The right choice matters!

SMT solving substantially improves model-based debugging of spreadsheets.

Simon Außerlechner, Sandra Fruhmann,
Wolfgang Wieser, Birgit Hofer, Raphael Spörk,
Clemens Mühlbacher, and Franz Wotawa

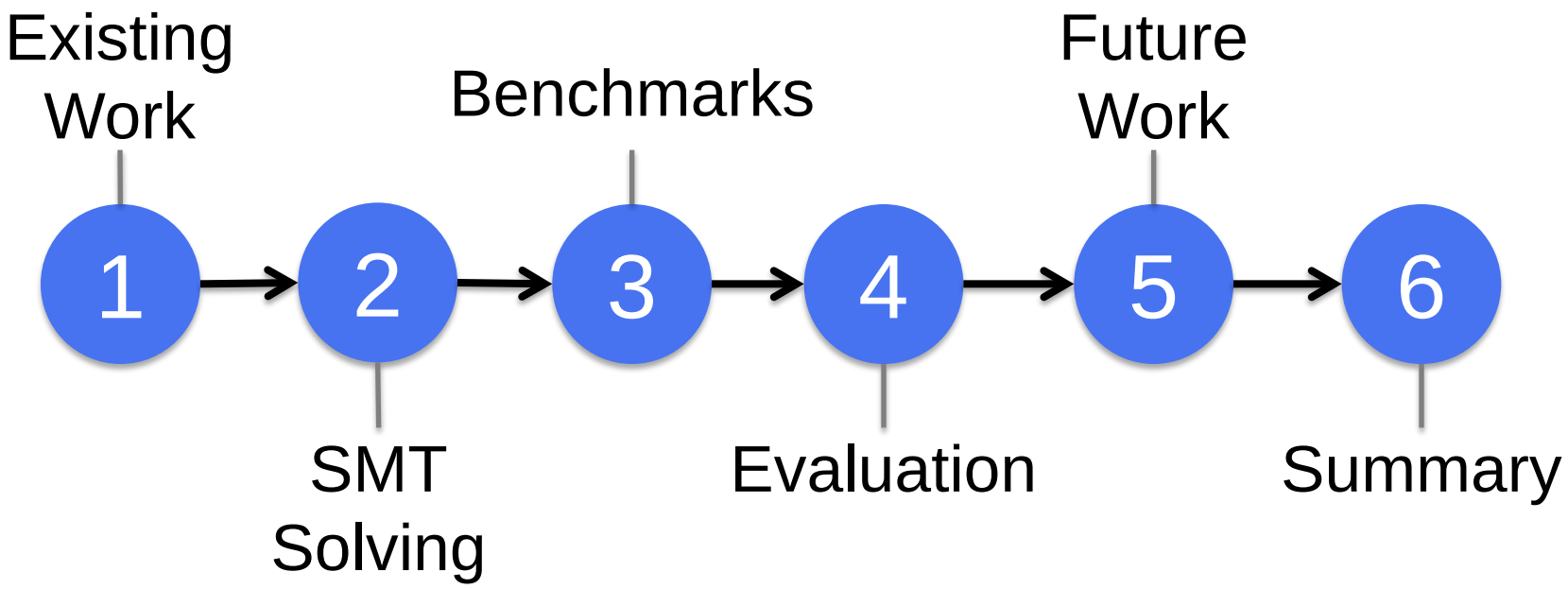
2013 13th Int. Conference on Quality Software (QSIC), July 2013, Nanjing, China

Fault Localization in Spreadsheets

- What is Spreadsheet fault localization?
- Why important?
  - Spreadsheet users outnumber programmers
  - Basis for decisions, but error prone
  - Hard to debug spreadsheets
- Attempts to use model-based debugging
  Limitations
  → No support for real numbers
  → Solving time

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Outline



Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Existing work

- Abreu et al.:“Constraint-based Debugging of Spreadsheets” [ARW12]
  Minion constraint solver
- Jannach and Engler: “Toward model-based debugging of spreadsheet programs” [JE10]
  Choco constraint solver

Limitations
No Support for Real Numbers
Solving Time

[ARW12] Rui Abreu, André Riboira, and Franz Wotawa. “Constraint-based Debugging of Spreadsheets.” 15th Iberoamerican Conference on Software Engineering, 2012, pp. 1–14.

[JE10] Dietmar Jannach and Ulrich Engler. “Toward model-based debugging of spreadsheet programs.” 9th Joint Conference on Knowledge-Based Software Engineering 2010, pp. 252–264.

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Running Example

Faulty Spreadsheet

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	23	15	$345,00
3	Jones	35	34	69	17	$1.173,00
4	**Total**	58	65	92		

Formula View

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	=SUM(B2)	15	=D2*E2
3	Jones	35	34	=SUM(B3:C3)	17	=D3*E3
4	**Total**	=SUM(B2:B3)	=SUM(C2:C3)	=SUM(D2:D3)		

Source: EUSES Spreadsheet Corpus

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Test Cases for Spreadsheets

- **Input cells:** cells that do not reference other cells

  I = {B2=23, C2=31, E2=15, B3=35, C3=34, E3=17}

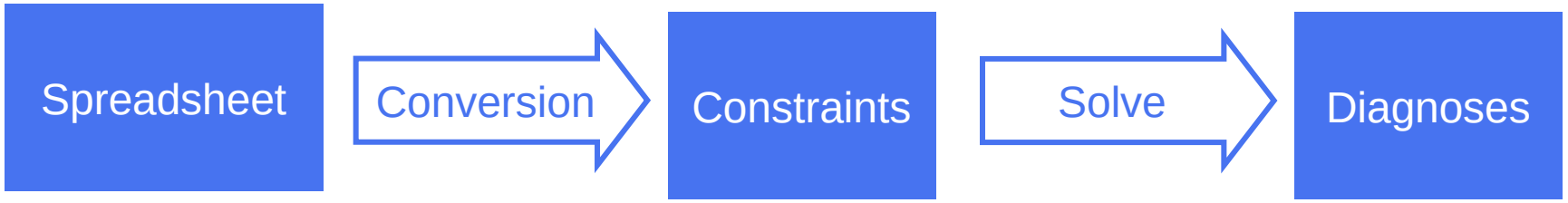
- **Output cells:** any formula cell, determined by user

  O = {B4=58, C4=65, D4=123, F2=810, F3=1173}

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	23	15	$345,00
3	Jones	35	34	69	17	$1.173,00
4	**Total**	58	65	92		

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Constraint-based Debugging (ConBug)



	D
1	**Total**
2	=SUM(B2)
3	=SUM(B3:C3)
4	=SUM(D2:D3)

AB(f_D2) ∨ (D2 == B2)
AB(f_D3) ∨ (D3 == B3+C3)
AB(f_D4) ∨ (D4 == D2+D3)
…

Test case:
D4 == 123
…

For single faults:
SUM(f_D2, f_D3, f_D4)==1

For complete example:

{D2}, {D4,F2}, …

[ARW12] Rui Abreu, André Riboira, and Franz Wotawa. “Constraint-based Debugging of Spreadsheets.” 15th Iberoamerican Conference on Software Engineering, 2012, pp. 1–14.

Satisfiability modulo theories (SMT) solver

- Based on Boolean satisfiability
- Enhanced by different theories (e.g. Integer, Real)
- Solves decision problem
  - Determine diagnoses via MCSes and MCSes-U algorithms [LS08, LS09]
- Z3 solver
  - Speedup in computation time
  - Handling of real numbers

[LS08] M. H. Liffiton and K. A. Sakallah. “Algorithms for computing minimal unsatisfiable subsets of constraints,” J. Autom. Reason., vol. 40, no. 1, pp. 1–33, Jan. 2008.

[LS09] M. H. Liffiton and K. A. Sakallah.“Generalizing core-guided max-sat,” 12th Int. Conference on Theory and Applications of Satisfiability Testing (SAT ’09). Springer-Verlag, 2009, pp. 481–494.

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Spreadsheet Corpora (1)

- Limitations of EUSES spreadsheet corpus [FR05]
  - Many small spreadsheets
  - Missing input values
  - Obsolete file formats
  - No faulty versions

[FR05] Marc II Fisher and Gregg Rothermel. “The EUSES Spreadsheet Corpus: A Shared Resource for Supporting Experimentation with Spreadsheet Dependability Mechanisms.” 1st Workshop on End-User Software Engineering. 2005, pp. 47–51.

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Spreadsheet Corpora (2)

	Mod. EUSES corpus	Integer corpus
Source	Filtering from EUSES	Artificially created + existing spreadsheets
Domain	Reals	Integer
Mutants	622 mutants	220 mutants
	created with spreadsheet mutation operators [AE09]	
Complexity	6 to 4,000 formulas AVG: 225 formulas	7 to 233 formulas AVG: 39 formulas

Original spreadsheets + mutants publicly available

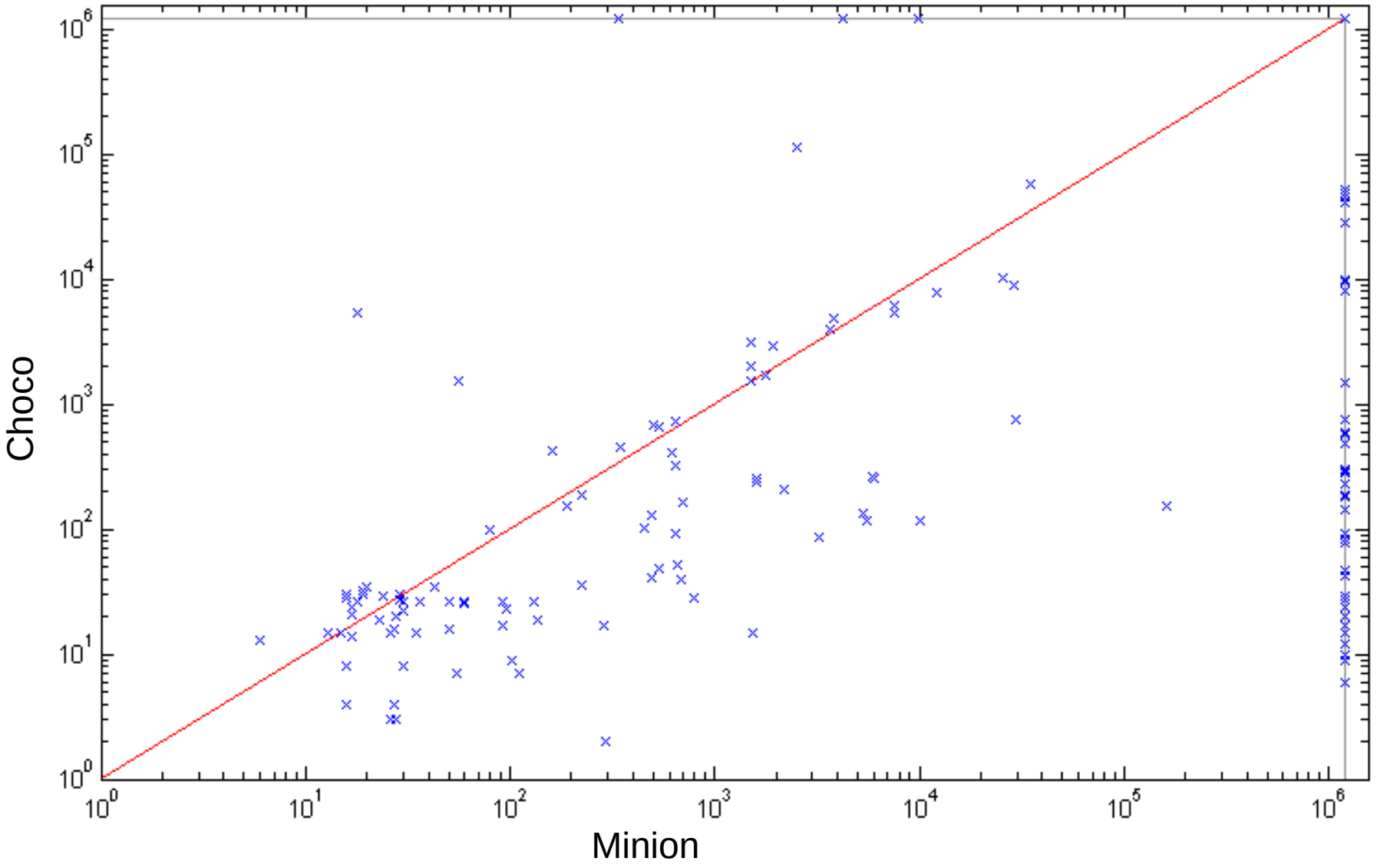
Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Empirical Evaluation

- Runtime comparison
  - Minion vs. Choco vs. Z3 (MCSes and MCSes-U)
    Integer corpus
  - Integers vs. Real numbers
    Integer corpus + Real number version
- Reduction quality
  - Modified EUSES corpus

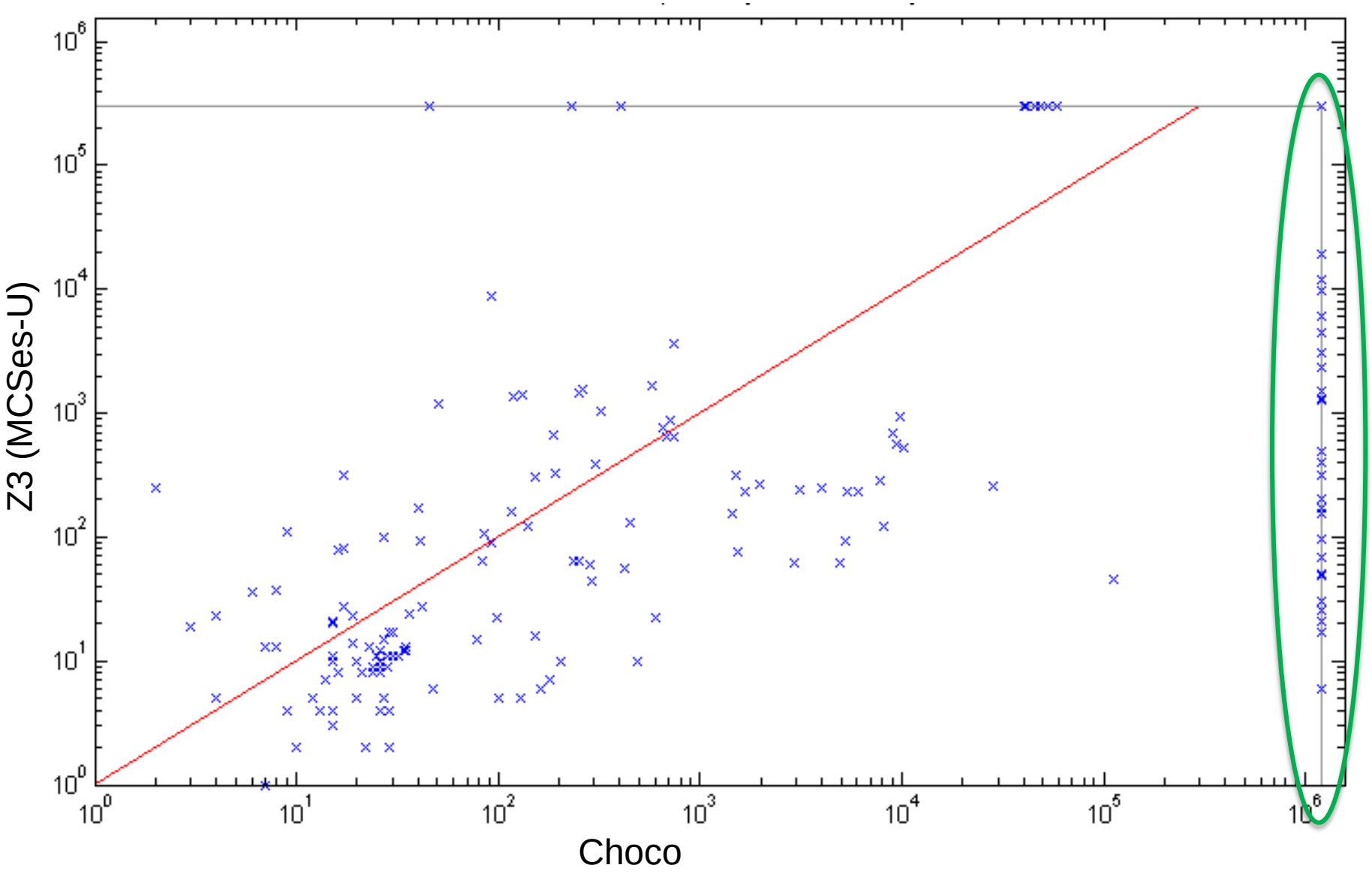
Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Runtime comparison [in ms]: Choco vs. Minion



Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Runtime comparison [in ms]: Z3 vs. Choco



Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Detailed runtime comparison

- Choco slightly faster than Minion
- Z3 faster than Minion and Choco
  - Approx. 6 times faster for Integer corpus
- Integer vs. Real domain
  - Integer spreadsheets approx. 2.6 times faster than Reals

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Reduction quality

- $\text{Reduction} = 1 - \frac{\text{Cells in diagnoses}}{\text{Formula cells}} \times 100\%$
- Evaluation basis
  - Subset of mod. EUSES corpus
  - 183 of 622 spreadsheets
- Results
  - 12 spreadsheets: timeout 5 minutes
  - ~ 80 % reduction
  - ~ 500 ms avg. solving time (Z3)

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Future Work and Open Challenges

- Scalability problem (Hypertree complexity)
- User interface and user acceptance study
- Explanations for the user (traceability)
- Dealing with vague user input

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)

Using Z3 as solver …

… makes it possible to use model-based debugging for spreadsheets

- In practice (Real numbers) and
- For interactive debugging (Solving time)

Außerlechner, Fruhmann, Wieser, Hofer, Spörk, Mühlbacher, Wotawa: „The right choice matters! SMT solving substantially improves model-based debugging of spreadsheets.“ 2013 13th Int. Conference on Quality Software (QSIC)