Progress in the Analysis of Algorithms at Stanford
October, 1987 to October, 1988
Donald E. Knuth, Principal Investigator

I'm pleased to report that the past year has seen several noteworthy advances in our knowledge about algorithms, thanks to work done in this research project, which is being supported in part by the National Science Foundation and the Office of Naval Research. Besides myself, the project has supported the visits of Philippe Flajolet, Boris Pittel, and Herbert Wilf to Stanford, and the thesis research of five graduate students has also been supported.

1. The most significant event of the year was the discovery by Flajolet, Knuth, and Pittel of new techniques for the analysis of random graphs. These new methods use the theory of complex variables to solve several open problems that previous researchers were unable to resolve with the traditional techniques of mathematical statistics (martingales, laws of large numbers, etc.). We had proved in 1987 that the expected length of the first cycle in an evolving graph on n vertices is asymptotically $Kn^{1/6}$ for some constant K; this result was quite surprising because such an order of growth had not been seen in other problems. This year we clarified the method and extended it in several directions. We evaluated the constant $K = 2.0337$, and showed that the standard derivation of the first cycle length is proportional to $n^{1/4}$. We found that the expected length of the second cycle is asymptotically equal to $\frac{1}{3}Kn^{1/6}\ln n$; in general, the expected length of the kth cycle is asymptotically

$$\frac{2}{6^k}\,\frac{(\ln n)^{k-1}}{(k-1)!}\,Kn^{1/6}$$

for any fixed k. We discovered the expected size of the component that contains the first cycle, and we determined its limiting probability distribution. We showed that the first component with more than one cycle has expected size of order $n^{2/3}$; the total length of all cycles formed before such a component occurs is proportional to $n^{1/3}$. We showed that a random graph with m edges and n vertices is bipartite with probability

$$\left(\frac{1-\lambda}{1+\lambda}\right)^{1/4} e^{\lambda/2}, \qquad \text{if} \qquad \lim_{m,n\to\infty}\frac{2m}{n} = \lambda < 1.$$

And we showed that a random graph with $\frac{1}{2}n$ edges and n vertices has a "bicyclic" component with probability

$$1 - \sqrt{\tfrac{2}{3}} + O(n^{-1/2}).$$

All of these results, and several more of a similar nature, follow from our new method of analysis, which we believe will help us to resolve other difficult problems that have previously resisted solution. These results appear in Stanford report CS 1219, "The first cycles in an evolving graph," September, 1988.

2. Another significant achievement of the year was the completion of the 640-page textbook *Concrete Mathematics* by Ronald L. Graham, Donald E. Knuth, and Oren

Patashnik. This book contains more than 500 exercises, many of which are "mini research results" that have not been published before. Although the book is a textbook used in upper-level undergraduate courses, it contains new procedures for solving recurrences and evaluating sums in closed form, as well as new expositions of important techniques never before treated in textbooks. Therefore we believe the book will advance the art of algorithmic analysis in hundreds of universities.

3. Knuth and Pittel studied "A recurrence related to trees" [to appear in *Proceedings of the American Mathematical Society*, December, 1988], an extension of results previously obtained by Knuth and Schönhage. We showed, for example, that the solution to the recurrence

$$x_n = n^\beta + \frac{1}{n-1} \sum_{0<k<n} \binom{n}{k} \left(\frac{k}{n}\right)^{k-1} \left(\frac{n-k}{n}\right)^{n-k-1} x_k$$

has the asymptotic form

$$x_n = \frac{\Gamma(\beta - 1/2)}{\sqrt{2}\,\Gamma(\beta)}\, n^{\beta+1/2} + O(n^\beta) + O(n \log n)$$

when $\beta > \frac{1}{2}$. The properties of an interesting family of polynomials called "tree polynomials" were developed en route to this result.

4. Floyd and Knuth studied the properties of machines that can only add, subtract, and test the sign of numbers. We showed that such machines can actually multiply, divide, and calculate greatest common divisors quite efficiently. In particular, the operation $x^m \bmod n$ (which is fundamental to the RSA method of data encryption/decryption) can be performed by doing only $(\log m)(\log n)$ additions and subtractions on nonnegative integers less than $2n$, using fewer than 10 registers. Therefore an efficient and inexpensive device can be designed to do RSA codes. (A Stanford CS report entitled "Addition machines" will be published in November, 1988.)

5. Boris Pittel resolved a longstanding open problem about stable marriages by proving that the expected number of stable marriages (when men and women have independently random choices) is asymptotic to $e^{-1} n \ln n$. Previously, the only thing known about this quantity was that it is ≥ 1; nobody had ever been able to prove that the expected number was unbounded, or that it did not grow exponentially fast. A Stanford CS report about this result is now in preparation.

6. One of the main goals in the research proposal for this project was the creation of the Stanford Graphbase, a database of interesting test cases that can be used as benchmarks for combinatorial algorithms. This database is taking shape nicely, and it should be ready for distribution to other researchers in 1989. Knuth lectured about this work-in-progress at a large SIAM conference in June, and at several other universities; response from the audience was enthusiastic on all occasions, so there is good reason to believe that the database will find wide use.

7. Flajolet discovered a way to determine the asymptotic properties of the coefficients of the Hadamard product of two generating functions $f(z)$ and $g(z)$, when f or g have

algebraic singularities on their circle of convergence. This extends his important work of 1987 with Odlyzko, when they contributed a new fundamental technique for studying the coefficients of generating functions (improving on and simplifying the classical method of Darboux).

8. Knuth and Wilf discovered a simple rule for the divisibility properties of generalized binomial coefficients, extending a classical theorem of Kummer. ["The power of a prime that divides a generalized binomial coefficient," submitted for publication.]

9. Knuth, Papadimitriou, and Tsitsiklis studied the complexity of the algorithm for eliminating dominated strategies in certain games. ["A note on strategy elimination in bimatrix games." *Operations Research Letters* **7** (1988), 103–107.]

10. Knuth found relatively prime 15-digit numbers a and b such that the sequence defined by

$$A_0 = a. \quad A_1 = b. \quad A_{n+2} = A_{n+1} + A_n$$

contains no prime numbers. It had previously been believed that a and b must be much larger (30 digits or more) for this to be possible. ["A Fibonacci-like sequence of composite numbers," to appear in *Mathematics Magazine*.]

11. Knuth found a simple way to compute permutation groups with given generators, and showed that was reasonably efficient. ["Efficient representation of perm groups," to appear in *Combinatorica*.]

12. Knuth discovered a curious new associative operation $m \circ n$ on negative integers, with the property that $mon \sim \sqrt{5}\,mn$ when m and n are large. ["Fibonacci multiplication," *Applied Mathematics Letters* **1** (1988), 57–60.]

13. Knuth and Pittel recently discovered that a random graph with $\frac{1}{2}n$ edges and n vertices has positive (but very small) probability of having more than one bicyclic component. This disproved a conjecture by students of R. M. Karp, who had noticed no such graphs in their empirical tests. The behavior of large random graphs with $\le (\frac{1}{2} - \epsilon)n$ edges or $\ge (\frac{1}{2} + \epsilon)n$ edges has been well understood for a long time, but the properties when the number of edges is near $\frac{1}{2}n$ are just now beginning to submit to analysis. We are continuing to explore such critical graphs, because Karp has just discovered a new way to compute transitive closures of random directed graphs, with an efficiency that depends heavily on the behavior of random *undirected* graphs in the critical region.

14. Five graduate students (Chen, Feder, Gangoli, Haddad, Patashnik) are being supported by this project as they complete their thesis research. In each case the work is proceeding well, and we expect the theses will be completed next year. Of particular interest is Chen's discovery that a new heuristic invented to study the size of backtrack trees proves to be unexpectedly effective at also finding the longest paths in such trees. In one experiment, his heuristic improved the search time by more than five orders of magnitude.