Using activity traces to characterize programming behaviour beyond the lab

Gail Murphy*
Petcharat Viriyakattiyaporn*
David Shepherd†

* This talk contains copyright pictures obtained under license. The license associated with this talk does not apply to these pictures.
experimentation is hard:

lack of access to experienced programmers

participating programmers may not be representative

small tasks/code bases may not be representative of actual tasks/code bases
field study

controlled experiment

From “A comparative study of three program exploration tools”, de Alwis et al 2007
markers of a behaviour

field study

controlled experiment

collect trace data

frequency of behaviour
an example

several research tools target local context investigations (e.g., RELO)
1. develop a marker

Local Context Measure \( (LCM) = 2 \)
2. apply marked on controlled data

3 traces from controlled study with source code available
3. validate

does LCM characterize code navigation behaviours of interest?

are similar behaviours seen in field data?
3. validate

does LCM characterize code navigation behaviours of interest?

predicted LCM in 70% of 10 significant edit points for which we conducted screen video analysis

5 traces from second controlled study with source code available
3. validate

are similar behaviours seen in field data?

5 traces from a field study with no source code available
limitations

measure may not be complete

measure may depend on programming style

trace data contents limit behaviours investigated
need to improve empirical techniques

indicators of programming behaviours
defined over activity traces
may help scale programming investigations