# Configuration Description

## Genetic Algorithm

- **Population size:** The GA used a population size of 50
- **Selection:** Tournament Selection was used with a selection size of 3
- **Crossover:** Single-point crossover
- **Mutation Rate:** 0.05
- **Termination Criteria:** Maximum of 100 generations

## Genetic Algorithm with ILS

- **Population size:** The GA used a population size of 50
- **Selection:** Tournament Selection was used with a selection size of 3
- **Crossover:** Single-point crossover
- **Mutation Rate:** 0.05
- **ILS:** Offspring were subjected to 100 iterations
- **Perturbation:** Random bit-flipping up to half of the given chromosome
- **HillClimb:** Neighbourhood based on single-item flips
- **Termination Criteria:** Maximum of 100 generations

# Local Search and Justification

Iterated Local Search was introduced with a hill climb algorithm as the local search. The neighborhood was defined by the set of solutions reachable based on flipping a single item. The hill climb algorithm iteratively moved over neighbors and moved to any solution with a higher fitness. The algorithm then terminated when no further improvement could be found.

Hill climb was selected for its simple implementation and the knapsack neighborhood structure made it a suitable choice for the local search. However, there was a problem with the hill climb that produced an infinite loop. A penalty factor in the calculateFitness was introduced to prevent the local search to choose any better fitness of a solution.

# Experimental Setup

**GA**

- **Initial Solution Generation:** Random generation of chromosomes (binary vectors, 0 = item excluded, 1 = item included).
- **Selection Method:** Tournament selection (n = 3).
- **Crossover Operator:** Single-point crossover.
- **Mutation Operator:** Bit flipping with a 0.05 mutation probability.
- **Termination Criteria:** Maximum number of 100 generations.

**GA + ILS**

- **Initial Solution Generation:** Random generation of chromosomes (binary vectors, 0 = item excluded, 1 = item included).
- **Selection Method:** Tournament selection (n = 3).
- **Crossover Operator:** Single-point crossover.
- **Mutation Operator:** Bit flipping with a 0.05 mutation probability.
- **Local Search Algorithm:** Hill climb with neighborhood based on flipping single items.
- **Perturbation:** Randomly flipping up to 50% of the bits in a chromosome.
- **ILS Iterations per Offspring:** 100
- **Termination Criteria:** Maximum of 100 generations.

# Results Table

| Problem Instance | Algorithm | Seed | Best Solution | Known Optimum | Runtime (seconds) |
|---|---|---|---|---|---|
| f1_l-d_kp_10_269 | GA | 46130034 | 295 | 295 | 0.001705 |
| | GA + ILS | 46130034 | 295 | 295 | 2.9779 |
| f2_l-d_kp_20_878 | GA | 46130034 | 922 | 1024 | 0.00162 |
| | GA + ILS | 46130034 | 978 | 1024 | 3.71943 |
| f3_l-d_kp_4_20 | GA | 46130034 | 35 | 35 | 0.000951 |
| | GA + ILS | 46130034 | 35 | 35 | 2.63213 |
| f4_l-d_kp_4_11 | GA | 46130034 | 23 | 23 | 0.000942 |
| | GA + ILS | 46130034 | 23 | 23 | 2.62764 |
| f5_l-d_kp_15_375 | GA | 46130034 | 557 | 4810694 | 0.001361 |
| | GA + ILS | 46130034 | 557 | 4810694 | 3.11308 |
| f6_l-d_kp_10_60 | GA | 46130034 | 52 | 52 | 0.001172 |
| | GA + ILS | 46130034 | 52 | 52 | 2.91718 |
| f7_l-d_kp_7_50 | GA | 46130034 | 107 | 107 | 0.001018 |
| | GA + ILS | 46130034 | 107 | 107 | 2.73772 |
| f8_l-d_kp_23_10000 | GA | 46130034 | 9757 | 9767 | 0.00165 |
| | GA + ILS | 46130034 | 9754 | 9767 | 3.70868 |
| f9_l-d_kp_5_80 | GA | 46130034 | 130 | 130 | 0.001016 |
| | GA + ILS | 46130034 | 130 | 130 | 2.67236 |
| f10_l-d_kp_20_879 | GA | 46130034 | 942 | 1025 | 0.001559 |
| | GA + ILS | 46130034 | 920 | 1025 | 3.64463 |
| knapPI_1_100_1000_1 | GA | 46130034 | 0 | 9147 | 0.004455 |
| | GA + ILS | 46130034 | 0 | 9147 | 10.568 |

# Test Analysis

**Null Hypothesis:** The average best solution found by GA is equal to the average best solution found by GA+ILS.
**Alternative Hypothesis:** The average best solution found by GA is less than the average best solution found by GA+ILS.

| Algorithm | Sample Best Mean (Runtime) | Sample Standard Deviation (Runtime) | Sample Size |
|---|---|---|---|
| GA | 0.00145 | 0.00043 | 11 |
| GA + ILS | 3.293 | 0.540 | 11 |

$$s_{p^2} = ((11 - 1) * 0.00043^2 + (11 - 1) * 0.540^2) / (11 + 11 - 2) = 0.147739$$

$$s_p = \sqrt{0.147739} = 0.384$$

$$z = (0.00145 - 3.293) / (0.384 * \sqrt{(1/11 + 1/11)}) = -18.99$$

$$\alpha = 0.05$$

$$z_{critical} = 1.645$$

Since the z-score (-18.99) is less than the critical value (1.645), we reject the null hypothesis.

**Conclusion**

The conclusion of this test is that the mean runtime performance between the GA and GA + ILS is significantly different. GA + ILS has a much higher runtime than the standard GA, therefore the null hypothesis is rejected.

# Critical Analysis

In some cases, the GA provided better solutions than the GA + ILS for certain instances. The inverse of this is also true. The results supporting this can be found in the Results Table above. Therefore, the substantial increase in runtime for GA + ILS coupled with the fact that it offers none to marginally better improvement on solution, raises the question if GA + ILS is practically meaningful. In the case of runtime and performance being prioritized, the GA would be a more suitable choice. If performance prioritization is not an issue, GA + ILS will be a suitable fit. This can ensure the solution space and population is properly searched, which can lead to better results.