Kex in 2021: Ups and Downs

Azat Abdullin

December 14, 2021

Kex

- white box fuzzer for JVM bytecode
- based on symbolic execution
- test generation for Java and Kotlin

What happened in 2021

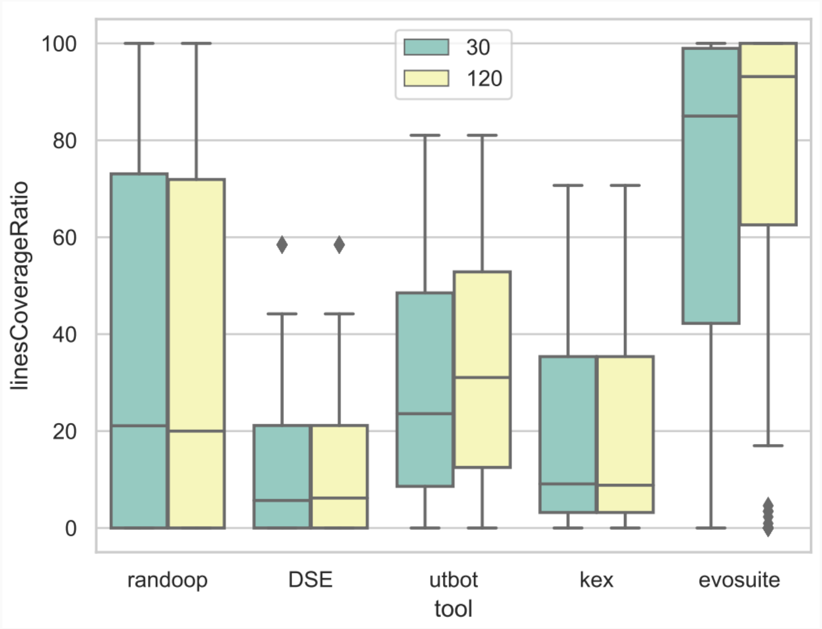
1. Participation in SBST Java tool competition 2021[1]
2. Work towards better Java standard library support
3. Reanimator evaluation

[1]Panichella S. et al. Sbst tool competition 2021 //2021 IEEE/ACM 14th International Workshop on Search-Based Software Testing (SBST). – IEEE, 2021. – C. 20-27.

SBST 2021

- Automatic test case generation competition
- evaluation on 6 projects with 98 benchmarks
- each tool evaluated on 30 and 120 second time budgets

SBST 2021 results



SBST 2021: Kex results[2]

- Kex was ranked fifth with score of 44.21
- Kex achieved any coverage only on one project
    - average coverage of ~20%
- Kex failed on 5 out of 6 projects
    - 1 project failed because of unhandled ASM error
    - 2 projects failed because Kfg encoutered some unexpected bytecode
    - 2 projects failed because Kex required too much RAM
- Kex (and Reanimator) failed on some of the more complex language features (abstract classes, inner classes, etc.)
- Kex required too much of disk space

Main teakeaway: Kex had a low level of maturity

[2] Abdullin A., Akhin M., Belyaev M. Kex at the 2021 SBST Tool Competition //2021 IEEE/ACM 14th International Workshop on Search-Based Software Testing (SBST). – IEEE, 2021. – C. 32-33.

SBST 2021: implications

- Kfg and Kex were optimized w.r.t. RAM usage:
  - Kfg currently uses ~2 times less RAM
  - Kex uses only one copy of each classes of PUT
- Kex and Reanimator were extended to support some new language features
- applied for SBST 2022 Java tool competition

	30s	**120s**
`line coverage`	21.70%	25.29%
`branch coverage`	14.69%	17.95%

CHANGE
AHEAD

Java standard library support

- Java standard library is used almost everywhere
- despite having access to standard library sources, Kex struggles to simulate it
- many of the standard library methods and classes can be approximated in SMT

Intrinsics library[3]

Intrinsics for basic operations and checks:

- `assertions` and `assumes`
- `unknown` values with no constraints
- array operations:
  - `contains` checks
  - array generation methods
- etc.

[3]https://github.com/vorpal-research/kex-intrinsics

kex-rt[4]

Proof-of-concept implementation:

- wrappers for primitive types
- string builders
- some collections (all based on `ArrayList` approximation)
- utility methods from `Arrays` and `System` classes

Kex substitutes all Java runtime operations with kex-rt analogs if they are available

[4]https://github.com/AbdullinAM/kex-rt

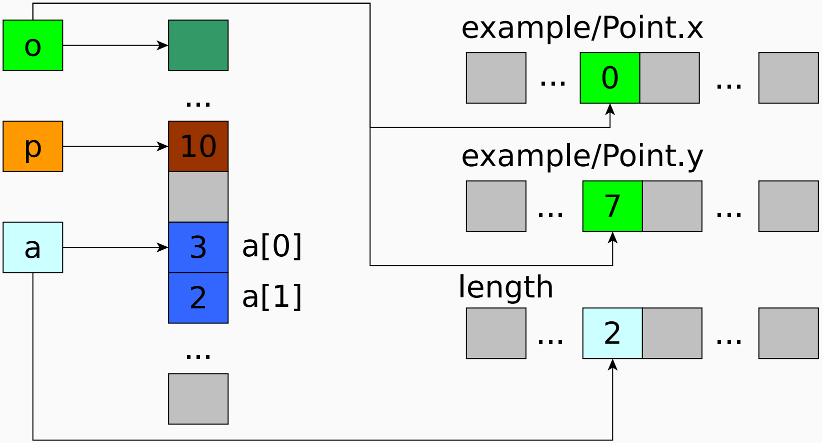
Exmaple of `ArrayList::add` method

```
@Override
public void add(int index, E element) {
  AssertIntrinsics.kexNotNull(elementData);
  int oldLength = elementData.length;
  elementData = CollectionIntrinsics.generateObjectArray(
    oldLength + 1,
    i -> {
      if (i < index) return elementData[i];
      else if (i == index) return null;
      else return elementData[i - 1];
  });
  elementData[index] = element;
}
```

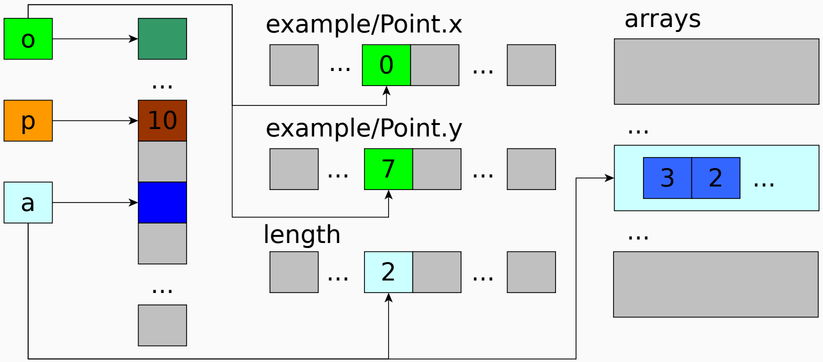
SMT support of intrinsics

- arrays are now represented as SMT arrays
- $\exists$ and $\forall$ quantors for array operations
- λ expressions for array generation
- experimented with SMT string theory (unsuccessfully)

SMT support of intrinsics



SMT support of intrinsics



Java standard library support: takeaways

- prototype implementation
    - limited in expressivness
    - limited number of supported classes
- no thorough evaluation
    - experiments show a small increase in coverage

CHANGE
AHEAD

Reanimator

- an approach to generate valid code snippets using only public API
    - can't produce invalid objects
- works in reasonable time
- applicable in any automatic test generation tool
- can be used in any programming language

Reanimator at the end of 2020

- working prototype
- evaluation:
  - testing with Kex on open source projects from github
  - testing on random objects
- can successfully and efficiently generate 70% of target objects on average

Problem: evaluation is not repreesntative enough

Reanimator: current state

- implemented as part of Tardis[5] tool
- compared with its default test generator — Evosuite[6]

	60s	120s	300s	600s
`tardis + evosuite`	13.96%	15.71%	18.50%	19.60%
`tardis + reanimator`	13.84%	15.99%	17.84%	19.30%
`kex + reanimator`	24.57%	25.29%	25.43%	27.61%

[5] Braione P., Denaro G. SUSHI and TARDIS at the SBST2019 Tool Competition //2019 IEEE/ACM 12th International Workshop on Search-Based Software Testing (SBST). – IEEE, 2019. – C. 25-28.

[6] Fraser G., Arcuri A. Evosuite: automatic test suite generation for object-oriented software //Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering. – 2011. – C. 416-419.

Reanimator: what to do

???

CHANGE
AHEAD

Kex related projects

- **Darya Grechishkina “Loop backstabbing for Kex”**
- **Vladislav Feofilaktov “Spider”**
- Petr Menshov “Effectiveness of paths search algorithms in Concolic Testing”
    - based on the prototype from Andrey Bychkov “Conteau: Concolic Testing Augmented”
- Golubev Kirill “SymFPU for Boolector”
- Viktor Korotkih “KexUI”
- Ramis Sahibgareev “Kfg pass manager”

Conclusion

TODO()

Contact information

azat.abdullin@jetbrains.com

JET BRAINS RESEARCH

POLYTECH
Peter the Great
St. Petersburg Polytechnic
University