



ECMAScript 2015 compliant RegExp modification for JONI

Bachelor thesis for Thomas Wolkenstein
E-Mail: thomas@wolkenstein.com

Graal.js is a JavaScript (ECMAScript 2015) interpreter, implemented in Java. It uses specialization to optimize its execution, e.g. specialization on data types. While JavaScript lacks data types and provides only one generic array type, internally, Graal.js provides several different array implementations, speculating on data types or other features of the array.

Graal.js uses a port of the JONI library for Regular Expression (RegExp) parsing and matching. This library is ECMAScript 5.1 compliant, but misses the new Unicode-related features of ECMAScript 2015. In this thesis, those features should be implemented. Goal is to pass the respective tests of the Test262 and the TestV8 test suites.

The scope of this thesis is as follows:

- Modify the existing port of JONI from the OpenJDK to provide the Unicode features of ECMAScript 2015.
- Providing those modifications in a way that is appropriate to publish them open source.
- Show correctness of the implementation by passing relevant tests from Test262 or TestV8.
- Avoid significant performance regressions.

Explicit non-goals are:

- Improving performance of RegExp parsing/matching by adopting a whole new execution model.
- Fixing issues that might be caused by the JDK8 not supporting the latest Unicode features.

The work's progress should be discussed with the supervisor at least every 2 weeks. Please note the guidelines of the Institute for System Software when preparing the written thesis.

The deadline for the written thesis is August 11th, 2016.

Supervisor: Dr. Matthias Grimmer, Dr. Christian Wirth