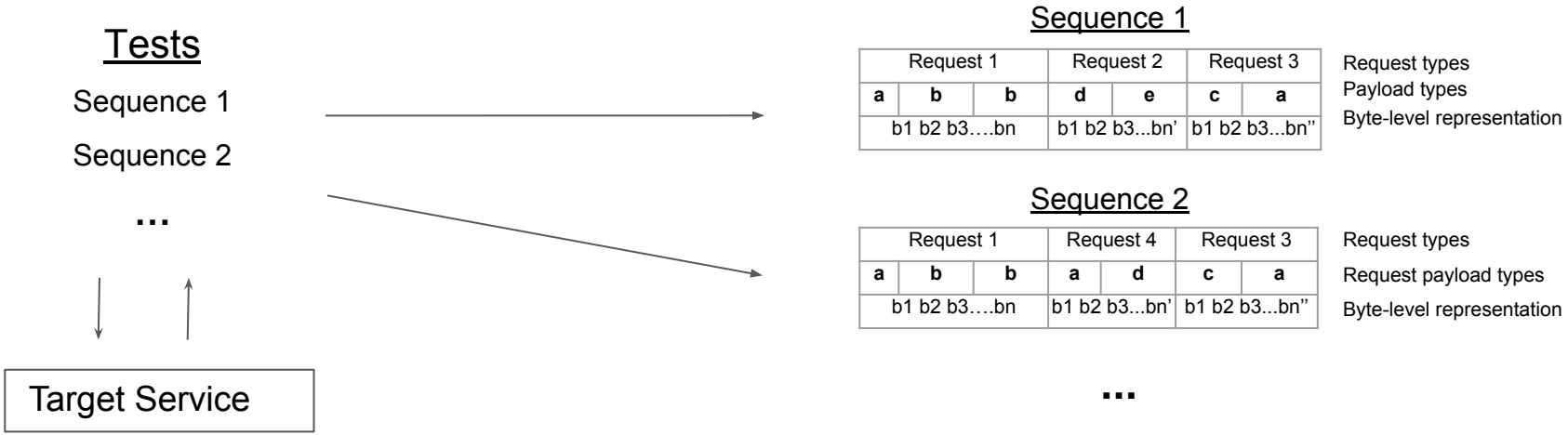
Pythia: Grammar-Based Fuzzing of REST APIs with Coverage-guided Feedback and Learning-based Mutations

w/ Roxana Geambasu, Baishakhi Ray (Columbia University) and Patrice Godefroid, Marina Polishchuk (Microsoft Research)

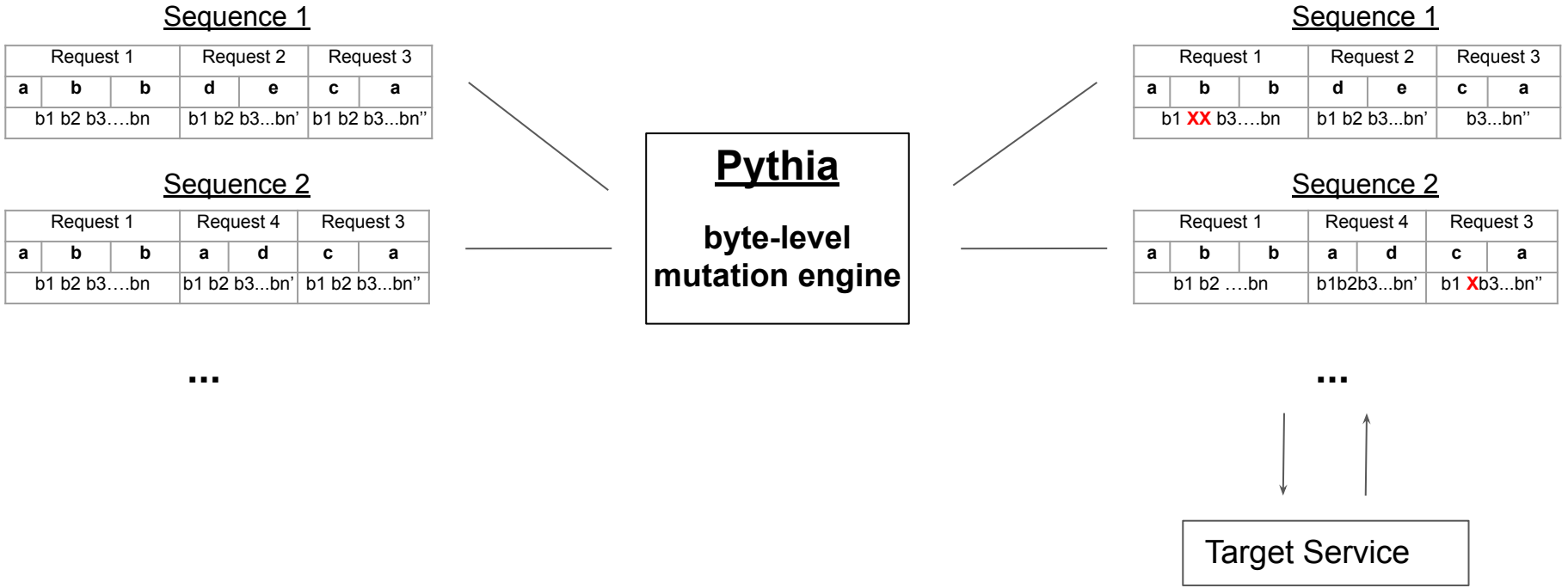
Challenge

- Most payload types have a fixed set of possible values
- Can we augment this set of possible values?

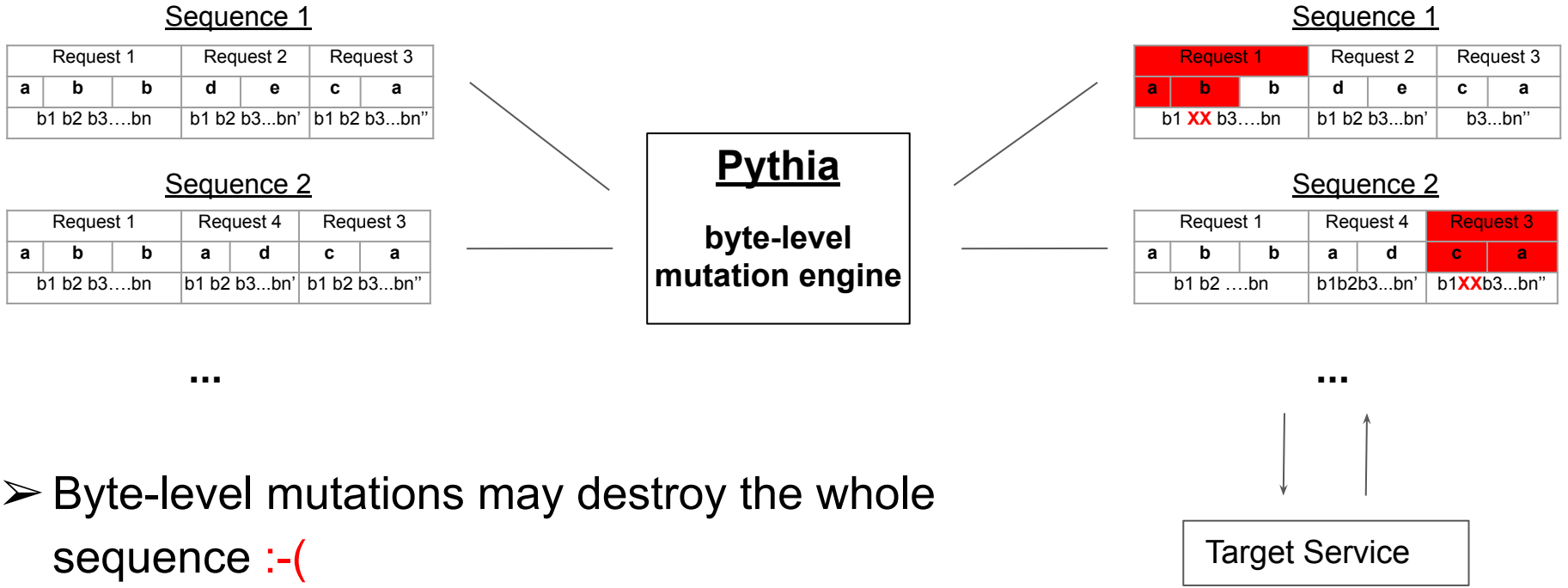
Example



First approach: Random byte-level mutations



First approach: Random byte-level mutations



➢ Byte-level mutations may destroy the whole sequence :-(

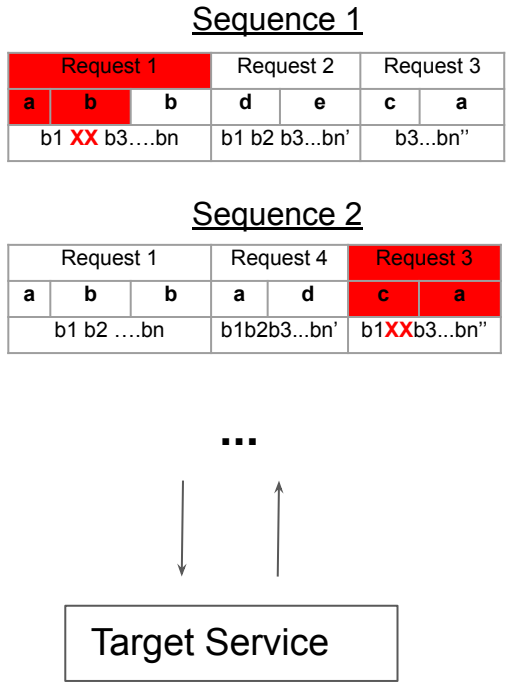
First approach: Random byte-level mutations

```
Old rule @382:'E'
New rule @382:' \xa4 '

Sent: 'POST /api/v4/projects/997190/repository/branches
HTTP/1.1\r\n...\r\nPRIVATE-TOKEN: DRiX47nu\xa4
P2ARa4APFrf\r\n\r\n{"ref":"master","branch":"string"}\r\n'

Recv: 'HTTP/1.1 401 Unauthorized\r\nDate: Thu, 31 Oct 2019
06:19:53 GMT...\r\n\r\n{"message":"401 Unauthorized"}'

Parsing exception:: Exception Parsing Response item: 'name'
```



➢ Byte-level mutations may destroy the whole sequence :-(

Second approach: AST-level mutations

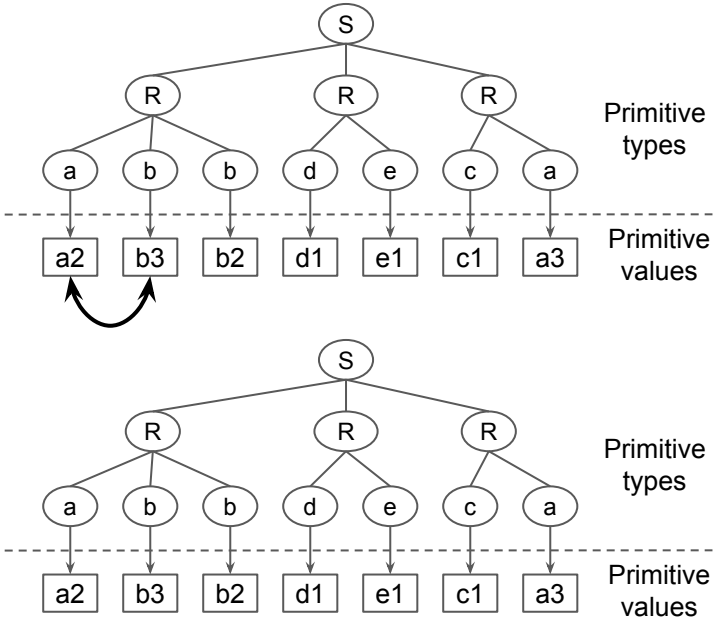
Sequence 1

Request 1			Request 2		Request 3	
a	**b**	**b**	**d**	**e**	**c**	**a**
b1 b2 b3….bn			b1 b2 b3...bn'		b1 b2 b3...bn''	

Sequence 2

Request 1			Request 4		Request 3	
a	**b**	**b**	**a**	**d**	**c**	**a**
b1 b2 b3….bn			b1 b2 b3...bn'		b1 b2 b3...bn''	

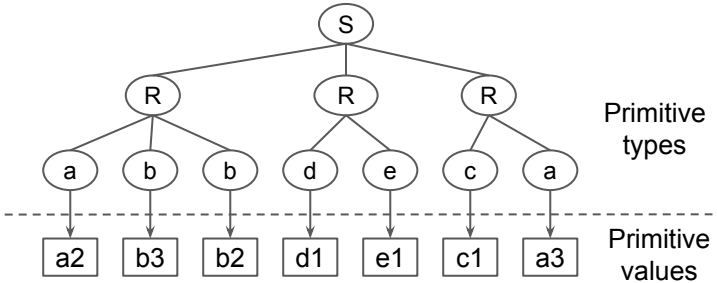
...



Second approach: AST-level mutations

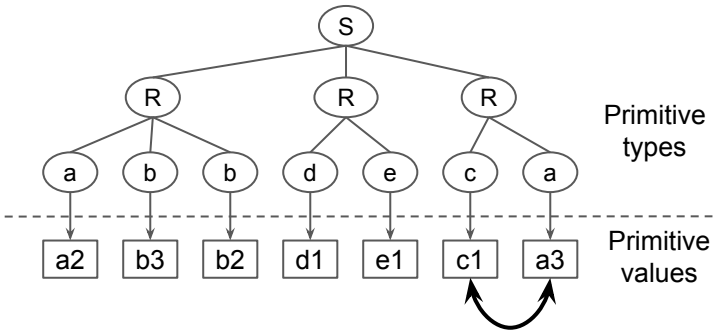
Sequence 1

Request 1			Request 2		Request 3	
a	**b**	**b**	**d**	**e**	**c**	**a**
b1 b2 b3….bn			b1 b2 b3...bn'		b1 b2 b3...bn''	



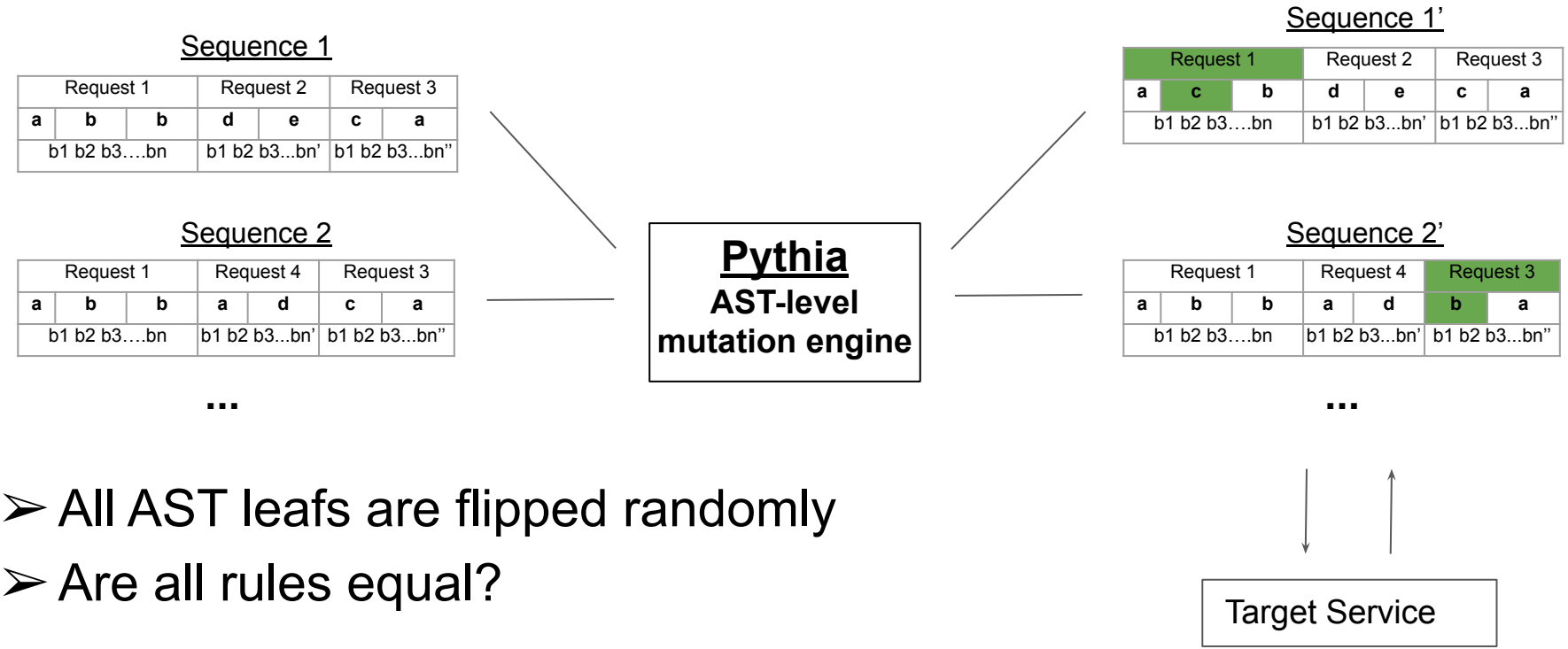
Sequence 2

Request 1			Request 4		Request 3	
a	**b**	**b**	**a**	**d**	**c**	**a**
b1 b2 b3….bn			b1 b2 b3...bn'		b1 b2 b3...bn''	



...

Second approach: AST-level mutations

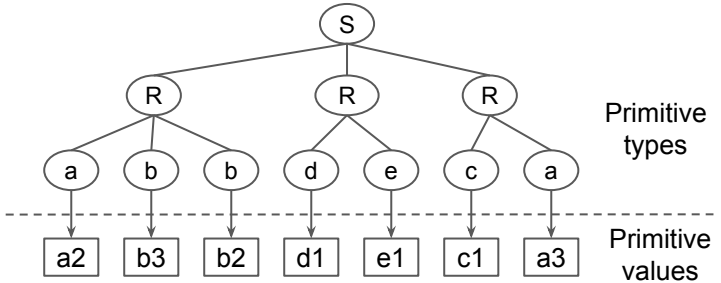


- All AST leafs are flipped randomly
- Are all rules equal?

Third approach: Probabilistic AST-level mutations

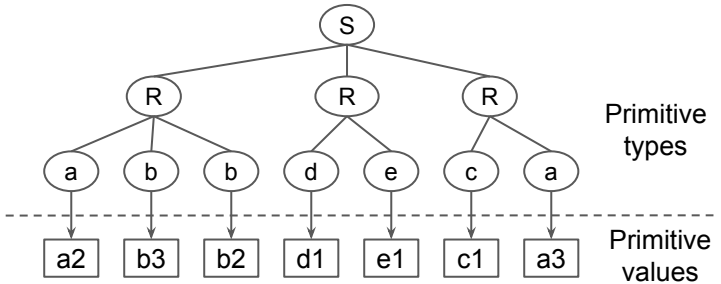
Sequence 1

Request 1			Request 2		Request 3	
a	**b**	**b**	**d**	**e**	**c**	**a**
b1 b2 b3….bn			b1 b2 b3...bn'		b1 b2 b3...bn''	



Sequence 2

Request 1			Request 4		Request 3	
a	**b**	**b**	**a**	**d**	**c**	**a**
b1 b2 b3….bn			b1 b2 b3...bn'		b1 b2 b3...bn''	



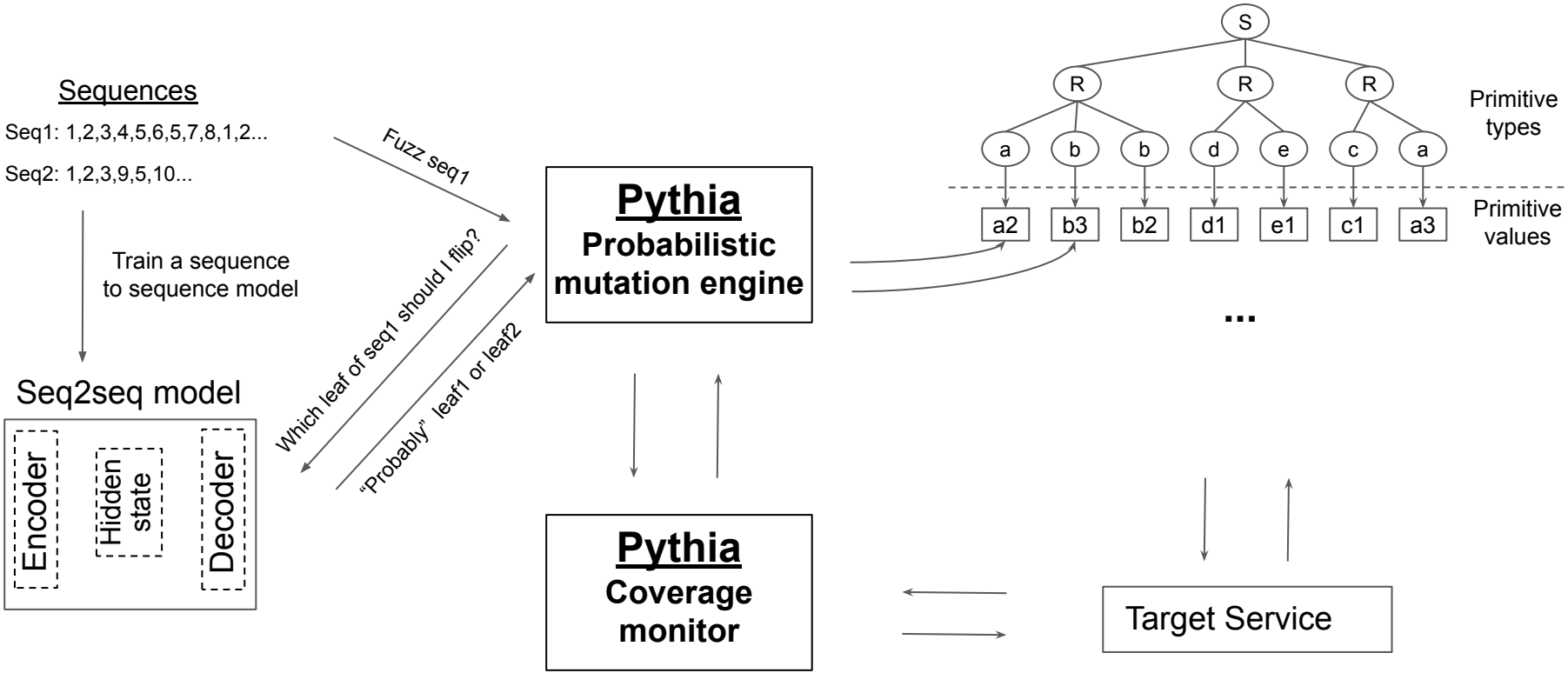
...

Production rules

1:	S → S + R
2:	R → R + P
3:	P → P + a
4:	a → a2
5:	P → P + b
6:	b → b3
5:	P → P + b
7:	b → b2
8:	R → e
1:	S → S + R
2:	R → R + P
	...

1:	S → S + R
2:	R → R + P
3:	P → P + a
9:	a → a2
5:	P → P + b
10:	b → b3
	...

Third approach: Probabilistic AST-level mutations



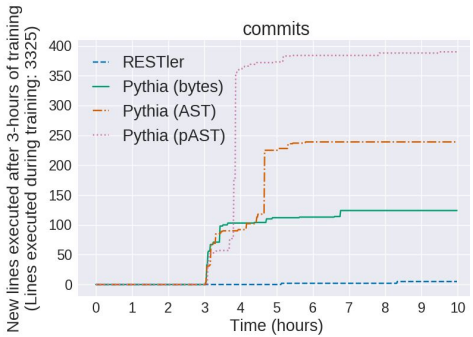
Selected evaluation results

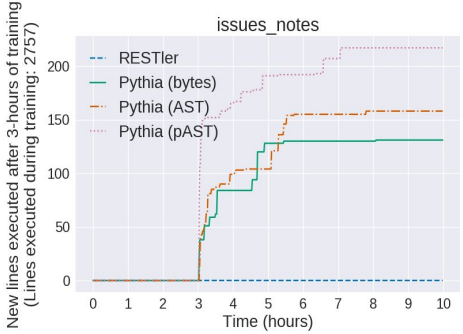
- Q1: Are tests generated by Pythia increasing code coverage?
- Q2: Can Pythia find new errors?

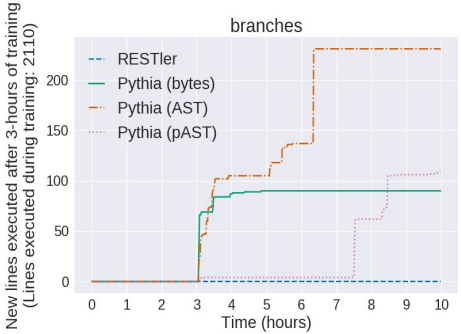
Case study: Four Gitlab APIs

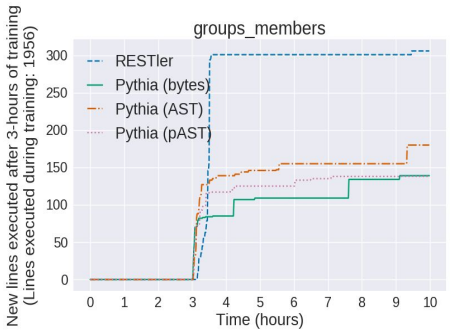
API	Total requests	Avg. values per primitive type	Avg. dependencies per request type	Path dependencies	Body dependencies
Commits	15	11	1.7	Yes	Yes
Issues	25	20	1.8	Yes	No
Groups	53	2	1.4	Yes	No
Branches	8	2	1.5	Yes	No

Code coverage (Q1)



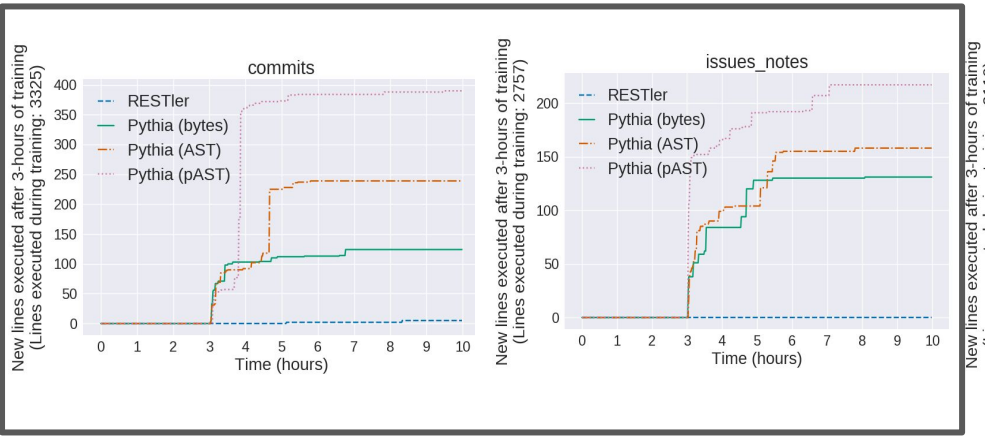


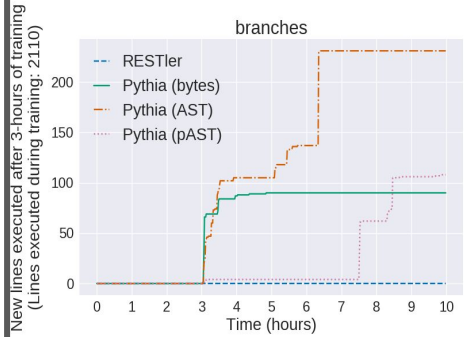


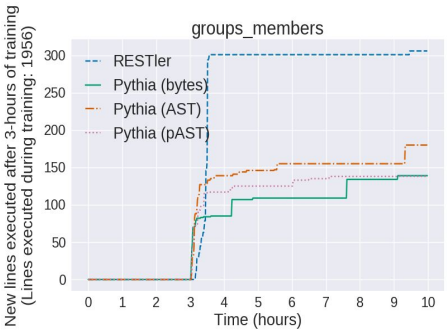


API	Total requests	Avg. values per primitive type	Avg. dependencies per request type	Path dependencies	Body dependencies
Commits	15	11	1.7	Yes	Yes
Issues	25	20	1.8	Yes	No
Groups	53	2	1.4	Yes	No
Branches	8	2	1.5	Yes	No

Code coverage (Q1)

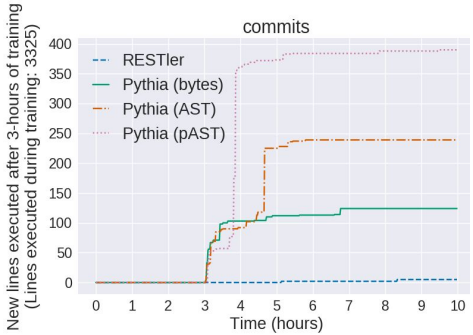


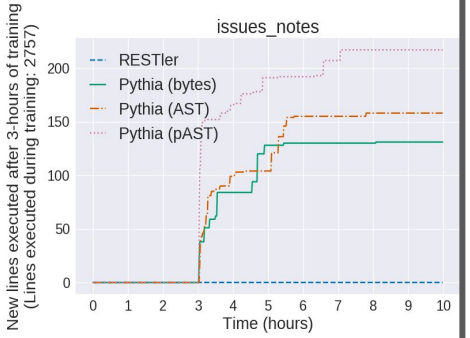


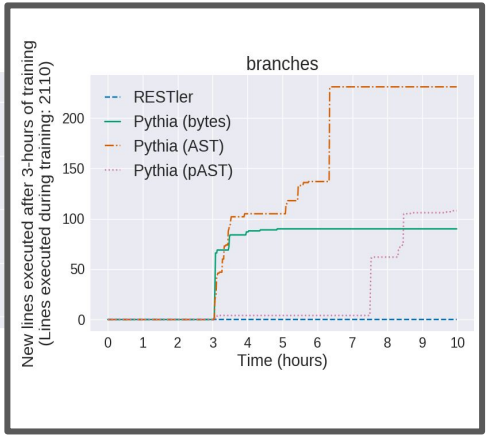


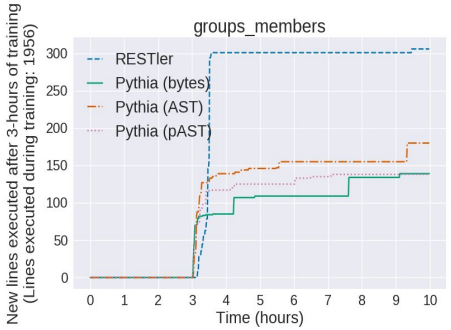
API	Total requests	Avg. values per primitive type	Avg. dependencies per request type	Path dependencies	Body dependencies
Commits	**15**	**11**	**1.7**	**Yes**	**Yes**
Issues	**25**	**20**	**1.8**	**Yes**	**No**
Groups	53	2	1.4	Yes	No
Branches	8	2	1.5	Yes	No

Code coverage (Q1)



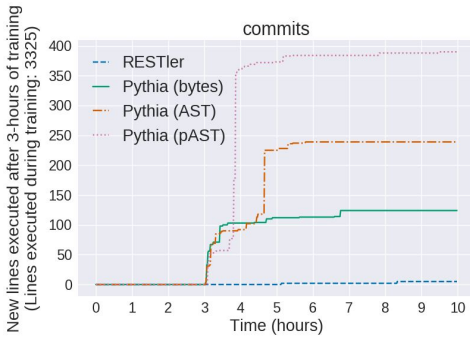


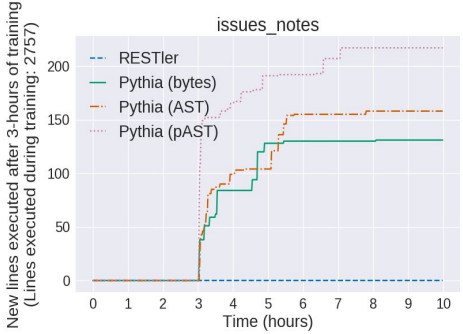


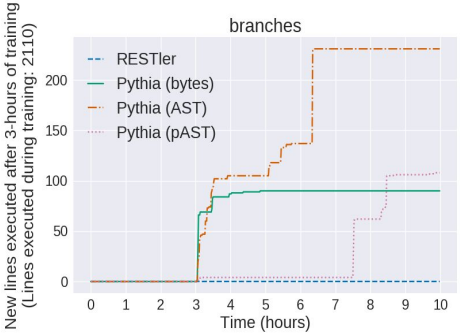


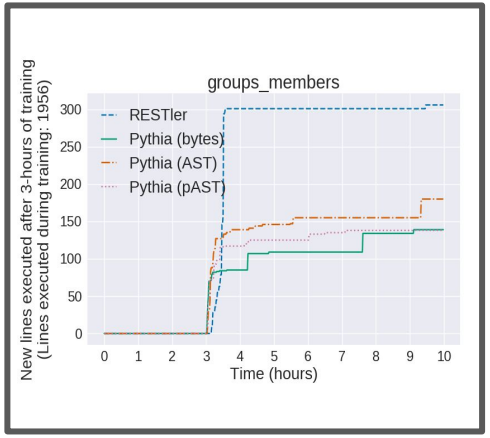
API	Total requests	Avg. values per primitive type	Avg. dependencies per request type	Path dependencies	Body dependencies
Commits	15	11	1.7	Yes	Yes
Issues	25	20	1.8	Yes	No
Groups	53	2	1.4	Yes	No
Branches	**8**	**2**	**1.5**	**Yes**	**No**

Code coverage (Q1)









API	Total requests	Avg. values per primitive type	Avg. dependencies per request type	Path dependencies	Body dependencies
Commits	15	11	1.7	Yes	Yes
Issues	25	20	1.8	Yes	No
Groups	**53**	**2**	**1.4**	**Yes**	**No**
Branches	8	2	1.5	Yes	No

New errors (Q2)