A Study of Fitness Proportional Mate Selection Schemes in Genetic Algorithms

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 this paper, I introduce simple models to study mate selection in the context of GA. Allowing individuals to search for mates is an approach to model interdependent fitnesses of population members. This introduces another source of selection pressure, and the resulting GA hence forms a more complex system in which individuals' fitnesses depend on both the environment and other population members. I will show that mate selection plays a crucial role in GA's search power.

In this paper, four mate selection schemes in the selection-for-mating step of a simple GA are proposed as follows:

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.

The testbed employed is an incompatible small royal road function IS_1 as shown in Table 1.

Table 1: Incompatible small royal road function IS_1.

s_1 = 11111***************; c_1 =10
s_2 = 00000***************; c_2 = 5
s_3 = *****11111**********; c_3 =10
s_4 = *****00000**********; c_4 = 5
s_5 = **********11111*****; c_5 =10
s_6 = **********00000*****; c_6 = 5
s_7 = ***************11111; c_7 =10
s_8 = ***************00000; c_8 = 5

This function involves a set of schemata $S = \{s_1, \ldots, s_8\}$ and the fitness of a bit string x is defined as

$$F(x) = \sum_{s \in S} c_s \sigma_s(x),$$

where each c_s is a value assigned to the schema s as defined in the table; $\sigma_s(x)$ is defined as 1 if x is an instance of s and 0 otherwise. In this function, the fitness of the global optimum string (20 1's) is $10 \times 4 = 40$.

The experiments performed here are based on one-point crossover rate 1, mutation rate 0.005, and population size 20 over 50 runs. Figure 1 shows the averaged best-so-far curves on function IS_1 for the four mate selection strategies.[1]

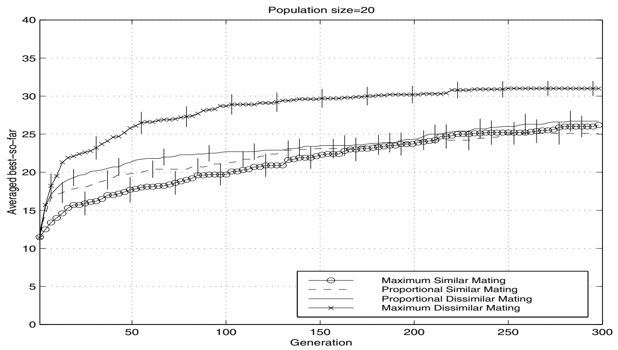

Figure 1: Best-so-far performance on IS_1.

We can see the maximum dissimilar mating results in better improvement than the other three. In (Huang, 2002) it is shown that, by suppressing hitchhiking and the founder effect, the maximum dissimilar mating retains more genetic variation in the population. The further exploration of the search space for the GA thus yields a better best-so-far performance.

Reference

Huang, C-F. (2002). *A Study of Mate Selection in Genetic Algorithms.* Doctoral dissertation. Ann Arbor, MI: University of Michigan, Electrical Engineering and Computer Science.

---

[1] The vertical bars overlaying the metric curves represent the 95-percent confidence intervals.