

AN OPTIMIZATION OF SET OF FEATURES BY GENETIC ALGORITHM

Biriukova M.M.

Postgraduate, Sumy State University

e-mail: margo_biriukova@ukr.net

An optimization method of set of features under information-extreme intellectual technology based on constructing an optimal feature space splitting into classes of equivalence during control system training is considered.

Genetic algorithms are stochastic search algorithms inspired by the biological phenomena of genetic recombination and natural selection. They simulate the evolution of objects representing solutions to a given problem. Unlike other search techniques, they simultaneously process a set of solutions and require no specific knowledge about the problem space to successfully search for good solutions. These characteristics make genetic algorithms[1] applicable to various search, optimization and machine-learning tasks. They have proved robust and efficient in solving complex problems, including learning to control dynamic systems.

Was performed optimization of set of features for estimation of student knowledge with a genetic algorithm in the framework of information-extreme intellectual technology.

The so-called “fitness function” is used to assign a quality measure to the solutions. The algorithm starts with a population of randomly created solutions, and iteratively improves them in steps called “generations”. In each generation, the solution vectors undergo selection and variation. In the selection phase, solutions generate descendants selectively according to their fitness values. The higher the fitness, the higher the probability of generating descendants, i.e. copies of the solution vectors that form a new population.

Fig.1. shows the changes fitness function, which used the average quality measure of the fittest individuals for 100 generations optimization.

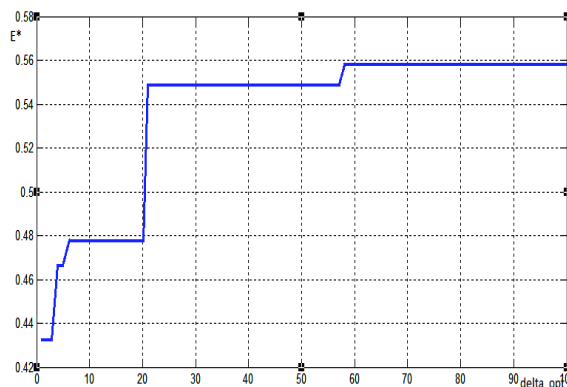


Fig.1. – Graph of the changes fitness function

Fig. 1 has demonstrated that the optimal features set were implemented on 58 generation and the final result of average quality measure represents $E^*=0,5579$. The optimal features set Σ^* were implemented by removing the 14 signs of the original set.

CONCLUSION

Thus, the proposed genetic algorithm of selection of features recognition considering management system of distance learning in the framework of information-extreme intellectual technology has greatly reduced the output of set of features and has increased the average quality measure.

REFERENCES

- [1] Bogdan Filipic, Tanja Urbancic, Viljem Krizman. A combined machine learning and genetic algorithm approach to controller design. – Engineering Applications of Artificial Intelligence 12 (1999) 405.