

GENERATING TESTS FOR MOBILE APPS FROM NATURAL LANGUAGE BUG REPORTS AND APP REVIEWS

SNEAK PREVIEW

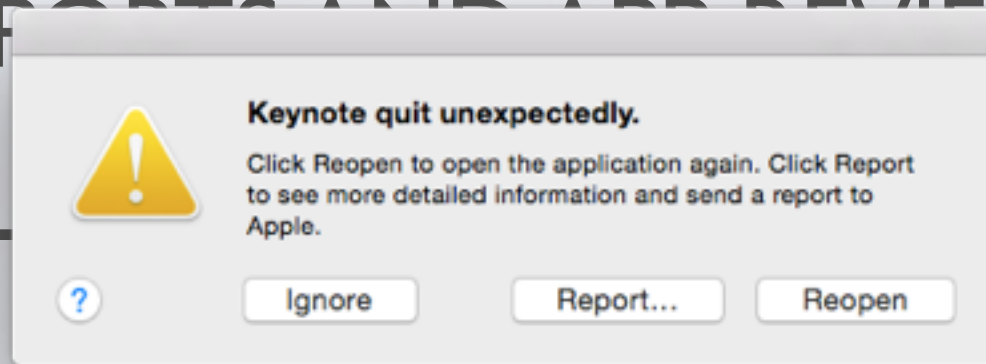
Alessandro (Alex) Orso

School of Computer Science – College of Computing
Georgia Institute of Technology

Joint work with Marcelo d'Amorim and Paola Spoletini

Partially supported by: NSF, IBM, MSR, and Google

GENERATING TESTS FOR MOBILE APPS FROM NATURAL LANGUAGE BUG REPORTS AND APP REVIEWS



Alessandro (Alex) Orso

School of Computer Science – College of Computing
Georgia Institute of Technology

Joint work with Marcelo d'Amorim and Paola Spoletini

Partially supported by: NSF, IBM, MSR, and Google

GENERATING TESTS FOR MOBILE APPS
FROM NATURAL LANGUAGE BUG
REPORTS

**Field failures are
unavoidable!**

Antonio Orso
School of Computer Science – College of Computing
Georgia Institute of Technology

Joint work with Marcelo d'Amorim and Paola Spoletini

Partially supported by: NSF, IBM, MSR, and Google

GENERATED FROM RE

Problem Report for Firefox

Problem and system information:

Date/Time: 2007-05-16 16:00:01.424 -0400
OS Version: 10.4.9 (Build 8P135)
Report Version: 4

Command: firefox-bin
Path: /Applications/Firefox.app/Contents/MacOS/firefox-bin
Parent: WindowServer [59]

Version: 1.5 (1.5)

PID: 947
Thread: 0

Exception: EXC_BAD_ACCESS (0x0001)
Codes: KERN_PROTECTION_FAILURE (0x0002) at 0x0186af20

Thread 0 Crashed:
0 LibSystem.B.dylib 0x00018d5c vfprintf + 40

Please describe what you were doing when the problem happened:

Your report will help Apple improve this software. Your personal information is not sent with this report. You will not be contacted in response to this report. For Apple product support, visit www.apple.com/support or contact your Apple dealer.

? Send to Apple...

TEAMS BUGS VS

F

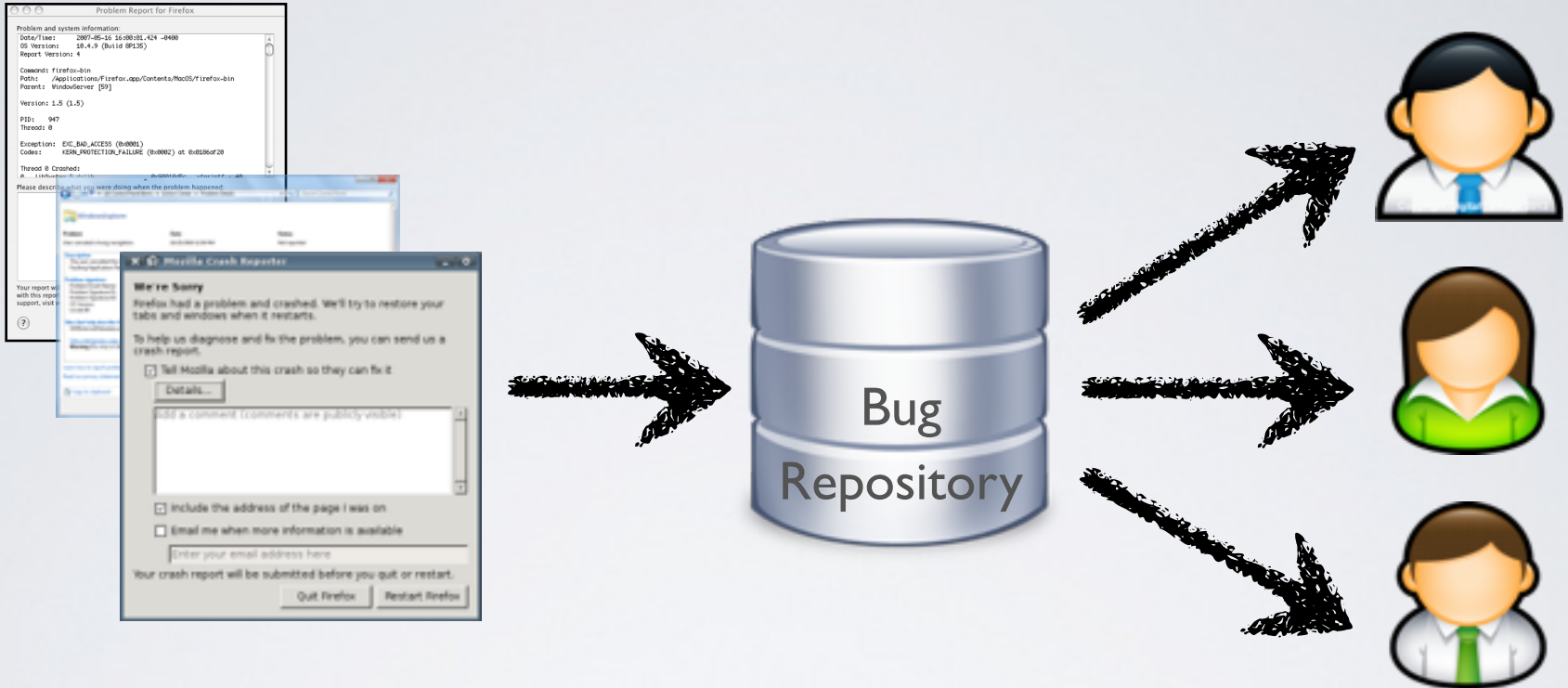
School

John W

puting

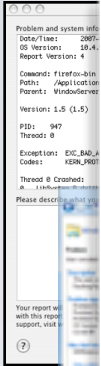
oletini

TYPICAL DEBUGGING PROCESS



Very hard to
(1) reproduce
(2) debug

TYPICAL DEBUGGING PROCESS



Recent survey of Apache, Eclipse, and Mozilla developers:

Information on *how to reproduce field failures* is the most valuable, and difficult to obtain, piece of information for investigating such failures.

[Zimmermann10]

OVERARCHING GOAL: help developers

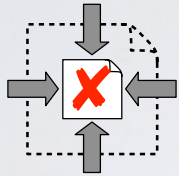
- (1) *investigate* field failures,
- (2) *understand* their causes, and
- (3) *eliminate* such causes.

OUR WORK SO FAR



Recording and replaying executions

[icsm 2007, icse 2007]



Input minimization

[woda 2006, icse 2007]



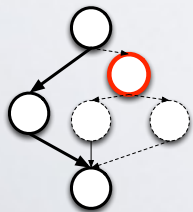
Input anonymization

[icse 2011]



Mimicking field failures

[icse 2012, ase 2013, icst 2014]

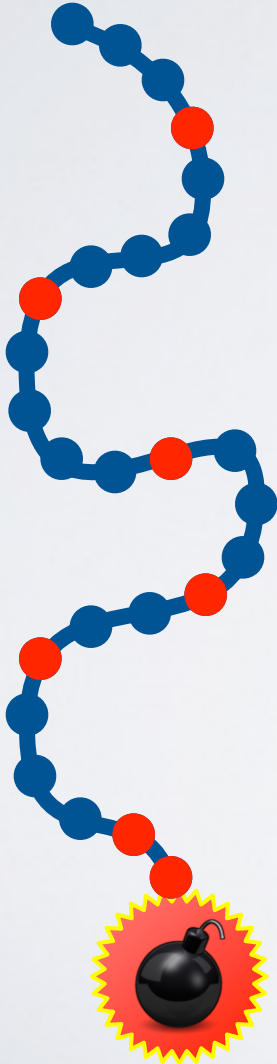


Locating and explaining field failures

[issta 2012, issta 2013, ASE 2014, TOSEM 2015, TR]

MIMICKING FIELD FAILURES

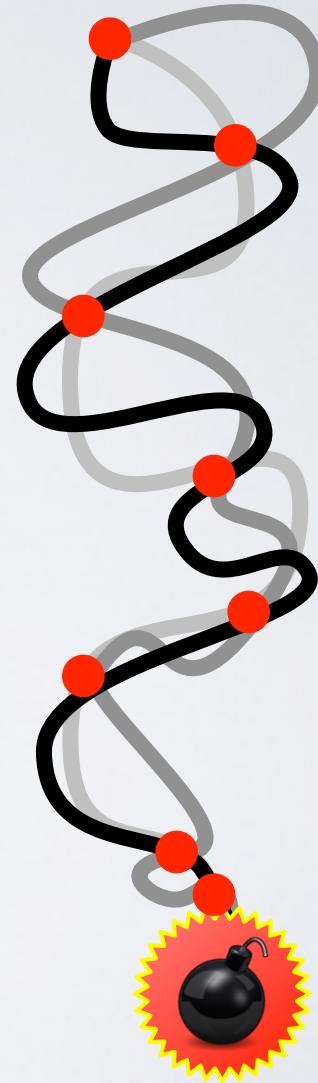
User run (**R**)



Relevant events
(breadcrumbs)



Mimicked run (**R'**)

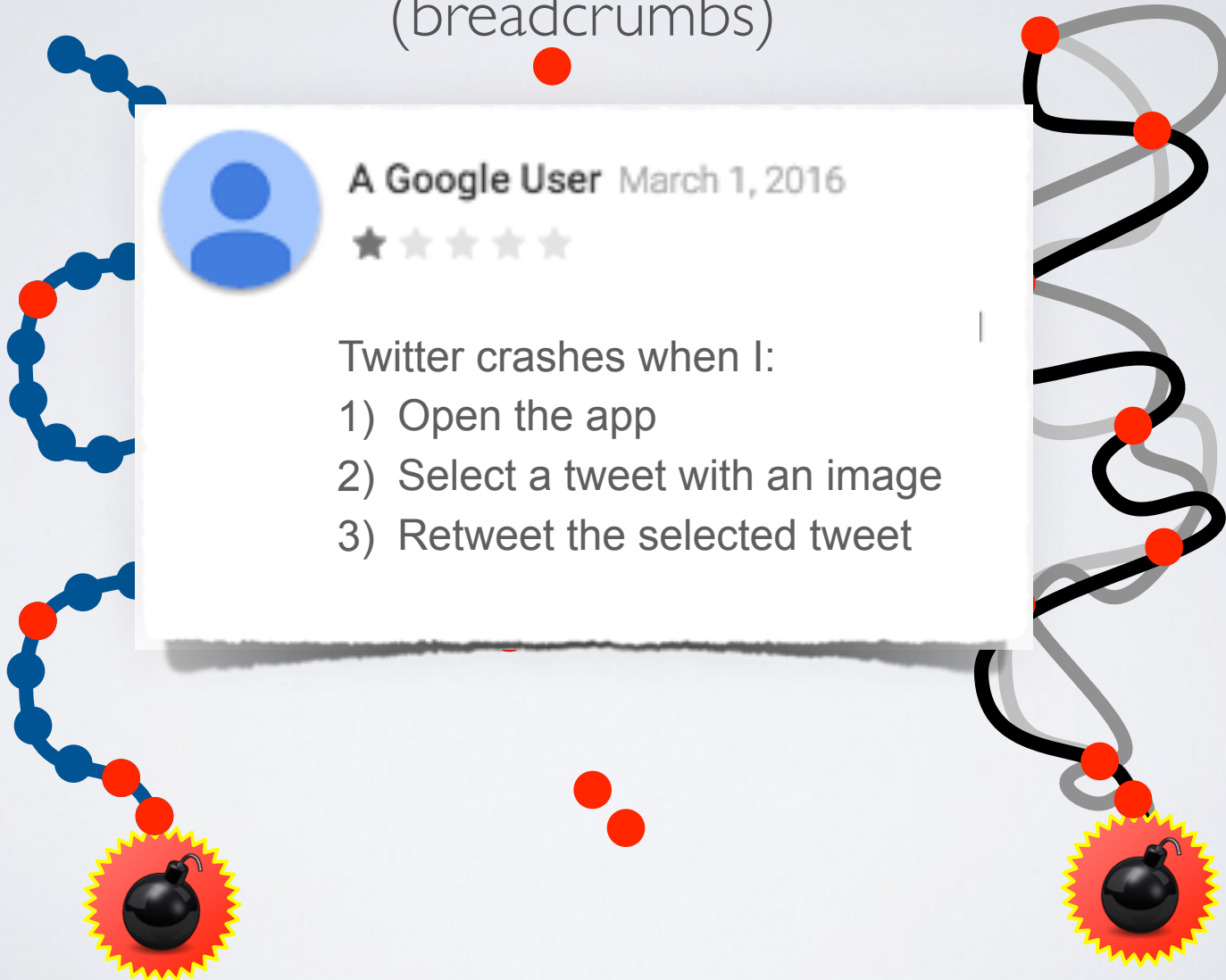


MIMICKING FIELD FAILURES

User run (**R**)

Relevant events
(breadcrumbs)

Mimicked run (**R'**)

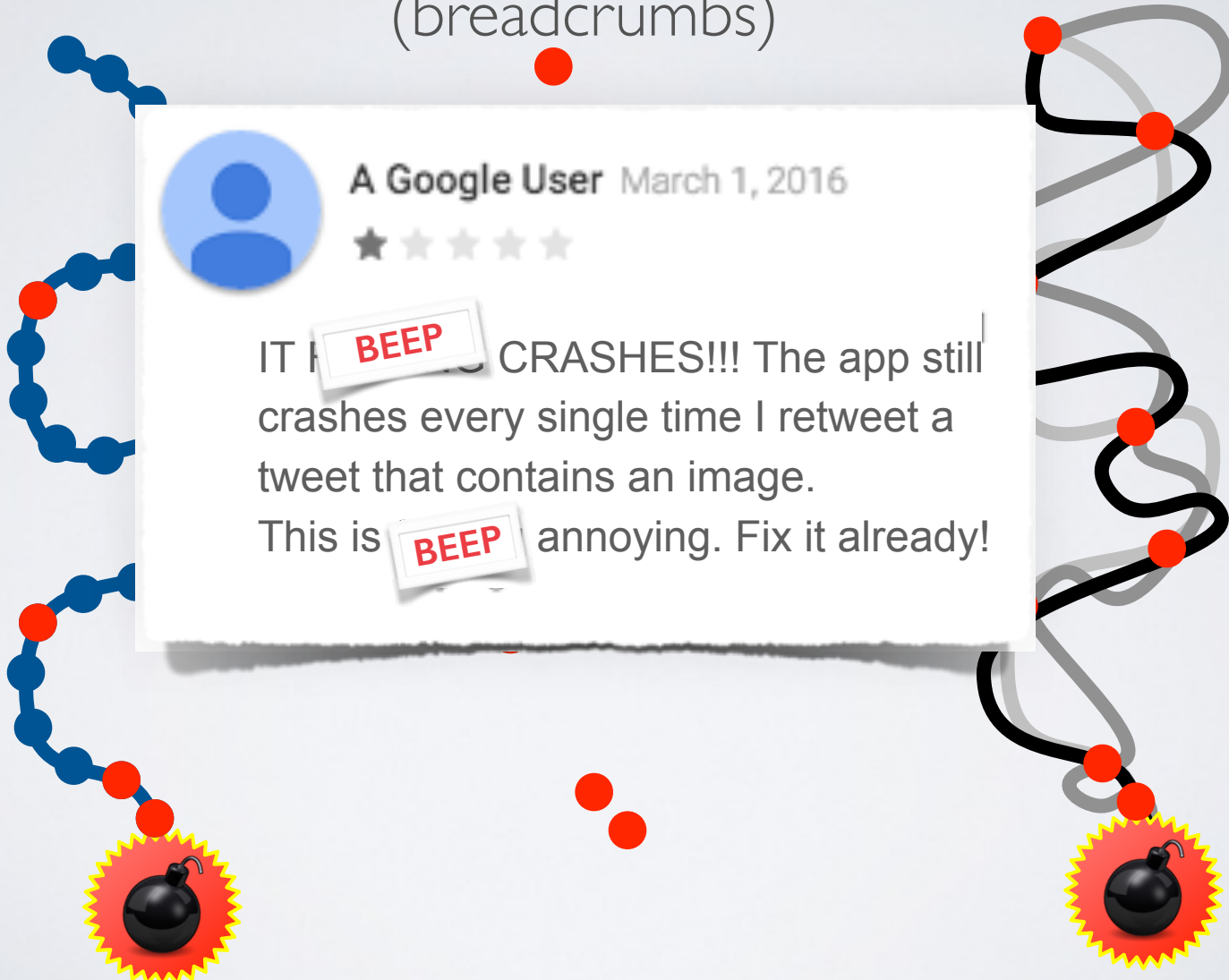


MIMICKING FIELD FAILURES

User run (**R**)

Relevant events
(breadcrumbs)

Mimicked run (**R'**)



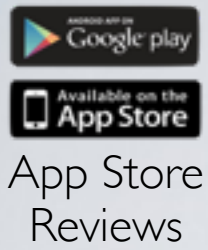
A Google User March 1, 2016



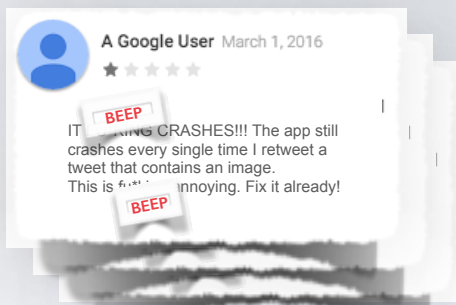
IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.

This is **BEEP** annoying. Fix it already!

OVERALL VISION



Reviews Classifier



Bug Report Miner



App Code



App GUI



Domain Knowledge

App Analyzer
Code Analyzer
GUI Analyzer



Test Steps

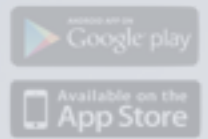


Actions Dictionary

App Test Generator

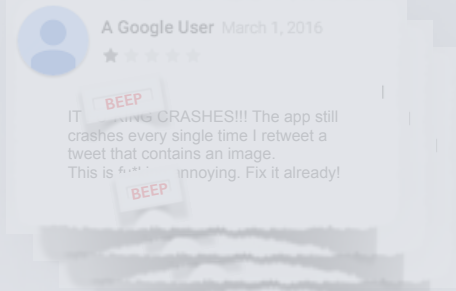


Test Cases



App Store
Reviews

Reviews
Classifier



Bug Reports



Report
iner



Test Steps



App Test
Generator



Test Cases



App Code



App GUI



Domain
Knowledge



Actions
Dictionary

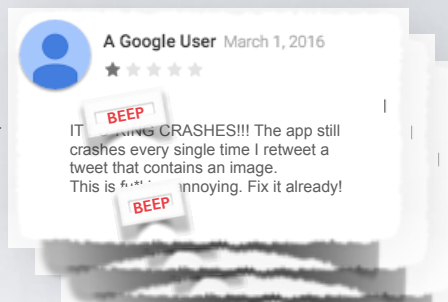




App Store
Reviews



**Reviews
Classifier**

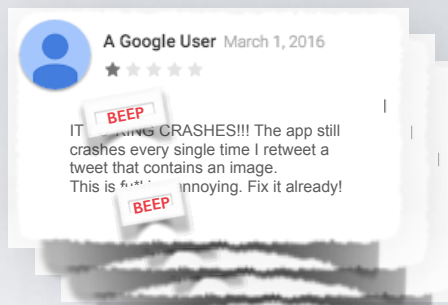


Bug Reports

Bug Report, Feature Request, or Simply Praise? On Automatically Classifying App Reviews

Walid Maalej
University of Hamburg
Hamburg, Germany
maalej@informatik.uni-hamburg.de

Hadeer Nabil
University of Hamburg
Hamburg, Germany
nabil.hadeer@gmail.com



Bug Reports



Bug Report Miner



Test Steps



A Google User March 1, 2016



IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.

This is **BEEP** annoying. Fix it already!



A Google User March 1, 2016



IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image. This is **BEEP** annoying. Fix it already!

Bug Reports



Bug Report Miner



Test Steps



A Google User March 1, 2016



IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.

This is **BEEP** annoying. Fix it already!



A Google User March 1, 2016

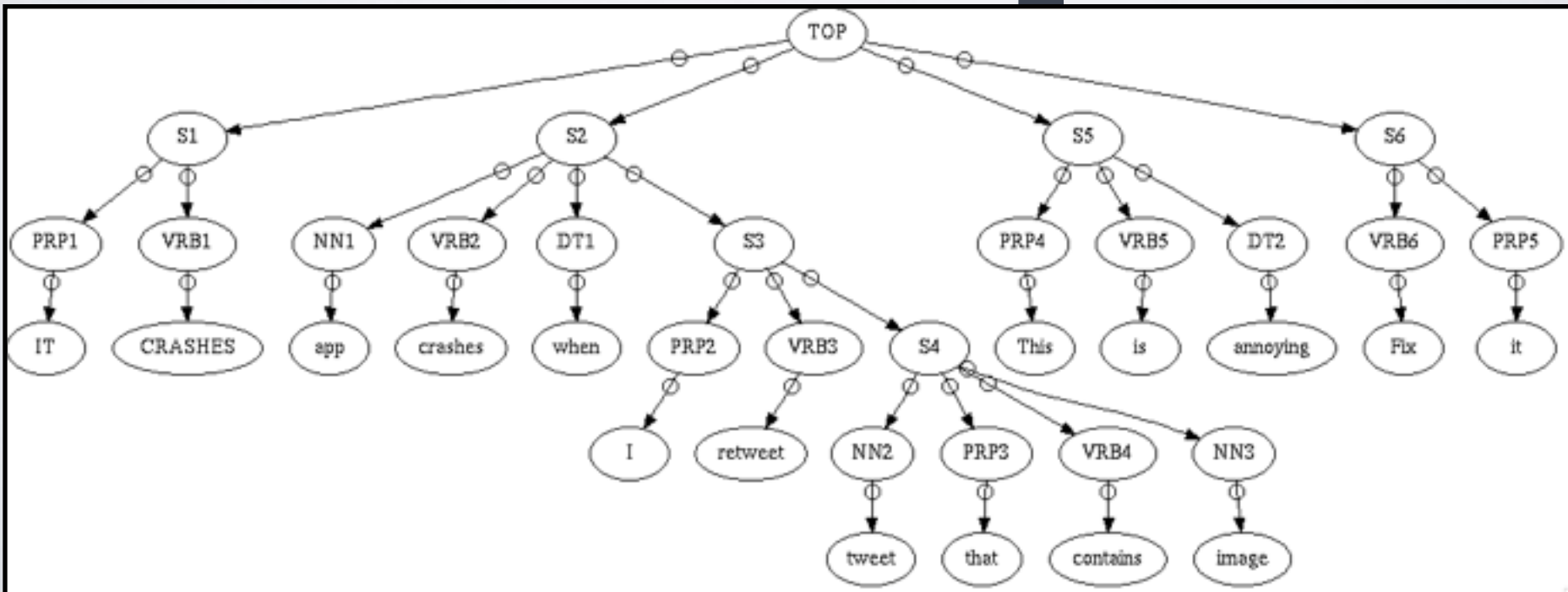


IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.
This is **BEEP** annoying. Fix it already!

Bug Reports



Bug Report Miner





A Google User March 1, 2016



IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.

This is **BEEP** annoying. Fix it already!



A Google User March 1, 2016

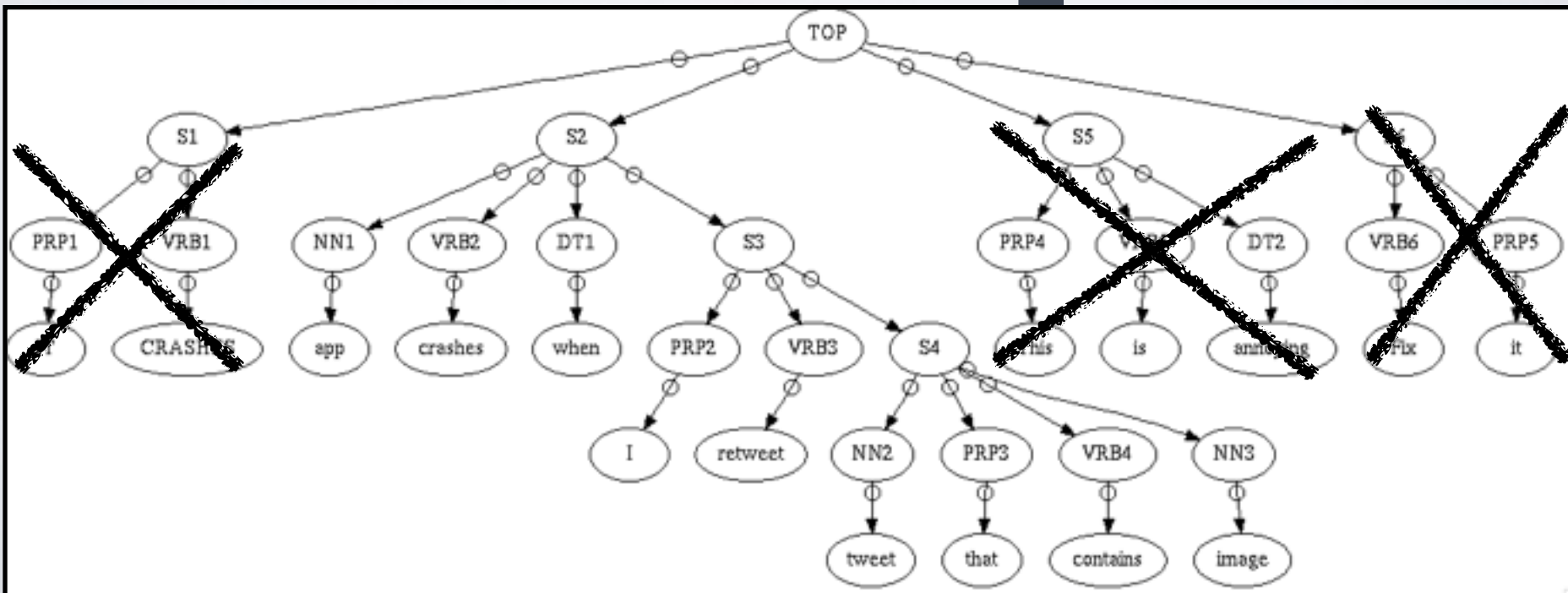


IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.
This is **BEEP** annoying. Fix it already!

Bug Reports



Bug Report Miner





A Google User March 1, 2016



IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.

This is **BEEP** annoying. Fix it already!



A Google User March 1, 2016

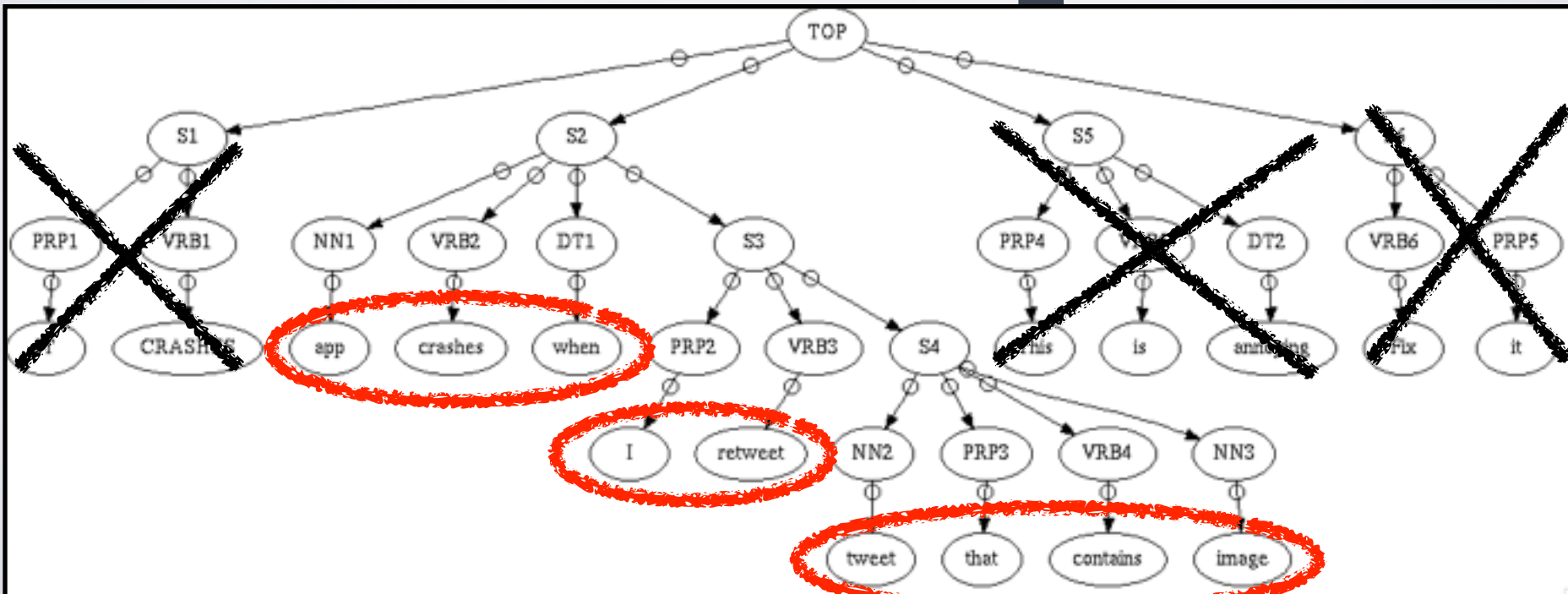


IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.
This is **BEEP** annoying. Fix it already!

Bug Reports



Bug Report Miner





A Google User March 1, 2016



IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image.

This is **BEEP** annoying. Fix it already!

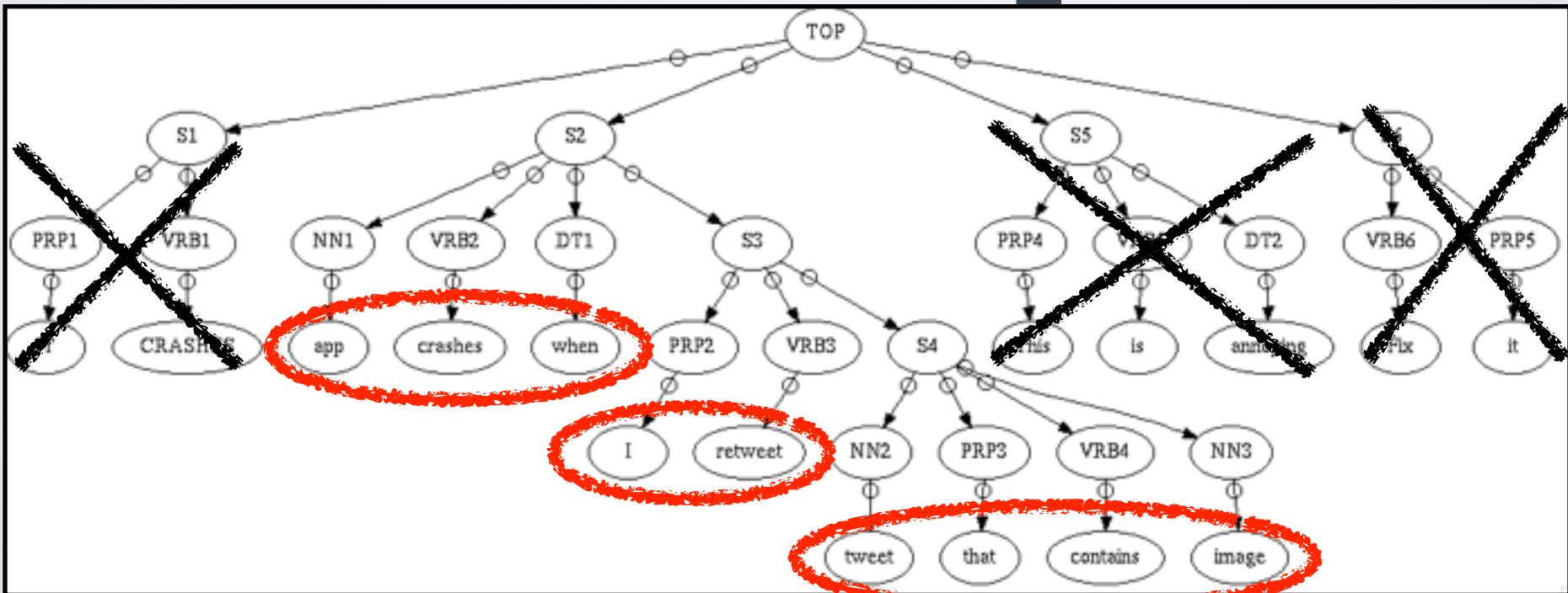


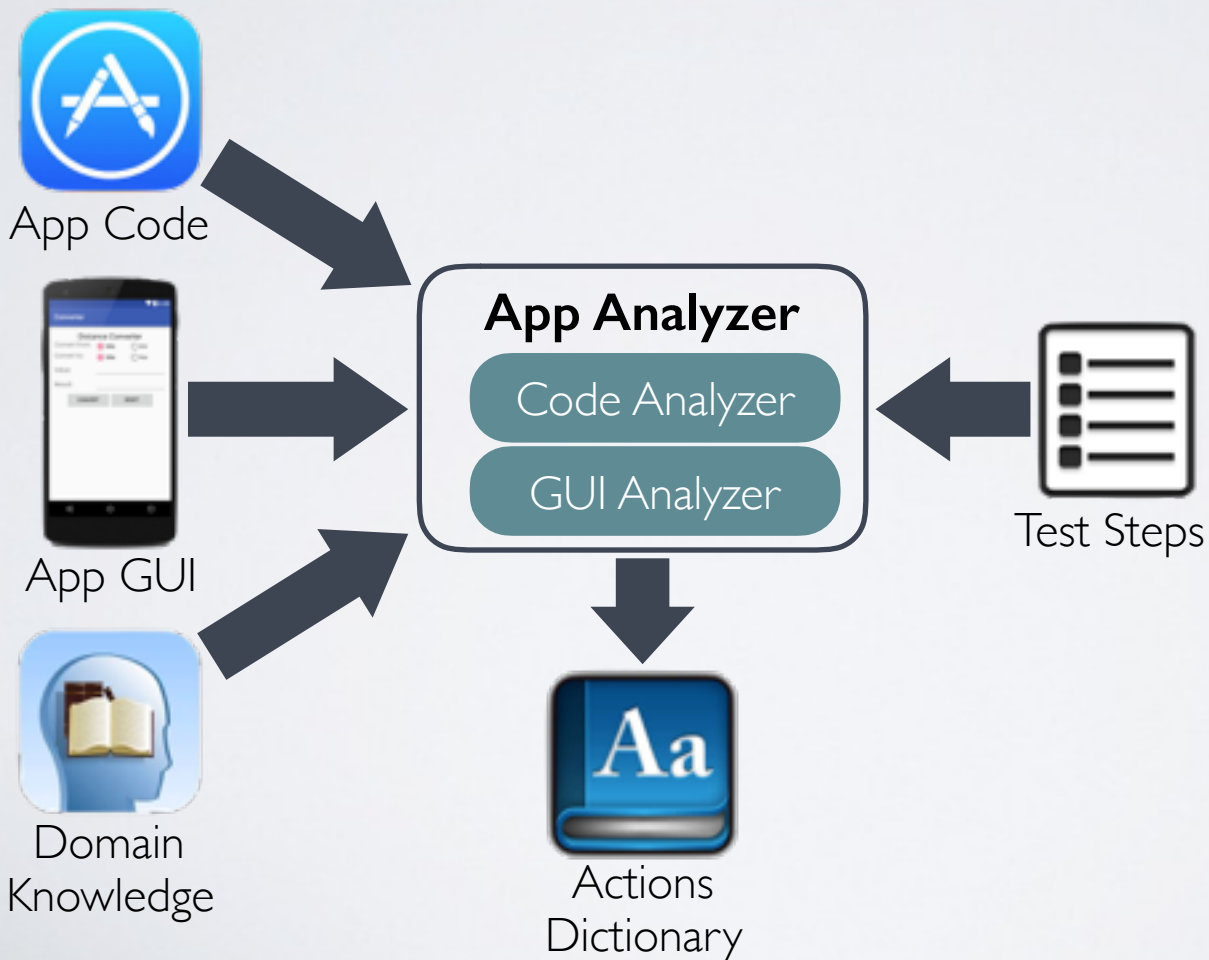
A Google User March 1, 2016

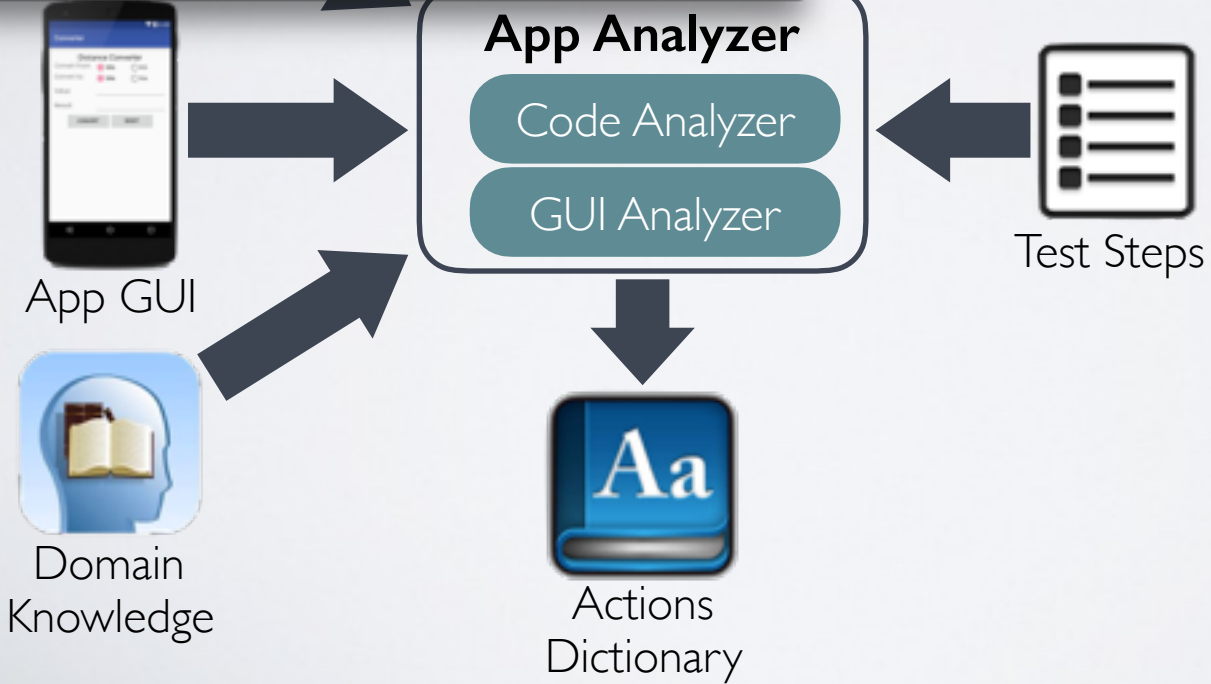
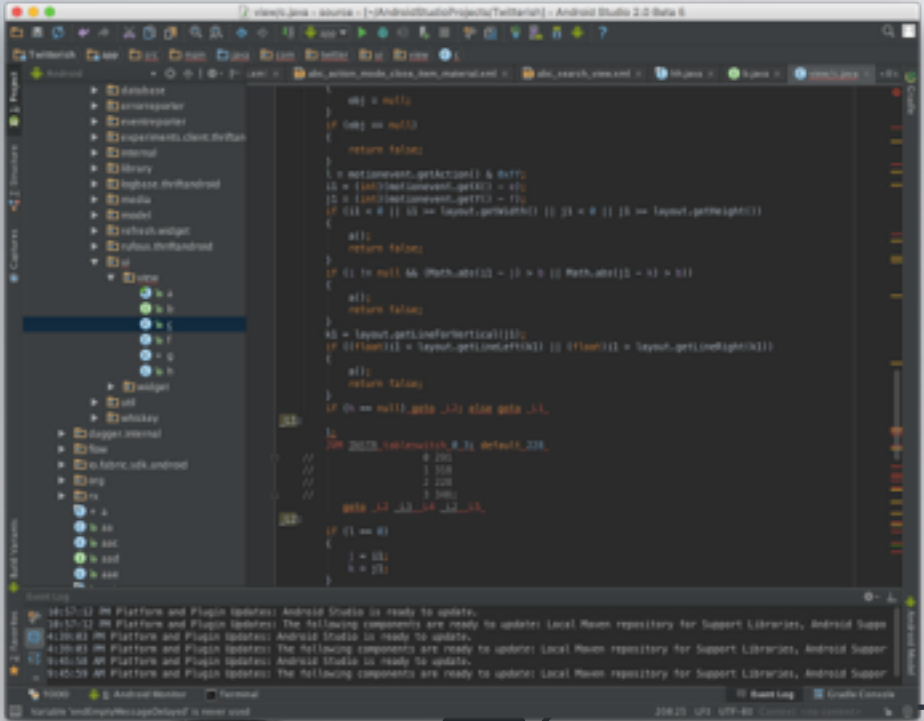


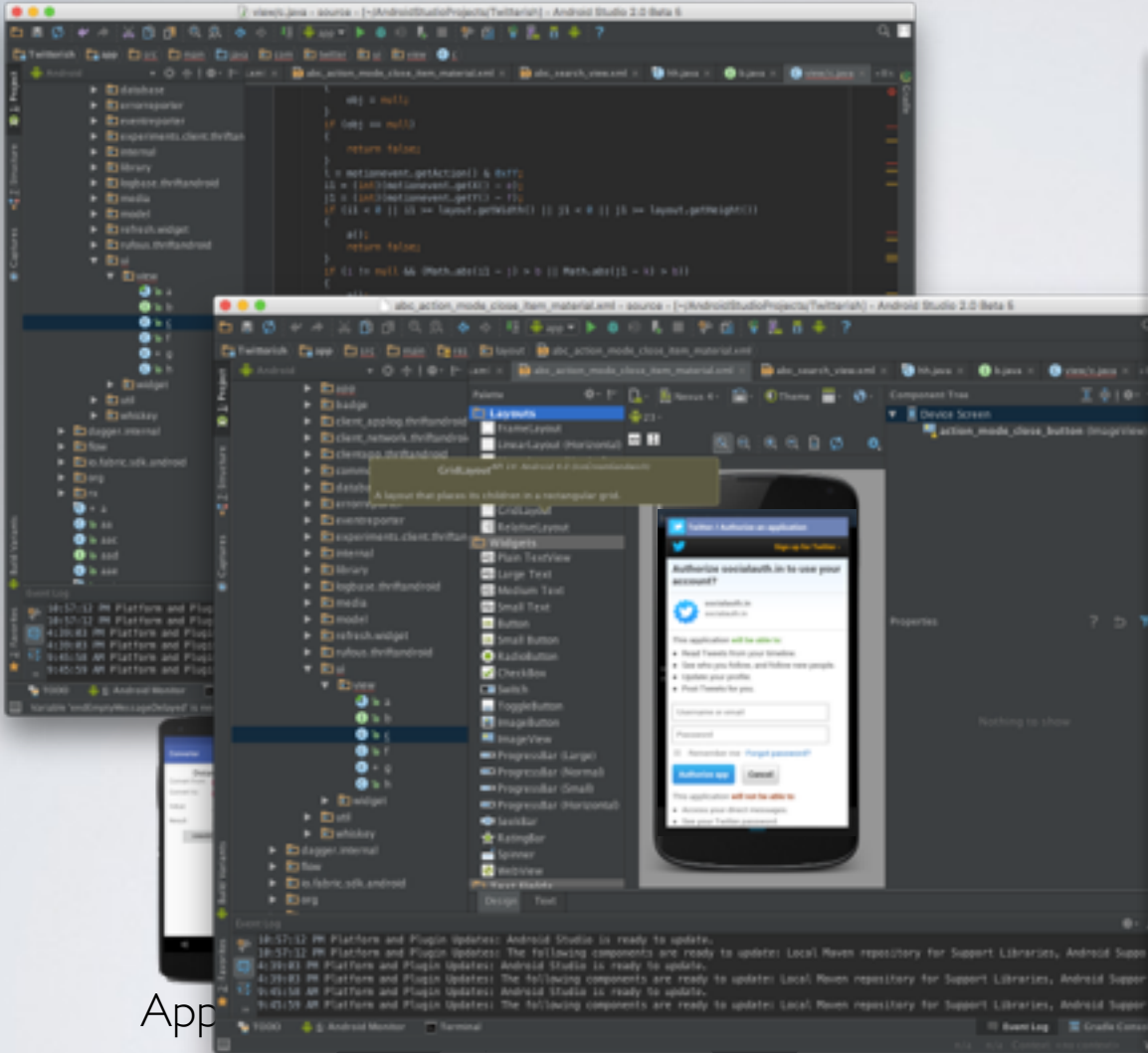
IT **BEEP** CRASHES!!! The app still crashes every single time I retweet a tweet that contains an image. This is **BEEP** annoying. Fix it already!

- 1) Open twitter
 - 2) Select tweet with image
 - 3) Retweet
- ➔ Crash









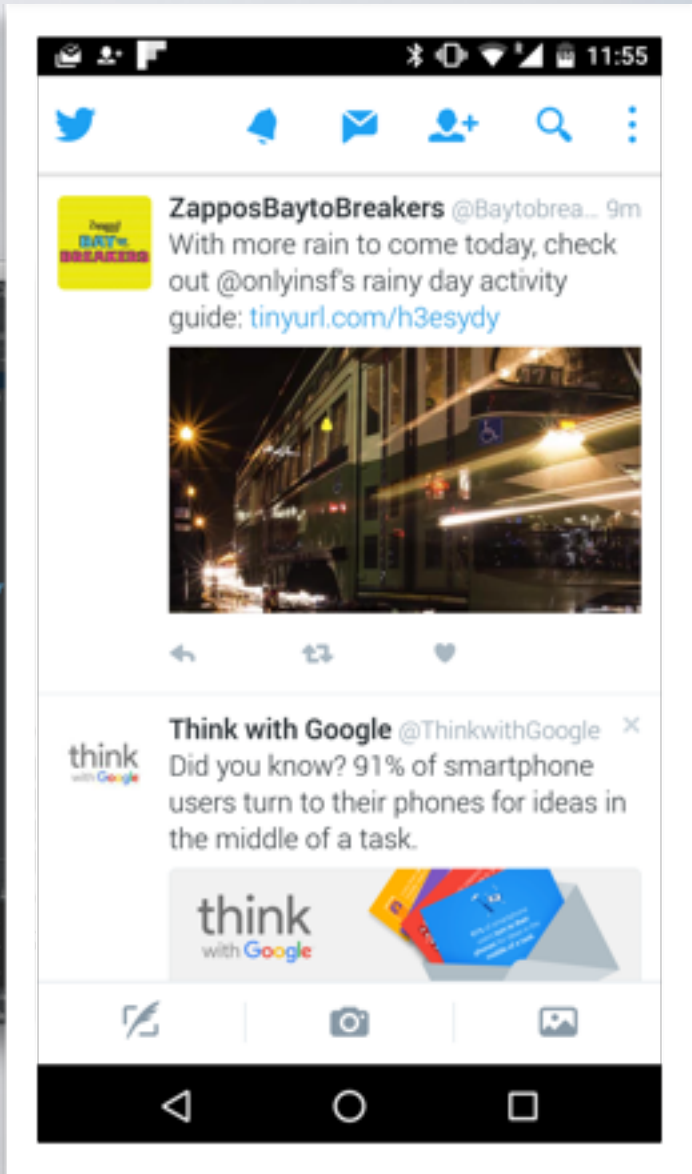
App

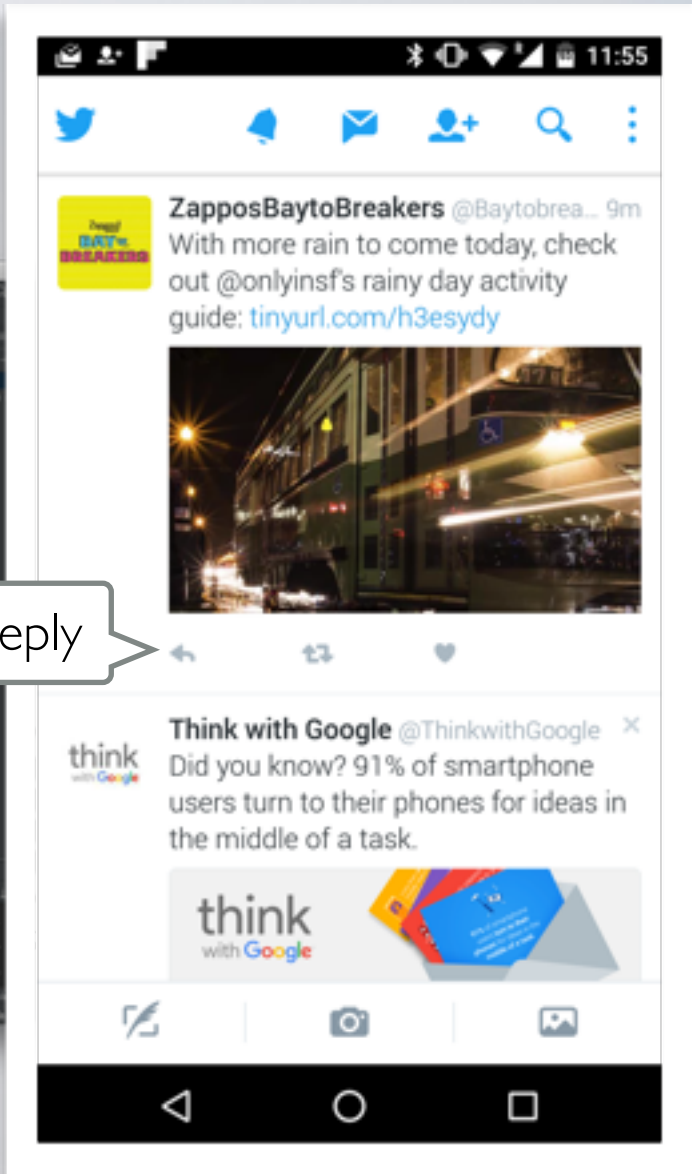
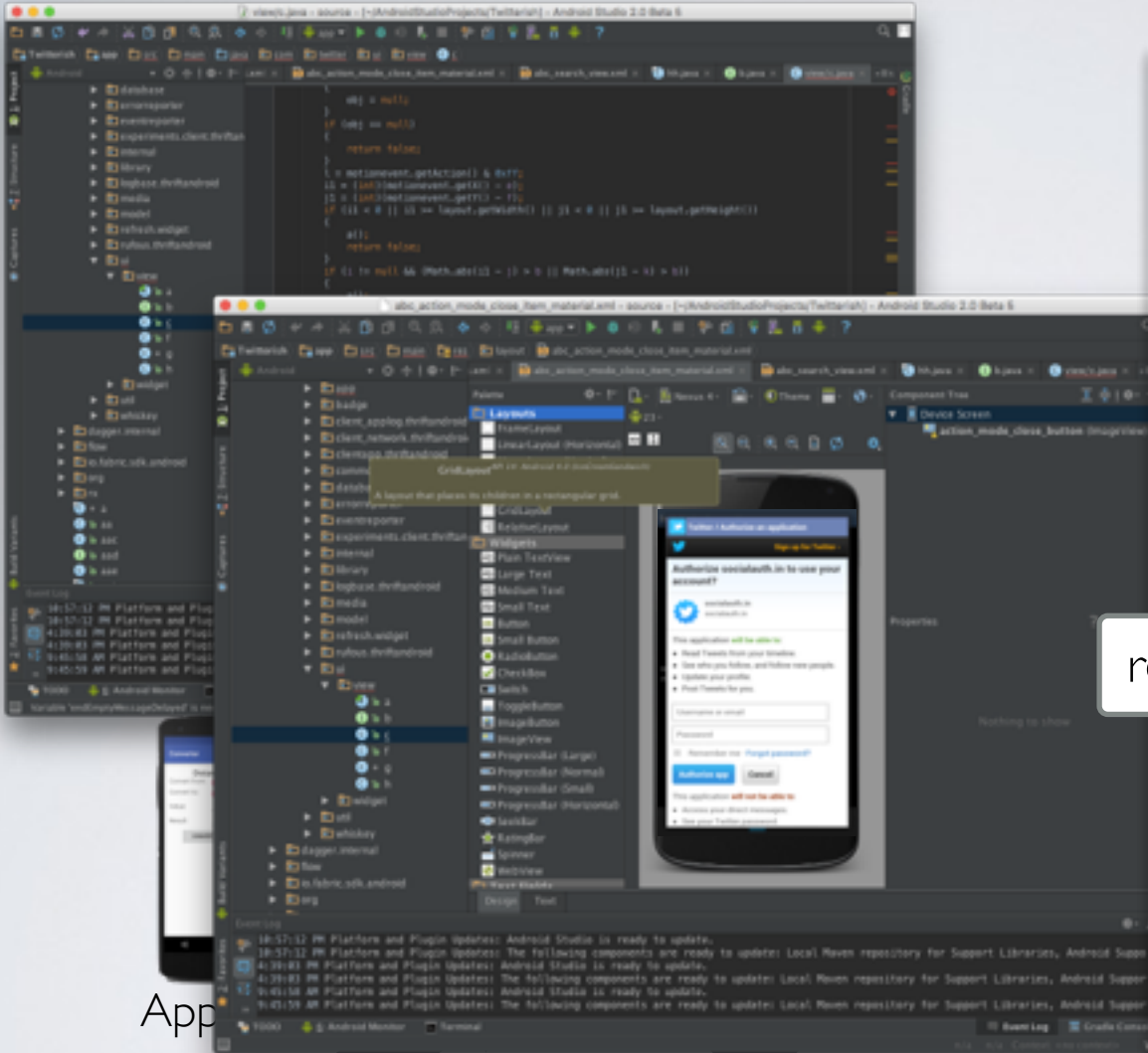


Domain Knowledge



Actions Dictionary





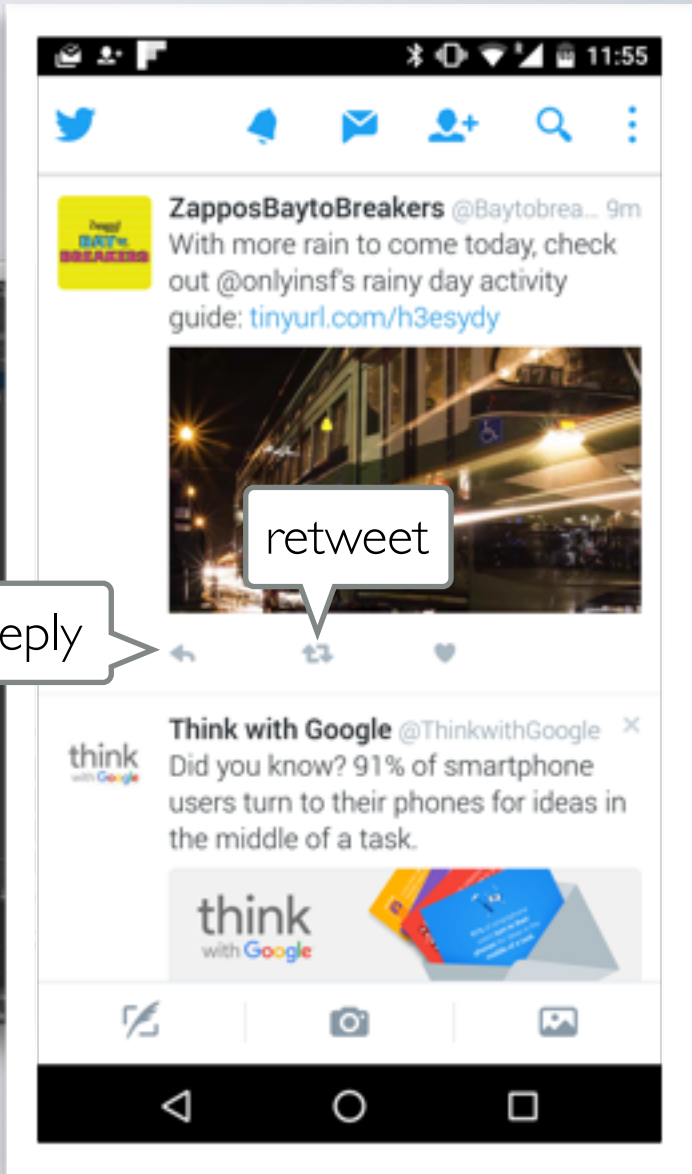
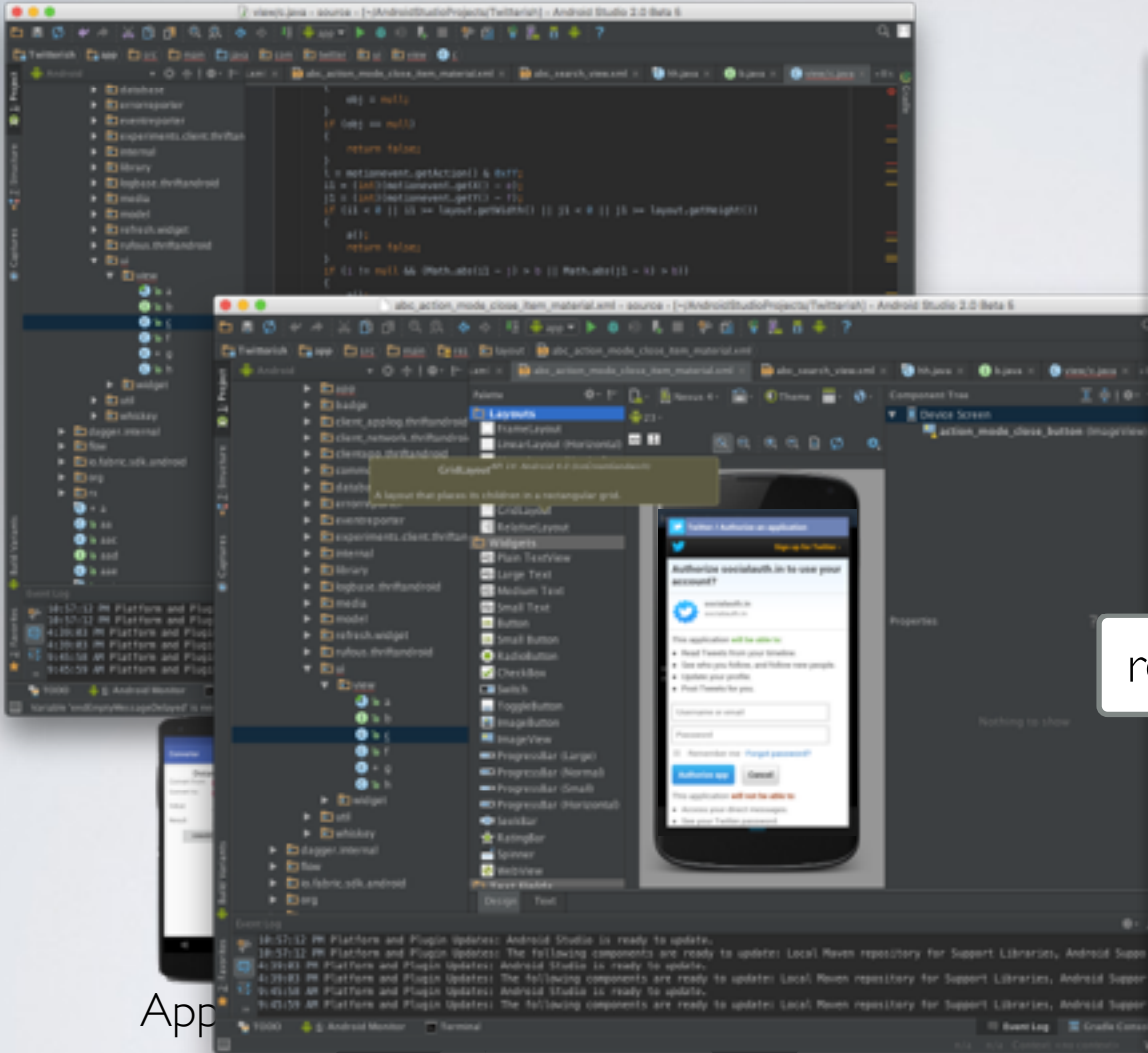
App



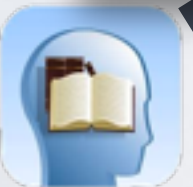
Domain Knowledge



Actions Dictionary



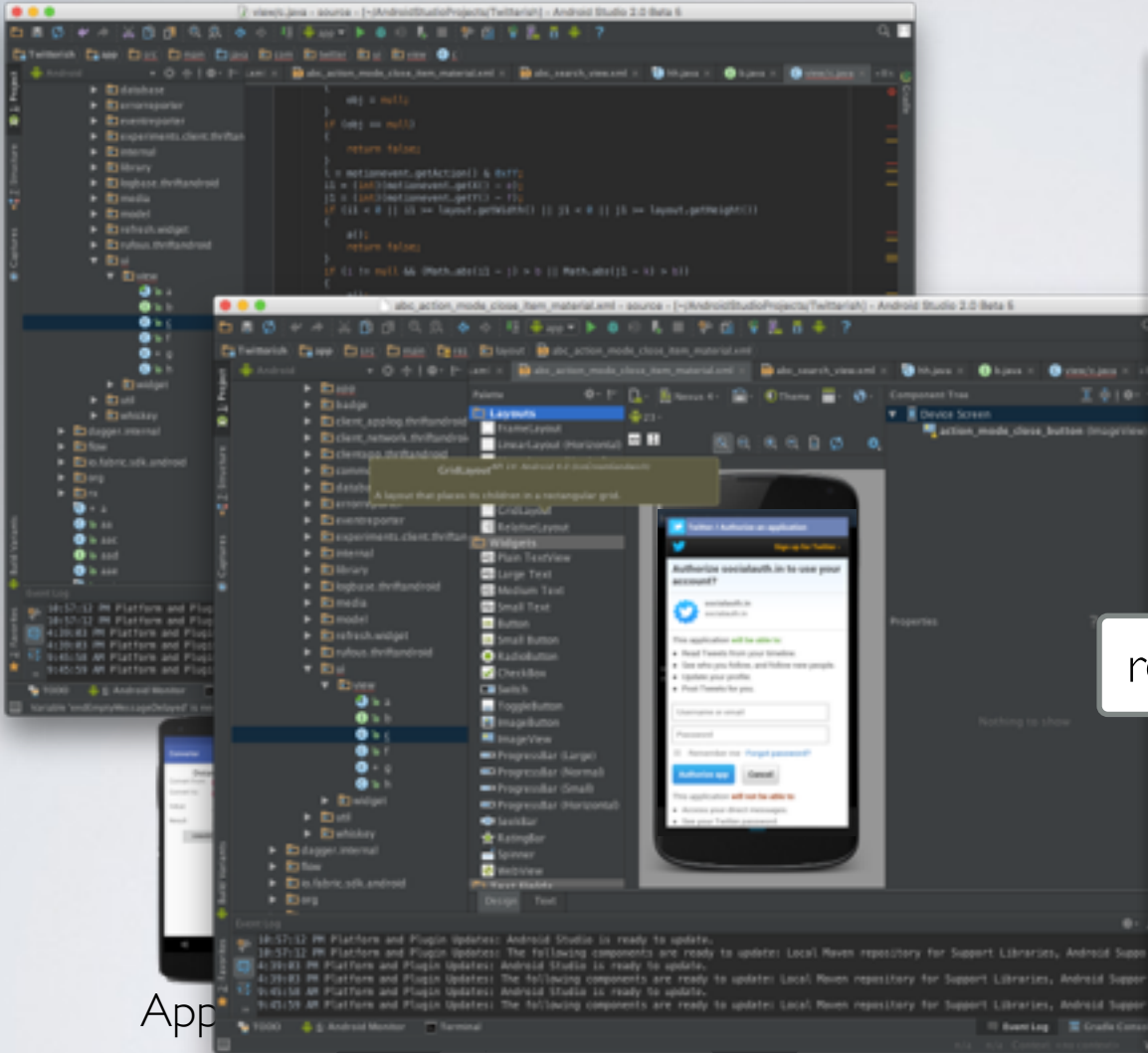
App



Domain Knowledge



Actions Dictionary



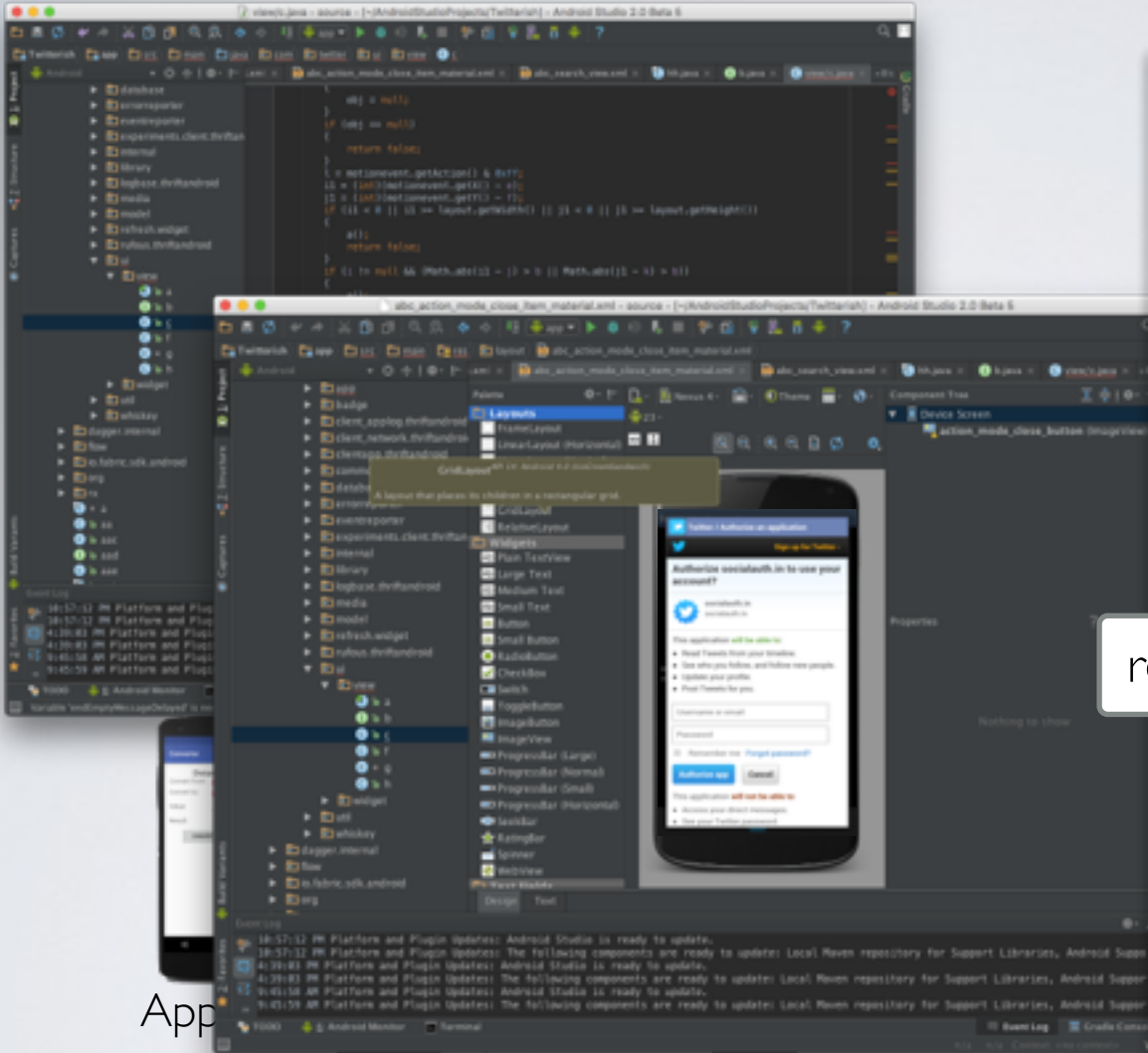
App



Domain Knowledge



Actions Dictionary



notifications



retweet

reply

like

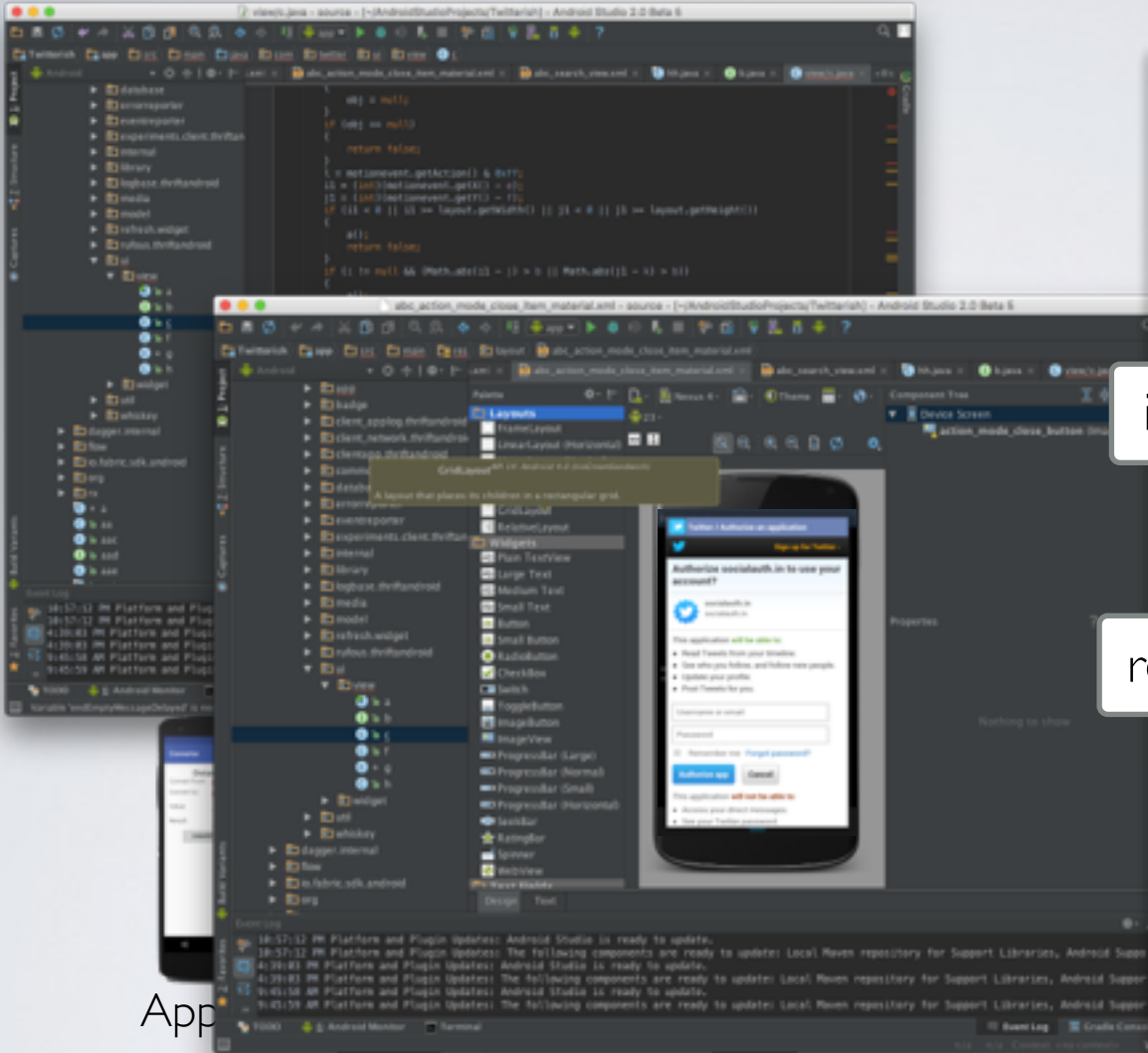
App



Domain Knowledge



Actions Dictionary



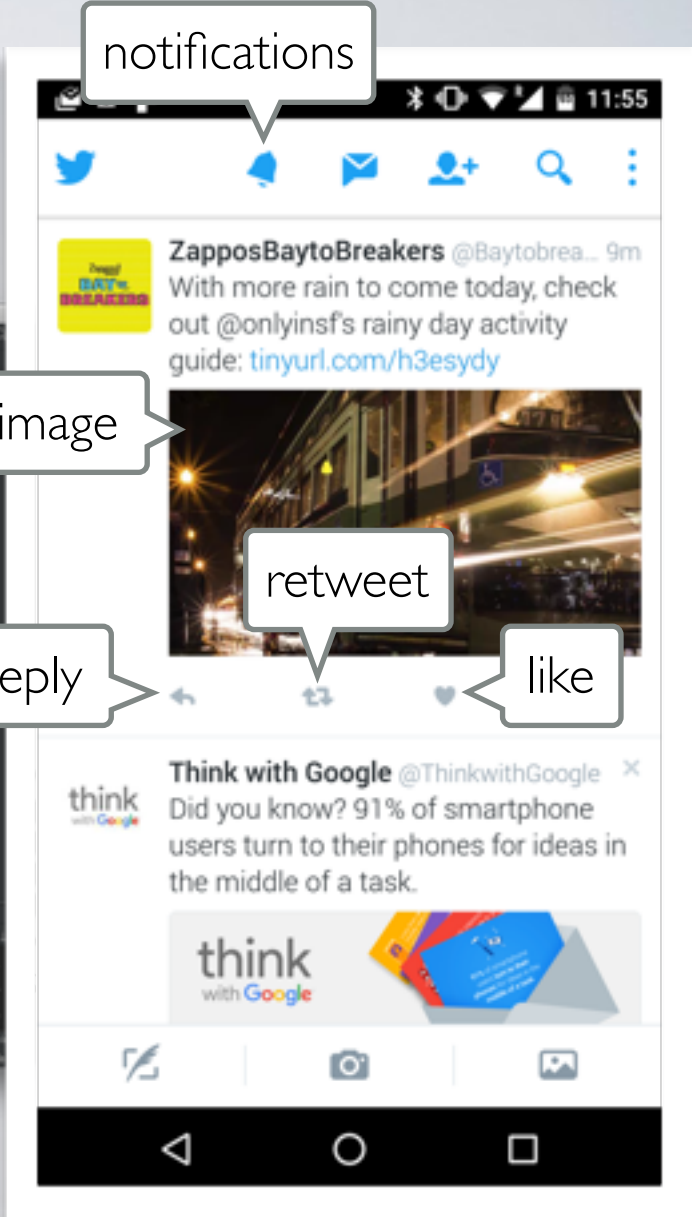
notifications

image

retweet

reply

like



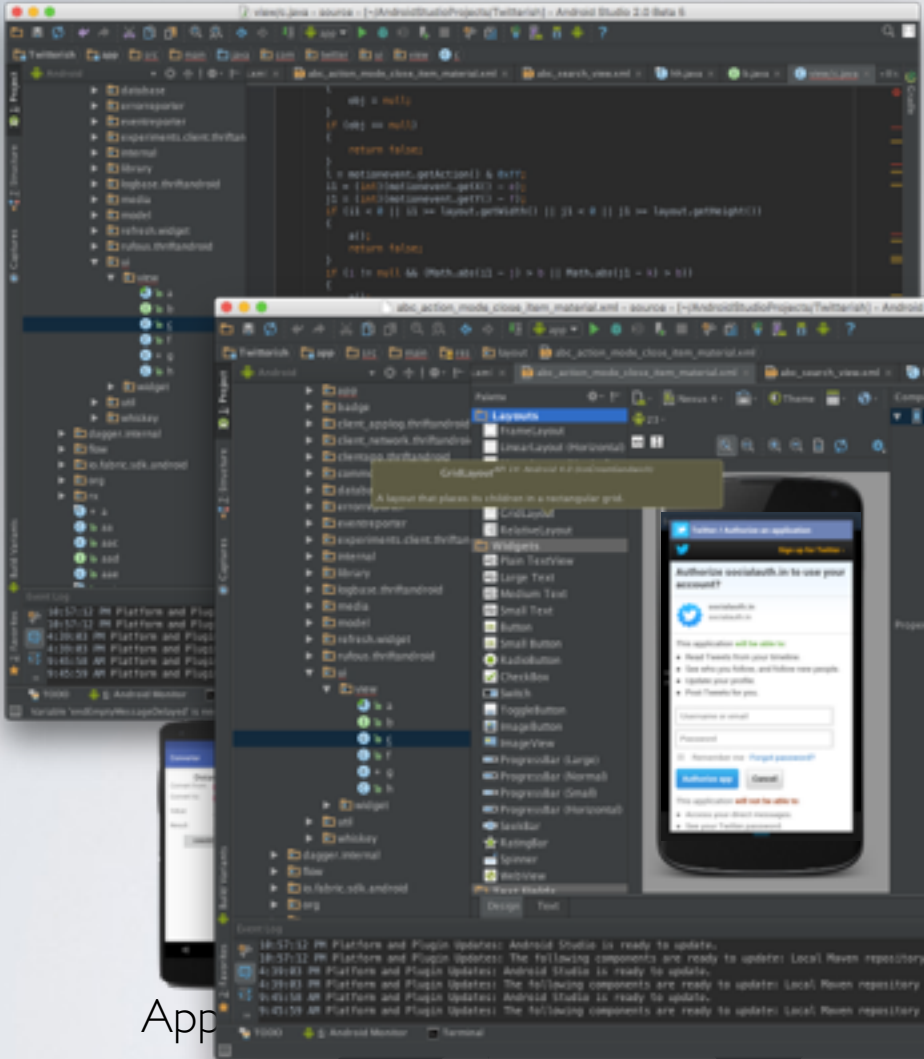
App



Domain Knowledge



Actions Dictionary



App

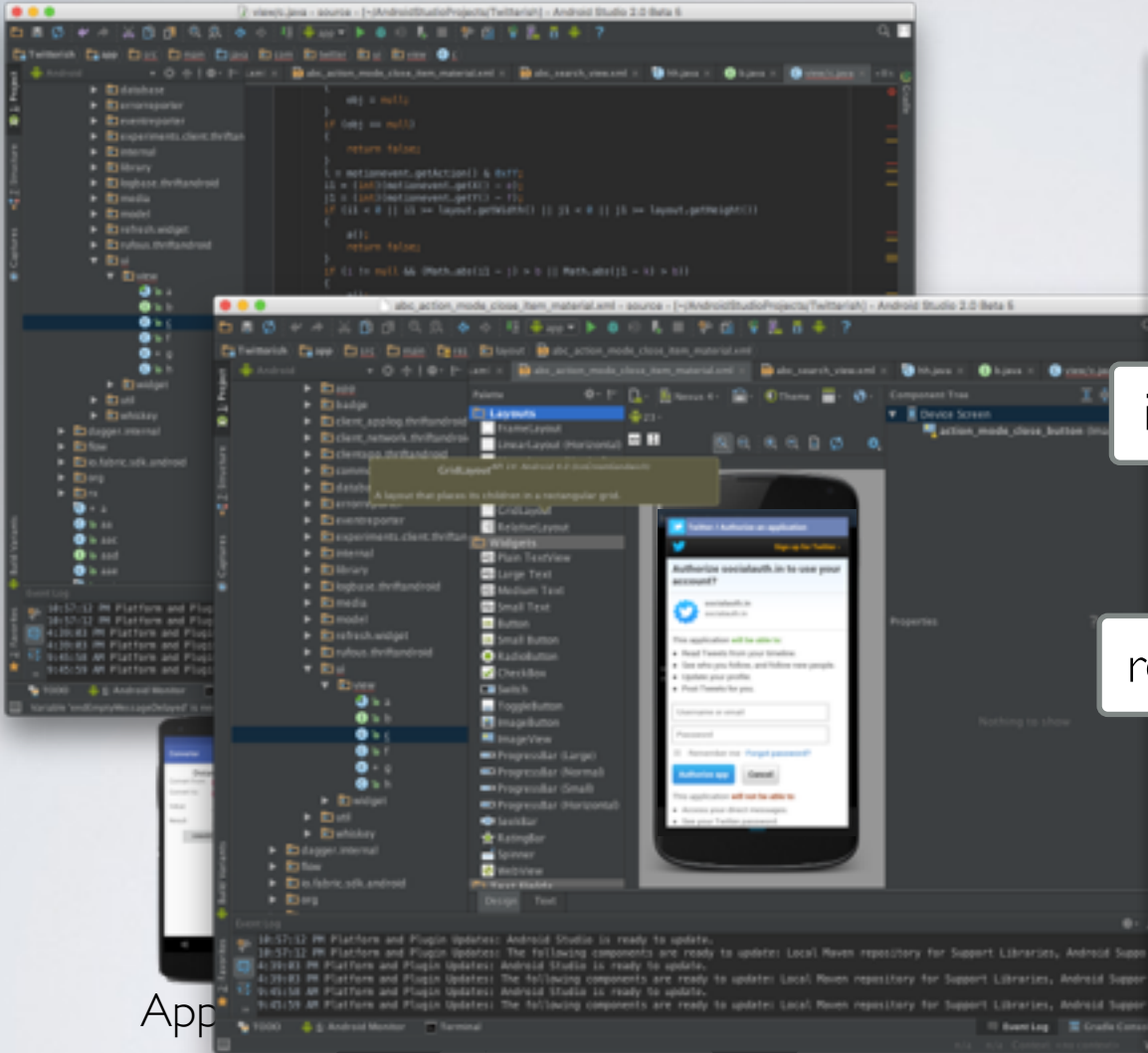


Domain Knowledge



Actions Dictionary

<action1> <parameter*> <trigger>
 <action2> <parameter*> <trigger>
 ...



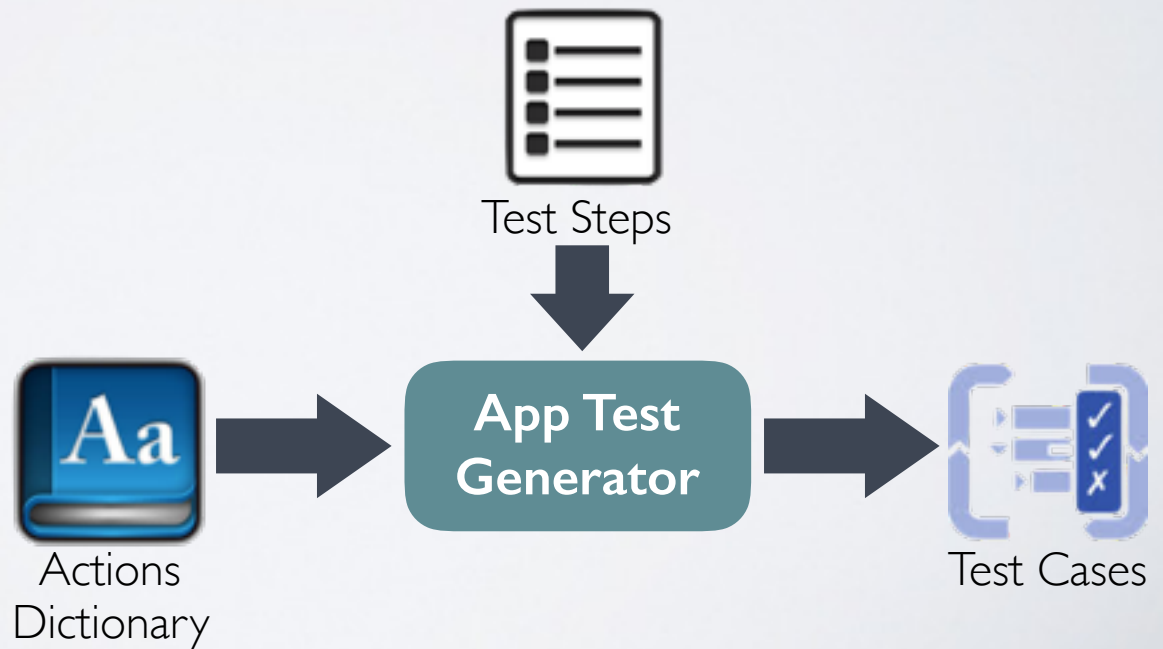
App



Domain Knowledge

<open twitter> <> <start com.twittwer.app.twitter>
 <retweet> <current tweet> <click  >
 ...

Dictionary



- 1) Open twitter
 - 2) Select tweet with image
 - 3) Retweet
- ➔ Crash

Test Steps



Actions
Dictionary




App Test
Generator



Test Cases

- 1) Open twitter
 - 2) Select tweet with image
 - 3) Retweet
- ➔ Crash

<open twitter> <> <start com...twitter>
<retweet> <current tweet> <click  >
...

Actions
Dictionary


Test Steps

App Test
Generator



Test Cases

- 1) Op
- 2) Sel
- 3) Re
- ➔ Cr


<open twitter> <> <start com...twitter>
<retweet> <current tweet> <click  >
...



Actions
Dictionary



- 1) Op
- 2) Sel
- 3) Re
- ➔ Cr


<open twitter> <> <start com...twitter>
 <retweet> <current tweet> <click  >
 ...



Actions
Dictionary



- 1) Op
- 2) Sel
- 3) Re
- ➔ Cr


<open twitter> <> <start com...twitter>
 <retweet> <current tweet> <click  >
 ...



Actions
 Dictionary



- 1) Op
- 2) Sel
- 3) Re
- ➔ Cr

<open twitter> <> <start com...twitter>
 <retweet> <current tweet> <click  >
 ...



Actions
Dictionary



CURRENT AND FUTURE WORK

CURRENT AND FUTURE WORK



CURRENT AND FUTURE WORK

A person is walking away from the camera down a long, straight asphalt road that stretches into the distance. The road is flanked by a dry, desert-like landscape with sparse vegetation and low mountains in the background under a clear sky. The person is wearing a light-colored shirt, dark pants, and a backpack.

- **Review Classifier**: reuse
- **Bug Report Miner**: very preliminary version
- **App Analyzer**: previous work on feature identification and cross-platform issues
- **App Test Generator**: previous work on Android testing

OPEN QUESTIONS/CHALLENGES

- Is it feasible? What are the metrics of success?
- What technologies are best suited for the tasks involved? Can we use them in a black-box fashion?
- Can we leverage multiple similar reviews/bug reports?
- Can we do this interactively (as the users enter their reviews)?

