Equipe-Projet Commune TAO
INRIA Saclay et LRI (UMR CNRS 8623)
Bât. 490, Université Paris-Sud
91405 Orsay Cedex (France)

Marc Schoenauer, PhD, HDR
Directeur de Recherche 1ère classe, INRIA

Report on the PhD dissertation submitted by Martin Pilát
August 5. 2013

Martin Pilát's PhD dissertation, entitled *Evolutionary Algorithms for Multiobjective Optimization* is centered on the use of surrogate models in Multi-Objective Evolutionary Optimization. There is no need to emphasize the importance of Multiobjective Optimization, as it is clear that most problems of the real world, if not all, are in fact multiobjective (e.g. optimizing both quality and cost of manufactured parts, duration and cost of industrial processes, ...). Furthermore, whereas Evolutionary Algorithms have demonstrated their efficiency in solving real-world problems (in particular multiobjective ones), the price to pay for their flexibility is a large computational cost, estimated in terms of number of function evaluations, whereas each one frequently take minutes or hours in real industrial applications. Surrogate models aim at decreasing such cost by building on the fly an easy-to-compute proxy for the objective function, thus decreasing the overall cost of the optimization, provided the quality is the solution is preserved. This is the context of Martin Pilát's work, an important hot topic today in Evolutionary Computation, or more generally in Optimization at large.

The dissertation is made of four parts. After the usual Introduction, part II is dedicated to presenting the state-of-the-art, part III introduces Martin Pilát's contributions, and part IV is the concluding section.

After sketching the contents of the dissertation, Chapter 2 in the Introduction states the background, describes several indicators, and rapidly introduces the benchmark functions that will be used in all experiments, though referring to the original publications for their actual formal definitions. Note that, to my taste, this Chapter also lacks some high level description of the research goals of the PhD.

Part II (State-of-the-art) describes in turn the different models for surrogate modeling that have been used in Martin Pilát's work, from Linear Regression (LR) to SVM (and the "kernel trick"), Multi-Layer Perceptrons, RBFs, and Gaussian Processes. Chapter 4 introduces some Evolutionary Multiobjective Optimization Algorithms (EMOAs), and Chapter 5 surveys previous works in surrogate-based Evolutionary Multiobjective Optimization.

In these introductory and state-of-the-art chapters, I was sometimes slightly surprised by the briefness of the presentations, probably because I am used to more self-contained dissertations. For instance, in Chapter 4, the single-objective framework is only sketched, and it is not clear that the version of CMA-ES that is used in the multiobjective case is a specific version of the complete CMA-ES. However, let us be clear: all information that is relevant for Martin Pilát's PhD work is there, or cited with proper references, and this briefness hence does not at all impact the quality of Martin Pilát's work. Also note that more details about some specific methods are also given along the dissertation, when needed, with, similarly, correct and complete references. On another point of view, I was also expecting some a priori discussions about the merits of this or that method at the end of Chapters 3, 4 and 5, that would have nicely introduced the issues that will prove relevant in the contributed work described in the following chapters.

But let us discuss now the most important part of the dissertation, that presents Martin Pilát's contributions. Chapter 6 and 7 are dedicated to the most original and significant contribution, the Aggregated Surrogate Model (ASM) for EMOAs. Inspired by some work by Ilya Loshchilov (one of my former PhD students, to make things clear), the idea of ASM for multiobjective optimization is to build a single function that somehow represents the Pareto dominance relationship, instead of building one surrogate model for each of the objectives as done in all previous works. Martin Pilát's proposition aimed at counterbalance some issues with Loshchilov's work, mainly the sensitivity to the scales of the objectives, and the arbitrary values given in the surrogate models to unknown individuals that could possibly lie closer to the true Pareto front. The main idea here is to use the Euclidian distance to the current Pareto front as the objective of the regression, with one additional parameter that allows to tune the weights of individuals depending on their distance to a given individual, and thus the locality of the model. This allows to use the surrogate model for some local search around each offspring (with some probability), thus enhancing the EMOA with some surrogate-based memetic component.

In Chapter 6, several models are tried (LR, SVM-regression, MLP), and experimentally compared using the benchmark functions introduced earlier, demonstrating statistically significant speedups on several problems, though no clear setting can be said to always perform best. Note that the document also contains an original evaluation of the modeling computational overhead that leads to a nice validation of the approach (w.r.t. computational costs), even for not-so-costly objective functions. On the other hand, maybe some deeper insight into the actual effect of local optimization could have been experimentally investigated, even though the global effects are clear enough.

Chapter 7 demonstrates that these results can be further improved by using another surrogate model to pre-select the most promising offspring from a pool of recently created offspring. A global model is created for this purpose, and again several possible surrogate models are investigated, demonstrating the importance of the choice of this model. Also, more sophisticated approaches allowing the algorithm to select more than one child are introduced. A new way of presenting the results, similar to that used by Lochshilov, is added to the previous hypervolume tables, making it easier to visually compare different algorithms.

Chapter 8 addresses an important issue raised by the previous results, that of model selection: some adaptive mechanism is proposed, that computes several models every time one is needed, and selects the best one according to some error criterion. Four different criteria are experimentally investigated. Results show that the answer is problem-dependent: for some problems, the algorithm can indeed choose the best model, while for other problems there is no clear 'optimal' choice.

Along similar line of thought that tries to relieve the user from educated guesses for different components of the algorithms, Chapter 9 proposes to use an EMOA to tune the hyper-parameters of some classifiers, after multi-objectivization of the problem: two additional measures, closely linked with the classification error, are added as new objectives, and the results demonstrate that this can indeed be helpful. There results, though they use some algorithms proposed earlier in the dissertation, fall a little outside of the main line of research of Martin Pilát, and require to be developed further. However, they also demonstrate the broadness of focus that Martin Pilát has acquired during his PhD.

Overall, Martin Pilát's work represents a clear contribution to the field of surrogate models for EMOAs. All along this PhD work, the experiments are clearly and precisely described, and all results are carefully statistically validated, thoroughly analyzed and discussed. And the different measures that are used for comparative results in the different chapters, even though making the comparisons between the different chapters a little uneasy, only reflect the evolution of the candidate's experience and understanding of his field, and his will to compare to others' results.

My main criticism regarding Martin Pilát's work is that many hyper-parameters of the proposed algorithms are set rather arbitrarily. This is clearly the case for the results in Chapters 6 and 7. And though Chapter 8 addresses partly this issue by choosing the surrogate model on-line, the experiments presented there still use some arbitrary values for their hyper-parameters. This is rather paradoxical, considering the discussion at the beginning of Chapter 9, that rightly states that the performances of all classification algorithms are highly sensitive to the setting of their hyper-parameters, something that is true as well for Optimization Algorithms at large. Furthermore, the technique proposed in Chapter 8 for model selection could also have been used for hyper-parameter tuning.

But it is clear that a PhD has to be completed in some limited time. Hence not all research directions that appear during the course of the work can be explored, and some have to be left for further work ... by the candidate or its followers. In particular, I am aware that a precise comparison with Loshchilov's work, that is briefly mentioned, without any precise result, at the end of Chapter 7, has been done in depth since the dissertation was written.

It should also be mentioned here that the dissertation is very clearly written, easy and pleasant to read. As already said, the bibliography is up-to-date and complete.

In summary, the criticisms expressed above should not hide the fact that this dissertation represents a huge body of excellent experimental work, containing original ideas that do advance the field of evolutionary multi-objective optimization. I hence strongly recommend that Martin Pilát is awarded the PhD of Mathematics and Physics at the Charles University in Pragua.

Marc Schoenauer
Head of TAO, INRIA project-team on Machine Learning and Optimization
Scientific Deputy Director, INRIA Saclay Île-de-France Research Center.