

# Incorporating Linkage Learning into the GeLog Framework

Tim Fühner and Gabriella Kókai

Various modifications were applied to the *GeLog* framework in order to significantly enhance its abilities<sup>2</sup>. *GeLog* combines two approaches, inductive logic programming and evolutionary computing [1]. Inductive logic programming (ILP) aims at detecting correlations of pieces of data [2]. This is done by inducing over a data set whose objects' relation is already known. Thus, a hypothesis that matches this training data is searched for, assuming that all other data instances are correctly classified by the hypothesis. *GeLog* searches the hypothesis space by means of a genetic algorithm, an optimization technique which utilizes recombination and selection as observed in nature[3]. Also, the data representation resembles representations found in genetics; the objective representation (phenotype) differs from its encoding in the search space (genotype), which most commonly is a string of characters of a discrete alphabet (genes).

Investigations on the dynamics of genetic algorithms have shown that tight linkage, i.e. the clustering of genes which contribute to the quality of the solution, is an important issue [4]. It was long assumed that individuals in genetic algorithms would eventually evolve towards tighter linkage. However, later investigations demonstrated that selection counteracts linkage learning. This made it necessary to tame the forces of selection [5, 6].

After different approaches that incorporate linkage learning were thoroughly reviewed and compared, the modifications necessary to employ linkage learning into the *GeLog* system were implemented. Furthermore, techniques that decelerate selection by maintaining a high level of diversity have been investigated in order to profit from the effects of linkage learning.

Finally, it could be shown that two experiments, which both proved to be hard for the original version of *GeLog*, can be solved using the enhanced version. The excellent results achieved by the modified version of *GeLog* show that the system has improved significantly. These results will have a significant impact on our future investigations on linkage learning and building block processing in genetic algorithms.

## References

- [1] Gabriella Kókai. *GeLog—A System Combining Genetic Algorithm with Inductive Logic Programming*. In *Proc of the International Conference on Computational Intelligence, 7th Fuzzy Days LNCS*, pages 326–345, Dortmund, 2001. Springer Verlag.
- [2] Nada Lavrač and Sašo Džeroski. *Inductive Logic Programming: Techniques and Applications*.
- [3] John H. Holland. *Adaptation in Natural and Artificial Systems*. PhD thesis, University of Michigan, Ann Arbor, 1975. Ellis Horwood, New York, 1994.
- [4] Dirk Thierens and David E. Goldberg, *Mixing in Genetic Algorithms*. In *Proceedings of the 5 th International Conference on Genetic Algorithms*, pages 38–55, San Mateo, CA, 1993. Morgan Kaufmann.
- [5] David E. Goldberg, Bradley Korb, and Kalyanmoy Deb. *Messy genetic algorithms: Motivation, analysis, and first results*. *Complex Systems*, 3(5):493–530, 1990.
- [6] George Harik. *Learning linkage to efficiently solve problems of bounded difficulty using genetic algorithms*. PhD thesis, Ann Arbor, Michigan, 1997.

---

<sup>2</sup>This work is supported by the grants of Bayerischer Habilitationsförderpreis 1999.