Stanford Exploration Project, Report SERGEY, May 1, 2000, pages 649–??

Stanford Exploration Project, Report SERGEY, May 1, 2000, pages 649–??

Short Note

Reproducible research - results from SEP-100

Marie L. Prucha, Robert G. Clapp, Sergey Fomel, Jon Claerbout, and Biondo Biondi[1]

INTRODUCTION

SEP has been striving to create reproducible research for many years. Our first attempts at reproducible documents began with the introduction of interactive documents (Claerbout, 1990). We then moved on to putting SEP reports on CDROMs and using "cake" (Nichols and Cole, 1989) so that the results could be recreated using the author's own processing flow. Later we updated to the GNU make system (Schwab and Schroeder, 1995). Now SEP reports are available online and can be downloaded. Sponsors can see exactly how most of the figures in each paper are created, and those with the type of environment we have at SEP can recreate most of the figures themselves. This paper will explain what we consider reproducible research to be, why we go through the effort of making our research reproducible, how we test reproducibility, and the results of the testing on our last report, SEP-100.

REPRODUCIBLE RESEARCH

Reproducible research can be easily recreated by people other than the author. Whenever possible, the authors of SEP papers make their figures either easily reproducible (ER) or conditionally reproducible (CR). An easily reproducible figure can be deleted and quickly rebuilt by anyone interested in the result. To recreate the figure, the author provides the necessary input data, makerules, parameter files, and source code. A conditionally reproducible figure can be reproduced, but it may require input data that is not provided because of size or proprietary issues, special system requirements such as parallel processing or special software, or it may take more than 20 minutes to rebuild. Both ER and CR figures assume that the environment they are being recreated in has X-windows, SEPlib, SEP makerules, and Fortran77, Fortran90 and C compilers. If a figure can't be reproduced, such as a hand-drawn picture, it is marked NR for non-reproducible.

[1]**email:** marie@sep.Stanford.EDU, bob@sep.Stanford.EDU, sergey@sep.Stanford.EDU, jon@sep.Stanford.EDU, biondo@sep.Stanford.EDU

IMPORTANCE OF REPRODUCIBILITY

We feel that reproducibility is important for helping other researchers understand and build from our research. Any other researcher can see exactly how each reproducible figure is created. Reproducibility also allows the author to go back to old work and try new ideas. Many SEP students rely heavily on the reproducibility of their old reports when thesis-writing time arrives.

TESTING REPRODUCIBILITY

Each SEP report is tested for reproducibility. Each paper is tested in four categories: ER reproducibility, CR reproducibility, clean rules, and independence. ER figure reproducibility is tested by deleting all ER figures and attempting to rebuild them using the author's makerules. All CR figures are checked to make certain that they do have accurate makerules. The clean rule check means that no files except the ER figures should be created in the build process, no files other than the results exist in the results directory, and no object or binary files exist after the clean rules are run. Finally, no makerules may attempt to access binaries outside of the paper directory. These requirements are considerably stricter than they have been in past years.

RESULTS FROM THE LAST REPORT

The testing of SEP-100 was completed before it was made available as an electronic document last spring. All testing was carried out on an IRIX 6.5 operating system. The ER testing was done first. The results are shown in the following table.

Testing Round	No. of ER figs	No. that failed	% that passed
1	95	7	92.6%
2	93	2	97.8%
3	93	0	100.0%

The CR testing was carried out on a paper-by-paper basis. If one CR figure failed, the paper failed. The results are in the following table.

Testing Round	No. of CR figs	No. of papers	No. that failed	% that passed
1	105	22	10	54.5%
2	105	22	2	90.9%
3	105	22	0	100.0%

The NR figures were examined by the editors of each paper to determine if the figure should be considered non-reproducible or if the author needed to try to make it reproducible. There were 48 NR figures before the reproducibility testing for the ER and CR figures began and 50 NR figures after the testing was complete.

CONCLUSIONS

Overall, the reproducibility of SEP-100 was good, even before the testing. Testing on this report (SEP-102) is in progress.

REFERENCES

Claerbout, J. F., 1990, Active documents and reproducible results: SEP–**67**, 139–144.

Nichols, D., and Cole, S., 1989, Device independent software installation with CAKE: SEP–**61**, 341–344.

Schwab, M., and Schroeder, J., 1995, Reproducible research documents using GNUmake: SEP–**89**, 217–226.