# Preventing Errors Before They Happen The Checker Framework



http://CheckerFramework.org/

Twitter: @CheckerFrmwrk

Live demo: http://CheckerFramework.org/live/

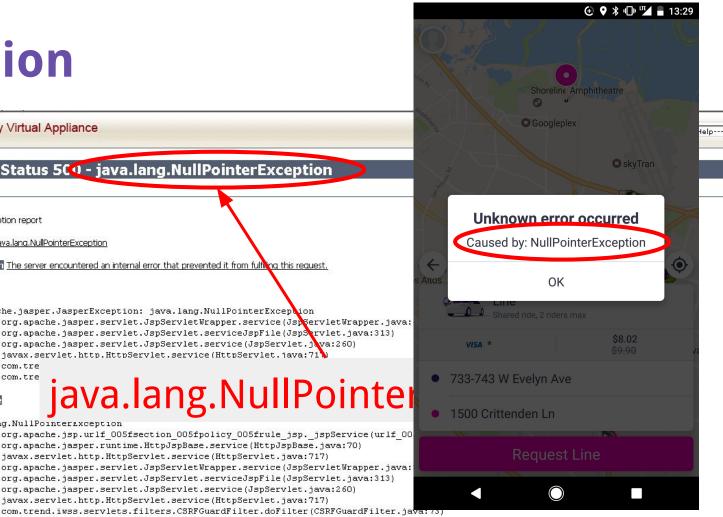
Werner Dietl, University of Waterloo



#### **Motivation**



com.trend.iwss.servlets.filters.AuthFilter.doFilter(AuthFilter.java:377)



#### **Cost of software failures**

**\$312 billion per year** global cost of software bugs (2013) **\$300 billion** dealing with the Y2K problem

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



#### Software bugs can cost lives

1997: **225** deaths: jet crash caused by radar software

1991: 28 deaths: Patriot missile guidance system

2003: 11 deaths: blackout

1985-2000: >8 deaths: Radiation therapy

2011: Software caused 25% of all medical device recalls



#### **Outline**

- Verification approach: Pluggable type-checking
- Tool: Checker Framework
- How to use it
- Creating a custom type system



```
Type checking prevents many errors
int i = "hello";
```

Type checking doesn't prevent enough errors

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

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



```
Type checking prevents many errors
int i = "hello";
```

```
Type checking (NullPointerException System.console().readLine();
```

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



```
Type checking prevents many errors
int i = "hello";
```

Type checking doesn't prevent enough errors

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

Nullness Checker demo Fixed



```
Type checking prevents many errors
int i = "hello";
```

Type checking doesn't prevent enough errors

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



```
Date date = new Date();
myMap.put(date, "now");
date.setSeconds(0); // round to minute
myMap.get(date);
```



```
Date date = new Date();
myMap.put(date, "now");
date.setSeconds(0); // round to minute
myMap.get(date);
   Corrupted map
```



dbStatement.executeQuery(userInput);



dbStatement.executeQuery(userInput);

SQL injection attack

Initialization, data formatting, equality tests, ...



# **SQL** injection attack

Goal: don't execute user input as a SQL command

```
private String wrapQuery(String s) {
  return "SELECT * FROM User WHERE userId='" + s + "'";
}
```

```
If a user inputs his name as: 'or 'x'='x
the SQL query is: ... WHERE userID=' or 'x'='x'
```

To prevent errors: sanitize user data before use



```
void op(String in) {
  executeQuery(in);
op(userInput);
```



```
void op(String in) {
  executeQuery(in);
op(userInput);
```

Where is the defect?



```
void op(String in) {
  executeQuery(in);
op(userInput);
```

Where is the defect?



```
void op(String in) {
Where is the defect?
  executeQuery(in);
        Can't decide without specification!
op(userInput);
```



## **Specification 1: untainted parameter**

```
void op(@Untainted String in) {
  executeQuery(in);
op(userInput);
```



## **Specification 1: untainted parameter**

```
void op(@Untainted String in) {
  executeQuery(in);
op(userInput); // error
```



## **Specification 2: tainted parameter**

```
void op(@Tainted String in) {
  executeQuery(in);
op(userInput);
```



## **Specification 2: tainted parameter**

```
void op(@Tainted String in) {
  executeQuery(in);
                              // error
op(userInput);
```



# **Preventing SQL injection**

Goal: don't execute user input as a SQL command

```
private String wrapQuery(String s) {
  return "SELECT * FROM User WHERE userId='" + s + "'";
    If a user inputs his name as: ' or 'x'='x
     the SQL query is: ... WHERE userID=' or 'x'='x'
```

@Tainted = might be untrusted user input

<sub>23</sub> **Ountainted** = sanitized, safe to use



# Solution: Pluggable Type Checking

- 1. Design a type system to solve a specific problem
- Write type qualifiers in code (or, use type inference) void foo (@Nullable Date date) {

```
date.setSeconds(0); // compile-time error
```

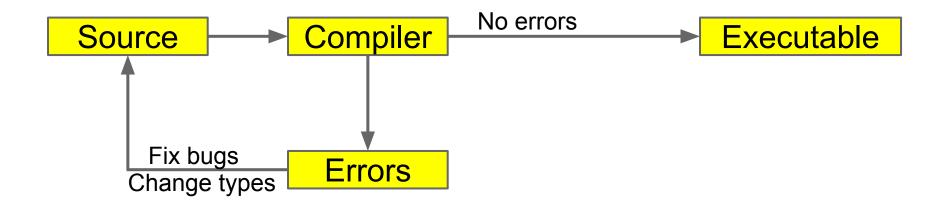
3. Type checker warns about violations (bugs)

```
% javac -processor NullnessChecker MyFile.java
```

MyFile.java:149: dereference of possibly-null reference bb2
 allVars = bb2.vars;

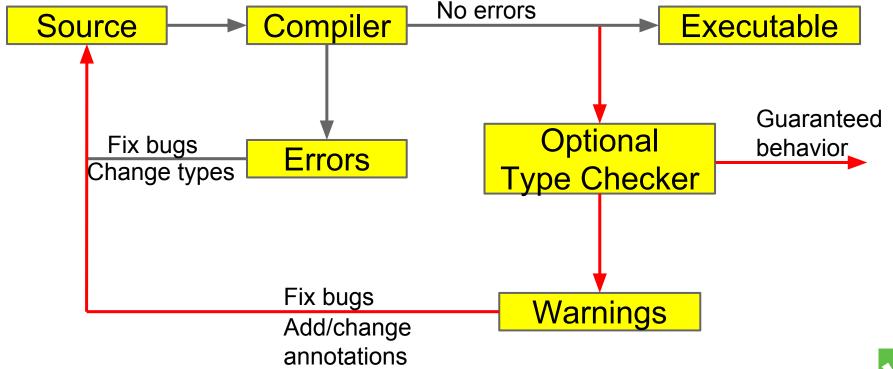


# **Type Checking**



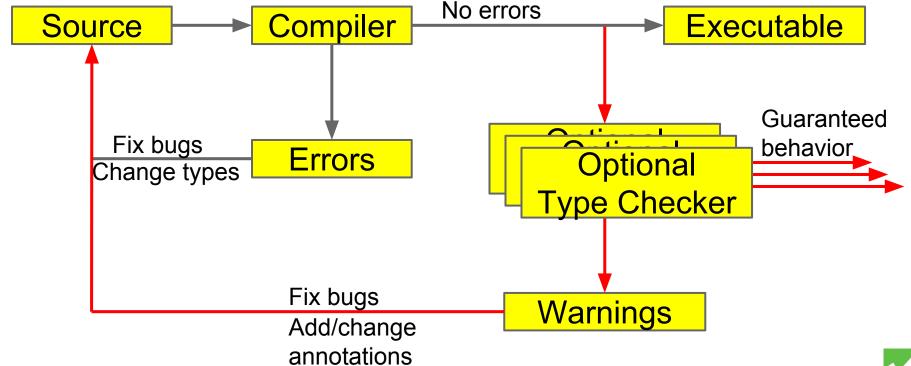


# **Optional Type Checking**





# **Optional Type Checking**





## Prevent null pointer exceptions

Type system that statically guarantees that: the program only dereferences known non-null references

#### Types of data:

@NonNull reference is never null
@Nullable reference may be null



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



#### Where is the defect?

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



#### Where is the defect?

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



```
Where is the defect?
```

```
String op(Data in) {
  return "transform: " + in.getF();
      Can't decide without specification!
String s = op(null);
```



## **Specification 1: non-null parameter**

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



## **Specification 1: non-null parameter**

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



## **Specification 2: nullable parameter**

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



## **Specification 2: nullable parameter**

```
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
  - E.g., could reduce number of run-time checks
- Possible negatives:
  - Must write the types (or use type inference)
  - False positives are possible (can be suppressed)



#### The Checker Framework

A framework for pluggable type checkers "Plugs" into the OpenJDK or OracleJDK compiler

javac -processor MyChecker ...

Standard error format allows tool integration



## Ant, Maven, Gradle integration

```
cpresetdef name="jsr308.javac">
  <javac fork="yes"</pre>
   executable="${checkerframework}/checker/bin/${cfJavac}" >
    <!-- JSR-308-related compiler arguments -->
    <compilerarg value="-version"/>
    <compilerarg value="-implicit:class"/>
 </javac>
                                       <dependencies>
</presetdef>
                                         ... existing <dependency> items ...
                                         <!-- annotations from the Checker Framework:
                                                nullness, interning, locking, ... -->
                                           <dependency>
                                             <groupId>org.checkerframework
                                             <artifactId>checker-qual</artifactId>
                                             <version>1.9.7
```

</dependency>

</dependencies>

## Eclipse, IntelliJ, NetBeans integration

```
public class Test {
     public static void main(String[] args) {
     Console c = System.console();
     c.printf("Test");
                                                        public class Test {
                                                            public static void main(String[] args) {
                                                        Console c = System.console();
              @ Javadoc 📵 Declaration 🧳 Search 📮
Problems ⊠
                                                         dereference of possibly-null reference c c.printf("Test");
0 errors, 1 warning, 0 others
Description

 Marnings (1 item)

                                                   🥊 Problems 🛭 🍘 Javadoc 📵 Declaration 🔗 Search 📮 Console 🗷 Task
       dereference of possibly-null reference c
          c.printf("Test");
                                                   0 errors, 1 warning, 0 others
                                                    Description
                                                                                                      Resource
                                                    dereference of possibly-null reference c
                                                                                                       Test.java
                                                              c.printf("Test");
```



#### Comparison: other nullness tools

	Null pointer errors		False	Annotations
	Found	Missed	warnings	written
Checker Framework	9	0	4	35
FindBugs	0	9	1	0
Jlint	0	9	8	0
PMD	0	9	0	0
Eclipse, in 2017	0	9	8	0
Intellij (@NotNull	0	9	1	0
default), in 2017	3	6	1	925 + 8

41 Checking the Lookup program for file system searching (4kLOC)



#### Live demo: http://CheckerFramework.org/live/

#### Checker Framework Live Demo

Write Java code here:

```
import org.checkerframework.checker.nullness.qual.Nullable;
class YourClassNameHere {
    void foo(Object nn, @Nullable Object nbl) {
        nn.toString(); // OK
        nbl.toString(); // Error
    }
}
```

Choose a type system: Nullness Checker ▼

Check

#### **Examples:**

Nullness: NullnessExample | NullnessExampleWithWarnings

MapKey: <u>MapKeyExampleWithWarnings</u>

Interning: InterningExample | InterningExampleWithWarnings



#### **Example type systems**

```
Null dereferences (@NonNull)
   >200 errors in Google Collections, javac, ...
Equality tests (@Interned)
   >200 problems in Xerces, Lucene, ...
Concurrency / locking (@GuardedBy)
   >500 errors in BitcoinJ, Derby, Guava, Tomcat, ...
Fake enumerations / typedefs (@Fenum)
   problems in Swing, JabRef
```



#### String type systems

```
Regular expression syntax (@Regex)
   56 errors in Apache, etc.; 200 annos required
printf format strings (@Format)
   104 errors, only 107 annotations required
Method signature format (@FullyQualified)
   28 errors in OpenJDK, ASM, AFU
Compiler messages (@CompilerMessageKey)
   8 wrong keys in Checker Framework
```



#### Security type systems

Command injection vulnerabilities (@OsTrusted)

5 missing validations in Hadoop

Information flow privacy (@Source)

SPARTA detected malware in Android apps



It's easy to write your own type system!



#### **Checkers are usable**

- Type-checking is familiar to programmers
- Modular: fast, incremental, partial programs
- Annotations are not too verbose
  - @NonNull: 1 per 75 lines
  - @Interned: 124 annotations in 220 KLOC revealed 11 bugs
  - **@Format**: 107 annotations in 2.8 MLOC revealed 104 bugs
  - Possible to annotate part of program
  - Fewer annotations in new code
- Few false positives
- First-year CS majors preferred using checkers to not
- Practical: in use in Silicon Valley, on Wall Street, etc.



## What a checker guarantees

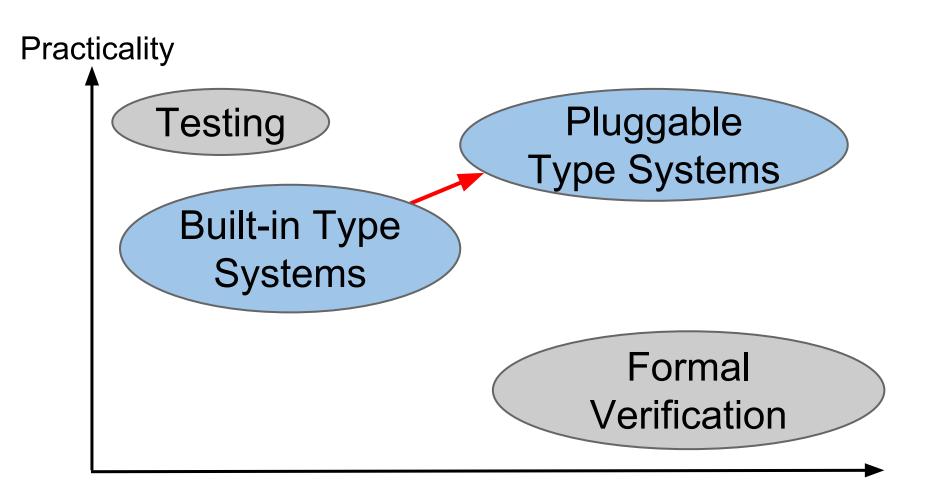
The program satisfies the type property. There are:

- no bugs (of particular varieties)
- no wrong annotations
- Caveat 1: only for code that is checked
  - Native methods (handles reflection!)
  - Code compiled without the pluggable type checker
  - Suppressed warnings
    - Indicates what code a human should analyze
  - Checking part of a program is still useful
- Caveat 2: The checker itself might contain an error



```
= Addr \rightarrow Obj
                                                                              Heap
       Formalizations
                                                                   \iota \ \in \ \mathtt{Addr}
                                                                                                  = Set of Addresses \cup {null<sub>a</sub>}
                                                                       ∈ Obj
                                                                                                  = Type, Fields
                                                                        \in {	extbf{r}}Type
                                                                                                  = OwnerAddr ClassId<\(\bar{r}\)Type>
      ∈ Program ::= Class, ClassId, Expr Fs
                                                                        \in
                                                                              Fields
                                                                                                  = FieldId \rightarrow Addr
             Class
                                   class ClassId < TVarId
Cls ∈
                                                                       \in
                                                                              OwnerAddr
                                                                                                  = Addr \cup \{any_a\}
                                   extends ClassId< Typ
                                                                                                  = TVarId rType; ParId Addr
                                                                              <sup>r</sup>Env
                                   { FieldId SType; Met
             <sup>s</sup>Type
                                   SNType | TVarId
                                                                                                              h, {}^{r}\Gamma, e_0 \rightsquigarrow h', \iota_0
             sNType
                                   OM ClassId < Type>
                                                                                                                   \iota_0 \neq \mathtt{null}_a
             OM
                                                      h, {}^{r}\Gamma, e_0 \rightsquigarrow h_0, \iota_0
                                                                                           OS-Read \frac{\iota = h'(\iota_0) \downarrow_2 (f)}{h, {}^r\Gamma, e_0.f \leadsto h', \iota}
             Meth
                                                          \iota_0 \neq \mathtt{null}_a
             MethSig
                                                     h_0, {}^{\mathbf{r}}\Gamma, e_2 \rightsquigarrow h_2, \iota
             Purity
             Expr
                                                                                                 \Gamma \vdash e_0 : N_0 \qquad N_0 = u_0 C_0 < >
                                   Expr.MethId<sType>(Expr)
                                                                                                        \mathtt{T}_1 = \mathit{fType}(\mathtt{C}_0, \mathtt{f})
                                   new SType (SType) Expr
                                                                                                        \Gamma \vdash e_2 : N_0 \triangleright T_1
                           εГ
             Env
                           ::= TVarId sNType; ParId sType
 h \vdash {}^{\mathbf{r}}\Gamma : {}^{\mathbf{s}}\Gamma
 h \vdash \iota_1 : dyn({}^{\mathfrak s}N, h, {}^{\mathfrak l}_{1,1})
 h \vdash \iota_2 : dyn(^{\mathfrak{s}}\mathsf{T}, \iota_1, h(\iota_1) \downarrow_1)
                                                  \implies h \vdash \iota_2 : dun({}^{\mathtt{s}}\mathtt{N} \triangleright^{\mathtt{s}}\mathtt{T}.h.{}^{\mathtt{r}}\Gamma)
 ^{s}N = u_{N} C_{N} < >
                                             u_N = this_u \Rightarrow {}^{\mathbf{r}}\Gamma(this)
                                                          dom(C) = \overline{X}
                                                                                                        free(^{s}T) \subseteq \overline{X} \circ \overline{X'}
 free(^{s}T) \subseteq dom(C_{N})
                                   DYN-
                                            dyn(^{s}T, \iota, ^{r}T, (\overline{X'} ^{r}T'; \_)) = ^{s}T[\iota'/this, \iota'/peer, \iota/rep, any, /any, /^{r}T/X, /rT'/X']
```





#### Demo: regular expression errors

```
@Regex = valid regular expression
  OK: "colou?r"
  NOT: "1) first point"
@Regex(2) = has 2+ capturing groups
  OK: "<u>((</u>Linked<u>)</u>?Hash<u>)</u>?Map"
  OK: "(http|ftp)://([^/]+)(/.*)?"
  NOT: "(brown|beige)"
```



#### 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 String conten PatternSyntaxException
 Pattern pat = Pattern.compile(regex);
 Matcher mat IndexOutOfBoundsExceptionon
   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);
```



#### Since Java 5: declaration annotations

Only for declaration locations:

```
@Deprecated
class Foo {
  @Getter @Setter private String query;
  @SuppressWarnings("unchecked")
  void foo() { ... }
```



#### But we couldn't express

A <u>non-null</u> reference to my data

An <u>interned</u> string

A <u>non-null</u> List of <u>English</u> strings

A <u>non-empty</u> array of <u>English</u> strings



#### With Java 8 Type Annotations we can!

```
A non-null reference to my data
   @NonNull Data mydata;
An interned String
   @Interned String query;
A non-null List of English Strings
   @NonNull List<@English String> msgs;
A non-empty array of English strings
   @English String @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, ...



# Java 6 & 7 compatibility (or avoid dependency on Checker Framework)

Annotations in comments:

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

(Requires use of jsr308-langtools compiler.)



#### **Annotating external libraries**

When type-checking clients, need library spec. Can write manually or automatically infer Two syntaxes:

- As separate text file (stub file)
- Within its .jar file (from annotated partial source code)



#### **Checker Framework facilities**

- Full type systems: inheritance, overriding, ...
- Generics (type polymorphism)
  - Also qualifier polymorphism
- Qualifier defaults
- Pre-/post-conditions
- Warning suppression



#### **Brainstorming new type checkers**

What runtime exceptions 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, ...



What runtime exceptions to prevent?

What properties of data should always hold?



What runtime exceptions to prevent?

NullPointerException

What properties of data should always hold?



What runtime exceptions to prevent?
NullPointerException

What properties of data should always hold?

@NonNull references always non-null



What runtime exceptions to prevent? NullPointerException What properties of data should always hold? @NonNull references always non-null What operations are legal and illegal? Dereferences only on @NonNull references



What runtime exceptions to prevent?

What properties of data should always hold?



What runtime exceptions to prevent?

PatternSyntaxException, IndexOutOfBoundsException

What properties of data should always hold?



What runtime exceptions to prevent?

PatternSyntaxException, IndexOutOfBoundsException

What properties of data should always hold?

Whether a string is a regex and number of groups



What runtime exceptions to prevent?

PatternSyntaxException, IndexOutOfBoundsException

What properties of data should always hold?

Whether a string is a regex and number of groups

What operations are legal and illegal?

Pattern.compile with non-@Regexp, etc,



#### New type system

```
What runtime exceptions to prevent?

1
```

What properties of data should always hold?



#### New type system

```
What runtime exceptions to prevent?

1
```

What properties of data should always hold?



#### New type system

What runtime exceptions to prevent?

1

What properties of data should always hold?



## New type system

What runtime exceptions to prevent?

1

What properties of data should always hold?

What operations are legal and illegal?



### New type system

What runtime exceptions to prevent?

1

What properties of data should always hold?

What operations are legal and illegal?



## Building a checker is easy

#### Example: Ensure encrypted communication

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



## Building a checker 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 {}
```



# **Encrypted Checker Demo**

Let's build it!



## **Testing infrastructure**

jtreg-based testing as in OpenJDK

Lightweight tests with in-line expected errors:

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

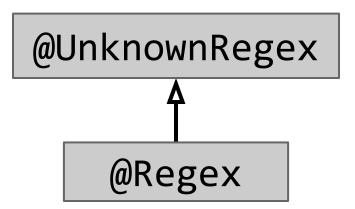


- 1. Qualifier hierarchy
  - defines subtyping
- 2. Type introduction rules
  - types for expressions
- 3. Type rules
  - checker-specific errors
- 4. Flow-refinement
  - better types than the programmer wrote



- 1. Qualifier hierarchy
  - subtyping, assignments

```
@SubtypeOf(UnknownRegex.class)
public @interface Regex {
```





- 2. Type introduction rulestypes for expressions
- @ImplicitFor( trees = {
   Tree.Kind.NEW\_CLASS,
   Tree.Kind.NEW\_ARRAY, ... })
  @DefaultQualifierInHierarchy
  @DefaultForUnannotatedCode({
   DL.PARAMETERS, DL.LOWER\_BOUNDS })



Data d = new Data();

- 3. Type rules
  - checker-specific errors

```
void visitSynchronized(SynchronizedTree node) {
    ExpressionTree expr = node.getExpression();
    AnnotatedTypeMirror type =
        getAnnotatedType(expr);
    if (!type.hasAnnotation(NONNULL))
        checker.report(Result.failure(...), expr);
```



synchronized(xxx) {

- 4. Flow-refinement
  - better types than the programmer wrote

if (x != null) {
 x.f = ...; // valid



#### **Dataflow Framework**

#### Goal: Compute properties about expressions

- More accurate types than the user wrote
- Foundation for other static analyses
  - e.g. by Google error-prone and Uber NullAway

#### Dataflow Framework user provides

- What are we tracking?
- What do operations do?
- What are intermediate results?

Dataflow Framework does all the work!



## **Tips**

- Start by type-checking part of your code
- Only type-check properties that matter to you
- Use subclasses (not type qualifiers) if possible
- 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>



### Verification

- Goal: prove that no bug exists
- **Specifications**: user provides
- False negatives: none
- False positives: user suppresses warnings
- Downside: user burden

# **Bug-finding**

- Goal: find some bugs at low cost
- Specifications: infer likely specs
- False negatives: acceptable
- False positives: heuristics focus on most important bugs
- **Downside**: missed bugs



## **Checker Framework Community**

Open source project:

https://github.com/typetools/checker-framework

- Monthly release cycle
- >13,000 commits, 75 authors
- Welcoming & responsive community
- Google Summer-of-Code participant



## **Checker Framework plans**

#### More type systems:

- Immutability
- Index-out-of-bounds
- Optional<T> type
- Determinism
- Signed vs. unsigned numbers

Type inference

Combined static & dynamic enforcement



# Pluggable type-checking improves code

Checker Framework for creating type checkers

Featureful, effective, easy to use, scalable

Prevent bugs at compile time

Create custom type-checkers

Improve your code!

http://CheckerFramework.org/

