Impact of Crossover Bias in Genetic Programming

Nicholas Freitag McPhee, M. Kirbie Dramdahl, David Donatucci
Division of Science and Mathematics
University of Minnesota, Morris
Morris, MN USA-56267
{mcphee, dramd002, donat056}@morris.umn.edu

ABSTRACT

In tree-based genetic programming with sub-tree crossover, the parent contributing the root portion of the tree (which we refer to as the *root parent*) often contributes more to the semantics of the resulting child than the other parent (the *non-root parent*). In previous research, we found that when the root parent had greater fitness than the non-root parent, the fitness of the child tended to be better than if the reverse were true. Here we explore the significance of that asymmetry by introducing the notion of *crossover bias*, which allows us to bias the system in favor of having the more fit parent be the root parent. To better understand the impact of this bias, we implemented several levels of crossover bias, including 0% bias (root individual chosen randomly, as in traditional sub-tree crossover), 100% bias (the stronger parent is always chosen to be the root parent), 50% bias (bias implemented in half the cases, and the other half chosen randomly), and reverse bias (the weaker parent is always chosen as root parent).

We applied crossover bias to a variety of problems. In most cases we found that using crossover bias either improved performance or had no impact. Our results do, however, indicate the possibility that crossover bias may increase selection pressure and premature convergence – undesirable behavior, as it encourages a genetic programming run to arrive at a solution too quickly, in the process potentially excluding more accurate solutions for a more generalized one.

Our results also demonstrate that the effectiveness of crossover bias is somewhat dependent on the problem, and significantly dependent on other parameter choices. In particular it appears that crossover bias has the largest impact when selection pressure is weaker, and the differences in the fitness of the parents is thus likely to be larger. We also found that the use of elitism reduced the influence of crossover bias. It's possible that crossover bias acts to some degree as an "elitism" operator, making it more likely that the semantics of more fit individuals are copied into the next generation; thus if traditional elitism is being employed this effect is less visible. Another possible explanation for this is that if the most fit individuals are automatically being carried over, there is perhaps less need to produce new, fitter individuals via crossover, reducing or even eliminating the usefulness of crossover bias. Other factors which we found to have potential impact on the effectiveness of crossover bias were tournament size, population size, and possibly the difference in parental fitness.

GECCO'15, July 11-15, 2015, Madrid, Spain.