

An Analysis of Diploidy and Dominance in Genetic Algorithms

Dan Simon
Cleveland State University
Department of Electrical and Computer Engineering
Cleveland, Ohio
d.j.simon@csuohio.edu

Abstract—The use of diploidy and dominance in genetic algorithms (GAs) has long been used to improve performance in time-varying optimization problems. Diploidy increases diversity in GAs by allowing recessive genes to survive in a population and become active at some later time when changes in the environment make them more desirable. This paper suggests an intuitive way to implement diploidy and presents some mathematical analyses of fitness proportional selection to justify its use in time-varying problems. An extension of the classical schema theorem for diploid GAs is presented. The mathematical analyses are geared towards the One Max problem, and assume a GA with selection and mutation only (no crossover). The analyses confirm that diploidy increases diversity, and provide some quantitative results for diversity increase as a function of the GA population characteristics.

Key Words – Genetic algorithms, optimization, diploidy, dominance, diversity

I. INTRODUCTION

The universe in which we live is nonstationary, so optimization problems are often time varying. This makes optimization more difficult than it would be otherwise. We can find a solution to an optimization problem under the assumption of stationarity, but the problem may change in the next instant and our solution will no longer be valid.

However, a solution that works well in one environment often works well in another environment too. In addition, if the environment of our optimization problem changes from A to B , then it very well may change back to A at some point in the future. Therefore, our optimal solution in environment A has some worth in the future, even if the environment is changing and the solution is invalid at the present time.

Most computer-based GAs are haploid. Haploid GAs involve the use of a single stranded chromosome to represent the solution to a problem. Haploid GAs have been shown to be useful for a variety of difficult optimization problems. However, most complex biological organisms have two strands of chromosomes. This fact makes it natural to investigate the use of diploid GAs (double stranded chromosomes) for optimization problems. Diploid GAs have two alleles (bits) at each locus (position) of the chromosome. Therefore something must determine which allele appears in the phenotype (expression) of the chromosome. The value of the expressed allele is determined through dominance. One allele value is dominant and the other is recessive. The diploid GA needs to carry twice as much information as the haploid GA. This allows a larger number of recessive

(low fitness) alleles to be maintained in the population. If the problem environment changes and the currently optimal solution becomes poor, the diploid GA can adapt and find the new optimum more quickly than a haploid GA.

A variety of methods for implementing diploid GAs have been suggested in the past [5], [8], [11]. In this paper we use a natural and straightforward method. At each locus we have two alleles, and each allele can be either a 0 or a 1. We therefore have four possible combinations of the two alleles. The *expressed* allele at the locus is a function of the two *given* alleles at that locus. The expressed allele will be equal to the dominant value if either of the two given alleles is equal to the dominant value. The expressed allele will be equal to the recessive value if both of the two given alleles are equal to the recessive value. For example, if 1 is dominant and 0 is recessive then the expressed allele is given as

$$A_e = A_1 \text{ OR } A_2 \quad (1)$$

where A_1 and A_2 are the two given alleles.

A diploid GA needs some way for dominance to evolve. Equation (1) assumes that 1 is the dominant allele value. However, we do not know if 1 should be dominant or not. The GA needs some way to change the dominant values between 0 and 1 at each position as the GA proceeds. A global approach to dominance can be used by evaluating the average fitness of each allele value at each locus [3]. A dominance map can also be implemented as part of the chromosome itself [2], or it can be implemented as a separate allele at each locus [10] (essentially a dominance chromosome). This latter approach can be simplified by using a ternary alphabet rather than a binary alphabet. In this representation an allele can take on the value 0, 1, or 1_0 . The value 0 is expressed as a 0, and the values 1 and 1_0 are both expressed as a 1. The allele 1 dominates 0, and 0 dominates 1_0 . A shift in dominance is therefore modeled as replacing a 1 with a 1_0 , or vice versa. Table I shows the expressed allele value as a function of the two given alleles.

This paper does not address the evolution of dominance. Although this is not a trivial problem for diploid GAs, we assume in this paper that a correct method for determining dominance is implemented. We will assume that the number of 0s in our population is given by m_0 , the number of 1s in our population is given by m_1 , a 1 dominates a 0 at each

TABLE I
EXPRESSED ALLELE VALUES AS A FUNCTION OF GIVEN ALLELES

		Allele #1		
		1	0	1 ₀
Allele #2	1	1	1	1
	0	1	0	0
	1 ₀	1	0	1

locus, each expressed 0 bit contributes f_0 to the fitness of its chromosome, each expressed 1 bit contributes f_1 to the fitness of its chromosome, and $f_1 > f_0$.

Section II derives a discrete-time dynamic equation for the proportion of low-fit alleles in a diploid GA, and uses it to show that diploid GAs have more diversity (i.e., more low-fit alleles) than haploid GAs in a scaled One Max problem. The analysis shows that maintenance of a proportion P of low-fit alleles requires a mutation rate of $p_m = P$ for haploid GAs, but it requires a mutation rate of only $p_m = P^2$ for diploid GAs. Section III derives a diploid version of the classical schema theorem and uses it to show that the increase in diversity provided by diploidy is more pronounced as P decreases, and as the fitness ratio of high-fit to low-fit alleles increases. Section IV provides some concluding remarks.

II. DIPLOIDY INCREASES DIVERSITY IN GAS

In this section we derive a discrete-time dynamic equation for the proportion of low-fit alleles in a diploid GA. We then use this equation to show that diploid GAs have more diversity (i.e., more low-fit alleles) than haploid GAs in the scaled One Max problem. The analysis shows that maintenance of a proportion P of low-fit alleles requires a mutation rate of $p_m = P$ for haploid GAs, but it requires a mutation rate of only $p_m = P^2$ for diploid GAs.

The classical schema theorem [7] says that for standard fitness proportionate selection in a binary haploid GA the following approximation holds.

$$m(H, t+1) \geq m(H, t) \frac{f(H)}{\bar{f}(t)} \left[1 - p_c \frac{\delta}{l-1} - p_m o(H) \right] \quad (2)$$

where $m(H, t)$ is the number of instances of schema H at generation t , $f(H)$ is the average fitness of schema H at generation t , $\bar{f}(t)$ is the average fitness of the entire population at generation t , p_c is the crossover probability, δ is the defining length of H , l is the number of bits in each chromosome, p_m is the probability of mutation per bit, and $o(H)$ is the order of H .

Note that we are using the original schema theorem as derived by John Holland [9]. Although this theorem has fallen into disfavor among many GA researchers, it nevertheless remains true. It has fallen into disfavor not because it is false, but because it has too often been used to incorrectly extrapolate results beyond a single generation in finite populations [13], [14]. If the population size is infinite, then the schema theorem implies that fit, short, low-order schemata receive exponentially increasing numbers of instances. More exact versions of the schema theorem can

be derived for finite populations [12], [13], [15], but the extension of this paper to these other versions of the schema theorem remains for future work.

A. Proportion of low-fit schema

The use of diploidy and dominance increases diversity in GAs. That is, more low-fit schemata will be retained in diploid GAs than in haploid GAs. This increased diversity balances the GA more towards exploration (rather than exploitation) of the fitness landscape, and helps GAs adapt to changing environments more quickly and effectively.

To analyze this in detail, consider a simplified haploid GA where the number of 0s and 1s are given by m_0 and m_1 respectively, and their fitnesses are f_0 and f_1 respectively. This is the scaled One Max problem [1]. Further suppose that the GA operates by selection and mutation only (no crossover). Although crossover is often considered as an essential component of GAs, there is evidence that its role is often overemphasized and that selection and mutation are the primary mechanisms of successful evolution [1, p. 96].

Given the assumptions above, the number of 0s and 1s during the next generation is given by

$$\begin{aligned} m_0(t+1) &= \frac{f_0 m_0(t)}{\bar{f}(t)} (1 - p_m) + \frac{f_1 m_1(t)}{\bar{f}(t)} p_m \\ m_1(t+1) &= \frac{f_1 m_1(t)}{\bar{f}(t)} (1 - p_m) + \frac{f_0 m_0(t)}{\bar{f}(t)} p_m \end{aligned} \quad (3)$$

We see that the number of bits is proportional to its fitness times the probability that a mutation does *not* take plus, plus the fitness of the competing bit times the probability that a mutation *does* take place. From these equations we can derive the proportion of 0s in the population as a function of time.

$$\begin{aligned} P(t+1) &= \frac{m_0(t+1)}{m_0(t+1) + m_1(t+1)} \\ &= \frac{P(t)(1 - p_m) + r(1 - P(t))p_m}{P(t) + r(1 - P(t))} \end{aligned} \quad (4)$$

where the ratio of the higher fitness to the lower fitness is $r = f_1/f_0 \geq 1$. Now consider a diploid GA population with m_0 0s and m_1 1s for a total of M bits. The use of diploidy means that this population will combine to produce a total of $M/2$ expressed values. A combination of two 0s will produce an expressed value of 0, and a combination of a 1 with any other bit will produce an expressed value of 1. The expected number of times that two given 0s will combine is given as

$$\begin{aligned} E(\text{two 0s}) &= (\# \text{ bit pairs}) \times P(\text{two zeros}) \\ &= \frac{M}{2} \frac{m_0(m_0 - 1)}{M(M - 1)} \\ &\approx \frac{m_0^2}{2M} \end{aligned} \quad (5)$$

Similarly, the expected number of times that two given 1s will combine, and the expected number of times that one 0

and one 1 will combine, is approximately equal to

$$E(\text{two 1s}) \approx \frac{m_1^2}{2M} \quad (6)$$

$$E(\text{one 0 and one 1}) \approx \frac{M}{2} - \frac{m_0^2}{2M} - \frac{m_1^2}{2M}$$

With a population of M alleles, a chromosome has a chance of f/\bar{f} of being selected for the next generation, where f is the expressed value of the chromosome. However, since each chromosome must be combined with another chromosome in order to determine its expressed value, each selected chromosome must be selected twice in order to maintain the population at a size of M alleles. If two 0s are at a locus, then the expressed value of the locus is f_0 . If two 1s are at a locus, then the expressed value of the locus is f_1 . If one 0 and one 1 are at a locus, then the expressed value of the locus is f . Given these data, we can write the number of 0s and 1s during the next generation as

$$m_0(t+1) = 2 \frac{f_0}{f(t)} \frac{m_0^2(t)}{2M} (1-p_m) + 2 \frac{f_1}{f(t)} \frac{m_1^2(t)}{2M} p_m + \frac{f_1}{f(t)} \frac{M^2 - m_0^2(t) - m_1^2(t)}{2M} \quad (7)$$

$$m_1(t+1) = 2 \frac{f_0}{f(t)} \frac{m_0^2(t)}{2M} p_m + 2 \frac{f_1}{f(t)} \frac{m_1^2(t)}{2M} (1-p_m) + \frac{f_1}{f} \frac{M^2 - m_0^2(t) - m_1^2(t)}{2M}$$

From these equations we can derive the proportion of 0s in the population as a function of time.

$$P(t+1) = \frac{2P^2(t)(1-p_m) + 2rQ^2(t)p_m + rv(t)}{2P^2(t) + 2rQ^2(t) + 2rv(t)} \quad (8)$$

$$v(t) \equiv 1 - P^2(t) - Q^2(t)$$

where $Q = 1 - P$ is the proportion of 1s in the population. Figures 1 and 2 show the evolution of $P(t)$ as a function of generation number t for different fitness ratios r and mutation rates p_m . The figures show $P(t)$ for the haploid GA (4) and diploid GA (8), and each GA begins with the initial condition $P(0) = 1/2$. These figures show that the diploid GA always has a larger proportion of less-fit schema than the haploid GA. In other words, the diploid GA increases diversity, allowing the GA to be more adaptable to changing environments.

B. Steady-state analysis

The steady state solution of $P(t)$ for the the haploid GA (4) is given by

$$P_h = \frac{q - \sqrt{q^2 - 4(r-1)rp_m}}{2(r-1)} \quad (9)$$

$$q \equiv r + rp_m + p_m - 1$$

where the subscript h refers to the haploid GA. The steady state solution of $P(t)$ for the the diploid GA (8) is given by the solution of the equation

$$(1-r)P_d^3 + (r-1)(1-p_m)P_d^2 + 2rp_mP_d - rp_m = 0 \quad (10)$$

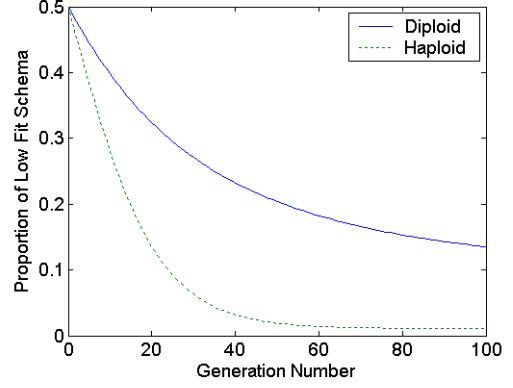


Fig. 1. Proportion of low-fit schema when fitness ratio $r = 1.1$ and mutation rate $p_m = 0.001$

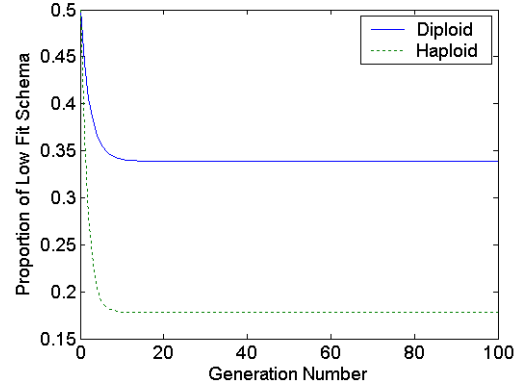


Fig. 2. Proportion of low-fit schema when fitness ratio $r = 2.0$ and mutation rate $p_m = 0.1$

where the subscript d refers to the diploid GA. Figures 3 and 4 show the steady state value of $P(t)$ for the haploid GA and the diploid GA as functions of r and p_m . Figure 5 shows the difference of the two plots (i.e., Figure 4 minus Figure 3). Again it is seen that the diploid GA results in a higher proportion of low-fit schema, which allows the GA to be more adaptable to changing environments.

Consider the limiting case as $r \rightarrow \infty$. Note that this implies $f_0 \rightarrow 0$ and coincides with the One Max problem [1]. In this case (9) shows that the steady-state value of P_h becomes

$$\lim_{r \rightarrow \infty} P_h = p_m \quad (11)$$

The limiting value of (10) can be approximated (for small P_d) as

$$\lim_{r \rightarrow \infty} P_d = \frac{\sqrt{p_m} - p_m}{1 - p_m} \quad (12)$$

This can be solved for p_m as

$$p_m = \frac{P_d^2}{(P_d - 1)^2} \quad (13)$$

$$\approx P_d^2 \quad \text{for small } P_d$$

Equation (11) shows that in the limiting case for a haploid GA, a mutation rate of p_m gives a steady-state proportion of

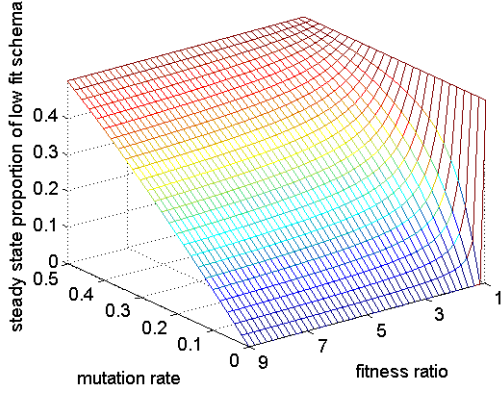


Fig. 3. Steady-state proportion of low-fit schema for the haploid GA

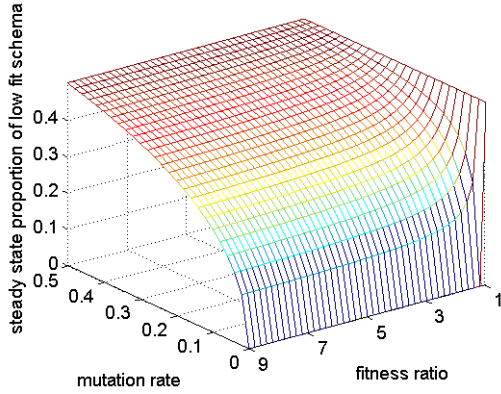


Fig. 4. Steady-state proportion of low-fit schema for the diploid GA

low-fit schema that is equal to p_m . On the other hand, (13) shows that for a diploid GA the same mutation rate gives a steady-state proportion that is equal to $\sqrt{p_m}$. In other words, the same mutation rate gives a higher proportion of low-fit schema for a diploid GA than for a haploid GA. This is in agreement with Holland's results, which were obtained by a much different type of analysis [9, p. 115].

III. SCHEMA ANALYSIS FOR DIPLOID GAS

In this section we present an analysis of the schema theorem that is modified for diploid GAs. This is based on [6] but this section contains less approximation, more detail, and some specific results. We show that the use of diploid GAs results in a higher probability of selection for low-fit schemata, and a lower probability of selection for high-fit schemata. Moreover, this difference between diploid and haploid GAs increases as P decreases and as r increases. As with the previous section, this analysis is geared towards the scaled One Max problem.

For haploid GAs, the schema theorem can be written as (2), which is repeated here.

$$m_h(H, t+1) \geq m_h(H, t) \frac{f_h(H)}{f_h(t)} \left[1 - p_c \frac{\delta}{l-1} - p_m o(H) \right] \quad (14)$$

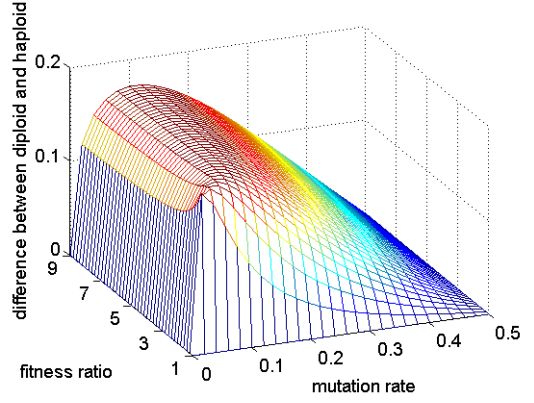


Fig. 5. Difference between the diploid and haploid GAs of the steady-state proportion of the less fit schema (Figure 4 minus Figure 3)

where the subscript h is used to refer to a haploid GA. As in the previous section, suppose that each 0 in a chromosome contributes an incremental fitness of f_0 to the chromosome, and each 1 in a chromosome contributes an incremental fitness of f_1 . Then the average fitness of the haploid population is given by

$$\bar{f}_h = \left(\frac{m_0 f_0 + m_1 f_1}{m_0 + m_1} \right) l \quad (15)$$

where m_0 and m_1 are the respective number of 0s and 1s in the population, and l is the number of bits in each chromosome. Now consider $f_h(H)$, the fitness of schema H in the haploid population. Each 0 in H contributes f_0 to the fitness, and each 1 in H contributes f_1 to the fitness. Each * (undefined bit) in H could be realized as either a 0 or a 1, and could therefore contribute either f_0 or f_1 to the fitness. The proportion of times that a * will be realized as a 0 is equal to P , the proportion of 0s in the population. The proportion of times that a * will be realized as a 1 is equal to $Q = 1 - P$. The average fitness of H is therefore given by

$$f_h(H) = o_0 f_0 + o_1 f_1 + o_* (f_0 P + f_1 Q) \quad (16)$$

where o_0 , o_1 , and o_* are the number of 0s, 1s, and *s in H . Equation (14) can therefore be written as

$$m_h(H, t+1) \geq m_h(H, t) \frac{o_0 f_0 + o_1 f_1 + o_* (f_0 P + f_1 Q)}{l(m_0 f_0 + m_1 f_1) / (m_0 + m_1)} \times \left[1 - p_c \frac{\delta}{l-1} - p_m o(H) \right] \quad (17)$$

Now consider a diploid GA. In this case the schema theorem (2) is written as

$$m_d(H, t+1) \geq m_d(H, t) \frac{f_d(H)}{f_d(t)} \left[1 - p_c \frac{\delta}{l-1} - p_m o(H) \right] \quad (18)$$

where the subscript d is used to refer to a diploid GA. From (5) we recall that the expected number of expressed 0s is equal to $m_0^2 / 2M$, where $M = m_0 + m_1$ is the total number of bits in the population. Therefore the expected number of expressed 1s is equal to $M/2 - m_0^2 / 2M$ (recall that there are

$M/2$ expressed bits in a population of M bits). The average fitness of the diploid GA population is therefore given by

$$\begin{aligned}\bar{f}_d &= \left[\frac{m_0^2}{2M} f_0 + \left(\frac{M}{2} - \frac{m_0^2}{2M} \right) f_1 \right] l / (M/2) \quad (19) \\ &= \left(\frac{m_0^2 f_0 + M^2 f_1 - m_0^2 f_1}{M^2} \right) l\end{aligned}$$

Now consider $f_d(H)$, the fitness of schema H . The proportion of times that a 0 in H contributes f_0 to the fitness is equal to P , and the proportion of times that a 0 in H contributes f_1 to the fitness is equal to Q . Since the number of 0s in H is equal to o_0 , the 0s will contribute a total of $(f_0 P + f_1 Q) o_0$ to the fitness of H . A 1 in H will always contribute f_1 to the fitness. Since the number of 1s in H is equal to o_1 , the 1s will contribute a total of $f_1 o_1$ to the fitness. The proportion of the time that a * in H will be realized as a 0 is equal to P , and the proportion of the time that a * in H will be realized as a 1 is equal to Q . The average fitness contributed by a * is therefore equal to $P(f_0 P + f_1 Q) + Q f_1 = f_0 P^2 + f_1(1 - P^2)$. Since the number of *s in H is equal to o_* , the *s will contribute a total fitness of $[f_0 P^2 + f_1(1 - P^2)] o_*$ to H . The average fitness of H is therefore given by

$$f_d(H) = [f_0 P^2 + f_1(1 - P^2)] o_* + (f_0 P + f_1 Q) o_0 + f_1 o_1 \quad (20)$$

Equation (18) can therefore be rewritten by substituting (19) for $\bar{f}_d(t)$ and (20) for $f_d(H)$.

To compare the growth in the number of instances of schema H between the haploid and diploid GAs, suppose that $m_h(t) = m_d(t)$ (i.e., the number of instances of H in the haploid and diploid GAs are the same at generation t). We can divide (18) by (17) to obtain

$$\begin{aligned}\frac{m_d(H, t+1)}{m_h(H, t+1)} &= \frac{(q_3 o_* + (P + rQ) o_0 + r o_1) (P + rQ)}{q_3 [o_0 + r o_1 + o_* (P + rQ)]} \\ q_3 &\equiv P^2 + r(1 - P^2) \quad (21)\end{aligned}$$

where we have used the fact that $r = f_1/f_0$. Note that the mutation and crossover terms in (17) and (18) cancel each other out when we perform the division. If P is small then the above equation can be approximated as

$$\begin{aligned}\frac{m_d(H, t+1)}{m_h(H, t+1)} &\approx \frac{rQ^2 o_0 + rQ o_1 + rQ o_*}{o_0 + r o_1 + rQ o_*} \quad \text{if } P \ll 1 \\ \frac{m_d(H, t+1)}{m_h(H, t+1)} &> 1 \quad \text{if } \frac{o_1}{o_0} < \frac{(1-P)^2 r - 1}{Pr} \quad (22)\end{aligned}$$

In other words, in populations with small proportions of low-fit individuals (small P), the schema instances in a diploid GA will increase more than in a haploid GA if the proportion of high-fit to low-fit alleles (o_1/o_0) is less than a threshold value. This threshold is plotted in Figure 6. Values *below* the surface of the figure are values of o_1/o_0 that will result in greater schema increase in a diploid GA than in a haploid GA. Note from the figure that the o_1/o_0 threshold increases as P decreases. In other words, the increased schema growth of a diploid GA will be more pronounced when P is small. The figure also shows that the o_1/o_0 threshold increases as

the fitness ratio r increases. In other words, the increased schema growth of a diploid GA will be more pronounced when the fitness difference between low-fit and high-fit schemata is larger. Both of these results agree with intuition. A diploid GA increases diversity, and the diversity increase is more pronounced when the proportion of low-fit schema is small, and when the difference between high and low fitnesses is larger.

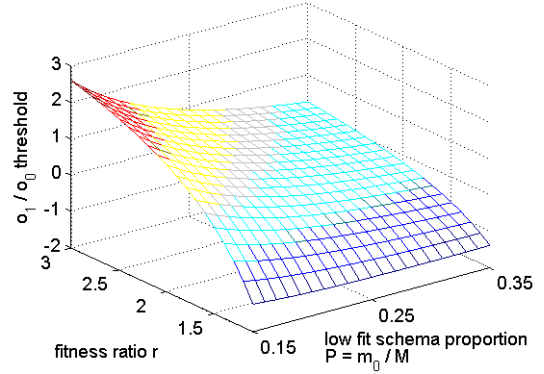


Fig. 6. o_1/o_0 threshold below which a diploid GA will grow a schema faster than a haploid GA

A. A limiting case of increased schema growth

Now consider the limit $r \rightarrow \infty$. In this case (21) can be approximated as

$$\frac{m_d(H, t+1)}{m_h(H, t+1)} = \frac{(1 - P^2)(o_1 + o_* Q) + Q^2 o_0 + o_1 P(P - 1)}{(1 - P^2)(o_1 + o_* Q)} \quad (23)$$

From this equation we can see that

$$\begin{aligned}\frac{m_d(H, t+1)}{m_h(H, t+1)} &> 1 \quad \text{if } Q^2 o_0 + o_1 P(P - 1) > 0 \\ \Rightarrow P &< \frac{o_0}{o_0 + o_1} \quad (24)\end{aligned}$$

In other words, for large r a diploid GA will result in greater schema growth if the schema has a greater proportion of 0s than the population. The converse is also true; a diploid GA will result in a *lower* schema growth if the schema has a *smaller* proportion of 0s than the population.

B. Example

A simple example can be used to illustrate the result at the end of the previous section. Suppose we have a four-member population, each with two bits. The population is distributed as follows.

- One instance of 00
- One instance of 01
- One instance of 10
- One instance of 11

Each 0 contributes a value of 0 to the fitness of an individual, and each 1 contributes a value of 1 to the fitness of an individual (i.e., $f_0 = 0$ and $f_1 = 1$). We see that there are an equal number of 0s and 1s in the population, so

$P = 1/2$. Table II shows (for both haploid and diploid GAs) the fitness of each individual, the average fitness of the population, and the expected number of selections of each individual for the next generation. Note that the expected number of selections of individual 00 is greater for the diploid GA. This is in accordance with the results of the previous subsection, because individual 00 has a greater proportion of 0s (100%) than the population has (50%). However, the expected number of selections of individual 11 is greater for the haploid GA. This is because individual 11 has a smaller proportion of 0s (0%) than the population has. However, the expected number of selections of individuals 01 and 10 are the same for both the haploid and diploid GAs. This is because these individuals have the same proportion of 0s (50%) as the population has.

TABLE II

ILLUSTRATIVE EXAMPLE OF HAPLOID AND DIPLOID SELECTION. $f(H)$ = INDIVIDUAL FITNESS, \bar{f} = AVERAGE POPULATION FITNESS, AND $E(H)$ = EXPECTED NUMBER OF SELECTIONS.

Chromosome	Haploid GA			Diploid GA		
	$f(H)$	\bar{f}	$E(H)$	$f(H)$	\bar{f}	$E(H)$
00	0	1	0	1	3/2	2/3
01	1	1	1	3/2	1	1
10	1	1	1	3/2	1	1
11	2	2	2	2	4/3	4/3

IV. CONCLUSION

Diploid GAs have long been used to improve performance relative to haploid GAs in time-varying optimization problems. Simulation results demonstrating this advantage can be found in a number of publications [4], [16], [17]. This paper uses the scaled One Max problem to provide additional theoretical basis for the superior time-varying performance of diploid GAs. Diploidy increases diversity in GAs by allowing recessive genes to survive in a population and then become active at some later time when changes in the environment make them more fit. This paper takes an intuitive implementation of diploidy and presents some mathematical analyses of fitness proportional selection to justify the use of diploidy in time-varying optimization problems. Maintenance of a proportion P of low-fit alleles requires a mutation rate of $p_m = P$ for haploid GAs, but it requires a mutation rate of only $p_m = P^2$ for diploid GAs. Our derivation of a diploid version of the classical schema theorem shows that the additional diversity provided by diploidy is more pronounced as P decreases, and as the fitness ratio of high-fit to low-fit alleles increases.

One drawback of diploidy is that the mechanics of a diploid GA requires twice as much computational effort as the mechanics of a haploid GA because we have twice as many alleles to deal with. However, for most real world GAs (complicated optimization problems) the bulk of the computational effort lies in fitness evaluation, and this effort is the same for diploid GAs as for haploid GAs.

This paper has analyzed GAs with selection and mutation only (no crossover), and has used John Holland's original schema theory. Future work could focus on extending the results of this paper to GAs with crossover, and to more exact versions of schema theory as given in [12], [13], [15].

REFERENCES

- [1] T. Back, *Evolutionary theorems in theory and practice*, Oxford University Press, 1996.
- [2] D. Bagley, *The behavior of adaptive systems which employ genetic and correlation algorithms*, doctoral dissertation, University of Michigan, Ann Arbor, MI, 1967.
- [3] A. Brindle, *Genetic algorithms for function optimization*, doctoral dissertation, University of Alberta, Edmonton, Canada, 1981.
- [4] R. Calabretta, S. Nolfi, D. Parisi, and R. Galbati, Diploid robots adapting to fast changing environments, *Proceedings of the 8th International Conference on Artificial Neural Networks*, Skovde, Sweden, 1998, 1145-1150.
- [5] C. Fonteix, F. Bicking, E. Perrin, and I. Marc, Haploid and diploid algorithms, a new approach for global optimization: compared performances, *International Journal of Systems Science*, 26(10), Oct. 1995, 1919-1933.
- [6] D. Goldberg and R. Smith, Nonstationary function optimization using genetic algorithms with dominance and diploidy, *International Conference on Genetic Algorithms*, Cambridge, MA, 1987, 59-68.
- [7] D. Goldberg, *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, 1989.
- [8] F. Greene, A new approach to diploid/dominance and its effects on stationary genetic search, *Proceedings of the Fifth Annual Conference on Evolutionary Programming*, Cambridge, MA, 1996, 171-176.
- [9] J. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press, 1975.
- [10] R. Hollstien, *Artificial genetic adaptation in computer control systems*, doctoral dissertation, University of Michigan, Ann Arbor, MI, 1971.
- [11] V. Klepikov, K. Mahotilo, and S. Sergeev, Modification of Holland's reproductive plan for diploid populations, *Proceedings of International Conference on Artificial Neural Networks and Genetic Algorithms*, Ales, France, 1995, 337-339.
- [12] R. Poli, Recursive conditional schema theorem, convergence and population sizing in genetic algorithms, *Foundations of genetic algorithms 6*, W. Martin and W. Spears (eds.), 2001, 143-163.
- [13] R. Poli, Exact schema theory for genetic programming and variable-length genetic algorithms with one-point crossover, *Genetic Programming and Evolvable Machines*, 2(2), June 2001, 123-163.
- [14] C. Reeves and J. Rowe, *Genetic algorithms: Principles and perspectives*, Kluwer Academic Publishers, 2003.
- [15] C. Stephens and H. Waelbroeck, Schemata evolution and building blocks, *Evolutionary Computation*, 7(2), Summer 1999, 109-124.
- [16] S. Uyar and A. Harmanci, Investigation of new operators for a diploid genetic algorithm, *Proceedings of the SPIE - The International Society for Optical Engineering*, 3812, Denver, CO, 1999, 32-43.
- [17] D. Weile and E. Michielssen, The control of adaptive antenna arrays with genetic algorithms using dominance and diploidy, *IEEE Transactions on Antennas and Propagation*, 49(10), Oct. 2001, 1424-1433.