Effects of Modifying the Expressive Power of the Language of Child Producing Programs in Autoconstructive Genetic Programming

Neil Singh (Y4628661)

Extended Abstract of Open University MSc Dissertation Submitted 6th March, 2012

Introduction

Genetic programming (GP) is a computer science technique for automatically generating programs that solve a given problem. This project focuses on a form of GP called *autoconstructive GP* (AGP).

GP is inspired by biological evolution and potential solutions (programs) are known as *individuals*. It begins by creating a population of random individuals. Then the initial population is evolved by *genetic operators* that reproduce and mutate individuals, and some of the *fittest* individuals are *selected* to survive to the next population. This evolutionary cycle continues until an adequately fit individual is found or a fixed number of cycles passes.

The choice of genetic operators is set by a GP system's parameters which are typically predetermined by a GP researcher based on the results of preliminary experiments. This is done because a GP system's performance depends on the combination of its parameters and the problem being tackled. However, due to the large number of possible parameter values and complex interactions between parameters, the preliminary experiments can be laborious and are not guaranteed to result in optimal parameters. Consequently, some researchers have investigated automating parameter setting, and AGP was developed as a result.

In AGP, instead of using pre-determined genetic operators, individuals are responsible for problem solving and for producing their own children. Each individual has an in-built genetic operator whose mechanism is evolved along with the individual, the idea being that the genetic operators will evolve to suit the problem.

Implicit in AGP is the concept of *reproductive competence*. Early on in an AGP run, not all individuals will successfully reproduce and random individuals have to be created to fill the next population. Accordingly, an AGP system is said to be reproductively competent when its individuals can fill the next population with children and no random individuals need to be created. Preliminary work, carried out by other researchers, suggests that AGP systems first have to achieve reproductive competence and only then do they begin to evolve individuals to solve the problem.

This project investigates the effects of modifying the expressive power — the set of enabled instructions in a program's programming language — of the child

producing programs of individuals in an AGP system. The hypothesis is that the expressive power determines the space of all possible child producing programs, and, consequently, it should be possible to reduce the expressive power to change the space to contain a higher proportion of competent child producing programs.

Method

The project's research method involves carrying out a set of controlled laboratory experiments in the form of computer simulations. An AGP system is implemented that runs in one of three *execution modes* and which is set to solve one of two problems (*problem 1* and *problem 2*). Both of the problems task a system with evolving programs that output specific values for predefined inputs.

In execution mode 1, individuals are represented by a single program responsible for problem solving and child production. It is this mode which most closely resembles the operation of AGP systems of research works that inform this project. In modes 2 and 3, individuals are represented by two programs; one for problem solving and another for child production. Furthermore, mode 3 reduces the expressive power of the language of child producing programs by disabling instructions identified to be least frequently used by competent child producing programs in an execution of the AGP system in mode 2 and on problem 1.

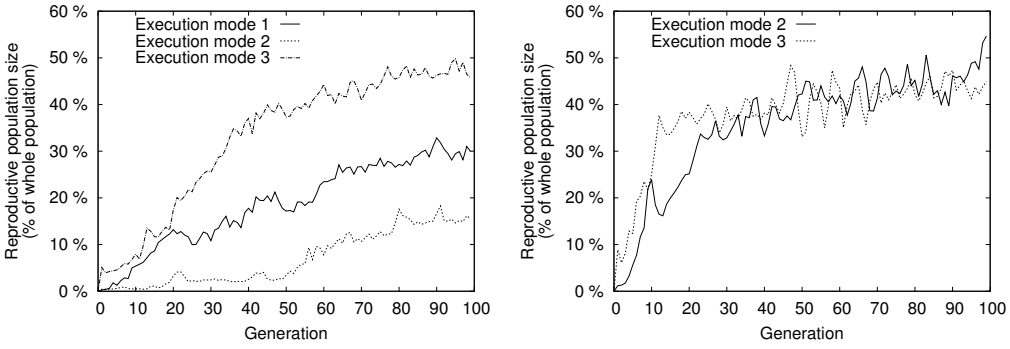
Three experiments are carried out. The first investigates the side effects of representing individuals with two programs instead of one. The second investigates the effects on reproductive competence caused by reducing the expressive power of child producing programs. The third attempts to generalise on the results of the second experiment by repeating that experiment except with another problem; the first and second experiments tackle problem 1, the third tackles problem 2.

Results

The following figures plot a measure of reproductive competence — the percentage of individuals in a generation's population that successfully produce children — against the generation number for experiments 1 & 2 (left) and experiment 3 (right).

The results of experiments 1 and 2 show that the switch from one program per individual (execution mode 1) to two (execution mode 2) harmed reproductive competence, but that the loss was more than made up for by reducing the expressive power of child producing programs (execution mode 3).

The results of experiment 3 show little difference between execution modes 2 and 3.



Plots of reproductive population size versus generation for experiments 1 & 2 (execution modes 1, 2 & 3 on problem 1, left) and experiment 3 (execution modes 2 & 3 on problem 2, right). Values averaged across 20 independent runs for each execution mode.

Analysis

The primary conclusion drawn from the project's results is that the expressive power of the language of child producing programs does affect the reproductive competence of an AGP system. Furthermore, the results hint that the space of reproductively competent child producing programs may be dependent on the problem being tackled, and this is why no difference was seen between execution modes 2 and 3 of experiment 3.

However, the project's conclusions are tentative because, due to practical limitations, too few runs of the AGP system were carried out for the results to be statistically significant.

Discussion

The implication of the project's tentative conclusions is that the expressive power of the language of child producing programs in an AGP system can be considered as a system parameter which can be tuned to improve performance, but that the parameter needs to be tuned on a problem by problem basis.

The project's outcome is interesting but more work is needed to substantiate and expand on it. Future research areas could involve investigating the relationship between the space of reproductively competent programs and the problem being solved; whether classes of problems exist for which the space of reproductively competent programs is similar; and whether the expressive power could be reduced in a dynamic and self-adaptive way.