

# Mathematics Awaits: Commentary on “Genetic Programming and Emergence” by Wolfgang Banzhaf\*

Lee Altenberg

Ronin Institute; Associate Editor, *BioSystems*; altenber@hawaii.edu

July 15, 2013

## Abstract

Banzhaf provides a portal to the subject of emergence, noting contentious concepts while not getting sucked into fruitless debate. Banzhaf refutes arguments against downward causation much as Samuel Johnson kicks a stone to refute Berkeley — by pointing to concrete examples in genetic programming, such as the growth of repetitive patterns within programs. Repetitive patterns are theoretically predicted to emerge from the evolution of evolvability and robustness under subtree exchange. Selection and genetic operators are co-equal creators of these emergent phenomena. GP systems entirely formal, and thus their emergent phenomena are essentially mathematical. The emergence of Lagrangian distributions for tree shapes under subtree exchange, for example, gives a glimpse of the possibilities for mathematical understanding of emergence in GP. The mathematics underlying emergence in genetic programming should be pursued with vigor.

KEYWORDS: evolvability, robustness, subtree exchange, mathematics, matrix theory, Lagrange distribution

Banzhaf’s lucid essay provides a very useful portal to the subject of emergence, emergent phenomena, and their related concepts: hierarchy, parts, wholes, upward and downward causation, and mechanism. Banzhaf is careful to note where the concepts are still surrounded by debate, and to warn us of where the “battlefields” are. Of course it is tempting to be lured into battle (that’s why they are battles) but Banzhaf astutely avoids this temptation, and after having sketched out the conceptual territory of emergence, brings us to the manifestations of emergence in genetic programming.

Regarding the debate about downward causation, Banzhaf’s approach is reminiscent of Samuel Johnson’s approach to “Bishop Berkeley’s ingenious sophistry to prove the nonexistence of matter.” “Johnson answered, striking his foot with mighty force against a large stone, till he rebounded from it — ‘I refute it *thus*.’ ” (Boswell, 1792, p. 286). Similarly, Banzhaf addresses claims of the nonexistence of downward causation by pointing to the rich phenomena observed in genetic programming as a result of

---

\*Banzhaf (2013), *Genetic Programming and Evolvable Machines*, in press.

selection, a clear influence of the whole upon the parts: it is the whole program that determines its value for the objective function, and the parts we observe are only those that selection has kept in existence, based on the values of the objective function.

Banzhaf provides a concrete illustration of emergence in genetic programming: the growth of repetitive patterns within programs as their evolution proceeds. He conjectures “that these patterns are emergent, a result of the presence of downward causation realized through selection.” A dynamical theory for the emergence of repetitive patterns and its mathematical underpinnings are sketched in two of Banzhaf’s citations ([1] and [2] Altenberg 1994b,a) — my own early work on the evolution of evolvability or robustness in genetic programming.

The mechanism is simply this: the genetic operator, subtree exchange, provides a second level of replicator besides the whole program — the subtree. While the replication rate of *whole* programs is determined by their fitness, the replication rate of *parts* of programs is determined by their effect on the *change* in fitness: patterns whose addition to a program have a higher than average chance of increasing fitness will increase their copy number *within programs* during early periods of evolution when the mean fitness of the population is increasing; patterns that increase the robustness of programs under action of the genetic operators will increase their copy number during later periods of evolution when the mean fitness of the population is stagnating.

What should be emphasized is that selection and the genetic operators are co-equal partners in the creation of these emergent phenomena. There is a tendency to consider selection as ‘signal’ and genetic operators as variation-producing ‘noise’, and so selection is given credit for any meaningful outcomes. But genetic operators are not mere variation-producing noise; they are specific actions on the structures representing the search space, and the relationships between these structures and the objective function typically have abstract structure themselves. To view the emergent phenomena in genetic programming as primarily a result of selection, is, I believe, a diversion from the real source: the *relationship* between selection and the genetic operators.

Emergence in genetic programming has the feature that it not only occurs ‘before our eyes’, but also ‘in our hands’ — GP systems are typically entirely formal, and we know and have control over every detail of the objects in the system. As completely formal systems, the emergent phenomena we observe are essentially mathematical in nature. Their mathematical nature tempts us to believe that the mathematics of emergence in genetic programming may be solvable, i.e. that one can characterize necessary or sufficient conditions, quantify relationships, classify behaviors, etc.

Emergence often appears mysterious and intractable, so it is helpful to revisit, as a model for transparent tractability in ‘bottom up’ emergence, one very familiar example: matrix theory. Beginning with a bag of  $n^2$  numbers, we make them into a ‘whole’ by ordering them with two indices,  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, n$ , to yield a matrix  $\mathbf{A} := [A_{ij}]_{i,j=1,\dots,n}$ . Interaction between these parts is manifest in the matrix multiplication operator, a simple sum of products:

$$[\mathbf{Ax}]_i = \sum_{j=1}^n A_{ij}x_j. \tag{1}$$

From this simple structure and interactions come all the panoply of emergent phenom-

ena of linear algebra: eigenvalues, eigenvectors, spectral radius, numerical range, rank, diagonalizability, irreducibility, invertibility, Jordan canonical form, Frobenius normal form, etc. The relationship between the parts  $A_{ij}$  and these emergent properties occupies the pages of many mathematics papers.

The ‘bottom up’ emergence in genetic programming — how the parts in a program map to the objective function — is like (1), except that the simplicity and repetition of the interactions is cast off, and we are exposed to an infinite space of parse trees. But this unruly space is no less formal and no less observable than the simple matrix, and their mathematical properties are simply awaiting our discovery. One example I would point to is the identification of Lagrangian distributions as the emergent distribution of tree shapes under subtree exchange by Poli et al. (2007). If, as a result of this commentary, the next analytic theorem on emergent structures in genetic programming comes one day earlier, it shall have been of value.

## References

- Altenberg, L. 1994a. Emergent phenomena in genetic programming. In Sebald, A. and Fogel, L., editors, *Evolutionary Programming — Proceedings of the Third Annual Conference*, pages 233–241. World Scientific Publishing. ISBN 9810218109.
- Altenberg, L. 1994b. The evolution of evolvability in genetic programming. In Kinnear, K. E., editor, *Advances in Genetic Programming*, chapter 3, pages 47–74. MIT Press, Cambridge, MA.
- Banzhaf, W. 2013. Genetic programming and emergence. *Genetic Programming and Evolvable Machines*. <http://dx.doi.org/10.1007/s10710-013-9196-7>.
- Boswell, J. 1792. *The Life of Samuel Johnson, LL. D.*, volume 1. Sir Isaac Pitman & Sons, LTD, Edinburgh, Roger Ingpen (1907) edition. [ark:/13960/t2n58rf0x](http://dx.doi.org/10.1007/s10710-013-9196-7).
- Poli, R., Langdon, W. B., and Dignum, S. 2007. On the limiting distribution of program sizes in tree-based genetic programming. In *Genetic Programming: 10th European Conference, EuroGP 2007 Valencia, Spain, April 11-13, 2007 Proceedings*, pages 193–204. Springer.