

# Apache Lucene

**22nd Sept 2010:** Harmony passes 100% of Lucene core tests with Lucene trunk.

Additionally all contrib modules and their tests pass, except the 'remote' contrib.

## How to

1. Build Lucene
  - a. Checkout lucene trunk from subversion: `svn co http://svn.apache.org/repos/asf/lucene/dev/trunk`
  - b. Change to the lucene directory
  - c. Ensure JAVA\_HOME points to harmony. Harmony needs to be at least r999725.
  - d. Compile lucene: `ant -Dbuild.compiler=extJavac`
    2. Run tests
  - e. `ant test-core -Dbuild.compiler=extJavac`

## Contrib modules

All of the contrib modules except 'remote' are also functional. For example, to build the xml-query-parser:

1. Change to the contrib/xml-query-parser directory
2. Compile: `ant -Dbuild.compiler=extJavac`
3. Run tests: `ant test -Dbuild.compiler=extJavac`

## Limitations

1. Lucene currently does not have Harmony-specific support for the [unmap hack](#) in MMapDirectory.
  - a. `FSDirectory.open` will return `SimpleFSDirectory` on Windows always, even 64-bit.
2. Because Harmony uses ICU for Unicode support, it implements a newer version of Unicode (5.2) than other Java implementations.
  1. If your index was built with another implementation, rebuild your index with Harmony to ensure queries will match.
  2. Alternatively, you can use Lucene's ICU analysis module (`modules/analysis/icu`) which does not rely on the JRE's Unicode support for text analysis.