OPPONENT'S REVIEW OF A DISSERTATION THESIS

Author: Mgr. Martin Pilát, MFF UK, Prague
Title: Evolutionary Algorithms for Multiobjective Optimization
Opponent: Petr Pošík, Ph.D., CTU FEE, Dept. of Cybernetics, petr.posik@fel.cvut.cz

The main topic of the dissertation thesis is the use of surrogate models in multiobjective evolutionary algorithms (MOEAs). Surrogate models, which enable to decrease the optimization cost of evolutionary algorithms (EAs), are a hot topic in the evolutionary community. Multiobjective optimization (MOO) is another area where evolutionary (and other population-based) algorithms excel. The crossover of these two areas is thus a very curent topic, and a well-chosen research area with a lot of space for new scientific results and with a potentially high impact to real-world applications.

The text of the thesis spans almost 130 pages, and has a nice logical structure, where the first chapters build the foundations for the latter ones. The first 40 pages contain the introduction, the description of related machine learning models, multiobjective optimization, and the state of the art in surrogate-based MOEAs. The actual contribution is then described in 4 chapters on 80 pages.

Regrading the scientific contribution, in chapter 6, the author first describes the distance-based surrogate model, introduces a weighting scheme which allows for building local models, and shows that the surrogate modeling can greatly decrease the number of needed evaluations to reach a Pareto set approximation of certain level, and that the use of local models can decrease it even further. In chapter 7, the author equips his algorithm with another component, the so called pre-selection, which further improves the results. In general, the proposed algorithm usually does not find a better Pareto set approximation than an ordinary MOEA, but it approaches the real Pareto set much faster, which is IMHO an important contribution. In chapter 8, the author discusses a very interesting topic—the choice of the right model for the given task. He introduces and compares several simple selection strategies, and shows that the use of dynamic model selection improves the robustness of the results.

Finally in chapter 9, the author presents a real-world application of his algorithm: the tuning of parameters of classification models using the multiobjectivization approach. This section is the only one to which I have some objections. The results presented in this section show some advantage of the proposed algorithm, however, they can be much better. First, I think that the parameters chosen for tuning are not selected well, especially in case of the multilayer perceptron (MLP) where the optimizer does not tune the parameters of the model itself, but rather the parameters of the model training algorithm, and if the algorithm is robust enough, then the effects of tuning are not that discernible. Also, the measure chosen for the comparison (the classifier error after certain number of evaluations) does not show the potential of the algorithm; if the number of evaluations needed to find a classifier with certain precision was chosen as the measure, the results may be much better.

In general, the text is written in a very good English, contains only a few mistakes, typos, or missing articles, and is very easy to read. The main point of each paragraph is always very clear, the concepts, ideas, and results presented in the text are nicely documented by the figures and tables. Although it is clear, that the results presented in the thesis are only a fraction of all the results that must have been obtained during the extensive testing, I highly appreciate that the author presents only the main results, and aggregates them nicely in support of his hypotheses and claims, often using statistical hypotheses testing to compare the results. He does not jump to conclusions prematurely, and always looks at the data from several angles. I especially appreciate section 6.2: many authors settle for mere statement that the surrogate modeling is suitable for time-consuming fitness functions, while the author presents some estimates of the time overhead related to the surrogate modeling; the results of this section clearly shows what the time demands of the evaluation function must be so that the surrogate-based algorithm can be better in terms of the wall-clock time.

To summarize, the author proved his ability to work systematically on a given problem using the scientific method, and showed a good potential for further work in research. The results presented in this thesis are interesting and practically usable for researchers and practitioners world-wide, which is exemplified by the numerous publications of the author on various international conferences. I only think it is a big shame that the author has only 1 journal publication because the thesis contains enough material for 3 or 4 journal articles.

Given all other requirements has been fulfilled, in my opinion, the author shall be awarded with the title Ph.D.

Additional questions:

1. On pages 44 and 45, you wrote: "...the memetic operator moves the individuals closer to the Pareto front and hopefully even finds new non-dominated solutions." Does the discovery of new non-dominated solutions really happen during the use of the memetic operator? What proportion of new non-dom. solutions is introduced by the memetic operator, compared to the number of all new ND solutions generated by the whole EA? Does this proportion depend on the used surrogate model?

2. On page 54, you wrote: "...linear regression gives better results than support vector regression and multilayer perceptrons" when used as a surrogate model for the ZDT benchmark suite. You also provide possible explanations for this phenomenon, but one possible explanation is missing. The ZDT benchmark problems contain the Pareto optimal solutions on the search space boundary. When optimizing any linear function on a bounded domain, you always find the optimum on the domain boundary. Isn't this the reason why the linear surrogate model works so well for the ZDT problems?

3. In sec. 8.3, you use your LSPS-MOEA algorithm to show the behavior of various model selection strategies. Are the selectors used in the local search phase only, in the pre-selection phase only, or in both? You also stated that the models used in LS and PS phases shall be different. Is this condition fulfilled, if the same model selector is used in both phases?

In Prague, August 16, 2013 — Ing. Petr Pošík, Ph.D., opponent