

Foundation

Wednesday, June 20, 2012
Software Engineering

Evolution of Research at CWI/SWAT

- DSL
- Analysis
- Refactoring
- Repository Mining
- Java, PHP, Lua, ...
- Forensics (NFI)
- Financial Auditing (PwC)
- Gaming (IC3D)
- GPUs (VU)
- HPC (Bergen)
- Rascal
- Generalized Parsing
- Pattern Matching
- Tree traversal
- Relational Calculus
- Java Bridge
- Eclipse Plugin
## Rascal Results so far

<table>
<thead>
<tr>
<th>Technology</th>
<th>Usability</th>
<th>Application domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types, locations, resources, visit, pattern-directed invocation, ...</td>
<td>Layered design: learn on demand</td>
<td>Java, PHP, Lua analysis &amp; refactoring</td>
</tr>
<tr>
<td>New GLL parsing</td>
<td>Eclipse IDE features</td>
<td>Repository mining</td>
</tr>
<tr>
<td>Native support for visualization</td>
<td>REPL</td>
<td>Metrics</td>
</tr>
<tr>
<td>Libraries Math, Statistics, JDBC, SVN, ...</td>
<td>Interactive Tutor</td>
<td>Various DSLs</td>
</tr>
</tbody>
</table>

Wednesday, June 20, 2012
These considerations motivate today’s topic: Understanding Software
Understanding Software
Understanding Software
Understanding Software
Understanding Software

Mark v.d. Brand: Meta-Modeling
Arie van Deursen: Testing
Mike Godfrey: Artifact Provenance
Jurgen Vinju: Meta-Programming
Stéphane Ducasse: Visualization
Understanding Software

Krzysztof Czarnecki: Variability
Ralf Lämmel: Technological Spaces

Ecosystems
Evolution

Mark v.d. Brand: Meta-Modeling
Arie van Deursen: Testing
Mike Godfrey: Artifact Provenance
Jurgen Vinju: Meta-Programming
Stéphane Ducasse: Visualization

Wednesday, June 20, 2012
Understanding Software

Oege de Moor: Business Intelligence
Joost Visser: Benchmarking Quality
Krzysztof Czarnecki: Variability
Ralf Lämmel: Technological Spaces

Business

Ecosystems

Evolution

Mark v.d. Brand: Meta-Modeling
Arie van Deursen: Testing
Mike Godfrey: Artifact Provenance
Jurgen Vinju: Meta-Programming
Stéphane Ducasse: Visualization

Wednesday, June 20, 2012
Software

Word cloud of today’s topics

Wednesday, June 20, 2012
Incomplete list of former SWAT and UvA colleagues
The SWAT Team
Enjoy
Understanding Software