Detecting Critical Bugs in SMT Solvers Using Blackbox Mutational Fuzzing

[ESEC/FSE ’20]

张灵毓

Outline

1. Background
   - Critical bugs in SMT solver
   - Novelty

2. Approach
   - Fuzzing technique
   - Instance minimization

3. Evaluation

Background

Critical bug

SMT solver : Given a well-formed formula φ in first-order logic, SMT solver decides whether there exists a satisfying solution for φ.

Bugs in SMT solver

Table 1: Classes of bugs in SMT solvers. GT stands for ground truth and SR for solver result.

GT \ SR	sat	unsat	unknown	Crash
sat		A	C	D
unsat	B		C	D

Novelty

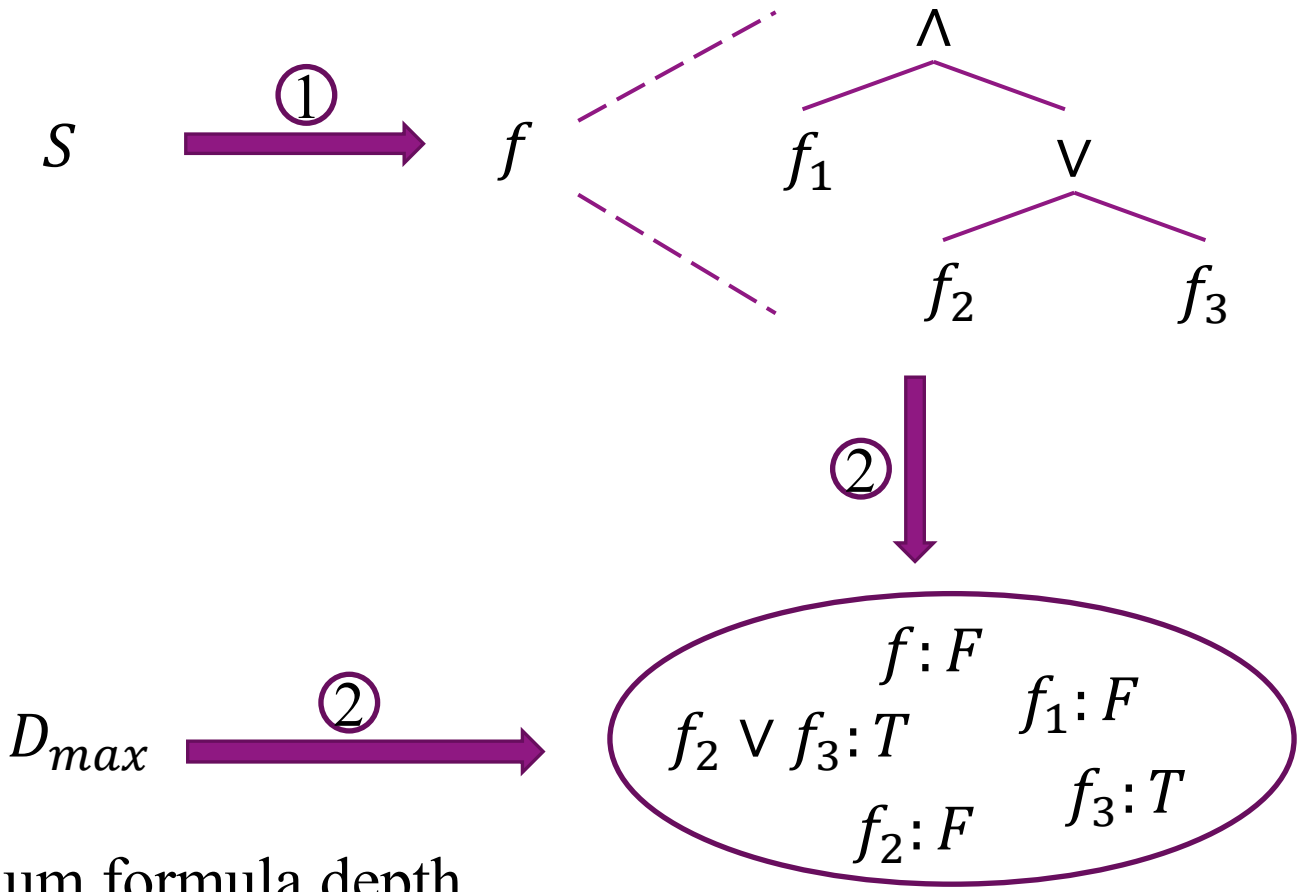
- Does not require a grammar to synthesize instances from scratch.

- Generate satisfiable instances from any given seed.

Approach

Fuzzing technique

Seed fragmentation

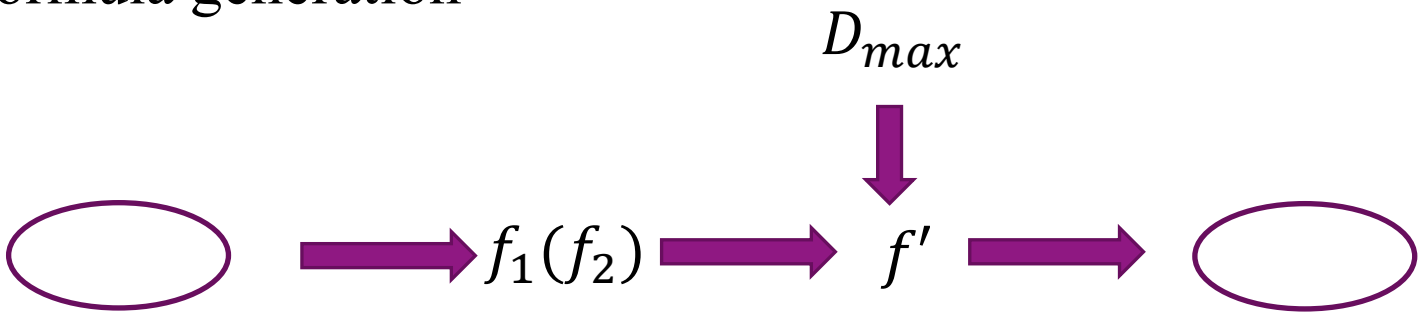


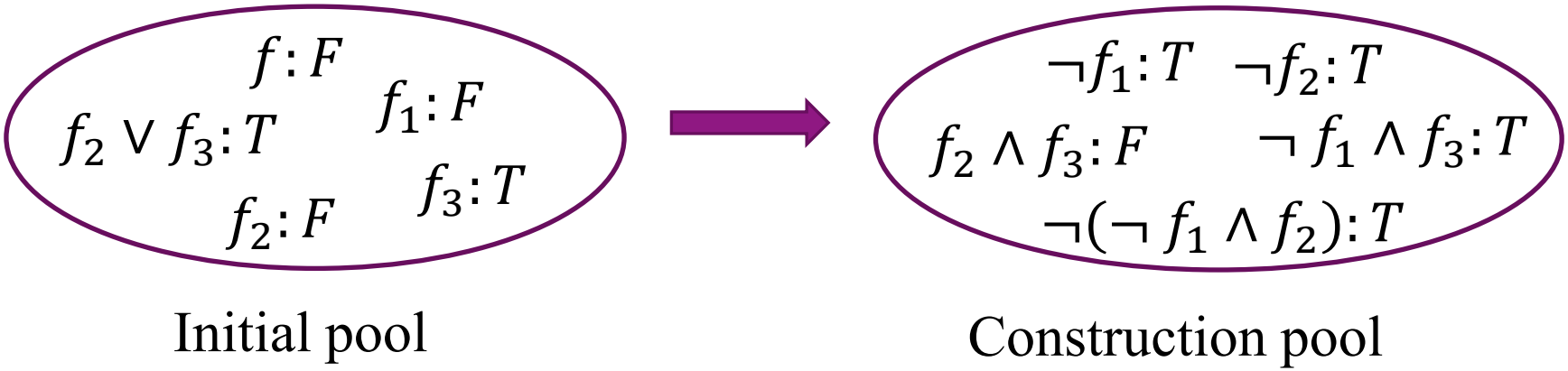
Maximum formula depth

Initial pool

Fuzzing technique

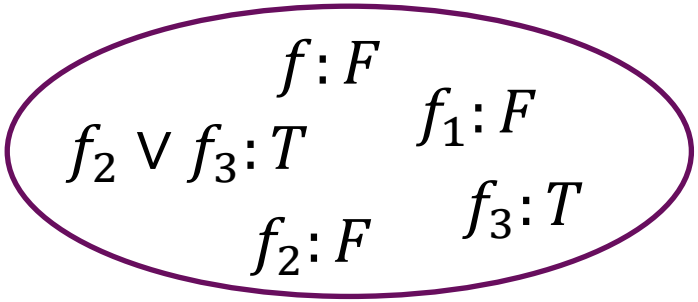
Formula generation



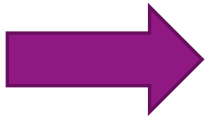


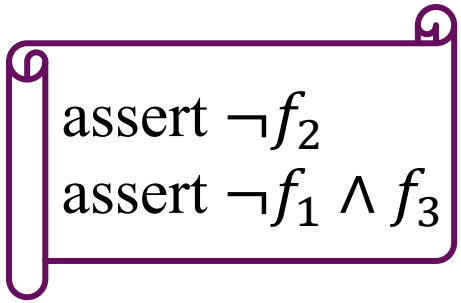
Fuzzing technique

Instance generation

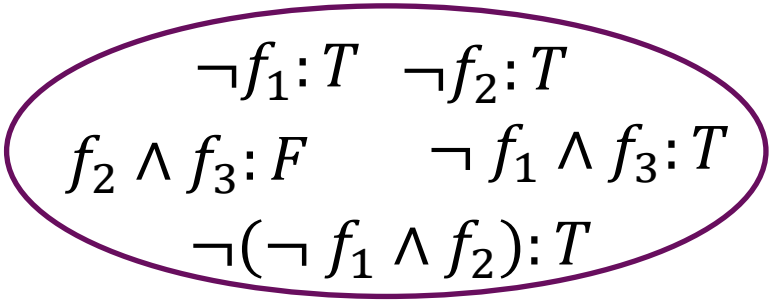


Initial pool





New instance



Construction pool

A_{max}: Maximum assertion number

Instance minimization

Motivation: Fuzzing technique generates large instances, complicating debugging.

Goal: Minimized instance should still reveal bug in buggy solver.

Method: Binary search to find minimized D_{max} and A_{max} .

Evaluation

Evaluation

New critical bugs

Table 3: Previously unknown, unique, and confirmed critical bugs found by STORM in the tested SMT solvers.

SMT Solver	Incremental Mode	Logics	Critical Bugs
MathSAT5		QF_FP QF_BVFP	2
Yices2		QF_UFIDL QF_UF	2
Yices2	✓	QF_UFIDL QF_UFLRA	2
Z3		QF_UFLIA QF_BV UF LIA QF_BVFP QF_LIA	8
Z3	✓	QF_FP QF_S	3
Z3str3		QF_S	6
Z3-AS		AUFNIRA QF_NIA AUFLIRA QF_NRA	4
Z3-AS	✓	AUFNIRA	1
Z3-DBG		QF_NIA	1

Evaluation

Impact of A_{max} and D_{max}

- The larger the easier to find bugs.
- The impact of A_{max} is more significant.

Instance minimization

- More reliable and median reduction of 82.7%
- Treat predicates as atomic blocks (cannot minimized further)

Evaluation

Table 5: Code coverage increase as more instances are generated by STORM.

Generated Instances	Line Coverage	Function Coverage
0	58219	26256
100	66945	30498
200	67063	30524
300	67119	30547
400	67208	30598
500	67759	30861

Thanks