1.0 Introduction

At its inception Holland's genetic algorithm (GA) featured fixed length strings which optimized static fitness functions [Holland 1975, 1992]. As GA research matures and GAs are applied to more varied and more difficult problems there is a necessity to develop techniques that better address dynamic fitness functions. This seems a natural progression, but it is not always easy to know in what direction these algorithms might be best extended to be generally applicable.

Two areas of interest are variable size representations and non-coding regions. It has been conjectured that non-coding regions, since they would not be subject to selection pressure would be capable of mutating freely and thus chance onto useful new schemata. A prerequisite of the ability to search the space freely is the lack of selection pressure; if care is not taken in their design, GAs characteristically lose variability through convergence.

There are two common usages of the term "convergence" in GA research. One meaning is homogeneity across the population, meaning all individuals are identical; the other is discovering the solution, as in "...the maintenance of a diverse population after convergence is reached." [Raich & Ghaboussi 1997]. The former meaning will be used for the remainder of this paper.

The loss of diversity in any region of the chromosome incapacitates crossover in that region and prevents any exploration by mutation except in the immediate vicinity (in Hamming distance) of the converged values [Jones 1997]. While convergence in some regions of a population's DNA is clearly advantageous, e.g. that portion which codes for the Krebs cycle, total convergence across the population is precarious except in a static environment. The phrase "premature

convergence" [Schaffer et al. 1991, Davidor 1991, Sakamoto & Goldberg 1997] seems to imply that convergence is desirable but that it should somehow be delayed. In fact, the phrase is usually used to mean that a GA has converged to something different from the desired value. However, when a GA is being used in the context of an unpredictable dynamic fitness function, convergence (meaning loss of genetic diversity) is never desirable. On the contrary, a population which has converged has lost a tremendous amount of potentially valuable information, and most of its capacity to adapt.

1.1 Two challenges for adaptive systems

Holland describes two dilemmas for any adaptive system; stability/plasticity and exploration/exploitation. The plasticity/stability dilemma is the challenge of being able to adapt to new situations without losing the ability to handle previously mastered situations. This is a famous and widespread difficulty for back-propagation nets, which for all their popularity remain abysmal at incorporating new information without needing retraining on old information; this characteristic has been dubbed "catastrophic forgetting" [French 1992, Robins 1995]. The exploration/exploitation dilemma is the tension between using information already acquired to exploit the current environment versus exploring the environment to discover better strategies for exploitation in the future. If a system spends too much energy exploring, it may not survive to use the knowledge it gains; on the other hand if it spends too much energy attempting to profit, it may fail to discover some much better strategy and so fail in the long run. This paper reports experiments which demonstrate that the addition of non-coding regions may ameliorate both dilemmas.

1.2 Chemostats and swapping in bacteria

It has been found that *e. coli* bacteria variably express the information in their chromosomes [Adams 1995]. Bacterial DNA has an active center section in which protein synthesis is carried out and inactive distal regions where it is not. The notion is that the organism cannot afford to manufacture every protein that it might ever need; it is more efficient to manufacture some subset of all those it codes for -- the trick is to decide which to manufacture when using negligible information processing. According to Adams [see also Rosenzweig *et. al.* 1994], when the bacteria is well fed, say with a steady supply of lactose, it simply eats and reproduces. When it is starved, by, for instance, replacing the lactose with glucose, it copies sections of DNA from its inactive to its active region. When it swaps in a section that codes for a protein that metabolizes glucose and so is again well-nourished, it goes back to eating and reproducing. This is a reasonably simple procedure (remarkably similar to paging in a virtual memory system) that can provide great benefits.

The inclusion of regions which are not expressed, commonly called introns, has a number of advantages. It has been shown to reduce schemata loss due to crossover [Levenick 1991]. It is also thought to provide a mechanism for evolution in general by allowing intron regions to vary widely and so search more of the space [Gould 1989]. In organisms which have the capability to swap introns into an active region (thus converting them to exons), it greatly enhances flexibility. These latter two mechanisms could provide solutions to both of Holland's dilemmas.

Swappers: Introns promote flexibility, diversity and invention.

2.0 Experiments

Inspired by Adams' description of *e. coli,* I decided to model these organisms, admittedly at a very crude and oversimplified level, by implementing a GA with the capacity to swap segments into an active region. I used this GA to test the following hypothesis:

Hypothesis 1: In a dynamic environment, swapping will confer adaptive advantage.

Hypothesis 2: Swapping will help retain diversity.

Hypothesis 3: Swapping can retain diversity in non-coding regions indefinitely.

2.1 Experiment 1: Demonstration of the adaptive advantage of swapping.

A very simple dynamic fitness function was designed to model a chemostat where two nutrients, lactose and glucose, were available in different quantities at different times. The total amount of nutrient available was a constant. Namely, input(lactose)+input(glucose)=2001 at all times. Each experiment began with input(lactose)=2000, input(glucose)=1. This fitness was maintained for 5 generations. Then 1/5 the sum (namely, 400) was added to the one and subtracted from the other each generation until input(lactose)=1, input(glucose)=2000. This fitness was then maintained for 29 generations. Then the values were adjusted by the same process back to input(lactose)=2000, input(glucose)=1 and that level was maintained for 29 generations. This oscillation was repeated throughout the run (see Figure 1).

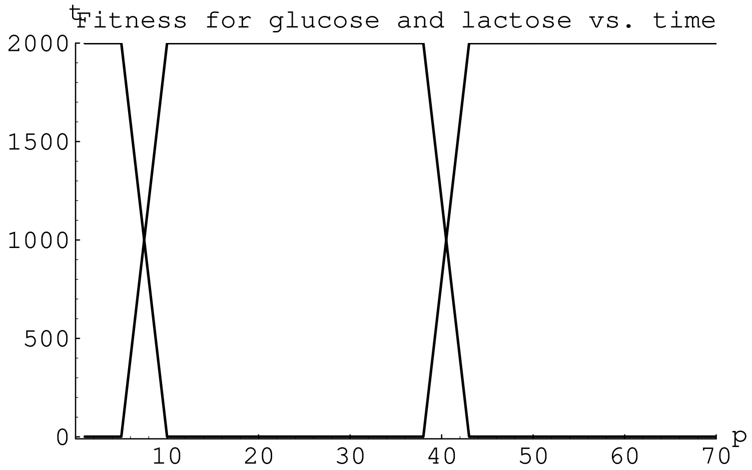

Figure 1: A graph of the concentration of (and thus contribution to fitness of) two nutrients, glucose and lactose, over time. This graph illustrates one cycle; the same cycle is repeated indefinitely.

The schema representing the string that coded for the protein (or metabolite) that allowed the bacteria to metabolize lactose was arbitrarily assigned the value 00000000, which will be denoted 0; and that for glucose , 11111111, denoted 1. An individual with the metabolite for lactose in the first region and one for glucose in the second would thus have the genotype 0000000011111111, and be denoted 01. Individuals were of one of two types; those with the capacity to swap pieces of their chromosome into an active region, called Swappers, and those without this capacity, called Standards.

2.1.1 The GA used for experiments

The GA used was based on Holland's original algorithm with a few enhancements to allow modelling bacteria with the capacity to swap regions. Experiments were conducted using a population of size 5000, fitness-proportional selection of 2% of the population, no interbreeding between Swappers and Standards, replacement of the worst individuals and a mutation probability of 1% per bit in new individuals. Crossover was performed bitwise with a 1%

probability of crossing at each bit (Table 1 summarizes these settings). The actions performed each generation were: swapIfHungry, select, breed, replace, and evaluate.

GA mechanisms	
population size	5000
initial population	11, 01, 10, and 00 in equal numbers
selection	fitness-proportional, 2%/generation
replacement	worst
representation	2 segments of 8 bits
operators	mutation and variable crossover
crossover	bitwise, 1% per bit
mutation	1%/bit new individuals only
reproduction only between two Swappers or two Standard individuals	
fitness function	variable, with noise <10

Table 1: Parameters and methods for the GA used in Experiment 1.

2.1.2 Swapping

Swapping, in Experiment 1, worked as follows. Like Standard individuals, each Swapper had two regions. Unlike Standard individuals, Swappers only produced metabolites in one region, called the active region; the other region was an intron, i.e. it was not expressed as a protein. The fitness of a Swapper was determined by the value of its active region. If a Swapper individual got nearly no nourishment during a generation, it swapped (i.e. exchanged) the string in its active region with that in its inactive region.

2.1.3 Fitness Function for Standard individuals

The fitness of Standard individuals was the sum of the contributions of the two regions of their chromosome. Table 2 lists the fitness function if the payoff were 400 for lactose, 1600 for glucose, and 1 for all other strings.

Region 1	Region 2	Standard fitness
lactose	lactose	800
lactose	glucose	2000
glucose	lactose	2000
glucose	glucose	3200
lactose	xxxxxxxx	401
xxxxxxxx	glucose	1601
xxxxxxxx	xxxxxxxx	2

Table 2: Fitness for Standard individuals. Here xxxxxxxx is any sequence not coding for the lactose or glucose metabolite.

2.1.4 Fitness Function for Swapper individuals

The fitness of a Swapper was computed differently. The leftmost region of the Swappers was considered the active region and fitness was calculated as a function of this region. The question now arises: What should the production rate of the active region in a swapper be? The purpose of this experiment was to demonstrate the adaptive advantage of swapping by comparing the proportion of Standard and Swapper individuals in the steady state. Thus, the experiment's outcome may hinge on the choice of the production rate for the active region. Recall that we are modelling the production of two metabolites; one that metabolizes lactose and one that metabolizes glucose. Perhaps the active region in a Swapper should have the same production rate as one of the two active regions in a Standard individual (i.e. the ratio of Swapper to Standard production rates per region is 1:1); this would set the total rate of Swapper metabolite production to one half that of a Standard individual. Or perhaps it should be the sum of the production of the two regions (i.e. the Swapper:Standard production rate per region is 2:1); this

would make a Swapper capable of producing the same total amount of metabolite as a Standard.

Likely some value between 1 and 2 would be reasonable, as it is sensible that there would be some inefficiency in using only one as opposed to two production regions; and it is also likely that a cell with only one active region would be able to produce more than half as much metabolite a cell with two since some factor besides the number of coding regions likely limits total metabolite production. There are many possible constraints or limiting factors for total production capacity for a cell. These include: 1) input of raw materials across the cell wall; 2) available energy; 3) transcription speed; and, 4) geometric considerations. Depending on which of these (or some other) constraint was the limiting factor, different ratios of Standard:Swapper production rates would be sensible. For instance if bacteria have a single speed for transcribing DNA, and neither energy or raw materials ever slow the manufacture of metabolites, then the appropriate ratio would be 1:1. Or, if transcription speed could be increased if only enough of the necessary molecules could be transported to the transcription site, and those molecules never got in each others way, then a ratio of 2:1 would be more appropriate. Table 3 shows the minimum (1:1) and maximum (2:1) values for Swapper individuals with all possible genotypes.

Region 1	Region 2	Standard fitness	min Swapper fitness	max Swapper fitness
lactose	lactose	800	400	800
lactose	glucose	2000	400	800
glucose	lactose	2000	1600	3200
glucose	glucose	3200	1600	3200
lactose	xxxxxxxx	401	400	800
xxxxxxxx	glucose	1601	2	2
xxxxxxxx	xxxxxxxx	2	2	2

Table 3: Fitness for Standard individuals. Here xxxxxxxx is any sequence not coding for a lactose or glucose metabolite.

2.1.5 Results of Experiment 1.

To test Hypothesis 1, populations of 5000 individuals were constructed with equal numbers of four genotypes: 00, 01, 10, and 11; and equal numbers of Standard and Swapper individuals. Trials were run until either 1000 generations had elapsed or only one type of individual remained. Various values for the production rate of Swappers were tested; the ratio between the production rates of Standards and Swappers ranged between 1 and 2 (min and max in Table 3). Pilot experiments had shown that values near 2 caused rapid loss of Standard individuals, so in the interest of efficiency the value for the ratio was varied in a manner similar to binary search. The values used were: {2, 1.5, 1.250, 1.125, 1.062, 1.031, 1.015, 1.007, 1.003, 1.002, 1.001}. Each value was tested 10 times. The results are summarized in Table 4 and graphed in Figure 2.

Ratio	Last Gen	Who Prevailed
2.000	69	Swappers
1.500	91	Swappers
1.250	108	Swappers
1.125	107	Swappers
1.062	109	Swappers
1.031	118	Swappers
1.015	121	Swappers
1.007	120	Swappers
1.003	139	Swappers
1.002	165	Swappers
1.001	74	Standard

Table 4: Results of Experiment 1 with various ratios of Swapper to Standard production efficiency. Trials were repeated 10 times. Each continued until either all the Swapper or all the Standard individuals were eliminated. Last Gen is the last generation in which any of the eliminated type survived.

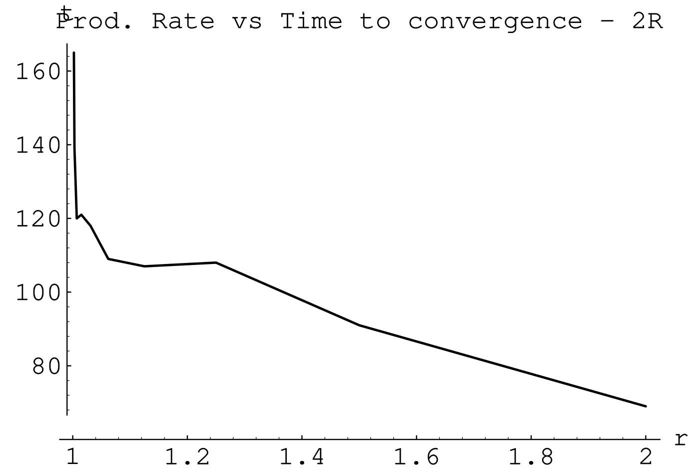

Figure 2: A graphical representation of Table 4; time to convergence to all Swappers.

Experiment 1 showed unequivocally that swapping confers adaptive advantage if the production efficiency of the single active region was even 0.2% more that that of one region of a Standard individual. In other words, if the total metabolite production rate of Swappers was at least 0.501 that of Standard individuals, Swappers drove Standard individuals to extinction in every trial. Thus, although there are many unknowns with respect to limits on production, under many sets of assumptions this mechanism could enhance survival in dynamic environments. Hypothesis 1 was therefore confirmed.

2.2 Experiment 2: Competition between different schemata

Experiment 2 investigates a more subtle question, whether introns are capable of effectively exploring the space of possible configurations. Hypothesis 2 could be expected to be confirmed if one assumed that no selection pressure is applied to non-coding regions, and there is nothing constraining genetic drift there. In that case, these regions could serve as an area to explore parts of the space far from the current situation, to create novel solutions, or, more generally, to retain diversity.

Experiment 2 was designed to test whether Swappers were superior to Standard individuals in retaining diversity in at least some segments of their chromosomes. A third region, again consisting of 8 locations, was appended to both Swapper and Standard individuals. Unlike Regions 1 and 2, this region, Region 3, was initialized to random bits. The fitness of Standard individuals was amended to be the sum of all three regions. When Swappers swapped, they did so between the active region and one of the two inactive regions with equal probability.

Populations of Swappers and Standards were compared for the length of time each retained copies of an arbitrarily chosen pattern, 01010101 -- called stripes, which modelled the coding of a metabolite for, say, fructose. Since this pattern does not increase or decrease fitness, it is as good as any other to check for the retention of diversity. Why is the retention of wide diversity desirable? Consider an *e. coli* population that having evolved in an environment which is sometimes rich in glucose and sometimes rich in lactose, has become adapted to metabolizing those two sugars. Occasionally it might find itself in an environment rich in some other nutrient, and poor in both glucose and lactose. If every coding region in every individual has converged to values that produce either the glucose, or the lactose metabolite, it will do much worse than another population which has retained enough diversity among its members that it happens to have a fructose metabolite in at least one of them. Thus the adaptability of our population in the face of unexpected changes (and the attendant requirement to innovate), can be gauged, at least in a crude manner, by simply counting the number of stripes remaining after some fixed time; or perhaps better, by recording the last generation in which any copies of that pattern existed in a fixed length run. Two different populations can be compared for flexibility by comparing values for each measure.

It is perhaps worth pointing out that the time of the last appearance of a stripe in the population, is different from the time that the stripe count is first 0. If we were counting stripes in a region that was doing a purely random walk under the effects of mutation we would expect to see variations about some mean. But that is not the only force at work here. Selection and replacement also play a role. Assume no mutation and an initial population of Standard individuals with an equal number of exactly two genotypes, 00 and 11. Since the population size in these experiments is 5000, the expected number of stripes initially is 5000/256 (which is about 19.5). Assume a population includes exactly 20 stripe regions and that 10 are associated with 00s and the other 10 with 11s. If, through the action of the GA, all the 00s are replaced by 11s then there will still be ~20 stripes. But, if initially all the stripes were on strings with 00s, then there would be none left after the population converges to 11s. During experiments the number of stripes observed ranged between 0 and 70. If the value of the stripe counter reaches 0 it does not necessarily mean that there is selection pressure against stripes; on the other hand if once the last stripe disappears no other is ever seen, then there is reason to suspect that the forces of convergence are at work.

2.2.1 Results of Experiment 2

As predicted, Swapper populations consistently outperformed Standard populations in retention of stripes, thus demonstrating another advantage of non-coding regions. Standard populations lost all stripes by generation 288 on average in 10 runs, whereas Swapper populations averaged 1322 generations before stripes disappeared for good. Thus, Hypothesis 2 was confirmed. Unfortunately, no Swapper population ever retained stripes for even 2000 generations, so Hypothesis 3 was not confirmed and in terms of long-term adaptability this mechanism was a failure. On inspection all regions in the Swapper population eventually converged to either 0's

or 1's, albeit more slowly than in the Standard population.

Because of the action of swapping, the initially random Region 3 values did not stay in Region 3. Once they were swapped to the active region (Region 1) they were as likely next to be swapped to Region 2, as back to Region 3; henceforth the 0 or 1 in Region 3 competed with and eliminated the random, unrewarded strings there. Experiment 3 was designed to remedy this problem.

2. 3 Experiment 3: Swap in but not out

A fourth region was added to Swappers, so that each individual had an active region and three inactive regions. Region 2 and Region 3 were initialized to 00, 01, 10, or 11 in equal numbers (like Region 1 and Region 2 in Experiment 2) and Region 4 was initialized at random (like Region 3 in Experiment 2). When swapping occurred, one of Region 2, Region 3, and Region 4 was copied with equal likelihood into the active region, but Region 1 was never copied back out. The runs from Experiment 2 were repeated, with the prediction that because the active region would never be copied to Region 4, it would remain random and so Swapper populations would retain stripes indefinitely, unlike Standard populations which inevitably converged and so eliminated them. This would confirm Hypothesis 3.

2.3.1 Results of Experiment 3

The prediction was not born out. Stripes never persisted past generation 5000 and the average time to the loss of the last stripe was 3360 generations. Again the problem was the growth of frequency of 0 and/or 1 (the metabolites for glucose and lactose) in Region 4. The mechanism that produced the selection pressure is apparent on a moment's reflection.

Region 4 was initialized at random, so as well as expecting about 20 individuals to have stripes in that region initially we would also expect there to be about 20 of all 255 other bit patterns there, including 0 and 1. When a Swapper fails to get any nourishment and so swaps it may swap in from Region 4. Sometimes the bit pattern in Region 4 codes for the metabolite that will grant the individual high fitness. This individual will then be more likely to reproduce and less likely to be eliminated than one with stripes in Region 4.

By way of example, consider two individuals, 0011 and 001S (where S stands for stripe, the code for the metabolite for fructose), in an environment rich in lactose (0). Both have 0 in the active region and both do equally well. Imagine the environment changes so there is no more lactose (0) but glucose (1) instead. Assume both individuals swap in from Region 4, so they now are: 1011 and S01S. The former has fitness 2000 and the latter fitness 2. Thus, in the next generation the latter will swap again; but in the current generation selection pressure favors the former. Thus the selection pressure and the reason 0s and 1s eventually predominated in Region 4.

2.4 Experiment 4: Differential probabilities of swapping.

The probability of swapping from Region 4, denoted P{R4s}, was varied from the value of 1/3=0.333 that was used in Experiment 3. The probability of swapping from Region 2 or 3 was equal: P{R2s}=P{R3s} and P{R2s}+P{R3s}+P{R4s}=1. Runs were repeated 10 times to 10,000 generations with various values for P{R4s}, namely {0.333, 0.3, 0.266, 0.233, 0.2, 0.166, 0.133, 0.1, 0.064, 0.032, 0.016, 0.008}.

2.4.1 Experiment 4: Results

As may be seen from the Table 5, shown graphically in Figures 3 and 4, this experiment confirms Hypothesis 3 when the probability of swapping in information from Region 4 was reduced.

P{R4s}	Number of trials in which stripes exist at g=10000	Time last stripe seen at (averaged)
0.333	0	3360
0.300	0	3410
0.266	0	4375
0.233	0	5789
0.200	0	7369
0.166	2	7637
0.133	7	6876
0.100	6	9665
0.064	8	9700
0.032	9	9885
0.016	8	9965
0.008	8	9910

Table 5: The number of trials out of 10 which had at least one stripe at generation 10,000, and the average generation at which the last stripe existed (in those runs which had no stripes at generation 10,000), for various probabilities of swapping from Region 4 (P{R4s}).

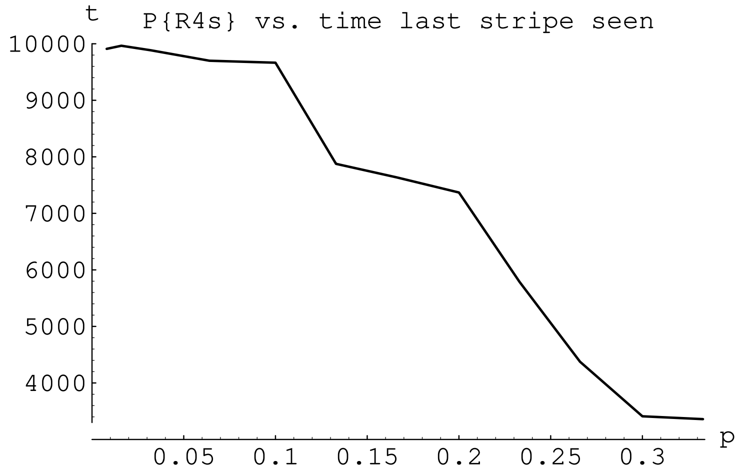

Figure 3: A graphical representation of the middle column of Table 4. With P{R4s} < 0.1 diversity was always retained.

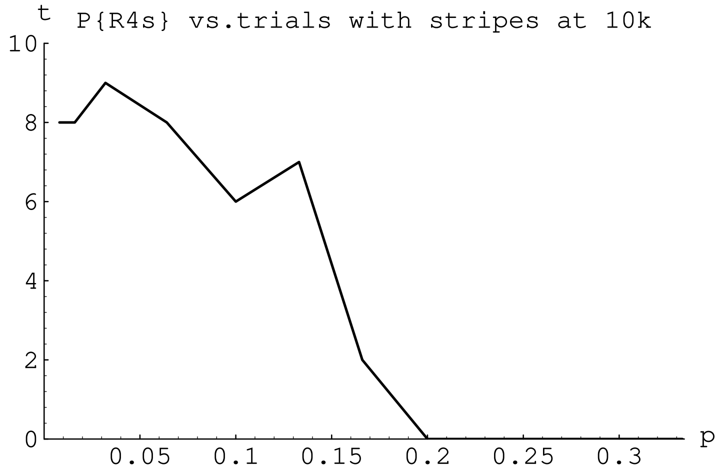

Figure 4:A graphical representation of the right column of Table 4. With P{R4s} > 0.2 no trial ever retained stripes for 10,000 generations.

2.5 Summary of Results

These experiments were designed and carried out to test three hypotheses :

Hypothesis 1: In a dynamic environment, swapping will confer adaptive advantage.

Hypothesis 2: Swapping will help retain diversity.

Hypothesis 3: Swapping can retain diversity in non-coding regions indefinitely.

Experiment 1 demonstrated the selective advantage of swapping introns into a single active

region over the standard of having all regions active. So long as that single region was able to produce at least 0.2% more metabolite than a single standard region (yielding a total production rate of 0.501 that of a standard individual) the Swappers always eliminated Standard individuals within 200 generations in repeated experiments. This confirmed Hypothesis 1.

Experiment 2 showed that individuals implementing swapping retained genetic diversity longer than standard individuals. This confirmed Hypothesis 2.

Experiment 3 tested the GA's ability to retain diversity indefinitely by copying introns into the active region, but never back out. This experiment failed to confirm Hypothesis 3 because unexpected selection pressure was generated by the chance occurrence of strings coding for one of the two prevalent nutrients in the region which was intended to remain free of selective pressure (Region 4).

Experiment 4 tested various values for P{R4s} (the probability of swapping in from Region 4). It demonstrated that lower probabilities (< 0.1) prevented convergence in that region and so allowed it to drift freely. This confirmed Hypothesis 3. Presumably the small remaining selection pressure for 0 and 1 in Region 4 was less than the loss rate due to mutation and replacement.

3.0 Future work

This paper reported simple experiments designed to test three elementary hypotheses relating to the benefits of swapping as a means of utilizing introns, or non-coding regions in evolutionary algorithms. There are numerous directions in which this research could be extended. These include:

1. An exploration of different dynamics. The experiments reported here always alternated between providing a large amount of only simulated glucose or lactose in a fixed pattern. Varying the relative concentrations of the two nutrients, the number of nutrients, and the length of time they are available would provide a large number of additional scenarios in which to test intron mechanisms.
2. Variable sized populations and/or endogenous fitness. The populations in these experiments were all of static size 5000. As numerous GA researchers have discovered, the population dynamics of fixed-sized populations are only one small, idiosyncratic corner of the larger space of variable-sized populations. With fixed sized populations a variant with higher fitness inevitably eliminates another with even slightly lower fitness. For instance, consider the inevitable, sudden extinctions of Standard individuals in Experiment 1. If the total amount of nutrient in the chemostat were divided between the bacteria which can metabolize it (thus making the fitness function endogenous), and the possibility of population explosions and/or crashes were thus introduced, a completely different set of challenges and insights would arise. This would allow the study of the proportion of individuals with various strategies for manipulating their genetic material which survived under different scenarios.
3. Mathematical analysis. It would be informative to perform an analysis to determine if the values derived empirically here conform to what one would expect. This would also allow

prediction about the performance of other potentially useful mechanisms.

4. Hunger as a non-binary signal. The Swappers in these experiments swapped the first generation in which they were getting less than some minimal amount of nutrient. Instead, each could gradually become hungry when nutrient uptake were less than energy expended and gradually become satisfied when it exceeded that value. Such hunger provides a primitive history mechanism (since it integrates a function of energy uptake over some period of the past). Gradually increasing hunger could gradually increase the likelihood of swapping and potentially provide more sophisticated performance.
5. Variable P{RXs} as a function of hunger. In cases of extreme hunger, individuals might increase mutation rates and/or the probability of swapping in low probability introns (where, presumably the most experimentation would have occurred).
6. Adding and deleting intron regions. It would be fascinating to study mechanisms that could implement the genesis of intron libraries. Two mechanisms spring to mind: 1) Add a copy of the current exon to the library at a very low frequency. Then, if an active region does not get replaced for a long time, it would be likely to be copied for later use. And, 2) Delete regions from the library at some low rate; to prevent it from increasing indefinitely. Together with this, one would want to introduce some cost function for increasing the size of the chromosome, and some mechanism for breeding individuals with different length chromosomes.

4.0 Conclusion

The experiments reported herein have shown unequivocally two advantages of non-coding regions. First, the capacity to convert introns to exons by copying them to a coding region provided a distinct selective advantage in a dynamic environment. Second, the addition of another intron region which was only rarely copied into the coding region, provided an effective means to retain diversity while maintaining performance. Together, these mechanisms offer an avenue of attack on the longstanding challenges of stability versus plasticity and exploitation versus exploration that Holland pointed out 25 years ago [Holland 1975]. Unexpressed genetic material (introns), given the possibility of being expressed at need, can solve the first problem; introns which are protected from selective pressure can search the space and stand ready for expression when some novel or rare situation arises, and so ameliorate the second problem. The mechanisms that implement these features in the model bacteria require no calculation, memory or intelligence, only the capacity to copy introns into the coding region when the bacteria is nutrient deprived. *E. coli* have long been known to follow nutrient gradients, so this decision process is clearly within their capacity, and is certainly within the capacity of GAs.

These simple experiments have demonstrated that introns can confer adaptive advantage in dynamic environments, but they pose more questions than they answer. Will this technique work with longer strings? How sensitive is it to the pattern of dynamics in the environment? How will it work in the context of dynamic population sizes? What is the cost of maintaining and copying introns? How should their number be determined? How amenable are various evolutionary algorithms to this technique? The way lies before us -- we have only to choose which paths to explore.

BIBLIOGRAPHY

Adams, J. (1995). Chemostat experiments. ECHO workshop presentation, Feb. 1995: Sevilleta, N.M.

Davidor, Y. (1991). A Naturally Occurring Niche & Species Phenomenon: The Model and First Results. *Proceedings of the Fourth ICGA,* 257-263. Kaufmann, San Mateo, CA.

French, R. M. (1992). Semi-distributed Representations and Catastrophic Forgetting in Connectionist Networks, *Connection Science*, 4, 365-377.

Gould, S. H. (1989). Through a Lens, Darkly: Do species change by random molecular shifts or natural selection? *Natural History*.

Holland, J.H. (1975). *Adaptation in natural and artificial systems.* Ann Arbor: University of Michigan Press.

Holland, J.H. (1992). *Adaptation in natural and artificial systems.* Cambridge. MA: MIT Press. Second edition.

Jones, T. (1995). Crossover, Macromutation, and Population-based Search. *Proceedings of the Sixth ICGA,* 73-80. Kaufmann, San Mateo, CA.

Levenick, J. R. (1991) Inserting Introns Improves Genetic Algorithm Success Rate: Taking a Cue from Biology. *Proceedings of the Fourth ICGA,* 123-127. Kaufmann, San Mateo, CA.

Raich, A. M., & J. Ghaboussi (1997). Autogenesis and Redundancy in GA Representation. Presented at a workshop at the Seventh ICGA.

Robins, A. (1995) Catastrophic forgetting, rehearsal and pseudorehearsal. *Connection Science*, 7, 123-146.

Rosenzweig, R. F., R. R. Sharp, D. S. Treves, J. Adams (1994). Microbial Evolution in a simple Unstructured Environment: Genetic Differentiation in *Escherichia coli*. Genetics, 137, 903-917.

Satamoto, Y. & D. E. Goldberg (1997). Takeover Time in a Noisy Environment. *Proceedings of the Seventh ICGA,* 160-165. Kaufmann, San Mateo, CA.

Schaffer, J. D., Eshelman, L. J. & Offutt, D. (1991). Spurious Correlations and Premature Convergence in Genetic Algorithms. In G. J. E. Rawlins (ed), *Foundations of GAs and Classifier Systems*, Kaufmann, San Mateo, CA.

Swappers: Introns promote flexibility, diversity and invention.

Abstract: This paper investigates non-coding and inactive regions in genetic algorithms. The ability to swap information into an active region confers increased flexibility and thus adaptive advantage in dynamic environments.

Key words: Genetic algorithms, non-coding regions, introns, convergence vs. diversity.

Jim Levenick
Willamette University, Computer Science
900 State St
Salem, OR. 97301
(503)-370-6486
levenick@willamette.edu