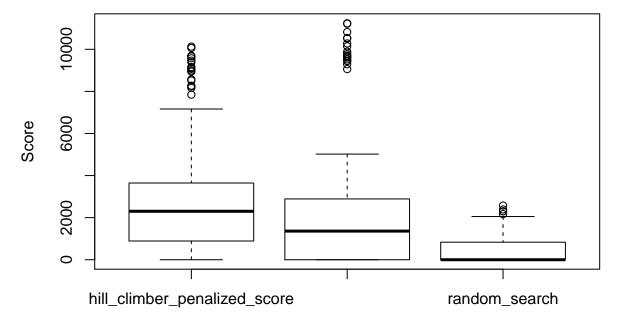
Knapsack experiments

November 2, 2016

When we remove negative values, it's easier to compare the penalized and cliff scores. Here we see that while random-restart-hill-climber remains the method with the most variation, it's upper values are now outliers. Regular hill climber's middle 50% of values are higher on average than random-restart's. Regular hill-climber could arguably be better because it returns zero with lower frequency. Random-search with cliff scoring has the lowest values on average, and returns zero with the most frequency.



Searcher

We will use a wilcox test to check our "eyeball statistics." The wilcox test shows that there is a lot of variation between the different methods. Significantly, regular hill-climber outperforms random-restart hill climber (this is likely due to the low max answers value, were we to conduct these tests with a greater value for maximum answers random restart hill climber would have a better competitive advantage to hill climber). Random-search performs significantly worse than both regular and random restart hill climber. The difference between random-search and the hill climbers is much more significant than the difference between the hill climbers.

```
pairwise.wilcox.test(dataSet$Non_negative_score, dataSet$Search_method)
```

```
##
## Pairwise comparisons using Wilcoxon rank sum test
##
```

```
## data: dataSet$Non_negative_score and dataSet$Search_method
##
## hill_climber_penalized_score
## random_restart_hill_climber_penalized_score 0.0017
## random_search <2e-16
## random_restart_hill_climber_penalized_score
## random_restart_hill_climber_penalized_score
## random_search <2e-16
##
## P value adjustment method: holm</pre>
```