Mining Markov Network Surrogates to Explain the Results of Metaheuristic Optimisation*

Alexander E.I. Brownlee, Aidan Wallace, and David Cairns

University of Stirling sbr@cs.stir.ac.uk
http://www.cs.stir.ac.uk/~sbr

Abstract. Metaheuristics are randomised search algorithms that are effective at finding "good enough" solutions to optimisation problems. However, they present no justification for the generated solutions, and are non-trivial to analyse. We propose that identifying which combinations of variables strongly influence solution quality, and the nature of that relationship, represents a step towards explaining the choices made by the algorithm. Here, we present an approach to mining this information from a "surrogate fitness function" within a metaheuristic. The approach is demonstrated with two simple examples and a real-world case study.

Keywords: metaheuristics · surrogates · optimisation · explainability

1 Introduction

Metaheuristic algorithms exploit randomised search to approximate optimal solutions to difficult optimisation problems such as workforce scheduling or making thousands of design decisions for a house. *Explainable AI* has made advances towards explaining the decisions of black-box systems like deep neural networks [10,12]. However, there is little like this for metaheuristics. Approaches that come close include systematic analysis of the relationships between variables and objectives [5] and using solutions arising from the search process to seed classic sensitivity analysis [18]. Deb et al [9] proposed "innovization" to yield additional problem-based knowledge alongside the generated optimal solutions, by seeking common principles among Pareto-optimal solutions. More recently, Urquhart [15] used MAP-Elites to increase trust by addressing the common complaint from end users, when presented with a solution constructed by a metaheuristic, that they themselves had no role in the solution's construction. An archive of high performing but diverse solutions are found using MAP-Elites, and presented via an interactive decision making tool.

For an optimisation problem, a decision maker might seek two key insights to accept solutions found by metaheuristics. I1: *Does the solution solve the problem, or have we found an error or loophole in the problem's definition?* I2: *What characteristics of the solution are crucial to optimality, and which are simply artefacts of the stochastic processes inherent to metaheuristics, that could be amended for aesthetic or implementation purposes?* In this paper, we propose that such explanation can be achieved by identifying which combinations of variables strongly influence solution quality, and which can be ignored; the ideal values for the variables; and interactions between variables. We do this by following an approach to mining "surrogate fitness functions" described in [4].

Surrogate fitness functions [7, 11] are well-established for improving metaheuristic search efficiency. A computationally cheap model is trained in parallel with the optimization, replacing calls to a costly fitness function such as a long-running simulation [6, 7, 14]). We exploit a little-used additional benefit: a surrogate is an explicit model of the problem, as seen through the solutions visited by the algorithm. Given that the initial motivation for using the surrogate was to improve the speed of the search, this model is effectively "for free".

This paper revisits the Markov network Fitness Model (MFM), a probabilistic model for bit string encoded problems originally developed for the Estimation of Distribution Algorithm, DEUM [13]. The relationship between MFM parameters and global optima for a given problem can be exploited to yield explanations in the form of characteristics that make a solution high in fitness.

2 Markov Network Fitness model

We begin with a brief recap of the Markov Network Fitness Model surrogate, further detailed in [3]. Let $\Omega = \{0, 1\}^n$ be the search space (i.e., bit string encoded solutions). $f(x) \Rightarrow \mathbb{R}$ is the fitness function and $X = (X_1, \ldots, X_n)$ is the variable vector. $X_i = x_i$ denotes that variable X_i has value x_i, and $x = x_1 \ldots x_n$ denotes a joint configuration of X. $V_k(x)$ is a Walsh transformation [1], which encodes the values of a binary variable x_i from $[0, 1]$ into $[-1, 1]$, and groups multiple variables as their product. The Markov Network Fitness Model specifies, for each solution $x = x_1, \ldots, x_n$, a negative log relationship between the fitness function and the Walsh transformation of the variables:

$$-\, lnf(x) = U(x) = \sum_K \alpha_K V_K(x) \tag{1}$$

With a large enough population of solutions and their fitnesses, (1) yields a system of equations that can be solved using a least-squares approximation to estimate the α_K (this stage can be seen as a linear regression problem [16]). With the parameters specified, (1) becomes a model of the fitness function, which we can use to predict the fitness $f(x)$ for solutions.

3 Mining the MFM

[2, 3] detailed how MFM α values can yield insights into fitness and the region around the optima. Equation (1) specifies a negative log relationship between energy and fitness, so minimising energy is equivalent to maximising fitness. For a univariate term, $V_i(x)$ (corresponding to a single x_i), if $\alpha_i > 0$, setting $x_i = 0$ will minimise energy and thus maximise fitness and, if $\alpha_i < 0$, setting $x_i = 1$ will maximise fitness. For terms with two variables $V_{i,j}(x)$, then $\alpha_{i,j} > 0$ requires x_i $!= x_j$ to maximise fitness; $\alpha_{i,j} < 0$ requires $x_i == x_j$. So, the signs of the α_K point towards values taken by variables in the globally optimal solutions. The magnitude indicates the sensitivity of $f(x)$ to the values taken by each clique.

Two examples from [2, 3] focus on a toy benchmark problem, *Checkerboard*, and a biocontrol problem. With Checkerboard, the goal is to maximise the number of cells with oppositely-valued neighbours when the bit string is laid out in a grid. The univariate terms (corresponding to each variable x_i) were all around zero, implying that cells could either be 0 or 1 in the optima. All pairwise coefficients for a 25-bit checkerboard (Figure 1) are positive: i.e., neighbouring cells should take opposite values. The coefficients that are double the magnitude of

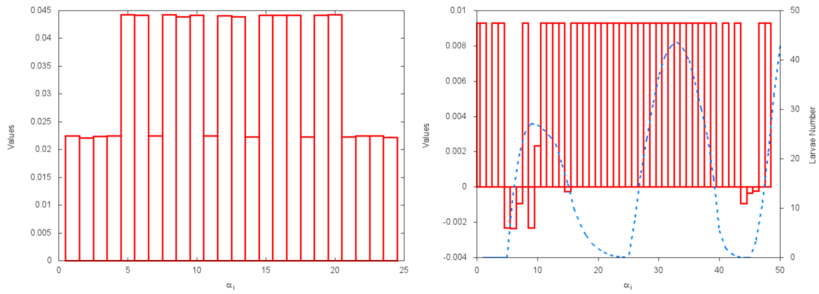

Fig. 1. Pairwise coefficients for checkerboard (left) and Univariate coefficients for the biocontrol problem and larval lifecycle (right)

the others correspond to the pairs of cells in the centre of the checkerboard, where they might be expected to have most influence.

The bio-control problem seeks to minimise insect larvae growth on mushrooms by choosing optimal times to spray the crop with nematode worms. Solutions are encoded as 50 bits representing times at which the bio-control spray is applied or not. Most of the univariate coefficients (each corresponding to one bit) (Figure 1) are positive, indicating no spraying. The few negative coefficients, when spray should be applied, coincide with growth points in the life cycle of the pest being targeted (represented by the blue dotted line).

For both problems, MFM coefficients, determined using only a few hundred randomly generated solutions, have a clear relationship with the underlying problem. The coefficients point towards optimal solutions and show sensitivity of the objective to particular variables or variable interactions.

4 Case Study: Cellular Windows

Our case study seeks to optimise the location of windows on a Chicago commercial building's southern façade; the goal is a design which minimises energy use and capital construction cost. This was previously studied in [8, 17, 19]. The external wall is divided into 120 cells which may be glazed, in a 15x8m grid. Fig 2 shows the fully glazed building.

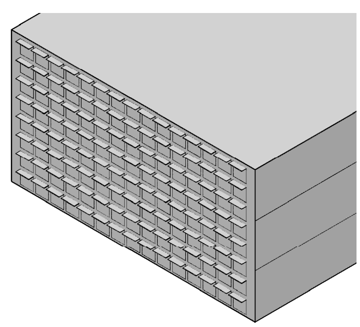

Fig. 2. Fully glazed façade

Objective 1: Minimise Energy. The unweighted sum of energy for heating, cooling and lighting over a year, computed by the EnergyPlus building simulation package. This is non-trivial: incoming sunlight reduces electric lighting demand but solar gain increases cooling and decreases heating energy demand. Heat loss through glass at night has the opposite effect. A single simulation run takes 1-2 minutes on an Intel i7 CPU: the original motivation for the use of surrogates to speed up the optimisation.

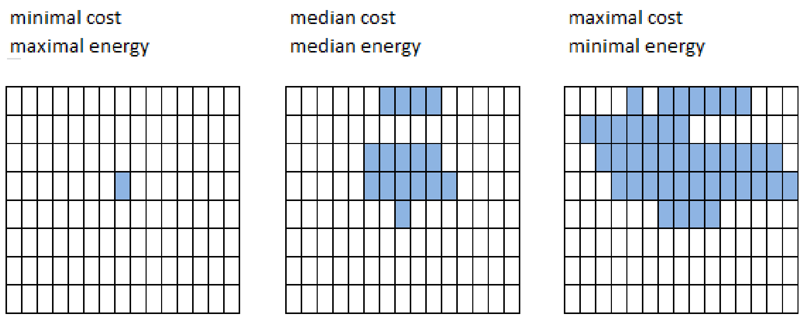

Fig. 3. Minimal, median and maximal cost solutions from the best attainment curve. White cells are unglazed, blue cells are glazed.

Objective 2: Minimise Cost. Construction cost for specified window configuration. A straightforward linear function of the number of windows.

Variables and Encoding. The wall is divided into 120 cells in a 15 x 8 grid. Each cell may be glazed or unglazed. This translates into a 120 variable bit string. A bit is true for a glazed cell and false otherwise.

5 Optimisation Results

Comparisons and analysis of results from several multi-objective evolutionary algorithms applied to this problem can be found in [8, 19]. The present paper focuses on mining a surrogate model, rather than on the optimisation process, so for convenience we replicate the best optimisation results from [19]. The specific algorithm is unimportant for this work and could be substituted by another that uses fitness to drive the search; however NSGA-II was found to perform well for this problem [8]. The algorithm used binary tournament selection; 100% crossover rate using uniform crossover; single bit-flip mutation for each new solution; population size 30 and a stopping criterion of 5000 unique evaluations.

Minimal, median and maximal cost solutions from the approximated Pareto front from a typical run are plotted in Figure 3. The substantial range of capital costs reflect the extra expense of glazing. The range of energy consumption is more modest, but still around 6% of the maximal consumption, representing considerable savings in emissions and energy costs over the building's life.

The approximated Pareto-optimal trade-off and the specific designs in each solution are already of great value to a decision maker, but themselves raise some questions. It would appear that, as a result of the algorithm's randomness, it missed the lowest cost solution (no glazing). It also produced odd glazing shapes on the higher-cost solutions. It would be helpful to know what the impact might be of making small aesthetic changes to these solutions.

One approach to explaining the contribution of individual glazing cells to optimality is a local sensitivity analysis [19]. For selected solutions, each cell was flipped from glazed to unglazed (or vice-versa), and the change in energy use determined. This is illustrated in Figure 4. The local sensitivities help to identify cells that were glazed or unglazed as a result of noise coming from using

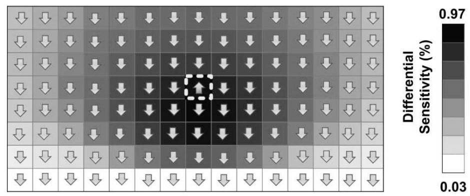

Fig. 4. Local sensitivities around the minimal cost solution. Arrows indicate the direction of change in energy consumption caused by mutating that bit in the solution and the shading shows the magnitude of the change.

a stochastic algorithm, and those which could be changed without impacting negatively on the objectives. However, the approach has the disadvantage that it requires further runs of the building performance simulation.

6 Mining the Surrogate

We constructed two MFMs for this problem; one each for energy and cost. We now mine these surrogates for explanations. The structure for the MFM (the neighbourhoods for each x_i) was fixed. Two sets of experiments were performed using different structures for the MFM.

Lattice structure. Initially, a grid structure was adopted, based on the intuition that glazing one cell would impact neighbouring cells. The MFM included 120 univariate V_K (one term per cell), and 240 pairwise V_K representing neighbouring cell pairs. Thus, there are 361 parameters in the model (including a constant), and a training population of around 1.1x this value should be used to obtain a good model [3]. The fittest (lowest energy or cost) 400 of the first 1000 solutions visited by NSGA-II were used as training data for one MFM per objective, for each of 30 repeat runs. The r^2 values comparing the predicted objective values with the true objective values from the simulation were 0.982 for energy and 0.997 for cost.

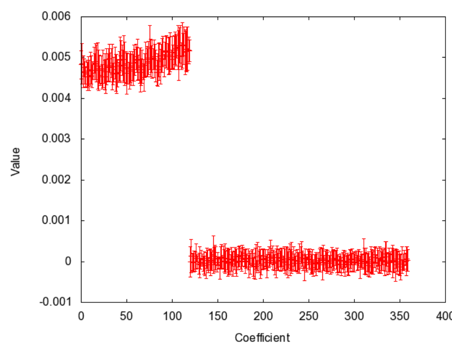

Fig. 5. Mean α_K values for energy. 1-120 are univariate α_K; 121-360 are pairwise α_K between neighbouring cells on the façade. Error bars are one standard deviation.

The mean and standard deviation for each α_K in the MFM was calculated over all the energy MFMs and all the cost MFMs. These values are plotted in Figure 5 for energy (for cost, the plot is similar, but with the higher points level rather than showing a gradual increase). The jump at α 120 coincides with the change from univariate α_Ks to the pairwise α_Ks. For both energy and cost objectives, the pairwise α_K values are all near zero. This means that they have little to no influence on either objective: it would seem that only the univariate α_K have any influence on the objectives (having non-zero values in the MFMs for both objectives) and our intuition on the appropriate structure was incorrect.

0.0053	0.0051	0.0053	0.005	0.0053	0.005	0.005	0.0049	0.0048	0.0051	0.005	0.005	0.0052	0.0051
0.0054	0.0054	0.005	0.0053	0.0053	0.0051	0.0054	0.005	0.0049	0.0049	0.0051	0.0048	0.0049	0.0051
0.0053	0.0056	0.0052	0.0051	0.0053	0.0051	0.005	0.0053	0.0049	0.0048	0.0051	0.0049	0.0052	0.0053
0.0052	0.0054	0.0051	0.0054	0.0053	0.0053	0.0048	0.0052	0.0051	0.0051	0.0051	0.0053	0.0051	0.0054
0.0056	0.0054	0.0056	0.0054	0.0053	0.005	0.0051	0.0051	0.0051	0.0049	0.0053	0.0049	0.0056	0.0053
0.0058	0.0054	0.0055	0.0054	0.0055	0.0053	0.0053	0.0052	0.0052	0.0053	0.0052	0.0056	0.0054	0.0053
0.0057	0.0052	0.0056	0.0057	0.0055	0.0056	0.0056	0.0057	0.0055	0.0055	0.0055	0.0055	0.0057	0.0058
0.0057	0.0054	0.0058	0.0058	0.0059	0.0057	0.0056	0.0057	0.0058	0.0058	0.0058	0.0059	0.0057	0.0057

Fig. 6. Mean coefficient values for energy MFM, arranged to match the locations of cells on the façade. Blue cells have high values, white medium, and red have low values.

Univariate structure. We repeated the process using a univariate structure (i.e., 120 α_Ks, one per variable) for both MFMs. The model was trained on the fittest 140 of the first 1000 solutions visited in the optimisation run. Median r^2 for the models on 400 unseen solutions was 0.993 for energy and 0.998 for cost. The mean values for these coefficients for energy are shown in Fig 6 as a grid corresponding to the façade. Cell colour shows each coefficient's value relative to the others: high values being blue, through white to red. For both objectives, all α_k are positive.

For the **energy** objective, there is a clear (though small) bias towards the upper right (East) of the façade, also visible in the rising values in Fig 6, suggesting that glazing should be concentrated in that region. This matches the local sensitivity analysis (Fig 4), but has the benefit that no additional evaluations were needed, and is rooted in models representing solutions spanning several generations rather than just the last one. It is also concordant with the real-world problem: high central glazing allows maximum penetration of daylight with less glazing, balancing heat gain+loss, and lighting needs (although it is less clear specifically how much glass there should be or precisely to place it, thus motivating the use of optimisation). A slight bias East also catches more early morning sunlight. Not illustrated due to space, for **cost**, the magnitudes of the α_k are highly similar, suggesting that optimal solutions should be unglazed and individual cells make equal contributions to cost. This matches with the problem definition, whereby an equal cost is associated with each cell.

7 Conclusion

We can make a step toward explainability for optimisation by metaheuristics by reporting relationships between problem variables and sensitivity of the objectives to them. This brings several benefits:

- Knowing the sensitive variables, solutions can be adjusted for factors not considered by the optimisation, knowing the likely impact on optimality. e.g., odd window shapes made more visually appealing.
- If the returned solutions match the conclusions drawn from the model, the decision maker can have added confidence in the optimality of the results.
- The model can point towards global optima long before the algorithm converges. With the glazing problem, the model suggested the overall glazing shape after the first 1000 of a 5000 solution run. This could show where algorithm has missed the global optimum, and with long-running simulations, make early suggestions of optima: particularly helpful if this indicates flaws in the problem formulation.

We applied the MFM surrogate to a glazing optimisation problem. Analysis of the model coefficients yields simple explanations for the optimisation results. This builds on earlier work [2, 3] showing that MFM coefficients point towards the global optima of benchmark functions. Considerably more needs done to generalise the concept of surrogate model mining to a wider range of problems and representations. In particular, how best to layout the visualisations is crucial. What this work has done is set out the possibility that surrogates can be used as the basis of explaining metaheuristic optimisation results.

References

1. Bethke, A.: Genetic Algorithms as Function Optimizers. Ph.D. thesis, University of Mitchigan, Ann Arbor, MI (1980)
2. Brownlee, A.E.I., Wu, Y., McCall, J.A.W., Godley, P.M., Cairns, D.E., Cowie, J.: Optimisation and fitness modelling of bio-control in mushroom farming using a Markov network EDA. In: Proc. GECCO. pp. 465–466. Atlanta, GA, USA (2008)
3. Brownlee, A., McCall, J., Zhang, Q.: Fitness Modeling With Markov Networks. IEEE T. Evolut. Comput. **17**(6), 862–879 (2013)
4. Brownlee, A.: Mining Markov network surrogates for value-added optimisation. In: Proc GECCO Companion (2016)
5. Brownlee, A., Wright, J.: Solution analysis in multi-objective optimization. In: Proc. Building Simul. and Optim. Conf. pp. 317–324. Loughborough, UK (2012)
6. Brownlee, A.E.I., Regnier-Coudert, O., McCall, J.A.W., Massie, S., Stulajter, S.: An application of a GA with Markov network surrogate to feature selection. Int. J. Syst. Sci. **44**(11), 2039–2056 (2013)
7. Brownlee, A.E., Wright, J.A.: Constrained, mixed-integer and multi-objective optimisation of building designs by NSGA-II with fitness approximation. Applied Soft Computing **33**, 114–126 (2015)
8. Brownlee, A.E., Wright, J.A., Mourshed, M.M.: A multi-objective window optimisation problem. In: Proc. GECCO. pp. 89–90. ACM, Dublin, Ireland (2011)
9. Deb, K., Srinivasan, A.: Innovization: Discovery of innovative design principles through multiobjective evolutionary optimization. In: Knowles, J., et al. (eds.) Multiobjective Problem Solving from Nature: From Concepts to Applications. pp. 243–262. Springer, Berlin, Heidelberg (2008)
10. Hendricks, L., et al.: Generating visual explanations. In: Euro. C. on Comp. Vis. pp. 3–19 (2016)
11. Jin, Y.: Surrogate-assisted evolutionary computation: Recent advances and future challenges. Swarm Evol. Comput. **1**(2), 61–70 (2011)
12. Ribeiro, M., et al.: Why should I trust you?: Explaining the predictions of any classifier. In: Proc SIGKDD Conf Kn Disc & DM. pp. 1135–1144 (2016)
13. Shakya, S., McCall, J., Brownlee, A.: DEUM - distribution estimation using Markov networks. In: Shakya, S., Santana, R. (eds.) Markov Networks in Evolutionary Comp., pp. 55–71. Springer (2012)
14. Tresidder, E., Zhang, Y., Forrester, A.I.J.: Optimisation of low-energy building design using surrogate models. In: Proc. Build. Sim. Conf. pp. 1012–1016 (2011)
15. Urquhart, N., Guckert, M., Powers, S.: Increasing trust in meta-heuristics using MAP-elites. In: Proc. GECCO Comp. pp. 1345–1348 (2019)
16. Valentini, G., Malagò, L., Matteucci, M.: Optimization by ℓ1-Constrained Markov fitness modelling. In: Proc. LION (LNCS 7219), pp. 250–264. Springer (2012)
17. Wright, J., Mourshed, M.: Geometric optimization of fenestration. In: Proc. IBPSA Building Simulation. pp. 920–927. Glasgow, Scotland (2009)
18. Wright, J.A., Wang, M., Brownlee, A.E., Buswell, R.A.: Variable convergence in evolutionary optimization and its relationship to sensitivity analysis. In: Proc. IBSPA BSO 2012, pp. 102–109. Loughborough University© IBPSA-England (2012)
19. Wright, J.A., et al.: Multi-objective optimization of cellular fenestration by an evolutionary algorithm. J. Build. Perform. Sim. **7**(1), 33–51 (2014)