

Evolutionary Algorithms with Extended Fitness

Dimitar Kazakov

Department of Computer Science, University of York,
Heslington, York YO10 5DD, UK,

kazakov@cs.york.ac.uk,

WWW home page: <http://www-users.cs.york.ac.uk/~kazakov/>

Abstract. The notion of fitness has been assigned various meanings, of which only the oldest, expressing individual reproductive success, has been explicitly used in Evolutionary Algorithms (EAs) so far. This paper suggests that the use of other well-known definitions borrowed from biology that are based on the success with which genes replicate and propagate themselves in the gene pool could be beneficial to EAs by replacing *ad-hoc* scaling techniques with means of maintaining genetic diversity that draw clear analogies with nature. The resulting improvement in performance is demonstrated on the problem of continuous function optimisation.

1 Introduction

Since Herbert Spencer introduced the term ‘survival of the fittest’ to describe the process of natural selection, the notion of fitness has been assigned various meanings, of which only the oldest, expressing individual reproductive success, has been explicitly used in Evolutionary Algorithms (EAs) so far. This paper suggests that the use of other well-known definitions borrowed from biology that are based on the success with which genes replicate and propagate themselves in the gene pool could be beneficial to EAs and demonstrates it on the problem of continuous function optimisation.

EAs, despite their large variety, are all based on the same basic principles. They maintain a population of individuals representing candidate solutions to an optimisation problem. For each generation, a *fitness* reflecting estimated or actual quality is assigned to each solution (individual). A next generation of individuals is obtained by sampling the current so that individuals with higher fitness are favoured. Finally, the new generation is subjected to genetic operations such as crossover and mutation that aim at introducing a variety of new individuals. Then the cycle is repeated until some termination condition is fulfilled [Gol89].

The fitness of an individual may measure the quality of the solution it proposes in absolute terms, say, as a scalar representing the value of a function that the EA is trying to maximise. In other cases, e.g., when it is normalised, fitness only represents the relative quality of an individual with respect to the rest.

Fitness could be an even more abstract concept only reflecting the rank of the individual in the population.

Whether the goal of the EA is to provide a single best solution or a number of these, its principle remains the same: to store copies of the best one(s) aside from the main population, and wait until better ones are produced. For that ever to happen, it is essential that the EA should be capable of producing individuals that have not been seen in the previous generations. While the genetic operators, such as crossover and mutation, are the EA components that introduce new individuals in the population, their success depends on preserving sufficient genetic variety in it. Evolution, whether in nature or as implemented in EAs, can be seen as a dynamic process driven by two factors: (natural) *selection*, which favours the survival of the fittest, and *genetic variation*, which introduces new individuals, some of which could potentially outperform the best so far. Neither factor is sufficient on its own: without selection, the search for the best individual will become completely random; without genetic variation, there will be nothing that selection can act upon as a uniform population of identical individuals reaches a dead end.

EAs employ several techniques that study the individual's fitness in the context of the whole population and modify it to preserve the balance between the forces of selection and those increasing genetic variation. Fitness *scaling* is used to reduce the risk of clones of one 'superindividual' taking over the whole population in the early stages of the search when the individuals' fitness is very varied and generally low. Also, in a population with a minimum of variety in the fitness, scaling helps emphasise the existing differences and promote the best individuals more strongly. In either case, scaling aims at normalising the differences between the fitness of individuals with respect to the extremes present in the population. This could be done in a number of ways, from using a linear scaling function to ranking individuals according to their fitness and substituting rank for the original fitness.

Crowding and niching are other techniques used to modify fitness in order to avoid overpopulating parts of the search space in favour of others. In general, these techniques are based on partitioning the range of all individuals into *niches* and then reducing the fitness of individuals in overpopulated (*crowded*) niches [Mah95].

2 Extended Fitness

In all of the above discussion fitness represents the quality of an entire individual, of all components of a solution assembled and put to the test in its integrity. However, the combination of selection and variation on which evolutionary algorithms are based would produce better results than random search (variation alone) only if 'good' individuals favoured by selection contain good building blocks, that is, such that appear in the best solution(s) one is searching for. In other words, if evolution is seen as search for the best set of inherited features, this search should be at least partially decomposable. Any success when using

EAs is based on this assumption. On the other hand, many designs superior to the ones existing in nature were never produced by evolution, as it would require the concerted change of many elements.

In biology, the most important division among supporters of evolution until the middle of the twentieth century and one that persists to the present day, although unevenly represented, is based on the way variation is introduced in the genes of the offspring. The theory that has become known as Lamarckian evolution states that variation is guided by parents' experience, i.e., it introduces new inherited features that are likely to increase the offspring's fitness. The nowadays prevailing Darwinists on the other hand believe that evolution is 'blind', i.e., based on unguided variation, and it is left to selection to separate the wheat from the chaff. Both schools of thought though had in common the view that the entity that competes for survival is the individual with its whole set of genes.

Since the 1960s, Neo-Darwinism has brought a shift of perspective putting individual genes at the centre of attention and stating that each gene can be seen as competing for survival with all other genes present in the genetic pool (i.e., the entire population) whereas individuals are only an intermediate step, a vessel ensuring the replication of the gene. The higher the individual's fitness, the better the chance that a gene it carries will be replicated. It is the gene that is subject to selection, and its overall potential to spread in the population is given by the combined reproductive success of all individuals carrying that gene [Daw82]. We shall modify the title of Dawkins's book, and call *extended fitness* any definition of fitness that goes beyond the reproductive success of the individual, and takes into account, in various ways, the reproductive success of that part of the population with which the individual in question shares genes.

As in population genetics, fitness[2]¹ can be defined to measure the relative advantage that a gene gives to its carrier with respect to all other genes that can appear in the same locus. Fitness of the whole genome, here denoted as *fitness*[3], then can be computed as the averaged contribution of all its loci [Fal81]. This could be illustrated on the example of maximising the two-dimensional function in Fig. 1. Let us for instance consider the population {**bw**, **bx**, **bz**, **cw**, **cy**, **dx**, **dz**, **ew**, **ey**, **ez**}. If the fitness of individuals is used for selection, the top-ranking individual in the population, **ez**, will be also the one that does not contain any of the values (alleles) of the two genes forming the best solution **cx**. On the other hand, it can be seen that for the given population, and each gene, of all alleles actually present it is allele **c** for Gene 1 and allele **x** for Gene 2 that have the highest fitness[2]. Indeed, the average fitness (*af*) of all individuals in which allele **c** appears is $(7+7)/2 = 7$, as compared to those containing alleles **b** with $af = (7+1+1)/3 = 3$, **d** with $af = (7+1)/2 = 4$, and $af = (10+1+1)/3 = 4$ for allele **e**. The superiority of allele **x** can be established in the same way.

In the context of Lamarckian evolution, the so computed fitness can be used in the selection phase to favour carriers of the best alleles or for a guided mutation

¹ Here indices are used to describe the various meanings of fitness as introduced by Dawkins [Daw82].

Gene 2	z	1	1	1	1	10
	y	1	1	7	1	1
	x	1	7	100	7	1
	w	1	1	7	1	1
	v	1	1	1	1	1
		a	b	c	d	e
		Gene 1				

Fig. 1. Example fitness function of a two-dimensional optimisation task. The range of each parameter encoded as one gene is shown along the axes.

that produces these alleles more often than any other one. In a Darwinian set-up, the EA can scale up the fitness of individuals that carry genes with high fitness. The effect on the population in either case would be an increased number of copies of these genes, which should also increase the chance of finding the best individual(s). One can also expect that some individuals with low personal merits (fitness) that would normally die out would now survive just because they are related to (share genes with) other, more successful individuals.

Let us define the fitness[3] of an individual as the simple average of the fitness[2] values for all loci. So, for instance, the fitness[3] of individual **ez** would be $(4+4)/2 = 4$. Fig. 2 shows how the landscape of the optimisation task from Fig. 1 changes when fitness[3] is used.

The local maximum **ez** has now been flattened, and individuals containing either of the alleles that form the global maximum **cx** have been valued more highly. This has been guaranteed by the fact that the close neighbours of **cx** — those at Manhattan distance one — also have an increased individual fitness. Since such close neighbours by definition share many genes, they will mutually boost each other’s fitness[2], and therefore, fitness[3] as well. This could be expected to help the survival of a part of the population in the neighbourhood of the individual fitness maxima, even if better individuals have been found elsewhere. As a result, copies of many or all of the genes representing the maximum at hand will await crossover to put them together. Also, maintaining a number of individuals close to a maximum increases the chances to find that maximum as a result of a small mutation.

The task of continuous function optimisation seems well suited to demonstrate the potential benefits of using fitness[3] in EAs, and will be used here as a testbed. Fig. 3 shows the exact algorithm computing this type of extended fitness.

Gene 2	z	1	3.5	1	4	4
	y	1	1	5.5	1	4
	x	1	5	100	5.5	1
	w	1	3	5	1	3.5
	v	1	1	1	1	1
		a	b	c	d	e
		Gene 1				

Fig. 2. Fitness[3].

3 Kin Selection and Inclusive Fitness

Increasing the relative weight of a poor individual's fitness means reducing the reproductive success of its richer relatives. This could have twofold implications: preserving a greater variety of genes (alleles) in the pool and slowing down the climb towards a (possibly local) peak in the fitness function.

From a gene's point of view, such altruism towards relatives (also known as *kin selection*) would have clear benefits, provided the fitness sharing scheme used increased in average the probability of survival of the individuals that have this gene. As in EAs the population size is fixed, this can only happen by reducing the chances of survival of the rest of the population.

Neither classical Darwinism nor EAs at their simplest (cf. Simple Genetic Algorithm [Gol89]) leave any room for altruistic behaviour: any inherited feature that hinders one's fitness—and this is the very nature of altruism—is bound to fade and disappear from the gene pool. On the other hand, Hamilton shows the feasibility of such behaviour on numerous examples from nature [Ham64b] and goes on to explain it through a mathematical model of the impact altruism has on the reproductive success of genes [Ham64a]. Hamilton defines *inclusive fitness* (to which Dawkins, in his own notation, refers as fitness[4]) as the relative change in the reproductive success of an individual and all its relatives (weighted by the appropriate coefficient of *relatedness* r expressing the proportion of shared genes) that a given behaviour of this individual brings about with respect to another, baseline behaviour used for comparison.

Omitting the formulae, we could illustrate the main idea on the case of an individual who dies to save the life of his three children ($r = 0.5$). Such behaviour minimises the loss of this individual's genes, as only one copy of them will disappear from the environment, as opposed to $3 \times r = 1.5$ (in average) in the case of a selfish parent who lets his children die. Altruism in such circumstances, far from being against the spirit of natural selection, promotes the survival of one's

```

procedure evaluateExtendedFitness[3]
for each chromosome C do
|extFitness(C) = 0
| for each locus L do
| | extFitness(C,L) = 0
| | numberOfMatches=0
| | for each chromosome C' do
| | | if locus L in C = locus L in C'
| | | | increment numberOfMatches
| | |_ |_ extFitness(C,L) += fitness(C')
| | extFitness(C,L) /= ( numberOfMatches * chromosomeLength )
|_ |_ extFitness(C) += extFitness(C,L)

```

Fig. 3. Procedure computing fitness[3].

own genes. If such altruism is encoded as an inherited feature, it will also be promoted by evolution for the same reason.

We have already studied kin selection in the case of a simulated two-dimensional ecosystem where a species of mobile agents seek stationary resources (food and water) and are capable of reproducing if the couple can afford the cost, a one-off payment that is subtracted from each parent's stock of food and water and is given to the offspring created. A part of the population carries a gene forcing its bearer to share some of its resources when requested by a poorer individual. The amount of help is proportional to the degree of (perceived or actual) kinship (relatedness r), so, for instance, one gives a sibling or parent, with whom one shares in average 50% of the genes, half the amount of what one would give an identical twin [TK02].

We have shown that under certain conditions, e.g., when perfect knowledge of the degree of kinship (the proportion of genes shared) between two individuals is available, an inherited feature promoting kin selection is feasible and is preserved by evolution. It does not take over the whole population, but rather exists side by side 'selfish' individuals. In all experiments carried out in which altruism survived, the ratio between altruistic and selfish individuals always converged on to approximately 50% : 50% regardless of its starting value. If some inherited features are unobservable, and the rest are used to estimate the degree of kinship, the *sharing function* specifying the help (share of one's fitness) that one provides to a relative of a given degree becomes a major factor in determining whether altruism should survive [TK02].

In the above simulation, sharing had two easily identifiable benefits for one's genes, either of which had to do with a non-linearity in the relationship between the amount of resource an agent has and its fitness. The first case is when an agent is not able to afford the daily cost of living, and would die out unless is helped by a relative. The other case has to do with the amount needed to mate and have offspring. Helping a poorer relative to afford the cost of reproduction means that the genes the helper shares with the helped have twice as

```

procedure evaluateExtendedFitness[34]
for each chromosome C do
|extFitness(C) = 0
| for each locus L do
| | extFitness(C,L) = 0
| | numberOfMatches=0
| | for each chromosome C' do
| | | if locus L in C = locus L in C'
| | | | increment numberOfMatches
| | |__|_ extFitness(C,L) += fitness(C') * r(C,C')
| | extFitness(C,L) /= ( numberOfMatches * chromosomeLength )
|_ |_ extFitness(C) += extFitness(C,L)

```

Fig. 4. Procedure computing fitness[34].

much chance of replicating, while for its remaining genes that chance remains unchanged.

Since the population size of an EA is typically a small fraction of the search space, the use of fitness[2] in the definition of fitness[3] assumes that the effect a gene has on the reproductive success of its carrier is relatively independent from the context (i.e., the carrier's other genes). While that may be the case for features such as good sight in the animal world, it is easy to see how in the case of continuous function optimisation the importance one assigns to the fitness of an individual with whom one shares some genes decreases with the increasing difference between the two genomes. Indeed, when decoded, such different genomes are likely to correspond to different parts of the landscape of the function being optimised, hence the knowledge of its value in one area would be of little or no relevance to the discovery of maxima in another (with periodic functions being an important exception). Inclusive fitness provides the instrument to measure the effect one genotype has on the *average* reproductive success of all genes in another genotype. The averaging character of this interaction comes from the fact that it is impossible in general to recognise the individual genes one carries. However, in EAs, this is not the case as the complete genotypes of all individuals are available. The EA designer can use that to his/her benefit.

The idea of inclusive fitness will be exploited here to define fitness[34], modifying the definition of fitness[2] to account for the link between genomic distance and fitness relevance discussed at the beginning of Section 2. Now, for two individuals sharing a gene (allele), the contribution one makes to the other's fitness[2] will be reduced by the factor r , the degree of relatedness (see Fig. 4). The overall effect of fitness[34] can be seen as a combination of fitness[3] and fitness[4] (hence the name).

4 Related Work

The definitions of extended fitness introduced here can be compared with explicit fitness sharing, proposed by Goldberg and Richardson [GR87]. In this case, the *shared fitness* $f'(i)$ of an individual is computed from the original fitness $f(i)$ as follows:

$$f'(i) = \frac{f(i)}{\sum_{j=1}^n sh(d(i, j))} \quad (1)$$

In Equation 1, d represents the distance between two individuals. The sharing function sh is of the type

$$\begin{aligned} sh(d) &= 1 - (d/\sigma)^\alpha && \text{if } d < \sigma \\ &= 0 && \text{otherwise} \end{aligned} \quad (2)$$

In the above equation, the constant α is typically set to 1, and σ is the distance within which the presence of other individuals leads to one's fitness being reduced in order to avoid crowding. Despite its name, fitness sharing only reflects the impact crowding around limited resources has on one's fitness. It is clearly different from our last definition of extended fitness as fitness sharing, unlike extended fitness, does not take into account the fitness of neighbours.

To highlight the differences, fitness sharing was applied to the example used in Section 2 with parameters $\alpha = 1$ and $\sigma = 3$. The resulting values displayed in Fig. 5 were rescaled to preserve the total fitness of the population unchanged.

5 Experimental Evaluation

The functions shown in Fig. 6–7 were chosen for optimisation for their properties: a narrow peak of the global maximum, and one or more local maxima, reachable through hill climbing for most of the independent variable range. The solution was encoded as a 16-digit binary chromosome from the interval $< 0, 6 >$; then a genetic algorithm was run 100 times and the best ever individual in each run recorded. (More details on the experiments: population size = 20, generations = 130, crossover rate = 75%, mutation = 0.5%.) These individuals are displayed in Fig. 8–12 for the function in Fig. 6 and the following definitions of fitness: individual fitness with ranking, fitness[34] and fitness[3]; the last two series of experiments were using fitness[34], resp. fitness[3] to select 50% of the population, and unscaled fitness for the other half.

The results in Table 1 and Figures 8–12 show that fitness[3] in average outperformed ranking. Even for fitness[34], which has virtually the same average score as ranking, it can be seen from Table 2 that of the three modes ($0 < x < 0.15$, $0.15 < x < 1$, $1 < x \leq 7$), fitness[34] finds the best two more often and more accurately than ranking, which misses one of the modes altogether. Experiments with the function in Fig. 7 show similar results (see Table 3) where the difference between the probabilities with which each method finds the best mode is statistically significant w.r.t. a T-test with $\alpha = 0.001$.

Gene 2	5 z	1	1.28	1	1.02	10.98
	4 y	1	1	5.98	1	1.02
	3 x	1	7.17	100	5.98	1
	2 w	1	1.10	7.17	1	1.28
	1 v	1	1	1	1	1
		a	b	c	d	e
		1	2	3	4	5
		Gene 1				

Fig. 5. Fitness sharing: Manhattan distance is used.

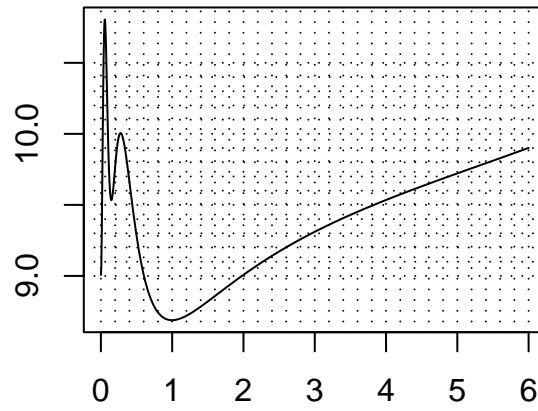


Fig. 6. $f_1(x) = \sin(x + 0.3)^{-2} * e^{0.5x} - 4x + 10$

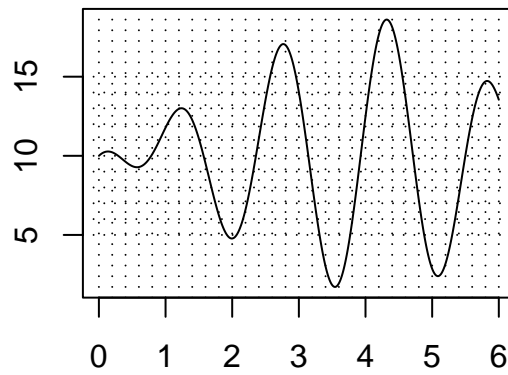


Fig. 7. $f_2(x) = \sin(4x) * (\exp(x/2) - 4x) + 10$.

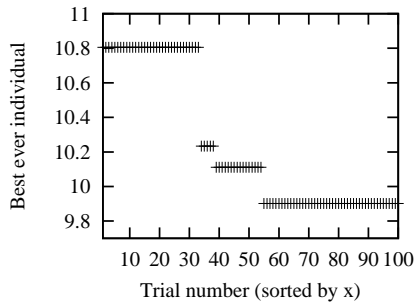


Fig. 8. Baseline experiment: GA with ranking.

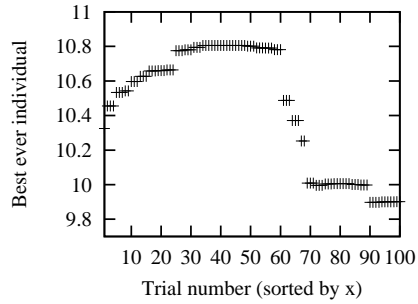


Fig. 10. GA with fitness[3].

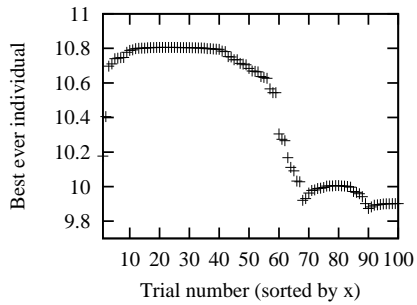


Fig. 9. GA with fitness[34].

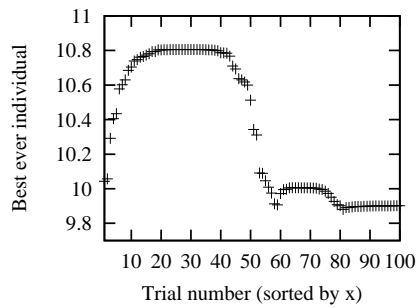


Fig. 11. 50% fitness[3] and 50% unscaled fitness.

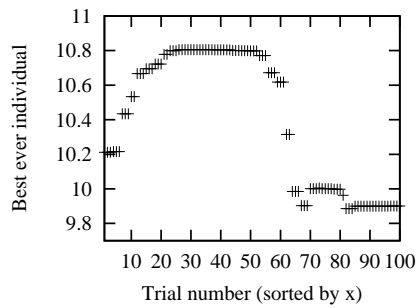


Fig. 12. 50% fitness[34] and 50% unscaled fitness.

Table 1. Average values of best ever individuals (over 100 runs).

	Rank.	F[34]	F[3]	$F[34]^{50\%}$	$F[3]^{50\%}$
Score	10.329	10.250	10.446	10.333	10.400
StErr	0.038	0.040	0.036	0.040	0.039

Table 2. Distribution of best ever individuals by modes for $f_1(x)$ (same 100 runs).

	Mode 1	Mode 2	Mode 3
F[34] (indiv.)	67	22	11
F[34] (av.score)	10.668	9.978	9.895
Ranking (indiv.)	54	0	46
Ranking (av.score)	10.547	–	9.901

Table 3. Distribution of best ever individuals by modes for $f_2(x)$.

	2nd highest mode	highest mode
F[34] (indiv.)	4	96
F[34] (av.score)	17.057	18.556
Rank. (indiv.)	27	73
Rank. (av.score)	16.909	18.609

6 Conclusions

The definitions of extended fitness proposed here are based directly on principles of evolution observed in nature or their simple modifications. Using extended fitness replaces altogether the need for scaling, as currently used. The effects of extended fitness were studied on the case of genetic algorithms, but the approach is equally applicable to evolutionary programming. The approach is promising, as in the experiments, extended fitness seems to show improved convergence on the global maximum mode at the cost of a small drop in the accuracy with which the actual global maximum is found. This is an exciting find that would require more experiments with various type of data (discrete, finite range) and other difficult to optimise, deceptive, functions [GKH92].

Acknowledgement

Part of this work was done during the author's research term visit to the Department of Intelligent Systems at the Josef Stefan Institute in Ljubljana, Slovenia. An almost identical version of the paper was submitted to ICML in Feb 2003.

References

- [Daw82] R. Dawkins. *The Extended Phenotype*. Oxford University Press, 1982.
- [Fal81] D.S. Falconer. *Introduction to Quantitative Genetics*. Longman, London, second edition, 1981.
- [GKH92] D. E. Goldberg, D. Kalyanmoy, and J. Horn. Massive multimodality, deception, and genetic algorithms. Technical Report 92005, Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign, 1992.
- [Gol89] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
- [GR87] D. E. Goldberg and J. Richardson. Genetic algorithms with sharing for multimodal function optimization. In *Proc. of the Second Intl. Conference on Genetic Algorithms*, pages 41–49, 1987.
- [Ham64a] W. D. Hamilton. The genetical evolution of social behaviour I. *Journal of Theoretical Biology*, 7:1–16, 1964.
- [Ham64b] W. D. Hamilton. The genetical evolution of social behaviour II. *Journal of Theoretical Biology*, 7:17–52, 1964.
- [Mah95] S. Mahfoud. *Niching Methods for Genetic Algorithms*. PhD thesis, University of Illinois, Urbana-Champaign, 1995.
- [TK02] H. Turner and D. Kazakov. Stochastic simulation of inherited kinship-driven altruism. *Journal of Artificial Intelligence and Simulation of Behaviour*, 1(2), 2002.