

TestR: R language test driven specification

Petr Maj*, Tomas Kalibera, Jan Vitek

Purdue University, West Lafayette, IN

*Contact author: peta.maj82@gmail.com

Keywords: Formal Specification, Testing, Regression, Compatibility Testing

Every computer language that seeks global adoption and a large user base requires a precise specification. Language references, such as the *ECMAScript* reference for *JavaScript*, are mostly used to fulfill this aim. They provide a concise and systematic description of the language and are easy to read and understand by the programmers. However, similar to other forms of documentation, references can lag behind the actual implementation. This problem is even more pronounced in rapidly evolving systems where the costs of keeping the references synchronized with the implementation are simply too high. *R* is no exception to this rule. While it provides an excellent help system and reference guide, these sometimes disagree with the implementation and lack the clarity and precision of a formal specification.

TestR attempts to provide another solution to the specification problem. It is a large collection of tests aiming to map all features and corner cases of the *R* language in a way similar to *Java* Technology Compatibility Kit. The tests are structured by language features and come with textual descriptions so that they can be easily used as a human-readable behavioral specification, complete with examples. The tests provide a regression suite for the *R* implementation and this validation will be essential to allow identifying incompatible language changes. The tests can be run with different *R* implementations to allow for compatibility testing. Comparison of these findings will be presented.