The On-Line Encyclopedia of Integer Sequences

N. J. A. Sloane*

1. Introduction

This article gives a brief introduction to the On-Line Encyclopedia of Integer Sequences (or OEIS). The OEIS is a database of nearly $90,000$ sequences of integers, arranged lexicographically. The entry for a sequence lists the initial terms (50 to 100, if available), a description, formulae, programs to generate the sequence, references, links to relevant web pages, and other information.

2. To consult the database

Since 1996, an electronic version [20] has been accessible via the Internet, at the URL

http://www.research.att.com/~njas/sequences/.

If a list of numbers is entered there, the reply will display the entries for all matching sequences.

For example, suppose you were trying to count the ways to insert parentheses into a string of n letters so that the parentheses are balanced and there are at least two letters inside each pair of parentheses. The outer pair of parentheses is to be ignored. For $n = 3$ and 4 there are respectively 3 and 11 possibilities:

$$
\begin{aligned}
n = 3 : \quad & abc, (ab)c, a(bc); \\
n = 4 : \quad & abcd, (ab)cd, a(bc)d, ab(cd), (ab)(cd), (abc)d, \\
& a(bcd), ((ab)c)d, (a(bc))d, a((bc)d), a(b(cd)).
\end{aligned}
$$

Further work shows that for $n = 1, \ldots, 5$ the numbers are 1, 1, 3, 11, 45. Entering these into the web page produces nine matching sequences, but they are sorted, with the most probable match appearing first. Indeed, this entry tells you that these are the numbers (sequence A1003) arising from "Schröder's second problem", and are also known as "super-Catalan numbers".

The reply (an abridged version is shown in Figure 1) gives 21 references, ranging from Schröder (1870) [17] to articles published electronically in the last few years. There is an

---
*Neil J. A. Sloane is with AT&T Shannon Labs, Florham Park, NJ. His email address is njas@research.att.com.

explicit formula:

$$a(n) = \frac{1}{n} \sum_{k=0}^{n-2} \left(\begin{array}{c} 2n-k-2 \\ n-1 \end{array} \right) \left(\begin{array}{c} n-2 \\ k \end{array} \right) , \quad n > 1 ,$$

a recurrence:

$$a(1) = a(2) = 1 ,$$

$$(n+1)a(n+1) = 3(2n-1)a(n) - (n-2)a(n-1) , \quad n > 1 ,$$

programs to produce the sequence in Maple and Mathematica, and much more.

There's no other reference work that will carry out this kind of search.

The Encyclopedia can also be consulted via email. There are two addresses. Sending email to sequences@research.att.com with a line in the body of the message saying

$$\text{lookup} \quad 1 \quad 1 \quad 3 \quad 11 \quad 45$$

will trigger the same search that the web page performs, only now the results are sent, almost immediately, via email. Superseeker (superseeker@research.att.com) carries out a more sophisticated analysis and tries hard to find an explanation for the sequence, even if it is not in the database. If the simple lookup fails, Superseeker carries out many other tests, including:

- applying over 130 transformations to the sequence, including the binomial, Euler, Möbius, etc. transforms [1], and checking to see if the result is in the database

- applying Padé approximation methods to try, for example, to express the nth term as a rational function of n (using the "gfun" package of Salvy and Zimmermann [16], the "guesss" program of Derksen [5] and the "RATE" program of Kratthentaler [8])

- checking to see if changing one or two characters produces a sequence in the database.

Since Superseeker carries out a nontrivial amount of calculation, users are asked to submit only one sequence per hour.

The electronic version of the Encyclopedia had its origins in the books [18] (1973) and [21] (1995). Disk space is cheap, and the present incarnation (excluding illustrations) contains about 72 times as much data as the 1995 book. The history of the Encyclopedia is described in more detail in [19].

3. Applications

Most people use the Encyclopedia to identify a sequence, as illustrated above. It has been around long enough so that there is a good chance that your sequence will be there. If not, you will see a message encouraging you to submit it.

Most of these applications are unspectacular, akin to looking up a word in a dictionary (cf. [2]). One encounters a sequence in the middle of a calculation, perhaps

$$1 \quad 2 \quad 4 \quad 6 \quad 10 \quad 12 \quad 16 \quad 18 \quad 22 \quad 28 \quad 30 \quad \dots ,$$

and one wants to know quickly what it is — preferably a formula for the n-th term (in this case it is probably $\text{prime}(n) - 1$, A6093) or a recurrence. Successful applications of this type usually go unremarked. Some are more dramatic: there is a web page[†] that lists several hundred articles that acknowledge help from the OEIS. One quotation will serve to illustrate this. Emeric Deutsch of Poytechnic University, Brooklyn, said in a recent email message: "... your database is invaluable. For example, for a certain sequence a_n, using Maple I found the first 100 or so indices i for which a_i is odd. Only the OEIS could tell me that the sequence of these i's is a known sequence related to the Thue-Morse sequence. Of course, this had to be followed by further reading and proof."

The other main application is to find out the current status of work on a problem — for example, the search for Mersenne primes (see A43), the enumeration of Hadamard matrices (A7299), Latin squares (A315) or meanders (A5316) of order n, the latest information about the decimal expansion of π (A796) or, better, its continued fraction expansion (A1203).

Of course people trying to solve puzzles or IQ tests find the database useful. A5228 is a classic:

$$1 \quad 3 \quad 7 \quad 12 \quad 18 \quad 26 \quad 35 \quad 45 \quad 56 \quad 69 \quad 83 \quad \dots$$

There are also some less obvious applications. One is in simplifying complicated integer-valued expressions. You might, for example, have encountered the sum

$$\sum_{k=0}^{n} \binom{4n+1}{2n-2k} \binom{n+k}{k} .$$

There are powerful methods for evaluating such sums [12], [13], but it doesn't take long to work out the first few terms: 1, 12, 240, 5376, and to look them up in the database. In this

---

[†]http://www.research.att.com/∼njas/sequences/cite.html

case you would have been lucky. The reply suggests that this is sequence A6588, $4^n \begin{pmatrix} 3n \\ n \end{pmatrix}$, and supplies, with references, the binomial coefficient identity you were hoping for.

Another application is in proving inequalities. You might suspect that $\sigma(n) < n\sqrt{n}$ for $n > 2$, where σ is the sum-of-divisors function (A203). If the initial terms of $[n\sqrt{n}] - \sigma(n)$ (where $[\,]$ denotes the "floor" function) are submitted to the database, the reply suggests that this is A55682, and points you to a reference that gives a proof of your inequality.

I cannot resist mentioning sequence A57641, which gives the values of

$$[H_n + \exp(H_n)\log(H_n)] - \sigma(n)$$

for $n \geq 1$, where H_n is the harmonic number $\sum_{i=1}^{n} 1/i$. Lagarias [9], extending earlier work of Robin [15], has shown that this sequence is nonnegative if and only if the Riemann hypothesis holds!

Although the database contains a number of sequences of both of the above types, I have not made a systematic search through reference works such as [7], [11] and it would be nice to get many more examples.

The database can also be used to save space when referring to particular sequences. When introducing the Motzkin numbers, for example, instead of giving the definition, references and the first few terms it is simpler just to say "... the Motzkin numbers M_n (sequence A1006 of [20])".

One can also search the database for sequences that mention a particular name (Riemann, say), and there is a separate alphabetical index, useful for keeping track of all sequences on a certain topic — e.g. the entry for *groups* lists abelian (A688), primitive permutation (A19), transitive permutation (A2106), simple (A5180), total number (A1), ...

In the past year the main look-up page has been translated into 28 languages, with the goal of making it easier to use throughout the world. The entries from the database still appear in English, but the headings in the replies and the error messages have also been translated.

4. The database

To be included in the database a sequence should be integer-valued, well-defined and interesting.

The main sources are combinatorics, number theory and recreational mathematics, but most branches of mathematics are represented (e.g. A27623, the number of rings with n

elements), and there are hundreds of entries from chemistry and physics (e.g. A8253, the coordination sequence for diamond: the number of carbon atoms that are n bonds away from a particular carbon atom).

Sequences of rational numbers are entered as a linked pair giving numerators and denominators separately. The Bernoulli numbers B_n are A27641/A27642.

Triangular arrays of numbers are read row-by-row, so that Pascal's triangle gives A7318:

$$1, \ 1, 1, \ 1, 2, 1, \ 1, 3, 3, 1, \ 1, 4, 6, 4, 1, \ \ldots .$$

Square arrays are read by anti-diagonals, so that the Nim-addition table

$$
\begin{array}{cccccc}
0 & 1 & 2 & 3 & 4 & 5 & \ldots \\
1 & 0 & 3 & 2 & 5 & 4 & \ldots \\
2 & 3 & 0 & 1 & 6 & 7 & \ldots \\
3 & 2 & 1 & 0 & 7 & 6 & \ldots \\
& \ldots & & \ldots & & \ldots & \ldots
\end{array}
$$

produces A3987:

$$0, \ 1, 1, \ 2, 0, 2, \ 3, 3, 3, 3, \ 4, 2, 0, 2, 4, \ \ldots .$$

Most well-defined submissions get accepted, since an open-door policy seems the best. The amazing coincidences of the Monstrous Moonshine investigations [4], for example, make it difficult to say that a particular sequence, no matter how obscure, will never be of interest.

Sequences that are discouraged are those that depend on an arbitrary and large parameter: primes of the form $x^2 + y^2 + 2003$, say, whereas primes of the form $x^2 + y^2 + 1$ form a perfectly acceptable sequence (A79545).

The Encyclopedia currently receives between $10,000$ and $12,000$ downloads per day. The rate of arrival of *new* sequences has remained constant at about $10,000$ per year for the past seven years, with roughly the same number of comments and updates. To keep this flood of information from getting out of control, people are asked to use a web form[‡] when submitting new sequences or comments.

For most of its life the Encyclopedia has been maintained by the author, but in the past year a board of associate editors has been formed to help with the work. There is also a group of regular users who constantly send corrections and extensions, and help maintain the accuracy of the entries.

[‡]http://www.research.att.com/∼njas/sequences/submit.html

Even so, much remains to be done. There are more journals and e–print servers now than ever, and the trained eye sees integer sequences everywhere. I still discover articles in the library or on the web where authors have published sequences without sending them to the Encyclopedia. If you come across an integer sequence in your own work or elsewhere, please submit it to the Encyclopedia!

Of course, accuracy is a major concern in maintaining the database. The entries in [18] and [21] were checked very thoroughly, and almost all the errors that have been discovered in those books were already present in the sources from which the sequences were taken. As the number of sequences has increased in recent years, it has become more difficult to check them all. However, the number of users has also increased, and a large number of the entries carry a comment that the sequence has been extended (or sometimes, corrected and extended) by someone. Contributors see a reminder that the standards are those of a mathematics reference work, and all submissions should be carefully checked. So, on the whole, users can be confident that the sequences are correct. The keywords "uned" and "obsc" indicate sequences that have not yet been edited or for which the definition is unclear. These serve both as warnings to users and to indicate places where volunteers can help.

One of the pleasures of maintaining the database is seeing the endless flow of new sequences. I will end by mentioning a few recent examples:

Home primes (A37274), [6]: $a(n)$ is the prime reached when you start with n, concatenate its prime factors, and repeat until a prime is reached ($a(n)$ is defined to be -1 if no prime is ever reached, although it is conjectured that this never happens). E.g. $8 = 2 \times 2 \times 2 \rightarrow 222 = 2 \times 3 \times 37 \rightarrow 2337 = 3 \times 19 \times 41 \rightarrow 31941 \rightarrow \ldots \rightarrow$ (after 13 steps) 3331113965338635107, a prime; so $a(8) = 3331113965338635107$:

$$1 \quad 2 \quad 3 \quad 211 \quad 5 \quad 23 \quad 7 \quad 3331113965338635107 \quad 311 \quad 773 \quad 11 \quad 223 \quad 13 \quad \ldots.$$

The EKG sequence (A64413), [10]: $a(1) = 1$, $a(2) = 2$ and, for $n \geq 3$, $a(n)$ is the smallest natural number not in $\{a(k) : 1 \leq k \leq n-1\}$ with the property that $\gcd\{a(n-1), a(n)\} \geq 2$:

$$1 \quad 2 \quad 4 \quad 6 \quad 3 \quad 9 \quad 12 \quad 8 \quad 10 \quad 5 \quad 15 \quad 18 \quad 14 \quad 7 \quad 21 \quad 24 \quad 16 \quad 20 \quad 22 \quad 11 \quad \ldots.$$

Lacing a shoe (A78601), [14]: Number of ways to lace a shoe that has n pairs of eyelets. The lace must follow a Hamiltonian path through the 2n eyelets, and at least one of the neighbors of every eyelet must be on the other side of the shoe.

$$1 \quad 3 \quad 42 \quad 1080 \quad 51840 \quad 3758400 \quad 382838400 \quad 52733721600 \quad \ldots.$$

A "bootstrap" sequence (A79000), [3]: $a(n)$ is taken to be the smallest positive integer greater than $a(n-1)$ which is consistent with the assertion "n is a member of the sequence if and only if $a(n)$ is odd."

$$1 \quad 4 \quad 6 \quad 7 \quad 8 \quad 9 \quad 11 \quad 13 \quad 15 \quad 16 \quad 17 \quad 18 \quad 19 \quad 20 \quad 21 \quad 23 \quad 25 \quad 27 \quad 29 \quad 31 \quad \ldots.$$

References

[1] M. Bernstein and N. J. A. Sloane, Some canonical sequences of integers, *Linear Alg. Appl.*, **226**–**228** (1995), 57–72; Erratum **320** (2000), 210. [arXiv: math.CO/0205301]

[2] B. Cipra, Mathematicians get an on-line fingerprint file, *Science*, **265** (22 July, 1994), 473.

[3] B. Cloitre, N. J. A. Sloane and M. J. Vandermast, Numerical analogues of Aronson's sequence, *J. Integer Sequences*, **6** (2003), #03.2.2. [arXiv: math.NT/0305308]

[4] J. H. Conway and S. P. Norton, Monstrous moonshine, *Bull. London Math. Soc.*, **11** (1979), 308–339.

[5] H. Derken, *An algorithm to compute generalized Padé-Hermite forms*, Report **9403**, Department of Mathematics, Catholic University Nijmegan, 1994.

[6] P. De Geest, *Home primes < 100 and beyond*, published electronically at www.worldofnumbers.com/topic1.htm, 2003.

[7] H. W. Gould, *Combinatorial Identities*, Morgantown, WV, 1972.

[8] C. Krattenthaler, *RATE-A Mathematica guessing machine*, available electronically from http://euler.univ-lyon1.fr/home/kratt/rate/rate.html.

[9] J. C. Lagarias, An elementary problem equivalent to the Riemann hypothesis, *Amer. Math. Monthly*, **109** (2002), 534–543.

[10] J. C. Lagarias, E. M. Rains and N. J. A. Sloane, The EKG sequence, *Experimental Math.*, **11** (2002), 437–446. [arXiv: math.NT/0204011]

[11] D. S. Mitrinović, J. Sándor and B. Crstici, *Handbook of Number Theory*, Kluwer, Dordrecht, 1996.

[12] I. Nemes, M. Petkovšek, H. S. Wilf and D. Zeilberger, How to do Monthly problems with your computer, *Amer. Math. Monthly*, **104** (1997), 505–519.

[13] M. Petkovšek, H. S. Wilf and D. Zeilberger, $A = B$, Peters, Wellesley, MA, 1996.

[14] B. Polster, What is the best way to lace your shoes?, *Nature*, **420** (Dec. 5, 2002), 476.

[15] G. Robin, Grandes valeurs de la fonction somme des diviseurs et hypothèse de Riemann [Large values of the sum-of-divisors function and the Riemann hypothesis], *J. Math. Pures Appl.*, **63** (1984), 187–213.

[16] B. Salvy and P. Zimmermann, Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable, *ACM Transactions on Mathematical Software*, **20** (1994), 163–177.

[17] E. Schröder, Vier combinatorische Probleme [Four combinatorial problems], *Zeit. f. Math. Phys.*, **15** (1870), 361–376.

[18] N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, NY, 1973.

[19] N. J. A. Sloane, My favorite integer sequences, in *Sequences and their Applications (Proceedings of SETA '98)*, edited by C. Ding, T. Helleseth and H. Niederreiter, Springer–Verlag, London, 1999, pp. 103–130. [arXiv: math.CO/0207175]

[20] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at http://www.research.att.com/∼njas/sequences/, 2003.

[21] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, 1995.

Figure 1: Part of the reply when the sequence 1, 1, 3, 11, 45 is submitted to the On-Line Encyclopedia. Many references, links and comments have been omitted to save space.

Greetingsfromthe [On-LineEncyclopediaofIntegerSequences](#) !

Matches(uptoalimitof30)foundfor1131145:

```
ID Number: A001003
Sequence:  1,1,3,11,45,197,903,4279,20793,103049,518859,2646723,
           13648869,71039373,372693519,1968801519,10463578353,
           55909013009,300159426963,1618362158587,8759309660445,
           47574827600981,259215937709463,1416461675464871
Name:      Schroeder's second problem (generalized parentheses); also called
              super-Catalan numbers or little Schroeder numbers.
Comments:  a(n) = number of ways to insert parentheses in a string of n symbols. The
              parentheses must be balanced but there is no restriction on the number
              of pairs of parentheses. The number of letters inside a pair of
              parentheses must be at least 2. Parentheses enclosing the whole string
              are ignored.
           Stanley gives several other interpretations for these numbers.
References P. J. Cameron, Some sequences of integers, Discrete Math., 75 (1989),
              89-102.
           I. M. H. Etherington, Some problems of non-associative combinations,
              Edinburgh Math. Notes, 32 (1940), 1-6.
           T. S. Motzkin, Relations between hypersurface cross ratios, and a
              combinatorial formula for par titions of a polygon, for permanent
              preponderance, and for non-associative products, Bull. Amer. Math.
              Soc., 54 (1948), 352-360.
           E. Schroeder, Vier combinatorische Probleme, Zeit. f. Math. Phys., 15
              (1870), 361-376.
           R. P. Stanley, Enumerative Combinatorics, Cambridge, Vol. 2, 1999;
              see page 178.
           ....................................
Links:     H. Bottomley, Illustration of initial terms
           E. Pergola and R. A. Sulanke,
              Schroeder Triangles, Paths, and Parallelogram Polyominoes, J. Integer
              Sequences, 1 (1998), #98.1.7.
Formula:   Recurrence: a(1) = a(2) = 1; for n > 1, (n+1)a(n+1) =
              3(2n-1)a(n)-(n-2)a(n-1). G.f.: (1/4)*(1+x-sqrt(1-6*x+x^2)).
           For n > 1, a(n) = (1/n) * sum_{k = 0 .. n-2}
              binomial(2*n-k-2,n-1)*binomial(n-2,k).
Example:   a(3) = 3: abc, a(bc), (ab)c; a(4) = 11: abcd, (ab)cd, a(bc)d, ab(cd),
              (ab)(cd), a(bcd), a(b(cd)), a((bc)d), (abc)d, (a(bc))d, ((ab)c)d.
Maple:     t1:=(1/4)*(1+x-sqrt(1-6*x+x^2)); series(t1,x,40);
Math'ca:   Sch[ 1 ]=Sch[ 2 ]=1;Sch[ n_Integer ]:=Sch[ n ]=(3(2n-3)Sch[ n-1 ]
              -(n-3)Sch[ n-2 ])/n; Array[ Sch[ # ]&, 20 ]
See also:  See A000108, A001190, A001699, A000081 for other ways to count
              parentheses. Cf. A000311, A010683, A065096.
Keywords:  nonn,easy,nice
```