

Exploring the SMT-LIB Benchmark Library

Hans-Jörg Schurr, François Bobot, Mathias Preiner,
Aina Niemetz, Clark Barrett, Pascal Fontaine,
and Cesare Tinelli

TACAS — Turin, Italy
April 16, 2026



What Is It For?

- Used to answer
 - Which benchmarks were solved by other system, but not by us?
 - Does my new method for `str.suffixof` really work?
 - Which solver deserves an award at the SMT competition?

Challenges

- **34 614 311** queries
- **495 177** benchmarks
- How do you answer these questions?

Cataloging the Benchmarks

We provide a **database** with all benchmark metadata we could find!

- A single SQLite file that
 - can used from most languages,
 - or just via raw SQL.
- Around 5.5 GiB of of data
- Benchmark and query metadata!
- Results from 20 years of SMT-COMP!



Brand New Catalog
FREE
for You & ...

Cataloging the Benchmarks

We provide a **database** with all benchmark metadata we could find!

- A single SQLite file that
 - can used from most languages,
 - or just via raw SQL.
- Around 5.5 GiB of data
- Benchmark and query metadata!
- Results from 20 years of SMT-COMP!



2025 Release <https://zenodo.org/records/16290040>

Artifact <https://zenodo.org/records/18192493>

Explore <https://explore.smt-lib.org>

Brand New Catalog
FREE
for You & Me

Part I

Database Structure



Benchmarks are commonly referred to by their file path.

non-incremental /
isIncremental

Benchmarks are commonly referred to by their file path.

non-incremental / UFNIA /
isIncremental logic

Benchmarks are commonly referred to by their file path.

non-incremental / UFNIA / 2019 - Preiner /
isIncremental logic date name

family

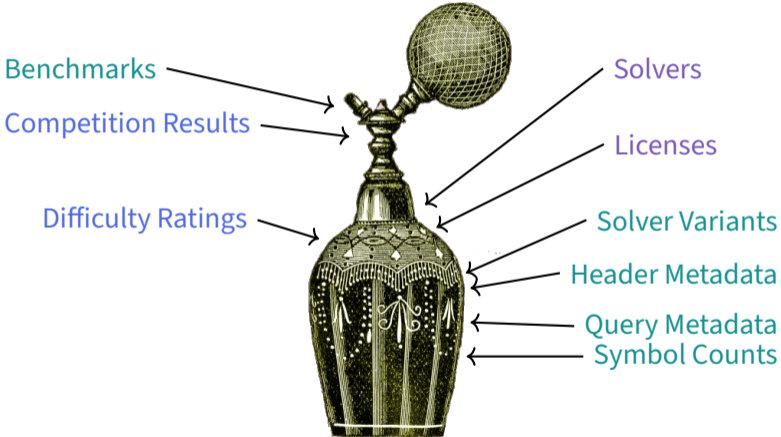
Benchmarks are commonly referred to by their file path.

non-incremental / UFNIA / 2019 - Preiner / partial/t3_rw617.smt2
isIncremental logic date name family name

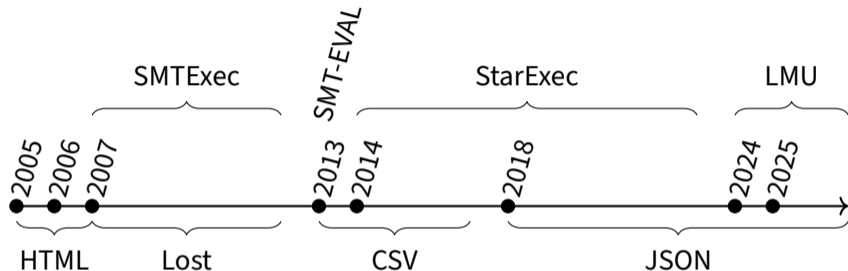
Anatomy of an SMT-LIB Benchmark

```
(set-info :smt-lib-version 2.6)
(set-logic UFNIA)
(set-info :source |
  Generated by: Mathias Preiner
  Generated on: 2019-03-22
  Application: Verifying bit-vector rewrite rule candidates.
  Target solver: CVC4, Z3, Vampire |)
(set-info :license "https://creativecommons.org/licenses/by/4.0/")
(set-info :category "crafted")
(assert [...])
[...]
(set-info :status unknown)
(check-sat)
(exit)
```

What's Inside?



SMT Competitions



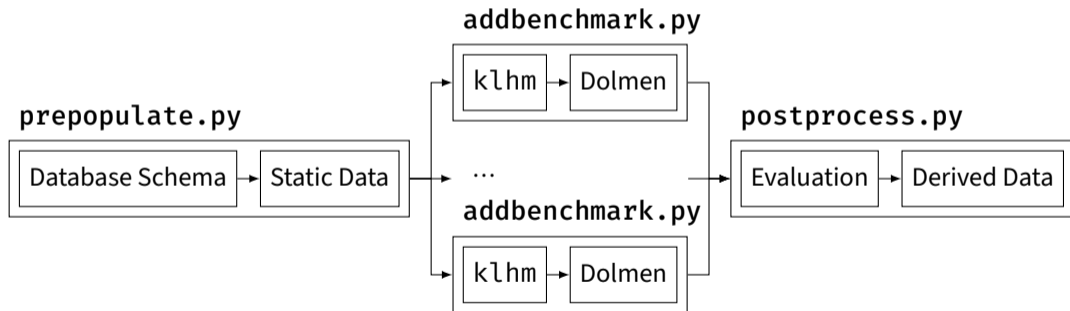
- Available data varies by year
- We often don't have access to benchmark archives!
- We match benchmarks using:
 1. Name
 2. Family
 3. Logic
 4. Incremental?

We built **Klammerhammer** to extract **benchmark metadata**.

- Standalone tool written in Zig
- Produces JSON
 - one object for each query
- Handles push/pop
- The **symbol list** is hardcoded
- Performs a single linear scan, however
- compression uses **zstd** (quadratic!)

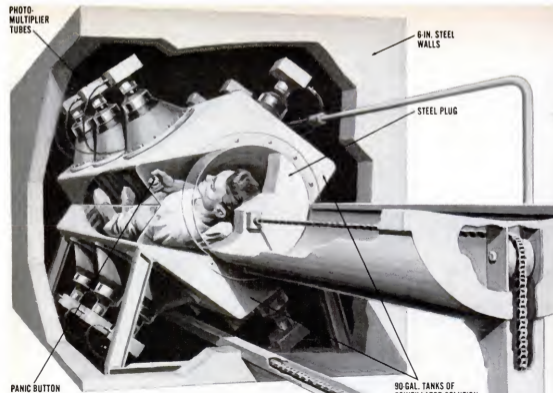
How We Build The Database

Pipeline implemented as a set of Python scripts. Hand-curated data is pure Python.

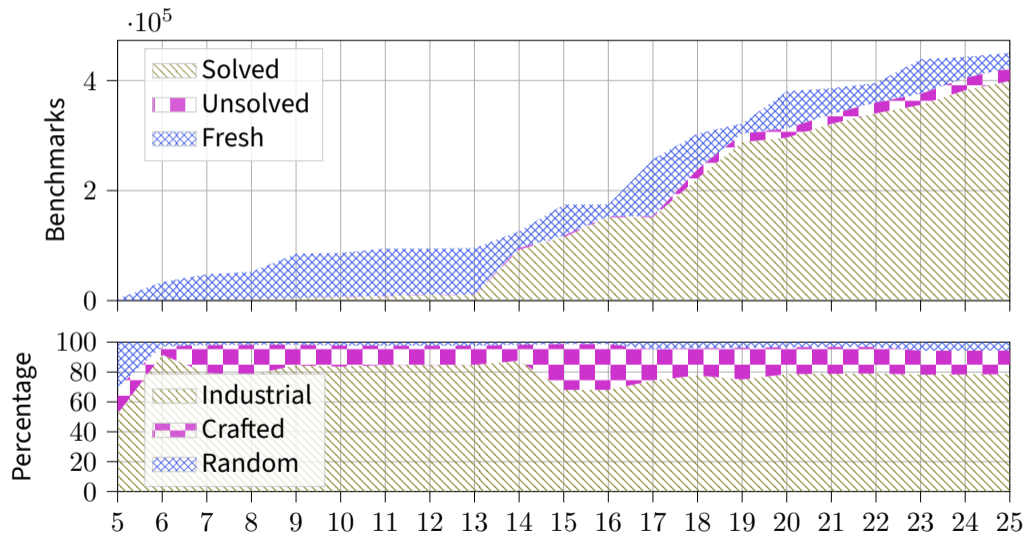


Part II

Case Study



New Non-Incremental Benchmarks



Challenge: how to work with historic competition results?

Changes Between Competitions

- Competition hardware
- Benchmarks
- Timeouts
- Memory limits

Challenge: how to work with historic competition results?

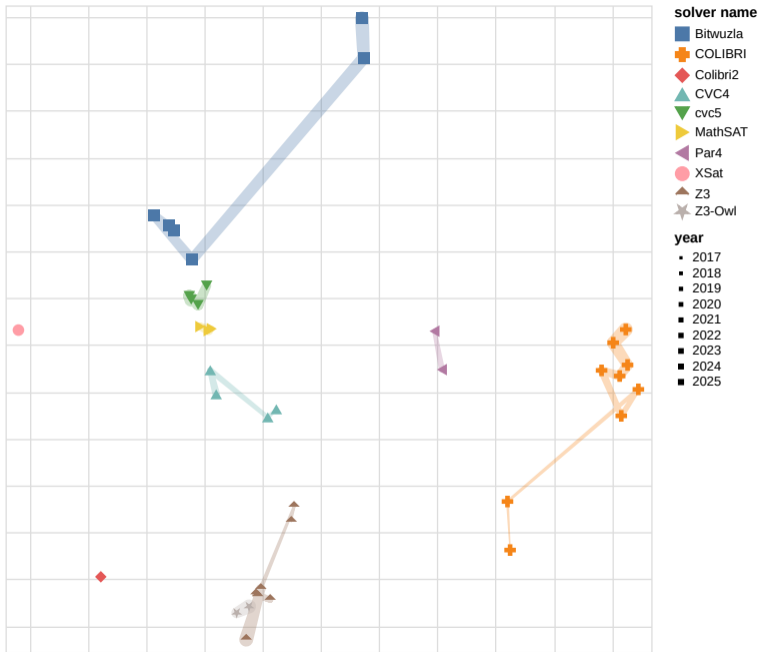
Changes Between Competitions

- Competition hardware
- Benchmarks
- Timeouts
- Memory limits

Example: Isomaps

1. Combine solver variants
2. Identify common benchmarks of all pairs
3. Compute cosine distance of runtimes
4. Compute 2D Isomap: two points are close if their distance is small

QF_FP

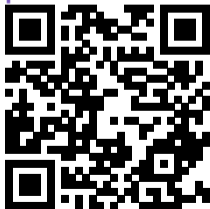


Conclusion

Future Releases

- Yearly release with benchmarks
 - Schema might change
 - Will include new SMT-COMP results
- We want you to play around with it
- and welcome feature requests!

Explore Online



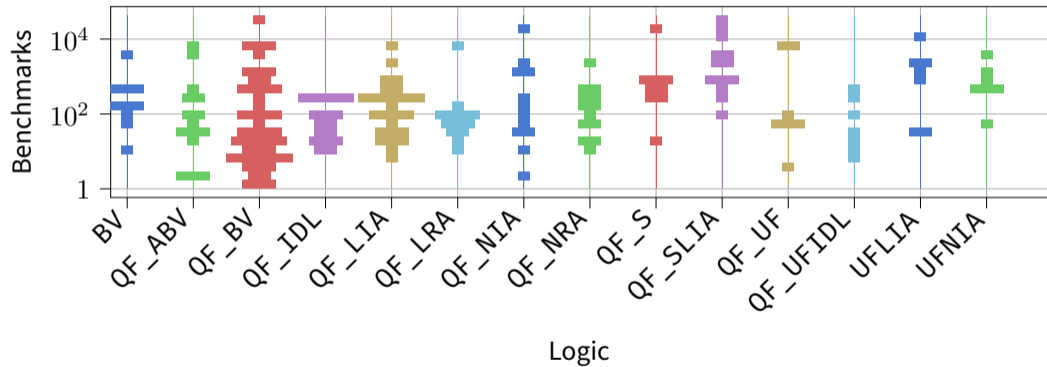
Fun Facts

- Most queries: **2 630 828**
- Largest query: **1.9 GB**
(**48.20 MB** compressed)
- Deepest term: **3 515 188**
- Best solver name:
Tiffany de Wintermonte

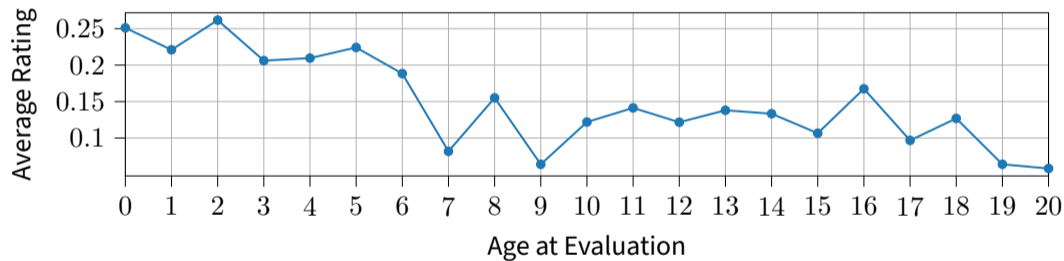


Part III
Bonus Slides

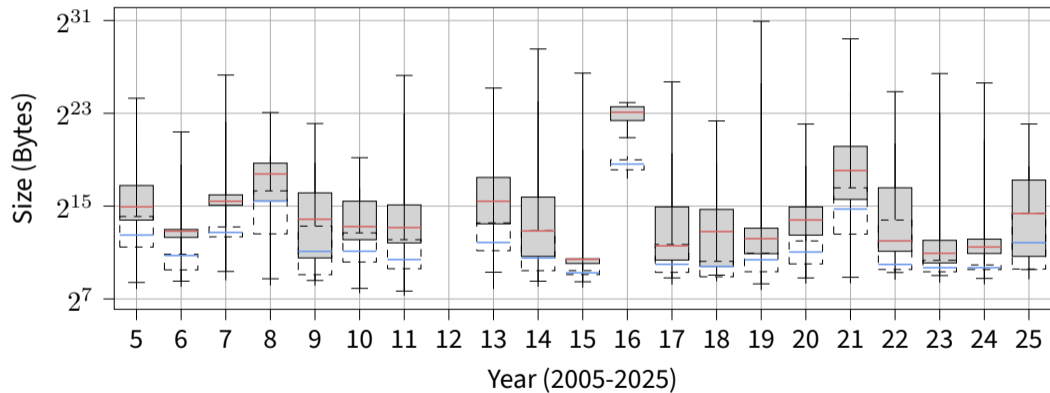
Histograms



Ratings By Age



New Benchmark Sizes



Solvers

