A Large Scale Study of Programming Languages and Code Quality in Github

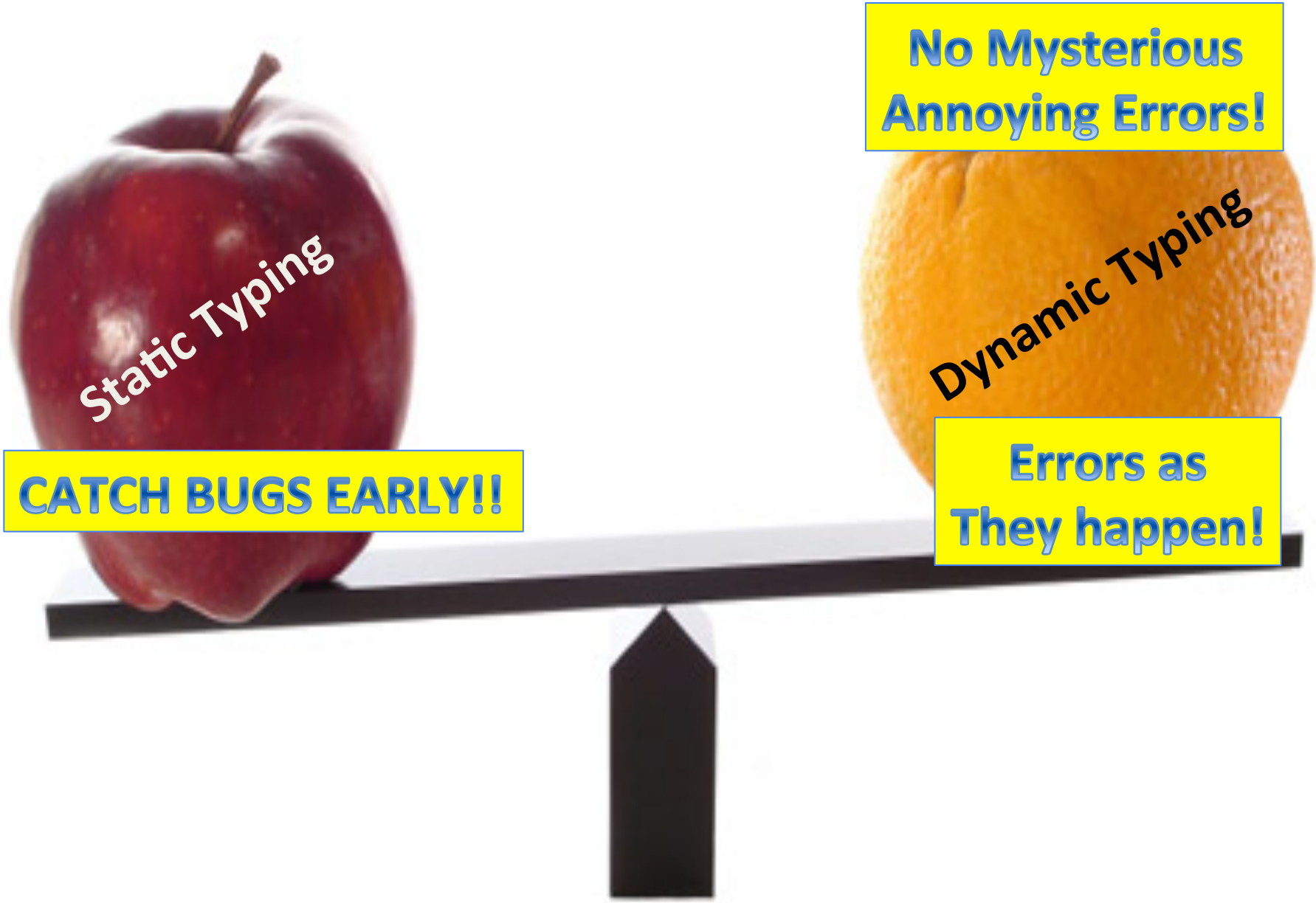
Baishakhi Ray

Daryl Posnett

Vladimir Filkov

University of California, Davis.

What is the right programming language for my job?



What is the effect of programming languages on software quality?

Ideally, we do controlled studies

Hanenberg, Prechelt, Kleinschmager, Robbes…

In this paper, we use Github.

Lot of data…use regression for control.

Study Subjects

- 729 popular GitHub projects.
  - Linux, Git, Php-src, Jquery, Mongodb etc.
  - 80M SLOC, 29K authors, 1.5M commits
- 17 languages:
  - C, C++, C#, Objectve-C, Go, Java, JavaScript, CoffeeScript, TypeScript, Ruby, Php, Python, Perl, Clojure, Erlang, Haskell, Scala

Punch Line

Yeah, a bit.

Static Typing, Strong Typing, Memory management and functional programming matter, but…

Not much effect observed.

SOFTWARE > DEVELOPER

Languages don't breed bugs, PEOPLE breed bugs, say boffins

You say C++, I say Python, you say JavaScript, I say Erlang ...

By Richard Chirgwin, 6 Nov 2014

52

Beginner's guide to SSL certificates

If you want to spark a religious war, express an unshakeable preference for a programming language, and by preference, make your favourite something relatively obscure, like Erlang. It turns out, according to a study by a bunch of UC Davis boffins, the differences in code quality between languages are pretty small.

To be presented at the ACM's Foundations of Software Engineering (FSE) symposium in Hong Kong in mid-November, the paper, *A Large Scale Study of Programming Languages and Code Quality in Github*, even analysing 729 projects, 80 million lines of code, 1.5 million commits from 29,000 or so authors in 17 languages only found "modest" effects coming from the language software's written in.

While strong typing is a little better than weak typing, and in functional languages static typing is a little better than dynamic typing, and managed memory is better than unmanaged, the authors (Baishakhi Ray, Daryl Posnett, Vladimir Filkov, and Premkumar Devanbu of the UC Davis department of comp. sci) find that "modest effects arising from language design are overwhelmingly dominated by the process factors such as project size, team size, and commit size".

RELATED STORIES

MoD releases code to GitHub: Our Ideaworks... sort of

Super-villains of C sought for WORLD CONQUEST plan

UK consumers particularly prone to piss-poor patching

Something for the Weekend, Sir? Sign off my IT project or I'll PHONE your MUM

Poodle Admins! Never mind POODLE, there're NEW OpenSSL bugs to splat

MOST READ | MOST COMMENTED

Facebook's plain English data policy: WE'LL SELL YOU LIKE A PIG at a fair

SLURP! Get your TONGUE around our LOLLIPOP - Google

Could YOU identify these 10 cool vintage mobile phones?

Rosetta probot drilling DENIED: Philae has its 'LEG in the AIR'

VINYL is BACK and you can thank Sonos for that

SPOTLIGHT

Microsoft: It's TIME at LAST. Yes - .NET is going OPEN and X-PLATFORM

Microsoft .NET released from its Windows chains... but what ABOUT MONO?

Drupalgeddon megaflaw raises questions over CMS bods' crisis mgmt

Groupon flees from army of angry GNOMES: Trademark bid for 'Gnome' tab scrapped

..meanwhile, in the American Media…

INFOWORLD TECH WATCH
By Paul Krill | Follow

About
Informed news analysis every weekday

Functional languages rack up best scores for software quality

Credit: iStockphoto

A study of GitHub projects and the languages used to build them finds that certain language characteristics are more likely to result in better software

InfoWorld | Nov 7, 2014

Language design makes a difference in software quality, and functional languages offer an edge when it comes to building quality software, a study of programming languages and code quality in GitHub reveals.

Researchers at the University of California, Davis, recently published their findings,

MORE LIKE THIS

10 things you should know about Apple's Swift

Scala founder: Language due for 'fundamental rethink'

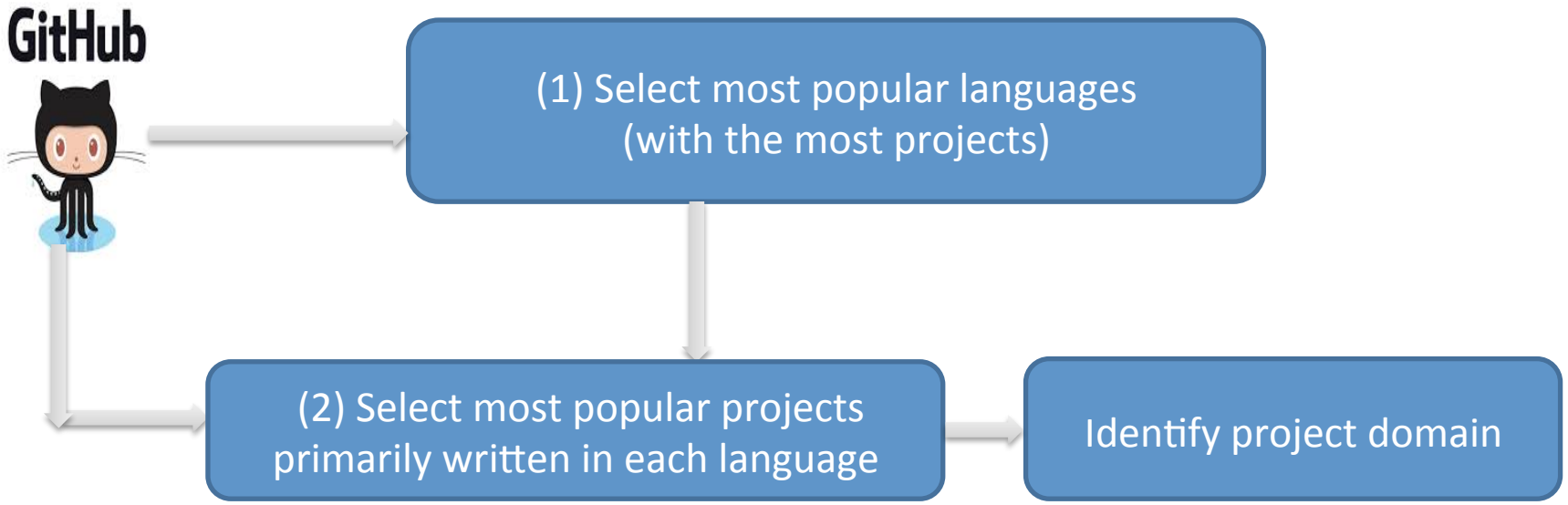

New NSA-funded programming language could close long-standing security holes

Yeah, by a bit.

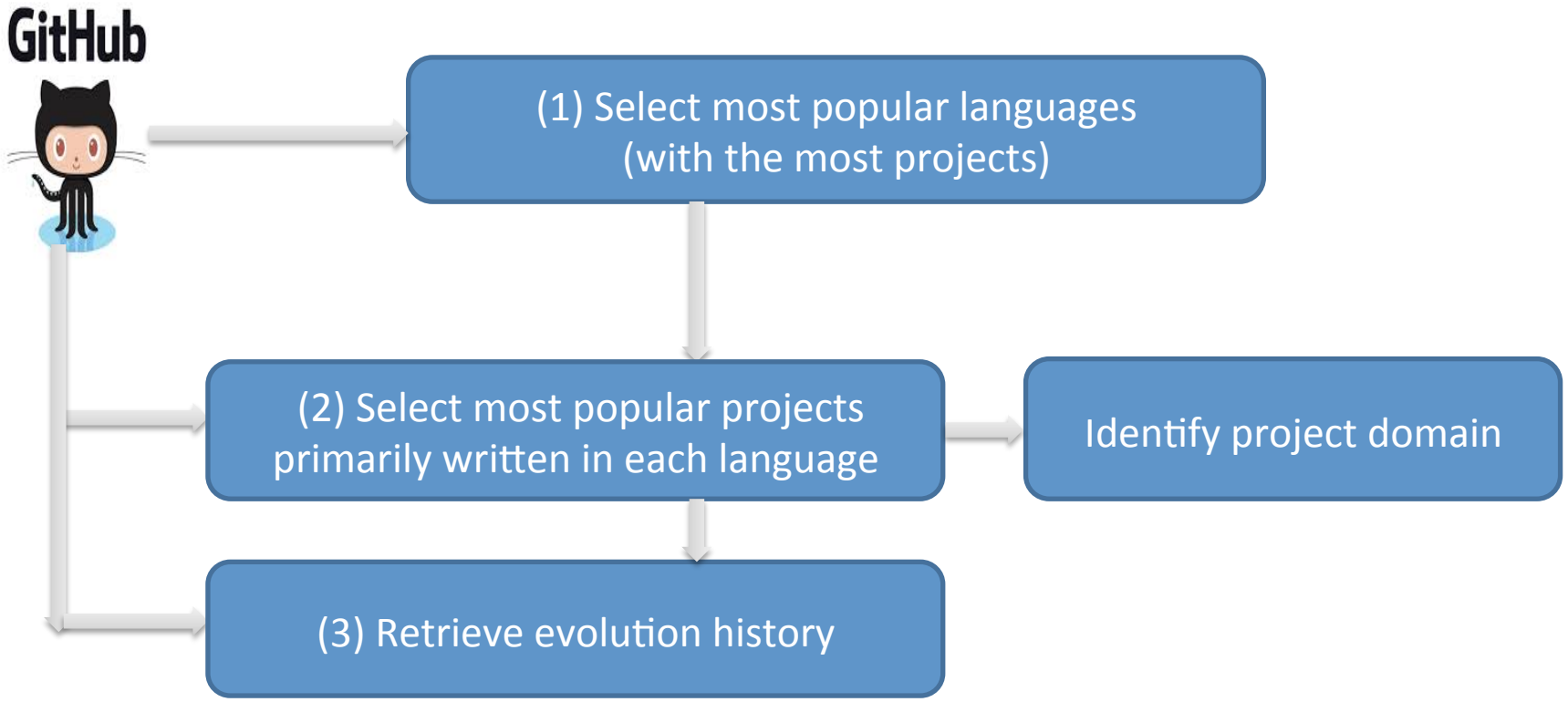
Methodology: Mining GitHub

(1) Select most popular languages
(with the most projects)

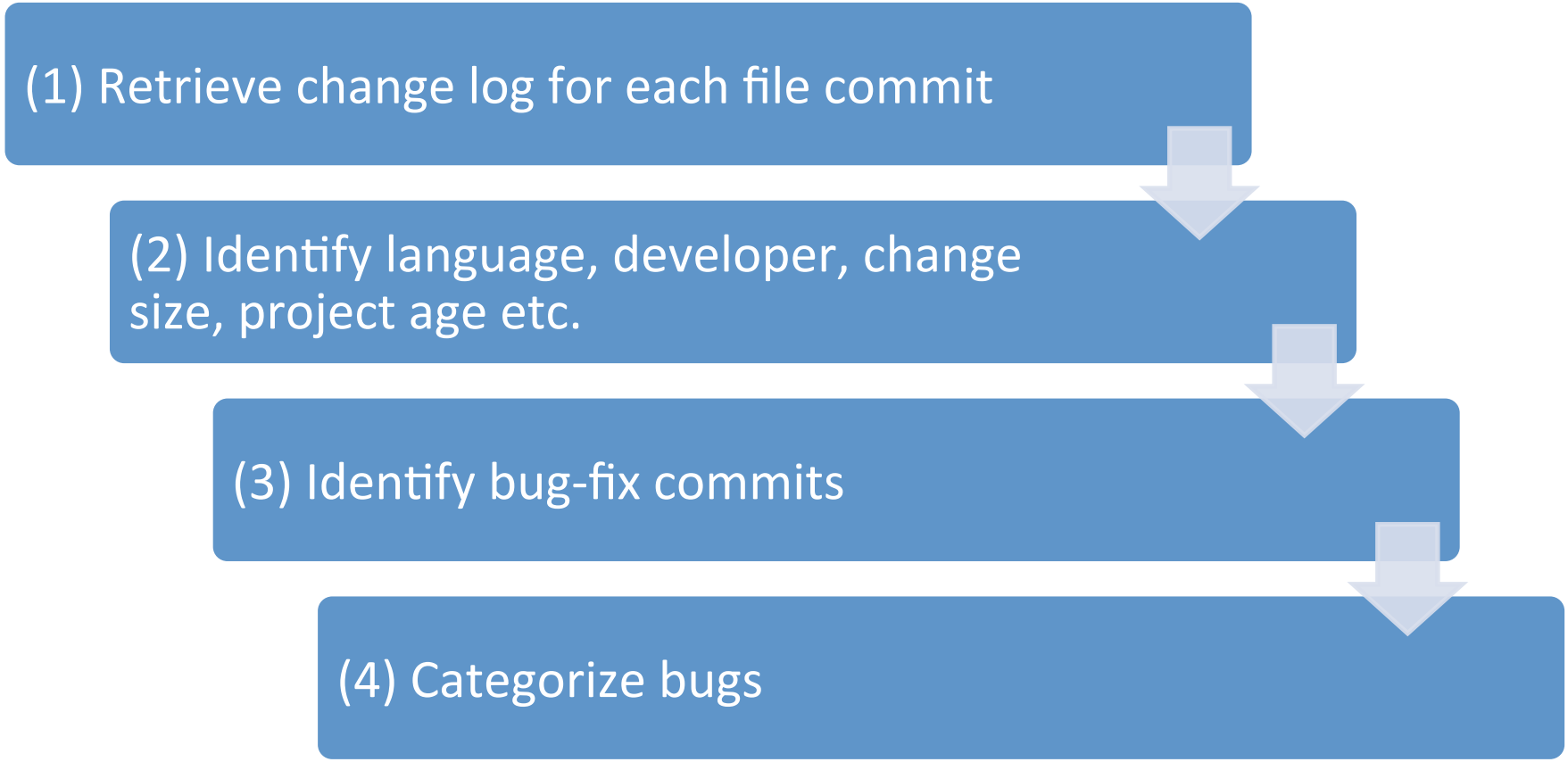
Methodology: Mining GitHub



Methodology: Mining GitHub



Methodology: Mining Repository



Methodology: Triangulation

- Regression analysis
  - Effect of language on bugs while controlling for *code size*, *project age*, *contributors* etc.
- Visualization
- Text Analysis

Results

RQ1. Are some languages more defect prone than others?

Less bug-prone than average

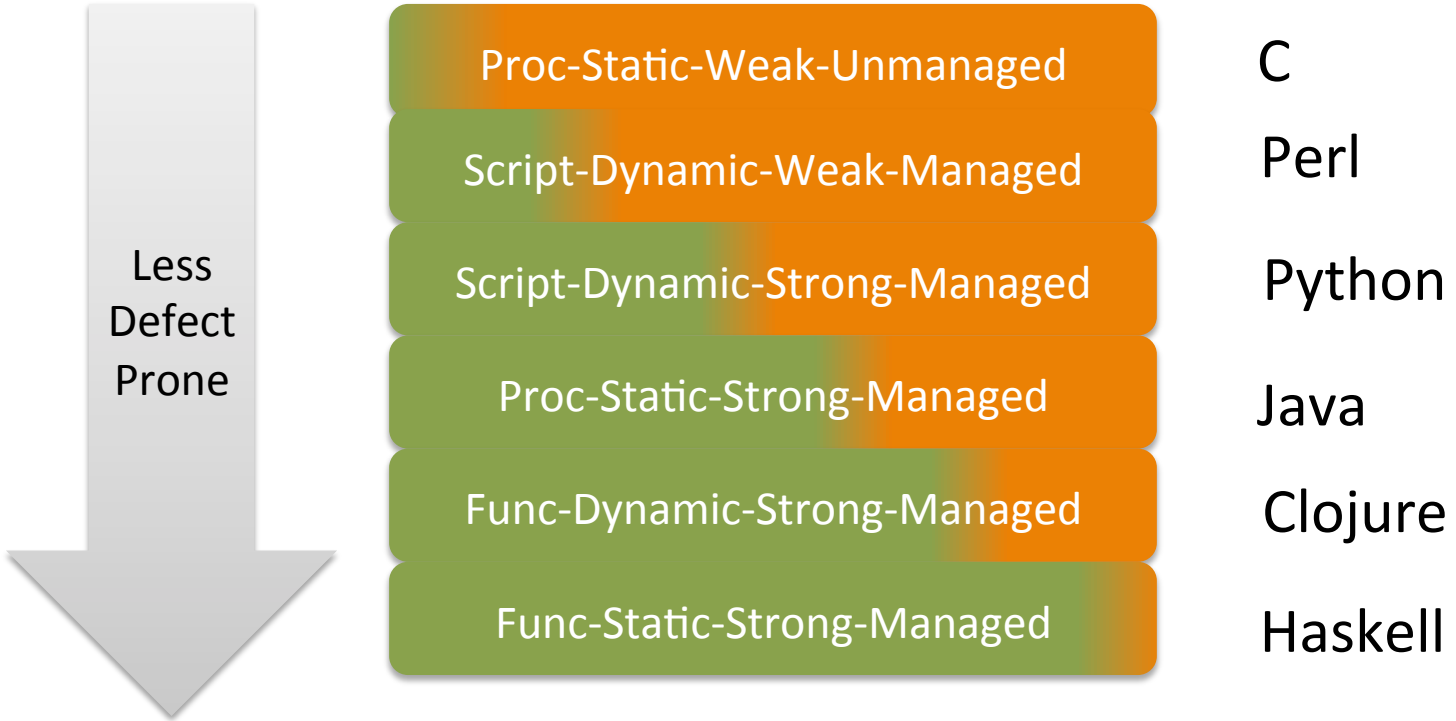
More bug-prone than average

Some languages have a greater association with defects than other languages, although ***the effect is small.***

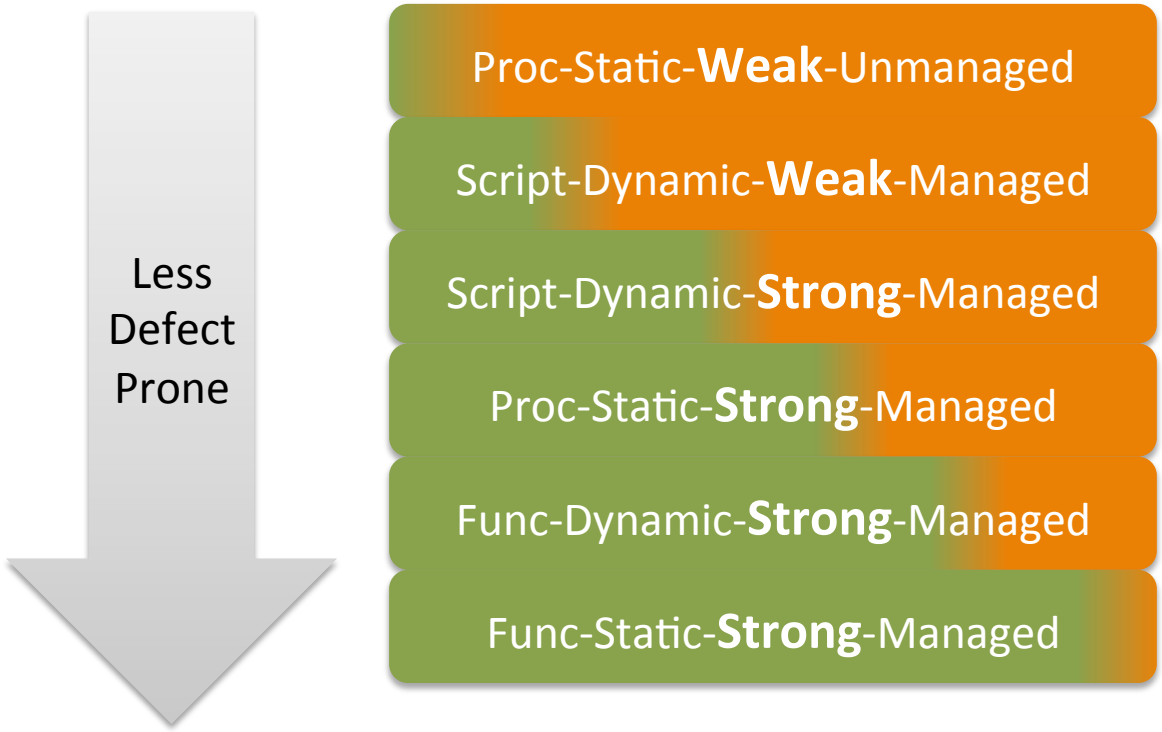
RQ2. Which language properties relate to defects?

- Procedural / Scripting / Functional ?
- Static / Dynamic ?
- Strong / Weak ?
- Managed / Unmanaged ?

RQ2. Which language properties relate to defects?

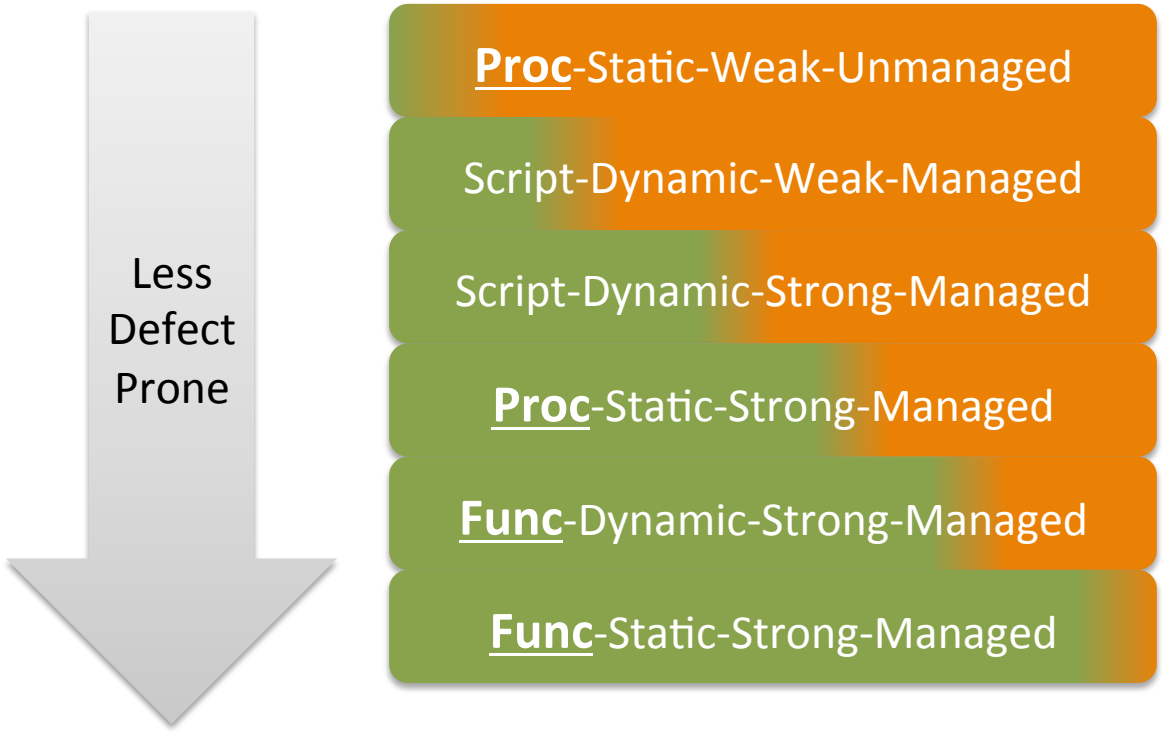


RQ2. Which language properties relate to defects?



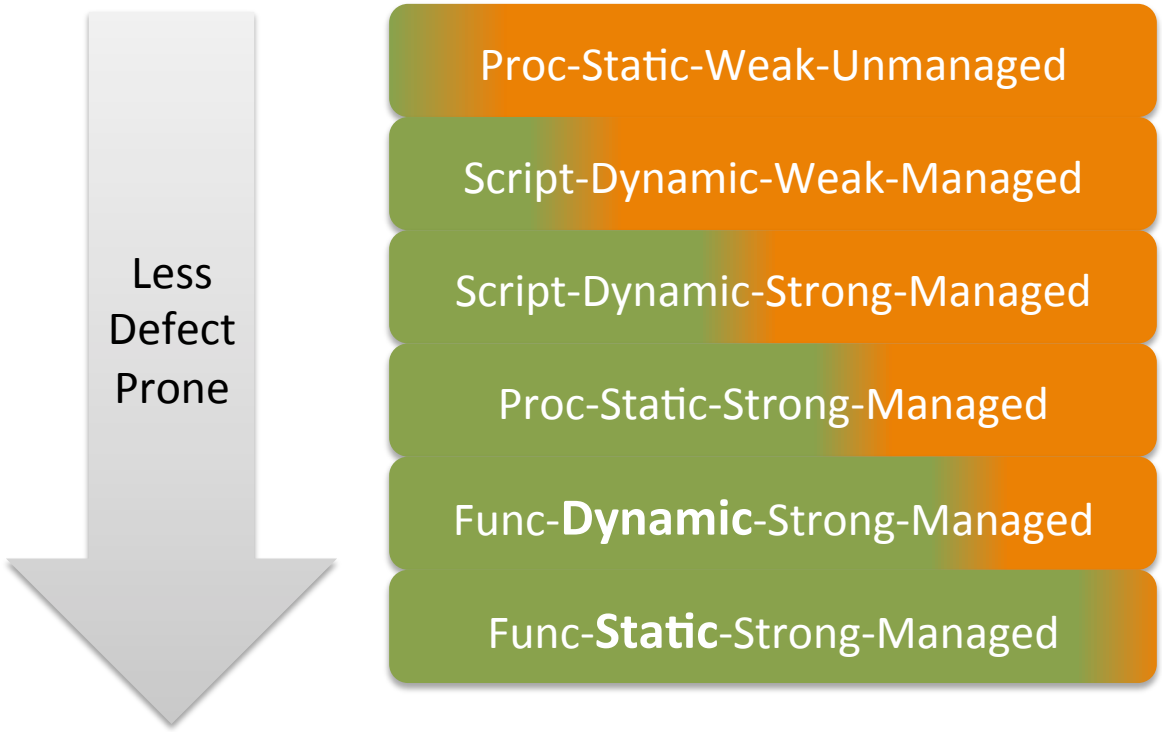
Strong typing is a bit better than weak tying

RQ2. Which language properties relate to defects?



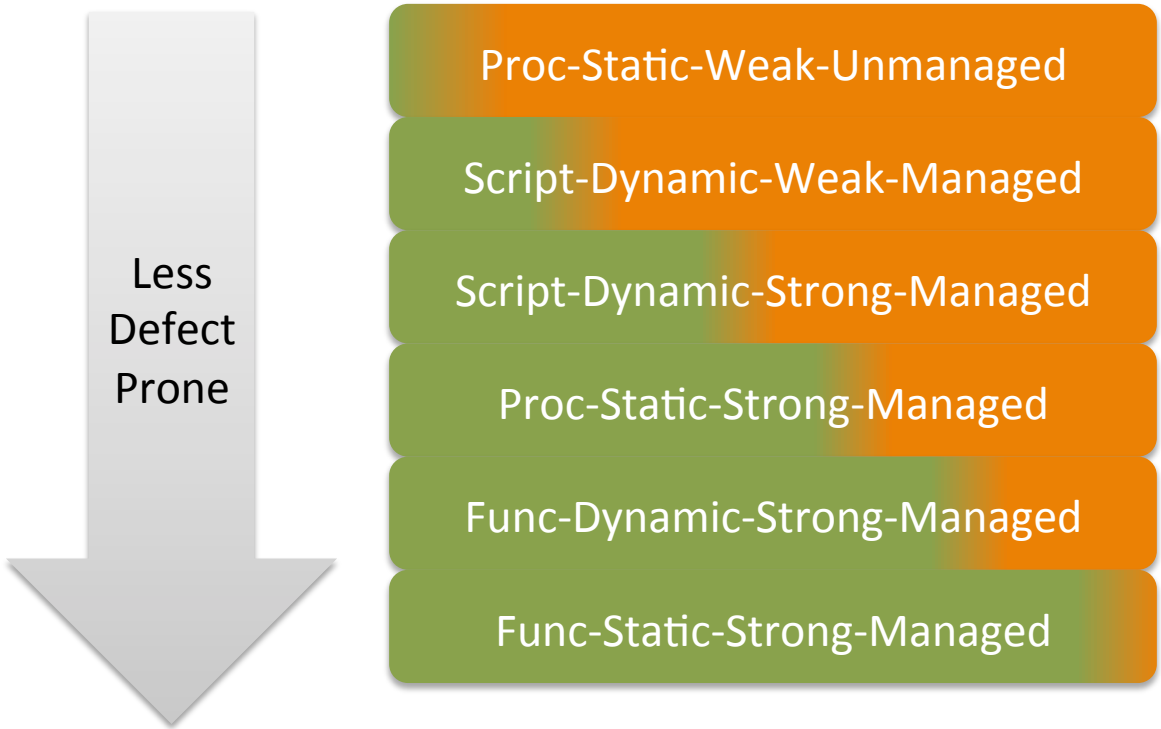
Functional Languages > Procedural Languages

RQ2. Which language properties relate to defects?



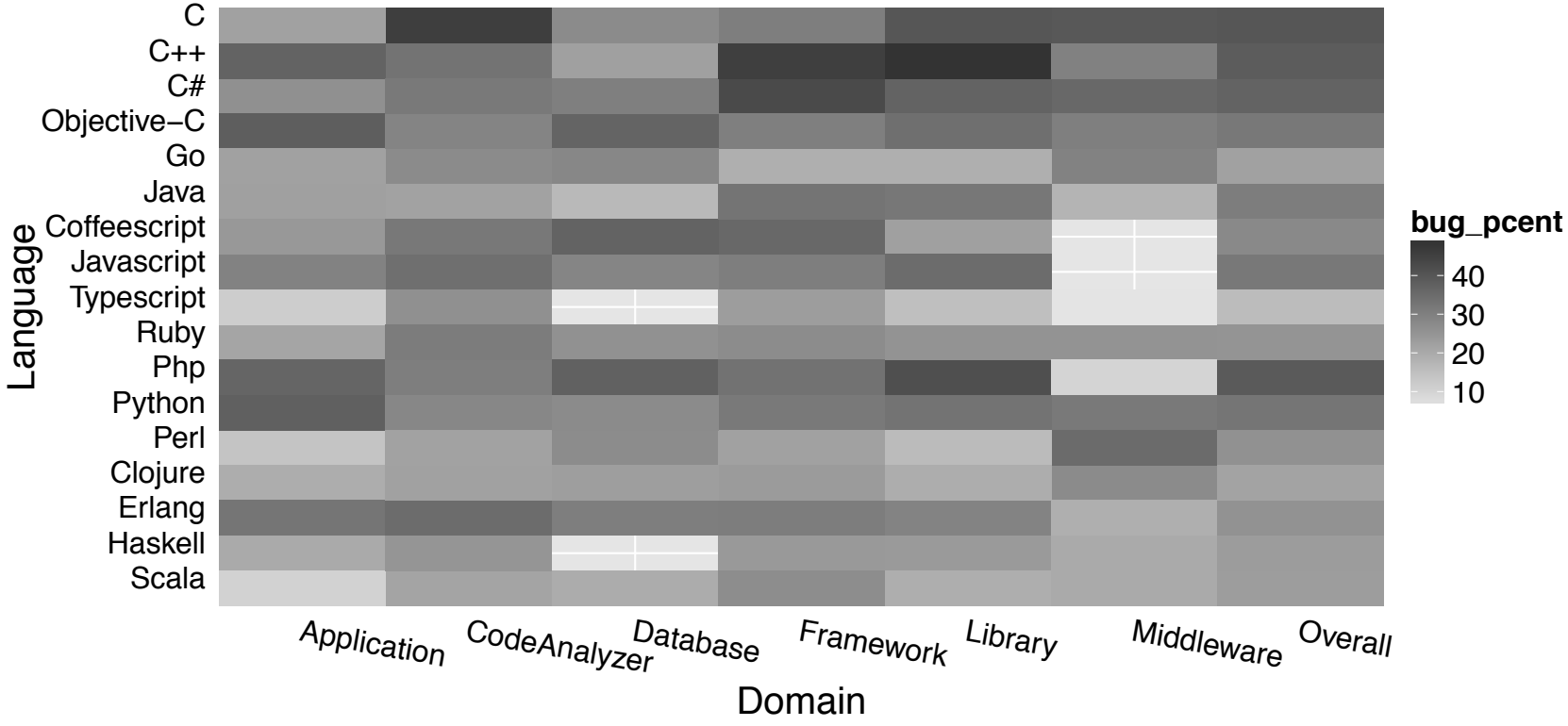
(For Functional) Static Typing > Dynamic Typing

RQ2. Which language properties relate to defects?



There is a small but significant relationship between language class and defects.

RQ3. Does language defect proneness depend on domain?

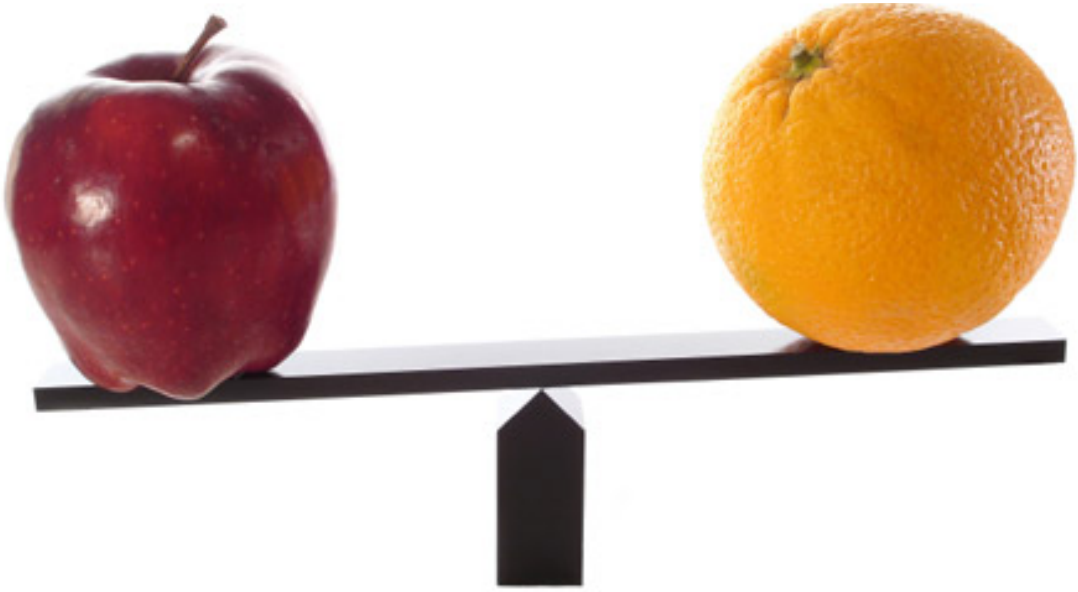


There is no general relationship between domain and language defect proneness.

Threats to Validity

- The personality of programmers choosing the language may matter more.
- Type of languages are classified based on file extensions.
- Focus on development time bugs.
- Bugs are categorized by the choice of keywords.
- External tools may increase bug detection in certain languages.

Summary



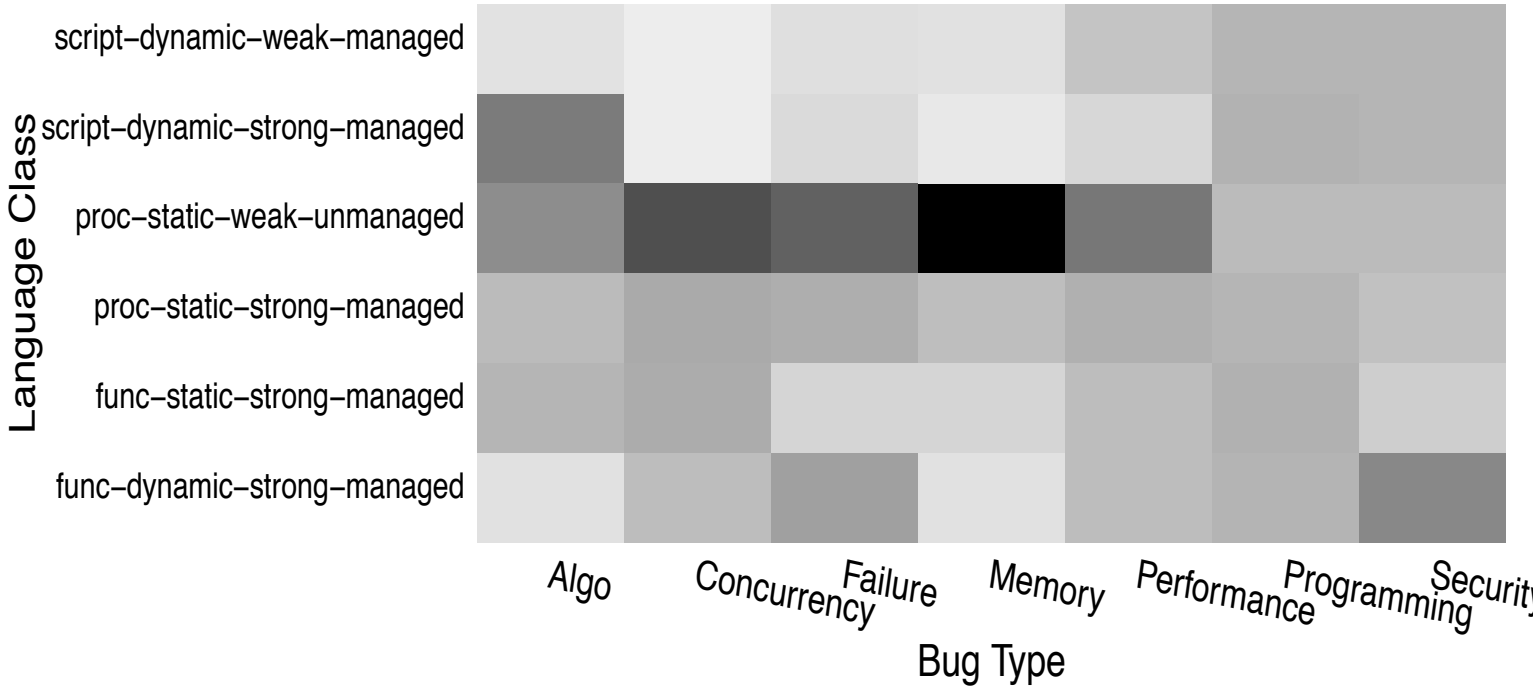
“Language Features matter, but, not much.”

Thanks to NSF CCF-1247280 and CCF-1446683
AFOSR FA955-11-1-0246

RQ4. What is the relation between language & bug category?

Bug Type	Count (%)
Algorithm	0.11
Concurrency	1.99
Memory	5.44
Security	2.01
Performance	1.55
Failure	3.77
Unknown	1.04
Programming	88.53

RQ4. What is the relation between language & bug category?



Language matters more for specific categories than it does for defects overall.