Geometric Surrogate-Based Optimisation for Permutation-Based Problems

Alberto Moraglio
School of Computing
University of Kent
Canterbury, UK
a.moraglio@kent.ac.uk

Yong-Hyuk Kim
Dept. Comp. Sci. & Eng.
Kwangwoon University
Seoul, Korea
yhdfly@kw.ac.kr

Yourim Yoon
Sch. Comp. Sci. & Eng.
Seoul National University
Seoul, Korea
yryoon@soar.snu.ac.kr

ABSTRACT

In continuous optimisation, surrogate models (SMs) are used when tackling real-world problems whose candidate solutions are expensive to evaluate. In previous work, we showed that a type of SMs – radial basis function networks (RBFNs) – can be rigorously generalised to encompass combinatorial spaces based in principle on *any arbitrarily complex underlying solution representation* by extending their natural geometric interpretation from continuous to general metric spaces. This direct approach to representations does not require a vector encoding of solutions, and allows us to use SMs with the most natural representation for the problem at hand. In this work, we apply this framework to combinatorial problems using the permutation representation and report experimental results on the quadratic assignment problem.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

General Terms

Theory

Keywords

Representations, Radial-Basis Functions, Surrogate Model Optimization

1. MODEL-BASED OPTIMIZATION

The traditional procedure of surrogate model based optimisation (SMBO) [2] is outlined in Algorithm 1.

SMBOs are naturally suited to continuous optimization or to discrete problems when solutions are vectors of integers, as there are many statistical techniques that can be used to build a surrogate model of the fitness landscape from data-points based on these representations.

In many optimization problems, natural solution representations are not real or integer vectors, but they can be permutations, variable-length sequences, trees, graphs or any arbitrarily complex structures. The choice of an adequate representation for the problem at hand is often critical to the success of the search algorithm, whichever particular search algorithm is to be used. In particular, although, permutations can be regarded as a special type of vectors, they generally cannot be effectively treated as vectors as the relevant information they encode about the problem at hand is in the order of the elements, rather than in the absolute values at each location in the vector.

GECCO'11, July 12–16, 2011, Dublin, Ireland.
ACM 978-1-4503-0690-4/11/07.

Algorithm 1 Surrogate Model Based Optimisation

1: Sample uniformly at random a small set of candidate solutions and evaluate them using the expensive objective function (initial set of data-points)
2: **while** limit number of expensive evaluations not reached **do**
3: Construct a new surrogate model using all data-points available
4: Determine the optimum of the surrogate model by search, e.g., using an evolutionary algorithm (this is feasible as the model is cheap to evaluate)
5: Evaluate the solution which optimises the surrogate model in the problem with the expensive objective function (additional data-point available)
6: **end while**
7: Return the best solution found (the best in the set of data-points)

2. GENERALIZED RBFNS

Is there a systematic and rigorous way to adapt a surrogate model to the target representation which does not require us to rethink the surrogate model, or make ad-hoc adaptation to the model, for any target representation considered however complex it is? In [3], the authors proposed a generalization of radial basis function networks [1] which answers in the affirmative the question above.

A radial basis function (RBF) is a real-valued function whose value depends only on the distance from some given points $\mathbf{c}_i$. The distance considered is usually the Euclidean distance. The most frequently used are Gaussian functions of the form $\phi(\mathbf{x}) = \exp(-\|\mathbf{x} - \mathbf{c}\|^2)$. Radial basis functions are used to build function approximations of the form (i.e., RBFNs): $y(\mathbf{x}) = w_0 + \sum_{i=1}^{N} w_i \, \phi(\|\mathbf{x} - \mathbf{c}_i\|)$. RBFNs can be trained to approximate the unknown fitness landscape by choosing the centers $\mathbf{c}_i$ to coincide with the known points $\mathbf{x}_i$ and finding the weights w_i such that the unknown function interpolates exactly the known data-points. They can be found by using simple linear algebra, which involves a matrix inversion, to solve a system of linear equations. The trained model can be queried with unseen points to forecast their function values.

All aspects of RBFNs that allow us to use them as surrogate models, i.e., model definition and representation, train-

ing, querying and searching of RBFNs can be naturally generalized from Euclidean spaces to general metric spaces, by replacing the Euclidean distance with a generic metric. Therefore, the generalized model applies to any underlying solution representation once a distance function rooted on that representation is provided (e.g., Swap distance on permutations). In particular, this method can be used *as it is* to learn in principle any function mapping directly complex structured representations to reals without introducing any arbitrary ad-hoc adaptation to the RBFNs. There is no special requirement of pre-processing the target representation and shoehorn it in a vector of features.

3. EXPERIMENTS

Experiments were carried out on three standard quadratic assignment problem (QAP) instances (kra32, tho30 and nug30, where the number in the name indicates the instance size), and on two unimodal problems on permutations of size 30 based on the Hamming distance (unih30) and the Swap distance (unis30). As distance functions between permutations as base for the instantiation of the SMBO, we used the Hamming distance ($SMBO_H$) and the Swap distance ($SMBO_S$). The number of total available expensive function evaluations was set to $n = 100$, 10 of which are initial data-points. To search the surrogate model we used a memetic algorithm on permutations with cycle crossover, swap mutation and local search based on the 2-opt neighborhood. We compared the SMBOs with a standard genetic algorithm (GA) and with random search (RS) applied directly on the problem with the expensive objective function. We gave all algorithms in the comparison exactly the same number of expensive objective functions, and we did 50 independent runs of each experiment.

Table 1 reports the results of the comparison. These results consistently rank the $SMBO_H$ as the most effective algorithm, followed by the GA, then by the $SMBO_S$ and finally followed by the RS. An initial conclusion that can be drawn is that the surrogate based on the Hamming distance really helps in locating better solutions given the same amount of expensive fitness evaluations, both on the QAP instances and on the unimodal ones. A rather surprising result is that the surrogate model based on the Swap distance does not seem to be very effective, as it is better than the RS but worse than the GA. This is surprising because the Hamming distance and the Swap distance are related distances. It is also noteworthy that the $SMBO_H$ is better than the $SMBO_S$ on the unimodal landscape under the Swap distance (unis30), which we introduced in the test-bed of problems because intuitively this can be thought as the easiest landscape to optimize for the $SMBO_S$.

To have a better picture of the working mechanism of the SMBO algorithms we did a number of analyses. Firstly, to make sure that the distances chosen as the basis for the SMBO are suited to the problem at hand, we did a static analysis of the predictive power of the surrogate models when they are used in isolation from the SMBO. We found that on all instances the Hamming model produces better prediction than the Swap model. Furthermore, a fitness-distance correlation analysis showed that the Swap distance is more suited than the Hamming distance to the QAP instances, as it obtains higher correlation. However, this is not reflected in the performance. So, the static prediction power of a model is a better indicator of which distance to use with the SMBO. We performed other analyses not reported here to understand how the hardness of the surrogate models and the special topological features of the spaces under the two distances affect the overall SMBO performance. However, after these analyses, it is still eluding us exactly what particular property is which the Hamming distance has and that the Swap distance has not that makes the former distance more suitable to be used effectively with an SMBO.

Table 1: Results for Random Search (RS), Genetic Algorithm (GA), $SMBO_H$ and $SMBO_S$ on QAP instances (kra32, tho30, nug30) and unimodal ones (unih30, unis30). Best, mean and standard deviation over 50 independent runs of the best fitness found by each algorithm are reported.

Algorithm	Best	Mean	SD
kra32			
RS	24156	25008.036	434.752
GA	23440	24625.032	586.598
$SMBO_H$	**22590**	**24094.320**	616.714
$SMBO_S$	23848	24833.920	452.219
tho30			
RS	190256	196662.822	3211.198
GA	**180274**	194389.912	4414.218
$SMBO_H$	180860	**193415.440**	4629.032
$SMBO_S$	186172	195231.920	3534.208
nug30			
RS	7350	7618.000	84.598
GA	7296	7558.360	126.650
$SMBO_H$	**7276**	**7500.240**	97.363
$SMBO_S$	7328	7563.880	94.384
unih30			
RS	24	25.796	0.784
GA	22	25.044	1.263
$SMBO_H$	**17**	**20.840**	1.554
$SMBO_S$	21	25.180	0.953
unis30			
RS	19	21.657	0.871
GA	17	20.906	1.294
$SMBO_H$	**15**	**18.480**	1.825
$SMBO_S$	19	21.040	0.937

4. CONCLUSIONS AND FUTURE WORK

We tested for the first time the new SMBO framework on the permutation representation. The SMBO based on the Hamming distance performed best, whereas the SMBO based on the Swap distance did less well even on problems which we would have expected it to do well. We will investigate the causes of this further in future work. Furthermore, we will test the permutation-based SMBO framework on real-world scheduling problems, and experiment with more complex representations.

5. REFERENCES

[1] L. C. Jain. *Radial Basis Function Networks*. Springer, 2001.

[2] D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21(4):345–383, 2001.

[3] A. Moraglio and A. Kattan. Geometric generalisation of surrogate model based optimisation to combinatorial spaces. In *Proceedings of the 11th European Conference on Evolutionary Computation in Combinatorial Optimisation*, 2011. (to appear).