Sam Trenter

Evolutionary Algorithms
Homework 1
In the following homework, I am asking you questions and I want you to think through the answers. Definitions of terms that do not directly answer the questions are not worth points. If you do not answer the question that I am asking, you will receive a zero.

> 1. Suppose you developed a new fitness function for onemax that returned the size of the string if the string was all 1s but returned 0 in all other cases. What would the effect be if you swapped this fitness function for the one used in your first project? What type of search would it
> become?

That fitness function would basically make your "GA" a brute force search. It would be like that because there is no "evolutionary pressure" to keep better parents alive and all parents would be treated equally. What I would guess would happen is your population would randomly swap bits until one got the full stringsize filled with "1"s .

> 2. Usually when we initialize a population in a GA, we do so with uniform distribution (ie: at random). However, it is also possible to initialize the population with higher-quality individuals than simply random ones. What benefits might we get from this? What negatives could occur?

So a benefit is that you could greatly speed up your GA by giving it a individual that is already fairly good at solving your problem, and that can make it much easier to generate new individuals who solve the problem slightly better. Two issues that I can see with this approach are that you could create a lack of diversity in the population because all individuals may become based of the "seeded" individuals. The second issue is that this could help you fall into a local minima/maxima. (depending on whether you are maximizing or minimizing something)

> 3. Suppose you had a population of size 50 in a genetic algorithm. You use tournament selection with tournament size of k=8. What effect would this have on diversity of the population compared to a tournament size of k=2. Explain your answer.

This should lead to lower diversity. for example if out of 50, 10 individuals were genetically similar and solved to problem well. every time you select 8, you are very likely to select a one out of the ten genetically similar individuals. This means that when you do combine their genetics by whatever means you choose your new individuals may be very similar to their parents. Now with k=2 you still have a chance of getting one of those ten individuals but you are much more likely to grab an individual who is not in the 10 very genetically similar individuals.

We can actually show the extreme of this, what if you set k=50? if you can only select each individual once you will only select the best individual, meaning that the next generation will just be copies of the best individual. This would be near zero diversity (not zero because of possible mutation).

> 4. In many GAs, instead of the population being in one large group, the population is instead spread out amongst smaller separate populations that only have limited transfers between them. What is this model called and in particular what is the benefit of it? Can you describe any
> downsides?

they are called MGA or MPGA, for multi population genetic algorithm.

I'm guessing a downside is that it may take much longer to train because you are not using the full "power" of your population to head towards a solution. A benefit should be that you should be much more adverse to becoming completely stuck in a local minima/maxima. This is because one population may get stuck in one, but you have other populations that may not. This would also help with increasing diversity, much like how the Galapagos finches each adapted to fit the needs of their specific island, you separate populations will each probably find a unique way to solve the problem.

> 5. Your GA seems to often lose the best individuals during recombination. You want to ensure that at least some of these best solutions are kept around. What replacement method would you use? Explain your answer.

Fitness based replacement, simply keep the top "N" percent and replace the rest, or you can select an "N" amount to replace every generation. I think this can also be called elitism based replacement. Basically you just replace the worst performers in your generation. You really just want to make sure that some top performers survive to keep their genes moving on.

> 6. Intron’s are sections of code that do nothing. The rise of intron’s in a GP population can lead to bloat. Suppose you developed an algorithm to go through GP code and remove all the intron’s. Please describe what you think the effect of doing so would be on your evolutionary algorithm when solving problems.

It would probably be bad depending on the problem we were trying to solve. Introns only provide no value in the sense that they currently don't contribute to an individuals fitness. The important part about introns is that they have massive potential when "activated" to provide some sort of benefit. Introns act like the thousands of "unused" genes humans have that seem useless now but may become very helpful during a mutation.