Testing the tests with Mutalk in 2024

<u>guillermo.polito@inria.fr</u> Guillermo Polito - ESUG'24













Quick About Me

- **Now:** Researcher at Inria Lille
- Pharo Contributor since ~2010

- **Keywords:** compilers, testing, test generation
- **Interests:** tooling, benchmarking, 日本語, board games, batman, concurrency

If any of that interests you, come talk to me!



guillermo.polito@inria.fr @guille





Automated Tests

SetTest >> testAdd

| aSet |
"Context"
aSet := Set new.

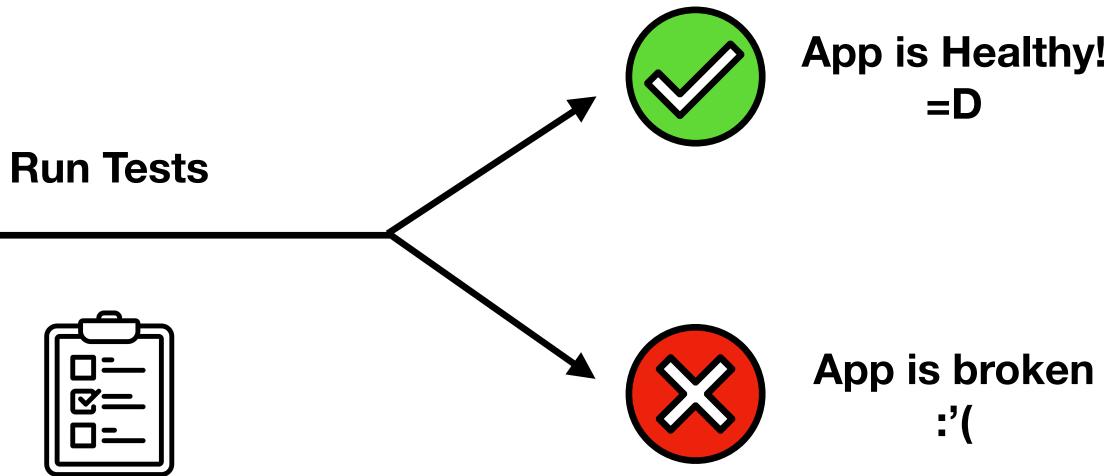
"Stimuli" aSet add: 5. aSet add: 5.

"Check"
self assert: aSet size equals: 1.

in this context when this happens then this should happen

We love tests





What is a good test?

"A good test is a test that catches bugs"

- me







Mutalk: Mutation testing for Pharo

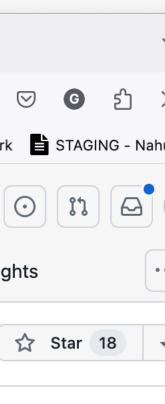
- Presented at ESUG'09

• Pharo 9 to 12 !



• Originally developed in Pharo 1.1 in Argentina (Chillo, Brunstein, Wilkinson)

Pharo-contributions/mutalk: N	lui× +			
C A https://github.	com/pharo-contributions/mu as Argentinas 🎧 Pharo GH 🎧		druid 🌎 mutalk 🚺 T	Ē ̀A ★
OperationPharo-contributionsMutalkodeOIssues7IllPull requests		Q Type 🛛 to searc		 I + → ○ I1 6 Security └< Insights
mutalk Public		🖍 Edit Pins 👻	⊙ Unwatch 7 -	♥ Fork 14 ▼ Star 18
ੇ master 👻 ਿ 1 Branch 🛇 6 Tags	Q Go to file	t H	<> Code -	About
guillep Update mutalkCI.yml 🗙		7a75b07 · yesterday	🕒 560 Commits	Mutation Testing in Pharo Smalltalk
.github/workflows	added pharo12 to Cl		6 months ago	🗘 Readme
src	Update src/MuTalk-CI-Tests	s/MTCITonelJso	5 days ago	-/- Activity
workflow	Update mutalkCI.yml		yesterday	Custom properties
] .project	add repository properties		6 years ago	☆ 18 stars⊙ 7 watching



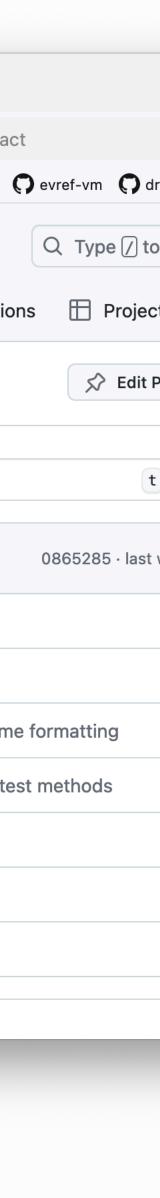
Coverage vs Mutations

30% Code Coverage



4% **Mutation** Coverage

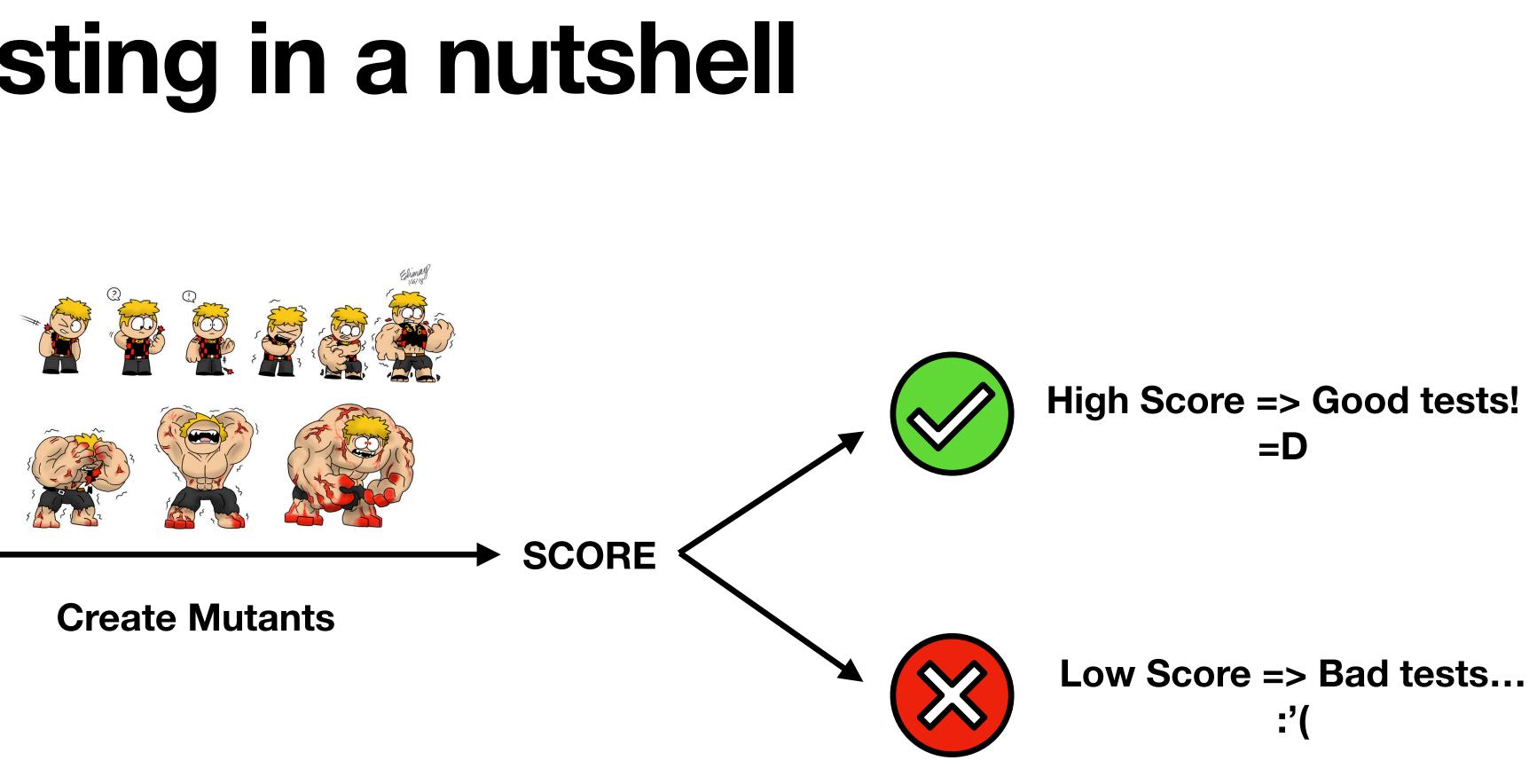
O pharo-contributions/Artefact: ArX	× +
$\leftarrow \rightarrow C$ \bigcirc A https://github.co	om/pharo-contributions/Artefa
🗋 Tech 🗋 Research 🗋 Admin 🗋 Japanese 🗋 Cosas A	Argentinas 🌔 Pharo GH 🌔 vm
pharo-contributions / Artefact	
<> Code ③ Issues 6 13 Pull requests	🖓 Discussions 🕟 Acti
Artefact Public	
우 development 우 2 Branches ♡ 6 Ta	Q Go to file
🏠 astares Update README.md 🗸	
.github/workflows	Update build.yml
.smalltalkci	Add CI and coverage
docs	Add documentation and son
src	PDFParagraphTest has not t
🗋 .project	first backup
LICENSE	Initial commit
README.md	Update README.md
C README MIT license	



Mutation testing in a nutshell







Tests





These kind of mutant

- Introduce artificial bugs
- See if the test suite detects them

• What kind of bugs?



Competent developer hypothesis

• Developers are capable people

- Mistakes/bugs are small details easily overseen. E.g.,
 - Missing +/-1 in a loop
 - An inverted conditional
 - Signed/unsigned

Thousands of simple mutations

- Control flow bugs
- Arithmetic bugs
- Logic bugs
- Overflow bugs
- Typoes

× – 🗆

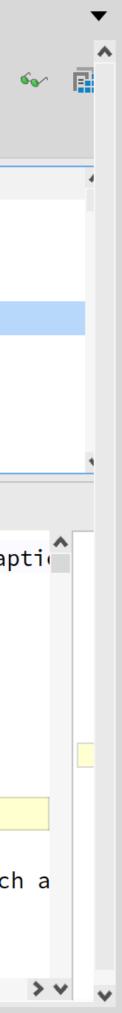
a MTGeneralResult (3090 mut...

Su	irviving Mutants	Killed Mutants	Те			
#	Results					
1	Replace #ifTrue: re	eceiver with false ir	ו Mic			
2	Replace detect: bl	ock with [:each tru	ue] v			
3	Replace #ifFalse: r	eceiver with false i	n Mi			
4	Replace #ifTrue:if	alse: receiver with	fals			
5	Decrease a literal	integer in Package	>>#b			
6	Replace #ifTrue: w	vith #ifFalse: in Base	eline			
7		~=' in MicInlineDeli				
8	Replace #or: with	#and: in MicFileRes	ouro			
	mathblock: a	aString first	_ine			
	colf row	. MathOnoning				
	self raw: MathOpeningBlo					
	aCol do: [:each					
	se	•				
		raw: each ke	y;			
		raw: '='.				
	ea	ch key = #cap	tio			
		ifTrue: [aC				
		ifFalse: ["				
		self raw:				
	separa	atedBy: [sel	t r			
	< COLT DOWN					

Inspector on 3090 mutants, 1599 killed, 1491 alive...

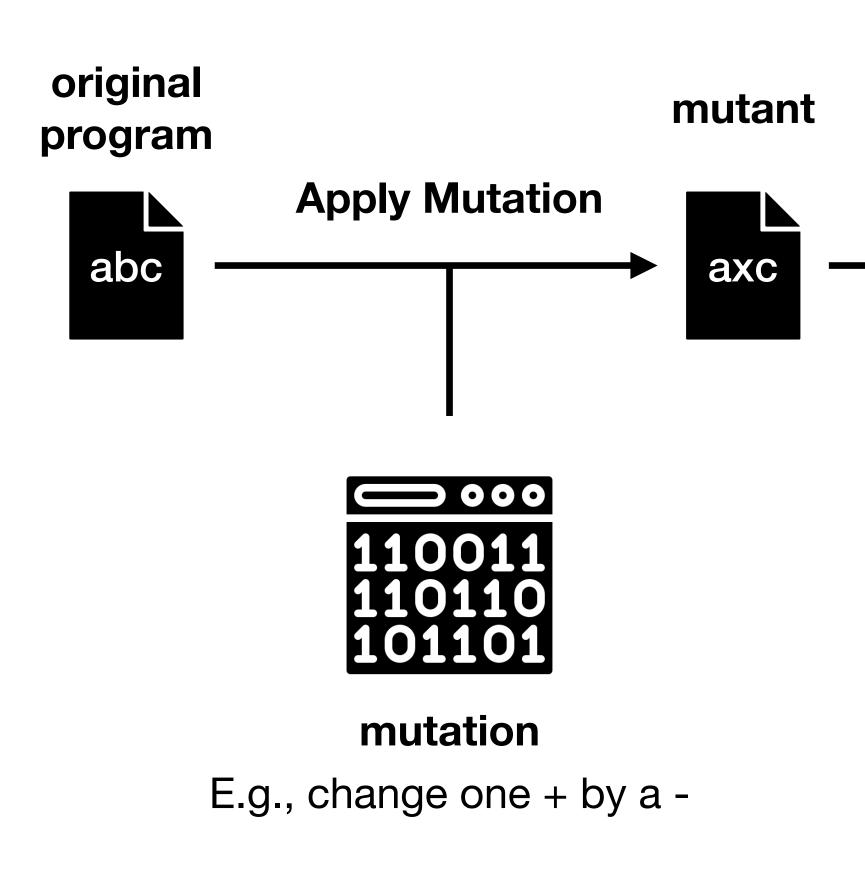
rminated Mutants	Excluded Tests	Raw	Breakpoints	Meta
AnchorBlock>>#pri	ntOn:			
	: in Class>>#docum	entExan	npleCode	
	ss>>#initializeDelimi			
	extualBuilder>>#ma		: firstLineAssocia	tions:wit
uildMicroDownUsin				
	croDownUsing:withC	ommer	nt:	
er class>>#initialize[C			
ceReference>>#load	Children			

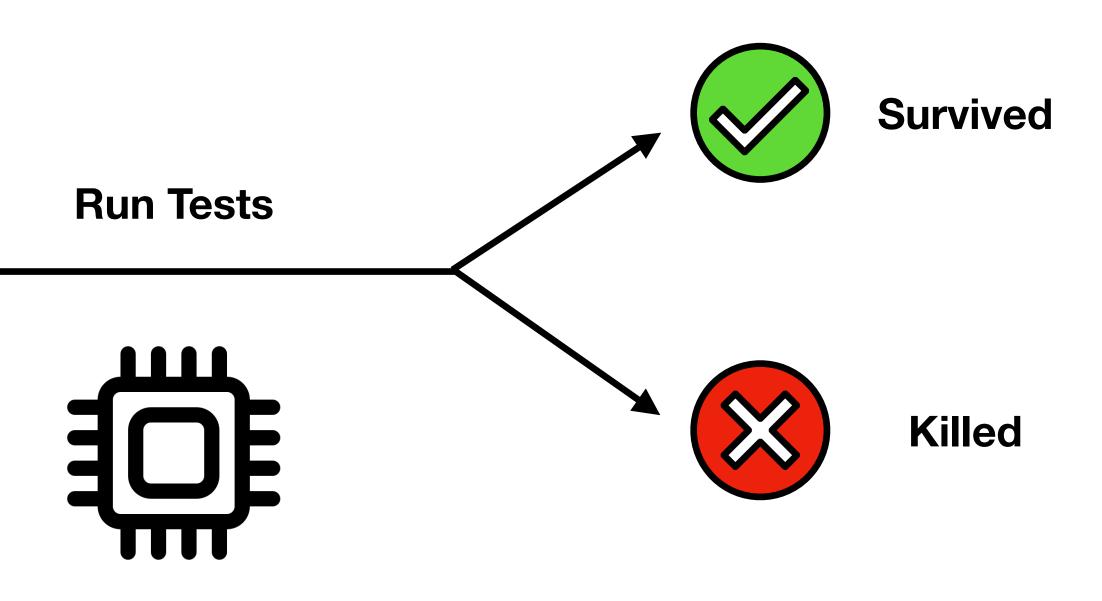
eAssociations: aCol withCaption: a	<pre>mathblock: aString firstLineAssociations: aCol withCa</pre>
ckMarkup.	self raw: MathOpeningBlockMarkup.
	aCol
	do: [:each
	self
	raw: each key;
	raw: '='.
n	false
ionBlock value]	ifTrue: [aCaptionBlock value]
that we can put format such as bo [·]	ifFalse: ["so that we can put format suc
ch value value]]	self raw: each value value]]
aw: '&'].	<pre>separatedBy: [self raw: '&'].</pre>
	<pre>solf nowLine</pre>



💙 💥 🚺

Mutation Analysis





The insight

• Survived mutants were either

- 1. not covered
- 2. not asserted
- 3. or semantically equivalent. E.g. A+B=A-B if B=0 always



Mutation Score

Run each mutation independently

• Score:



#Killed #Mutants

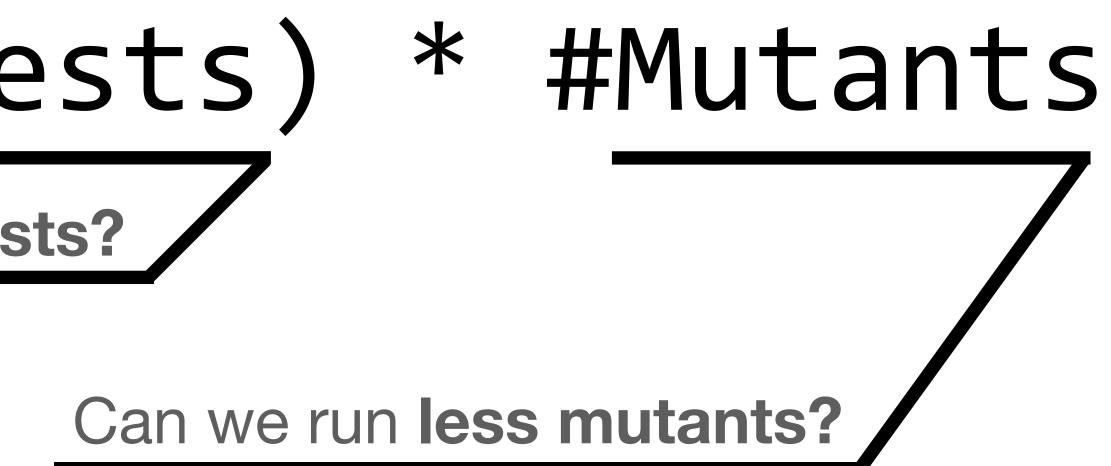
THE Problem of Mutation Analysis

Runtime = Time(tests) * #Mutants

Optimizing of Mutation Analysis

Runtime = Time(tests) * #Mutants

Can we run less tests?



Ideally... no impact on score!



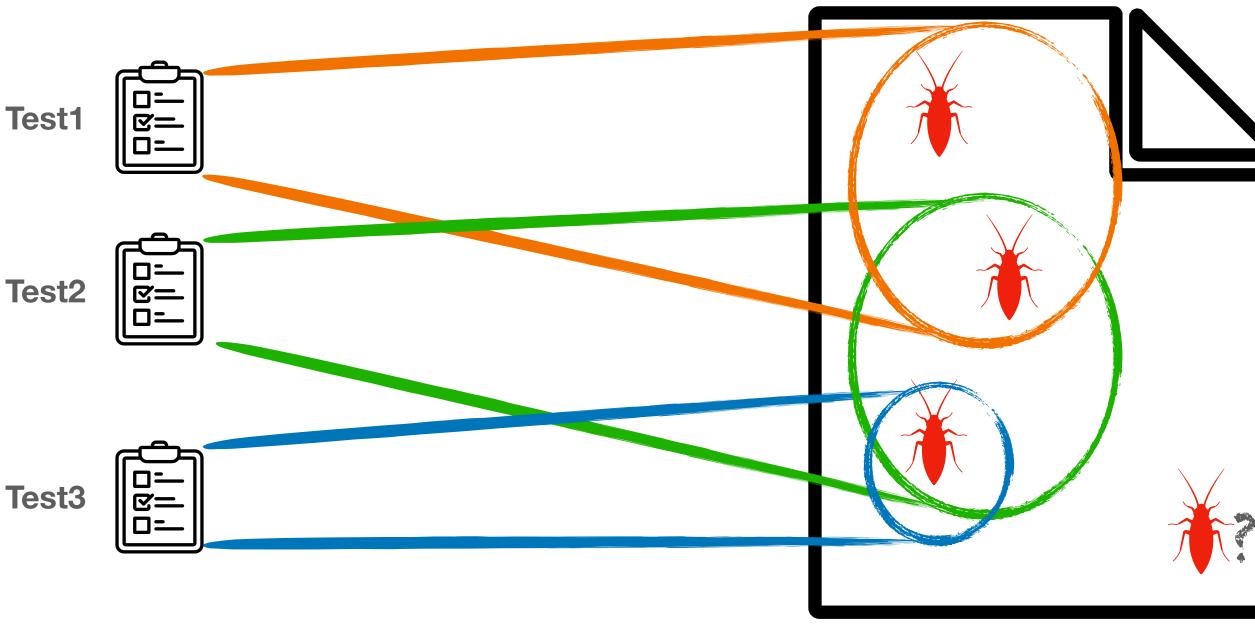
Selecting Mutants and Tests from Coverage

- Run only tests covering mutants
- Only mutate covered code

Test2

Test3

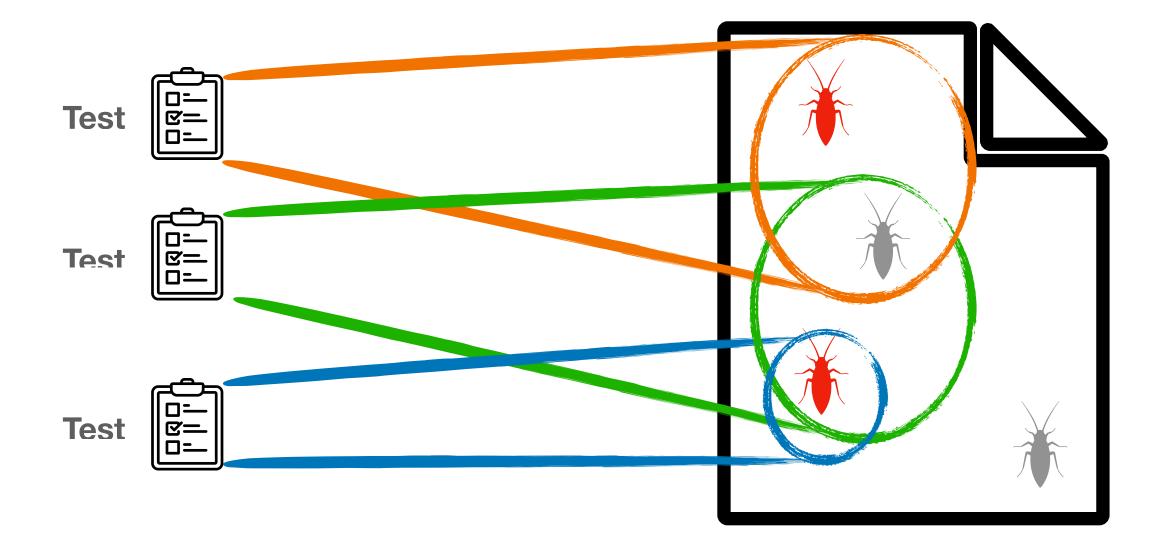
• What if lots of mutants remain?

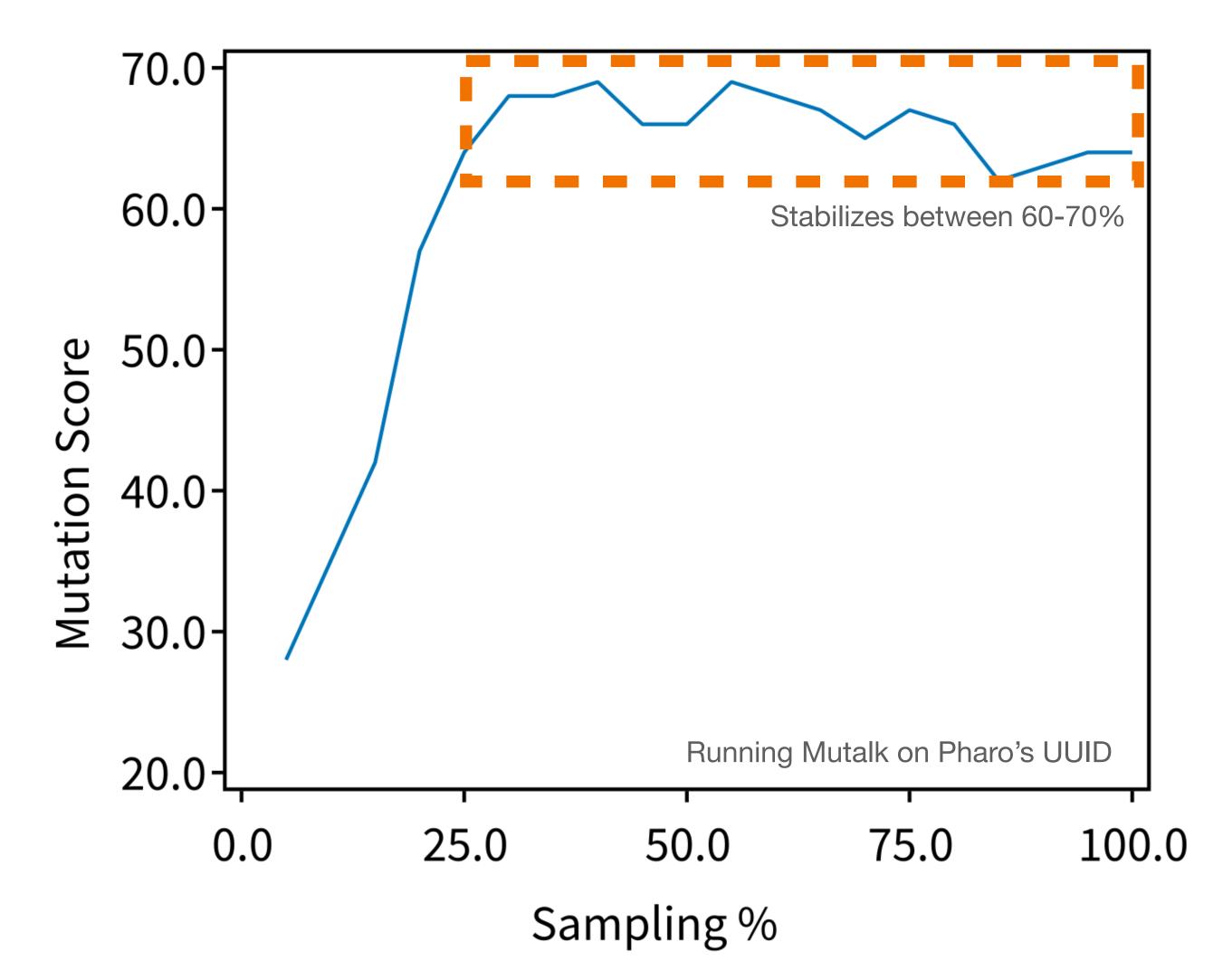




Random Sampling of Mutants

- Lots on research on it
- Aggressively reduces run time

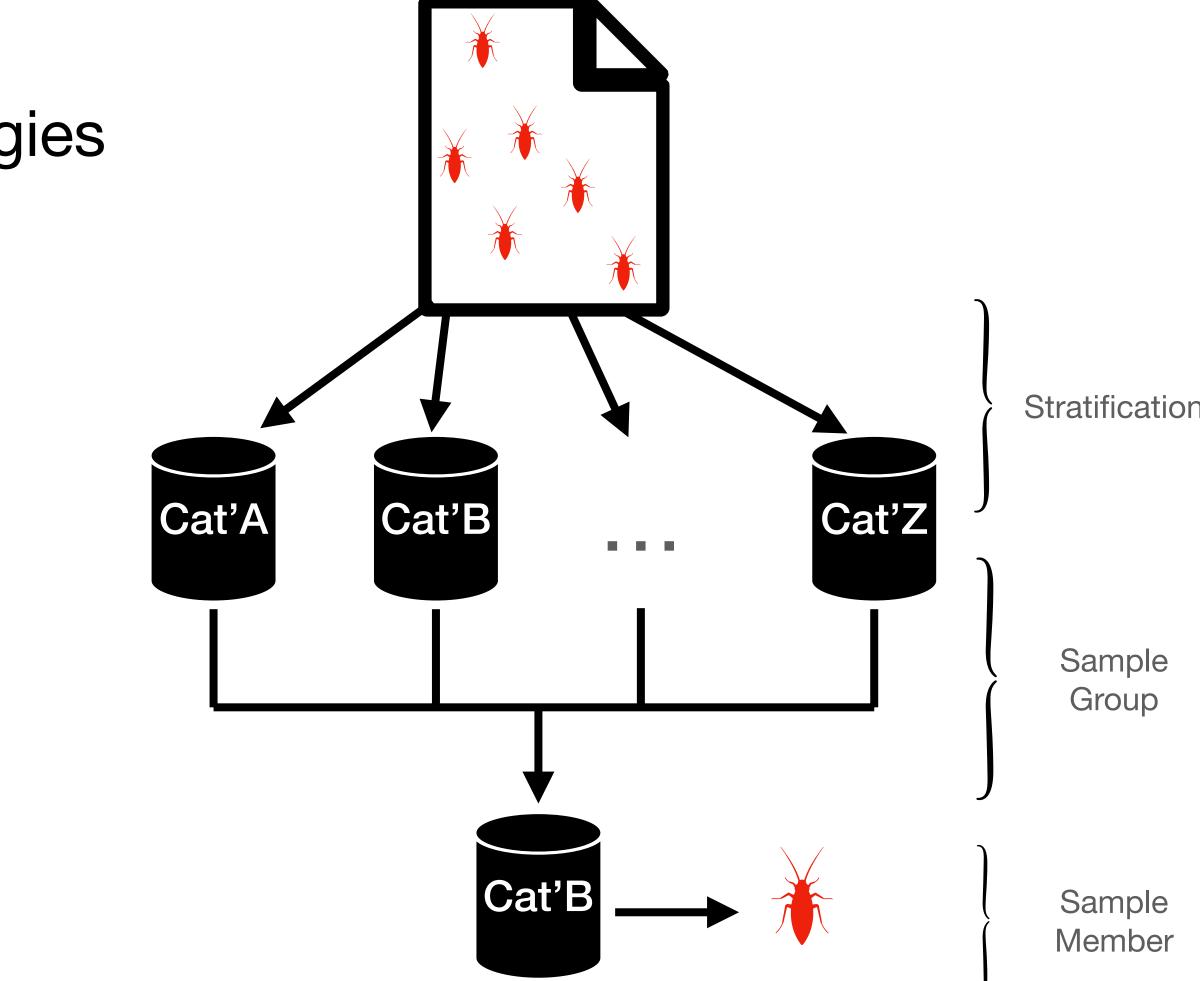




Guiding Random Sampling of Mutants

- Stratified sampling with different strategies
 - per class, method, ... lacksquare

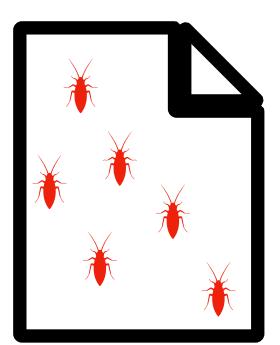
- Taking a sample:
 - Take a group at random 1.
 - 2. Take a member from the group



When to Stop Sampling: Budgets!

- Fixed budgets: # or % of mutants
 - Too many: No win in runtime!
 - Too few: Unrealistic mutation score...

- We need a practical solution:
 - **Time budget!**



How many?



Github actions integration

- Github actions
- + Microdown
- Configure budgets, different run modes...

```
name: MutalkCI
on:
  pull_request_target:
    branches: [ main ]
    paths-ignore:
    - 'README.md'
    - 'resources/**'
jobs:
  mutation_testing:
    uses: pharo-contributions/mutalk/.github/workflows/mutalk-workflow.yml@v2.5.0
```





- Merge branch 'Alamvic:main' into main
- Merge branch 'Alamvic:main' into main



github-actions (bot) commented 3 days ago

Mutation results

• Commit: 82aace5

Mutants	Number of Mutants
Alive	161
e Killed	834
Terminated	0

Mutation score: 83



Mutation Matrix and Heatmap

- Fine-tune for your project
- Advanced mutation analysis

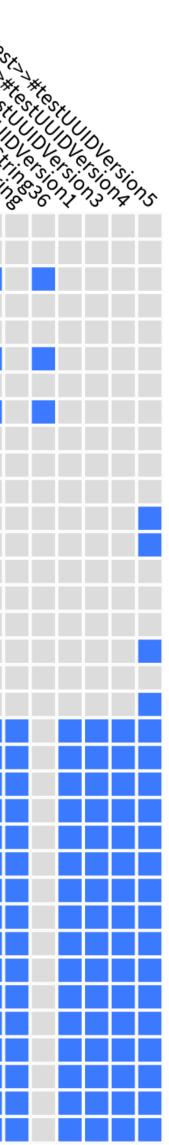
- Identify
 - **Trivial** mutations: easy to kill
 - Redundant tests and mutants
 - Randomization strategy

2	
1	MTRemoveCaretOperator
5	MTReplaceMinusWithPlusMutantOperator
8	MTReplacePlusWithMinusMutantOperator
3	MTReplaceEqualWithNotEqualOperator
1	MTReplaceAndArgumentWithTrueOperator
0	MTReplaceAndReceiverWithTrueOperator
0	MTReplaceAndWithEqvMutantOperator
14	MTReplaceAndWithNandMutantOperator
1	MTReplaceAndWithOrMutantOperator
1	MTReplaceAndWithFalseOperator
4	MTReplaceIfFalseReceiverWithFalseOperator
4	MTReplaceIfFalseReceiverWithTrueOperator
4	MTReplaceIfFalseWithIfTrueMutantOperator
1	MTRemoveNotOperator
1	MTReplaceGreaterWithLessMutantOperator
8	MTReplaceLessWithGreaterMutantOperator
1	${\sf MTReplaceMultiplicationWithDivisionMutantOperator}$
1	MTReplaceDoBlockWithEmptyBlockOperator
3	MTReplaceIdentityWithNegationOfIdentity
7	MTReplaceNotEqualWithEqualOperator
0	MTReplaceIfTrueReceiverWithFalseOperator
7	MTReplaceIfTrueReceiverWithTrueOperator
7	MTReplaceIfTrueWithIfFalseMutantOperator
0	

eNodeIdentifie hdomGenerato ‡nextRandom1 #nextRandom16 nerator>>#ne #nextCounter1 #nextCounter16 class>>#defai or class>>#nex dDependencies dDependencies HA1NameBased HA1NameBasec HA1NameBasec HA1NameBased HA1NameBased HA1NameBased HA1NameBased HA1NameBased ID>>#readFrom ID>>#readFrom: ID>>#readFrom: ID>>#readFrom: ID>>#readFrom ID>>#readFrom ID>>#readFrom ID>>#readFrom: ID>>#readFrom: ID>>#readFrom: ID>>#readFrom:

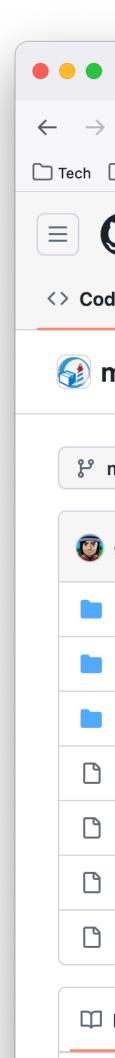
Replace #ifFalse: receiver with true in UUID>>#readFrom: Replace #ifFalse: with #ifTrue: in UUID>>#readFrom:

77

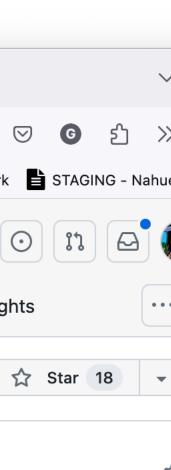


Talk to me

- If you want to use it!
- If you want to contribute!
- You need specific operators



pharo-contributions/mutalk:	Mut× +		
C A https://github	.com/pharo-contributions/mutalk	-vm 🎧 druid 🎧 mutalk 🚺 T	🗄 🖏 🚖 Tableaunoir 🗋 Stages 🗋 work
pharo-contributions / mutalk			+ •
ode 💿 Issues 7 11 Pull requests	s 🖓 Discussions 🕞 Actions 🗄	Projects 1 🛛 Wiki	! Security 🗠 Insig
mutalk Public	S Edit I	Pins 👻 💿 Unwatch 7 🗣	- २४ Fork 14 - ▼
master 👻 가 Branch 📀 6 Tags	Q Go to file	t + <> Code •	About
guillep Update mutalkCI.yml 🗙	7a75b07 · yes	sterday 🕒 560 Commits	Mutation Testing in Pha
.github/workflows	added pharo12 to Cl	6 months ago	Readme
src	Update src/MuTalk-CI-Tests/MTCITonel	Jso 5 days ago	-\- Activity
workflow	Update mutalkCI.yml	yesterday	 Custom properties 18 stars
.project	add repository properties	6 years ago	7 watching
.smalltalk.ston	Exclude tests	4 months ago	ੳ 14 forks
.travis.yml	add Travis Cl	3 years ago	Report repository
README.md	update readme	3 months ago	Releases 6
README		∅ :≡	v2.4.0 Latest on May 6 + 5 releases





Mutalk 2024

- Working from Pharo9 to Pharo12 (13?)
- Practical: Budgets, random selection and CI

- **New** features: test filters, logging, new operators
- Fixes, improvements, cleanups

• Thanks to all people contributing over the years!

Coverage vs Mutation

- Average has 100% coverage
- What if we mistake the / by a *?

CollectionTest >> testAverage self assert: #(1 2 3) average

Collection >> average ^ self sum / self size