Statistical Science
1993, Vol. 8, No. 1, 76-77

Missing Pieces, Derandomization and Concluding Remarks

J. Michael Steele

Abstract. The most salient omissions from the survey provided by the preceding articles are reviewed and ameliorated by reference to recent publications.

Key words and phrases: Chen-Stein method, derandomization, expander graphs, percolation theory, pseudorandom graphs, random graphs.

No survey is ever complete, and completeness is especially elusive for a survey of a rapidly evolving subject like the interface of probability and the theory of algorithms. In just the last few months there has been stunning progress on transparent proof techniques, which, in a nutshell, are methods that (in some versions) allow one to test the validity of alleged proofs by applying tests that will fail with probability 1/2 unless the proof is valid. Some of the foundations that underlie this development are touched on by Feigenbaum and Lagarias and by Feigenbaum in this issue. Even so, an honest sketch of the new ideas in transparent proof techniques would require a substantial excursion into the most modern bits of computational complexity theory. Rather than take on that excursion, we have to be content with referring the reader to the journalistic accounts of Kolata (1992) and Cipra (1992) and the more technical discussion of Johnson (1992). The latter contains pointers to the appropriate scientific literature.

If one studies the engineering of these recent advances in transparent proof techniques, one finds many structures that were first supported by probabilistic thinking. As it happens, one also finds that considerable effort was subsequently invested to replace most of the original probabilistic scaffolding by constructions that could be called purely deterministic. To probabilists, this passion for excising randomized constructions seems curious, but many computer scientists and combinatorists feel themselves to be on philosophically shaky ground with randomized constructions. Thus, in many contexts, such constructions are accorded only the status of a pure existence proof, and, almost always, they are seen as lacking at least some of the pristine virtue of deterministic constructions.

The classic way to atone for the easy virtue of probabilistic constructions has been to supply deterministic replacements, and a famous illustration of this tradition arises in the discussion of expander graphs in the article by Maggs in this issue. Though it is shockingly easy to show the existence of all sorts of expander graphs via probability, many researchers seemed to breathe a sigh of relief when deterministic constructions for good expanders were developed, even when the new constructions called on methods as sophisticated as those used by Lubotzky, Phillips and Sarnak (1988).

Although most of the discoveries provoked by the urge to replace a randomized construction with a deterministic one have turned out to be rather specialized, there are recent developments that change this situation and offer enticing suggestions of an important new theory of *derandomization*. Instead of pursuing clever ad hoc constructions, the new idea is to look for algorithmic procedures that replace the very steps employed in the randomized construction. This shift in perspective turns out to be very fruitful, and a good sense of its power can be found in Chapter 15 of the important new book *The Probabilistic Method* by Alon, Spencer and Erdös (1992). Further, because of the remarkable utility of the Lovász Local Lemma in the more subtle probabilistic constructions, the recent algorithmization of the Local Lemma by Beck (1992) and Alon (1992) offers a compelling validation of the derandomization concept.

In addition to its useful discussion of derandomization, the volume of Alon, Spencer and Erdös (1992) also provides charming introductory treatments of at least two other topics that may seem underrepresented in this survey, *graph algorithms* and *random graphs*. The latter topic is also well covered in the treatise *Random Graphs* by Bollobás (1985), which is a bible for any serious student of random graphs. From the

J. Michael Steele is Professor, Department of Statistics, The Wharton School of the University of Pennsylvania, 3000 Steinberg Hall-Dietrich Hall, Philadelphia, Pennsylvania 19104-6302.

probabilists' point of view, two of the most important recent links to these areas are to percolation theory and to the Chen–Stein method of Poisson approximation. Much of the recent progress in percolation theory is beautifully introduced by Grimmett (1989), and the recent treatise on Poisson approximation of Barbour, Holst and Janson (1992) surely belongs on the bookshelf of anyone interested in probability, graphs or discrete algorithms.

Another important sphere of probability in the service of algorithms that some may see as underrepresented here is the *analytical* probabilistic analysis of algorithms. This area can be characterized by the use of explicit combinatorial calculations and generating functions to calculate the means and variances of algorithm running times and other features of interest. Historically, the critical parts of such calculations tend to be more closely connected with real and complex analysis than with probability theory, but the language of probability always drives the problem formation and increasingly contributes to the analysis. Much of this tradition springs from seminal work of Donald Knuth, with many illustrations of the central themes found in his now classic books *The Art of Computer Programming*, vols. 1–3 (Knuth, 1973). A more recent and introductory treatment that is sympathetic in approach is the text of Bender and Williamson (1991), which also can be commended for the insights it offers into asymptotic analyses assisted by generating functions. Another recent volume that anyone involved with the analytical tradition should read is Wilf (1990), which christens "generatingfunctionology" as a field in itself and also offers up many of the field's secrets in a way in which they can be enjoyably mastered.

A final volume that deserves mention here is the recent collection *Probabilistic Combinatorics and Its Applications*, edited by Bollobás (1991). The seven essays in this collection are all of great interest to the field, and each points toward many lively research topics. In particular, the essay by F. R. K. Chung (1991) provides quite another perspective on derandomization theory and illustrates many of the subtleties that perplex investigators who examine randomness in a quest to find acceptable surrogates.

ACKNOWLEDGMENTS

Research supported in part by NSF Grants DMS-88-12868, DMS-92-11634, ARO Grants DAAL03-89-G-0092, DAAL03-91-G-0110 and NSA Grant MDA-904-H-2034.

REFERENCES

ALON, N. (1992). A parallel algorithmic version of the Local Lemma. *Random Structures Algorithms* 2 367–377.

ALON, N., SPENCER, J. and ERDÖS, P. (1992). *The Probabilistic Method*. Wiley, New York.

BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation*. Clarendon Press, Oxford.

BECK, J. (1992). An algorithmic approach to the Lovász Local Lemma I. *Random Structures Algorithms* 2 343–366.

BENDER, E. A. and WILLIAMSON, S. G. (1991). *Foundations of Applied Combinatorics*. Addison-Wesley, Reading, MA.

BOLLOBÁS, B. (1985). *Random Graphs*. Academic, New York.

BOLLOBÁS, B., ed. (1991). *Probabilistic Combinatorics and Its Applications*. Amer. Math. Soc., Providence, RI.

CHUNG, F. R. K. (1991). Constructing random-like graphs. In *Probabilistic Combinatorics and Its Applications* (B. Bollobás, ed.) 21–56. Amer. Math. Soc., Providence, RI.

CIPRA, B. A. (1992). Theoretical computer scientists develop transparent proof technique. *SIAM News* 25 1.

FEIGENBAUM, J. (1993). Probabilistic algorithms for defeating adversaries. *Statist. Sci.* 8 26–30.

FEIGENBAUM, J. and LAGARIAS, J. C. (1993). Probabilistic algorithms for speedup. *Statist. Sci.* 8 20–25.

GRIMMETT, G. (1989). *Percolation*. Springer, New York.

JOHNSON, D. S. (1992). The NP-completeness column: An ongoing guide, 23rd ed.: The tale of the second prover. *J. Algorithms* 13 502–524.

KNUTH, D. E. (1973). *The Art of Computer Programming* **1–3**. Addison-Wesley, Reading, MA.

KOLATA, G. (1992). New short cut found for long proofs. *New York Times*, April 6, Section C.

LUBOTZKY, A., PHILLIPS, R. and SARNAK, P. (1988). Ramanujan graphs. *Combinatorica* 8 261–277.

MAGGS, B. M. (1993) Randomly wired multishape networks. *Statist. Sci.* 8 70–75.

WILF, H. S. (1990). *Generatingfunctionology*. Academic, New York.