**It's your turn to make a presentation. Believe in your own likeability, and then go easy on yourself. It's been said that performing is just the art of keeping a large group of people from coughing.**

**Minneapolis Star Tribune, May 26, 2007**

# Was that Thinking?

## Issues understanding development time with non-intrusive data collection

**Christine A. Halverson**
**IBM Research**
**Social Computing Group**

# Overview

- Understanding *Programmer* Productivity
- A look at intrusive and non-intrusive data collection examples
- Why it matters
- A call for engagement

# Programmer Productivity

- **Not easy to measure or quantify**
- **Affected by different factors**
  - Size of programs
  - Complexity of programs
    - Cognitive, mathematical, systemic
  - Communication between programmers
  - Time constraints
  - Social factors
- **Some more amenable to study**

# Approaches

- **Manual data gathering**
  - Programmer reflection and self report
    - Diaries, journals, notes
  - Researcher motivated reports
    - Surveys, interviews relying on post programming memories
  - Objective Observations
- **Automated data gathering**
  - Instrument applications and machine behaviors.
- **Hybrid methodologies**

# Examples

- **Combine observations with automated data gathering**
- **Study at Pittsburgh Supercomputing Center 2005**
  - SUMS, observation
- **Pilot study 2007**
  - Hackystat, observation, Camtasia

# Sample Automated Data

- **The data we get is somewhat scarce**
- **Easy to collect (once code is written)**


| | | | |
|---|---|---|---|
| U6 | 2005-05-25 09:46:40 | shell | command: a.out |
| U6 | 2005-05-25 09:46:49 | shell | command: vim |
| U6 | 2005-05-25 09:47:14 | shell | command: vim |

# Observations

- **Difficult to collect**
- **More detailed**
- **Covers areas not on the computer**

UPC User

Source

Subj: U6 Date: 5/25 Coder: JW

Date & Time

State*1: ED

Has an Editor window open

9:45a LK SCR, PG U/D, WR SP, LK HO (Global alignment)

9:46a LK SCR, HNK, HOK, CL 1, LS [quitting the editor took him back to a shell prompt], RUN, ED (Output.txt), HNK

9:47a GL SP, HOK, O (Edmiston_final.c), PG U/D, CUR D, Left hand on mouth, POINT (MIN statement)

Looks at the screen, pages up then down, writes on the scratch pad, looks at the handout (from learning UPC) on Global alignment

Subj: U6 Date: 5/25 Coder: JW

State*1: ED

Editor window open

9:45a LK SCR, PG U/D, WR SP, LK HO (Global alignment)

Looks at the screen, pages up then down, writes on the scratch pad, looks at the handout (from learning UPC) on Global allignment

9:46a LK SCR, HNK, HOK, CL 1, LS [quitting the editor took him back to a shell prompt], RUN, ED (Output.txt), HNK

Looks at screen, hands not on keyboard, hands on keyboard, closes window 1, back to shell prompt, types ls, runs the code, opens editor with output.txt, takes hands off keyboard

9:47a GL SP, HOK, O (Edmiston_final.c), PG U/D, CUR D, Left hand on mouth, POINT (MIN statement)

Glances at the scratch pad, hands on keyboard, Opens Edmiston_final.c file, pages up and down, moves the cursor down in the file, moves left hand to mouth, points to MIN statement on screen

# Map Automated to Observed data

9:45a LK SCR, PG U/D, WR SP, LK HO (Global alignment)

9:46a LK SCR, HNK, HOK, CL 1, LS [quitting the editor took him back to a shell prompt], RUN, ED (Output.txt), HNK

9:47a GL SP, HOK, O (Edmiston_final.c), PG U/D, CUR D, Left hand on mouth, POINT (MIN statement)

| | | |
|---|---|---|
| 09:46:40 | shell | command: a.out |
| 09:46:49 | shell | command: vim |
| 09:47:14 | shell | command: vim |

# Two separate inferences

- Trying to work out a problem – possibly around global alignment.
- Looks through code, making notes, probably thinking
- Running code (possibly to just observe behavior) before opens file, searches for the right place, then ponders the screen, pointing at a MIN statement

- Running code, looking at output
- Opening file (to edit)

**Thinking**
**Running code**

**Running code**
**Writing code**

# Different Quantifications

9:46a LK SCR, HNK, HOK, CL 1, LS [quitting the editor took him back to a shell prompt], RUN, ED (Output.txt), HNK

9:47a GL SP, HOK, O (Edmiston_final.c), PG U/D, CUR D, Left hand on mouth, POINT (MIN statement)

| | | |
|---|---|---|
| 09:46:40 | shell | command: a.out |
| 09:46:49 | shell | command: vim |
| 09:47:14 | shell | command: vim |

Run, open and view output 46:30?

Glance at scratch pad 47:00

Opens program file 47:15?

Moves around in file

ponders file 47:30?

Run, produces output file at 46:40

Opens output file 46:49

Open program file 47.14

# Differences Matter

- 09:46:40 runs code
- 09:46:49 starts editor
- 09:47:14 starts editor

1. Run, open and view output 46:30?
2. Glance at scratch pad 47:00
3. Opens program file 47:15?
4. Moves around in file
5. ponders file 47:30?

- 9 sec running code
- Opens editor
- 25 seconds later starts 2nd editor

- 10 sec running code
- 20 sec looking at output
- 15 sec looking at notes
- Open program file
- 15 sec moves around
- 30 sec Ponders file

Is subject editing? Writing code?
Looking at code? Looking at output?
Was that Thinking?

Looking at code, not editing.
Looks like thinking through the problem – looking at output 20
looking at notes 15
ponders file 30
**1m 5s**

# Another example

Edit Eclipse C:/DWR/ntkr/chrome/notetaker/libjs/eventutils.js 05/22/2007-09:46:30 {unit-type=file, subtype=Open, unit-name=eventutils.js}

Edit Eclipse C:/DWR/ntkr/chrome/notetaker/libjs/eventutils.js 05/22/2007-09:46:34 {modified=true, to-buff-name=C:/DWR/ntkr/chrome/notetaker/libjs/eventutils.js, from-buff-name=C:/DWR/ntkr/chrome/notetaker/tablelist.xml, subtype=BufferTransition}

Edit Eclipse C:/DWR/ntkr/chrome/notetaker/libjs/eventutils.js 05/22/2007-09:46:43 {}

Edit Eclipse C:/www/xampp/htdocs/libjs/eventutils.js 05/22/2007-09:47:32 {unit-type=file, subtype=Open, unit-name=eventutils.js}

Open file

Edit file

Copied from tablelist.xml

Save file

Open file

9:46 30-9:46:43 Editing  9:46:43-9:47:32 Running code.

# Why it matters

- **Development time and phase varies**
- **Still missing much of the complexity and human interaction that is involved**
  - Does a gap indicate: thinking, getting help, going to lunch? Or all 3?
- **How do you know if a tool intervention really made the difference?**
- **How do you quantify the difference that it made?**

# A Call to arms

- How can we make unobtrusive, automatic data collection more detailed?
- Can we integrate screen capture/attached cameras to pick up when nothing is happening to engage the automated data collection?
- Can we build studies that combine automated and observational data and determine patterns of behavior to better make inferences?

# Comments? Questions? Ideas?

*This material is based upon work supported by the Defense Advanced Research Projects Agency under its Agreement No. HR0011-07-9-0002.*