

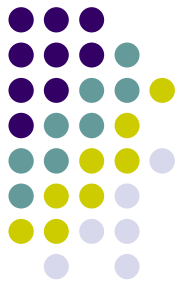
SPREADSHEET TESTING



Presentation to ISACA Northern England
June 23, 2004

Patrick O'Beirne
Systems Modelling Ltd

www.sysmod.com



Current concerns

- The ranks of “sorcerer’s apprentice” user-programmers will swell rapidly, giving many who have little training or expertise in how to avoid or detect high-risk defects tremendous power to create high-risk defects. (Barry Boehm)
- FDA 21 CFR Part 11 Warning letters on spreadsheet validation (LabCompliance.com)
- UK Customs and Excise (can enforce compliance!) The largest VAT error found in a spreadsheet in 2001 was around one million pounds.

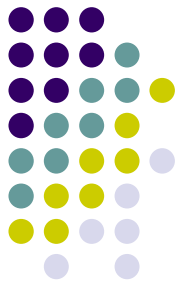
So, what’s the problem?



What's the problem?

- Spreadsheets not taken seriously
- Used by “near experts” → overconfidence
- Anybody can use them ... and change them ...
- Usage from Q&D to important decisions
- <http://panko.cba.hawaii.edu/ssr> Ray Panko (Errors in 91% of 54 spreadsheets, cell error rate 2% - 12.5%)
- www.eusprig.org/stories.htm European Spreadsheet Risks Interest Group

So, how can we test spreadsheets?



Different kinds of testing

- What is your role?
- checking against business requirements
- replicate a “shadow” model
- running test case scenarios
- auditing tools

Next.. how to get test data?



Test data sets

- Obtain reference data sets and corresponding reference results from a reputable source (“oracle”)
- Simple data sets – all 0, all 1; invalid, empty, boundary, extreme data
- Macro to clear all input cells, store standard data sets
- Previously known inputs and proven outputs
- Scenarios, sensitivity analysis

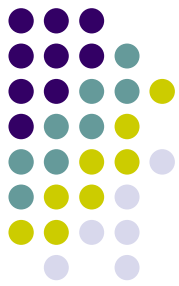
Next...an example of a test harness, XLSior

F12 Failed

Test:	Description:	Id: 4		Status:	Failed				
	Lookups			Run at:	2004-04-10 22:10				
Data:	Comment	Range	Value type	Substitute From	Notes	Data range	Value type	Value	
	Known values in lookup	0.783973247	Range	=A55:B78		MiscCalcs!C11:D34	Range	A55:B78	
	Known values for index	5	Range	=D55:D66		MiscCalcs!F14:F25	Range	D55:D66	
Conditions :	Comment	Value1	Operator	Value2	Outcome	Condition	Result of Value1	Result of Value2	Result of condition
	Test contents of column 1	0.201655819	=	11	Passed	MiscCalcs!G14=A55	11	11	TRUE
		0.438449109	=	12	Passed	MiscCalcs!G15=A56	12	12	TRUE
		0.783973247	=	13	Passed	MiscCalcs!G16=A57	13	13	TRUE
		0.096727468	=	14	Passed	MiscCalcs!G17=A58	14	14	TRUE
		0.111167046	=	15	Passed	MiscCalcs!G18=A59	15	15	TRUE
		0.443094454	=	16	Passed	MiscCalcs!G19=A60	16	16	TRUE
		0.796597038	=	17	Passed	MiscCalcs!G20=A61	17	17	TRUE
		0.689762752	=	18	Passed	MiscCalcs!G21=A62	18	18	TRUE
		0.128488174	=	19	Passed	MiscCalcs!G22=A63	19	19	TRUE
		0.849004106	=	20	Passed	MiscCalcs!G23=A64	20	20	TRUE
		0.319921131	=	21	Failed	MiscCalcs!G24=A65	20	21	FALSE



Test:	Description:	Id: 5		Status:	Passed				
	Sumproduct			Run at:	2004-04-10 22:10				
Data:	Comment	Range	Value type	Substitute From	Notes	Data range	Value type	Value	
	all zero in months	=MiscCalcs!B11:B34	Direct	0		MiscCalcs!B11:B34	Direct	0	
	1 in elt 6	0.5	Direct	1		MiscCalcs!B16	Direct	1	
	all zero in vals	=MiscCalcs!C11:C34	Direct	0		MiscCalcs!C11:C34	Direct	0	
	1 in elt 6	0.320320916	Direct	1		MiscCalcs!C16	Direct	1	
Conditions :	Comment	Value1	Operator	Value2	Outcome	Condition	Result of Value1	Result of Value2	Result of condition
	weighted average	11.92200205	=	1	Passed	MiscCalcs!L11=1	1	1	TRUE



Report of the most recent XLSior tests

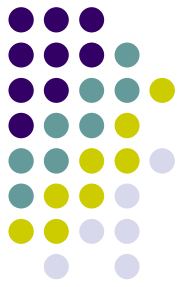


Updated at 2004-04-10 22:10

		2004-04-10 22:10			
	Tests last run at:				
	Number of tests run:	5			
	Number passed:	4			
	Number failed:	1			
Id	Description	Status	Sheet	Cell	
1	Zero wage inflation	Passed	X~TestProject	\$A\$12	
2	dividends of 1 per month	Passed	X~TestProject	\$A\$25	
3	Test construction months	Passed	X~TestProject	\$A\$38	
4	Lookups	Failed	X~TestMisc	\$A\$12	
5	Sumproduct	Passed	X~TestMisc	\$A\$33	

www.XLSior.com

(C) 2004 Patrick O'Beirne, sysmod.com



Usability testing

- Can & do the users access the help – online (in cells, instruction worksheets) or written?
- How adequately does the help resolve different user groups' questions about using the spreadsheet?
- Any user problem comments like *slow, time-consuming, awkward, tedious, intensive, unclear, not sure, confused, bewildering, hard to navigate / follow / find, repetition, duplication*?

Next .. how can we reduce the incidence of errors?

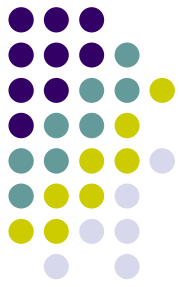


Good practice - design

- Build in testability
- Clear, consistent layout to manage complexity
- Simplify formulas to aid understanding
- Use colour coding for description, not decoration
- Controlled printing, formatted, date/time stamped
- Change control

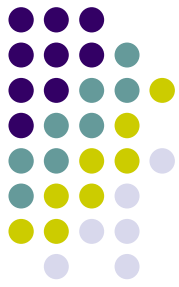
Next .. data validation

Good practice – data control



- Use cross-checks, control totals, and validations
- Provide data entry validation, error trapping
- Data control for externally linked data
- Protect cells and worksheets for authorised users
- Hide unused toolbars, menus

Next: How to go about a model review



Risk assessment, scoping

- Is testing necessary – is the risk worth the effort?
- What evidence is there for the ability of the developer?
- What evidence is there of a mature design methodology?
- Evidence of verification of calculation methods?
- Version control / change management?
- What evidence is there of a test plan?
- Has the model been tested, and regression tested?
- How good is the documentation if any?

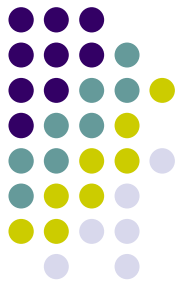
Next: high level review



High level review

- Get an authorised working copy and scan it for viruses
- Auditing tools provide a static analysis and overview
- Look for evidence of correct use of cell/sheet protection.
- Create a diagram of the sheet and database structures to aid understanding.
- Form an opinion of the need for detailed review
- Prioritise risk given your time and resources

Next: detailed review



Detailed review

- Check formulas and their clones for arithmetic and semantic correctness.
- Excel commands: select errors, precedents, dependents
- Auditing toolbar, formula evaluation (Excel 2002+)
- Known error-prone functions: lookup, financial, ...
- Does the model work as expected?
- Checklist includes calculation, precision, circular, range name usage, array formulas, complexity, embedded constants, incomplete structure, mixed units, hidden data, ...

Next: examples of XLAnalyst, SpACE



Help	Summary Potential Risk Report	Results	Info	Example address	Example formula
Factors suggesting a high risk of an error					
?	Circular References	Not Found			
?	Cells Displaying A Number But Storing Text	Found		TextNum!\$C\$4	200, Referenced by: TextNum!\$C\$6
?	Mixed Formulas And Values	Found		FormConst!\$C\$6	=SUM(C3:C5)-50
?	Formulas Evaluating To An Error	Found		FormError!\$C\$6	=C5/C3
?	Vlookups Expecting An Ordered List	Found		VLook3!\$B\$5	=VLOOKUP(B4,vlkSales,2)
?	Hlookups Expecting An Ordered List	Found		HLook3!\$C\$4	=HLOOKUP(B4,hlkSales,2)
Factors suggesting a significant risk of an error					
?	Links To External Workbooks	Not Found			
?	Presence Of Very Hidden Sheets	Found		VHidSht2	
?	Hidden Rows Or Columns	Found		HidRCIRow 4 is hidden	
?	"=+" Construct	Found		EqPlus!\$C\$5	=+C3-C4
?	Conditional Formatting	Found		CondForm!\$C\$3	300
?	Use Of Pivot Tables	Found		PivTble!\$D\$29:\$H\$3	Name : pivRefreshEg
Factors suggesting complex logical modelling					
?	Array Formulas	Found		ArrForm!\$D\$4	{=AVERAGE(IF((H3:H33>0),H3:H33))}
?	Nested If Statements	Found		NestIf!\$E\$13	=IF(NOT(ISBLANK(D13)),IF(D13=\$D\$3,\$E\$3,
?	Use Of Sumif	Found		SumIf!\$D\$5	=SUMIF(H3:H22,C4,K3:K22)
?	Use Of Database Functions (Dsum Etc)	Found		Dfunc!\$D\$5	=DSUM(\$H\$2:\$K\$22,"Value",F5:F6)
?	Use Of Indirect	Found		Indir!\$D\$6	=IF(ISERROR(INDIRECT(D3)),0,INDIRECT(D:
Measures					
?	Longest Formula	Above Limit	115	NestIf!\$E\$27	=IF(NOT(ISBLANK(D27)),IF(ISERROR(VLOO
?	Most Complex Formula	Above Limit	13	NestIf!\$E\$27	=IF(NOT(ISBLANK(D27)),IF(ISERROR(VLOO
?	Total Number Of Formulas	Above Limit	145		
?	Total Number Of Unique Formulas	Above Limit	37		
?	Workbook Size	Within Limit	82 Kb		
?	No Of Worksheets	Above Limit	19		
?	Total All Lines of VBA Code	Above Limit	71/20	71 Lines	In 20 Components
?	Largest Formula Result	Above Limit	3320	Dfunc!\$K\$24	=SUBTOTAL(9,K3:K23)



Directory Analysis

Summary Details

Local Disk (C:) | \$WIN_NT\$.~B | Clients | divelog | DOCS | Documents and Settings | E | EBT | FC | forte4j | IMNnq_NT | Inetpub | j2sdkee1.3.1 | java | jdk1.3.1_02 | Kpcms | MCVCF

Open Spread Sheet | Print Analysis | Preview Analysis | Save

Drag a column header here to group by that column

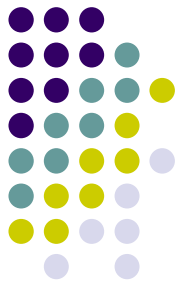
File Name	Status	Size	Sheets	Formulas	Occupied
Eusprig inv ver 1.xls	✓	37888	3	165	365
Ready reckoner.xls	✓	37888	3	216	388
esempioTechniques.xls	✓	44544	2	308	515
test02.xls	✓	53248	12	61	508
DEMOSPCE.XLS	⚠	57856	5	265	945
rubric.xls	✓	64512	3	21	111
2003Eval.xls	✓	81920	4	68	2409
dirxlsf.xls	✓	96256	1	0	4073
Conf2004.xls	✓	99840	6	872	1143
SCANXL.XLS	✓	100864	1	0	645
horrors4,xlAnalyst.xls	✓	104448	19	145	848
Eusprig2003.xls	✓	118784	6	703	2145
Eusprig2003.xls	✓	128000	8	736	2338
aeroparts.xls	✓	137216	11	7	2067
AquaPart.xls	✓	193536	16	98	667
ExcelShortcuts.xls	✓	247808	1	0	1358
nestedif.XLS	✓	396800	1	0	3436
LDUTIL.xls	✓	485888	2	0	415
Total: 44		Avg: 93568	Avg: 4	Avg: 494	Avg: 1244

Analysis Options

- Recurse Sub-directories
- Apply Validation Template
- Customise Grid Columns
- View Analysis History

Close

www.SpreadsheetAuditing.com



Software tools

- [SpACE methodology](#)
- [EXChecker](#)
- [Operis Analysis Kit](#)
- [Spreadsheet Professional](#)
- [Spreadsheet Detective](#)
- [XLAnalyst](#)
- [Code Tracer](#)
- [Rainbow Analyst](#)
- [XDrill drill-down](#)
- [Navigator Utilities](#)
- [xlNavigator flowcharter](#)
- [Synkronizer comp/diff](#)
- [ActiveData data analysis](#)

This list maintained at: <http://www.sysmod.com/sslinks.htm>

Monthly newsletter at: <http://www.sysmod.com/praxis/>



Summary

- What is the risk & potential impact?
- What are you looking for?
- Test at different levels
- Improve through better practice
- Reinforce training through review