why is software so hard? and what can we do about it?



Daniel Jackson Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology Cambridge, MA

Accenture · India Delivery Center · November 29, 2007

how's our personal software?

"Cosmotronic Software Unlimited Inc. does not warrant that the functions contained in the program will meet your requirements or that the operation of the program will be uninterrupted or error-free. However, Cosmotronic Software Unlimited Inc. warrants the diskette(s) on which the program is furnished to be of black color and square shape under normal use for a period of ninety (90) days from the date of purchase."

"Cosmotronic Software Unlimited Inc. does not warrant that the functions contained in the program will meet your requirements or that the operation of the program will be uninterrupted or error-free. However, Cosmotronic Software Unlimited Inc. warrants the diskette(s) on which the program is furnished to be of black color and square shape under normal use for a period of ninety (90) days from the date of purchase."

"We don't claim Interactive EasyFlow is good for anything ... if you think it is, great, but it's up to you to decide. If Interactive EasyFlow doesn't work: tough. If you lose a million because Interactive EasyFlow messes up, it's you that's out of the million, not us. If you don't like this disclaimer: tough. We reserve the right to do the absolute minimum provided by law, up to and including nothing. This is basically the same disclaimer that comes with all software packages, but ours is in plain English and theirs is in legalese."

ACM Software Engineering Notes, Vol. 12, No. 3, 1987

Apple

"Except for the limited warranty on media ... software is provided "as is", with all faults and without warranty of any kind..."

Apple

"Except for the limited warranty on media ... software is provided "as is", with all faults and without warranty of any kind..."

Google

"as is, with no warranties whatsoever"

Apple

"Except for the limited warranty on media ... software is provided "as is", with all faults and without warranty of any kind..."

Google

"as is, with no warranties whatsoever"

Microsoft

"substantially in accordance with the accompanying materials, for a period of 90 days..."

is your PC secure?

typical patch size

• 100MB

typical time to download

[,] 10 minutes

average time to infection*

• 4 minutes

* [Windows XP, default firewall settings] Unprotected PCs Fall To Hacker Bots In Just Four Minutes Gregg Keizer; Nov 30, 2004; <u>http://www.techweb.com/wire/security/54201306</u> From: Security Absurdity: The Complete, Unquestionable, And Total Failure of Information Security Noam Eppel; <u>http://securityabsurdity.com</u>

















maybe government's doing better?

US government report, 2006

1

United States Government Accountability Office



Report to Congressional Requesters

March 2006

FINANCIAL MANAGEMENT SYSTEMS

Additional Efforts Needed to Address Key Causes of Modernization Failures

sample failures

-

Navy enterprise resource planning

• \$1B wasted on systems that don't interoperate

NASA financial systems

- [,] after 12 years and \$120M spent, on third attempt expected to cost \$1B
- still cannot produce auditable financial statements

Department of Veterans' Affairs

- · supplies not available for patients due to bad inventory control
- implementation halted after spending \$250M

FBI modernization attempts

reacting to 9/11

- had to send photos of suspected hijackers by fax
- no PCs for most employees, no secure email for images

Trilogy

- new network, thousands of PCs, software system ("VCF")
- contract awarded to SAIC

National Research Council report, 2004

- · agents can't take copies of cases into the field
- no bookmarking or history to help navigation, no sorting

outcome

\$600M later, no system; Sentinel (\$425M) planned for 2009

© Daniel Jackson 2007

maybe critical systems are better?

runaway cannons

South Africa, October 2007

- antiaircraft cannon kills 9 soldiers and injures 14 others
- cause not known, but software suspected



http://blog.wired.com/defense/2007/10/robot-cannon-ki.html

air-traffic control



A radar system that was supposed to warn low-flying planes of nearby obstacles was plagued with problems and fixed nationwide only after a 1997 fatal airplane crash on Guam, according to a published report. In some cases, programming errors caused the Minimum Safe-Altitude Warning system not to operate over wide areas, including near busy airports such as those in Chicago and Dallas-Ft. Worth. In other cases, **false alarms were so numerous that air traffic controllers placed cardboard over warning speakers to silence the noise**. The Federal Aviation Administration was warned about the trouble after a business jet crashed outside Washington in 1994, but it did not take decisive action to resolve it until after a Korean Air jumbo jet slammed into a hill on approach to Guam in August 1997, killing 228.

AP, Oct 1999; <u>http://ns.gov.gu/guam/indexmain.html</u>

most aviation deaths from "controlled flight into terrain"



how did we get here? dtech/dt and criticality creep

storage costs



magnetic disks, US\$/gigabyte



from Frans Kaashoek and Jerome Saltzer, *Topics in the Engineering of Computer Systems*, to appear.

operating system growth

size in millions of lines of code



from Frans Kaashoek and Jerome Saltzer, *Topics in the Engineering of Computer Systems*, to appear.

texas A&M bonfire





bonfire history

traditional began in 1928

small bonfire at annual football game

grew in size and complexity each year

' in 1990's required crane to erect

November 18, 1999

• collapsed killing 12 people



the collapse





http://www.fayengineering.com/structural.html

fundamental challenges: context, state space, coupling

software as system component

a software system is a component

- interacts with physical environment
- and organizational context of operators & users

sources of defects

- < 3% of software failures due to bugs in code
- >90% from poor understanding of requirements

consequences

- requirements analysis is critical
- not just function, also assumptions





what happened

Airbus A320, Warsaw 1993



- Airbus A320, Warsaw 1993
- aircraft landed on wet runway



- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work



- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



what happened

- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



why
what happened

- · Airbus A320, Warsaw 1993
- · aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled

why

airborne ↔

disabled



what happened

- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled

why

airborne ⇔ disabled

airborne \Leftrightarrow not WheelPulse \Leftrightarrow disabled



what happened

- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



why

airborne ⇔ disabled airborne ⇔ not WheelPulse ⇔ disabled ENV

what happened

- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



why

airborne	\Leftrightarrow	disabled
airborne ⇔ n	<mark>ot Whe</mark> elPulse	$e \Leftrightarrow disabled$
ENV	MAG	CHINE

what happened

- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



why

airborne	\Leftrightarrow	disabled
airborne ⇔ n	<mark>ot Whe</mark> elPulse	e ⇔ disabled
ENV	MAG	CHINE ✓

what happened

- Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



why

airborne	\Leftrightarrow	disabled
airborne ⇔ n	<mark>ot Whe</mark> elPulse	e ⇔ disabled
ENV	× MAG	CHINE ✓

what happened

- · Airbus A320, Warsaw 1993
- aircraft landed on wet runway
- aquaplaned, so brakes didn't work
- pilot applied reverse thrust, but disabled



why

airborne \Leftrightarrow disabledairborne \Leftrightarrow not WheelPulse \Leftrightarrow disabledENV \checkmark MACHINE \checkmark

simplified; for full analysis, see Report on the Accident to Airbus A320-211 Aircraft in Warsaw on 14 September 1993, Main Commission Aircraft Accident Investigation, http://www.rvs.uni-bielefeld.de/publications/Incidents/DOCS/ComAndRep/Warsaw/warsaw-report.html

state space complexity

software systems have huge state space

- in lifetime, small proportion covered
- ' in testing, hardly any covered

implications

- "Program testing can be used to show the presence of bugs, but never to show their absence!"
- often running in uncharted territory

*E.W. Dijkstra, Structured programming (EWD268) http://www.cs.utexas.edu/users/EWD/



mechanical watch

state space is actually large

many cogs, many positions

but rotational symmetry

if works in one position,
 likely to work in others

likely failure mode

[,] cogs wear down or break

unlikely failure mode

design error causes error at 3:05pm



software watch

extract from Harel's watch model

states & transitions

many states

- some symmetry
- but many cases remain

likely failure mode

[,] design flaw in code

unlikely failure mode

, code wears out



a fault tree tool





counting structures

suppose you're building a fault tree analyzer

how many fault trees?

- with n nodes, can make nⁿ⁻² trees
- so for 10 nodes, $10^8 = 100$ million trees
- actually much worse sharing, AND/OR, etc

how about relations?

- ' table with C columns over N elements $2^{\ensuremath{\mathsf{N}^\mathsf{C}}}$ values
- so database with 3 tables, 3 columns, 3 elements has 2⁸¹ values!
- checking 1 billion/sec, would take about 100 million years

alternative to covering states?

"reliability growth modelling"

- determine operational profile
- , pick random inputs weighted by profile

how long to test for?

- · for probability of failure on demand (pfd) of 0.001
- with 99% confidence
- need about 6,600 demands without failure
- rises dramatically if failures have occurred

implication

• need huge number of tests for high confidence

© Daniel Jackson 2007

coupling

what is coupling?

- , when components of a system affect each other
- [,] damages reliability, makes changes hard

physical components

• coupled in simple and predictable ways

software components

 coupled in complex and unpredictable ways





USS Yorktown, 1997

what happened

- bad data entered into spreadsheet
- divide-by-zero crashes application
- entire network went down
- ship dead in water for 3 hours



Government Computer News / July 13, 1998 Software glitches leave Navy Smart Ship dead in the water Gregory Slabodkin, <u>http://www.gcn.com/print/17_17/33727-1.html</u>

dependences in internet explorer



graph from <u>http://www.spinellis.gr/blog/20031003</u> for Netscape story see: Competing on Internet Time: Lessons From Netscape & Its Battle with Microsoft by Michael A. Cusumano and David B. Yoffie gdi32.dll

user32.dll

what can we do? now and future



	today	future
requirements	pay attention to context explicit modelling designations & definitions	end-to-end arguments
state space	simplicity automated testing	model-based exploration
decoupling	safe languages data abstraction dependence diagrams	dependency management

pay attention to context

construct a context diagram

- · all flows in and out of the system
- all users, operators, stakeholders



explicit modelling

construct lightweight, precise models

object models are most useful



designations & definitions

be clear about the meaning of terms

- designations: connect requirements to the world
- definitions: new terms from old ones

example

- a designation: shelved(b): book b is on a shelf in the library
- a definition: shelved(b) = owned(b) and not onLoan(b)

recommended reading

Software Requirements and Specifications: A Lexicon of Principles, Practices and Prejudices. Michael Jackson. Addison Wesley, 1995.



end-to-end arguments





simplicity



"I gave desperate warnings against the obscurity, the complexity, and overambition of the new design, but my warnings went unheeded. I conclude that there are **two ways** of constructing a software design: **One way is to make it so simple there are obviously no deficiencies and the other way is to make it so complicated that there are no obvious deficiencies**"



Tony Hoare, Turing Award Lecture, 1980

"Simplicity does not precede complexity, but follows it"



automated testing



write your tests so they can be **automated** exploit code to **generate** as many tests as you can

● ● ●			Java – FilterTreeT	est.java – Eclip	se SDK – /Users	/dnj/eclipse	
] ॐ • • • • • • • • • • • • • • • • • •	ا 🖨 [🕹 • 🕞 🖉 🖢	• 🔆 •	╚╸ᠿ∙╺े∙				
Package Explorer Hierarchy 🔂 JUnit 😒	🗆 🗐 🚺 List	t.java	NonEmptyList.java	J Set.java	J ListSet.java	🔊 Main.java	FilterTre
Finished after 0.068 seconds 🛛 🕀 🔐 🔂 🛛	🗞 🔒 🗒 🗸 🎽 🛛 po	ackage	filter;		·		
Runs: 1/1 🛛 Errors: 0 🖾 Fail	ures: 0 🛛 🕀 in	nport j	ava.text.ParseExc	eption;			
▼ 📴 filter.FilterTreeTest [Runner: JUnit 3]	pu	stat Date	lass FilterTreeTe ic final int <i>KEY</i> Filter f07, f07b, erTree tree;	= ExifDirect	tory.TAG_DATE1		
			<pre>ic void setUp () f07 = new DateFil f07b = new DateFil f07b = new DateFil f0107 = new DateF f010107 = new DateF f010107 = new Date f010207 = new Date tree = new Filter tree.add (f07); tree.add (f0107); tree.add (f010107);</pre>	ter (<i>KEY</i> , ") lter (<i>KEY</i> , ") ilter (<i>KEY</i> , ") ilter (<i>KEY</i> , eFilter (<i>KEY</i> eFilter (<i>KEY</i>) Tree ();	Jan 1, 2007", 'Jan 1, 2007", Jan 1, 2008", "Jan 1, 2007" "Feb 1, 2007" Y, "Jan 1, 200	"Dec 31, 200 "Dec 31, 2008 ', "Jan 31, 20 ', "Feb 30, 20)7", "Jan 2, 2	07"); 8"); 007"); 007"); 2007");
≡ Failure Trace			ic void testGetSu assertEquals (f01 assertEauals (f01	07, tree.get			
r	Cor	nsole 🖾	<u></u>				
	stermin	nated > Fi	terTreeTest [IUnit] /Syster	m/Library/Frame	works/JavaVM frame	work /Versions / 1.5	0/Home/bin/

simulating 200 years of sunlight

can we do this for software?

www.wilhelm-research.com

Category: Desktop Inkjet Printers

September 25, 2006 (page 3 of 8)

Epson Stylus Pro 3800 – Print Permanence Ratings (preliminary¹)

Black-and-white prints made with Epson UltraChrome K3 inkset and the "Advanced Black and White Print Mode"

Note: The Display Permanence Ratings given here are based on long-term testing with the previous generation of UltraChrome inks. WIR testing to date with UltraChrome K3 inks indicates that significant increases in Display Permanence Ratings for black-and-white prints can be expected because the three-level, highly-stable carbon pigment based black inks in the UltraChrome K3 inkset largely replace the cyan, magenta, and yellow color inks in B&W prints when they are made with the "Advanced Black and White Print Mode." Very high stability inks such as these require extended test times; tests are continuing and this webpage will be updated regularly. "> 150 Years" means "greater than 150 years," and that tests are continuing.

Pete Turner and his wife Reine at their home in Wainscott, New York with an Epson print of "Fronds" printed with UltraChrome K3 inks and Epson Premium Luster Paper. Long known for his photographs made with Kodachrome film, Turner used a Nikon D1X digital camera to take this photograph in 2004.



Display Permanence Ratings and All	oum/Dark St	orage Perma	nence Ratir	I GS (Years Before Noticeal	ble Fading and	/or Changes i	n Color Balar	nce Occur) ²
Paper, Canvas, or Fine Art Media Printed with UltraChrome K3 Pigment Inks	Finite Framed Finite Framed Round at 75 F & 50					Resistance to High Humidity ⁽⁸⁾	Resistance to Water ⁽⁹⁾	
Epson Premium Glossy Photo Paper (250)	>135 years	>135 years	>76 years	>300 years	now in test	very high	high	no
Epson Premium Luster Photo Paper (250)	>95 years	>218 years	>58 years	>200 years	now in test	very high	high	yes
Epson Premium Semimatte Photo Paper (250)	>76 years	>170 years	>57 years	>200 years	now in test	very high	high	yes
Epson UltraSmooth Fine Art Paper	>205 years	>300 years	>138 years	>300 years	now in test	very high	moderate ⁽¹¹) no
Somerset Velvet for Epson (255 and 505 gsm)	>90 years	>168 years	>60 years	>200 years	now in test	very high	moderate ⁽¹¹) some
Somerset Velvet for Epson w/ PremierArt™ Spray ⁽¹²⁾	>200 years	>200 years	>141 years	>200 years	now in test	very high	moderate ^{(**}) some
Epson Velvet Fine Art Paper	>115 years	>125 years	>112 years	>200 years	now in test	very high	moderate ⁽¹¹) some
Epson Velvet Fine Art Paper w/ PremierArt ^m Spray ⁽¹²⁾) > 178 years	>145 years	>118 years	>200 years	now in test	very high	moderate ⁽¹¹) no
Epson Watercolor Paper Radiant White	>200 years	>200 years	>200 years	>200 years	now in test	very high	moderate ^{(**}) yes
Epson Enhanced Matte Paper ⁽¹³⁾	>110 years	>110 years	>110 years	110 years	now in test	very high	moderate ⁽¹¹) yes
PremierArt™ Water Resistant Canvas for Epson	>105 years	>177 years	>76 years	>200 years	now in test	very high	moderate ^{(**}) no
PremierArt™ Water Resistant Canvas for Epson w/PremierArt™ Print Shield Spray ⁽¹²⁾	>150 years	>196 years	>100 years	>200 years	now in test	very high	moderate ⁽¹¹	o no
PremierArt™ Water Resistant Canvas for Epson w/PremierArt™ Eco Print Shield Coating ⁽¹²⁾		>150 years	>100 years	now in test	now in test	very high	moderate ^{(**}) no

This document originated at <www.wilhelm-research.com> File name: <WIR_Ep3800_2006_09_

model-based exploration with alloy

testing the Galileo fault tree analyzer

- used by NASA on space station
- ' generated 250,000 trees (all 4-event)
- found 8 faults (tool), 3 (spec), 3 (oracle)

Mondex smartcard

- ' developed by NatWest Bank
- formal specification by Logica UK Ltd
- analysis with Alloy by Tahina Ramanandro
- all scenarios in scope of 5 (cards, users, etc)

K. Sullivan, J. Yang, D. Coppit, S. Khurshid, D. Jackson Improving Software Assurance by Bounded Exhaustive Testing International Symposium on Software Testing and Analysis, 2004



achieving decoupling

for system architects

· choose a safe language: Java, not C++

for programmers

- use real data abstractions, not just objects
- all fields should be private

for designers, programmers, testers

- construct a dependence diagram
- identify dependency liabilities
- focus testing on module interactions



learning about dependences

from Parnas's classic paper, 1979

After studying a number of such systems, I have identified some simple concepts that can help programmers to design software so that subsets and extension are more easily obtained. These concepts are simple if you think about software in the way suggested by this paper. Programmers do not commonly do so.

David L. Parnas. Designing software for ease of extension and contraction. IEEE Transactions on Software Engineering, SE-5, 2 (1979)

dependency management

controlling dependen

- tool extracts depende
- checks conformance t

dences ndences from code	org. a	org.	or	ju o	jio	org.apac	org.apache.to	org.apac	org.apache.to	org.apac	org.apache.t	org.ap	org.apach	org.apache	rg.apache.tools.ant.taskdefs.conditio	org.apache.tools.ant.types.selectors	org.apache.tools.ant.taskdefs.email	org.apache.	g.apache.tools.ant.taskdefs.compiler	org.apache.tools	org.apache.tools.ant.taskdefs.cvslib
ce to architecture	org.apache.tools.bzip2	org.apache.tools.mail	org.apache.tools.tar	org.apache.tools.zip	org.apache.tools.ant	org.apache.tools.ant.input	org.apache.tools.ant.util.regexp	he.tools.ant.types	org.apache.tools.ant.util.facade	org.apache.tools.ant.filters	org.apache.tools.ant.filters.util	org.apache.tools.ant.util	org.apache.tools.ant.helper	org.apache.tools.ant.listener	t.taskdefs.conditio	nt.types.selectors	ant.taskdefs.email	org.apache.tools.ant.taskdefs	taskdefs.compiler	org.apache.tools.ant.taskdefs.rmic	ant.taskdefs.cvslib
org.apache.tools.bzip2	1																				
org.apache.tools.mail		•		an I		_	-	_		_		_		_	-						
org.apache.tools.tar			- 22	-				_		_		_		_							
org.apache.tools.zip	1		4	100					-	-				-							-
org.apache.tools.ant			1		N.	Х		Х				Х	Х		Х	Х	-		1		-
org.apache.tools.ant.input		-			X	i.													1-1	_	-
org.apache.tools.ant.util.regexp		-			XX		x			Х		Х				Х			1	_	
org.apache.tools.ant.types					^		~	X		~		~				^			11		
org.apache.tools.ant.util.facade org.apache.tools.ant.filters		-			х		Х		-			v								-	1
org.apache.tools.ant.filters.util		-			x		^	X X		X		X X					-				
org.apache.tools.ant.util					X		Х	x		~	x	~			Х						
org.apache.tools.ant.helper					X		~	~			~	X			~						
org.apache.tools.ant.listener		Х			X							X									
org.apache.tools.ant.taskdefs.condition		A			X							X X		- 27				X			
org.apache.tools.ant.types.selectors					X			Х				X			x			~			
org.apache.tools.ant.taskdefs.email		Х		=8	X		49-04	X	10-0	-	1.2	X	19-0		-	and an	12 au		1997 - 19		
org.apache.tools.ant.taskdefs	X	~	X	X	X	Х		X	X	Х	Х	X			х	X	x		Х	х	
org.apache.tools.ant.taskdefs.compilers		-	~	~	X			X		~		X			~	~	~	X		~	
org.apache.tools.ant.taskdefs.rmic					X			X				X						X X X			
org.apache.tools.ant.taskdefs.cvslib					X	-		X				X						V			

Lattix's LDM

10

0

and process matters too...

the importance of process

NOAA weather satellite at Lockheed Martin, September 2003





NOAA N-Prime Mishap Investigation, Final Report NASA, September 2004

conclusions

three challenges of software

requirements, state space complexity, coupling

powerful tools we have today

• models, test-case generation, dependency diagrams

in the future

- end-to-end cases
- model-based analysis and code generation
- automated dependency management



for more information

modelling and analysis with Alloy

[•] Software Abstractions, MIT Press, 2006

on requirements

 Software Requirements and Specifications: A Lexicon of Principles, Practices and Prejudices. Michael Jackson. Addison Wesley, 1995.

on decoupling

Designing software for ease of extension and contraction.
 David Parnas. IEEE Transactions on Software Engineering, SE-5, 2 (1979).

on programming

- [,] *Programming Pearls*. Jon Bentley. Addison Wesley, 1989.
- [•] Effective Java. Joshua Bloch. Addison Wesley, 2001

© Daniel Jackson 2007



ogic, Language, and Analysis

Software Abstractions