Signal, Noise, and Genetic Algorithms

Mike Rudnick
Oregon Graduate Institute
&
David E. Goldberg
University of Illinois at Urbana-Champaign

Oregon Graduate Institute
Department of Computer Science
and Engineering
19600 N.W. von Neumann Drive
Beaverton, OR 97006-1999 USA

Technical Report No. CS/E 91-013

June, 1991

Signal, Noise, and Genetic Algorithms

Mike Rudnick
Oregon Graduate Institute
&
David E. Goldberg
University of Illinois at Urbana-Champaign

Technical Report No. CS/91-013
Oregon Graduate Institute of Science and Technology
June 1991

Also available as
IlliGAL Report No. 91004
Illinois Genetic Algorithms Laboratory

Department of Computer Science and Engineering
Oregon Graduate Institute of Science and Technology
19600 NW von Neumann Dr.
Beaverton, OR 97006-1999

Signal, Noise, and Genetic Algorithms

Mike Rudnick
Department of Computer Science and Engineering
Oregon Graduate Institute
Beaverton, OR 97006-1999
rudnick@cse.ogi.edu

David E. Goldberg
Department of General Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801-2996
goldberg@vmd.cso.uiuc.edu

Abstract

Signal, noise, and the signal-to-noise ratio are defined, and a Walsh basis expression is derived for each. *Signal* is a measure of the force tending towards convergence within a competition partition. *Noise* is a measure of the force preventing the partition's signal from having effect. The *signal-to-noise ratio* is a measure reconciling these antagonistic effects. The conjecture is made that the relative magnitudes of signal and noise determine the degree to which convergence occurs in each competition partition.

1 Introduction

Genetic algorithms (GAs) function by sampling schema fitness, but because populations of modest size are generally used, schema fitness variance is a primary source of stochastic noise which can hamper correct evaluation of building blocks (Holland, 1973; De Jong, 1975). To account for this effect, Goldberg and Rudnick (Goldberg & Rudnick, 1991) have derived a Walsh basis expression for static schema fitness variance and shown how it can be used to perform static population sizing to control the probability of incorrect evaluation of building blocks.

In this note the previous work is extended. First, the conceptual context is set in section 2. Then, sections 3, 4, and 5 give rigorous definitions of signal, noise, and the signal-to-noise ratio, respectively. For each, a static Walsh basis expression is derived. Finally, section 6 briefly summarizes.

2 GAs and Noise

Each schema may be viewed as a partial solution defined by its fixed positions. Schemata are organized into competition partitions — competing partial solutions, or sets of schemata fixing the same bit positions. Thus, every complete solution, such as a member of a GA's population, belongs to exactly one schema within each competition partition. In effect, the partial solutions within a competition partition compete for representation in the GA's population of complete solutions.

Each schema has as its *schema fitness*, $f(\mathbf{h})$, the average fitness of its elements. The partial solutions within each partition will, in general, have a spread in their fitnesses, which can be thought of as the partition's selection pressure, or convergence signal. The greater the fitness spread among the partition's schemata (partial solutions), the greater the partition's signal strength, and the greater will be the convergence occurring within the partition. It is this spread in fitness which enables the GA to select for the better partial solutions in the partition.

The GA continually operates to enrich succeeding populations with respect to the fitness of the partial solutions represented within the population. It does this by preferentially selecting individual solutions having above-average fitness with respect to the current population. When some partitions have strong signals while other partitions have weak signals, the GA pays attention (through the mechanism of selection) to the strong signals at the expense of the weaker signals. In effect, the GA has only so much selection attention to distribute among the various competition partitions.

Thus from the point of view of a particular partition, all other partition's signals contribute to noise competing with its signal. The net effect is that the 'signal' from partitions with weak signals is lost among the loud 'noise' from partitions with strong signals. This happens when the difference between the strong and weak partition signals is large relative to the size of the population.

3 Signal

Schema fitness variance, or *collateral noise*, may be expressed as

$$\mathrm{var}(f(\mathbf{h})) \;=\; \overline{f^2(\mathbf{x})} - \overline{f(\mathbf{x})}^2, \tag{3.1}$$

where $\overline{f^2(\mathbf{x})}$ and $\overline{f(\mathbf{x})}^2$ are averages over $\mathbf{h}$.

Define the measure of the force tending toward convergence within a competition partition as the square root of the variance of the schema fitnesses of the schemata within the partition, and call it the *partition signal strength*, $S(\mathbf{J})$, or in mathematical form,

$$S^2(\mathbf{J}) \;=\; \mathrm{var}(f(\mathbf{J})). \tag{3.2}$$

The reason for the squaring is that variance is itself a squared measure.

Restating equation 3.2 in terms of the variance expression given in equation 3.1,

$$S^2(\mathbf{J}) \;=\; \overline{f^2(\mathbf{h})} - \overline{f(\mathbf{h})}^2, \tag{3.3}$$

where $\mathbf{h}$ varies over the schemata in $\mathbf{J}$. Tackling the first term in equation 3.3,

$$\overline{f^2(\mathbf{h})} \;=\; \frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}} f^2(\mathbf{h}), \tag{3.4}$$

from the definition of the mean, where $|\mathbf{J}|$ is the number of schemata in $\mathbf{J}$. Substituting the expression for $f(\mathbf{h})$ from the Walsh-schema transform for a uniform population (Bethke, 1981; Goldberg, 1989),

$$\overline{f^2(\mathbf{h})} \;=\; \frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}} \left(\sum_{j\in J_i(\mathbf{h})} w_j \psi_j(\mathbf{h}) \right)^2, \tag{3.5}$$

where $J_i(\mathbf{h})$ is the standard Walsh-schema transform index set generated by replacing $*$s with 0s in each schema template of each schema in $\mathbf{h}$'s partition, and interpreting the resulting binary strings as integers. Expanding the quadratic yields

$$\overline{f^2(\mathbf{h})} = \frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}}\sum_{j,k\in J_i(\mathbf{h})} w_j w_k \psi_{j,k}(\mathbf{h}). \tag{3.6}$$

The two-dimensional Walsh function, $\psi_{j,k}(\mathbf{h})$, may be defined as the product of two one-dimensional Walsh functions, or

$$\psi_{j,k}(\mathbf{h}) = \psi_j(\mathbf{h})\psi_k(\mathbf{h}) = \prod_{i=1}^{l}(-1)^{h_i j_i}\prod_{i=1}^{l}(-1)^{h_i k_i} = \prod_{i=1}^{l}(-1)^{h_i(j_i+k_i)}, \tag{3.7}$$

where subscripts identify bits and $*$s in h are replaced by 0s. Note that all schemata within a single partition share the same index set, i.e., $J_i(\mathbf{h}_j) = J_i(\mathbf{h}_k)$ for all $\mathbf{h}_j, \mathbf{h}_k \in \mathbf{J}$. Thus, the same Walsh coefficient products occur $|\mathbf{J}|$ times in the outer summation, but with possibly different signs due to the action of $\psi_{j,k}(\mathbf{h})$. In fact, because of the orthogonality of the Walsh basis and because each partition covers the genome space (the space of all genotypes), when $j \neq k$ an equal number of plus and minus terms occur for each Walsh product pair, $w_j w_k$, resulting in the elimination of all off-diagonal products. Thus, equation 3.6 reduces to

$$\overline{f^2(\mathbf{h})} = \sum_{j\in J_i(\mathbf{J})} w_j^2, \tag{3.8}$$

where $J_i(\mathbf{J})$ has been substituted for $J_i(\mathbf{h})$ by noting that all schemata in the same partition share the same index set.

The last term of equation 3.3 may be expanded analogous to equations 3.4, 3.5, and 3.6, resulting in

$$\overline{f(\mathbf{h})}^2 = \left[\frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}} f(\mathbf{h})\right]^2 = \left[\frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}}\sum_{j\in J_i(\mathbf{h})} w_j\psi_j(\mathbf{h})\right]^2. \tag{3.9}$$

Likewise, because $\mathbf{J}$ covers the space of all possible genotypes and Walsh functions are orthogonal, equation 3.9 simplifies to

$$\overline{f(\mathbf{h})}^2 = \left[\frac{|\mathbf{J}|}{|\mathbf{J}|}w_0\right]^2 = w_0^2, \tag{3.10}$$

as it must since the average of the averages of equal sized partition elements is simply the average of the underlying space, which for genome space is w_0.

A Walsh expression for a partition's squared signal, $S^2(\mathbf{J})$, may now be formed by substituting the Walsh basis expressions for $\overline{f^2(\mathbf{h})}$ and $\overline{f(\mathbf{h})}^2$ from equations 3.8 and 3.10, respectively, into equation 3.3, producing

$$S^2(\mathbf{J}) = \sum_{j\in J_i(\mathbf{J})} w_j^2 - w_0^2. \tag{3.11}$$

Since w_0 is in every $J_i(\mathbf{J})$, the effect of subtracting w_0^2 is to remove it from the equation altogether, which is equivalent to removing zero from the index set, and equation 3.11 may be restated as

$$S^2(\mathbf{J}) = \sum_{j\in J_i(\mathbf{J})-\{0\}} w_j^2, \tag{3.12}$$

where the minus sign in the index expression denotes set difference. As an example, the squared signal of partition $\mathbf{J} = ff*$ is $S^2(ff*) = w_2^2 + w_4^2 + w_6^2$. Equation 3.12 is a remarkably simple expression — a partition's squared signal strength is just the sum of the squares of the Walsh coefficients of order one or greater in the partition's index set. Further, it is completely general, in the sense that nothing has been assumed about the form of the fitness function.

4 Noise

Other partition's signals contribute to noise competing with the signal of the partition under consideration for the control of the GA's selection process. Define *partition root mean squared noise*, $C(\mathbf{J})$, as the square root of the average of the collateral noise values for each schema in the competition partition under consideration, or

$$C^2(\mathbf{J}) = \overline{\mathrm{var}(f(\mathbf{h}))} = \frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}} \mathrm{var}(f(\mathbf{h})), \tag{4.13}$$

where $\overline{\mathrm{var}(f(\mathbf{h}))}$ is the average collateral noise among the schemata in the partition.

Goldberg and Rudnick (Goldberg & Rudnick, 1991) derive $\mathrm{var}(f(\mathbf{h}))$ as

$$\mathrm{var}(f(\mathbf{h})) = \sum_{(j,k)\,\in\, J^2_{\oplus}(\mathbf{h})-J^2_i(\mathbf{h})} w_j w_k \psi_{j,k}(\mathbf{h}), \tag{4.14}$$

where $J^2_{\oplus}(\mathbf{h})$ denotes the set of index pairs for which the sum over $\mathbf{h}$ is nonzero, and $J_i^2(\mathbf{h})$ is the cross product of $J_i(\mathbf{h})$ with itself. Since equation 4.14 assumes the uniform population, the resulting expression for noise also assumes the uniform population. Substituting equation 4.14 into equation 4.13 yields

$$C^2(\mathbf{J}) = \frac{1}{|\mathbf{J}|}\sum_{\mathbf{h}\in\mathbf{J}} \sum_{(j,k)\in J^2_{\oplus}(\mathbf{h})-J^2_i(\mathbf{h})} w_j w_k \psi_{j,k}(\mathbf{h}). \tag{4.15}$$

As with equation 3.6, and for the same reason, the off-diagonal product terms are zero and can thus be eliminated from the index set. Note that the remaining (diagonal) entries in the inner summation's index set are exactly the elements which are not in $J_i(\mathbf{J})$. Thus, the index set is $\overline{J_i(\mathbf{J})}$, the complement of $J_i(\mathbf{J})$. As in equation 3.6 the outer summation cancels against the $1/|\mathbf{J}|$ term because each term is added $|\mathbf{J}|$ times, and equation 4.15 simplifies to

$$C^2(\mathbf{J}) = \sum_{j\in\overline{J_i(\mathbf{J})}} w_j^2. \tag{4.16}$$

Continuing the previous example, the squared noise of partition $\mathbf{J} = ff*$ is $C^2(ff*) = w_1^2 + w_3^2 + w_5^2 + w_7^2$.

Note that signal plus noise equals a constant determined by the particular fitness function used, or

$$S^2(\mathbf{J}) + C^2(\mathbf{J}) = \sum_{j\in\{1,2,\ldots,l-1\}} w_j^2. \tag{4.17}$$

Thus signal and noise can each be expressed in terms of the other, as in

$$C^2(\mathbf{J}) = \sum_{j\in\{1,2,\ldots,l-1\}} w_j^2 - S^2(\mathbf{J}) \tag{4.18}$$

and

$$S^2(\mathbf{J}) = \sum_{j \in \{1,2,\ldots,l-1\}} w_j^2 - C^2(\mathbf{J}). \tag{4.19}$$

Next the equations for signal and noise are combined to give a measure of which has the upper hand in a particular situation.

5 Signal-to-Noise Ratio

The *signal-to-noise ratio*, $R(\mathbf{J})$, is a measure reconciling the antagonistic effects of signal strength and noise. It is defined as

$$R(\mathbf{J}) = \frac{S(\mathbf{J})}{C(\mathbf{J})}. \tag{5.20}$$

Substituting in the expressions for squared signal and squared noise from equations 3.12 and 4.16 gives

$$R(\mathbf{J}) = \sqrt{\frac{\sum_{j \in J_i(\mathbf{J})-\{0\}} w_j^2}{\sum_{j \in \overline{J_i(\mathbf{J})}} w_j^2}}. \tag{5.21}$$

Completing the ongoing example for partition $\mathbf{J} = ff*$,

$$R(ff*) = \sqrt{\frac{w_2^2 + w_4^2 + w_6^2}{w_1^2 + w_3^2 + w_5^2 + w_7^2}}. \tag{5.22}$$

Note the way in which each Walsh coefficient other than w_0 occurs exactly once in equation 5.21, contributing to either squared signal (the numerator) or squared noise (the denominator). That w_0, the average fitness over the entire genome, does not participate in S, C, or R, makes sense since both signal and noise are composed exclusively of combinations of variances, in which the genome's average fitness plays no part. That equation 5.21 is so simple is yet another demonstration of how the Walsh basis respects competition partitions.

Finally, note that $R(\mathbf{J})$ is undefined for the competition partition containing order l schemata, since its noise is zero. Likewise, $R(\mathbf{J})$ for the competition partition whose single element is the order-zero schema is zero, since $S(\mathbf{J}) = 0$.

A partition's signal, noise, and signal-to-noise ratio each induce a total order on competition partitions. That is to say, each may be used to rank competition partitions into a linear sequence in which ties are possible. We conjecture the signal-to-noise ratio is a measure appropriate to reconciling the antagonistic effects of signal and noise. If true, the signal-to-noise ratio gives convergence order among competition partitions due to selection in the absence of operators. And of course, $R(ff\ldots f)$, the signal-to-noise ratio of the partition fixing all positions, gives the net convergence force acting on the GA population.

6 Conclusion

We have defined signal, noise, and the signal-to-noise ratio for competition partitions. For each a Walsh basis expression has been derived.

The signal-to-noise ratio has long been used in decision theory to make probabilistic statements about the likelihood of an event based on the relative magnitude of the event's signal and noise. Further research is needed to clarify the function of competition partition signal-to-noise ratio in GA convergence.

Acknowledgments

The authors wish to thank Kalyanmoy Deb and Dirk Thierens for useful discussions. This material is based upon work supported by the National Science Foundation under Grants CTS-8451610 and ECS-9022007 and by U.S Army Contract DASG60-90-C-0153. The first author acknowledges departmental support from the Oregon Graduate Institute Department of Computer Science and Engineering.

References

Bethke, A. D. (1981). Genetic algorithms as function optimizers (Doctoral dissertation, University of Michigan). *Dissertation Abstracts International*, *41(9)*, 3503B. (University Microfilms No. 8106101)

De Jong, K. A. (1975). An analysis of the behavior of a class of genetic adaptive systems (Doctoral dissertation, University of Michigan). *Dissertation Abstracts International*, *36(10)*, 5140B. (University Microfilms No. 76–9381)

Goldberg, D. E. (1989). Genetic algorithms and walsh functions: Part I, a gentle introduction. *Complex Systems*, *3*, 129–152.

Goldberg, D. E., & Rudnick, M. (1991). Genetic algorithms and the variance of fitness. *Complex Systems*. in press

Holland, J. H. (1973). Genetic algorithms and the optimal allocations of trials. *SIAM Journal of Computing*, *2(2)*, 88—105.