
A Markov Chain Analysis of Fitness Proportional Mate Selection Schemes in Genetic Algorithm

Chien-Feng Huang

Center for the Study of Complex Systems, 4477 Randall Lab.
University of Michigan, Ann Arbor, MI 48109
cfhuang@fiore.physics.lsa.umich.edu

In the research of Genetic Algorithms (GAs), many models focus on problems where each individual's fitness is independent of others. In (Huang, 2002a), simple models that implement mate selection in GAs were introduced to model interdependent fitnesses of population members. They have been studied by the Schema Theorem and some empirical results. In this paper, I conduct a Markov chain analysis to further investigate these models, based on the Markov model developed by Nix and Vose (1992). Although such models quickly become unwieldy with increasing population size or string length, they provide important insights into how mate selection plays a crucial role in GA's search power, and thus serve as guidelines for studying more realistic problems.

In (Huang, 2002a), the implementation of the four mate selection schemes in the selection-for-mating step of a simple GA is:

During each mating event, a fitness-proportionate selection is run to pick out the first individual. Then the Hamming distances of all population members to the first individual are calculated. The actual mate of the first individual is chosen according to the following four different schemes:

Maximum Similar Mating (MSM): The population member whose Hamming distance is the smallest is selected for mating.

Proportional Similar Mating (PSM): The probabilities of population members being selected are *reversely* proportional to their Hamming distances.

Proportional Dissimilar Mating (PDM): The probabilities of population members being selected are proportional to their Hamming distances.

Maximum Dissimilar Mating (MDM): The population member whose Hamming distance is the largest is selected for mating.

As indicated in (Huang, 2002b), the way of calculating the probability of the second individual being selected as a parent is the only part that needs to be reconsidered in the Nix and Vose model. Please see (Huang, 2002b) for the detailed derivation for these four mate selection schemes.

With the modified Markov models, I use the GAFO (GAs being used for function optimization) idea developed by De Jong et al. (1994) to examine the effects of different mate selection schemes on GA's performance. As an example, I use the test function $f(y) = \text{integer}(y) + 1$. Due to the computational limitation, I use the simplest possible case, string length 2, to proceed the investigation.

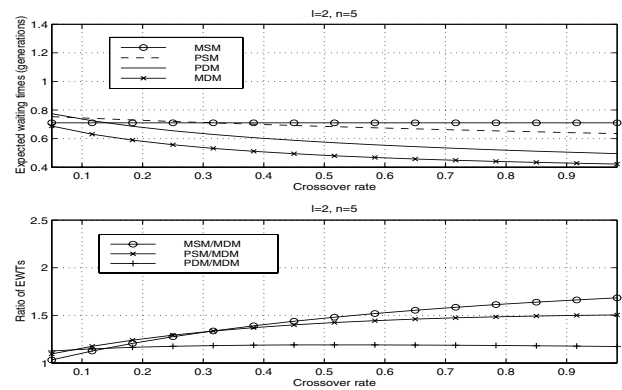


Figure 1: Interacting effects of mate selection and crossover.

Figure 1 shows the results obtained, based on crossover rates ranging from 0.05 to 1, mutation rate .1, and population size 5. The top plot is for the exact expected waiting times (EWTs) to the optimum and the bottom plot shows the ratios of the EWTs for the MSM, PSM, and PDM GAs to that for the MDM GA. One can see that MDM generally has the least EWTs than the other three. In particular, the dissimilar mating schemes demonstrate increasingly improved performance as crossover rate increases. This shows that both proper mate selection and crossover must operate together to enhance the power of simulating information exchange in GA's population.

Reference

- Huang, C.-F. (2002a). A Study of Fitness Proportional Mate Selection Schemes in Genetic Algorithms. *Technical Report CSCS-2002-002*. Ann Arbor, MI: University of Michigan, Electrical Engineering and Computer Science.
- Huang, C.-F. (2002b). A Markov Chain Analysis of Fitness Proportional Mate Selection Schemes in Genetic Algorithm *Technical Report CSCS-2002-003*. Ann Arbor, MI: University of Michigan, Electrical Engineering and Computer Science.
- Nix, A. E. and Vose, M. D. (1992). Modelling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence* #5, pp. 79-88.