



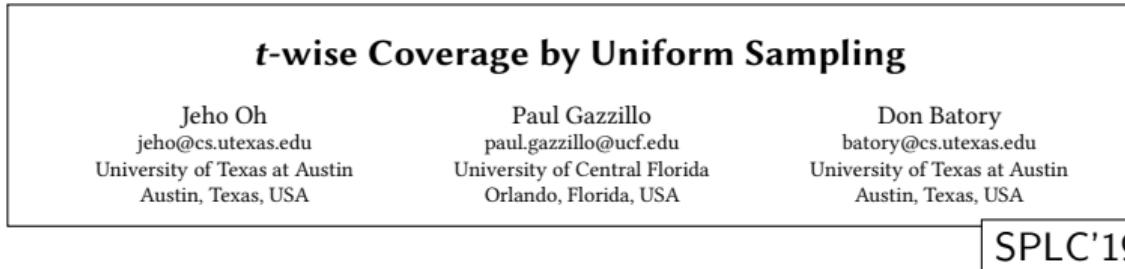
© Thomas Thüm

## UnWise: High T-Wise Coverage From Uniform Sampling

VaMoS'24 | Tobias Heß, Tim Jannik Schmidt, Lukas Ostheimer, Sebastian Krieter, Thomas Thüm | 08.02.2024

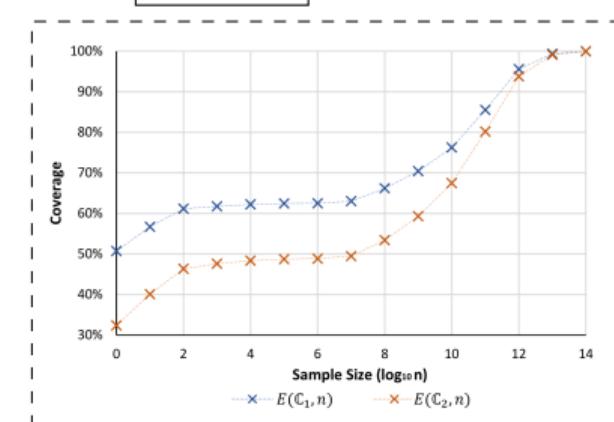
# The Big Picture

## Remember this?

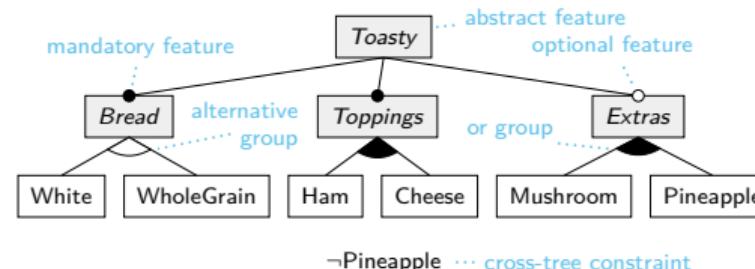


## Our Work

- ▶ better (uniform) samplers
- ▶ post processing
- ▶ larger evaluation, together with YASA and Baital



# Preliminaries



**Valid Configuration:** {White, Cheese, Ham, Mushroom}

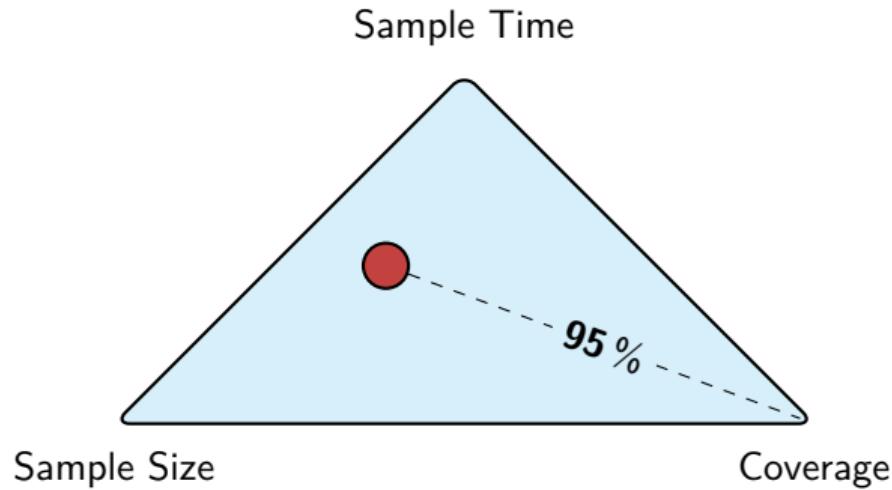
## Uniform Sampling

**Idea:** Draw configurations at random from a *urn of all configurations*.

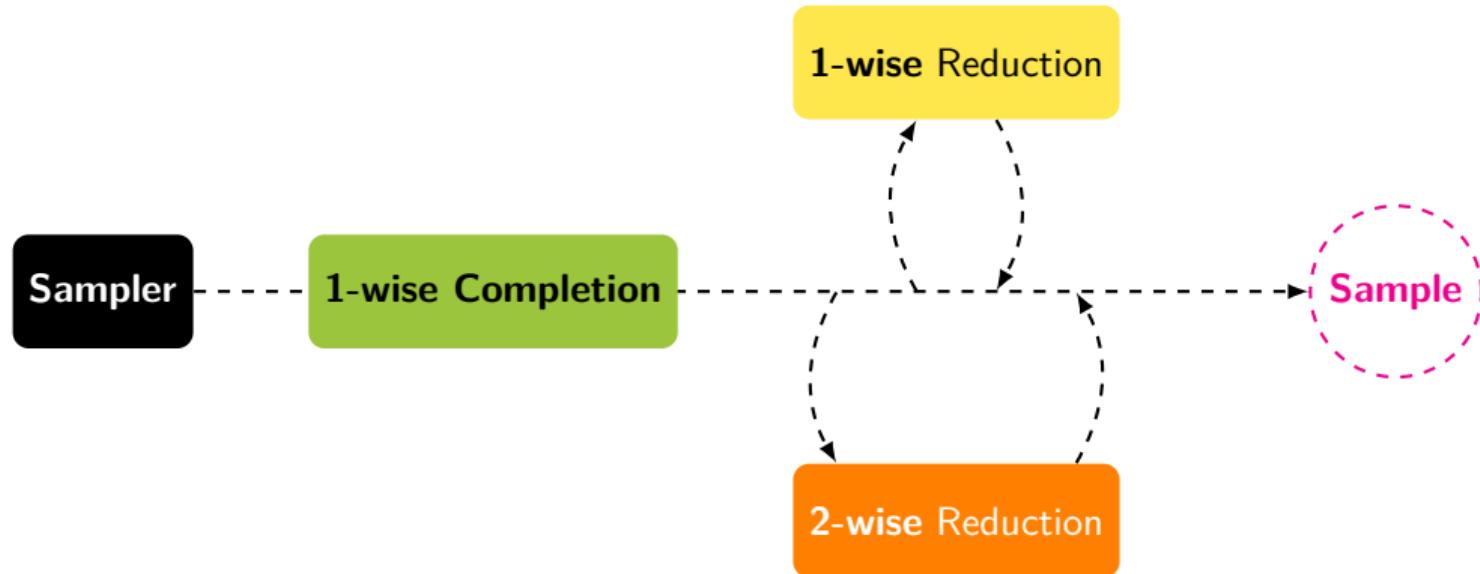
## t-Wise Sampling

**Goal:** Cover all t-wise interactions with at least one configuration

# Idea



# Approach



# Results

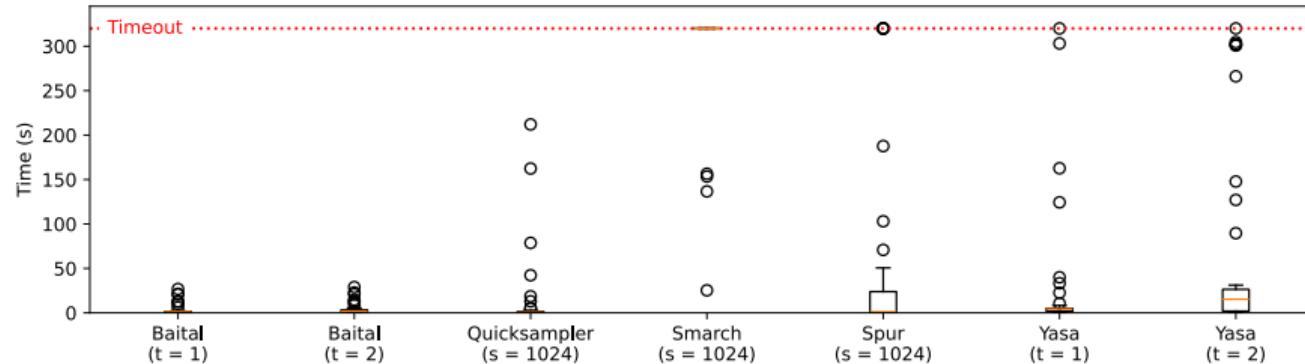
# How and What

- ▶ 300 s timeout
- ▶ (uniform) samplers: 1,024 configurations
- ▶ t-wise samplers:  $t = 1$  and  $t = 2$

## Questions

- ▶ **Scalability**, e.g., success, sampling time
- ▶ **Quality**, i.e., sample size, coverage
- ▶ **Post Processing Effects**, i.e., sample size, coverage

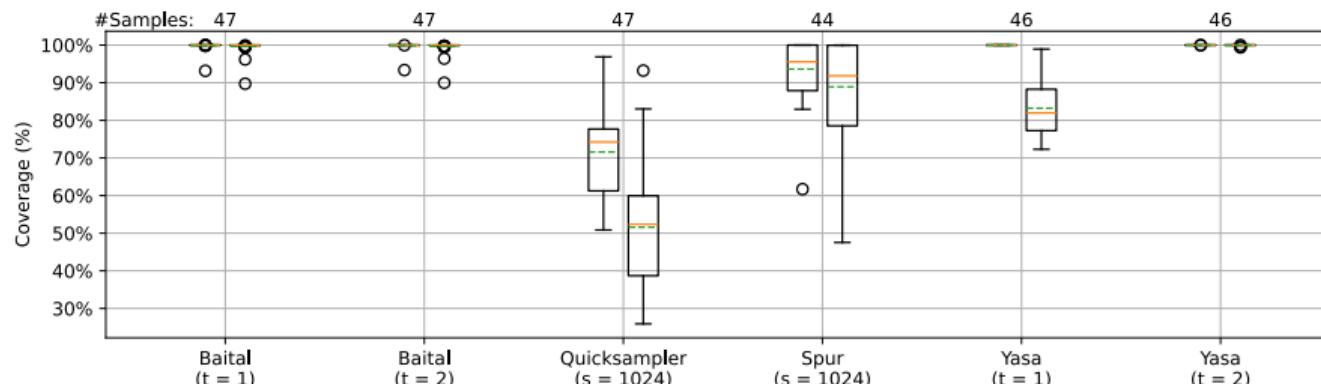
# Sampling Times



- ▶ Smarch times out for 45 / 49 models
- ▶ SPUR, YASA time out for embtoolkit-smarch (need a bit more time)
- ▶ Baital, Quicksampler sample all models

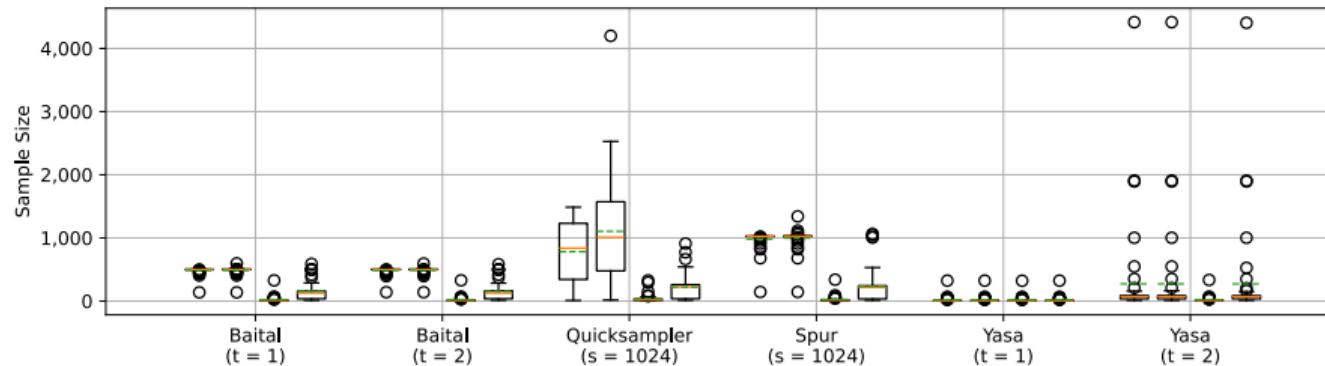
# Achieved Coverages

- ▶ **Smarch** excluded due to bad performance (4 / 49 models)
- ▶ Excluded automotive02v4 and freeetz as counting interactions timed out (24 h)



1-wise and 2-wise coverages achieved without post-processing

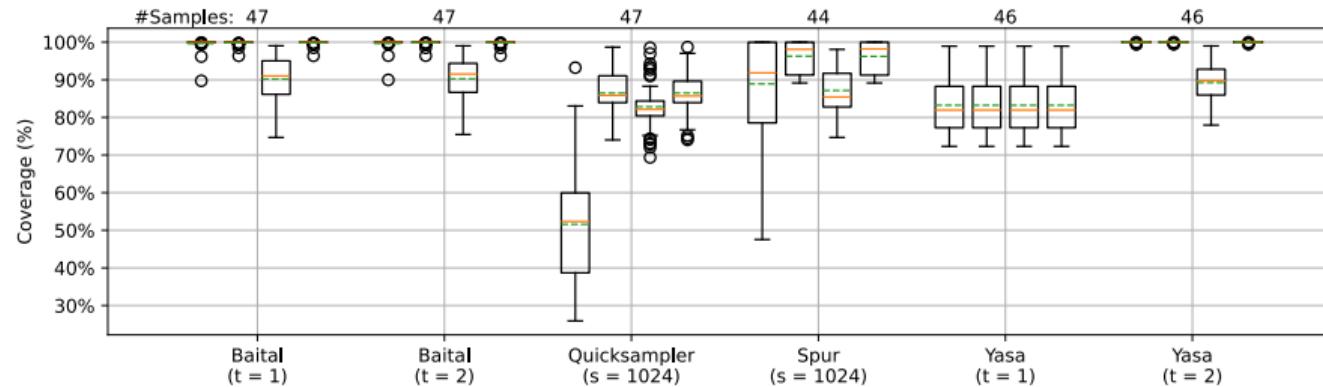
# Sampling Sizes



No post-processing, 1-wise compl., 1-wise compl. + 1-wise red., 1-wise compl. + 2-wise red.

- YASA is always best (= smallest sample size)

# Achieved Coverages



No post-processing, 1-wise compl., 1-wise compl. + 1-wise red., 1-wise compl. + 2-wise red.

## Conclusion

Just use YASA.

# Limited Sample Size

Idea: I can only test  $n$  configurations!

Sampler	No post-processing		Compl. + 2-wise Red.		To size 16
	Size	Cov	Size	Cov	Cov
Baital (t = 1)	484±55.8	99.63±1.6 %	156±155.3	99.83±0.6 %	<b>93.45±8.0 %</b>
Baital (t = 2)	484±55.9	99.65±1.5 %	156±154.7	99.83±0.6 %	<b>93.42±8.0 %</b>
Quicksampler	781±500.4	51.58±16.3 %	221±211.9	86.48±5.9 %	81.70±8.6 %
Spur	982±145.2	88.91±11.7 %	233±285.8	96.23±4.0 %	88.76±9.1 %
Yasa (t = 1)	19±48.2	83.23±6.9 %	19±48.2	83.23±6.9 %	80.88±6.7 %
Yasa (t = 2)	272±744.7	99.98±0.1 %	272±743.2	99.98±0.1 %	<b>93.21±8.2 %</b>

# Future Work

- ▶ More sophisticated post-processing, interaction counting
- ▶ Investigate coverage estimators (e.g., Baranov et al. ICSE'22)

**A Scalable t-wise Coverage Estimator**

Eduard Baranov Universite catholique de Louvain Belgium <a href="mailto:eduard.baranov@uclouvain.be">eduard.baranov@uclouvain.be</a>	Sourav Chakraborty Indian Statistical Institute India <a href="mailto:sourav@isical.ac.in">sourav@isical.ac.in</a>	Axel Legay Universite catholique de Louvain Belgium <a href="mailto:axel.legay@uclouvain.be">axel.legay@uclouvain.be</a>
Kuldeep S. Meel National University of Singapore India <a href="https://nus.edu.sg/~kumee/">https://nus.edu.sg/~kumee/</a>	Vinodchandran N. Variyam University of Nebraska-Lincoln USA <a href="https://www.unl.edu/cse/faculty/varyam/">https://www.unl.edu/cse/faculty/varyam/</a>	

ICSE'22

- ▶ Evolution (e.g., Pett et al. SPLC'23)

# Conclusion

