Editorial

for a special section on applications of evolutionary algorithms in chemistry

Already some decades ago in the 1950s in computer science emerged an idea, that for the solution of some problems one should search for an inspiration in nature. Classical deterministic approaches work wonderfully for some types of tasks. However, when one either does not know exactly the rules and laws of nature which are necessary to solve the task, or the complexity of the problem is too big, one has to resort to less explicit methods involving random features.

Two basic inspirations were found in nature: The modeling of brain processes brought artificial neural networks, used mostly in classification and prediction tasks, and the modeling of Darwinian evolutionary forces created evolutionary algorithms used mainly in optimization. The random features and a certain lack of strict logic were so uncongenial to most of scientists and engineers, that it took them some decades to recognize the efficiency of these methods.

But when the change came, it came rapidly. While a decade ago papers on application of evolutionary algorithms in chemistry could be counted on fingers, nowadays they are produced in hundreds per year. Applications range from design of new molecules, similarity search, search for conformation and protein folding, docking of molecules, QSAR, to chemometrics and spectra analysis and crystallography. Basically, any complex function, especially one with a lot of input variables, is suitable for optimization by evolutionary methods.

The selected papers in this issue deal mainly with position of atoms (whether in crystals, clusters or conformations of molecules). This selection is rather restricted, nevertheless it shows the mainstream direction of applications in chemistry.

There is a whole bunch of methods covered under umbrella term of evolutionary algorithms. The best known are the genetic algorithms, genetic programming, evolutionary strategies, and evolutionary programming. One can also merge these methods with other optimization methods like simulated annealing, tabu search, hill climbing, cellular computations, etc. The basic feature of evolutionary algorithms is their dealing with a group of possible solutions at once, each solution competing for existence with other solutions. Better solutions have a greater chance to be kept in memory and multiply, but the process has some random features. The solutions also undergo small random changes during the course of algorithm. Some of evolutionary algorithms also use so called crossover, where two solution are combined to create hopefully a better solution.

Theoretic foundations of all these methods are murky despite an enormous effort of researchers; when some theorems are proved, the restricting conditions are usually so fierce, that the result is useless for practical purposes. All these methods therefore give a big space for enthusiastic experimentation, changing internal control parameters of the methods, merging the methods and creating new variants. None of the methods can be declared as the best one generally, though there should be best method for a specific task. Which of the methods is best for which task is left for experimentation. The input variables can be either discrete or real variable. Discrete input is used manly in design of new molecules, when we establish, whether and what kind of atom or connection should be included and in selection of most substantial features - should the feature be considered or discarded as insubstantial? Real variables are used in all the selected papers dealing with applications. In all the selected papers the input variables are restricted to position of atoms and/or torsion angles of bonds. While in canonical genetic algorithms real variables were coded as a binary numbers and handled accordingly, there seems to be a general agreement, that in chemistry for problems dealing with positions of atoms it is more useful to deal with real variables directly.

Since evolutionary algorithms deal with many solutions at once, they are suitable not only for finding the best solution, but also for finding several suitable solutions. This is very handy when we look e.g. for all possible conformations. To keep the best solution from overtaking the whole population of solutions, a couple of new techniques is described in this issue in papers Predatory genetic algorithms by Manby et al. and in An Evolutionary Algorithm with Local Search and Classification for Conformational Searching by Frey.

The present special section of the issue is an attempt to show the current trends and typical applications of evolutionary algorithms in chemistry.

Bratislava, October 1998

Jiří Pospíchal and Vladimír Kvasnička