**Bug1**

diff -r /tmp/Codec\_1b/src/java/org/apache/commons/codec/language/Caverphone.java /tmp/Codec\_1f/src/java/org/apache/commons/codec/language/Caverphone.java

59c59

< txt = txt.toLowerCase();

---

> txt = txt.toLowerCase(java.util.Locale.ENGLISH);

diff -r /tmp/Codec\_1b/src/java/org/apache/commons/codec/language/Metaphone.java /tmp/Codec\_1f/src/java/org/apache/commons/codec/language/Metaphone.java

88c88

< return txt.toUpperCase() ;

---

> return txt.toUpperCase(java.util.Locale.ENGLISH) ;

diff -r /tmp/Codec\_1b/src/java/org/apache/commons/codec/language/SoundexUtils.java /tmp/Codec\_1f/src/java/org/apache/commons/codec/language/SoundexUtils.java

53c53

< return str.toUpperCase();

---

> return str.toUpperCase(java.util.Locale.ENGLISH);

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-1

--------------------------------------------------------------------------------

Revision ID (fixed version):

52d82d1dfff8c2b2ded9d843e0b03017af6d747c

--------------------------------------------------------------------------------

Revision date (fixed version):

2008-04-27 00:35:56 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-65

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-65

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.language.CaverphoneTest::testLocaleIndependence

--> junit.framework.ComparisonFailure: tr: expected:<[A]111111111> but was:<[1]111111111>

- org.apache.commons.codec.language.DoubleMetaphoneTest::testLocaleIndependence

--> junit.framework.ComparisonFailure: tr: expected:<[I]> but was:<[İ]>

- org.apache.commons.codec.language.MetaphoneTest::testLocaleIndependence

--> junit.framework.ComparisonFailure: tr: expected:<[I]> but was:<[İ]>

- org.apache.commons.codec.language.RefinedSoundexTest::testLocaleIndependence

--> junit.framework.AssertionFailedError: tr: 239

- org.apache.commons.codec.language.SoundexTest::testLocaleIndependence

--> junit.framework.AssertionFailedError: tr: The character is not mapped: İ

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.language.Caverphone

- org.apache.commons.codec.language.Metaphone

- org.apache.commons.codec.language.SoundexUtils

--------------------------------------------------------------------------------

**Bug2**

diff -r /tmp/Codec\_2b/src/java/org/apache/commons/codec/binary/Base64.java /tmp/Codec\_2f/src/java/org/apache/commons/codec/binary/Base64.java

446c446

< if (lineLength > 0) {

---

> if (lineLength > 0 && pos > 0) {

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-2

--------------------------------------------------------------------------------

Revision ID (fixed version):

b8c2d9d9dc9aab45f83cf49ac93cfa8546e4c08e

--------------------------------------------------------------------------------

Revision date (fixed version):

2009-07-13 22:33:28 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-77

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-77

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64InputStreamTest::testBase64EmptyInputStream

--> junit.framework.AssertionFailedError: EOF expected:<-1> but was:<13>

- org.apache.commons.codec.binary.Base64OutputStreamTest::testBase64EmptyOutputStream

--> junit.framework.AssertionFailedError: Streaming base64 encode

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64

--------------------------------------------------------------------------------

**Bug3**

diff -r /tmp/Codec\_3b/src/java/org/apache/commons/codec/language/DoubleMetaphone.java /tmp/Codec\_3f/src/java/org/apache/commons/codec/language/DoubleMetaphone.java

455c455

< } else if (contains(value, index + 1, 4, "IER")) {

---

> } else if (contains(value, index + 1, 3, "IER")) {

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-3

--------------------------------------------------------------------------------

Revision ID (fixed version):

a5dfe5cbc95d7f3ce0b4829756690c2cb8439f4c

--------------------------------------------------------------------------------

Revision date (fixed version):

2009-08-02 22:45:30 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-84

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-84

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.language.DoubleMetaphone2Test::testDoubleMetaphoneAlternate

--> junit.framework.ComparisonFailure: Test [19]=Angier expected:<AN[J]R> but was:<AN[K]R>

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.language.DoubleMetaphone

--------------------------------------------------------------------------------

**Bug4**

diff -r /tmp/Codec\_4b/src/java/org/apache/commons/codec/binary/Base64.java /tmp/Codec\_4f/src/java/org/apache/commons/codec/binary/Base64.java

225c225

< this(false);

---

> this(0);

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-4

--------------------------------------------------------------------------------

Revision ID (fixed version):

bcfef8906931a7929a6db14b9d82d387fd3f17f6

--------------------------------------------------------------------------------

Revision date (fixed version):

2010-03-27 00:02:22 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-89

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-89

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64Codec13Test::testEncoder

--> junit.framework.AssertionFailedError: Encoder test-1

- org.apache.commons.codec.binary.Base64Codec13Test::testBinaryEncoder

--> junit.framework.AssertionFailedError: BinaryEncoder test-1

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64

--------------------------------------------------------------------------------

**Bug5**

diff -r /tmp/Codec\_5b/src/java/org/apache/commons/codec/binary/Base64.java /tmp/Codec\_5f/src/java/org/apache/commons/codec/binary/Base64.java

585a586,588

> if (buffer == null || buffer.length - pos < decodeSize) {

> resizeBuffer();

> }

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-5

--------------------------------------------------------------------------------

Revision ID (fixed version):

800f0531068ebaf2f2d257bb1bd805781ddd4760

--------------------------------------------------------------------------------

Revision date (fixed version):

2010-06-01 21:52:33 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-98

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-98

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64InputStreamTest::testCodec98NPE

--> java.lang.NullPointerException

- org.apache.commons.codec.binary.Base64OutputStreamTest::testCodec98NPE

--> java.lang.NullPointerException

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64

--------------------------------------------------------------------------------

Bug6

diff -r /tmp/Codec\_6b/src/java/org/apache/commons/codec/binary/Base64InputStream.java /tmp/Codec\_6f/src/java/org/apache/commons/codec/binary/Base64InputStream.java

147a148

> int readLen = 0;

163a165

> while (readLen == 0) {

178c180,182

< return base64.readResults(b, offset, len);

---

> readLen = base64.readResults(b, offset, len);

> }

> return readLen;

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-6

--------------------------------------------------------------------------------

Revision ID (fixed version):

d74fc31604c805a47c44d7853f63a3b06ad6c016

--------------------------------------------------------------------------------

Revision date (fixed version):

2011-01-21 19:19:51 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-101

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-101

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64InputStreamTest::testCodec101

--> junit.framework.AssertionFailedError: Codec101: First read successful [c=0]

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64InputStream

--------------------------------------------------------------------------------

**Bug7**

diff -r /tmp/Codec\_7b/src/java/org/apache/commons/codec/binary/Base64.java /tmp/Codec\_7f/src/java/org/apache/commons/codec/binary/Base64.java

670c670

**< return StringUtils.newStringUtf8(encodeBase64(binaryData, true));**

**---**

**> return StringUtils.newStringUtf8(encodeBase64(binaryData, false));**

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-7

--------------------------------------------------------------------------------

Revision ID (fixed version):

954d995c5603b616c3c4a9ffb1823f36dd7ebcb0

--------------------------------------------------------------------------------

Revision date (fixed version):

2011-01-23 05:52:42 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-99

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-99

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64Test::testByteToStringVariations

--> junit.framework.ComparisonFailure: byteToString static Hello World expected:<SGVsbG8gV29ybGQ=[]> but was:<SGVsbG8gV29ybGQ=[

- org.apache.commons.codec.binary.Base64Test::testRfc4648Section10Encode

--> junit.framework.ComparisonFailure: expected:<Zg==[]> but was:<Zg==[

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64

--------------------------------------------------------------------------------

**Bug8**

diff -r /tmp/Codec\_8b/src/java/org/apache/commons/codec/binary/Base64.java /tmp/Codec\_8f/src/java/org/apache/commons/codec/binary/Base64.java

419,425d418

< void setInitialBuffer(byte[] out, int outPos, int outAvail) {

< if (out != null && out.length == outAvail) {

< buffer = out;

< pos = outPos;

< readPos = outPos;

< }

< }

diff -r /tmp/Codec\_8b/src/java/org/apache/commons/codec/binary/Base64InputStream.java /tmp/Codec\_8f/src/java/org/apache/commons/codec/binary/Base64InputStream.java

169,171d168

< if (c > 0 && b.length == len) {

< base64.setInitialBuffer(b, offset, len);

< }

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-8

--------------------------------------------------------------------------------

Revision ID (fixed version):

2c13032a16e37b7bc9e0272d25fc185b5c82b42e

--------------------------------------------------------------------------------

Revision date (fixed version):

2011-01-24 00:46:09 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-105

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-105

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64InputStreamTest::testCodec105

--> java.lang.ArrayIndexOutOfBoundsException: 2

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64

- org.apache.commons.codec.binary.Base64InputStream

--------------------------------------------------------------------------------

**Bug9**

diff -r /tmp/Codec\_9b/src/java/org/apache/commons/codec/binary/Base64.java /tmp/Codec\_9f/src/java/org/apache/commons/codec/binary/Base64.java

827c827

< long len = getEncodeLength(binaryData, MIME\_CHUNK\_SIZE, CHUNK\_SEPARATOR);

---

> long len = getEncodeLength(binaryData, isChunked ? MIME\_CHUNK\_SIZE : 0, CHUNK\_SEPARATOR);

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-9

--------------------------------------------------------------------------------

Revision ID (fixed version):

c5291ab1ca4d4d7a041874a66b1c05b7d5c02ccb

--------------------------------------------------------------------------------

Revision date (fixed version):

2011-01-26 23:40:25 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-112

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-112

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base64Test::testCodec112

--> java.lang.IllegalArgumentException: Input array too big, the output array would be bigger (6) than the specified maxium size of 4

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base64

--------------------------------------------------------------------------------

**Bug10**

diff -r /tmp/Codec\_10b/src/java/org/apache/commons/codec/language/Caverphone.java /tmp/Codec\_10f/src/java/org/apache/commons/codec/language/Caverphone.java

76c76

< txt = txt.replaceAll("^mb", "m2");

---

> txt = txt.replaceAll("mb$", "m2");

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-10

--------------------------------------------------------------------------------

Revision ID (fixed version):

41c68e9ef470696009d72133c7f05a20e2728e34

--------------------------------------------------------------------------------

Revision date (fixed version):

2011-03-01 17:56:14 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-117

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-117

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.language.CaverphoneTest::testEndMb

--> junit.framework.ComparisonFailure: Source: mbmb expected:<M[PM]1111111> but was:<M[MP]1111111>

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.language.Caverphone

--------------------------------------------------------------------------------

**Bug11**

diff -r /tmp/Codec\_11b/src/main/java/org/apache/commons/codec/net/QuotedPrintableCodec.java /tmp/Codec\_11f/src/main/java/org/apache/commons/codec/net/QuotedPrintableCodec.java

69a70

> private static final byte CR = 13;

70a72

> private static final byte LF = 10;

72a75

> private static final int SAFE\_LENGTH = 73;

114c117

< private static final void encodeQuotedPrintable(int b, ByteArrayOutputStream buffer) {

---

> private static final int encodeQuotedPrintable(int b, ByteArrayOutputStream buffer) {

119a123

> return 3;

131a136,142

> private static int getUnsignedOctet(final int index, final byte[] bytes) {

> int b = bytes[index];

> if (b < 0) {

> b = 256 + b;

> }

> return b;

> }

143a155,163

> private static int encodeByte(final int b, final boolean encode,

> final ByteArrayOutputStream buffer) {

> if (encode) {

> return encodeQuotedPrintable(b, buffer);

> } else {

> buffer.write(b);

> return 1;

> }

> }

151a172,174

> private static boolean isWhitespace(final int b) {

> return b == SPACE || b == TAB;

> }

174a198

> int pos = 1;

176a201,203

> for (int i = 0; i < bytes.length - 3; i++) {

> int b = getUnsignedOctet(i, bytes);

> if (pos < SAFE\_LENGTH) {

178,184c205

< for (byte c : bytes) {

< int b = c;

< if (b < 0) {

< b = 256 + b;

< }

< if (printable.get(b)) {

< buffer.write(b);

---

> pos += encodeByte(b, !printable.get(b), buffer);

186a208

> encodeByte(b, !printable.get(b) || isWhitespace(b), buffer);

189c211,214

< encodeQuotedPrintable(b, buffer);

---

> buffer.write(ESCAPE\_CHAR);

> buffer.write(CR);

> buffer.write(LF);

> pos = 1;

194a220,222

> int b = getUnsignedOctet(bytes.length - 3, bytes);

> boolean encode = !printable.get(b) || (isWhitespace(b) && pos > SAFE\_LENGTH - 5);

> pos += encodeByte(b, encode, buffer);

198a227,233

> if (pos > SAFE\_LENGTH - 2) {

> buffer.write(ESCAPE\_CHAR);

> buffer.write(CR);

> buffer.write(LF);

> }

> for (int i = bytes.length - 2; i < bytes.length; i++) {

> b = getUnsignedOctet(i, bytes);

199a235,237

> encode = !printable.get(b) || (i > bytes.length - 2 && isWhitespace(b));

> encodeByte(b, encode, buffer);

> }

229c267,270

< int u = Utils.digit16(bytes[++i]);

---

> if (bytes[++i] == CR) {

> continue;

> }

> int u = Utils.digit16(bytes[i]);

235c276

< } else {

---

> } else if (b != CR && b != LF) {

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-11

--------------------------------------------------------------------------------

Revision ID (fixed version):

7e8c20bd1b0593c3496a168be573aed40609a433

--------------------------------------------------------------------------------

Revision date (fixed version):

2012-03-07 15:34:01 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-121

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-121

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.net.QuotedPrintableCodecTest::testSkipNotEncodedCRLF

--> org.apache.commons.codec.DecoderException: Invalid URL encoding: not a valid digit (radix 16): 13

- org.apache.commons.codec.net.QuotedPrintableCodecTest::testSoftLineBreakDecode

--> org.apache.commons.codec.DecoderException: Invalid URL encoding: not a valid digit (radix 16): 13

- org.apache.commons.codec.net.QuotedPrintableCodecTest::testSoftLineBreakEncode

--> junit.framework.AssertionFailedError: expected:<...matics is the most b[=

- org.apache.commons.codec.net.QuotedPrintableCodecTest::testUltimateSoftBreak

--> junit.framework.AssertionFailedError: expected:<...There is no end to i[=

- org.apache.commons.codec.net.QuotedPrintableCodecTest::testTrailingSpecial

--> junit.framework.AssertionFailedError: expected:<... might contain sp=3D[=

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.net.QuotedPrintableCodec

--------------------------------------------------------------------------------

**Bug12**

diff -r /tmp/Codec\_12b/src/main/java/org/apache/commons/codec/binary/BaseNCodecInputStream.java /tmp/Codec\_12f/src/main/java/org/apache/commons/codec/binary/BaseNCodecInputStream.java

141a142,146

> @Override

> public long skip(long n) throws IOException {

> if (n < 0) {

> throw new IllegalArgumentException("Negative skip length");

> }

143a149,151

> final byte[] b = new byte[512];

> final int max = (int) Math.min(n, Integer.MAX\_VALUE);

> int total = 0;

144a153,166

> while (total < max) {

> int len = max - total;

> if (len > b.length) {

> len = b.length;

> }

> len = read(b, 0, len);

> if (len == EOF) {

> break;

> }

> total += len;

> }

>

> return total;

> }

151a174

> public int available() throws IOException {

157a181,186

> if (baseNCodec.eof) {

> return 0;

> } else {

> return 1;

> }

> }

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-12

--------------------------------------------------------------------------------

Revision ID (fixed version):

1fc451c79779bdbf788e8bfbec6dc031a47495d8

--------------------------------------------------------------------------------

Revision date (fixed version):

2012-03-19 18:33:42 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-130

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-130

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base32InputStreamTest::testSkipToEnd

--> junit.framework.AssertionFailedError: expected:<-1> but was:<183>

- org.apache.commons.codec.binary.Base32InputStreamTest::testCodec130

--> junit.framework.AssertionFailedError: expected:<[ello World]> but was:<[����

- org.apache.commons.codec.binary.Base32InputStreamTest::testSkipWrongArgument

--> junit.framework.AssertionFailedError: Expected exception: java.lang.IllegalArgumentException

- org.apache.commons.codec.binary.Base32InputStreamTest::testAvailable

--> junit.framework.AssertionFailedError: expected:<1> but was:<8>

- org.apache.commons.codec.binary.Base32InputStreamTest::testSkipPastEnd

--> junit.framework.AssertionFailedError: expected:<3> but was:<8>

- org.apache.commons.codec.binary.Base32InputStreamTest::testSkipBig

--> junit.framework.AssertionFailedError: expected:<3> but was:<8>

- org.apache.commons.codec.binary.Base64InputStreamTest::testSkipToEnd

--> junit.framework.AssertionFailedError: expected:<-1> but was:<255>

- org.apache.commons.codec.binary.Base64InputStreamTest::testCodec130

--> junit.framework.AssertionFailedError: expected:<[ello World]> but was:<[[��ܛ]>

- org.apache.commons.codec.binary.Base64InputStreamTest::testSkipWrongArgument

--> junit.framework.AssertionFailedError: Expected exception: java.lang.IllegalArgumentException

- org.apache.commons.codec.binary.Base64InputStreamTest::testAvailable

--> junit.framework.AssertionFailedError: expected:<1> but was:<8>

- org.apache.commons.codec.binary.Base64InputStreamTest::testSkipPastEnd

--> junit.framework.AssertionFailedError: expected:<6> but was:<8>

- org.apache.commons.codec.binary.Base64InputStreamTest::testSkipBig

--> junit.framework.AssertionFailedError: expected:<6> but was:<8>

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.BaseNCodecInputStream

--------------------------------------------------------------------------------

**Bug13**

Only in /tmp/Codec\_13f/src/main/java/org/apache/commons/codec/binary: CharSequenceUtils.java

diff -r /tmp/Codec\_13b/src/main/java/org/apache/commons/codec/binary/StringUtils.java /tmp/Codec\_13f/src/main/java/org/apache/commons/codec/binary/StringUtils.java

69a70,81

> public static boolean equals(final CharSequence cs1, final CharSequence cs2) {

> if (cs1 == cs2) {

> return true;

> }

> if (cs1 == null || cs2 == null) {

> return false;

> }

> if (cs1 instanceof String && cs2 instanceof String) {

> return cs1.equals(cs2);

> }

> return CharSequenceUtils.regionMatches(cs1, false, 0, cs2, 0, Math.max(cs1.length(), cs2.length()));

> }

diff -r /tmp/Codec\_13b/src/main/java/org/apache/commons/codec/language/DoubleMetaphone.java /tmp/Codec\_13f/src/main/java/org/apache/commons/codec/language/DoubleMetaphone.java

21a22

> import org.apache.commons.codec.binary.StringUtils;

244c245

< return doubleMetaphone(value1, alternate).equals(doubleMetaphone(value2, alternate));

---

> return StringUtils.equals(doubleMetaphone(value1, alternate), doubleMetaphone(value2, alternate));

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-13

--------------------------------------------------------------------------------

Revision ID (fixed version):

37ba197e62d6b60037d18afc33801e6221f1b8c6

--------------------------------------------------------------------------------

Revision date (fixed version):

2014-04-10 13:51:06 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-184

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-184

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.language.DoubleMetaphoneTest::testIsDoubleMetaphoneEqualBasic

--> java.lang.NullPointerException

- org.apache.commons.codec.language.DoubleMetaphoneTest::testCodec184

--> java.lang.NullPointerException

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.CharSequenceUtils

- org.apache.commons.codec.binary.StringUtils

- org.apache.commons.codec.language.DoubleMetaphone

--------------------------------------------------------------------------------

**Bug14**

diff -r /tmp/Codec\_14b/src/main/java/org/apache/commons/codec/language/bm/Lang.java /tmp/Codec\_14f/src/main/java/org/apache/commons/codec/language/bm/Lang.java

98c98

< private static final String LANGUAGE\_RULES\_RN = "org/apache/commons/codec/language/bm/lang.txt";

---

> private static final String LANGUAGE\_RULES\_RN = "org/apache/commons/codec/language/bm/%s\_lang.txt";

102c102

< Langs.put(s, loadFromResource(LANGUAGE\_RULES\_RN, Languages.getInstance(s)));

---

> Langs.put(s, loadFromResource(String.format(LANGUAGE\_RULES\_RN, s.getName()), Languages.getInstance(s)));

diff -r /tmp/Codec\_14b/src/main/java/org/apache/commons/codec/language/bm/PhoneticEngine.java /tmp/Codec\_14f/src/main/java/org/apache/commons/codec/language/bm/PhoneticEngine.java

31c31

< import java.util.TreeSet;

---

> import java.util.TreeMap;

338c338,339

< final Set<Rule.Phoneme> phonemes = new TreeSet<Rule.Phoneme>(Rule.Phoneme.COMPARATOR);

---

> final Map<Rule.Phoneme, Rule.Phoneme> phonemes =

> new TreeMap<Rule.Phoneme, Rule.Phoneme>(Rule.Phoneme.COMPARATOR);

361c362,370

< phonemes.addAll(subBuilder.getPhonemes());

---

> for (final Rule.Phoneme newPhoneme : subBuilder.getPhonemes()) {

> if (phonemes.containsKey(newPhoneme)) {

> final Rule.Phoneme oldPhoneme = phonemes.remove(newPhoneme);

> final Rule.Phoneme mergedPhoneme = oldPhoneme.mergeWithLanguage(newPhoneme.getLanguages());

> phonemes.put(mergedPhoneme, mergedPhoneme);

> } else {

> phonemes.put(newPhoneme, newPhoneme);

> }

> }

364c373

< return new PhonemeBuilder(phonemes);

---

> return new PhonemeBuilder(phonemes.keySet());

diff -r /tmp/Codec\_14b/src/main/java/org/apache/commons/codec/language/bm/Rule.java /tmp/Codec\_14f/src/main/java/org/apache/commons/codec/language/bm/Rule.java

157a158,160

> public Phoneme mergeWithLanguage(final LanguageSet lang) {

> return new Phoneme(this.phonemeText.toString(), this.languages.merge(lang));

> }

diff -r /tmp/Codec\_14b/src/main/resources/org/apache/commons/codec/language/bm/ash\_lang.txt /tmp/Codec\_14f/src/main/resources/org/apache/commons/codec/language/bm/ash\_lang.txt

0a1,206

> /\*

> \* Licensed to the Apache Software Foundation (ASF) under one or more

> \* contributor license agreements. See the NOTICE file distributed with

> \* this work for additional information regarding copyright ownership.

> \* The ASF licenses this file to You under the Apache License, Version 2.0

> \* (the "License"); you may not use this file except in compliance with

> \* the License. You may obtain a copy of the License at

> \*

> \* http://www.apache.org/licenses/LICENSE-2.0

> \*

> \* Unless required by applicable law or agreed to in writing, software

> \* distributed under the License is distributed on an "AS IS" BASIS,

> \* WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.

> \* See the License for the specific language governing permissions and

> \* limitations under the License.

> \*/

>

> // ASHKENAZIC

>

> // 1. following are rules to accept the language

> // 1.1 Special letter combinations

> zh polish+russian+german+english true

> eau french true

> [aoeiuäöü]h german true

> ^vogel german, true

> vogel$ german true

> witz german true

> tz$ german+russian+english true

> ^tz russian+english true

> güe spanish true

> güi spanish true

> ghe romanian true

> ghi romanian true

> vici$ romanian true

> schi$ romanian true

> chsch german true

> tsch german true

> ssch german true

> sch$ german+russian true

> ^sch german+russian true

> ^rz polish true

> rz$ polish+german true

> [^aoeiuäöü]rz polish true

> rz[^aoeiuäöü] polish true

> cki$ polish true

> ska$ polish true

> cka$ polish true

> ue german+russian true

> ae german+russian+english true

> oe german+french+russian+english true

> th$ german true

> ^th german true

> th[^aoeiu] german true

> mann german true

> cz polish true

> cy polish true

> niew polish true

> stein german true

> heim$ german true

> heimer$ german true

> ii$ russian true

> iy$ russian true

> yy$ russian true

> yi$ russian true

> yj$ russian true

> ij$ russian true

> gaus$ russian true

> gauz$ russian true

> gauz$ russian true

> goltz$ russian true

> gol'tz$ russian true

> golts$ russian true

> gol'ts$ russian true

> ^goltz russian true

> ^gol'tz russian true

> ^golts russian true

> ^gol'ts russian true

> gendler$ russian true

> gejmer$ russian true

> gejm$ russian true

> geimer$ russian true

> geim$ russian true

> geymer russian true

> geym$ russian true

> gof$ russian true

> thal german true

> zweig german true

> ck$ german+english true

> c$ polish+romanian+hungarian true

> sz polish+hungarian true

> gue spanish+french true

> gui spanish+french true

> guy french true

> cs$ hungarian true

> ^cs hungarian true

> dzs hungarian true

> zs$ hungarian true

> ^zs hungarian true

> ^wl polish true

> ^wr polish+english+german true

>

> gy$ hungarian true

> gy[aeou] hungarian true

> gy hungarian+russian true

> ly hungarian+russian+polish true

> ny hungarian+russian+polish true

> ty hungarian+russian+polish true

>

> // 1.2 special characters

> â romanian+french true

> ă romanian true

> à french true

> ä german true

> á hungarian+spanish true

> ą polish true

> ć polish true

> ç french true

> ę polish true

> é french+hungarian+spanish true

> è french true

> ê french true

> í hungarian+spanish true

> î romanian+french true

> ł polish true

> ń polish true

> ñ spanish true

> ó polish+hungarian+spanish true

> ö german+hungarian true

> õ hungarian true

> ş romanian true

> ś polish true

> ţ romanian true

> ü german+hungarian true

> ù french true

> ű hungarian true

> ú hungarian+spanish true

> ź polish true

> ż polish true

>

> ß german true

>

> // Every Cyrillic word has at least one Cyrillic vowel (аёеоиуыэюя)

> а cyrillic true

> ё cyrillic true

> о cyrillic true

> е cyrillic true

> и cyrillic true

> у cyrillic true

> ы cyrillic true

> э cyrillic true

> ю cyrillic true

> я cyrillic true

>

> // Hebrew

> א hebrew true

> ב hebrew true

> ג ebrew true

> ד hebrew true

> ה hebrew true

> ו hebrew true

> ז hebrew true

> ח hebrew true

> ט hebrew true

> י hebrew true

> כ hebrew true

> ל hebrew true

> מ hebrew true

> נ hebrew true

> ס hebrew true

> ע hebrew true

> פ hebrew true

> צ hebrew true

> ק hebrew true

> ר hebrew true

> ש hebrew true

> ת hebrew true

>

>

> // 2. following are rules to reject the language

> // Every Latin character word has at least one Latin vowel

> a cyrillic+hebrew false

> o cyrillic+hebrew false

> e cyrillic+hebrew false

> i cyrillic+hebrew false

> y cyrillic+hebrew+romanian false

> u cyrillic+hebrew false

>

> v[^aoeiuäüö] german false // in german "v" can be found before a vowel only

> y[^aoeiu] german false // in german "y" usually appears only in the last position; sometimes before a vowel

> c[^aohk] german false

> dzi german+english+french false

> ou german false

> aj german+english+french false

> ej german+english+french false

> oj german+english+french false

> uj german+english+french false

> k romanian false

> v polish false

> ky polish false

> eu russian+polish false

> w french+romanian+spanish+hungarian+russian false

> kie french+spanish false

> gie french+romanian+spanish false

> q hungarian+polish+russian+romanian false

> sch hungarian+polish+french+spanish false

> ^h russian false

diff -r /tmp/Codec\_14b/src/main/resources/org/apache/commons/codec/language/bm/gen\_lang.txt /tmp/Codec\_14f/src/main/resources/org/apache/commons/codec/language/bm/gen\_lang.txt

0a1,295

> /\*

> \* Licensed to the Apache Software Foundation (ASF) under one or more

> \* contributor license agreements. See the NOTICE file distributed with

> \* this work for additional information regarding copyright ownership.

> \* The ASF licenses this file to You under the Apache License, Version 2.0

> \* (the "License"); you may not use this file except in compliance with

> \* the License. You may obtain a copy of the License at

> \*

> \* http://www.apache.org/licenses/LICENSE-2.0

> \*

> \* Unless required by applicable law or agreed to in writing, software

> \* distributed under the License is distributed on an "AS IS" BASIS,

> \* WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.

> \* See the License for the specific language governing permissions and

> \* limitations under the License.

> \*/

>

> // GENERIC

>

> // 1. following are rules to accept the language

> // 1.1 Special letter combinations

> ^o’ english true

> ^o' english true

> ^mc english true

> ^fitz english true

> ceau french+romanian true

> eau romanian true

> eau$ french true // mp: I've added this

> eaux$ french true // mp: I've added this

> ault$ french true

> oult$ french true

> eux$ french true

> eix$ french true

> glou$ greeklatin true

> uu dutch true

> tx spanish true

> witz german true

> tz$ german+russian+english true

> ^tz russian+english true

> poulos$ greeklatin true

> pulos$ greeklatin true

> iou greeklatin true

> sj$ dutch true

> ^sj dutch true

> güe spanish true

> güi spanish true

> ghe romanian+greeklatin true

> ghi romanian+greeklatin true

> escu$ romanian true

> esco$ romanian true

> vici$ romanian true

> schi$ romanian true

> ii$ russian true

> iy$ russian true

> yy$ russian true

> yi$ russian true

> ^rz polish true

> rz$ polish+german true

> [bcdfgklmnpstwz]rz polish true

> rz[bcdfghklmnpstw] polish true

> cki$ polish true

> ska$ polish true

> cka$ polish true

> ae german+russian+english true

> oe german+french+russian+english+dutch true

> th$ german+english true

> ^th german+english+greeklatin true

> mann german true

> cz polish true

> cy polish+greeklatin true

> niew polish true

> etti$ italian true

> eti$ italian true

> ati$ italian true

> ato$ italian true

> [aoei]no$ italian true

> [aoei]ni$ italian true

> esi$ italian true

> oli$ italian true

> field$ english true

> stein german true

> heim$ german true

> heimer$ german true

> thal german true

> zweig german true

> [aeou]h german true

> äh german true

> öh german true

> üh german true

> [ln]h[ao]$ portuguese true

> [ln]h[aou] portuguese+french+german+dutch+czech+spanish+turkish true

> chsch german true

> tsch german true

> sch$ german+russian true

> ^sch german+russian true

> ck$ german+english true

> c$ polish+romanian+hungarian+czech+turkish true

> sz polish+hungarian true

> cs$ hungarian true

> ^cs hungarian true

> dzs hungarian true

> zs$ hungarian true

> ^zs hungarian true

> ^wl polish true

> ^wr polish+english+german+dutch true

>

> gy$ hungarian true

> gy[aeou] hungarian true

> gy hungarian+russian+french+greeklatin true

> guy french true

> gu[ei] spanish+french+portuguese true

> gu[ao] spanish+portuguese true

> gi[aou] italian+greeklatin true

>

> ly hungarian+russian+polish+greeklatin true

> ny hungarian+russian+polish+spanish+greeklatin true

> ty hungarian+russian+polish+greeklatin true

>

> // 1.2 special characters

> ć polish true

> ç french+spanish+portuguese+turkish true

> č czech true

> ď czech true

> ğ turkish true

> ł polish true

> ń polish true

> ñ spanish true

> ň czech true

> ř czech true

> ś polish true

> ş romanian+turkish true

> š czech true

> ţ romanian true

> ť czech true

> ź polish true

> ż polish true

>

> ß german true

>

> ä german true

> á hungarian+spanish+portuguese+czech+greeklatin true

> â romanian+french+portuguese true

> ă romanian true

> ą polish true

> à portuguese true

> ã portuguese true

> ę polish true

> é french+hungarian+czech+greeklatin true

> è french+spanish+italian true

> ê french true

> ě czech true

> ê french+portuguese true

> í hungarian+spanish+portuguese+czech+greeklatin true

> î romanian+french true

> ı turkish true

> ó polish+hungarian+spanish+italian+portuguese+czech+greeklatin true

> ö german+hungarian+turkish true

> ô french+portuguese true

> õ portuguese+hungarian true

> ò italian+spanish true

> ű hungarian true

> ú hungarian+spanish+portuguese+czech+greeklatin true

> ü german+hungarian+spanish+portuguese+turkish true

> ù french true

> ů czech true

> ý czech+greeklatin true

>

> // Every Cyrillic word has at least one Cyrillic vowel (аёеоиуыэюя)

> а cyrillic true

> ё cyrillic true

> о cyrillic true

> е cyrillic true

> и cyrillic true

> у cyrillic true

> ы cyrillic true

> э cyrillic true

> ю cyrillic true

> я cyrillic true

>

> // Every Greek word has at least one Greek vowel

> α greek true

> ε greek true

> η greek true

> ι greek true

> ο greek true

> υ greek true

> ω greek true

>

> // Arabic (only initial)

> ا arabic true // alif (isol + init)

> ب arabic true // ba'

> ت arabic true // ta'

> ث arabic true // tha'

> ج arabic true // jim

> ح arabic true // h.a'

> خ' arabic true // kha'

> د arabic true // dal (isol + init)

> ذ arabic true // dhal (isol + init)

> ر arabic true // ra' (isol + init)

> ز arabic true // za' (isol + init)

> س arabic true // sin

> ش arabic true // shin

> ص arabic true // s.ad

> ض arabic true // d.ad

> ط arabic true // t.a'

> ظ arabic true // z.a'

> ع arabic true // 'ayn

> غ arabic true // ghayn

> ف arabic true // fa'

> ق arabic true // qaf

> ك arabic true // kaf

> ل arabic true // lam

> م arabic true // mim

> ن arabic true // nun

> ه arabic true // ha'

> و arabic true // waw (isol + init)

> ي arabic true // ya'

>

> آ arabic true // alif madda

> إ arabic true // alif + diacritic

> أ arabic true // alif + hamza

> ؤ arabic true // waw + hamza

> ئ arabic true // ya' + hamza

> لا arabic true // ligature l+a

>

> // Hebrew

> א hebrew true

> ב hebrew true

> ג hebrew true

> ד hebrew true

> ה hebrew true

> ו hebrew true

> ז hebrew true

> ח hebrew true

> ט hebrew true

> י hebrew true

> כ hebrew true

> ל hebrew true

> מ hebrew true

> נ hebrew true

> ס hebrew true

> ע hebrew true

> פ hebrew true

> צ hebrew true

> ק hebrew true

> ר hebrew true

> ש hebrew true

> ת hebrew true

>

> // 2. following are rules to reject the language

>

> // Every Latin character word has at least one Latin vowel

> a cyrillic+hebrew+greek+arabic false

> o cyrillic+hebrew+greek+arabic false

> e cyrillic+hebrew+greek+arabic false

> i cyrillic+hebrew+greek+arabic false

> y cyrillic+hebrew+greek+arabic+romanian+dutch false

> u cyrillic+hebrew+greek+arabic false

>

> j italian false

> j[^aoeiuy] french+spanish+portuguese+greeklatin false

> g czech false

> k romanian+spanish+portuguese+french+italian false

> q hungarian+polish+russian+romanian+czech+dutch+turkish+greeklatin false

> v polish false

> w french+romanian+spanish+hungarian+russian+czech+turkish+greeklatin false

> x czech+hungarian+dutch+turkish false // polish excluded from the list

>

> dj spanish+turkish false

> v[^aoeiu] german false // in german, "v" can be found before a vowel only

> y[^aoeiu] german false // in german, "y" usually appears only in the last position; sometimes before a vowel

> c[^aohk] german false

> dzi german+english+french+turkish false

> ou german false

> a[eiou] turkish false // no diphthongs in Turkish

> ö[eaiou] turkish false

> ü[eaiou] turkish false

> e[aiou] turkish false

> i[aeou] turkish false

> o[aieu] turkish false

> u[aieo] turkish false

> aj german+english+french+dutch false

> ej german+english+french+dutch false

> oj german+english+french+dutch false

> uj german+english+french+dutch false

> eu russian+polish false

> ky polish false

> kie french+spanish+greeklatin false

> gie portuguese+romanian+spanish+greeklatin false

> ch[aou] italian false

> ch turkish false

> son$ german false

> sc[ei] french false

> sch hungarian+polish+french+spanish false

> ^h russian false

diff -r /tmp/Codec\_14b/src/main/resources/org/apache/commons/codec/language/bm/sep\_lang.txt /tmp/Codec\_14f/src/main/resources/org/apache/commons/codec/language/bm/sep\_lang.txt

0a1,105

> /\*

> \* Licensed to the Apache Software Foundation (ASF) under one or more

> \* contributor license agreements. See the NOTICE file distributed with

> \* this work for additional information regarding copyright ownership.

> \* The ASF licenses this file to You under the Apache License, Version 2.0

> \* (the "License"); you may not use this file except in compliance with

> \* the License. You may obtain a copy of the License at

> \*

> \* http://www.apache.org/licenses/LICENSE-2.0

> \*

> \* Unless required by applicable law or agreed to in writing, software

> \* distributed under the License is distributed on an "AS IS" BASIS,

> \* WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.

> \* See the License for the specific language governing permissions and

> \* limitations under the License.

> \*/

>

> // SEPHARDIC

>

> // 1. following are rules to accept the language

> // 1.1 Special letter combinations

> eau french true

> ou french true

> gni italian+french true

> tx spanish true

> tj spanish true

> gy french true

> guy french true

>

> sh spanish+portuguese true // English, but no sign for /sh/ in these languages

>

> lh portuguese true

> nh portuguese true

> ny spanish true

>

> gue spanish+french true

> gui spanish+french true

> gia italian true

> gie italian true

> gio italian true

> giu italian true

>

> // 1.2 special characters

> ñ spanish true

> â portuguese+french true

> á portuguese+spanish true

> à portuguese true

> ã portuguese true

> ê french+portuguese true

> í portuguese+spanish true

> î french true

> ô french+portuguese true

> õ portuguese true

> ò italian+spanish true

> ú portuguese+spanish true

> ù french true

> ü portuguese+spanish true

>

> // Hebrew

> א hebrew true

> ב hebrew true

> ג hebrew true

> ד hebrew true

> ה hebrew true

> ו hebrew true

> ז hebrew true

> ח hebrew true

> ט hebrew true

> י hebrew true

> כ hebrew true

> ל hebrew true

> מ hebrew true

> נ hebrew true

> ס hebrew true

> ע hebrew true

> פ hebrew true

> צ hebrew true

> ק hebrew true

> ר hebrew true

> ש hebrew true

> ת hebrew true

>

> // 2. following are rules to reject the language

>

> // Every Latin character word has at least one Latin vowel

> a hebrew false

> o hebrew false

> e hebrew false

> i hebrew false

> y hebrew false

> u hebrew false

>

> kh spanish false

> gua italian false

> guo italian false

> ç italian false

> cha italian false

> cho italian false

> chu italian false

> j italian false

> dj spanish false

> sce french false

> sci french false

> ó french false

> è portuguese false

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-14

--------------------------------------------------------------------------------

Revision ID (fixed version):

39d5df29fb768fd257f9d328b99f00bc69ec864a

--------------------------------------------------------------------------------

Revision date (fixed version):

2014-07-05 19:58:38 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-187

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-187

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.language.bm.PhoneticEngineRegressionTest::testCompatibilityWithOriginalVersion

--> junit.framework.AssertionFailedError: expected:<...dzn|bntsn|bnzn|vndzn[]> but was:<...dzn|bntsn|bnzn|vndzn[|vntsn]>

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.language.bm.Lang

- org.apache.commons.codec.language.bm.PhoneticEngine

- org.apache.commons.codec.language.bm.Rule

- src.main.resources.org.apache.commons.codec.language.bm.ash\_lang.txt

- src.main.resources.org.apache.commons.codec.language.bm.gen\_lang.txt

- src.main.resources.org.apache.commons.codec.language.bm.sep\_lang.txt

--------------------------------------------------------------------------------

**Bug15**

diff -r /tmp/Codec\_15b/src/main/java/org/apache/commons/codec/language/Soundex.java /tmp/Codec\_15f/src/main/java/org/apache/commons/codec/language/Soundex.java

188,192c188,190

< final char hwChar = str.charAt(index - 1);

< if ('H' == hwChar || 'W' == hwChar) {

< final char preHWChar = str.charAt(index - 2);

< final char firstCode = this.map(preHWChar);

< if (firstCode == mappedChar || 'H' == preHWChar || 'W' == preHWChar) {

---

> for (int i=index-1 ; i>=0 ; i--) {

> final char prevChar = str.charAt(i);

> if (this.map(prevChar)==mappedChar) {

193a192,194

> }

> if ('H'!=prevChar && 'W'!=prevChar) {

> break;

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-15

--------------------------------------------------------------------------------

Revision ID (fixed version):

28b7b958d39730fc278fd1919bf335c52133a1a2

--------------------------------------------------------------------------------

Revision date (fixed version):

2015-03-22 18:48:52 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-199

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-199

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.language.SoundexTest::testHWRuleEx1

--> junit.framework.AssertionFailedError: expected:<Y3[3]0> but was:<Y3[0]0>

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.language.Soundex

--------------------------------------------------------------------------------

**Bug16**

diff -r /tmp/Codec\_16b/src/main/java/org/apache/commons/codec/binary/Base32.java /tmp/Codec\_16f/src/main/java/org/apache/commons/codec/binary/Base32.java

99c99

< 25, 26, 27, 28, 29, 30, 31, 32,

---

> 25, 26, 27, 28, 29, 30, 31, // 50-57 P-V

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-16

--------------------------------------------------------------------------------

Revision ID (fixed version):

c82fe35c48bd0082c16644d16d82d1be79d6b9d1

--------------------------------------------------------------------------------

Revision date (fixed version):

2015-06-04 14:01:15 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-200

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-200

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.Base32Test::testCodec200

--> java.lang.IllegalArgumentException: pad must not be in alphabet or whitespace

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.Base32

--------------------------------------------------------------------------------

**Bug17**

diff -r /tmp/Codec\_17b/src/main/java/org/apache/commons/codec/binary/StringUtils.java /tmp/Codec\_17f/src/main/java/org/apache/commons/codec/binary/StringUtils.java

339c339

< return new String(bytes, Charsets.ISO\_8859\_1);

---

> return newString(bytes, Charsets.ISO\_8859\_1);

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-17

--------------------------------------------------------------------------------

Revision ID (fixed version):

d2f27093d7d95a07da901902f894d88b4ecc3e95

--------------------------------------------------------------------------------

Revision date (fixed version):

2017-03-26 17:22:08 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-229

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-229

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.StringUtilsTest::testNewStringNullInput\_CODEC229

--> java.lang.NullPointerException

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.StringUtils

--------------------------------------------------------------------------------

**Bug18**

diff -r /tmp/Codec\_18b/src/main/java/org/apache/commons/codec/binary/StringUtils.java /tmp/Codec\_18f/src/main/java/org/apache/commons/codec/binary/StringUtils.java

81c81

< return CharSequenceUtils.regionMatches(cs1, false, 0, cs2, 0, Math.max(cs1.length(), cs2.length()));

---

> return cs1.length() == cs2.length() && CharSequenceUtils.regionMatches(cs1, false, 0, cs2, 0, cs1.length());

Summary of configuration for Project: Codec

--------------------------------------------------------------------------------

Script dir: /defects4j/framework

Base dir: /defects4j

Major root: /defects4j/major

Repo dir: /defects4j/project\_repos

--------------------------------------------------------------------------------

Project ID: Codec

Program: commons-codec

Build file: /defects4j/framework/projects/Codec/Codec.build.xml

--------------------------------------------------------------------------------

Vcs: Vcs::Git

Repository: /defects4j/project\_repos/commons-codec.git

Commit db: /defects4j/framework/projects/Codec/active-bugs.csv

Number of bugs: 18

--------------------------------------------------------------------------------

Summary for Bug: Codec-18

--------------------------------------------------------------------------------

Revision ID (fixed version):

5ef5bd199a4e3df74479b4b363fa29b0b3092547

--------------------------------------------------------------------------------

Revision date (fixed version):

2017-03-26 21:43:36 +0000

--------------------------------------------------------------------------------

Bug report id:

CODEC-231

--------------------------------------------------------------------------------

Bug report url:

https://issues.apache.org/jira/browse/CODEC-231

--------------------------------------------------------------------------------

Root cause in triggering tests:

- org.apache.commons.codec.binary.StringUtilsTest::testEqualsCS1

--> java.lang.StringIndexOutOfBoundsException: String index out of range: 3

- org.apache.commons.codec.binary.StringUtilsTest::testEqualsCS2

--> java.lang.StringIndexOutOfBoundsException: String index out of range: 3

--------------------------------------------------------------------------------

List of modified sources:

- org.apache.commons.codec.binary.StringUtils

--------------------------------------------------------------------------------