

Landscape Properties of the 0-1 Knapsack Problem

Khulood Alyahya
School of Computer Science
University of Birmingham, B15 2TT, UK
kya020@cs.bham.ac.uk

Jonathan E. Rowe
School of Computer Science
University of Birmingham, B15 2TT, UK
J.E.Rowe@cs.bham.ac.uk

ABSTRACT

This paper studies two landscapes of different instances of the 0-1 knapsack problem. The instances are generated randomly from varied weight distributions. We show that the variation of the weights can be used to guide the selection of the most suitable local search operator for a given instance.

Categories and Subject Descriptors

F.2.0 [Theory of Computation]: Analysis of algorithms and problem complexity.

Keywords

Fitness Landscapes, Instance Difficulty, Constraint Handling

1. THE 0-1 KNAPSACK PROBLEM

The classical 0-1 knapsack problem (0-1 KP) is defined as follows: Given a knapsack of capacity C and a set of n items each with associated weight w_i and profit p_i , the aim is to find a subset of items that maximises $f(x) = \sum_{i=1}^n x_i p_i$, subject to $\sum_{i=1}^n x_i w_i \leq C$, where $x \in \{0, 1\}^n$, $C = \lambda \sum_{i=1}^n w_i$, and $0 \leq \lambda \leq 1$. The binary vector $x = (x_1, \dots, x_n)$ represents the decision variable where $x_i = 1$ when item i belongs to the subset and $x_i = 0$ otherwise. To avoid trivial instances, we only study instances where $w_i \leq C$ for all $i = 1, \dots, n$, and $\lambda < 1$.

We generated instances randomly with different correlation between the fitness function coefficient p_i and constraint coefficient w_i as in [4]. Based on the value of the correlation coefficient, an instance is classified as *uncorrelated*, *weakly correlated*, *strongly correlated*, *inverse strongly correlated*, or *subset sum*. For each instance type we varied the tightness of the capacity constraint by setting λ value between 0.1 to 0.9 with 0.1 step interval. Also, for each instance type, weights are drawn randomly from five different discrete probability distributions: uniform, normal, negatively skewed, positively skewed and bimodal distribution with peaks at both ends. The variation of the weights is then measured by their

coefficient of variation ($CV = \sigma/\mu$). Ten random instances were generated for each settings combination.

An infeasible solution x that violates the given constraint is penalised by a value $\text{Pen}(x) > 0$, while $\text{Pen}(x) = 0$ for a feasible solution x . The fitness functions after adding the penalty term is $f(x) = \sum_{i=1}^n x_i p_i - \text{Pen}(x)$. In this paper, we use the linear penalty function from [3] where the value of the penalty grows linearly with respect to the degree of constraint violation. We also add the term $\sum_{i=1}^n p_i$ to the penalty function as an offset term that insures that all infeasible solutions achieve lower fitness values than all feasible solutions [2]. The penalty after adding the offset term:

$$\text{Pen}(x) = \rho \left(\sum_{i=1}^n x_i w_i - C \right) + \sum_{i=1}^n p_i \quad (1)$$

where $\rho = \max \{p_i/w_i \mid i = 1, \dots, n\}$. We use this penalty function with all instance types except for *subset sum*. Since applying this penalty function to infeasible solutions in a *subset sum* instance assigns equal fitness values for all infeasible solution, thus, creating large plateaus in the landscape. We therefore choose this simple penalty function to use with *subset sum* instances:

$$\text{Pen}(x) = \sum_{i=1}^n x_i w_i - C \quad (2)$$

2. LANDSCAPE OF THE 0-1 KP

Landscape analysis allows us to study an optimisation problem in connection with a neighbourhood structure defined over the search space. The aim of such analysis is to lead to a better understanding of the optimisation problem structure. The gained insights could then be used to guide the design or selection of the best search operators or parameter values to use. In this paper we study two landscapes of the 0-1 KP induced by the Hamming 1 ($H1$) and Hamming 1+2 ($H1+2$) neighbourhood operators. The $H1$ operator is the simple 1-bit flip mutation operator, the neighbourhood size is thus $|N(x)| = n$. The neighbourhood of the $H1+2$ operator includes the Hamming one neighbours plus the Hamming two neighbours of the current solution. The neighbourhood size for this operator is $|N(x)| = n + (n(n-1)/2)$. Only small problem sizes were considered to allow exhaustive enumeration of the entire search space ($n = 14, 18$).

2.1 Number of Optima

A very large difference in the number of local optima was found between the two landscape. The $H1+2$ landscape has

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '15 July 11-15, 2015, Madrid, Spain

© 2015 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-3488-4/15/07.

DOI: <http://dx.doi.org/10.1145/2739482.2764688>

less number of local optima in all instance types and all values of CV and λ , with the largest difference being found in instances with small CV values. The number of optima in the $H1$ landscape was found to be highly and negatively correlated with the CV of the weights in the middle region between very high and very low constraints ($0.3 \leq \lambda \leq 0.7$) as depicted in Figure 1. In the $H1+2$ landscape, only *inverse strongly correlated* instances were found to have a strong correlation between the optima number and the CV with high and positive correlation across all constraint levels. The correlation in the *subset sum* and *strongly correlated* instances is strong and positive but only in the very constrained and very low constrained regions ($\lambda \leq 0.2$, $\lambda \geq 0.8$).

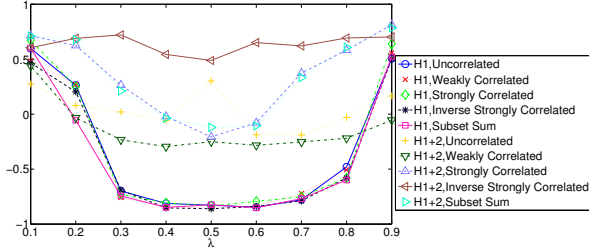


Figure 1: Correlation between the CV of the weights and the number of local optima versus λ ($n = 18$).

2.2 Basins of Attraction

We define the basin of attraction of an optimum as the set of points that leads to it after applying steepest ascent to them. The neighbours of a point are evaluated in an increasing order and the first best improving neighbour is always selected. The correlation between the optimum basin size and fitness has been conjectured to be related to problem difficulty. Previous studies have shown that landscapes, where fitter optima have larger basins, tend to be easier to search [5]. Figure 2 shows a very low positive and even negative correlation between the basin size and fitness in *inverse strongly correlated* and *subset sum* instances. This indicates that fitter optima do not necessarily have larger basins in these instances and even worse bad optima tend to have larger basins (when the correlation is negative). This could mean that these instances are difficult for local search.

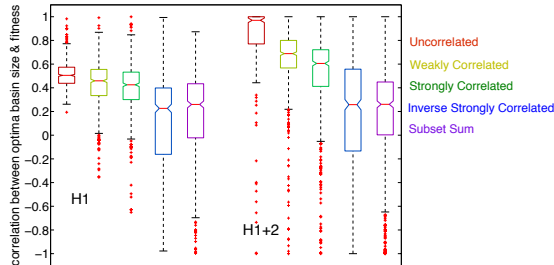


Figure 2: Correlation between basin size and fitness in the two landscapes ($n = 14$).

2.3 Local Search

We compare here the cost of finding the optimal value using a local search algorithm based on the underlying neigh-

bourhood of each landscape. Steepest ascent algorithm with random restart was run 30 times per instance and with each of the $H1$ and $H1+2$ neighbourhood operators. The cost is then calculated using the number of used fitness evaluations until the optimal value was found. *Inverse strongly correlated* and *subset sum* instances had the highest average cost of finding the optimal solution while *uncorrelated* instances had the lowest average cost. Results from section 2.2 suggests that the high cost in *inverse strongly correlated* and *subset sum* instances is due to steepest ascent being attracted to the large basins of local optima and having to perform many restarts until it could find the small basin of the optimal or one of the optimal solutions.

There is a trade-off between the number of fitness evaluations needed to explore the neighbourhood and the difference between the number of local optima between the two landscapes. When the difference is very large, it is better to use the $H1+2$ operator even though the size of the neighbourhood is much larger as Figure 3 shows ($CV < 0.3$). When the difference is moderate, it is better to use the $H1$ operator (in most instance types) even though the number of local optima in the $H1$ landscape is higher ($CV > 0.9$).

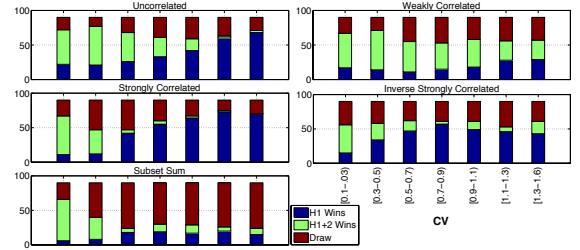


Figure 3: Number of instances where each operator performed significantly better and instances where no significant difference was found (Draw). Significance determined using Wilcoxon rank-sum test.

3. CONCLUSIONS

The number of local optima in the $H1$ landscape was found to be strongly and negatively correlated with weights CV in moderately constrained instances of the 0-1 KP. We showed that this correlation can be used to guide the choice of the best search operator. These results are similar to the results we obtained for the number partitioning problem [1].

4. REFERENCES

- [1] K. Alyahya and J. E. Rowe. Local optima and weight distribution in the number partitioning problem. In *PPSN XIII*, volume 8672, pages 862–871. 2014.
- [2] J. Gottlieb. On the feasibility problem of penalty-based evolutionary algorithms for knapsack problems. In *Applications of Evolutionary Computing*, volume 2037, pages 50–59. 2001.
- [3] Z. Michalewicz and J. Arabas. Genetic algorithms for the 0/1 knapsack problem. In *Methodologies for Intelligent Systems*, volume 869, pages 134–143. 1994.
- [4] D. Pisinger. Where are the hard knapsack problems? *Comput. Oper. Res.*, 32(9):2271–2284, 2005.
- [5] P. F. Stadler. Fitness landscapes. In *Biological Evolution and Statistical Physics*, volume 585, pages 183–204. 2002.