"A Handbook of Integer Sequences" Fifty Years Later

N. J. A. Sloane,
The OEIS Foundation Inc.,
11 So. Adelaide Ave., Highland Park, NJ 08904, USA
Email: njasloane@gmail.com

February 24, 7am, 2023

Before 1973 there was no database of integer sequences. Someone coming across the sequence $1, 2, 4, 9, 21, 51, 127, \ldots$ would have had no way of discovering that it had been studied since 1870 (today these are called the Motzkin numbers, and form entry A001006[1] in the database). Everything changed in 1973 with the publication of *A Handbook of Integer Sequences*, which listed 2372 entries. This article describes the fifty-year evolution of the database from the *Handbook* to its present form as *The On-Line Encyclopedia of Integer Sequences* (or *OEIS*), which contains $360,000$ entries, receives a million visits a day, and has been cited $10,000$ times, often with a comment saying "discovered thanks to the OEIS".

Integer Sequences

Number sequences arise in all branches of science: for example, $1, 1, 2, 4, 9, 20, 48, 115, \ldots$ gives the number of rooted trees with n nodes (A000081, see also Fig. 1), and in daily life: how many pieces can you cut a pancake into with n knife-cuts? The pieces need not all be the same size. That one is easy: $1, 2, 4, 7, 11, 16, \ldots, n(n+1)/2+1$ (A000124). But what is the answer for cutting up an (ideal) bagel or doughnut? That is a lot harder: with a sharp knife you might get a few terms, perhaps $1, 2, 6, 13, \ldots$, but probably not enough to guess the formula, which is $n(n^2 + 3n + 8)/6$ for $n > 0$. For that you would need to to consult the database: go to https://oeis.org and enter "cutting bagel", or go directly to A003600.

My fascination with these sequences began in 1964 when I was a graduate student at Cornell University in Ithaca, NY, studying neural networks. I had encountered a sequence of numbers, $1, 8, 78, 944, 13800, \ldots$, and I badly needed a formula for the n-th term, in order to determine the rate of growth of the terms (this would indicate how long the activity in this very simple neural network would persist).

I noticed that although several books in the Cornell library contained sequences somewhat similar to mine, as far as I could tell this particular sequence was not mentioned. I expected to have to analyze many related sequences, so in order to keep track of the sequences in these books, I started recording them on 3" × 5" file cards.

---

[1] Six-digit numbers prefixed by A refer to entries in the current version of the *Handbook*, *The On-Line Encyclopedia of Integer Sequences* [13].
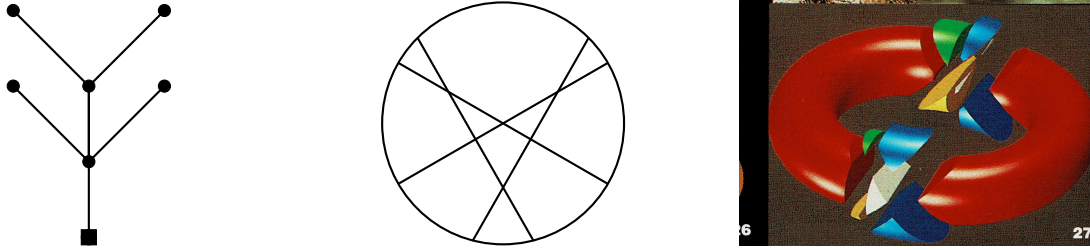
Figure 1: Left: one of 48 unlabeled rooted trees with 7 nodes (the root node is at the bottom); center: four cuts of a pancake can produce 11 pieces; right: three cuts of a bagel can produce 13 pieces.

The collection grew rapidly as I searched though more books, and once the word got out, people started sending me sequences. Richard Guy was an enthusiastic supporter right from the start. In 1973 I formalized the collection as *A Handbook of Integer Sequences*, which was published by Academic Press (Fig. 2). It contained 2372 entries.
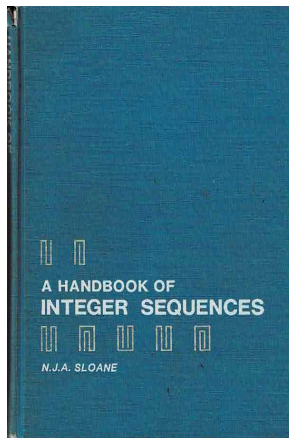

Figure 2: Front cover of the *Handbook*. The embossed figures show side views of the two ways of folding a strip of three (blank) stamps, and the five ways of folding a strip of four stamps. The full sequence begins $1, 1, 2, 5, 14, 38, 120, 353, 1148, 3527, \ldots$, A001011. No formula is known.

Once the book appeared, the flood of correspondence increased, and it took twenty years to prepare the next version. Simon Plouffe helped a great deal, and in 1995 Academic Press published our sequel, *The Encyclopedia of Integer Sequences*, with 5487 entries. From this point on the collection grew even more rapidly. I waited a year, until it had doubled in size, and then put it on the Internet, calling it *The On-Line Encyclopedia of Integer Sequences.*

In the rest of this article I will first say more about the evolution of the database: the *Handbook*, the 1995 *Encyclopedia*, the *On-Line Encyclopedia*, and the *OEIS Foundation*. The next sections describe the database itself: what sequences are—or are

not—included, how the database is used, the layout of a typical entry, the arrangement of the entries, and a fact sheet. The final sections describe some especially interesting sequences: Recamán's sequence, the iteration of number-theoretic functions, Gijswijt's sequence, lexicographically earliest sequences, the stepping stones problem, and stained glass windows, These last sections mention several open questions to which I would very much like to know the answers.

Notation. $a(n)$ denotes the n-th term of the sequence under discussion. $\sigma(n)$ is the sum of the divisors of n (A000203).

Evolution of the database

The *Handbook of Integer Sequences*

Once the collection had grown to a few hundred entries, I entered them on punched cards,[2] which made it easier to check and sort them. The *Handbook* was typeset directly from the punched cards. There were a few errors in the book, but almost all of them were caused by errors in the original publications. Accuracy was a primary concern in that book, as it is today in the OEIS.

The book was an instant success. It was, I believe, the world's first dictionary of integer sequences (and my original title said *Dictionary* rather than *Handbook*). Many people said "What a great idea", and wondered why no one had done it before. Martin Gardner recommended it in the *Scientific American* of July 1974. Lynn A. Steen, writing in the *American Mathematical Monthly* said "Incomparable, eccentric, yet very useful. Contains thousands of 'well-defined and interesting' infinite integer sequences together with references for each ... If you ever wondered what comes after $1, 2, 4, 8, 17, 35, 71, \ldots$, this is the place to look it up".

Harvey J. Hindin, writing from New York City, exuberantly concluded a letter to me by saying: "There's the *Old Testament*, the *New Testament*, and the *Handbook of Integer Sequences*."

I never did find the sequence that started it all in the literature, but I learned Pólya's theory of counting, and with John Riordan's help found the answer, which appears in [16] and A000435.

The *Encyclopedia of Integer Sequences*

Following the publication of the *Handbook*, a large amount of correspondence ensued, with suggestions for further sequences and updates to the entries. By the early 1990's over a cubic meter of new material had accumulated. A Canadian mathematician, Simon Plouffe, offered to help in preparing a revised edition of the book, and in 1995

---

[2]These were never called "punch cards" (sic). To anyone who worked with them in the 1960s, "punch cards" sounds like "grill cheese" (sic) for "grilled cheese", or "barb wire" (sic) for "barbed wire", both of which I have recently seen in print.

The Encyclopedia of Integer Sequence, by me and Plouffe, was published by Academic Press. It contained 5487 sequences, occupying 587 pages. By now punched cards were obsolete, and the entries were stored on magnetic tape.

The *On-Line Encyclopedia of Integer Sequences*

Again, once the book appeared, many further sequences and updates were submitted from people all over the world. I waited a year, until the size of the collection had doubled, to 10000 entries, and then in 1996 I launched *The On-Line Encyclopedia of Integer Sequences* (now usually called simply the *OEIS*) on the Internet. From 1996 until October 26, 2009, it was part of my homepage on the AT&T Labs website.

Incidentally, in 2004 the database was mentioned by the Internet website *slashdot* ("News for Nerds. Stuff that Matters"), and this brought so much traffic to my Bell Labs homepage that it briefly crashed the whole Bell Labs website. My boss was quite proud of this, since it was a rare accomplishment for the Mathematics and Statistics Research Center.

The OEIS Foundation

In 2009, in order to ensure the long-term future of the database, I set up a non-profit foundation, *The OEIS Foundation Inc.*, a 501(c)(3) Public Charity, whose purpose is to own, maintain and raise funds to support *The On-Line Encyclopedia of Integer Sequences* or *OEIS*.

On October 26, 2009, I transferred the intellectual property of *The On-Line Encyclopedia of Integer Sequences* to the Foundation. A new OEIS with multiple editors was launched on November 11, 2010.

Since then it has been possible for anyone in the world to propose a new sequence or an update to an existing sequence. To do this, users must first register, and then submissions are reviewed by the editors before they become a permanent part of the OEIS. Technically the OEIS is now a "moderated wiki".

I started writing this article on November 11, 2022, noting that this marked twelve years of successful operation of the online OEIS, and that the database is in its 59th year of existence.

The database today

What sequences are included?

From the very beginning the goal of the database has been to include all "interesting" sequences of integers. This is a vague definition, but some further examples may make it clearer. The database includes a huge number of familiar and unfamiliar sequences from mathematics (the prime numbers, $2, 3, 5, 7, 11, 13, \ldots$, A000040; the orders of noncyclic

simple groups, $60, 168, 360, 504, 660, 1092, \ldots$, A001034), computer science (the number of comparisons needed for merge sort, $0, 1, 3, 5, 8, 11, 14, \ldots$, A001855), physics (see "self-avoiding walks on lattices", Ising model, etc., e.g. A002921), chemistry (the enumeration of chemical compounds was one of the motivations behind Pólya's theory of counting, see e.g. A000602), and not least, from puzzles and I.Q. tests ($1, 8, 11, 69, 99, 96, 111, \ldots$, the "strobogrammatic" numbers, guess!, or see A000787; $4, 14, 23, 34, 42, 50, 59, \ldots$, the stops with numerical values for the A Train (8 Avenue express) in Manhattan, as of January 2023 A011554. The latter entry has links to a map and the train schedule).

Sequences that have arisen in the course of someone's work—especially if published—have always been welcomed. On the other hand, sequences that have been proposed simply because they were missing from the database are less likely to be accepted. And of course, the sequence must be well defined.

Very short sequences and sequences that are subsequences of many other sequences are not accepted. A sequence for which the only known terms are $2, 3, 5, 7$ would not be accepted since it is matched by a large number of existing sequences. The definition may not involve an arbitrary but large parameter (primes ending in 1 are fine, A030430, but not primes ending in 2023).

The OEIS Wiki has a section listing additional examples of what not to submit, as well as a great deal of information about the database that I won't repeat here, such as the meaning of the various keywords, the definition of the "offset" of a sequence, descriptions of the submission and editorial processes, and a list of over $10,000$ citations of the OEIS in the scientific literature.

Most OEIS entries give an ordered list of integers. But triangles of numbers are included by reading them row-by-row. For example, Pascal's triangle becomes 1, $1, 1$, $1, 2, 1$, $1, 3, 3, 1, \ldots$, A007318. Doubly-infinite square arrays are included by reading them by antidiagonals: the standard multiplication table for positive integers becomes 1, $2, 2$, $3, 4, 3$, $4, 6, 6, 4, \ldots$, A003991.

Sequences of fractions are included as a linked pair giving the numerators and denominators separately (the Bernoulli numbers are A027641/A027642). Important individual real numbers are included by giving their decimal or continued fraction expansions (for π see A000796 and A001203). A relatively small number of sequences of nonintegral real numbers are included by rounding them to the nearest integer, or by taking floors or ceilings (the imaginary parts of the zeros of Riemann's zeta function give A002410).

Two less obvious sources for sequences are binomial coefficient identities and number-theoretic inequalities. The values of either side of the identity

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

[9, (3.68)] give A036910. From the inequality $\sigma(n) < n\sqrt{n}$ for $n > 2$, [12, Sect. III.1.1.b], we get the integer sequence $\lfloor n\sqrt{n} \rfloor - \sigma(n)$, A055682. The point here is that if you want to find out if this inequality is known, you look up the difference sequence, and find

[A055682](#) and a reference to the proof. Many more sequences of these two types should be added to the database.

How the database is used

The main applications of the database are in identifying sequences or in finding out the current status of a known sequence. Barry Cipra has called it a mathematical analogue of a "fingerprint file". You encounter a number sequence, and wish to know if anyone has ever come across it before.

If your sequence is in the database, the reply will give a definition, the first 50 or so terms, and, when available, formulas, references, computer code for producing the sequence, links to any relevant web sites, and so on.

Figures 3 and 4 show what happens if you submit $1, 2, 5, 14, 42, 132, 429$, the first few Catalan numbers, one of the most famous sequences of all.

Figure 3: The result of submitting $1, 2, 5, 14, 42, 132, 429$ to the database. This figure shows the banner at the top of the reply. There are 26 matches, ranked in order of importance, the top match being the one we want, the Catalan numbers. A shortened version of the top match is shown in the next figure.

I could have chosen a simpler example, like the Fibonacci numbers, but I have a particular reason for choosing the Catalan numbers. When the OEIS was new, people would sometimes say to me that they had a sequence they were trying to understand, and would I show them how to use the database. At least twice when I used the Catalan sequence as an illustration, they said, why, that is my sequence, how on earth did you know? It was no mind-reading trick, the Catalan numbers are certainly the most common sequence that people don't know about. This entry is the longest—and one of the most important—in the whole database.

If we do not find your sequence in the database, we will send you a message inviting you to submit it (if you consider it is of general interest), so that the next person who

A000108	Catalan numbers: $C(n) = \text{binomial}(2n, n)/(n+1) = (2n)!/(n!(n+1)!)$.
	(Formerly M1459 N0577)
DATA	1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, ...
COMMENTS	These were formerly sometimes called Segner numbers.
	A very large number of combinatorial interpretations are known -
	see references, esp. R. P. Stanley, Catalan Numbers, Camb., 2015.
	This is probably the longest entry in the OEIS, and rightly so.
	The solution to Schröder's first problem: number of ways to insert n pairs
	of parentheses in a word of $n+1$ letters. E.g., for $n = 2$ there are 2 ways:
	((ab)c) or (a(bc)); for n=3 there are 5 ways: ((ab)(cd)), (((ab)c)d), ...
	...
REFERENCES	The large number of references and links demonstrates the ubiquity
	of the Catalan numbers.
	R. Alter, Some remarks and results on Catalan numbers, pp. 109-132
	in Proc. Louisiana Conf. Combinatorics, Graph Theory and
	Computer Science. Vol. 2, edited R. C. Mullin et al., 1971.
	M. Bona, ed., Handbook of Enumerative Combinatorics, CRC Press, 2015
	L. Comtet, Advanced Combinatorics, Reidel, 1974, p. 53.
	J. H. Conway & R. K. Guy, The Book of Numbers, Springer, 1995, 96-106.
	...
LINKS	Robert G. Wilson v, Table of n, a(n) for n = 0..1000
	...
	F. R. Bernhart, Catalan, Motzkin and Riordan numbers, Discr. Math.,
	Vol. 204, No. 1-3 (1999), 73-112.
	...
	W. G. Brown, Historical Note on a Recurrent Combinatorial Problem,
	Amer. Math. Monthly,. 72, No. 9 (1965), 973-977.
	...
FORMULA	Recurrence: $a(n) = 2 * (2 * n - 1) * a(n-1)/(n+1)$ with $a(0) = 1$.
	...
MAPLE	A000108 := n->binomial(2*n, n)/(n+1);
	...
MATHEMATICA	A000108[n_] := (2 n)!/n!/(n+1)!
	...
PARI	a(n)=binomial(2*n, n)/(n+1)
	...
KEYWORD	core,nonn,easy,nice
AUTHOR	N. J. A. Sloane

Figure 4: The entry for the Catalan number A000108. The full entry has over 750 lines, which have been edited here to show samples of the different fields.

comes across it will be helped, and your name will go on record as the person who submitted it.

The second main use of the database is to find out the latest information about a particular sequence.

Of course we cannot hope to keep all 360000 entries up-to-date. But when a new paper is published that mentions the OEIS, Google will tell us, and we then add links to that paper from any sequence that it mentions. People have told us that this is one of

the main ways they use the OEIS. After all, even a specialist in (say) permutation groups cannot keep track of all the papers published worldwide in that area. And if a paper in a physics or engineering journal happens to mention a number-theoretic sequence, for example, Google will notify us and we will record it.

There are also many other ways in which the database has proved useful.

For example, it is an excellent source of problems to work on. The database is constantly being updated. Every day we get thirty to fifty submissions of new sequences, and an equal number of comments on existing entries (new formulas, references, additional terms, etc.). The new sequences are often sent in by non-mathematicians, and are a great source of problems. You can see the current submissions at `https://oeis.org/draft`. Often enough you will see a sequence that is so interesting you want to drop everything and work on it. And remember that we are always in need of more volunteer editors. In fact anyone who has registered with the OEIS can suggest edits, you do not even need to be an official editor. We have been the source of many international collaborations.

There is also an educational side: several people have told us that they were led into mathematics through working as an editor. Here is a typical story.

Subject: Reminiscence from a young mathematician

"I wanted to relay a bit of nostalgia and my heartfelt thanks. Back in the late 1990s, I was a high school student in Oregon. While I was interested in mathematics, I had no significant mathematically creative outlet until I discovered the OEIS in the course of trying to invent some puzzles for myself. I remember becoming a quite active contributor through the early 2000s, and eventually at one point, an editor. My experience with the OEIS, and the eventual intervention of one of my high school teachers, catalyzed my interest in studying mathematics, which I eventually did at [...] College. I went on to a Ph.D. in algebraic geometry at the University of [...], and am currently at [...].

"I wanted to thank you for seriously engaging with an 18-year old kid, even though I likely submitted my fair share of mathematically immature sequences. I doubt I would have become a mathematician without the OEIS!"

A less-obvious use of the database is to quickly tell you how hard a problem is. I use it myself in this way all the time. Is the sequence "Catalan" or "Collatz"? If a sequence comes up in your own work, or when reviewing someone else's work, it is useful to know right away if this is a well-understood sequence, like the Catalan numbers, or if it is one of the notoriously intractable problems like the Collatz or $3x + 1$ problem (A006577).

Finally, the OEIS is a welcome escape when you feel the world is falling apart. Take a look at Scott Shannon's drawings of stained glass windows in A331452; or Jonathan Wild's delicate illustrations of the ways to draw four circles in A250001; or Éric Angelini's "1995" puzzle (A131744) or any of his "lexicographically earliest sequences" (A121053, A307720, and many more); or find better solutions to the Stepping Stones Problem (A337663). You can find brand new problems at any hour of the day or night by looking at the stack of recent submissions: but beware, you may see a problem there that will keep you awake for days. Or search in the database for phrases like "It appears that ...", or "Conjecture: ...", or "It would be nice to know more!"

Layout of a typical entry

This is a good place to mention some of the features of an OEIS entry. Most of the fields (see Figs. 3 and 4) are self-explanatory. At the top it tells you how many matches were found to your query (26 in the example). These are ranked in order of importance.

The DATA section shows the start of the sequence, usually enough terms to fill a few lines on the screen (typically 300 to 500 decimal digits). All terms listed must be known to be correct, and there cannot be any gaps. If the first n terms are known but the $n + 1$st is known only to be either 14 or 15, then the listing of the sequence must end with the nth term. In the case of Mersenne primes (A000043), it is common for later primes to be discovered before all smaller candidates have been tested. Until it is known for certain that the new discoveries are indeed the *next* terms, they cannot be added to the sequence (although of course they can be mentioned in comments).

Often one wants more terms than are shown in the DATA section, and in many cases the first link in the entry will point to a plain text file with perhaps 10000 or 20000 terms. That file will have a name like b001006.txt, and is called the "b-file" for the sequence. Some entries also have much larger tables, giving a million or more terms.

If you click the "graph" button near the top of the reply, you will be shown two plots of the sequence, and if you click the "listen" button, you can listen to the sequence played on an instrument of your choice. The default instrument is the grand piano, and the terms of the sequence would then be mapped to the 80 keys by reducing the numbers mod 80 and adding 1.

I conclude this section with a philosophical comment.

When you are seriously trying to analyze a sequence, and are prepared to spend any amount of time needed (searching for a formula or recurrence, for instance), you need all the help you can get, which is why we provide the b-files and other data files, and why we give computer programs in so many languages. This is also the reason we give as many references and links as possible for a sequence. Even if the reference is to an ancient or obscure journal, or one that has been accused as being "predatory", we still give the reference, especially for sequences that are not well-understood. The same thing holds for formulas, comments, and cross-references to other sequences. When you are desperate, you will accept help from anywhere. And do not forget "Superseeker", which invokes a collection of algorithms to try to analyze your sequence or to transform it into an OEIS entry.

Arrangement of the entries

The entries in the database are (virtually) arranged in two different ways, the first essentially chronological, the second lexicographic.

The first is by their *absolute* identification number, or A-number.[3] Once the collection reached a few hundred entries, I sorted them into lexicographic order and numbered

---

[3]The sequences in the 1973 and 1995 books were numbered N0001,..., M0001,... respectively.

them A1, A2, A3, …. A1 gives the number of symmetry groups of order n, A2 is the famous Kolakoski sequence, and so on. This numbering is still used today, only A1 has become [A000001](), A2 is [A000002](), ..., and as each new submission comes in it gets the next free number. Currently, sequences are being issued numbers around A360000. Rejected A-numbers are recycled, so there are no gaps in the order. We reached 100000 entries in 2004, and 250000 in 2015. The present growth rate is about 12000 new entries each year.

The second arrangement is a kind of lexicographic ordering. First I describe an idealized, theoretical, lexicographic order. Sequences of nonnegative numbers can be arranged in lexicographic (or dictionary) order. For example, sequences beginning $1, 2, 4, \ldots$ come before $1, 2, 5, \ldots$, $1, 2, 4, 3, \ldots$, $1, 3, \ldots$, etc., but after $1, 2, 3, \ldots$. Also $1, 2, 4, \ldots$ comes after the two-term sequence $1, 2$ (because blanks precede numbers).

More formally, we compare the two sequences term-by-term, and in the first position where they differ, whichever is smaller (or blank) is the lexicographically earlier one.

For sequences with negative terms, we ignore the signs and sort according to the absolute values.

Here is the actual ordering used in the OEIS. The sequences are arranged (virtually) into a version of lexicographic order, according to the following rules. First, delete all minus signs. Then find the first term that is greater than 1, and discard all the terms before it. What's left determines its position in the lexicographic order. For example, to place $-1, 0, 1, 1, \underline{2}, 1, 17, -3, -2, 6, \ldots$ in the ordering, we would ignore the terms before the underlined 2, and consider the sequence as beginning $2, 1, 17, 3, 2, 6, \ldots$.

Sequences that contain only 0s, 1s and −1s are sorted into lexicographic order by absolute value and appear at the beginning of the ordering. The first sequence in the database is therefore the zero sequence [A000004]().

In this way every sequence has a unique position in the ordering. The sequences have been sorted in this way since the 1960s. For the first ten years the punched card entries were physically sorted into this order.

When you look at any OEIS entry, [A005132]() say, towards the bottom you will see[4] two lines like

Sequence in context: A277558 A350578 A335299 * A064388 A064387 A064389

Adjacent sequences: A005129 A005130 A005131 * A005133 A005134 A005135

which tell you the three entries immediately before and after that entry in the lexicographic ordering, and the three entries before and after it in the A-numbering. The asterisks represent the sequence you are looking at. The first group can be useful if you are uncertain about a term in your sequence, the second in case you want to look at other sequences submitted around that time.

Today the sequences are stored internally in an SQLite database. However, the punched card format has been so useful that when you view a sequence, as in Fig. 4, it is still presented to you in something very like the old punched card format.

---

[4]If you don't see them, click on the A-number at the top of the entry.

Summary: "A Handbook of Integer Sequences" today

- Now *The On-Line Encyclopedia of Integer Sequences* or *OEIS*: https://oeis.org
- Accurate information about 360000 sequences.
- Definition, formulas, references, links, programs. View as list, table, graph, music!
- Traffic: 1 million hits/day.
- 30 new entries, 50 updates every day.
- Often called one of best math sites on the Web. Fingerprint file for mathematics.
- "Street creds": 10000 citations.
- A moderated Wiki, owned by OEIS Foundation, a 501(c)(3) public charity.
- Uses: to see if your sequence is new, to find references, formulas, programs.
- Catalan or Collatz? (Very easy or very hard?)
- Source of fascinating research problems;[5] low-hanging fruit from recent submissions.
- Accessible (free, friendly).
- Fun $(1, 2, 4, 6, 3, 9, 12, 8, 10, 5, 15, ...?)$. Interesting, educational. Escape.
- Addictive (better than video games).
- Has led many people into mathematics.
- One of the most successful international collaborations, a modest contribution towards world peace.
- Need editors.

Some favorite sequences

I'm sometimes asked for my favorite sequence. This is a difficult question. I'm tempted to reply by saying: If you were the keeper of the only zoo in the world, how would you answer that question? (Because that is roughly the situation I'm in.) Would you pick one of the exotic animals, a giraffe, a kangaroo, or a blue whale? Or one of the essential animals, like a horse, a cow, or a duck? If the question came from a visiting alien, of course, there is only one possible answer: a human being.

For sequences, the essential ones are the primes, the powers of 2, the Catalan numbers, or (especially if the question came from an alien with no fingers or toes), the counting sequence 0, 1, 2, 3, 4, ... (A001477).

But here I'll mention a few that are fairly exotic. The Recamán and Gijswijt sequences have simple recursive definitions, yet are astonishingly hard to understand.

---

[5] Look for "Conjecture", "It appears that", "It would be nice to", ...

Recamán's sequence (A005132)

This remarkable sequence has resisted analysis for over 30 years, even though we have computed an astronomical number of terms. It was contributed to the database by Bernardo Recamán Santos in 1991.

The definition is deceptively simple. The first term is 0. We now add or subtract 1, then we add or subtract 2, then add or subtract 3, and so on. The rule is that we always first try to subtract, but we can only subtract if that leaves a nonnegative number that is not yet in the sequence. Otherwise we must add.

Here is how the sequence starts. We have the initial 0. We can't subtract 1, because that would give a negative number, so we add 1 to 0. So the second term is 1. We can't subtract 2 from 1, so we add it, getting the third term $1 + 2 = 3$. Again we can't subtract 3, for that would give 0, which has already appeared, so we add 3, getting the fourth term $3 + 3 = 6$.

Now we must add or subtract 4, and this time we can subtract, because $6 - 4 = 2$, and 2 is nonnegative and a number that hasn't yet appeared. So at this point the sequence is $0, 1, 3, 6, 2$. Then it continues $7(= 2 + 5)$, $13(= 7 + 6)$, $20(= 13 + 7)$, $12(= 20 - 8)$, and so on. The first 16 terms are

$$0, 1, 3, 6, 2, 7, 13, 20, 12, 21, 11, 22, 10, 23, 9, 24, \ldots$$

When adding rather than subtracting, repeated terms are permitted (42 is repeated at the 24th term).

Edmund Harriss has found an elegant way to draw the sequence as a spiral on the number line. Start at 0, and when we subtract n, draw a semicircle of diameter n to the left from the last point, or to the right if we are adding n. Draw the semicircles alternately below and above the horizontal axis so as to produce a smooth spiral.

The main question about this sequence is: Does every positive number appear? What makes this sequence so interesting is that certain numbers (for reasons we do not understand) are extremely reluctant to appear. 4 does not appear until 131 steps, and 19 takes 99734 steps.

A group of us at AT&T Bell Labs worked on this in 2001, and found a way to greatly speed up the computation. Allan Wilks used it to compute the first 10^{15} terms, and found that 2406 (which had been missing for a long time) finally appeared at step 394178473633984.

At this point the smallest missing number was $852655 = 5 \cdot 31 \cdot 5501$. Benjamin Chaffin has continued this work, and in 2018 reached 10^{230} terms. However, 852655 was still missing, and there has been no progress since then.

Thirty years ago I thought that every number would eventually appear. Now I am not so sure. My current belief is that there are two possibilities: either there are infinitely many numbers that never appear, and 852655 just happens to be the smallest of them, but has no other special property. A similar phenomenon seems to occur when iterating various number-theoretic functions—see the next section. Or, every
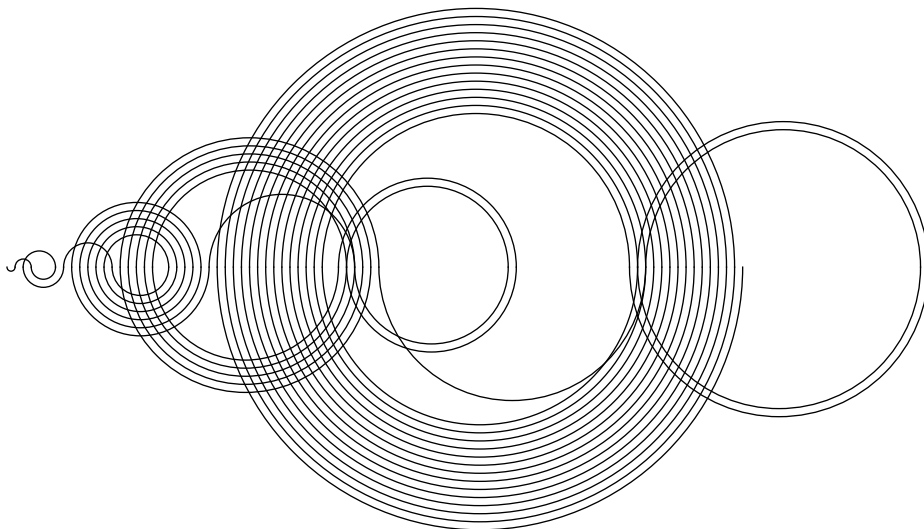
Figure 5: Harriss's drawing of the first 64 terms of Recamán's sequence. (The tiny initial semicircle, at the extreme left, is below the axis. It has diameter 1 and joins the points 0 and 1. It continues as a semicircle of diameter 2, above the axis, joining the points 1 and 3.)

number will eventually appear (just as presumably every one of Shakespeare's plays will eventually appear in the expansion of π in base 60), although we may never be able to extend the sequence far enough to hit 852655. For the latest information about this sequence (or any other sequence mentioned in this article), consult the OEIS.

Open question: Does 852655 appear in A005132?

Iteration of number-theoretic functions

Many mysterious sequences arise from the iteration of number-theoretic functions. A classic problem concerns the iteration of the function $f(n) = \sigma(n) - n$, the sum of the divisors of n that are less than n (technically, the sum of the "aliquot parts" of n) A001065. For an initial value of n, what happens to the sequence (or "trajectory") $n, f(n), f(f(n)), f(f(f(n))), \ldots$? All $n < 276$ terminate by entering a cycle (such n are called "perfect", "amicable", or "sociable" numbers), or reaching a prime, then 1, then 0.

But it appears likely that $n = 276$ and perhaps all sufficiently large even numbers, will never terminate [6], [10, §B6]. The trajectory of 276 is sequence A008892. At the time of writing, 2145 terms of this trajectory have been computed, and it is still steadily growing, term 2145 being a 214-digit number [8].

There are arguments which suggest that 276 will eventually terminate, and other arguments that suggest it will grow forever. It is surprising that even today mathematics

13

cannot resolve such a concrete question.

[A098007](#) gives the number of distinct terms in the trajectory of aa general initial number n, or -1 if the trajectory is unbounded. The value of [A098007](#)(276) is unknown.

If indeed 276 does go to infinity, it is natural to ask, how did 276 know it was destined to be the first immortal number under the map f? The answer may be that there are infinitely many immortal numbers, and 276 just happens to be the first. It got lucky, that's all! Just as 852655 got lucky in Recamán's problem.

A similar question, discussed by Richard K. Guy [10, §B41], which has received much less attention, concerns the map $g(n) = (\sigma(n) + \phi(n))/2$, where $\phi(n)$ is the Euler totient function [A000010](#). The trajectory may end at 1, a prime, or a fraction, or it may increase monotonically to infinity. Sequence [A292108](#) gives the number of steps in the trajectory, or -1 if the trajectory is infinite. All numbers $n < 270$ have finite trajectories, but it appears that 270 goes increases forever. The trajectory of 270 is [A291789](#). For this problem there is less doubt about what happens, because Andrew Booker has given a heuristic argument showing that almost all numbers go to infinity. What makes 270 the first immortal number under g? Again I suspect it just got lucky!

Open questions: Does the trajectory of 276 under f increase forever? What about the trajectory of 270 under g?

Gijswijt's sequence (A090822)

For this sequence it will be helpful to remember that chemists do not write HHO, they write H_2O, they do not write $AlAlAlSOOOOSOOOO$, they write $Al_3(SO_4)_2$. We will apply a similar compression to sequences of numbers, except that we indicate repetition by superscripts rather than by subscripts.

For this problem, when we look at a sequence of numbers, we want to write it in the form $XYY \dots Y$, or XY^k, where X and Y are themselves sequences of numbers, X can be missing, and the exponent k is *as large as possible.*

For example, we can write $1, 2, 2, 2, 2$ as XY^k, where $X = 1$, $Y = 2$, and $k = 4$. The highest k we can achieve for a sequence is called its "curling number". So $1, 2, 2, 2, 2$ has curling number 4. Think of an animal with its head looking to the left, with a very curly tail. X represents the head and body of the animal, and Y^k represents the curls in its tail.

Consider the sequence $3, 2, 4, 4, 2, 4, 4, 2, 4, 4$. We could take $X = 3, 2, 4, 4, 2, 4, 4, 2$ and $Y = 4$, getting XY^2, with $k = 2$, or we could take $X = 3$, $Y = 2, 4, 4$, getting XY^3, with $k = 3$, which is larger. So this sequence has curling number 3.

Remember that X may be missing. So the sequence with a single term 99, say, can be written as Y^1 where Y is the number 99, and it has curling number 1. The notion of curling number is independent of the base in which the numbers are written.

We are now ready to define Dion Gijswijt's absolutely brilliant sequence, which he sent to the OEIS in 2004.

The rule for finding the next term is simple: it is the curling number of the sequence so far. And you start with 1. That's the sequence!

So let's construct it. We start with 1, and the curling number of 1 is 1. So now we have $1, 1$. This has curling number 2, so now we have $1, 1, 2$. At each step we recompute the curling number, and make that the next term.

Here are the first few generations.

```
1
1  1
1  1  2
1  1  2  1
1  1  2  1  1
1  1  2  1  1  2
1  1  2  1  1  2  2
1  1  2  1  1  2  2  2
1  1  2  1  1  2  2  2  3
```

To go from line 6 to line 7, we took $Y = 1, 1, 2$. In line 8 we see the first 3, at the 9th term, and after a while a 4 appears at term 220.

But Gijswijt was unable to find a 5, and left that question open when he submitted the sequence. Some Bell Labs colleagues computed many millions of terms, but no 5 appeared.

Finally, over the course of a long weekend, Fokko van der Bult (a fellow student of Gijswijt's in Amsterdam) and I independently showed that there is a 5. In fact there are infinitely many 5's, but the first one does not appear until about term $10^{10^{23}}$. The universe would be cold long before any computer search would find it.

In the paper we wrote about the sequence [4], we also conjectured that the first time a number $N > 4$ appears is at about term

$$2 \uparrow (2 \uparrow (3 \uparrow (4 \uparrow (5 \uparrow \ldots \uparrow (N-1))))),$$

where the up-arrows ($\uparrow$) indicate exponentiation, a tower of exponents of height $N - 1$. This is a very slow-growing sequence.

A very recent manuscript by a student of Gijswijt's, Levi van de Pol [14], still under review, has extended our work, and may have proved the above conjecture.

I cannot resist adding a further comment about curling numbers, which if true shows that the Gijswijt sequence is in a sense universal. My "curling number conjecture" asserts that if any finite starting sequence is extended by the rule that the next term is the curling number of the sequence so far, then eventually the curling number will be 1.

If true, this implies that if the starting sequence contains no 1s, then the sequence eventually becomes Gijswijt's sequence [5, Th. 23]. In fact I conjecture that this is true for any starting sequence.

Open question: Is the curling number conjecture true?

Lexicographically Earliest Sequences

Although there is no space to discuss them in detail, there are many fascinating and difficult sequences in the OEIS whose definition has the form "Lexicographically Earliest Sequence of distinct positive numbers with the property that ...", where now we are using lexicographic in its pure sense, as mentioned when we were discussing the arrangement of sequences in the database. A favorite example is the EKG (or ECG) sequence A064413, whose definition is the lexicographically earlier infinite sequence of distinct positive numbers with the property that each term after the first has a nontrivial common factor with the previous term [11]. Other L.E.S. examples are the Yellowstone permutation A098550 [2], the Enots Wolley sequence A336957 (the name suggests the definition), and the Binary Two-Up sequence A354169 [7].

Open question: Show that the terms of the Enots Wolley sequence are precisely 1, 2, and all numbers with at least two distinct prime factors.

The Stepping Stones Problem (A337663)

This lovely problem was invented in 2020 by two undergraduates, Thomas Ladouceur and Jeremy Rebenstock. You have an infinite chessboard, and a handful of brown stones, which are worth one point each. You also have an infinite number of white stones, of values 2, 3, 4,..., one of each value. Suppose you have n brown stones. You start by placing them anywhere on the board. Now you place the white stones, trying to place as many as you can. The rules are that you can only place a white stone labeled k on a square if the values of the stones on the eight squares around it add up to k. And you must place the white stones in order, first 2, then 3, and so on. You stop when you cannot place the next higher-numbered white stone. The goal is to maximize the highest value that you place. Call this $a(n)$.

	9	5	10	11			
		4	1				
	12	8	3	2		16*	
			6	1	15		
		13	7	14			

Figure 6: A solution to the Stepping Stones problem for two starting stones. The high point $a(2) = 16$ here is indicated by an asterisk, as it is in the next three tables.

Say we start with $n = 2$ brown stones. There are infinitely many squares where they can be placed, but it turns out that the best thing is to place them so they are separated diagonally by a single blank square, as in Fig. 6. Now we start trying to place the white

stones. The 2 stone has to go between the two brown (or 1) stones, and then the 3 goes on a square adjacent to the 1 and the 2. There is now a choice for where the 4 goes, but the choice shown in Fig. 6 is the best. (After we have placed the 4, the neighbors of the 3 no longer add up to 3, but that is OK. It is only when we *place* the 3 that its neighbors must add to 3.) Continuing in this way, we eventually reach 16. There is nowhere to place the 17, so we stop. Ladouceur and Rebenstock showed, using a computer and considering all possible arrangements, that 16 is the highest value that can be attained with two starting stones. So $a(2) = 16$.

This is clearly a hard problem, since the number of possibilities grows rapidly with the number of brown stones. Only six terms of this sequence are known: $a(1)$ through $a(6)$ are $1, 16, 28, 38, 49, 60$. A solution for $n = 4$ found by Arnauld Chevallier is shown in Fig. 7. There are lower bounds for larger values of n which may turn out to be optimal. For $n = 7, \ldots, 10$ the current best constructions give $71, 80, 90, 99$. See A337663 for the latest information.

	35	18	36		23		21		32					
		17	1		14	9		12	20					
		34	16	15		5	4	8			26	27		
				31		10	1	3	19	25		1	28	
						11		2	6		33		29	
						24	13	22	1	7				
						37			30	38*				

Figure 7: A solution to the Stepping Stones problem for four starting stones.

We don't know how fast $a(n)$ grows. There have been a series of upper and lower bounds, initiated by Robert Gerbicz and Andrew Howroyd. The simple linear construction shown in Fig. 8 shows that $a(n) \geq 6(n-1)$ for $n \geq 3$.

	1										
		2	3	4	5	6	7	8	9		
			1			1			1	10	
		18*	17	16	15	14	13	12	11		

Figure 8: Every additional 1 on the middle row increases the number of white stones by 6, showing that $a(n) \geq 6(n-1)$ for $n \geq 3$.

By combining the constructions of Figs. 6 and 8, Menno Verhoeven obtained $a(n) \geq 6n+3$ for $n \geq 3$ (Fig. 9).

								25		
							24	**1**	26	
							23		27	
							22		28	
							21	**1**	29	
							20		30	
							19		31	
	9	5	10	11			18	**1**	32	
		4	**1**				17		33*	
12	8	3	2		16					
			6	**1**	15					
		13	7	14						

Figure 9: Combining the constructions of of Figs. 6 and 8 gives $a(n) \geq 6n+3$ for $n \geq 3$. The case $n = 5$ is shown. For other values of n, adjust the height of the "chimney" on the right.

The best lower bound for large n is due to Robert Gerbicz, who has shown by a remarkable extension of the construction in Figs. 8 and 9 that $\underline{\lim}_{n\to\infty} a(n)/n > 6$. (A preliminary version of his bound gives $a(n) > 6.0128\,n - 5621$ for all n, although the exact values of the constants have not been confirmed.) In his construction the "chimney" on the right of Fig. 9 gets expanded into a whole trellis.

One might think that with a sufficiently clever arrangement, perhaps extending the construction in Fig. 8 so that the path wraps around itself in a spiral, one could achieve large numbers with only a few starting stones. But a simple counting argument due to Robert Gerbicz shows this is impossible. The current best upper bound is due to Jonathan F. Waldmann, who has shown that $a(n) < 79n + C$ for some constant C. See A337663 for the latest information, including proofs of of the results mentioned here.

Open question: Improve the lower and upper bounds on $a(n)$. The lower bound looks especially weak.

Stained glass windows

In 1998 Poonen and Rubinstein [15] famously determined the numbers of vertices and cells in the planar graph formed from a regular n-gon by joining every pair of vertices by a chord. The answers are in A006561 and A007678. Lars Blomberg, Scott Shannon,
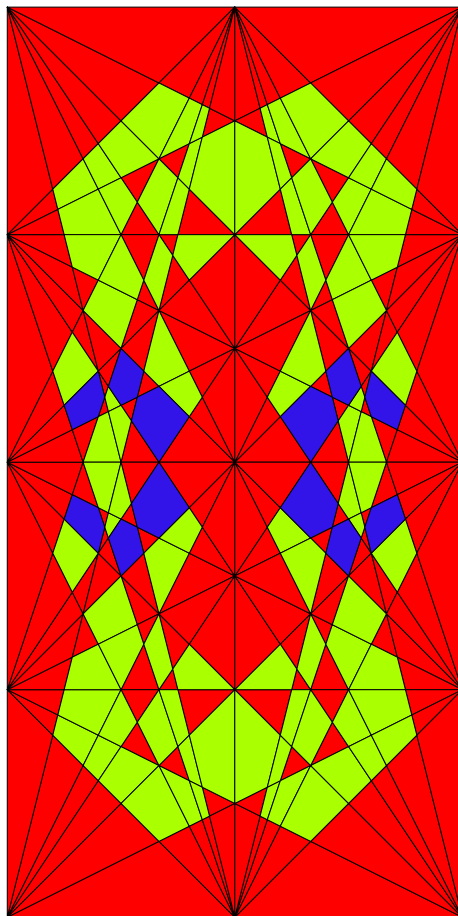
Figure 10: A 4 × 2 grid of squares with every pair of boundary points joined by a chord. The graph has 213 vertices and 296 cells. The cells are color-coded to distinguish triangles (red), quadrilaterals (yellow), and pentagons (blue).

and I have studied versions of this question when the regular n-gon is replaced by other polygons, for instance by a square in which n equally-spaced points are placed along each side and each pair of boundary points is joined by a chord. We also studied rectangles, triangles, etc. In most cases we were unable to find formulas for the numbers of vertices or cells, but we collected a lot of data, and the graphs, when colored, often resemble stained glass windows (see [3] and the illustrations in A331452 and other sequences cross-referenced there).[6] So we consoled ourselves with the motto: if we can't solve it, make art!

The most promising case to analyze seems to be the $n \times 2$ grid (although we did not succeed even there).

---

[6] There is no fee for downloading images from the OEIS, but if you use any of them, please credit the source!

Open question: How many vertices and cells are there in the graph for the $n \times 2$ grid, as illustrated for $n = 4$ in Fig. 10? Sequences A331763 and A331766 give the first 100 terms, yet even with all that data we have not found a formula.

The case of an $n \times n$ grid seems even harder. Figure 11 shows the 6×6 graph. Sequences A331449 and A255011 give the numbers of vertices and cells for $n \leq 42$. A334699 enumerates the cells by number of sides.
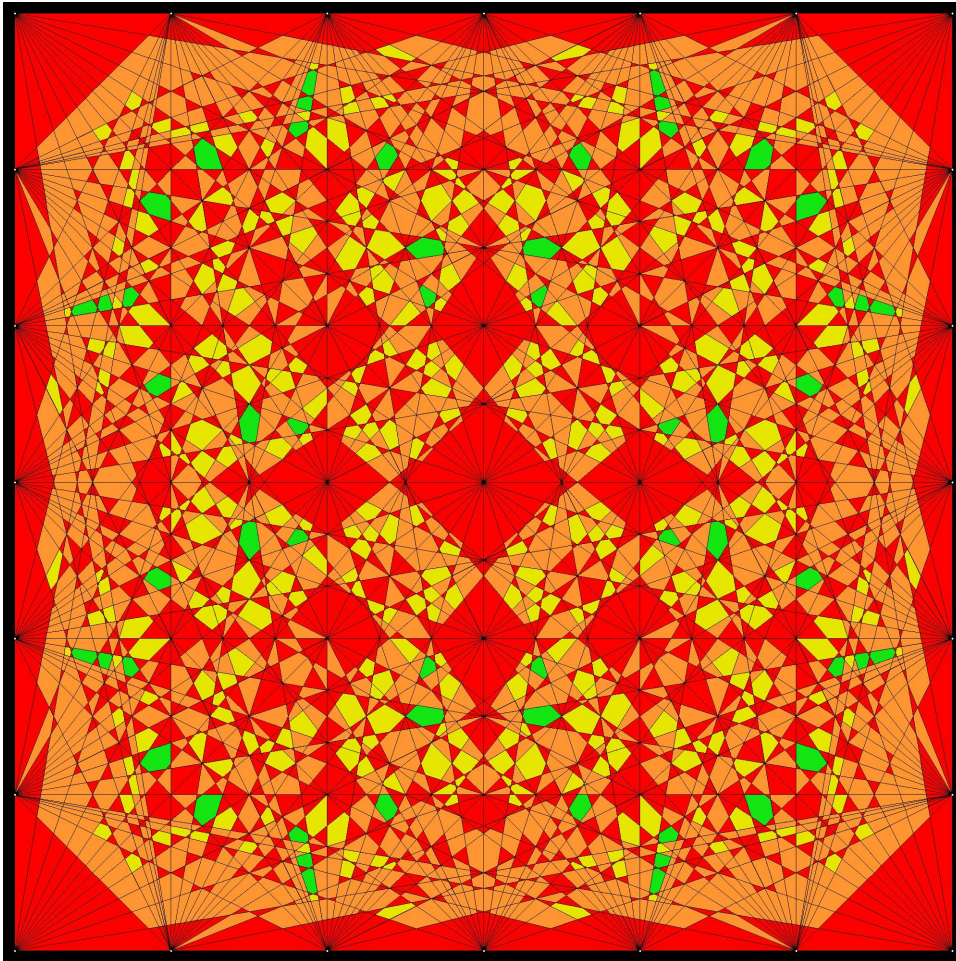
Figure 11: A 6×6 grid with every pair of boundary points joined by a chord. There are 4825 vertices and 6264 cells.

In the summer of 2022 Scott Shannon and I considered several other families of planar graphs. I cannot resist showing one of Shannon's graphs, a 16×16 grid, illustrating the 16th term of A355798 (Fig. 12). There are 61408 cells. Although Shannon has calculated 40 terms of this sequence, again no formula is known.
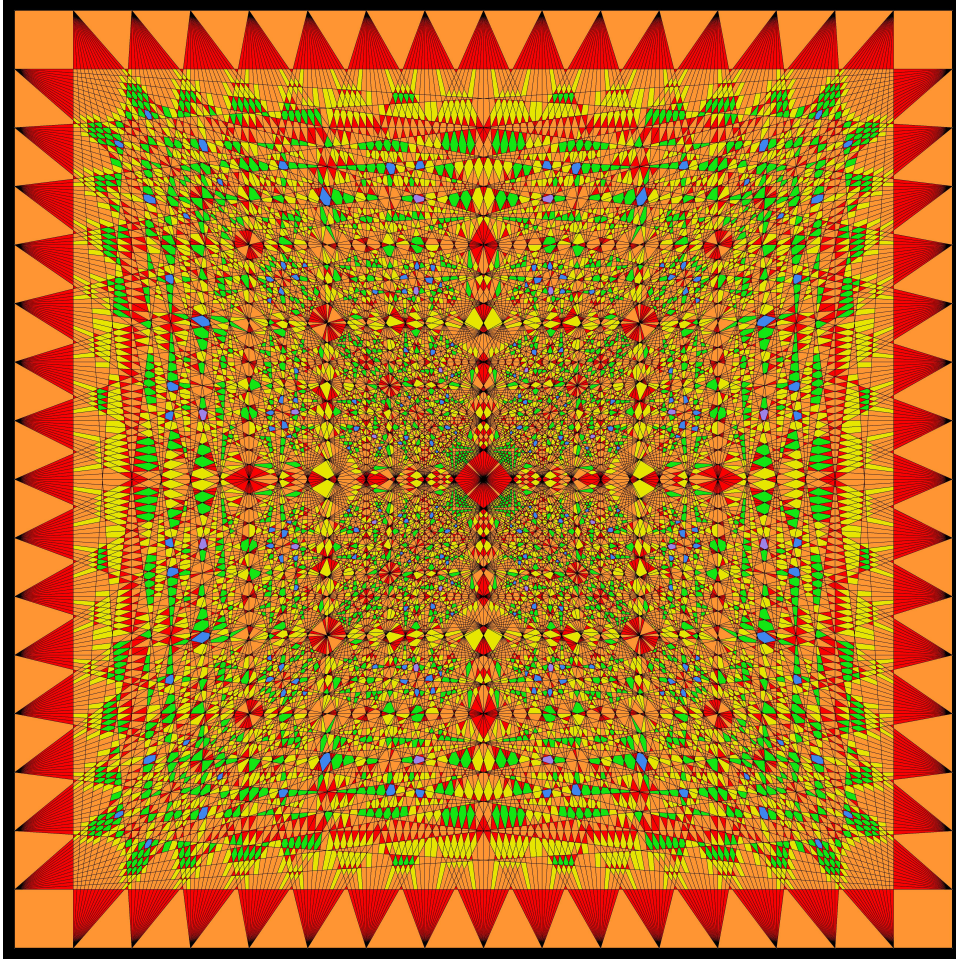
Figure 12: Scott Shannon's "Magic Carpet" graph, illustrating A355798(16).

Other sequences I would have liked to include

It is getting late, so I had better stop. Another time I will tell you about some very interesting sequences arising from the problem of dissecting a square to get a regular n-gon (A110312); from Gerrymandering (A341578, A348453, and many others): counting the ways in which circles can overlap (A250001); the inventory sequence A342585; Kaprekar's junction numbers (A006064, [1]); the kissing number problem (A001116, A257479); the neural network problem that started it all (A000435); and the squares in the plane problem (A051602).

And perhaps also meta-sequences such as A051070 ($a(n)$ is the nth term of A_n) and A107357 (the nth term is $1 +$ the nth term of A_n)!

A final comment: there are many videos on the Internet of talks I have given about sequences, including several that Brady Haran has made for the Youtube Numberphile channel (and that have been viewed over eight million times). See for example "Terrific

21

References

[1] M. A. Alekseyev and N. J. A. Sloane, On Kaprekar's Junction Numbers, *J. Combinat. and Number Theory*, 2023, in press.

[2] D. L. Applegate, H. Havermann, B. Selcoe, V. Shevelev, N. J. A. Sloane, and R. Zumkeller, The Yellowstone permutation, *J. Integer Seqs.*, **18** (2015), #15.6.7.

[3] L. Blomberg, S. R. Shannon, and N. J. A. Sloane, Graphical enumeration and stained glass windows, 1: Rectangular grids, *Integers*, Ron Graham Memorial Volume **21A** (2021), #A5.

[4] F. J. van de Bult, D. C. Gijswijt, J. P. Linderman, N. J. A. Sloane, and Allan Wilks, A slow-growing sequence defined by an unusual recurrence, *J. Integer Seqs.*, **10** (2007), #07.1.2.

[5] B. Chaffin, J. P. Linderman, N. J. A. Sloane and A. R. Wilks, On curling numbers of integer sequences, *J. Integer Seqs.*, **16** (2013), #13.4.3.

[6] K. Chum, R. K. Guy, M. J. Jacobson, Jr., and A. S. Mosunov, Numerical and statistical analysis of aliquot sequences, *Exper. Math.*, **29:4** (2020), 414–425.

[7] M. De Vlieger, T. Scheuerle, R. Sigrist, N. J. A. Sloane, and W. Trump, *The binary Two-Up sequence*, arXiv:2209.04108, 2022.

[8] FactorDB Database, factordb.com/sequences.php, 2023.

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

[10] R. K. Guy, *Unsolved Problems in Number Theory*, 3rd. ed., Springer, 2010.

[11] J. C. Lagarias, E. M. Rains, and N. J. A. Sloane, The EKG sequence, *Exper. Math.*, **11** (2002), 437–446.

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

[13] The OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, https://oeis.org, 2023.

[14] L. van de Pol, *The first occurrence of a number in Gijswijt's sequence*, arXiv:2209.04657, 2022.

[15] B. Poonen and M. Rubinstein, The number of intersection points made by the diagonals of a regular polygon, *SIAM J. Discr. Math.*, **11.1** (1998) 135–156.

[16] J. Riordan and N. J. A. Sloane, Enumeration of rooted trees by total height, *J. Austral. Math. Soc.*, **10** (1969), 278–282.