Summer Lecture Series 2003

Experimental Design & Methodology

Basic lessons in empiricism

Overview and References

Rafal Kicinger   R. Paul Wiegand
rkicinge@gmu.edu   paul@tesseract.org

Overview

The study of evolutionary algorithms (EAs) has had, like most scientific endeavors, to make use of basic empirical methods for the purposes of drawing useful conclusions during research. Unfortunately, while having a sound and functioning methodology for conducting empirical research is not merely important, but fundamentally necessary, such methods are easily discarded, such rigor easily lost, in the face of overwhelming research objectives and enthusiastic practitioners. Ironically, it is precisely at this point when one should be most adherent to basic research methodology.

In few places is the surprising lack of rigor more apparent than in the field of evolutionary computation (EC). Though inherently suited for experimental research of many forms and colors, the field seems to be surprisingly littered with papers that convey little to no convincing argumentation because their method was unsound, experiments poorly designed, or results presented badly. While this assessment may be overly critical, it is nevertheless true that a much more effective, directed, and justifiable research process is possible if one takes a step back and reexamines experimental design and methodology. Application of such practices is part and parcel to the conduct of science, especially in the case of empirical research.

This lecture attempts to provide two high-level services for its audience. First, it serves the role of providing background material for the remainder of the Evolutionary Computation Laboratory's 2003 Summer Lecture Series ("Understand Empirical Methods"). Second, it specifically undertakes exactly the task mentioned in the last paragraph: to take a step back and reexamine experimental design and methodology from the perspective of an EC researcher. It does this in two ways. First, by establishing a very basic, high-level discussion about methodology, experimental design, experiment execution, and presentation of experimental results, a basic framework for conducting empirical research is provided. Second, the lecture offers very specific examples from EC literature illustrating both good and bad methodological choices, and the ramifications of these choices.

General overview of the organization of the lecture is shown below:
Part I:	Methodology
Part II:	Designing Experiments
Part III:	Conducting Experiments
Part IV:	Presenting Experiments

References

Booth, W. C., Williams, J. M., & Colomb, G. G. (2003). The craft of research (Chicago guides to writing, editing, and publishing). Chicago, IL: University of Chicago Press.

Davidor, Y. (1991). Epistasis variance: a viewpoint on GA-hardness. In G. J. E. Rawlins (Ed.), Foundations of Genetic Algorithms I (pp. 23-35). San Mateo, CA: Morgan Kaufmann.

Leedy, P. D., & Ormrod, J. E. (2000). Practical research: planning and design: Prentice Hall.

Reeves, C. R., & Wright, C. C. (1995). An experimental design perspective on genetic algorithms. In L. D. Whitley & M. D. Vose (Eds.), Foundations of Genetic Algorithms 3. San Mateo, CA: Morgan Kaufmann.

Reeves, C. R., & Wright, C. C. (1995). Epistasis in genetic algorithms: an experimental design perspective. Proceedings of the 6th International Conference on Genetic Algorithms (ICGA-95), Pittsburgh, PA, USA.

Reeves, C. R., & Wright, C. C. (1995). Genetic algorithms and statistical methods: a comparison. Proceedings of the 1st IEE/IEEE International Conference on Genetic Algorithms for Engineering Systems: Innovations and Applications, Sheffield, UK.

Reeves, C. R., & Wright, C. C. (1997). Genetic algorithms versus experimental methods: a case study. Proceedings of the 7th International Conference on Genetic Algorithms, East Lansing, MI, USA.

Reeves, C. R., & Wright, C. C. (1999). Genetic algorithms and the design of experiments. In D. D. Lawrence & M. D. Vose & K. A. De Jong & L. D. Whitley (Eds.), Evolutionary Algorithms: Springer Verlag.