JOACHIM VON ZUR GATHEN (1992). A polynomial factorization challenge. *ACM SIGSAM Bulletin* **26**, 22–24. URL http://dx.doi.org/10.1145/130933.130939.

A Polynomial Factorization Challenge

JOACHIM VON ZUR GATHEN

Department of Computer Science, University of Toronto
Toronto, Ontario M5S 1A4, Canada
gathen@theory.toronto.edu

In the early 1970s, a major paradigm shift took place in algorithms research, away from experimental results to asymptotic analysis. Knuth popularized the "Big O" notation, and Hopcroft says in his 1986 ACM Turing Award (with Robert Tarjan) address: "During the 1960s, research on algorithms had been very unsatisfying. A researcher would publish an algorithm in a journal along with execution times for a small set of sample problems, and then several years later, a second researcher would give an improved algorithm along with execution times for the same set of sample problems. The new algorithm would invariably be faster, since in the intervening years, both computer performance and programming languages had improved. The fact that the algorithms were run on different computers and programmed in different languages made me uncomfortable with the comparison. It was difficult to factor out both the effects of increased computer performance and the programming skills of the implementors—to discover the effects due to the new algorithm as opposed to its implementation. Furthermore, it was possible that the second researcher had inadvertently tuned his or her algorithm to the sample problems ... I set out to demonstrate that a theory of algorithm design based on worst-case asymptotic performance could be a valuable aid to the practitioner."

The paradigm of worst case asymptotic analysis defines our science. I want to propose a complementary experimental paradigm, in the subarea of factoring polynomials. This new challenge must take Hopcroft's criticism into account, and it should not detract from my fundamental belief that asymptotic performance improvements are the ultimate goal of the algorithm designer.

My proposal is modelled on the situation in computational number theory, where asymptotic analysis is the usual paradigm. However, there is a vital experimental challenge in this theory: the Cunningham project, started in 1925. This contains a list of composite integers and the factorizations of most of them. The integers have the form $b^n \pm 1$ with $b \leq 12$. Many factorizations are given in Brillhart *et al.* (1988), and Wagstaff maintains this list, including ten integers with the "most wanted factorizations". The most recent success is the factorization of the ninth Fermat number $2^{512}+1$ by the new number field sieve of Lenstra *et al.* (1990). This question has provided title-page copy to the New York Times.

For factoring polynomials, there are three basic subproblems: univariate polynomials over finite fields and over $\mathbf{Q}$, and multivariate polynomials. A detailed account of progress is given in the surveys by Kaltofen (1986, 1992), who also made major

contributions to the third subproblem. For univariate polynomials over finite fields, asymptotic progress has occurred roughly in intervals of about a decade:

1967/70 Berlekamp, McEliece, Zassenhaus,

1980/81 Ben-Or, Cantor & Zassenhaus, Rabin.

Recent progress is reported in von zur Gathen & Shoup (1992). Over $\mathbf{Q}$, a similar pattern prevails:

1969/70 Berlekamp, Zassenhaus,

1982 Lenstra, Lenstra & Lovász.

To provide entertainment in the long waiting periods, I propose to have an experimental *Polynomial Factorization Challenge*, in the form of a list of polynomials with "most wanted factorizations". As a serious reason, this would tell us where the boundaries of "routine" and "special effort" factorization are, and which methods achieve the extremes. In particular, it will be interesting to see what the relative importance of algorithm design and implementation is. Due to the polynomial time nature of our algorithms, we can expect much more incremental progress than in the integer factorization industry.

THE CHALLENGE. *It consists of a finite list of integers $n_1 < n_2 < \cdots$. Denote by p_n the smallest prime number larger than $2^n \cdot \pi$. For an n in the list, the* CHALLENGE *is to factor $f_n = x^n + x + 1$ modulo p_n. A solution consists of the monic irreducible factors and their multiplicities.*

Although it takes only about $\log_2 n$ bits to communicate a particular challenge, we must consider the decimal input size to be $n^2 \log_{10} 2$. It is assumed that compositeness certificates for the integers between $2^n \cdot \pi$ and p_n and a primality certificate for p_n are easy to obtain. The list and recent progress will be regularly published in the SIGSAM Bulletin.

Using a Maple library routine, I have factored f_n modulo p_n for $n \leq 126$. The corresponding input size of over 4000 digits is already well beyond the reach of current integer factorization or primality testing methods. With more powerful machines (than the single Apple Macintosh IIfx I used) and more time (the longest factorization took 32 hours), this "brute-force" approach will carry somewhat further, and probably solve the first challenge:

CHALLENGE 1 (March 23, 1992): *All integers n between 127 and 200.*

Solutions to CHALLENGE 1 should be sent to the author, preferably by email in a format readable by a computer algebra system, such as Maple. You may want to include information about the hardware and software used, and the time it took.

In order to explain how I arrived at the formulation of the CHALLENGE, here are some questions I asked myself, and my answers.

Do we need such a challenge? No, we can continue to do business without it, but—yes!—it will provide a focus of practical efforts. It may also shed light on the practical implications of "polynomial time": how large can the input size be? How does that

compare to the input sizes for integer factorization and primality tests?

Over what domain should the polynomials be, and how many variables should they have? To focus on the most basic problems, the polynomials should be univariate. This is not meant to detract from the importance of multivariate factorization. As ground domains, the two possible choices are $\mathbf{Q}$ or finite prime fields. I favor the latter, partly because it is a subproblem for almost all factoring methods, and partly because I do not know interesting candidate polynomials over $\mathbf{Q}$.

Which polynomials should we choose? Some possibilities are $x^n \pm 1$, $x^n \pm 2$, $x^n \pm x \pm 1$, or cyclotomic polynomials. Some of these polynomials have, however, well-studied special multiplicative properties which permit methods that do not apply to general polynomials. I propose—somewhat arbitrarily—to use $x^n + x + 1$. Random polynomials are not in the spirit of the Cunningham project, and have the drawback of being more complicated to communicate. A further possibility would be to take combinatorial polynomials, such as the nth partition polynomial $Q_n = \sum_i Q_{ni} x^i$, where $Q_{ni} \in \mathbf{N}$ is the number of partitions of n into i positive integer summands. The additive properties of these polynomials are well-studied.

Modulo what primes do we want to factor? There does not seem to be a nice and sufficiently dense class of primes. Random primes are, again, not suitable. The first prime p larger than a^n for some small a, say $a = 2$, has the drawback that repeated squaring modulo p is somewhat special. Arnold Schönhage suggested to use $2^n \cdot \pi$.

ACKNOWLEDGEMENT. Many thanks go to Rafaela von zur Gathen for help with the computations.

References

J. BRILLHART, D. H. LEHMER, J. L. SELFRIDGE, B. TUCKERMAN, AND S. S. WAGSTAFF, JR., *Factorizations of $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ Up to High Powers*, vol. 22 of *Contemporary Mathematics.*. American Mathematical Society, 1988, 2nd edition.

J. VON ZUR GATHEN AND V. SHOUP, Computing Frobenius maps and factoring polynomials. In *24th Ann. ACM Symp. Theory of Computing*, Victoria BC, 1992, to appear.

E. KALTOFEN, Polynomial factorization 1982–1986. In *Computers in Mathematics*, ed. D. V. CHUDNOVSKY AND R. D. JENKS, Marcel Dekker, New York, 1990, 285–309.

E. KALTOFEN, Polynomial factorization 1987–1991. In *Proc. Latin'92*, São Paulo, Brazil, 1992, to appear.

A. K. LENSTRA, H. W. LENSTRA, JR., M. S. MANASSE, AND J. M. POLLARD, The number field sieve. In *Proc. 22nd ACM Symp. Theory Comput.*, 1990, 564–572.