

Adaptive Test Case Selection - Data Sets

These data sets provide historical information about test case executions and their results. It can be used to evaluate test case prioritization and selection methods, finding test cases most likely to fail during their next execution. Test cases are defined by their execution duration, their previous last execution time and results of their recent executions.

Two of these data sets are provided by ABB Robotics Norway, the other one was published by Google and is included here in its converted form.

Data Set	Test Cases	CI cycles	Verdicts	Failed
ABB Paint Control	89	352	25,594	19.36%
ABB IOF/ROL	1,941	320	30,319	28.43%
Google GSDTSR	5,555	336	1,260,617	0.25%

For the original GSDTSR data set repository see: *Sebastian Elbaum, Andrew McLaughlin, and John Penix, "The Google Dataset of Testing Results", <https://code.google.com/p/google-shared-dataset-of-test-suite-results>, 2014*

File overview

The three data sets: - `paintcontrol.csv`: ABB Paint Control data set - `iofrol.csv`: ABB IOF/ROL data set - `gsdtsr.csv`: Google Shared Dataset of Test Suite Results (GSDTSR)

The GSDTSR data set can be re-created from the original data: - `convert_gsdtsr.py`: Converts the original GSDTSR data `testShareData.csv.rev.gz` to `gsdtsr.csv` in the common file format. - `requirements.txt`: Necessary Python packages to run `convert_gsdtsr.py` - `testShareData.csv.rev.gz`: Original GSDTSR data with removed comma at beginning of file (see issue #1 at the original repository)

File Format

The data sets are provided as CSV files (delimiter ';'). Because we are interested in detecting failures, test results indicate failures by True and passed runs by False to simplify our software.

Column Name	Content
Cycle	The number of the CI cycle this test execution belongs to
Id	Unique numeric identifier of the test execution
Duration	Approximated runtime of the test case
LastRun	Previous last execution of the test case as date-time-string (Format: YYYY-MM-DD HH:ii)
LastResults	List of previous test results (Failed: True, Passed: False), ordered by ascending age. Lists are delimited by [].
Verdict	Test verdict of this test execution (Failed: True, Passed: False)