Which Unit Test Generation Method is Better? A Comparison Through Different Mutation Operators

Beatriz Souza

6/4/2019

Installing and loading packages

```
install.packages("ggpubr")
```

```
## Installing package into '/home/beatriz/R/x86_64-pc-linux-gnu-library/3.4'
## (as 'lib' is unspecified)
```

```
install.packages("ggplot2")
```

```
## Installing package into '/home/beatriz/R/x86_64-pc-linux-gnu-library/3.4'
## (as 'lib' is unspecified)
```

```
install.packages("reshape")
```

```
## Installing package into '/home/beatriz/R/x86_64-pc-linux-gnu-library/3.4'
## (as 'lib' is unspecified)
```

```
install.packages("reshape2")
```

```
## Installing package into '/home/beatriz/R/x86_64-pc-linux-gnu-library/3.4'
## (as 'lib' is unspecified)
```

```
install.packages("dplyr")
```

```
## Installing package into '/home/beatriz/R/x86_64-pc-linux-gnu-library/3.4'
## (as 'lib' is unspecified)
```

```
library("ggplot2")
library("reshape")
library("reshape2")
```

```
##
## Attaching package: 'reshape2'
```

```
## The following objects are masked from 'package:reshape':
##
##     colsplit, melt, recast
```

```r
library("ggpubr")
```

```
## Loading required package: magrittr
```

```r
library("dplyr")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:reshape':
##
##     rename
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Reading data

```r
evoResults1 <- read.csv('/home/beatriz/Desktop/FinaldataSet/evoResultsEvoAvg.csv', sep = ",", header = TRUE)
evoResults2 <- read.csv('/home/beatriz/Desktop/FinaldataSet/evoResults2EvoAvg.csv', sep = ",", header = TRUE)
evoResults3 <- read.csv('/home/beatriz/Desktop/FinaldataSet/evoResults3EvoAvg.csv', sep = ",", header = TRUE)

randoopResults1 <- read.csv('/home/beatriz/Desktop/FinaldataSet/randoopResultsRandoopAvg.csv', sep = ",", header = TRUE)
randoopResults2 <- read.csv('/home/beatriz/Desktop/FinaldataSet/randoopResults2RandoopAvg.csv', sep = ",", header = TRUE)
randoopResults3 <- read.csv('/home/beatriz/Desktop/FinaldataSet/randoopResults3RandoopAvg.csv', sep = ",", header = TRUE)

manualResults1 <- read.csv('/home/beatriz/Desktop/FinaldataSet/manualSuitesResultsPITAvg.csv', sep = ",", header = TRUE)
manualResults2 <- read.csv('/home/beatriz/Desktop/FinaldataSet/manualSuitesResults2PITAvg.csv', sep = ",", header = TRUE)
manualResults3 <- read.csv('/home/beatriz/Desktop/FinaldataSet/manualSuitesResults3PITAvg.csv', sep = ",", header = TRUE)
```

Merging data

```
evoResults12 <- merge(evoResults1, evoResults2, all = TRUE)
#evo12
evoResults123 <- merge(evoResults12, evoResults3, all = TRUE)
#evoResults123

randoopResults12 <- merge(randoopResults1, randoopResults2, all = TRUE)
#randoopResults12
randoopResults123 <- merge(randoopResults12, randoopResults3, all = TRUE)
#randoopResults123

manualResults12 <- merge(manualResults1, manualResults2, all = TRUE)
#manualResults12
manualResults123 <- merge(manualResults12, manualResults3, all = TRUE)
#manualResults123

evoAndRandoopDataSet <- merge(evoResults123, randoopResults123, all = TRUE)
completeDataSet <- merge(evoAndRandoopDataSet, manualResults123, all = TRUE)
#completeDataSet
```

Declaring common values

```
metrics1 = c("LineCoverage","MutationCoverage")
metrics2 = c("ReturnValsMutator","ConditionalsBoundaryMutator",
             "VoidMethodCallMutator","IncrementsMutator",
             "NegateConditionalsMutator","MathMutator", "InvertNegsMutator")

projects=c("BCEL","CLI","Codec", "CSV", "Email", "FileUpload", "Imaging", "Lang", "St
atistics", "Validator")
```

Melting data

```
completeDataSet.m1 <- melt(completeDataSet, id.var = "Tool", measure.vars = metrics1)
#completeDataSet.m1

completeDataSet.m2 <- melt(completeDataSet, id.var = "Tool", measure.vars = metrics2)
#completeDataSet.m2

completeDataSet.m3 <- melt(completeDataSet, id.vars = c("Tool", "Project"))
#completeDataSet.m3
```
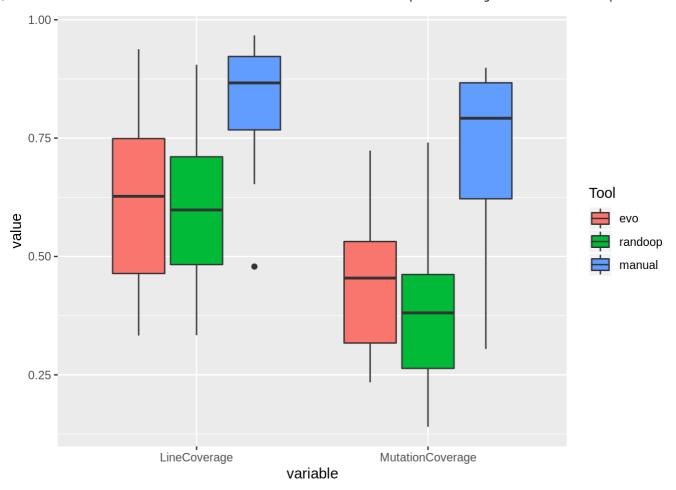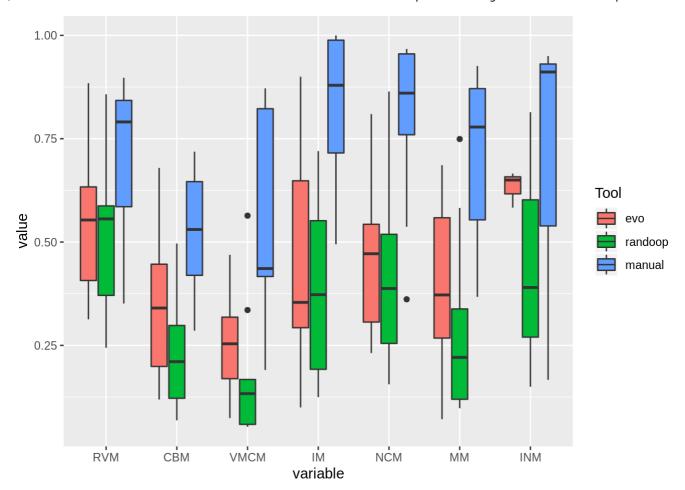
Ploting Box-plots

Line Coverage and Mutation Coverage

```
ggplot(data = completeDataSet.m1, aes(x=variable, y=value)) +
  geom_boxplot(aes(fill=Tool)) +
  scale_x_discrete(labels= c("LineCoverage", "MutationCoverage"))
```

PIT's Default Mutation Operators

```
ggplot(data = completeDataSet.m2, aes(x=variable, y=value)) +
  geom_boxplot(aes(fill=Tool)) +
  scale_x_discrete(labels= c(
             "RVM","CBM",
             "VMCM","IM",
             "NCM","MM", "INM"))
```

```
## Warning: Removed 24 rows containing non-finite values (stat_boxplot).
```

Line Coverage

Comparing Line Coverage per Tool

```
# If paired = FALSE, you perform a Mann-Whitney test
wilcox.test(manualResults123$LineCoverage, evoResults123$LineCoverage, paired = TRUE,
alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$LineCoverage and evoResults123$LineCoverage
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$LineCoverage, randoopResults123$LineCoverage, paired = T
RUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$LineCoverage and randoopResults123$LineCoverage
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$LineCoverage, randoopResults123$LineCoverage, paired = TRUE
, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$LineCoverage and randoopResults123$LineCoverage
## V = 32, p-value = 0.3477
## alternative hypothesis: true location shift is greater than 0
```

Mutation Coverage

Comparing Mutation Coverage per Tool

```
wilcox.test(manualResults123$MutationCoverage, evoResults123$MutationCoverage,  paire
d = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$MutationCoverage and evoResults123$MutationCoverage
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$MutationCoverage, randoopResults123$MutationCoverage,  p
aired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$MutationCoverage and randoopResults123$MutationCoverage
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$MutationCoverage, randoopResults123$MutationCoverage, paire
d = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$MutationCoverage and randoopResults123$MutationCoverage
## V = 53, p-value = 0.00293
## alternative hypothesis: true location shift is greater than 0
```

Mutation Operators

Comparing ReturnValsMutator Coverage per Tool

```
wilcox.test(manualResults123$ReturnValsMutator, evoResults123$ReturnValsMutator, pair
ed = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$ReturnValsMutator and evoResults123$ReturnValsMutator
## V = 54, p-value = 0.001953
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$ReturnValsMutator, randoopResults123$ReturnValsMutator,
paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$ReturnValsMutator and randoopResults123$ReturnValsMutator
## V = 54, p-value = 0.001953
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$ReturnValsMutator, randoopResults123$ReturnValsMutator, pai
red = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$ReturnValsMutator and randoopResults123$ReturnValsMutator
## V = 38, p-value = 0.1611
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$ReturnValsMutator, randoopResults123$ReturnValsMutator, pai
red = TRUE)
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$ReturnValsMutator and randoopResults123$ReturnValsMutator
## V = 38, p-value = 0.3223
## alternative hypothesis: true location shift is not equal to 0
```

Comparing ConditionalsBoundaryMutator Coverage per Tool

```
wilcox.test(manualResults123$ConditionalsBoundaryMutator, evoResults123$ConditionalsB
oundaryMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$ConditionalsBoundaryMutator and evoResults123$Conditionals
BoundaryMutator
## V = 53, p-value = 0.00293
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$ConditionalsBoundaryMutator, randoopResults123$Condition
alsBoundaryMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$ConditionalsBoundaryMutator and randoopResults123$Conditio
nalsBoundaryMutator
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$ConditionalsBoundaryMutator, randoopResults123$Conditionals
BoundaryMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$ConditionalsBoundaryMutator and randoopResults123$Conditional
sBoundaryMutator
## V = 53, p-value = 0.00293
## alternative hypothesis: true location shift is greater than 0
```

Comparing VoidMethodCallMutator Coverage per Tool

```
wilcox.test(manualResults123$VoidMethodCallMutator, evoResults123$VoidMethodCallMutat
or, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$VoidMethodCallMutator and evoResults123$VoidMethodCallMuta
tor
## V = 36, p-value = 0.003906
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$VoidMethodCallMutator, randoopResults123$VoidMethodCallM
utator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$VoidMethodCallMutator and randoopResults123$VoidMethodCall
Mutator
## V = 36, p-value = 0.003906
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$VoidMethodCallMutator, randoopResults123$VoidMethodCallMuta
tor, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$VoidMethodCallMutator and randoopResults123$VoidMethodCallMut
ator
## V = 35, p-value = 0.08203
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(randoopResults123$VoidMethodCallMutator, evoResults123$VoidMethodCallMuta
tor, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  randoopResults123$VoidMethodCallMutator and evoResults123$VoidMethodCallMut
ator
## V = 10, p-value = 0.9355
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(randoopResults123$VoidMethodCallMutator, evoResults123$VoidMethodCallMuta
tor, paired = TRUE)
```

```
##
##  Wilcoxon signed rank test
##
## data:  randoopResults123$VoidMethodCallMutator and evoResults123$VoidMethodCallMut
ator
## V = 10, p-value = 0.1641
## alternative hypothesis: true location shift is not equal to 0
```

Comparing IncrementsMutator Coverage per Tool

```
wilcox.test(manualResults123$IncrementsMutator, evoResults123$IncrementsMutator, pair
ed = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$IncrementsMutator and evoResults123$IncrementsMutator
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$IncrementsMutator, randoopResults123$IncrementsMutator,
paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$IncrementsMutator and randoopResults123$IncrementsMutator
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$IncrementsMutator, randoopResults123$IncrementsMutator, pai
red = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$IncrementsMutator and randoopResults123$IncrementsMutator
## V = 33, p-value = 0.3125
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$IncrementsMutator, randoopResults123$IncrementsMutator, pai
red = TRUE)
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$IncrementsMutator and randoopResults123$IncrementsMutator
## V = 33, p-value = 0.625
## alternative hypothesis: true location shift is not equal to 0
```

Comparing NegateConditionalsMutator Coverage per Tool

```
wilcox.test(manualResults123$NegateConditionalsMutator, evoResults123$NegateCondition
alsMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$NegateConditionalsMutator and evoResults123$NegateConditio
nalsMutator
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$NegateConditionalsMutator, randoopResults123$NegateCondi
tionalsMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$NegateConditionalsMutator and randoopResults123$NegateCond
itionalsMutator
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$NegateConditionalsMutator, randoopResults123$NegateConditio
nalsMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$NegateConditionalsMutator and randoopResults123$NegateConditi
onalsMutator
## V = 48, p-value = 0.01855
## alternative hypothesis: true location shift is greater than 0
```

Comparing MathMutator Coverage per Tool

```
wilcox.test(manualResults123$MathMutator, evoResults123$MathMutator, paired = TRUE, a
lternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$MathMutator and evoResults123$MathMutator
## V = 54, p-value = 0.001953
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123$MathMutator, randoopResults123$MathMutator, paired = TRU
E, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123$MathMutator and randoopResults123$MathMutator
## V = 55, p-value = 0.0009766
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123$MathMutator, randoopResults123$MathMutator, paired = TRUE,
  alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123$MathMutator and randoopResults123$MathMutator
## V = 48, p-value = 0.01855
## alternative hypothesis: true location shift is greater than 0
```

Comparing InvertNegsMutator Coverage per Tool

```
evoResults123WithoutNA <- evoResults123 %>% select(InvertNegsMutator) %>% filter(!is.
na(InvertNegsMutator))
#evoResults123WithoutNA
randoopResults123WithoutNA <- randoopResults123 %>% select(InvertNegsMutator) %>% fil
ter(!is.na(InvertNegsMutator))
#randoopResults123WithoutNA
manualResults123WithoutNA <- manualResults123 %>% select(InvertNegsMutator) %>% filte
r(!is.na(InvertNegsMutator))
#manualResults123WithoutNA

wilcox.test(manualResults123WithoutNA$InvertNegsMutator, evoResults123WithoutNA$Inver
tNegsMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123WithoutNA$InvertNegsMutator and evoResults123WithoutNA$Inve
rtNegsMutator
## V = 3, p-value = 0.625
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123WithoutNA$InvertNegsMutator, randoopResults123WithoutNA$I
nvertNegsMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123WithoutNA$InvertNegsMutator and randoopResults123WithoutNA
$InvertNegsMutator
## V = 6, p-value = 0.125
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(evoResults123WithoutNA$InvertNegsMutator, randoopResults123WithoutNA$Inve
rtNegsMutator, paired = TRUE, alternative = "greater")
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123WithoutNA$InvertNegsMutator and randoopResults123WithoutNA$Inv
ertNegsMutator
## V = 5, p-value = 0.25
## alternative hypothesis: true location shift is greater than 0
```

```
wilcox.test(manualResults123WithoutNA$InvertNegsMutator, evoResults123WithoutNA$Inver
tNegsMutator, paired = TRUE)
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123WithoutNA$InvertNegsMutator and evoResults123WithoutNA$Inve
rtNegsMutator
## V = 3, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
```

```
wilcox.test(manualResults123WithoutNA$InvertNegsMutator, randoopResults123WithoutNA$I
nvertNegsMutator, paired = TRUE)
```

```
##
##  Wilcoxon signed rank test
##
## data:  manualResults123WithoutNA$InvertNegsMutator and randoopResults123WithoutNA
$InvertNegsMutator
## V = 6, p-value = 0.25
## alternative hypothesis: true location shift is not equal to 0
```

```
wilcox.test(evoResults123WithoutNA$InvertNegsMutator, randoopResults123WithoutNA$Inve
rtNegsMutator, paired = TRUE)
```

```
##
##  Wilcoxon signed rank test
##
## data:  evoResults123WithoutNA$InvertNegsMutator and randoopResults123WithoutNA$Inv
ertNegsMutator
## V = 5, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
```