

Preventing Errors Before They Happen

<http://CheckerFramework.org/>



Werner Dietl
University of Waterloo
<https://ece.uwaterloo.ca/~wdietl/>



Joint work with Michael D. Ernst and many others.

Bug Evolution

Photo # NH 96566-KN (Color) First Computer "Bug", 1947

92

9/9

0800

Auton started

1000

stopped - auton ✓

$$\left\{ \begin{array}{l} 1.2700 \quad 9.037847025 \\ \quad \quad \quad 9.037846995 \text{ correct} \\ \hline 1.30476415 \end{array} \right.$$

~~13rd col (033) MP - MC~~

(033) PRO 2 2.130476415

correct 2.130476415

Relays 6-2 in 033 failed special speed test
in relay 10.000 test.

Relay 2145
Relay 3370

1100

Started Cosine Tape (Sine check)

1525

Started Multi Adder Test.

1545



Relay #70 Panel F
(moth) in relay.

1630

First actual case of bug being found.

Auton started.

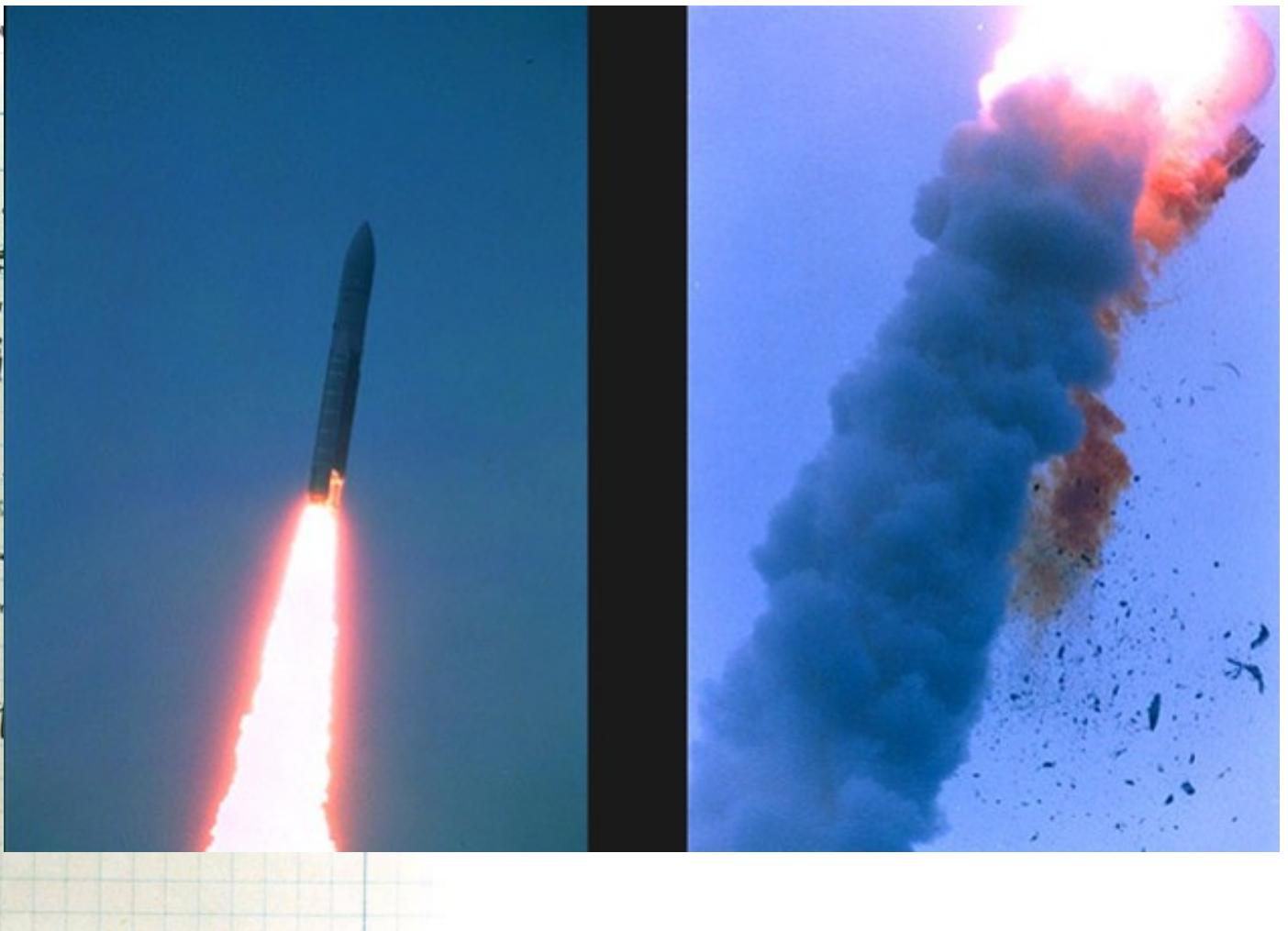
1700 closed down.

Bug Evolution

92

9/9

Photo # NH 96566-KN (Color) First Computer Bug
0800 Auton started
1000 " stopped - auton ✓
13° C (032) MP - MC 2.1307
(033) PRO 2 2.1304
couch 2.1306
Relays 6-2 in 033 failed
in relay
Relays changed
Started Cosine Tape (Sin)
1525 Started Multi Adder Test
1545 Relay (moth)
First actual case of bug
1630 Auton started.
1700 closed down.



<http://dlimages.businessweek.com/imageserve/0arRbTIdRSg3e/630x418.jpg>

Cost of software failures

\$312 billion per year global cost of software bugs (2013)

\$300 billion dealing with the Y2K problem

\$650 million loss by NASA Mars missions in 1999; unit conversion bug

\$500 million Ariane 5 maiden flight in 1996; 64 bit to 16 bit conversion bug

\$440 million loss by Knight Capital Group Inc. in 30 minutes in August 2012

Software bugs can cost lives

225 deaths in 1997: jet crash caused by radar software

28 deaths in 1991: Patriot missile guiding system failure

Radiation therapy

- at least three deaths in 1985-87 by Therac-25
- at least five deaths in 2000 by Multidata

11 deaths and \$6 billion estimated damages in 2003:
blackout in Northeast USA & Canada; 50 million people
without power; monitoring software bug

Software cause for quarter of all medical device recalls in
2011

November 2012 medical pump recall

Outline

- Pluggable type-checking
- Quick demo
- Architecture overview
 - Java 8 syntax for Type Annotations
 - Checker Framework
- Writing your own type system
- Nullness Checker in detail
- Our experience
 - SPARTA: Android Security
 - Verigames: Crowd-sourced Verification Games

Outline

- • Pluggable type-checking
- Quick demo
- Architecture overview
 - Java 8 syntax for Type Annotations
 - Checker Framework
- Writing your own type system
- Nullness Checker in detail
- Our experience
 - SPARTA: Android Security
 - Verigames: Crowd-sourced Verification Games

Java's type system is too weak

Type checking prevents many errors

```
int i = "hello"; // error
```

Type checking doesn't prevent enough errors

```
System.console().readLine();
```

```
Collections.emptyList().add("one");
```

```
dbStatement.executeQuery(userData);
```

Java's type system is too weak

Type checking prevents many errors

```
int i = "hello"; // error
```

Type checking can't prevent all errors

NullPointerException

```
System.console().readLine();
```

```
Collections.emptyList().add("one");
```

```
dbStatement.executeQuery(userData);
```

Java's type system is too weak

Type checking prevents many errors

```
int i = "hello"; // error
```

Type checking doesn't prevent enough errors

System **UnsupportedOperationException**

```
Collections.emptyList().add("one");
```

```
dbStatement.executeQuery(userData);
```

Java's type system is too weak

Type checking prevents many errors

```
int i = "hello"; // error
```

Type checking doesn't prevent enough errors

```
System.console().readLine();
```

Collections

SQL Injection Attacks!

```
dbStatement.executeQuery(userData);
```

Null pointer exception

```
String op(Data in) {  
    return "transform: " + in.getF();  
}  
...  
String s = op(null);
```

Where is the error?

Null pointer exception

```
String op(Data in) {  
    return "transform: " + in.getF();  
}  
...  
String s = op(null);
```

Where is the error?

Can't decide without specification!

Solution 1: Restrict use

```
String op(@NotNull Data in) {  
    return "transform: " + in.getF();  
}  
...  
String s = op(null);      // error
```

Solution 2: Restrict implementation

```
String op(@Nullable Data in) {  
    return “transform: ” + in.getF();  
                           // error  
}  
  
...  
  
String s = op(null);
```

Benefits of type systems

- **Find bugs** in programs
 - Guarantee the **absence of errors**
- **Improve documentation**
 - Improve code structure & maintainability
- Aid compilers, optimizers, and analysis tools
 - Reduce number of run-time checks

Possible negatives:

- Must write the types (or use type inference)
- False positives are possible (can be suppressed)

Extended type systems

- Null pointer exceptions (ICSE SEIP'11)
- Energy consumption (PLDI'11)
- Unwanted side effects (OOPSLA'04, '05, '12, ECOOP'08, ESEC/FSE'07)
- Unstructured heaps (ECOOP'07, '11, '12)
- Malformed input (FTfJP'12)
- UI actions not on event thread (ECOOP'13)
- Information leakage (CCS'14)

Extended type systems

- Null pointer exceptions (ICSE SEIP'11)
- Energy consumption (PLDI'11)
- Unwanted side effects (OOPSLA'04, '05, '12, ECOOP'08, ESEC/FSE'07)
- Unstructured heaps (ECOOP'07, '11, '12)
- Malformed input (FTfJP'12)
- UI actions not on event thread (ECOOP'13)
- Information leakage (CCS'14)



The Checker Framework

A framework for pluggable type checkers
“Plugs” into the OpenJDK compiler

```
javac -processor MyChecker ...
```

Eclipse plug-in, Ant and Maven integration

Input Format Validation

Demo: ensure that certain strings contain valid regular expressions.

Regular Expression Example

```
public static void main(String[] args) {  
    String regex = args[0];  
    String content = args[1];  
    Pattern pat = Pattern.compile(regex);  
    Matcher mat = pat.matcher(content);  
  
    if (mat.matches()) {  
        System.out.println("Group: " +  
            mat.group(1));  
    }  
}
```

Regular Expression Example

```
public static void main(String[] args) {  
    String regex = "PatternSyntaxException";  
    String content = "IndexOutOfBoundsException on  
    Pattern pat = Pattern.compile(regex);  
    Matcher mat = pat.matcher(content);  
  
    if (mat.find()) {  
        System.out.println("Group: " +  
            mat.group(1));  
    }  
}
```

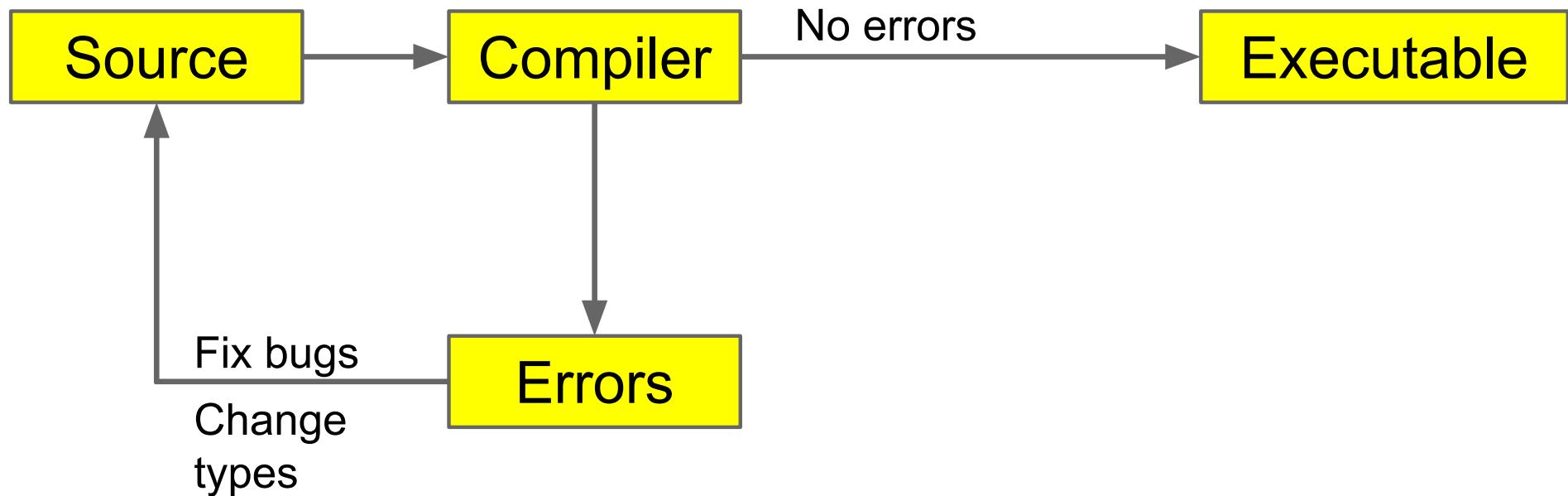
Fixing the Errors

Pattern.compile only on valid regex
Matcher.group(i) only if > i groups

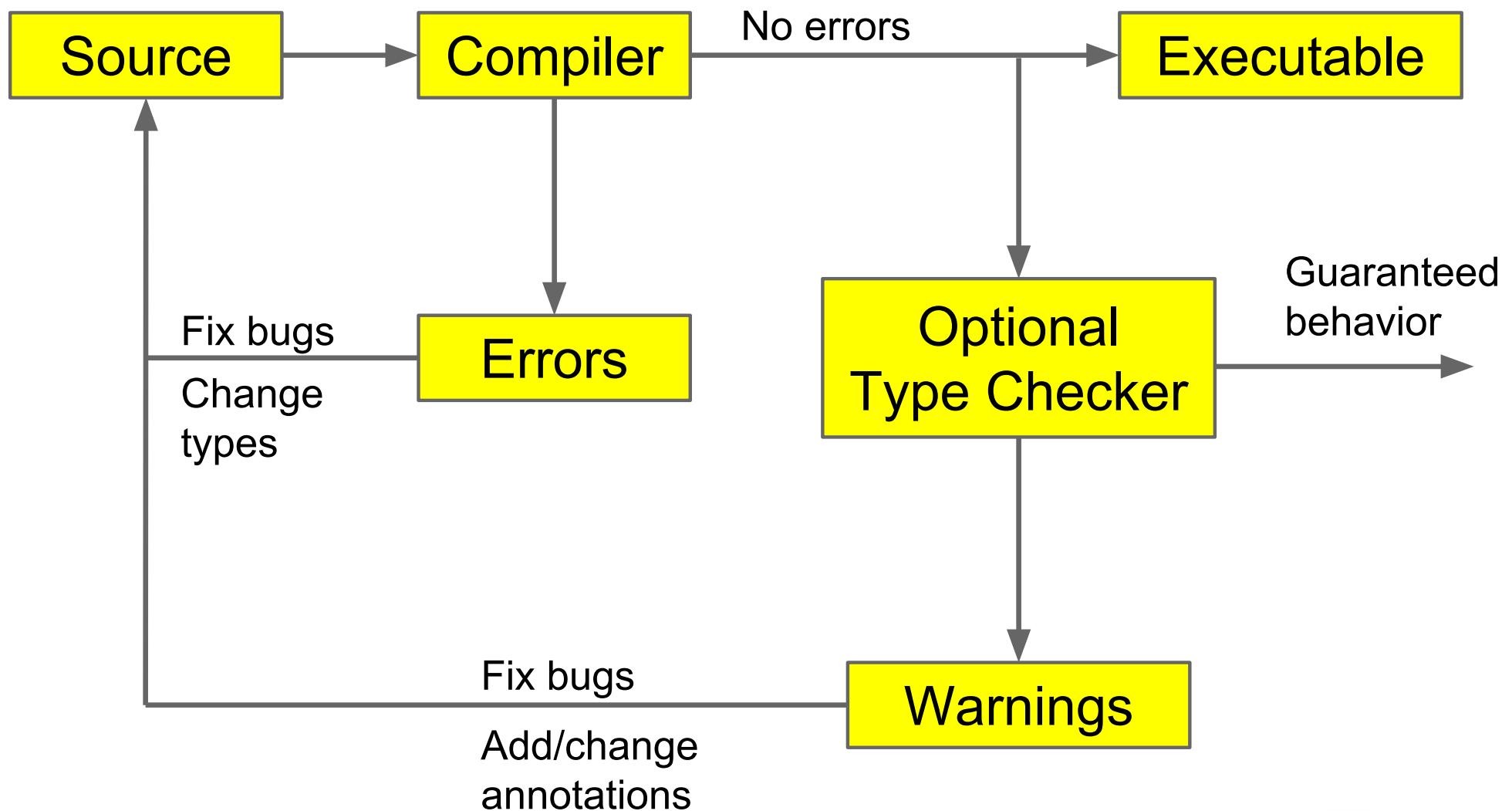
...

```
if (!RegexUtil.isRegex(regex, 1)) {  
    System.out.println("Invalid: " + regex);  
    System.exit(1);  
}  
...
```

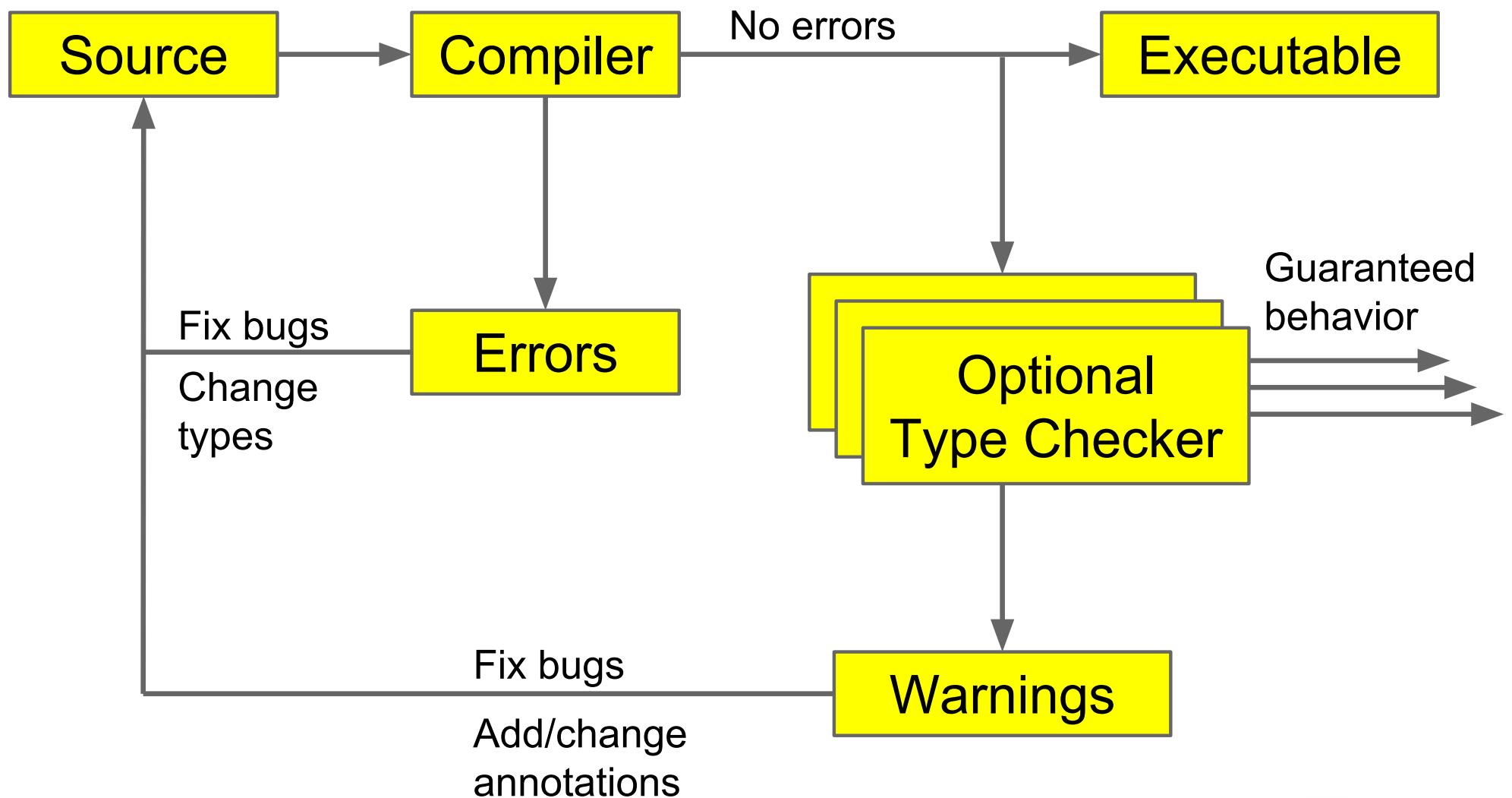
Type Checking



Optional Type Checking



Optional Type Checking



What bugs can you detect & prevent?

The property you care about:

Null dereferences

Mutation and side-effects

Concurrency: locking

Security: encryption,
tainting

Aliasing

Equality tests

Strings: localization,

Regular expression syntax

Signature format

Enumerations

Typestate (e.g., open/closed files)

The annotation you write:

@NonNull

@Immutable

@GuardedBy

@Encrypted

@Untainted

@Linear

@Interned

@Localized

@Regex

@FullyQualified

@Fenum

@State

Users can write their own checkers!

Checker Framework experience

Type checkers reveal important latent defects

Ran on >6 million LOC of real-world
open-source code

Found hundreds of user-visible failures

Improved design and documentation

Annotation overhead is low

Mean 2.6 annotations per kLOC [Dietl et al. ICSE'11]

Formalizations

$P \in \text{Program} ::= \overline{\text{Class}}, \text{ClassId}, \overline{\text{Expr}}$	$h \in \text{Heap}$	$= \text{Addr} \rightarrow \text{Obj}$
$\text{Cls} \in \text{Class} ::= \text{class ClassId} < \overline{\text{TVarId}} \text{ extends ClassId} < \overline{\text{sType}}$	$\iota \in \text{Addr}$	$= \text{Set of Addresses} \cup \{\text{null}_a\}$
$\{ \text{FieldId} \text{ } \overline{\text{sType}}; \text{ Met}$	$\circ \in \text{Obj}$	$= \overline{\text{rType, Fields}}$
	$\overline{\text{rT}} \in \overline{\text{rType}}$	$= \text{OwnerAddr ClassId} < \overline{\text{rType}}$
	$Fs \in \text{Fields}$	$= \text{FieldId} \rightarrow \text{Addr}$
	$\iota \in \text{OwnerAddr}$	$= \overline{\text{Addr} \cup \{\text{any}_a\}}$
	$\overline{\text{rGamma}} \in \overline{\text{rEnv}}$	$= \overline{\text{TVarId rType; ParId Addr}}$
$\overline{\text{sT}} \in \overline{\text{sType}} ::= \overline{\text{sNType}} \mid \overline{\text{TVarId}}$		$h, \overline{\text{rGamma}}, e_0 \rightsquigarrow h', \iota_0$
$\overline{\text{sN}} \in \overline{\text{sNType}} ::= \text{OM ClassId} < \overline{\text{sType}}$		$\iota_0 \neq \text{null}_a$
$u \in \text{OM} ::= h, \overline{\text{rGamma}}, e_0 \rightsquigarrow h_0, \iota_0$		
$m \in \text{Meth} ::= \iota_0 \neq \text{null}_a$		
$\text{MethSig} ::= h_0, \overline{\text{rGamma}}, e_2 \rightsquigarrow h_2, \iota$		
$w \in \text{Purity} ::= h' = h_2[\iota_0.f := \iota]$		
$e \in \text{Expr} ::= \text{OS-Upd} \frac{h' = h_2[\iota_0.f := \iota]}{h, \overline{\text{rGamma}}, e_0.f = e_2 \rightsquigarrow h'},$		
		$\overline{\text{Expr.MethId} < \overline{\text{sType}}}(\text{Expr}) \mid \text{new } \overline{\text{sType}} \mid (\overline{\text{sType}}) \text{ Expr}$
$\overline{\text{sGamma}} \in \overline{\text{sEnv}} ::= \overline{\text{TVarId sNType; ParId sType}}$		
$h \vdash \overline{\text{rGamma}} : \overline{\text{sGamma}}$		
$h \vdash \iota_1 : dyn(\overline{\text{sN}}, h, \overline{\text{rGamma}})$		
$h \vdash \iota_2 : dyn(\overline{\text{sT}}, \iota_1, h(\iota_1) \downarrow_1)$		
$\overline{\text{sN}} = u_N \ C_N \leftrightarrow$		
$u_N = \text{this}_u \Rightarrow \overline{\text{rGamma}}(\text{this})$		
$free(\overline{\text{sT}}) \subseteq \text{dom}(C_N)$		
DYN		

$$\text{OS-Read} \quad \frac{}{h, \overline{\text{rGamma}}, e_0.f \rightsquigarrow h', \iota}$$

$$\frac{\Gamma \vdash e_0 : N_0 \quad N_0 = u_0 \ C_0 \leftrightarrow T_1 = fType(C_0, f) \quad \Gamma \vdash e_2 : N_0 \triangleright T_1 \quad u_0 \neq \text{any} \quad rp(u_0, T_1)}{\Gamma \vdash e_0.f = e_2 : N_0 \triangleright T_1}$$

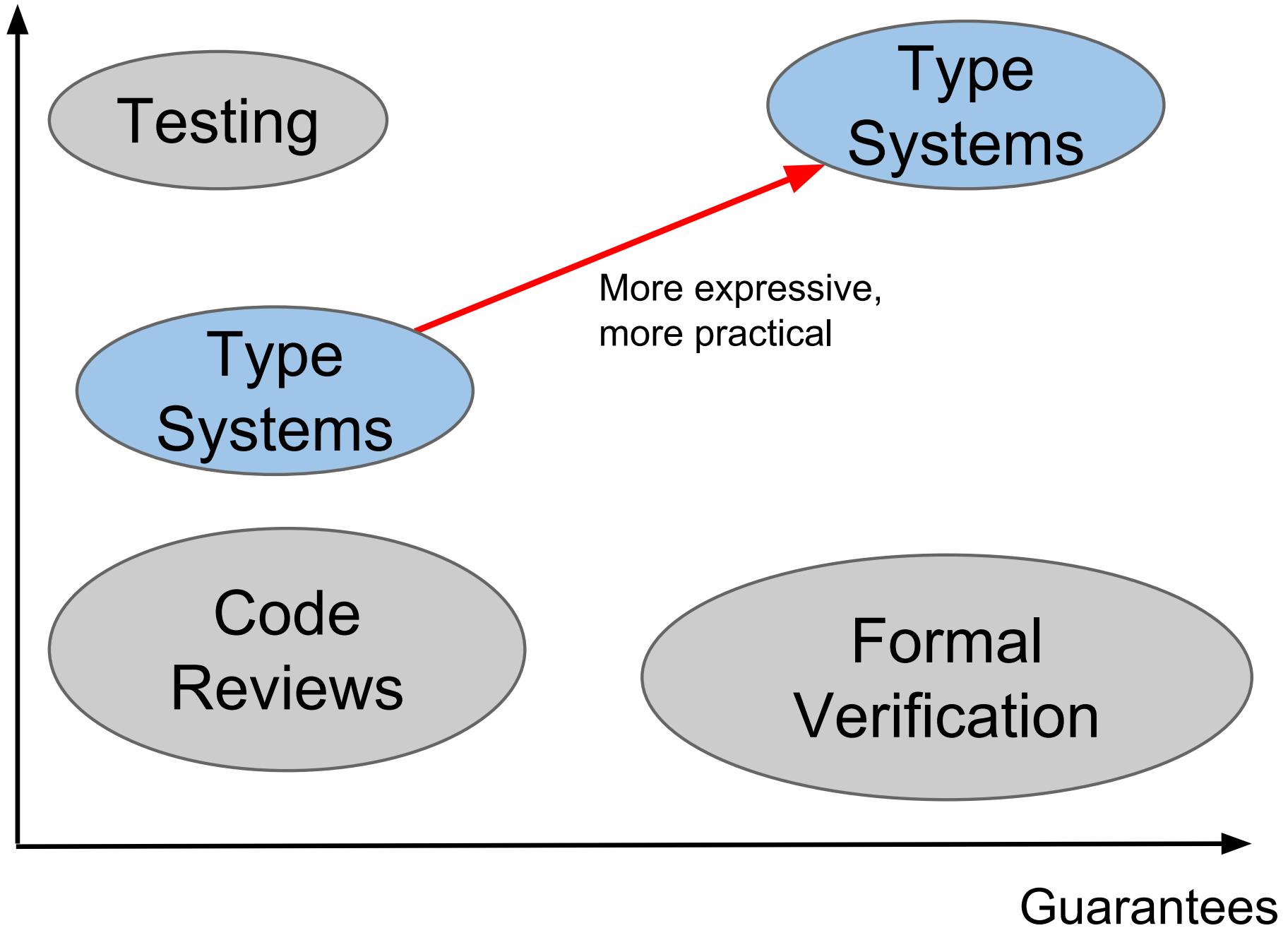
$$\text{GT-Read} \quad \frac{\Gamma \vdash e_0 : N_0 \quad N_0 = -}{\Gamma \vdash e_0.f : N_0 \triangleright fType(C_0, f)}$$

$$\left. \begin{array}{c} \Rightarrow h \vdash \iota_2 : dyn(\overline{\text{sN}} \triangleright \overline{\text{sT}}, h, \overline{\text{rGamma}}) \\ \overline{\text{rT}} = \iota' _ \leftrightarrow \quad \iota \vdash \overline{\text{rT}} \ r <: \iota' \ C < \overline{\text{rT}} \quad \iota \vdash \overline{\text{rT}} \ r <: \iota' \ C < \overline{\text{rT}} \Rightarrow \iota \vdash \overline{\text{rT}} \ r <: \overline{\text{rT}} \\ \text{dom}(C) = \bar{X} \quad \text{free}(\overline{\text{sT}}) \subseteq \bar{X} \circ \bar{X}' \end{array} \right\}$$

$$dyn(\overline{\text{sT}}, \iota, \overline{\text{rT}}, (\overline{\text{X}}' \ \overline{\text{rT}}'; -)) = \overline{\text{sT}}[\iota'/\text{this}, \iota'/\text{peer}, \iota/\text{rep}, \text{any}_a/\text{any}_u, \overline{\text{rT}}/\bar{X}, \overline{\text{rT}}'/\bar{X}']$$

Practicality

Allow reliable and secure
programming in practice



Outline



- Pluggable type-checking
- Quick demo
- Architecture overview
 - Java 8 syntax for Type Annotations
 - Checker Framework
- Writing your own type system
- Nullness Checker in detail
- Our experience
 - SPARTA: Android Security
 - Verigames: Crowd-sourced Verification Games

Since Java 5: annotations

Only for declaration locations:

@Deprecated

```
class Foo {
```

@Getter @Setter private String query;

@SuppressWarnings("unchecked")

```
void foo() { ... }
```

```
}
```

But we couldn't express

A non-null reference to my data

An interned String

A non-null List of English Strings

A read-only array of non-empty arrays of
English strings

With Java 8 Type Annotations we can!

A non-null reference to my data

```
@NotNull Data mydata;
```

An interned String

```
@Interned String query;
```

A non-null List of English Strings

```
@NotNull List<@English String> msgs;
```

A read-only array of non-empty arrays of English strings:

```
@English String @ReadOnly [] @NonEmpty [] a;
```

Java 8 extends annotation syntax

Annotations on all occurrences of types:

```
@Untainted String query;  
List<@NonNull String> strings;  
myGraph = (@Immutable Graph) tmp;  
class UnmodifiableList<T>  
    implements @ReadOnly List<T> {}
```

Stored in classfile

Handled by javac, javap, javadoc, ...

Array annotations

A **read-only** array of non-empty arrays of English strings:

```
@English String @ReadOnly [] @NonEmpty [] a;
```

Explicit method receivers

```
class MyClass {  
    int foo(@TParam String p) {...}  
    int foo(@TRecv MyClass this,  
            @TParam String p) {...}
```

No impact on method binding and overloading

Constructor return & receiver types

Every constructor has a return type

```
class MyClass {  
    @TReturn MyClass(@TParam String p) {...}
```

Inner class constructors also have a receiver

```
class Outer {  
    class Inner {  
        @TReturn Inner(@TRecv Outer Outer.this,  
                      @TParam String p) {...}}
```

CF: Java 6 & 7 Compatibility

Annotations in comments

```
List</*@NotNull*/ String> strings;
```

Voodoo comments for arbitrary source code

```
/*>>> import myquals.TRecv; */
```

```
...
```

```
int foo(/*>>> @TRecv MyClass this,*/
        @TParam String p) {...}
```

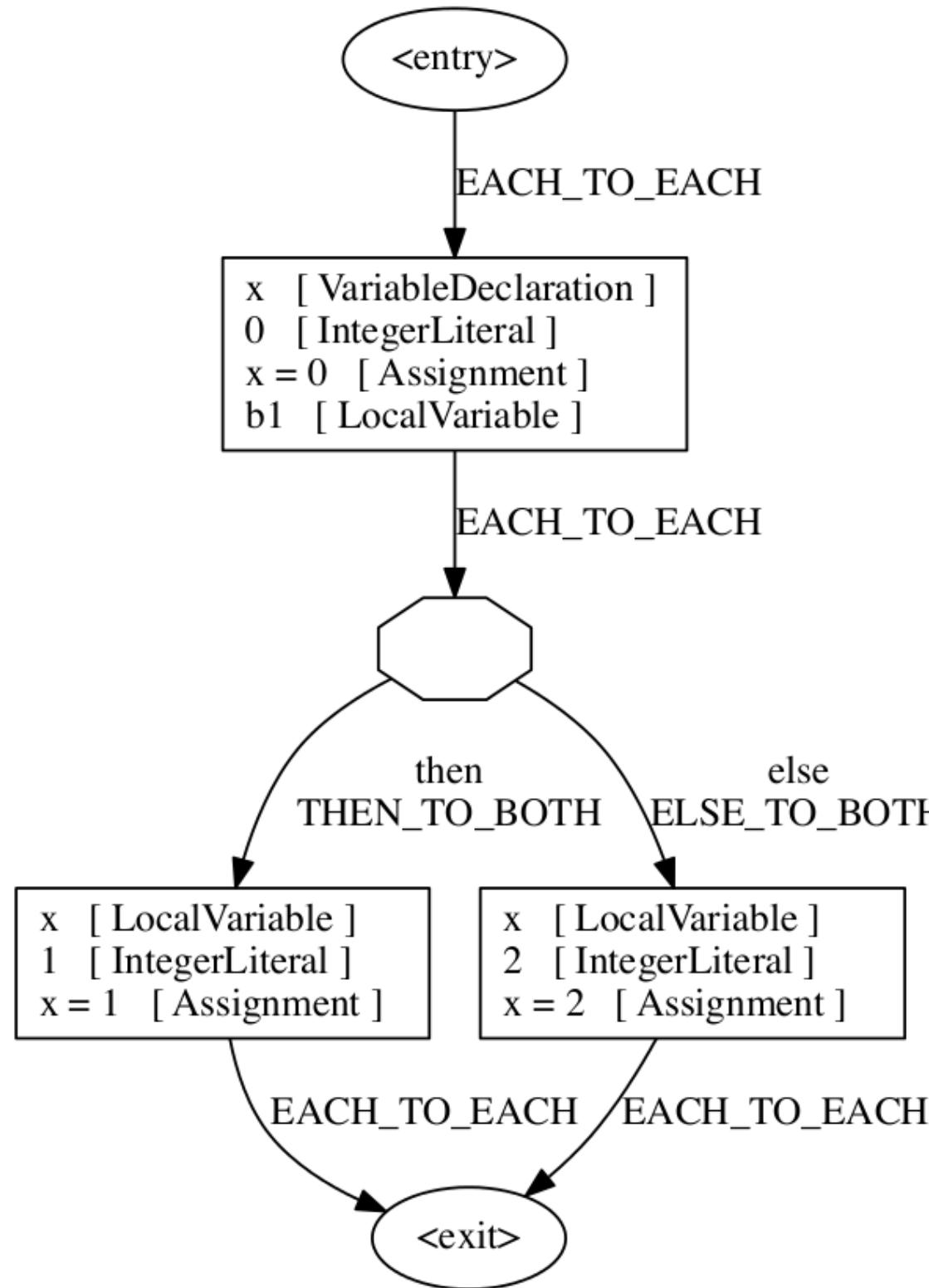
CF: stub files and annotated JDK

Allows annotating external libraries

- As separate text file (stub file)
checker-framework/checker/src/.../jdk.astub
- Within its .jar file (from annotated partial source code)
checker-framework/checker/jdk/...

Dataflow Framework

1. Translate AST to CFG
 - Standard multipass visitor over AST
2. Perform dataflow analysis over CFG with user-provided
 - a. Abstract value What are we tracking?
 - b. Transfer functions What do operations do?
 - c. Store What are intermediate results?
3. Allow queries about result



Checker Framework - “Framework”

- Full type systems: inheritance, overriding, ...
- Generics (type polymorphism)
 - Also qualifier polymorphism
- Flow-sensitive type qualifier inference
 - Infers types for local variables
 - reusable component
- Qualifier defaults
- Pre-/Post-conditions
- Warning suppression
- Testing infrastructure

Checker Framework type systems

The property you care about:

Null dereferences

Mutation and side-effects

Concurrency: locking

Security: encryption,
tainting

Aliasing

Equality tests

Strings: localization,

Regular expression syntax

Signature format

Enumerations

Typestate (e.g., open/closed files)

The annotation you write:

@NonNull

@Immutable

@GuardedBy

@Encrypted

@Untainted

@Linear

@Interned

@Localized

@Regex

@FullyQualified

@Fenum

@State

Outline



- Pluggable type-checking
- Quick demo
- Architecture overview
 - Java 8 syntax for Type Annotations
 - Checker Framework
- Writing your own type system
- Nullness Checker in detail
- Our experience
 - SPARTA: Android Security
 - Verigames: Crowd-sourced Verification Games

Building checkers is easy

Example: Ensure encrypted communication

```
void send(@Encrypted String msg) {...}  
@Encrypted String msg1 = ...;  
send(msg1); // OK  
String msg2 = ...;  
send(msg2); // Warning!
```

Building checkers is easy

Example: Ensure encrypted communication

```
void send(@Encrypted String msg) {...}  
@Encrypted String msg1 = ....;  
send(msg1); // OK  
String msg2 = ....;  
send(msg2); // Warning!
```

The complete checker:

```
@Target(ElementType.TYPE_USE)  
@SubtypeOf(Unqualified.class)  
public @interface Encrypted {}
```

Let's try this!

Defining a type system

1. Qualifier hierarchy
 - defines subtyping
2. Type introduction rules
 - types for expressions
3. Type rules
 - checker-specific errors
4. Flow-refinement
 - checker-specific flow properties

Testing Infrastructure

jtreg-based testing as in OpenJDK

Light-weight tests with in-line expected errors:

```
String s = "%+s%";  
//:: error: (format.string.invalid)  
f.format(s, "illegal");
```

Brainstorming new type checkers

What runtime exceptions do you wish to prevent?

What properties of data should always hold?

What operations are legal and illegal?

Type-system checkable properties:

- Dependency on values
- Not on program structure, timing, ...

Brainstorming

Outline

- Pluggable type-checking
- Quick demo
- Architecture overview
 - Java 8 syntax for Type Annotations
 - Checker Framework
- Writing your own type system
- • Nullness Checker in detail
- Our experience
 - SPARTA: Android Security
 - Verigames: Crowd-sourced Verification Games

Preventing Null-Pointer Exceptions

Basic type system:

<code>@Nullable</code>	might be null
<code>@NonNull</code>	non-null

`@Nullable`

`@NonNull`



Default is `@NonNull`

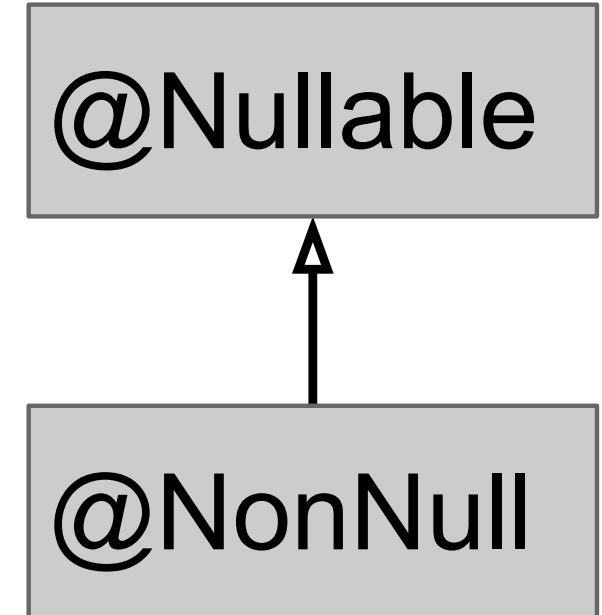
(Opposite of Java's assumption)

- Makes the dangerous case explicit
- Requires fewer annotations

CLIMB-to-top defaulting rule (applies to all type systems)

Top type is the default for:

- Casts
- Local variables
- instanceof
- iMplicit Bounds



Type refinement from assignments

```
myVar = new Foo();
```

Likewise for cast/instanceof expression

Top for implicit types allows every instantiation

Dynamic checks

```
if (x != null) {  
    x.hashCode();  
}
```

```
if (!RegexUtil.isRegex(userInput)) {  
    throw new RuntimeException(...);  
}  
Pattern p = Pattern.compile(userInput);
```

One check for null is not enough

```
if (myField != null) {  
    method1();  
    myField.hashCode();  
}
```

One check for null is not enough

```
if (myField != null) {  
    method1();  
    myField.hashCode();  
}
```

3 ways to express persistence across side effects:

`@SideEffectFree void method1() { ... }`

`@MonotonicNotNull myField;`

`@EnsuresNotNull("myField") method1(){...}`

Side effects

@SideEffectFree

Does not modify externally-visible state

@Deterministic

If called with == args again, gives == result

@Pure

Both side-effect-free and deterministic

The side-effect annotations are trusted, not checked

Lazy initialization and persistence across side effects

@MonotonicNonNull

Might be null or non-null

May only be (re-)assigned a non-null value

Purpose: avoid re-checking

Once non-null, always non-null

Method pre- and post-conditions

Preconditions:

@RequiresNonNull

Postconditions:

@EnsuresNonNull

@EnsuresNonNullIf

@EnsuresNonNullIf(expression="#1", result=true)

```
public boolean equals(@Nullable Object obj) { ... }
```

A non-null field might contain null

```
@NotNull String name;  
... myObject.name ...
```

Initialization

`@Initialized` (constructor has completed)

`@UnderInitialization(Frame.class)`

Its constructor is currently executing

`@UnknownInitialization`

Might be initialized or under initialization

Suppressing warnings

Because of Checker Framework false positives

`@SuppressWarnings("nullness")`

Use smallest possible scope (e.g., local var)

Write the rationale

```
assert x != null : "@AssumeAssertion(nullness);"
```

More: <http://types.cs.washington.edu/checker-framework/current/checker-framework-manual.html#suppressing-warnings>

How to get started

1. Write the specification

Search the Javadoc for occurrences of “null”

Replace the wordy English text by `@Nullable`

Can also search code, but no annos in methods

2. Run Nullness Checker: verify/complete spec

For each warning

Reason about why the code is safe

Express that reasoning as annotations

Consider improving the code’s design

Tips

- Write the spec first (and think of it as a spec)
- Avoid warning suppressions when possible
- Avoid raw types such as `List`; use
`List<String>`
- Start by type-checking part of your code
- Only type-check properties that matter to you
- Use subclasses (not type qualifiers) if possible

Outline

- Pluggable type-checking
- Quick demo
- Architecture overview
 - Java 8 syntax for Type Annotations
 - Checker Framework
- Writing your own type system
- Nullness Checker in detail
- Our experience
 - SPARTA: Android Security
 - Verigames: Crowd-sourced Verification Games



Evaluations

- Checkers reveal important latent bugs
 - Ran on >6 million LOC of real-world code
 - Found hundreds of user-visible bugs & mistakes
- Annotation overhead is low
 - Mean 2.6 annotations per kLOC
- Learning their usage is easy
 - Used successfully by first-year CS majors
- Building checkers is easy
 - New users developed 3 new realistic checkers

SPARTA: Static Program Analysis for Reliable Trusted Apps

Security type system for Android apps
Guarantees no leakage of private information

Part of
Automated Program Analysis
for Cybersecurity (APAC)
program.
See CCS'14 paper



Crowd-sourced verification

Make software verification easy and fun

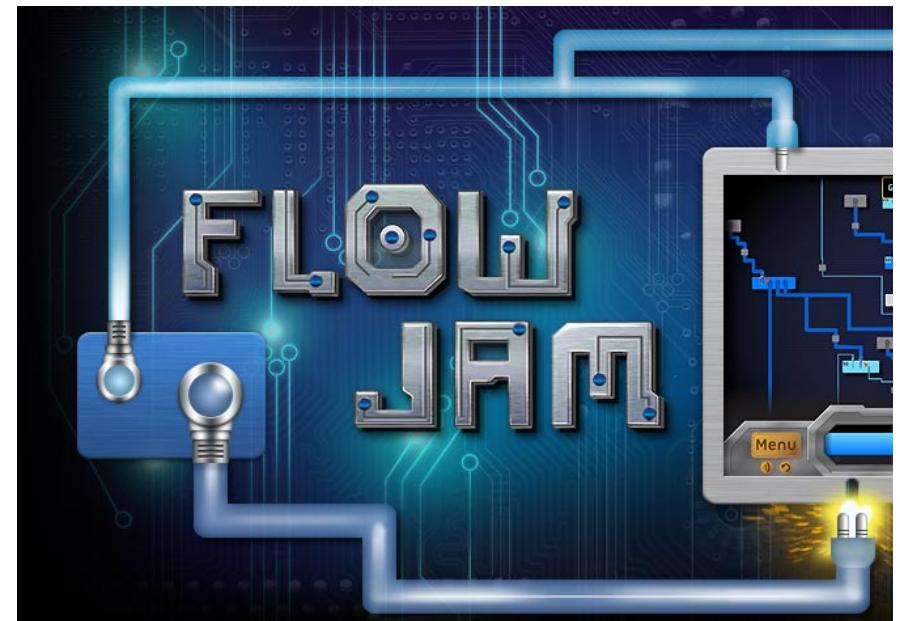
Make the game accessible to everyone

Harness the power of the crowd

Goal: Verify software while waiting

<http://verigames.com/>

FTfJP'12 paper



More at JavaOne'14

Using Type Annotations to Improve Code Quality

BoF, tonight, 19:00 to 19:45

Enhanced Metadata in Java SE 8

Conference talk, Wednesday, 16:30 to 17:30

Conclusions

Java 8 syntax for type annotations

Checker Framework for creating type checkers

- Featureful, effective, easy to use, scalable

Prevent bugs at compile time

Create custom type-checkers

Learn more, or download at:

<http://CheckerFramework.org/>