Ph.D. Thesis Review

Ph.D. candidate:	Ing. Michaela Drahošová
Thesis title:	Coevolution of Fitness Predictors in Cartesian Genetic Programming
Examiner:	Petr Pošík, Ph.D.
	CTU in Prague, Fac. of Electrical Engineering, Dept. of Cybernetics
	petr.posik@fel.cvut.cz

The submitted Ph.D. thesis deals with the application of fitness predictors in Cartesian Genetic Programming (CGP) using coevolution. CGP is a special form of GP, highly suitable for implementation in hardware platforms. Fitness predictors belong to methods that try to reduce the time complexity of fitness evaluation in evolutionary algorithms. These methods are nowadays studied very intensively. Thus, **the thesis topic is very timely, important, and interesting for the community**.

The thesis has the form of **5 collected articles with a short "introductory" part** which spans about 30 pages. This introduction has a logical structure, is written in a readable way in a good English, and gives a nice and concise overview of the background and contributions of the thesis. The 5 collected articles span about 60 pages, and systematically build on each other.

It is clear that **all parts of the thesis were published at high-end conferences** in the field. Sadly, the articles included in the thesis do not contain any journal publication; nevertheless, a journal submission has been made, and I have no doubts that the manuscript is going to be accepted.

From my point of view, the thesis contains the following **main original contributions**:

- the application of the concept of coevolution of fitness predictors to CGP;
- its adaptation and configuration such that it brings speedups compared to the original CGP; and
- the proposal and evaluation of fitness predictors with variable size (at least in the context of CGP).

To summarize, the thesis, in my opinion, shows that **the Ph.D. candidate is able to systematically and independently perform research** of a chosen problem, and shows 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.

Given all other requirements are fulfilled, I propose to award the candidate the title "Doctor of Philosophy".

Additional questions for a discussion:

1. I miss in the thesis a clear, concise, and complete description of all the differences of your application of fitness predictors to CGP compared to Schmidt and Lipson's application of fitness predictors to tree-based GP. Could you list, describe, and justify the differences?

2. In paper [A], page 49, in table 1 you present the tested values of all the parameters of your algorithm. How did you find your final configuration? Have you done an exhaustive search of all the combinations? Or parameter-wise sweeps? Or a random search?

3. When evaluating the candidate solutions, only a single fitness predictor is used each time. Is the evolving population of fitness predictors really needed? Have you considered the possibility to build a fitness predictor each time anew, using e.g. a kind of greedy constructive procedure which would add/delete a data point to/from the fitness predictor to maximally improve its fitness?

4. The formulas used as the SR benchmarks contain only a few numeric constants, if any. Although I understand that the tuning of numeric constants is not the primary aim of the thesis,

I would like to know the opinion of the Ph.D. candidate and responses on the following questions. What results could we expect if the formulas contained a larger number of numeric parameters? Which GP system (tree-based GP vs. CGP) is better equipped for the optimization of such numeric constants?

In Prague, August 28, 2017 Petr Pošík, Ph.D., examiner