

An example of experimental analysis of SLS algorithms

Andrea Roli

DEIS

Alma Mater Studiorum Università di Bologna – Italia
andrea.roli@unibo.it

Experimental setting

- Problem: N-queens
- Objective function: number of violations = $\sum_{i=1}^N \text{violations}(q_i)$
- Datafiles: see file data.zip. For each algorithm, the data file composed of 3 columns and n_{runs} rows. Each row contains: iteration to best, time to best, best solution.
- Tool: R (*The R Project for Statistical Computing*, www.r-project.org)

Loading files with R

```
data_bi <- read.table("data-queens-best_improvement.txt")  
data_si <- read.table("data-queens-stochastic_first_improvement.txt")  
data_mm <- read.table("data-queens-maxminconflict.txt")  
data_ts <- read.table("data-queens-tabusearch.txt")
```

Success ratio

Select only the rows corresponding to a success:

```
data_bi_success <- data_bi[data_bi[,3] == 0,]  
data_si_success <- data_si[data_si[,3] == 0,]  
data_mm_success <- data_mm[data_mm[,3] == 0,]  
data_ts_success <- data_ts[data_ts[,3] == 0,]
```

Success ratio of algorithm $\mathcal{A}$ is the size of the respective success array. With R:

```
bi_success <- dim(data_bi_success)[1]  
si_success <- dim(data_si_success)[1]  
mm_success <- dim(data_mm_success)[1]  
ts_success <- dim(data_ts_success)[1]
```

Summary of main statistics

```
summary(data_ts_success[,2])  
summary(data_mm_success[,2])
```

More information about distributions

Histograms:

```
hist(data_ts_success[,2])  
hist(data_mm_success[,2])
```

Boxplots:

```
boxplot(data_ts_success[,2], data_mm_success[,2])
```

Run-time and run-length distributions

The *run-time distribution* provides an estimation of the probability of finding a solution within a given time t .

We plot the estimated cumulative distribution of the run-time data:

```
plot(ecdf(data_ts_success[,2]),xlim=c(0,400),col.points=2,  
      main="run-time distribution",  
      xlab="time (sec.)",ylab="rtd")  
plot(ecdf(data_mm_success[,2]),col.points=3,add=TRUE)
```

A statistical test

To assess the significance of the results observed, one should apply a statistical test.

Example: MM better than TS if execution time is concerned.

Wilcoxon test (a non-parametric test):

```
wilcox.test(data_mm_success[,2], data_ts_success[,2],  
            alternative="less", conf.level=0.95)
```

Result: $p\text{-value} = 1.195e-11 < 0.05$, therefore MM is better than TS with confidence 95%