Single-objective optimization can be enhanced by adding auxiliary objectives, but how should we choose the most efficient ones, and when should we use the particular objective? A method designed to solve these issues was proposed in this work. The method is called EA+RL, which stands for an evolutionary algorithm (EA) controlled with reinforcement learning (RL). Further, an overview of existing approaches of objective selection is given and the EA+RL method, as well as the theoretical and practical results, are described.

There are several techniques that involve using some additional objectives in order to enhance performance of EAs. In the multiobjectivization technique [1], all the objectives are optimized simultaneously by some multiobjective algorithms (MOEAs). In this technique the objectives should be specially developed in order to increase the optimization performance. It was shown that adding an inefficient objective leads MOEAs to fail on the considered model problems [4].

The helper-objective approach also involves using MOEAs, but it requires a strategy of choosing the auxiliary objective to be optimized at the current population [2]. The strategy can be either random, or ad-hoc. The random one is general, but it does not take advantage of problem characteristics. At the same time, ad-hoc strategies can be efficient, but they lack generality.

The proposed EA+RL method incorporates auxiliary objectives into a single-objective EA [3]. It requires less computational effort than MOEA-based methods, which makes it more applicable to resource-consuming problems. The selection strategy used in EA+RL is problem independent and it allows to learn some features of the problem as well, thus the method seems to increase both efficiency and generality of the helper-objective approach.

In the EA+RL method, a RL agent interacts with an EA. The scheme of the method is shown in the Fig. 1. The agent selects an objective from the auxiliary objectives and the target one. Then, the selected objective is passed to the EA. The next generation is evolved using this objective as the fitness function. A numerical reward and some representation of the state are passed to the agent. The agent updates its strategy using the obtained information and the process repeats.
The goal of RL is to optimize the total amount of reward. In the EA+RL method, an immediate reward is based on the difference of average target objective values in two sequential iterations. It also can be the difference between the best target objective values in two sequential iterations. Hence, the total reward being maximized is roughly equivalent to the difference between the final and the initial values of the target objective.

It was shown that EA+RL method successfully ignores an inefficient auxiliary objective [4]. Random Mutation Hill Climber adjusted with Q-learning using greedy exploration strategy was considered. The target objective was OneMax and the auxiliary one was ZeroMax. It turned out that EA+RL with an inefficient auxiliary fitness function performs on par with a conventional EA, namely in $\Theta(n\log n)$ fitness function evaluations, where $n$ was the size of the OneMax problem.

It was also shown that the EA+RL method selects an efficient auxiliary objective. The case when the auxiliary objective was a fine-grained version of the target one was considered. A coarse-grained version of OneMax called XdivK was defined. The target objective was XdivK and a single auxiliary objective was OneMax. Exact expressions for the expected running time of RMHC solving the XdivK, as well as of the EA+RL method solving the XdivK+OneMax problem, were constructed. It was shown that the EA+RL method made optimization faster, and the speedup was exponential in $k$, where $k$ is a parameter of XdivK.

The proposed EA+RL method was also successfully used to increase the efficiency of test case generation. It was shown that the proposed algorithm generates a required test in less number of generations than the other considered methods [5].

To sum up, a method that selects between auxiliary objectives in evolutionary algorithms using reinforcement learning was proposed. Asymptotical bounds of its running time for the two different problems were obtained. The method was successfully applied to test case generation.

References