Measurement of Automation

Prepared and presented
by
Mark Fewster
Grove Consultants
Llwyyncynhwyr, Cwmdu
Llandeilo, SA19 7EW, UK
Tel: +44 1558 685180    Email: mark@grove.co.uk
Fax: +44 1558 685181    www.grove.co.uk
© Grove Consultants, 2006

Measurement of Automation

Contents

Some principles
Important automation metrics
Internal automation metrics
Guidelines and recommendations

© Grove Consultants, 2006
Useful measures

- **a useful measure:**
  “supports effective analysis and decision making, and that can be obtained relatively easily.”
  

- easy measures may be more useful even though less accurate (e.g. car fuel economy)
- ‘useful’ depends on objectives, i.e. what you want to know

---

A few principles

- any measure is better than no measure
  - need not be perfect, may not be adequate
- two measures better than one?
- presentation of measures is important
  - see “Visual Display of Quantitative Data”
- every measure has limitations
- any measure can be abused
- currency of measure must be meaningful
Measurement of Automation

Contents

Some principles

**Important automation metrics**

Internal automation metrics

Guidelines and recommendations

A particularly useful metric: EMTE

- **equivalent manual test effort (EMTE)**
  - a simple measure
  - easily understood
  - useful in several ways

- **test case EMTE = amount of effort required to run the test case manually**
  - can relate other measures to EMTE
  - more meaningful than plain averages
Important measures for automation

- **benefit**
  - related to the main objective(s) of automation
  - to prove that effort spent automating gives greater value than it would if spent running tests manually

- **recurring costs**
  - the largest and most frequent
    - build cost (effort to automate new tests)
    - failure analysis cost (effort investigating)
    - maintenance cost (effort updating testware)

Measures of benefit

- objective: same tests, less effort
  - e.g. 4 hours* to run 120 tests with 40 hrs EMTE
    saving = 36 hours for 1 test cycle

- relate target to cost of automation
  - e.g. cost to build = 40 minutes/test
    for 120 tests, cost to build = 80 hours
    2 test cycles required to break even
    4 test cycles required for 2 fold ROI

*4 hours effort to invoke automated tests and analyse failures
Measures of benefit

- objective: more thorough testing, same effort
  • e.g. 4 hours* to run 120 tests with 40 hrs EMTE
    saving = 36 hours
  • use the 36 hours to specify and run new manual tests, then employ usual metric(s) used to measure test thoroughness, e.g.
    – functional/structural coverage
    – number of test cases/conditions covered
    – additional (manual) execution time

*4 hours effort to invoke automated tests and analyse failures

Measures of benefit

- objective: unattended testing
  • e.g. hours of additional (unattended) machine use

- objective: faster testing (same testing, less time)
  • average number of tests run per hour of test effort

- objective: reduce testing cost
  • average effort per test case run

- objective: run more tests
  • number of test cases run

- objective: repeatable testing
  • number of repeated runs
Measures of build effort

- time taken to automate tests
  - hours to add new or existing manual tests
  - average across different test types
- proportion of equivalent manual test effort
  - e.g. 1 hour to automate 30 minute manual test
    - 2 times equivalent manual test effort

  Suggestion
  | Target: | < 2 times |
  | Trend:  | decreasing 10% per year |

Measures of failure analysis effort

- analysis effort for each test
  - captured in fault report
  - effort from first recognition through to resumption of test execution
  - average hours (or minutes) per failed test case
  - % tests failing

  Suggestion
  | Target: | 15 minutes |
  | Trend:  | stable |
Measures of maintenance effort

- **maintenance effort of automated tests**
  - percentage of test cases requiring maintenance
  - average effort per test case
  - percentage of equivalent manual test effort
  - frequency of software changes that impact automated tests

**Suggestion**

<table>
<thead>
<tr>
<th>Target</th>
<th>&lt; 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend</td>
<td>stable or decreasing</td>
</tr>
</tbody>
</table>

---

**Measurement of Automation**

**Contents**

- Some principles
- Important automation metrics
  - Internal automation metrics
- Guidelines and recommendations
Internal (to automation) measures

- **indicators of test automation health**
  - number of tests versus scripts
  - total execution time of tests
  - total EMTE
  - number of uses (executions)
  - number of software faults found
  - number of test faults found
  - number of new (automated) tests
  - number of new scripts

---

Attributes of test automation

- **reliability**
  - % automated tests with faults
  - no. test iterations due to test faults
  - no. false negatives (test failed but result OK)

- **flexibility**
  - time to test an emergency fix
  - time to identify and execute a specific subset of tests (e.g. for regression testing)
  - time to restore archived tests
Attributes of test automation

- **usability**
  - average time to become confident and productive
  - effort to perform main tasks (e.g. select/invoke tests, analyse failures)
  - users’ own rating
- **robustness**
  - no. tests that fail due to one software fault
  - frequency of test failures due to unexpected events

Attributes of test automation

- **efficiency**
  - effort to perform tasks (e.g. automating tests, preparing tests for execution, ascertaining results)
  - % test scripts used by at least $x$ tests
- **portability**
  - effort to prepare automated tests for execution in new environment
  - no. of different environments supported
  - effort to make tests run using a different tool
Measurement of Automation

Contents

Some principles
Important automation metrics
Internal automation metrics
Guidelines and recommendations

Collecting the facts

- do not collect information for the sake of it
- decide how often you need to report
- decide who needs the information
- decide what information is needed
- decide why that information is needed
- expect to refine the data collection

Source: Isabel Evans “Get your message across”
Presenting measures

- figures embedded in explanatory text
  - may not be read
- tables of figures
  - important information may not stand out
- graphs
  - particularly good for trends and comparison
- dash board of instruments
  - helps to characterise information

Figures embedded in text

The latest testing effort proved to be more difficult than planned. Of the 30 programs to be tested only 10 were delivered on time and 5 were delivered with only one week remaining of the test schedule.

The quality of the software was also poorer than expected. In the first week 47 faults were found of which 18 were of high severity and 19 medium severity. This compares with an expected total of 30 with a maximum of 10 high severity and 10 medium severity faults. During the next two weeks of testing a further 55 faults were found, 22 of them high severity and 16 medium severity.
Table of figures

<table>
<thead>
<tr>
<th>No. Programs</th>
<th>10</th>
<th>1</th>
<th>8</th>
<th>4</th>
<th>0</th>
<th>2</th>
<th>5</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days late</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

Inconsistent scales may be confusing

<table>
<thead>
<tr>
<th>Week no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>18</td>
<td>14</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Medium</td>
<td>19</td>
<td>11</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Low</td>
<td>10</td>
<td>6</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>31</td>
<td>24</td>
<td>15</td>
</tr>
</tbody>
</table>

Graphs

![Graphs](image-url)
Dash board (ideas)

Graphics

- **graphics reveal data**
  - use them to display complex ideas clearly, precisely and efficiently, but ... 
  
  “As to the propriety and justness of representing sums of money, and time, by parts of space, tho’ very readily agreed to by most men, yet a few seem to apprehend that there may possibly be some deception in it, of which they are not aware ...”

William Playfair, “The Commercial and Political Atlas” (1786)
source: Isabel Evans, “Get you message across”, Eurostar 2002
The Lie Factor

\[
\text{The Lie Factor} = \frac{\text{size of effect in graphic}}{\text{size of effect in data}}
\]

<table>
<thead>
<tr>
<th>Weeks</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of bugs</td>
<td>50</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
</tbody>
</table>

Using the diameter of the circle to show the data:

Using the area of the circle to show the data:

Presentation guidelines

- **use charts / graphs were possible**
  - make the information visually interesting
- **keep important information visible**
  - large and stuck to the wall
- **use (primary / bright) colours**
  - red signifies problems / risk
  - green signifies things are OK
- **keep information up-to-date**
Recommendations

- **measure benefit with respect to objectives**
  - these may change over time, must be kept challenging

- **measure recurring costs**
  - largest costs: build, failure analysis, maintenance
  - set targets for reduction (per test)

- **be credible**
  - don’t artificially inflate

---

**Measurement of Automation**

Summary

- **there are no perfect measures**
  - measuring something is better than no measure

- **objectives of automation indicate important metrics**

- **measure benefit**
  - most important (since costs are more obvious)

- **measure specific costs**
  - so you know where to improve