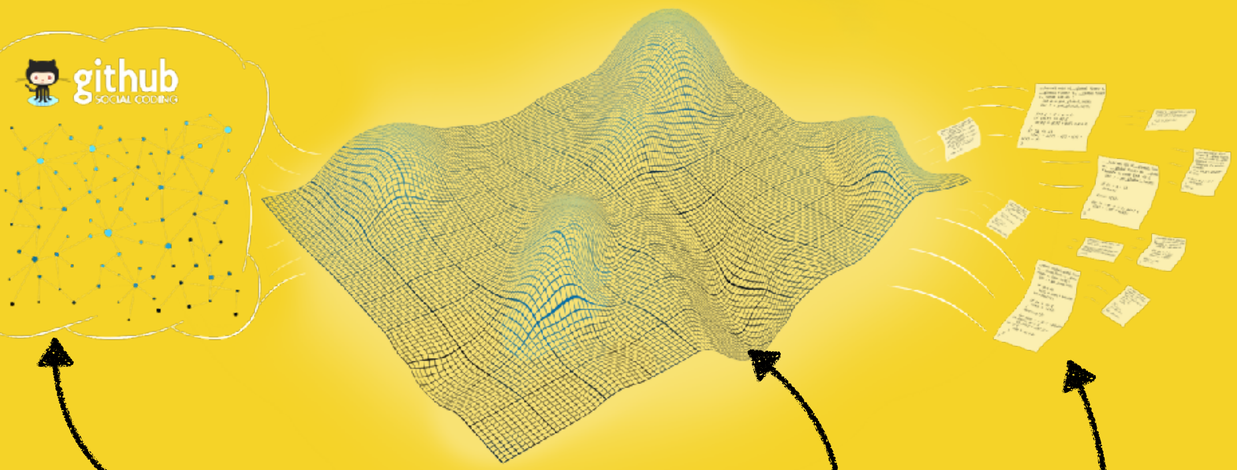
Progression Review

2017

Fork me on GitHub

Synthesizing Benchmarks for Predictive Modeling



mine code from web

model character distribution

generate executable benchmarks

Fork me on GitHub

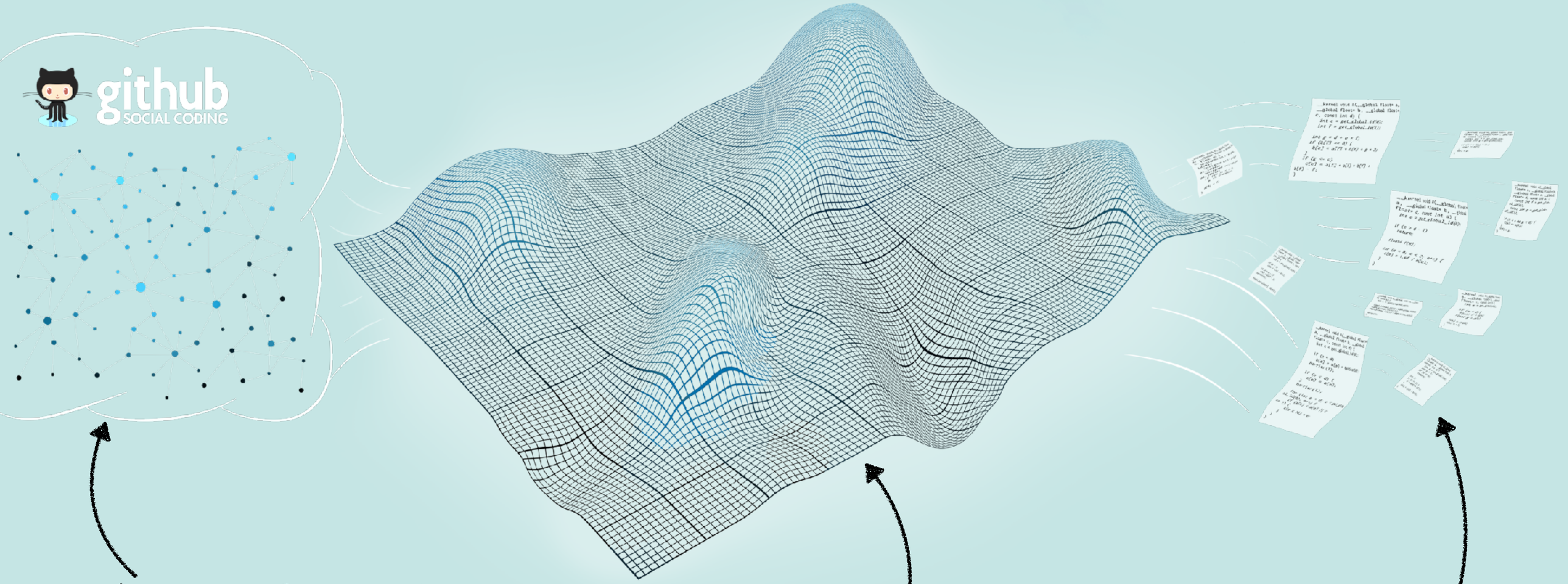
End-to-end Deep Learning of Optimization Heuristics

normalize input code

model token distribution

predict optimizations

Compiler Fuzzing through Deep Learning



Compiler Fuzzing through Deep Learning

Faster than CSmith

3x faster execution, 100x smaller size

Covers broader class of bugs

more expressive, not limited by SA

bugs in front, middle and back

Lower development cost

80x less code, learns model in 14 hours

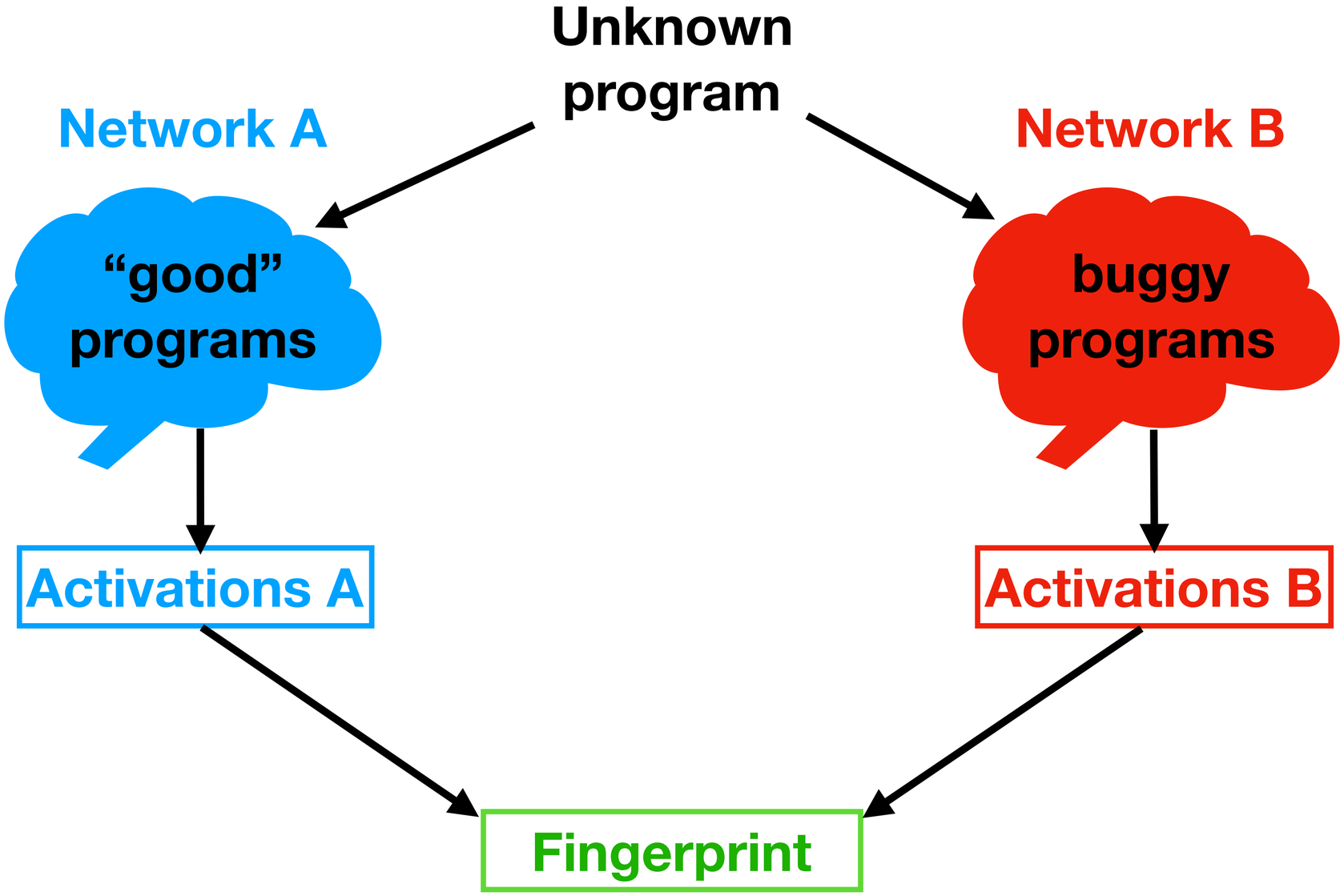
TODO List

Evaluate state-of-the-art
C and JavaScript
Compare DeepSmith to them

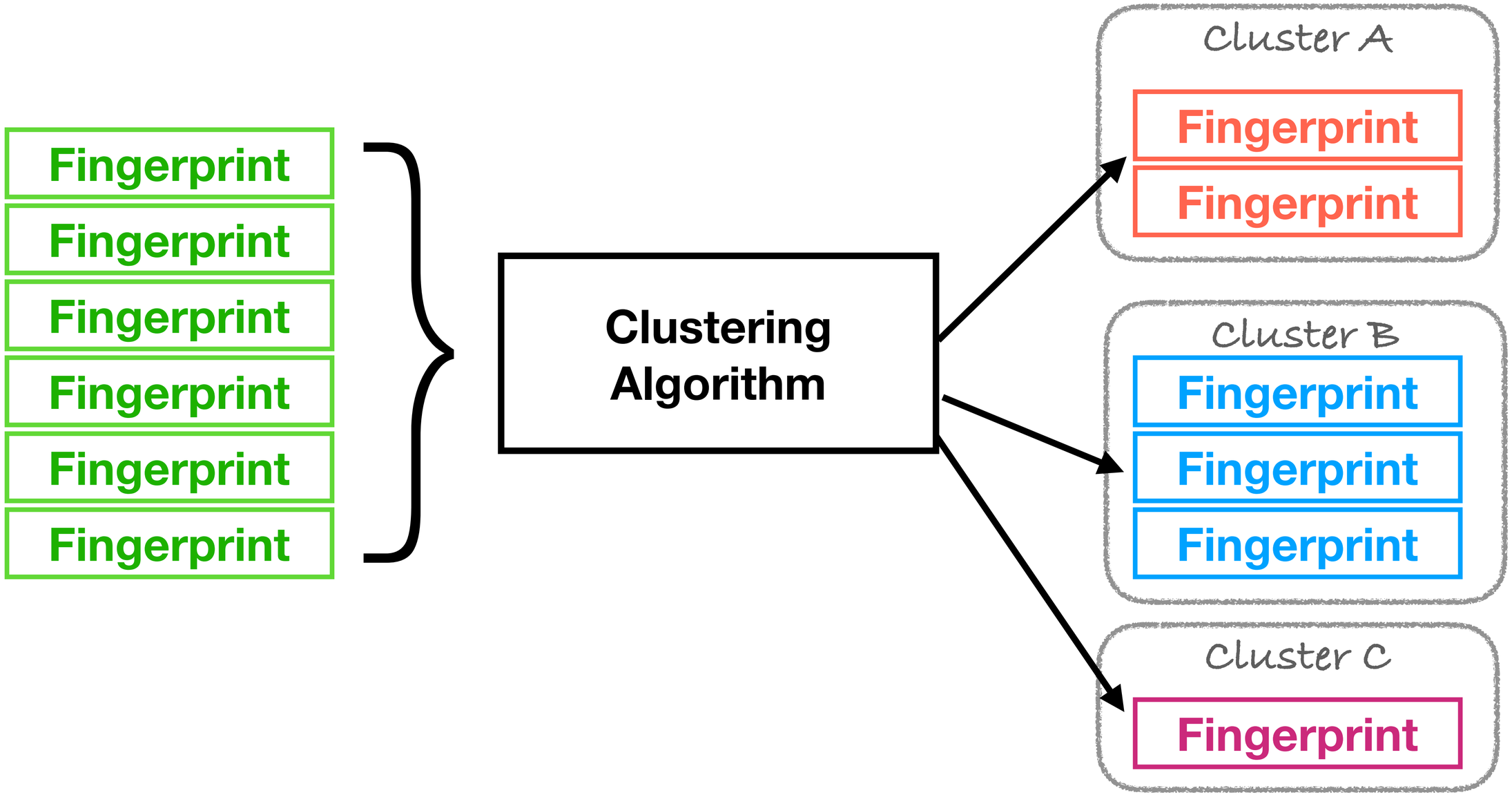
Fuzztest Compute Languages
GLSL and Solidium
OOPSLA'17 — no prior art

Build an ML bug dataset
Thousands of bug-causing examples

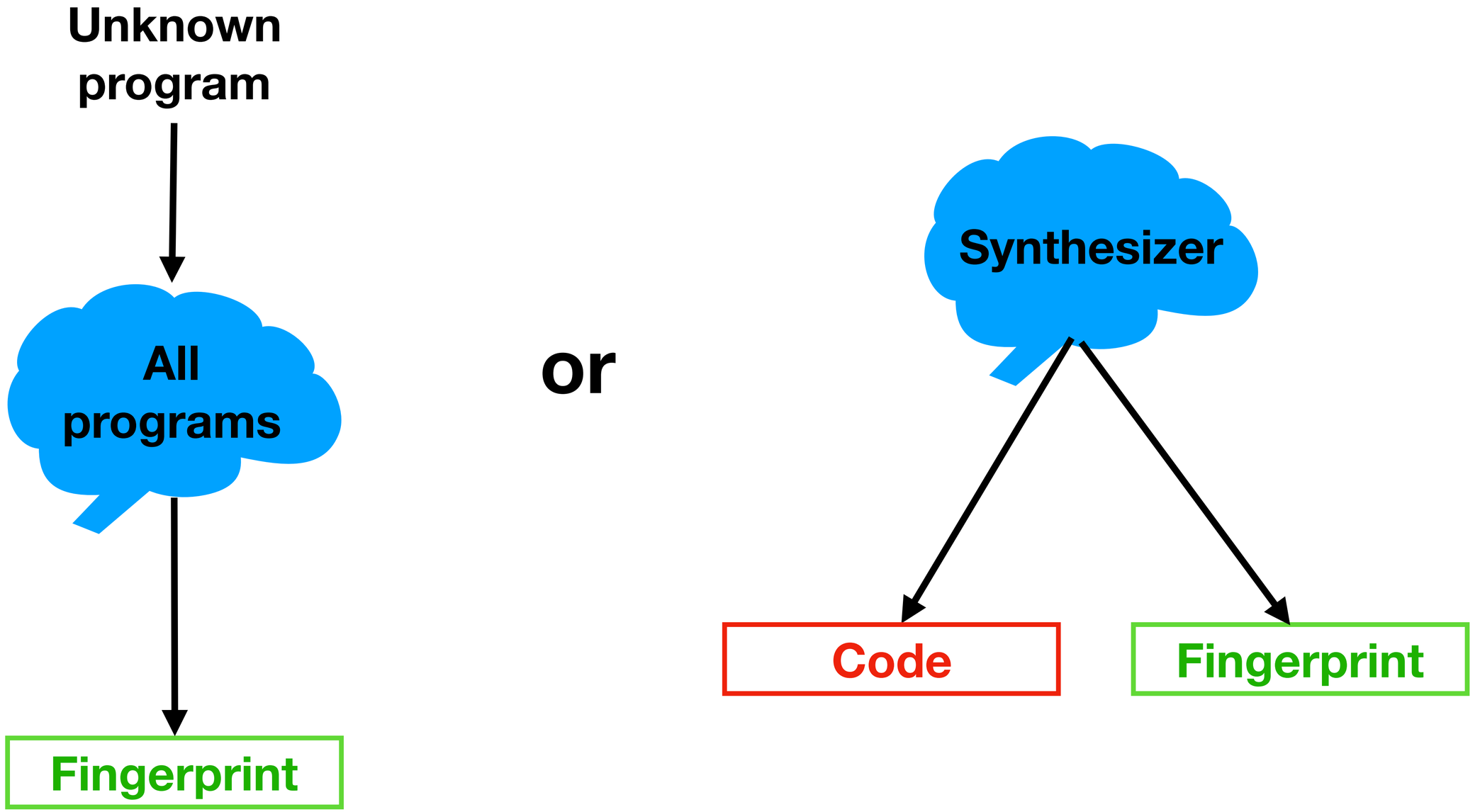
Program “Fingerprinting”



Program “Fingerprinting”



Program “Fingerprinting”



~~Program~~ “Fingerprinting”

Compiler

GCC Bug tracker:

64K bug reports

Hand labelled

With *fixes*

also...

1. GAN
2. Recursive Synthesis
3. Guided synthesis