Combinatorial Optimization through Statistical Instance-Based Learning

Orestis Telelis Panagiotis Stamatopoulos
Department of Informatics and Telecommunications
University of Athens
157 84 Athens, Greece
{telelis,takis}@di.uoa.gr

Abstract

Different successful heuristic approaches have been proposed for solving combinatorial optimization problems. Commonly, each of them is specialized to serve a different purpose or address specific difficulties. However, most combinatorial problems that model real world applications have a priori well known measurable properties. Embedded machine learning methods may aid towards the recognition and utilization of these properties for the achievement of satisfactory solutions. In this paper, we present a heuristic methodology which employs the instance-based machine learning paradigm. This methodology can be adequately configured for several types of optimization problems which are known to have certain properties. Experimental results are discussed concerning two well known problems, namely the knapsack problem and the set partitioning problem. These results show that the proposed approach is able to find significantly better solutions compared to intuitive search methods based on heuristics which are usually applied to the specific problems.

1. Introduction

An important issue to notice in combinatorial optimization (CO) problems that emerge from real world applications is that they exhibit some inherent structural and statistical properties. These properties constitute observable common knowledge for the humans that are in charge of solving the problem. The human experience transforms into heuristic tools for obtaining a satisfactory solution. In most heuristic algorithms, important knowledge concerning a particular problem is embedded in an abstracted and more generic form, so that it can be applied in multiple instances of the same problem model. This abstraction, however, is an obstacle in recognizing specific numerical and structural properties of the particular instance being solved.

In this paper, we propose a heuristic methodology for CO problems, which embeds simple machine learning methods for recognition and utilization of specific numerical problem properties. On every search state of the solving path, alternative choices are evaluated by a *Kernel Regression* (*KR*) scheme. The evaluation of each choice is an estimation of the expected objective function value, under the assumption that the solving path extends in favour of the respective choice. Solutions previously found by the proposed methodology are utilized as training examples, for evaluation of *partial* solutions on each search state. The huge search spaces of the CO problems faced in this work are not expected to supply consistent training sets. Therefore, the widely known *Leave One Out Cross Validation* (*LOOCV*) test adjusts the *KR* approximator's parameters with respect to the underlying training set, so that a minimal estimation error is achieved.

The application of machine learning towards achievement of optimized solutions is a relatively recent aspect. However, there have been some remarkable, as well as pioneering, works dealing with it. We discuss here some of them by reporting briefly their point of view.

In [3], a reinforcement learning system is described for learning evaluation functions of startup search states for a local search algorithm. Reinforcement learning [6] methods have gained great attention because of their state-reward policy, which seems to fit well in the search state paradigm imposed by problem solving. Interesting works have emerged, such as [12], where a job-shop scheduling problem is addressed through a reinforcement learning approach. In [10], scheduling control knowledge is accumulated through reinforcements and exploited for schedule repair. Machine learning has also been used for deciding the best search policy on a problem [4], as well as for configuring specific problem solving methods [11]. In [8], an analytical learning technique is used for heuristic induction. Four main research directions of statistical machine learning application to combinatorial optimization are surveyed in [2]. The approach presented in our paper shares a common part with the *search space understanding* direction be-

cause it gathers statistical information relative to properties of the search space during the solving process. It also lies in part within the *evaluation function learning* discipline, since a *KR* scheme is employed for the approximation of an evaluation function, which shares its optimum with the objective function of the problem.

To our knowledge, machine learning techniques have been mostly integrated in local search procedures. In this paper, we present a heuristic function which employs *KR* and is designed to cooperate with *solution constructive search methods* for global optimization. This aspect is of particular interest in solving CO problems, since constructive search methods are able to preserve the validity of a problem's constraints during the search. This is not the case with local search procedures. They often visit invalid search states, and thus it is harder to even find a feasible solution.

In the following, the proposed methodology is presented initially at a higher level and then, the machine learning based heuristic algorithm is described in more detail. The application of the approach on two specific problems, namely the knapsack problem and the set partitioning problem, is discussed next. Comments on the experimental results are made and, finally, the concluding remarks are presented and further work is briefly described.

2. An Overview of the Search Schema

Constructive search is the kind of solving procedure exploited in this work. By the term constructive search, we designate the construction of a solution to the CO problem by assigning a value to each decision variable in turn, and thus by searching a tree of variable assignments. This policy comes in contrast with the various local search techniques, such as hill climbing and simulated annealing, which alter an already known complete assignment at each step, in order to obtain a new one.

The algorithmic schema is iterative. An overview of the approach is presented in Fig. 1. If I is a problem instance to be solved, the first thing to do is to use some simple heuristic method to obtain initial solutions S. A simple method might be a *Depth First Search* (DFS), guided by a common heuristic that intuitively fits the problem. This step stops after a limited time interval, which suffices for obtaining some initial solutions. The set S is then used for the production of an initial training set $\mathcal{E}$ for the machine learning algorithm employed by the heuristic.

The first step of the iterative process shown in Fig. 1 is a preprocessing procedure, which adapts the *KR* approximator to the training set, in order to achieve higher prediction accuracy with respect to the underlying training set $\mathcal{E}$. The problem instance is then solved by some constructive search algorithm, guided heuristically by the *KR* supported heuristic. The search stops when some criterion, such as a time limit, is met. The training set $\mathcal{E}$ is augmented with information extracted from newly found solutions and the process is repeated. The number of iterations is subject to experimentation. It is important to note that there is no constraint dictating that solutions obtained in each iteration should be better than those found in the previous iterations. The absence of such a constraint contributes to the enrichment of the training set with feasible solutions of varying qualities, which contribute to a more detailed picture of the search space.

```
Find some solutions S to I
       using a "simple" method
Produce E of the KR-heuristic from S
Iterate until end criterion met
   Adapt KR approximator to E
       using LOOCV
   Obtain new solutions constructively
       with KR-Heuristic
   Augment E using new solutions
Return Best Solution
```

Figure 1. The overall algorithmic schema

3. Design Details and the Heuristic Function

In this section, various aspects of the heuristic algorithm are discussed concerning the instance-based learning method, the representation of training examples and the dynamic KR approximator selection.

3.1. Kernel Regression on Nearest Neighbours

The machine learning methodology exploited within the proposed framework belongs to the family of memory-based or instance-based methods [9]. Memory-based learning methods explicitly store all training examples they are shown. Only at prediction time do they perform non-trivial amounts of computation, which is their main disadvantage. We use the *Kernel Regression* (*KR*) method for approximating the value of a real function. The *KR* method is also known as *Locally Weighted Averaging*. There is a generic scheme for *KR*, which might be configured in many different ways as is the case for *locally weighted* learning methodologies [1].

The exact configuration of *KR* used in this work follows. The *KR* algorithm is used here for real function value estimation. The function being approximated can be considered as: $f : \Re^n \mapsto \Re$.

The training set for the *KR* algorithm contains pairs of input vectors to f and their corresponding f-outputs. Thus, if $\vec{x} \in \Re^n$ is an input vector to the function, the respective

training example contained in the training set will be the pair $\langle \vec{x}, f(\vec{x}) \rangle$.

Let $\vec{x_q}$ be a query to the *KR* algorithm. The prediction $\hat{f}(\vec{x_q})$ is calculated by the algorithm as follows:

$$\hat{f}(\vec{x_q}) = \frac{\sum_{i=1}^{k} K(d(\vec{x_i}, \vec{x_q}))f(\vec{x_i})}{\sum_{i=1}^{k} K(d(\vec{x_i}, \vec{x_q}))} \quad (1)$$

In formula (1), $d(\vec{x_i}, \vec{x_q})$ denotes the Euclidean distance of the query vector $\vec{x_q}$ from the i-th training example vector $\vec{x_i}$. The k parameter stands for the number of training examples nearest to $\vec{x}_q$, that contribute their known f-values to the prediction $\hat{f}(\vec{x_q})$.

The function $K : \Re \mapsto \Re$ is the *kernel* function, which assigns a smaller weight to the contribution of $f(\vec{x_i})$ to the sum, as much as greater is the distance of $\vec{x_i}$ from the query vector $\vec{x_q}$. Thus, the contribution of less significant values, i.e. values that correspond to more distant vectors, is punished. The kernel function to be used at each iteration of the search schema is determined dynamically by *LOOCV*, from a repertoire of available kernel functions. Dynamic selection of kernel function is a part of the *KR* approximator's adjustment, and will be discussed in a following paragraph in more detail. When the algorithm is presented with a query, all attributes are scaled down to the $[0, 1]$ range. This normalization helps avoiding the domination of large-ranged attributes in computations.

3.2. Representation of Training Examples

An important issue for the applicability of *KR* is the implicit definition of the f function, mentioned in paragraph 3.1, whose value is going to be estimated. The function input consists of vectors in the Euclidean space $\Re^n$ which describe feasible solutions to the CO problem. The function value for each of these vectors is the value of the objective function of the problem for the corresponding solution.

Each training example for the *KR* approximation scheme is a pair of a solution descriptive vector $\vec{F}$, known as the *features vector*, and the respective objective function value $obj(\vec{F})$. Thus, the training set $\mathcal{E}$ can be defined as

$$\mathcal{E} = \{\langle \vec{F}, obj(\vec{F}) \rangle \mid \vec{F}\text{: extracted from a solution}\} \ .$$

The features (i.e. the dimension values) of the features vectors are real arithmetic values that correspond to specific properties of the solution to the optimization problem. Each feature should be an aggregate function on the assignments of the problem variables. As in every step of the search a decision variable is selected to be assigned a value, the features of each vector should be calculated upon the solution path. The only limitations on the features that might belong in a features vector are imposed by the problem structure. As discussed later, a solution can be described through statistical information that is considered to be characteristic of the solution's quality in terms of the objective function value.

3.3. Selection of *KR* Approximator

The *Leave One Out Cross Validation* (*LOOCV*) test appears to be quite appropriate for adapting a *KR* approximator to the training set of each iteration in Fig. 1. A discussion on cross validation tests can be found in [7]. Within our study we have limited the selection of a proper *KR* approximator to the selection of a k-value and a kernel function K from a set of available kernel functions. A brief overview of commonly used kernel functions can be found in [1].

Each different pair of k-value (number of nearest neighbours contributing to the estimation) and kernel function yields a different *KR* approximator. For each candidate approximator, each training example is estimated, as if it was a novel example, using as training set the remaining training examples. The distance of this estimation from the actual target value is a measure of the error in prediction. The selected approximator is the one that yields the lowest average prediction error over all training examples.

As the proposed methodology is supposed to solve problems that enclose properties in a rather statistical than precisely defined manner, training data collected during the solving process are expected to be inconsistent. Search spaces of CO problems are extremely large and different solutions to a CO problem might belong to different neighbourhoods of the problem's search space. The features used for the description of the solutions are chosen empirically as representative of the problem's a priori known statistical properties. However, the selected features might prove to be insufficient for the discrimination of certain solutions. It is expected that some solutions might belong to different neighbourhoods, whereas their discrimination in terms of distance of features vectors might be impossible. Such inconsistency of the training data is handled via the dynamic approximator selection.

3.4. The Heuristic Function

During the construction of a potential solution, the system simultaneously constructs a path towards the bottom of the search tree. On each node of the tree visited, decisions must be taken, so that the next step down the tree is the most promising for the solution quality among all available choices. We describe the heuristic function which guides

```
getBestAssignment(E, P, A)
  For each assignment α ≡ ⟨x_j = v⟩ ∈ A
    P̂ = P ∪ {α}
    Calculate F⃗_P̂
    val_α = KR(E, F⃗_P̂)
  Choose β ∈ A such that val_β is optimum
  return β
```

Figure 2. The *KR* supported heuristic function

the search by exploiting the previous experience acquired by the system.

Let $\mathcal{P}$ be the so far constructed path during the search. This is a *partial* path. Each search step consists of two choices: the selection of an unassigned decision variable of the problem and the determination of a value to be assigned to it. Let $\mathcal{A}$ be the set of all possible (i.e. feasibility preserving) such assignments and, consequently, the set of all possible ways to augment the partial path $\mathcal{P}$. The heuristic function should dictate an assignment from $\mathcal{A}$ as the next step for the extension of $\mathcal{P}$.

As mentioned in paragraph 3.2, the training examples for *KR* are features vectors calculated upon feasible solutions of the problem, i.e. upon complete paths. However, even a partial path can be used to calculate such a vector, if the unassigned decision variables are ignored. If statistical or aggregate information is used to describe the partial path extension by using a features vector, then it is reasonable to prefer extensions whose features present similarity to these of the best known solutions contained in the training set. In this way, portions of the search space that have previously produced good solutions are explored further. Let $\hat{\mathcal{P}}$ be the partial path resulting after augmenting $\mathcal{P}$ with a choice from $\mathcal{A}$. The features vector $\vec{F}_{\hat{\mathcal{P}}}$ is calculated upon $\hat{\mathcal{P}}$. The *KR* approximation scheme is requested to produce an objective function value estimation for $\vec{F}_{\hat{\mathcal{P}}}$. The extension of $\mathcal{P}$ which yields the optimum estimation is preferred over all other choices. An overview is presented in Fig. 2.

4. Application on Two Problems

The heuristic methodology was tested on two well known CO problems, namely the *knapsack* and the *set partitioning* problems. These are described below.

Knapsack. Given the n-dimensional vectors: profits $\vec{P}$ with $P_j \in \mathcal{Z}^+$, weights $\vec{W}$ with $W_j \in \mathcal{Z}^+$, $\vec{X}$ a vector of binary decision variables and some capacity $C \in \mathcal{Z}^+$,

$$\text{maximize } Z = \sum_{j=1}^{n} P_j x_j$$

$$\text{subject to } \sum_{j=1}^{n} W_j x_j \leq C \ .$$

Set Partitioning (SP). Given a $m \times n$ binary matrix A, a n-dimensional cost vector $\vec{C}$ with $C_j \in \mathcal{Z}^+$ and the n-dimensional vector $\vec{X}$ of binary decision variables, we want to

$$\text{minimize } Z = \sum_{j=1}^{n} C_j x_j$$

$$\text{subject to } \sum_{j=1}^{n} A_{ij} x_j = 1, \quad i = 1 \ldots m \ .$$

4.1. Generation of Problem Instances

An important assumption mentioned in the very beginning of this paper is that the CO problems faced have some a priori known properties. For the previous problems, instances were generated that have such properties. In [11], a generation method for knapsack instances is briefly discussed. The $\vec{W}$ vector mentioned previously is determined from a normal distribution, while the fraction P_j/W_j is also calculated from a normal distribution. Each P_j element of the $\vec{P}$ vector is computed as $P_j = W_j \cdot P_j/W_j$. The knapsack instances created in this way tend to associate greater P_j values to greater W_j values. These problems have been shown to be more difficult.

For the SP problem, instances were generated in a similar way as for the knapsack problem. The number of 1s contained in the j column of $A_{m \times n}$ and the ratio of the column cost C_j to the number of 1s were determined by two different normal distributions, while the C_j quantity was computed as a dependent variable. Thus, columns containing more 1s, tend to have higher cost values C_j. Because the 1s in each column are decided via a uniform distribution, the mentioned property is slightly depressed by the satisfiability of the problem constraints.

For the SP and knapsack instances, optimum solutions were planted in a simple manner.

4.2. Selection of Features

The selection of features that assemble the features vectors is mostly important for the accuracy of the predictions of the *KR* scheme. Features that encapsulate some information relevant to the objective function's value are preferred, since they are expected to provide a quantitative partition of the search space into regions with expected objective function value. For each of the aforementioned problems, their features are presented, which were selected intuitively.

Knapsack. Three features constitute the features vector for the knapsack problem:

1. The mean value of the fraction P_j/W_j for js such that $X_j = 1$ in the solution.
2. The mean value of the profits P_j which participate in the objective function value, i.e. for js such that $X_j = 1$.
3. The weighted average of profits that participate in the objective function value. The contribution of each P_j profit to the average is weighted by the inverse corresponding W_j value.

Set Partitioning. For the SP problem, the features vector constituted of two features:

1. The mean value of the number of 1s contained in the columns of the binary matrix A that participate in a solution.
2. The mean value of the costs of the columns of matrix A that constitute the solution.

The j-th column of the matrix A is part of a solution if $X_j = 1$.

5. Experimental Results

Experiments were carried out on 10 instances for each problem. The instances were of varying sizes.

5.1. Experimental Configuration

For each problem, a heuristic solving method was chosen among the most commonly used. The chosen method provided the best results for the corresponding problem, within the experimentation time interval, and was employed for the construction of the initial training set. The results obtained by the proposed methodology were compared to those provided by the common method in the same time. Only solutions found by the common method within the first 10 minutes were considered for the construction of the initial training set. The common methods were applied for 4200 seconds. The overall running time of the iterative part of our methodology was arranged to last for 3600 seconds, so that summed to the initial training set construction time equals to 4200 seconds.

The system was provided a set of kernel functions to choose from, in order to configure a *KR* approximator, which would yield a minimum expected prediction error during the *LOOCV* test. We let the system choose among the following kernel functions: $K_a(d) = 1/d^a$ and $E_a(d) = 1/e^{d^a}$, for $a = 1, 2, 3$. The forms K_a and E_a lie among the most commonly used [1]. Particularly for the K_a form, rare $d = 0$ cases were handled by assigning the query vector an estimation equal to the known value for the corresponding nearest (zero distance) neighbour. Values of a greater than 3 were not considered, since they would yield very low kernel values, and thus, the *KR* estimation would be dominated by the contribution of one unique nearest neighbour.

The training set expands from one iteration to the next. In order to avoid cases where the *LOOCV* test would decide large values for k, and thus, slowing down the *KR* approximator, the size of the training set was kept stable to, at most, 25 training examples. At the end of each iteration, the training examples that represent solutions of worse qualities are removed.

The measurement of success for the experiments presented in this section is defined as the percentage of improvement achieved by our methodology towards the optimal solution, in comparison with the performance of the common method. Thus if c_o is the best solution found by the common method, ml_o is the best solution obtained by our methodology and opt is the optimal solution, performance is measured as

$$\alpha = 100 \times \frac{c_o - ml_o}{c_o - opt} \ .$$

5.2. Knapsack

For the knapsack problem, the simple heuristic policy of *"try the most profitable choice first"* on a DFS proved to be quite successful in obtaining soon some high quality solutions.

Twelve iterations, of 5 minutes each, were performed on each problem instance after the construction of the initial training set. *Limited Discrepancy Search* (LDS) [5] was used as constructive procedure for the proposed method.[1] The quality of solutions which were obtained for all instances significantly exceeded the quality of solutions found by the employed common heuristic method. Table 1 summarizes the results. The characteristics of each problem are depicted, namely the n parameter and an estimation μ of the average number of "items" that fit in the knapsack, calculated as the ratio of the capacity C to the mean value of the weights in vector $\vec{W}$.

5.3. Set Partitioning

The common method that provided the best results for the SP problem was the heuristic policy *"try the col-*

[1] LDS was also tested with common heuristics on both problem instances but did not outperform DFS

Instance		Performance			
n	μ	c_o	ml_o	opt	α
3000	15	53848	54856	57499	27.6
4000	20	73954	77976	78693	84.8
4000	40	140669	143079	151415	22.4
4000	51	179656	186341	189211	69.9
4000	23	82385	83882	88849	23.1
4000	30	107621	109785	113112	39.4
6000	30	108744	111043	116577	29.3
6000	20	70956	71750	76909	13.3
8000	15	56049	57100	60031	26.4
8000	23	90257	91352	96367	17.9

Table 1. Experimental results for the knapsack problem

Instance		Performance			
n	m	c_o	ml_o	opt	α
4000	15	1175	1063	898	40.4
4000	18	1474	1401	1113	20.2
4000	20	900	481	452	93.5
5000	18	15698	12563	10669	62.3
5000	25	993	910	544	18.5
5000	30	2720	2538	1734	18.4
6000	20	754	580	407	50.1
6000	25	1148	893	593	45.9
6000	30	1133	1088	779	12.7
7000	25	1128	905	569	39.8

Table 2. Experimental results for the SP problem

umn with the minimum cost first" with DFS. The proposed methodology performed significantly better on all instances. Table 2 depicts the characteristics of the SP instances and the performance of our methodology. The dimensions of the $A_{m\times n}$ binary matrix are also shown as the major characteristics for an SP instance.

Six iterations, of 10 minutes each, were performed on each SP instance. Remarkably better solutions were found by the proposed methodology using LDS, than those obtained by the common heuristic policy.

5.4. Behaviour of the Methodology

The experimentations on the SP and the knapsack problems gave a view of the behaviour of the methodology during the solving process. Figure 3 gives a low level view of a solving path followed by the heuristic function of Fig. 2

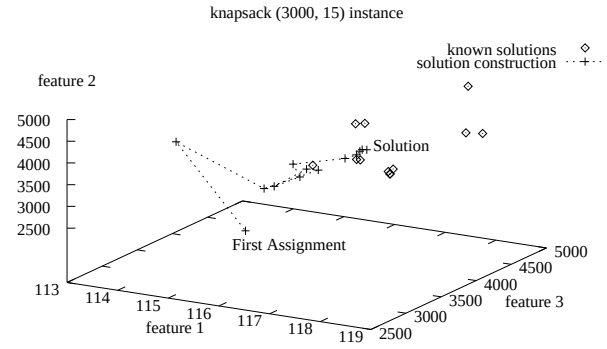

Figure 3. Solution construction path for a knapsack instance

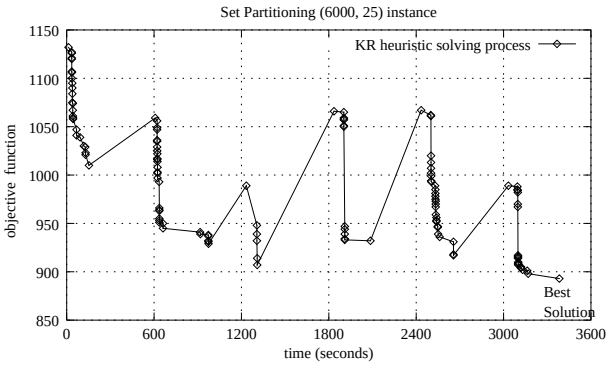

Figure 4. Solving process for an SP instance

for a knapsack instance, whereas the overall solving performance of the proposed iterative methodology is depicted in Fig. 4 on an SP instance.

Figure 3 depicts the features vectors for known solutions to the $\langle n = 3000, \mu = 15\rangle$ knapsack instance that belong to the training set. It also demonstrates the trajectory of a solution construction using the given training set. Every cross point in the trajectory represents a features vector calculated upon a partial path, after a new assignment to some decision variable is performed. At every search step, the heuristic function chooses the assignment which brings the features vector of the extended path closer to specific features vectors of the best known solutions.

The overall functionality of our methodology is demonstrated in Fig. 4, on the $\langle n = 6000, m = 25\rangle$ SP instance. Each ascending line in the diagram is a *border* between two subsequent iterations of the schema in Fig. 1, because only then a solution might be of worse quality than the previous one. Within the iterations each solution should be better than the previous found. The best solution is found in the last iteration, but this was not the case for all SP instances. This fact confirms the need for information acquisition and exploitation regarding the search space, in order to explore

its most promising portions.

6. Conclusions and Further Work

In this paper, we propose a heuristic methodology for combinatorial optimization, which employs instance-based learning and function approximation through kernel regression, for guiding any constructive search procedure. This work is not concerned with the achievement of feasible solutions to a problem (this issue is addressed successfully by sophisticated implementations of constructive search methods, e.g. backtracking), but with the guidance of search to promising regions of the search space, as far as optimality is concerned.

Problem models grown from real world applications usually enclose vast contents of numerical information, which can be statistically handled for the construction of optimized solutions. The objective functions of such problems are generally designed upon desirable facts and dictate the intuitive policy for their optimization. We suggest that known solutions to these problems are represented via statistical information calculated upon each solution's structural constituents. The proposed policy constructs a solution by minimizing its distance (in terms of its statistical properties) from the best (in terms of objective function value) known solutions, which lie nearby.

Experimental results were carried out on two widely used models of real world combinatorial problems, namely the knapsack and the set partitioning problems. These problems model important real world applications, such as nuclear waste packing and crew scheduling. The methodology performed satisfactory on these problems and obtained solutions whose quality exceeded the quality of solutions obtained by other heuristic methods, common for each of the problems.

Some directions for further research are drawn from questions that arise quite naturally. In our experiments, the proposed framework performs satisfactory for an initial training set created by some simple methods. *However, which is the proper way for systematically sampling initial solutions of a useful quality distribution from the problem's search space?* This is an important issue, which could possibly boost the performance of the heuristic function, since an initial set of solutions with known quality distribution is actually a detailed picture of the search space.

The complexity of the heuristic function depends on the size of the training set. Using big training sets slows down the search, while small sets provide little information about the search space. We have started to examine the option of partitioning the training set into consistent clusters, each of which represents a small portion of the search space. Each of these clusters is meant to be used as a separate training set, for searching the corresponding sections of the search space in a locally exhaustive manner.

As an aspect of future work, extended experimentations on a variety of optimization problems is expected to reveal valuable statistical features, strongly informative and representative of the corresponding search spaces.

References

[1] C. G. Atkeson, A. W. Moore, and S. Schaal. Locally weighted learning. *Artificial Intelligence Review*, 11(1–5):11–73, February 1997.

[2] J. Boyan, W. Buntime, and A. Jagota. Statistical machine learning for large-scale optimization. *Neural Computing Surveys*, 3:1–58, 2000.

[3] J. A. Boyan and A. W. Moore. Learning evaluation functions for global optimization and boolean satisfiability. In *Proceedings of the 15th National Conference on Artificial Intelligence AAAI-98*, pages 3–10, Madison, Wisconsin, July 1998. AAAI Press.

[4] D. J. Cook and R. C. Varnell. Maximizing the benefits of parallel search using machine learning. In *Proceedings of the 14th National Conference on Artificial Intelligence AAAI-97*, pages 559–564, Providence, Rhode Island, July 1997. AAAI Press.

[5] W. D. Harvey and M. L. Ginsberg. Limited discrepancy search. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence IJCAI-95*, pages 607–613, Montréal, Québec, Canada, August 1995. Morgan Kaufmann.

[6] L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[7] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence IJCAI-95*, pages 1137–1145, Montréal, Québec, Canada, August 1995. Morgan Kaufmann.

[8] S. Minton. An analytic learning system for specializing heuristics. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence IJCAI-93*, pages 922–929, Chambéry, France, August 1993. Morgan Kaufmann.

[9] T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.

[10] K. Miyashita. Learning scheduling control knowledge through reinforcements. *International Transactions in Operational Research*, 7(2):125–138, March 2000.

[11] M. J. Realff, P. H. Kvam, and W. E. Taylor. Combined analytical and empirical learning framework for branch and bound algorithms: The knapsack problem. *Artificial Intelligence in Engineering*, 13(3):287–300, July 1999.

[12] W. Zhang and T. G. Dietterich. A reinforcement learning approach to job-shop scheduling. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence IJCAI-95*, pages 1114–1120, Montréal, Québec, Canada, August 1995. Morgan Kaufmann.