Generating Primes

Eric Rowland

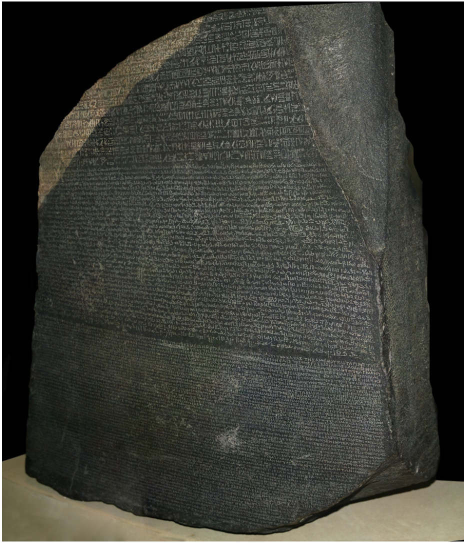
Mathematics Department
Tulane University, New Orleans, USA

↓

School of Computer Science
University of Waterloo, Waterloo, Canada

July 8, 2011

Main theme: Translation



Outline

1. Identifying and generating primes — a selective history
2. Interlude
3. A prime-generating recurrence

The sequence of primes

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, ...

Two questions:

- Is it easy to tell when a number is prime?
- Is it easy to generate primes?

(Not obviously.)

Primality testing

How to determine whether n is prime?

Trial division: Test divisibility by all numbers $2 \leq m \leq \sqrt{n}$.

Wilson's Theorem (Lagrange, 1773)

If $p \geq 2$, then p is prime if and only if p divides $(p-1)! + 1$.

Can the primality of a number be determined quickly?

In 2002, Agrawal, Kayal, & Saxena proved that "PRIMES is in P"!

If n has l digits, their algorithm determines whether n is prime in at most $c \cdot l^{12}$ steps.

Mersenne primes

A Mersenne prime is a prime of the form $2^k - 1$.
First few Mersenne primes:
$3 = 2^2 - 1, 7 = 2^3 - 1, 31 = 2^5 - 1, 127 = 2^7 - 1$.

Marin Mersenne (1588–1648)

If $2^k - 1$ is prime, then k must also be prime:

$$2^{ab} - 1 = (2^a - 1) \cdot \left(1 + 2^a + 2^{2a} + 2^{3a} + \cdots + 2^{(b-1)a}\right).$$

So each Mersenne prime is of the form $2^p - 1$.

GIMPS

Testing primality of $2^p - 1$ is (relatively) easy: Lucas–Lehmer test.

The Great Internet Mersenne Prime Search
is distributed computing project begun in 1996.
`http://mersenne.org`

All 47 known Mersenne primes:

$2^{2}-1, 2^{3}-1, 2^{5}-1, 2^{7}-1, 2^{13}-1, 2^{17}-1, 2^{19}-1, 2^{31}-1, 2^{61}-1, 2^{89}-1, 2^{107}-1, 2^{127}-1, 2^{521}-1,$
$2^{607}-1, 2^{1279}-1, 2^{2203}-1, 2^{2281}-1, 2^{3217}-1, 2^{4253}-1, 2^{4423}-1, 2^{9689}-1, 2^{9941}-1, 2^{11213}-1, 2^{19937}-1,$
$2^{21701}-1, 2^{23209}-1, 2^{44497}-1, 2^{86243}-1, 2^{110503}-1, 2^{132049}-1, 2^{216091}-1, 2^{756839}-1, 2^{859433}-1,$
$2^{1257787}-1, 2^{1398269}-1, 2^{2976221}-1, 2^{3021377}-1, 2^{6972593}-1, 2^{13466917}-1, 2^{20996011}-1, 2^{24036583}-1,$
$2^{25964951}-1, 2^{30402457}-1, 2^{32582657}-1, 2^{37156667}-1, 2^{42643801}-1, 2^{43112609}-1$

Largest known prime: $2^{43112609} - 1$.
It was discovered in August 2008 and has 12978189 decimal digits.

Sieve of Eratosthenes

Naive way to generate the sequence of primes:

~~1~~ 2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ~~20~~ ~~21~~ ~~22~~ 23 ~~24~~ ~~25~~ · · ·

Euler's polynomial

Several functions are known to generate primes.

In 1772, Euler observed that the polynomial $n^2 + n + 41$ is prime for $0 \leq n \leq 39$:

$$41, 43, 47, 53, 61, 71, \ldots, 1523, 1601.$$

But for $n = 40$ the value is $1681 = 41^2$.

Does there exist a polynomial $f(n)$ that only takes on prime values?

The constant polynomial $f(n) = 41$ does!

Prime-generating polynomials

What about a non-constant polynomial?

No. Suppose $f(n)$ is prime for all n; let $p = f(1)$.
Then $f(1 + pk) = f(1) + p \cdot \text{stuff}$, so p divides $f(1 + pk)$.

What about a multivariate polynomial?

Theorem (Jones–Sato–Wada–Wiens, 1976)

The set of positive values taken by the following degree-25 polynomial in 26 variables is equal to the set of prime numbers.

$$
\begin{aligned}
(k+2)\big(1 &- (wz+h+j-q)^2 \\
&- ((gk+2g+k+1)(h+j)+h-z)^2 \\
&- (2n+p+q+z-e)^2 \\
&- (16(k+1)^3(k+2)(n+1)^2+1-f^2)^2 \\
&- (e^3(e+2)(a+1)^2+1-o^2)^2 \\
&- ((a^2-1)y^2+1-x^2)^2 \\
&- (16r^2y^4(a^2-1)+1-u^2)^2 \\
&- (((a+u^2(u^2-a))^2-1)(n+4dy)^2+1-(x+cu)^2)^2 \\
&- (n+l+v-y)^2 \\
&- ((a^2-1)l^2+1-m^2)^2 \\
&- (ai+k+1-l-i)^2 \\
&- (p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m)^2 \\
&- (q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x)^2 \\
&- (z+pl(a-p)+t(2ap-p^2-1)-pm)^2\big)
\end{aligned}
$$

Corollary: If p is prime, then there is a proof that p is prime consisting of 87 additions and multiplications.

Prime-generating polynomials

This polynomial is an implementation of a primality test in the language of polynomials.

The first result of this kind was a degree-37 polynomial in 24 variables constructed by Yuri Matiyasevich in 1971.

Motivation was Hilbert's 10th problem:
Is there an algorithm to determine whether a polynomial equation has integer solutions?

Answer:
No. Any set of positive integers output by a computer program (running forever) can be encoded as the set of positive values of a polynomial.

A prime-generating double exponential

In 1947, William Mills proved the existence of a real number b such that $\lfloor b^{3^n} \rfloor$ is prime for $n \geq 1$.

Assuming the Riemann hypothesis, the smallest such b is

$$b = 1.30637788386308069046861449260260571 \cdots$$

and generates the primes

$$2, 11, 1361, 2521008887, 16022236204009818131831320183, \dots .$$

But the only known way of computing digits of b is by working backward from known large primes!

In 1964, C. P. Willans produced this formula for the nth prime:

$$p_n = 1 + \sum_{i=1}^{2^n} \left\lfloor \left(\frac{n}{\sum_{j=1}^{i} \left\lfloor \left(\cos \frac{(j-1)!+1}{j} \pi \right)^2 \right\rfloor} \right)^{1/n} \right\rfloor$$

But Willans' formula is built on Wilson's theorem!

$$\frac{(j-1)!+1}{j} = \begin{cases} \text{an integer} & \text{if } j = 1 \text{ or } j \text{ is prime} \\ \text{not an integer} & \text{if } j \geq 2 \text{ is not prime.} \end{cases}$$

$$\left\lfloor \left(\cos \frac{(j-1)!+1}{j} \pi \right)^2 \right\rfloor = \begin{cases} 1 & \text{if } j = 1 \text{ or } j \text{ is prime} \\ 0 & \text{if } j \geq 2 \text{ is not prime.} \end{cases}$$

$$\sum_{j=1}^{i} \left\lfloor \left(\cos \frac{(j-1)!+1}{j} \pi \right)^2 \right\rfloor = \pi(i) + 1.$$

$$\left\lfloor \left(\frac{n}{\pi(i)+1} \right)^{1/n} \right\rfloor = \begin{cases} 1 & \text{if } i < p_n \\ 0 & \text{if } i \geq p_n. \end{cases}$$

The cold, hard truth

In practice, none of those “generators” actually generate primes at all!

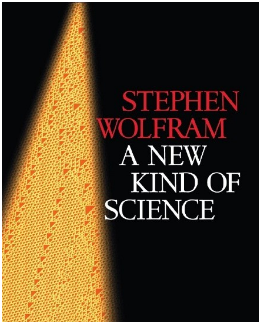
They are just *engineered*.

Are there “naturally occurring” functions that generate primes?

Outline

A New Kind of Science

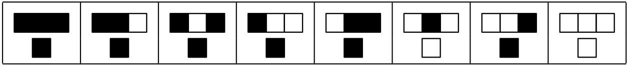
In 2002 Stephen Wolfram published *A New Kind of Science*.

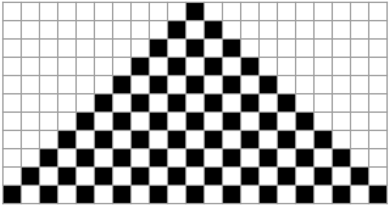


Simple programs are capable of complex behavior.

In particular, mathematics only considers a small subset of the possible programs that exist.

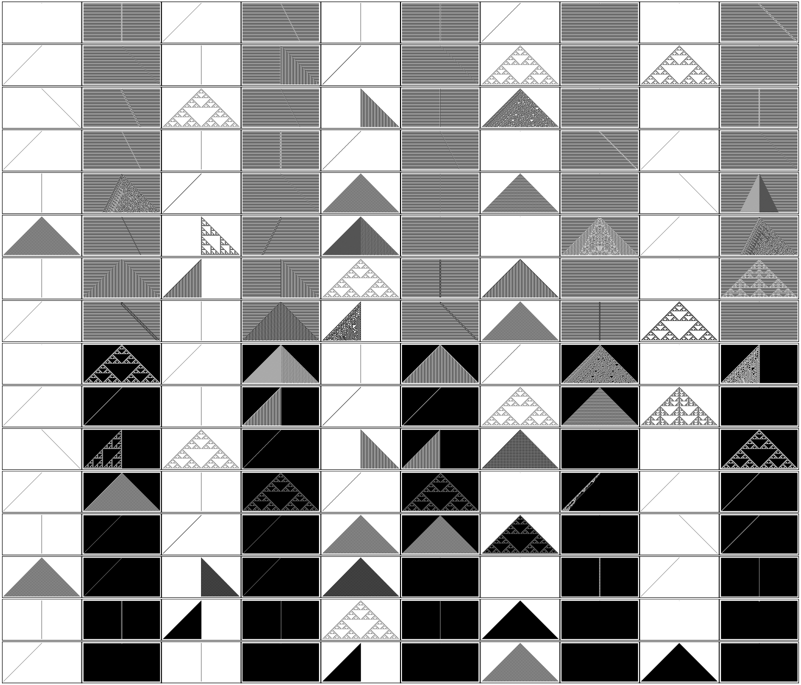
Cellular automata

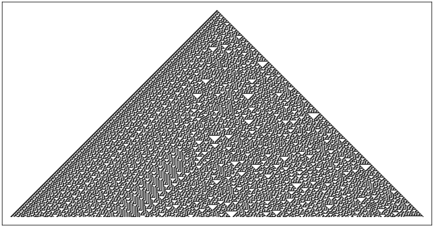


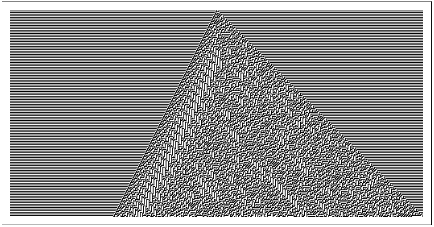


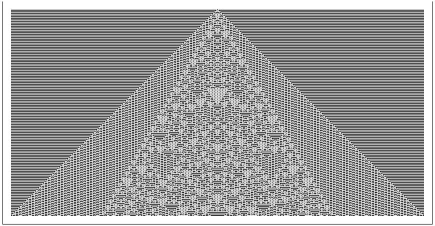
Certain cellular automata had been studied before.
For example, John Conway's "game of life".

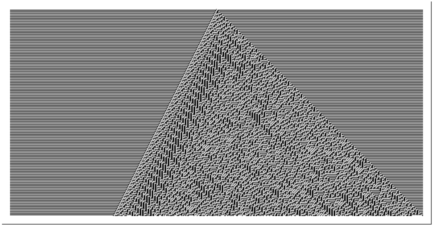
Wolfram's approach: Systematically look at all possible rules.

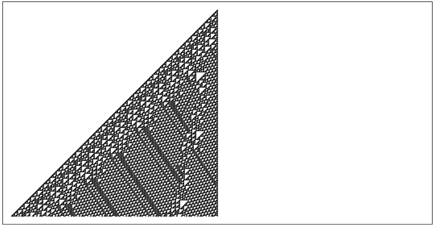


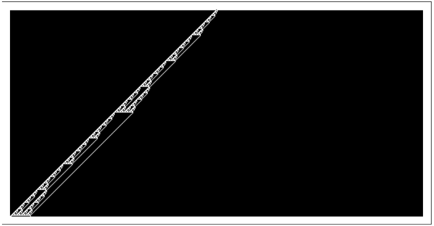






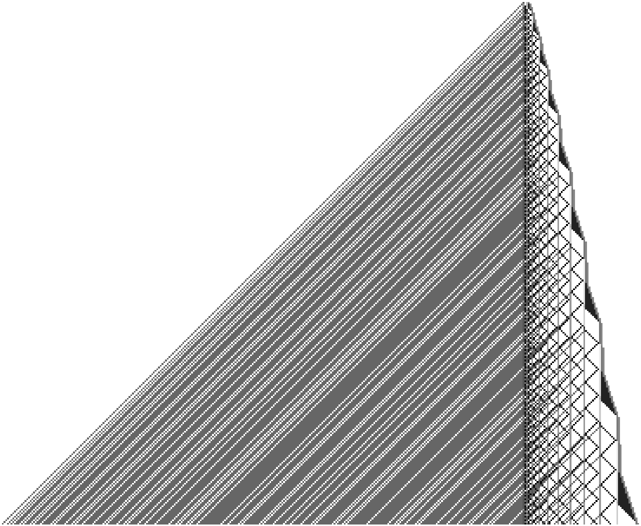






A prime-generating cellular automaton

A 16-color rule depending on 3 cells that computes the primes:



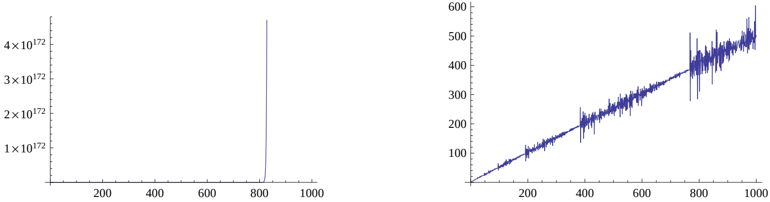
Outline

Recurrences

At the 2003 NKS Summer School, Matthew Frank decided to explore a different kind of system that evolves through time: integer recurrences.

Fibonacci recurrence:

$$a(n) = a(n-1) + a(n-2).$$



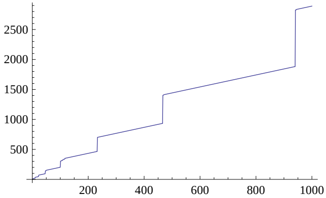
Hofstadter recurrence:

$$a(n) = a(n - a(n-1)) + a(n - a(n-2)).$$

A new recurrence

Frank systematically substituted several *Mathematica* functions into a template recurrence and looked at the pictures they generated.

One that caught his eye was this:



The recurrence was

$$a(n) = a(n-1) + \gcd(n, a(n-1))$$

with initial condition $a(1) = 7$.

First few terms

$$a(n) = a(n-1) + \gcd(n, a(n-1))$$

$a(1) = 7$
$a(2) = 7 + \gcd(2, 7) = 7 + 1 = 8$
$a(3) = 8 + \gcd(3, 8) = 8 + 1 = 9$
$a(4) = 9 + \gcd(4, 9) = 9 + 1 = 10$
$a(5) = 10 + \gcd(5, 10) = 10 + 5 = 15$
$a(6) = 15 + \gcd(6, 15) = 15 + 3 = 18$
$a(7) = 18 + \gcd(7, 18) = 18 + 1 = 19$
$a(8) = 19 + \gcd(8, 19) = 19 + 1 = 20$
$a(9) = 20 + \gcd(9, 20) = 20 + 1 = 21$
$a(10) = 21 + \gcd(10, 21) = 21 + 1 = 22$
$a(11) = 22 + \gcd(11, 22) = 22 + 11 = 33$
$a(12) = 33 + \gcd(12, 33) = 33 + 3 = 36$
$a(13) = 36 + \gcd(13, 36) = 36 + 1 = 37$
$a(14) = 37 + \gcd(14, 37) = 37 + 1 = 38$

Difference sequence $a(n) - a(n-1) = \gcd(n, a(n-1))$:

1, 1, 1, 5, 3, 1, 1, 1, 1, 11, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 23, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 47, 3, 1, 5, . . .

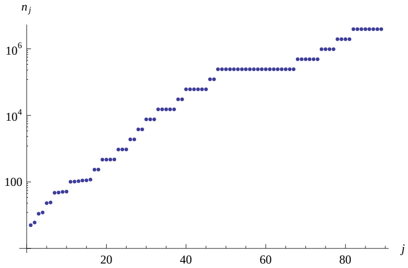
The sequence gcd(n, $a(n-1)$)

1, 1, 1, 5, 3, 1, 1, 1, 1, 11, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 23, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 47, 3, 1, 5, 3,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 101,
3, 1, 1, 7, 1, 1, 1, 1, 11, 3, 1, 1, 1, 1, 1, 13, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 233, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 467, 3, 1, 5, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, . . .

gcd(n, $a(n-1)$) appears to always be 1 or prime!

Key observations

A few years later I generated this plot:



logarithmic plot of n_j, the jth value of n for which $\gcd(n, a(n-1)) \neq 1$

Ratio between clusters is very nearly 2.

Each cluster is initiated by a large prime p.

Another key observation

n	$\gcd(n, a(n-1))$	$a(n)$
1		7
2	1	8
3	1	9
4	1	10
5	5	15
6	3	18
7	1	19
8	1	20
9	1	21
10	1	22
11	11	33
12	3	36
13	1	37
14	1	38
15	1	39
16	1	40
17	1	41
18	1	42
19	1	43
20	1	44

n	$\gcd(n, a(n-1))$	$a(n)$
21	1	45
22	1	46
23	23	69
24	3	72
25	1	73
26	1	74
27	1	75
28	1	76
29	1	77
30	1	78
31	1	79
32	1	80
33	1	81
34	1	82
35	1	83
36	1	84
37	1	85
38	1	86
39	1	87
40	1	88

n	$\gcd(n, a(n-1))$	$a(n)$
41	1	89
42	1	90
43	1	91
44	1	92
45	1	93
46	1	94
47	47	141
48	3	144
49	1	145
50	5	150
51	3	153
52	1	154
53	1	155
54	1	156
55	1	157
56	1	158
57	1	159
58	1	160
59	1	161
60	1	162

$a(n) = 3n$ whenever $\gcd(n, a(n-1)) \neq 1$.

Local structure

Lemma

Let $n_1 \geq 2$*. Let* $a(n_1) = 3n_1$*, and for* $n > n_1$ *let*

$$a(n) = a(n-1) + \gcd(n, a(n-1)).$$

Let n_2 *be the smallest integer greater than* n_1 *such that* $\gcd(n_2, a(n_2 - 1)) \neq 1$*. Then*

- $\gcd(n_2, a(n_2 - 1)) = p$ *is prime,*
- p *is the smallest prime divisor of* $2n_1 - 1$*,*
- $n_2 = n_1 + \frac{p-1}{2}$*, and*
- $a(n_2) = 3n_2$*.*

This lemma provides the inductive step.

Main result

Theorem (2008)

Let $a(1) = 7$*, and for* $n > 1$ *let*

$$a(n) = a(n-1) + \gcd(n, a(n-1)).$$

For each $n \geq 2$*,* $\gcd(n, a(n-1))$ *is either* 1 *or prime.*

Is the recurrence a “magical” producer of primes?

No.
Without the shortcut, $\frac{p-3}{2}$ consecutive 1s precede p.
With the shortcut, each step requires finding the smallest prime divisor of $2n - 1$.

Other initial conditions

Do all initial conditions produce only 1s and primes? No.
$a(1) = 532$ produces $\gcd(18, a(17)) = \gcd(18, 567) = 9$.
$a(1) = 801$ produces $\gcd(21, a(20)) = \gcd(21, 840) = 21$.

Conjecture

Let $n_1 \geq 1$ and $a(n_1) \geq 1$. For $n > n_1$ let

$$a(n) = a(n-1) + \gcd(n, a(n-1)).$$

Then there exists an N such that for each $n > N$ $\gcd(n, a(n-1))$ is either 1 or prime.

It would suffice to show that $a(n)/n$ always reaches 1, 2, or 3.

Nontrivial values of gcd(n, $a(n-1)$)

5, 3, 11, 3, 23, 3, 47, 3, 5, 3, 101, 3, 7, 11, 3, 13, 233, 3, 467, 3, 5, 3, 941, 3, 7, 1889, 3, 3779, 3, 7559, 3, 13, 15131, 3, 53, 3, 7, 30323, 3, 60647, 3, 5, 3, 101, 3, 121403, 3, 242807, 3, 5, 3, 19, 7, 5, 3, 47, 3, 37, 5, 3, 17, 3, 199, 53, 3, 29, 3, 486041, 3, 7, 421, 23, 3, 972533, 3, 577, 7, 1945649, 3, 163, 7, 3891467, 3, 5, 3, 127, 443, 3, 31, 7783541, 3, 7, 15567089, 3, 19, 29, 3, 5323, 7, 5, 3, 31139561, 3, 41, 3, 5, 3, 62279171, 3, 7, 83, 3, 19, 29, 3, 1103, 3, 5, 3, 13, 7, 124559609, 3, 107, 3, 911, 3, 249120239, 3, 11, 3, 7, 61, 37, 179, 3, 31, 19051, 7, 3793, 23, 3, 5, 3, 6257, 3, 43, 11, 3, 13, 5, 3, 739, 37, 5, 3, 498270791, 3, 19, 11, 3, 41, 3, 5, 3, 996541661, 3, 7, 37, 5, 3, 67, 1993083437, 3, 5, 3, 83, 3, 5, 3, 73, 157, 7, 5, 3, 13, 3986167223, 3, 7, 73, 5, 3, 7, 37, 7, 11, 3, 13, 17, 3, 19, 29, 3, 13, 23, 3, 5, 3, 11, 3, 7972334723, 3, 7, 463, 5, 3, 31, 7, 3797, 3, 5, 3, 15944673761, 3, 11, 3, 5, 3, 17, 3, 53, 3, 139, 607, 17, 3, 5, 3, 11, 3, 7, 113, 3, 11, 3, 5, 3, 293, 3, 5, 3, 53, 3, 5, 3, 151, 11, 3, 31889349053, 3, 63778698107, 3, 5, 3, 491, 3, 1063, 5, 3, 11, 3, 7, 13, 29, 3, 6899, 3, 13, 127557404753, 3, 41, 3, 373, 19, 11, 3, 43, 17, 3, 320839, 255115130849, 3, 510230261699, 3, 72047, 3, 53, 3, 17, 3, 67, 5, 3, 79, 157, 5, 3, 110069, 3, 7, 1020460705907, 3, 5, 3, 43, 179, 3, 557, 3, 167, . . .

Which primes appear?

$p = 2$ cannot occur.
But one suspects that all other primes do.
After ten thousand nontrivial gcds, the smallest odd prime that has not yet appeared is 587.

Theorem (Chamizo–Raboso–Ruiz-Cabello, 2011)

If $a(n) = a(n-1) + \gcd(n, a(n-1))$ *with* $a(1) = 7$*, then the difference sequence* $\gcd(n, a(n-1))$ *contains infinitely many distinct primes.*

Moreover, they obtained a simple characterization of the finite sequences of primes that appear for some initial condition.
For example, the sequence $17, 5, p$ does not occur for any prime $p > 3$.

It also follows that no sequence of primes occurs twice consecutively.

A variant

Benoit Cloitre looked at the recurrence

$$a(n) = a(n-1) + \mathrm{lcm}(n, a(n-1))$$

with $a(1) = 1$.

He observed that $\frac{a(n)}{a(n-1)} - 1$ seems to be 1 or prime for each $n \geq 2$:

2, 1, 2, 5, 1, 1, 1, 1, 5, 11, 1, 13, 1, 5, 1, 17, 1, 19, 1, 1, 11, 23, 1, 5, 13, 1, 1, 29, 1, 31, 1, 11, 17, 1, 1, 37, 1, 13, 1, 41, 1, 43, 1, 1, 23, 47, 1, 1, 1, 17, 13, 53, 1, 1, 1, 1, 29, 59, 1, 61, 1, 1, 1, 13, 1, 67, 1, 23, 1, 71, 1, 73, 1, 1, 1, 1, 13, 79, 1, 1, 41, 83, 1, 1, 43, 29, 1, 89, 1, 13, 23, 1, 47, 1, 1, 97, 1, 1, 1, 101, 1, 103, 1, 1, 53, 107, 1, 109, 1, 1, 1, 113, 1, 23, 29, 1, 59, 1, 1, 1, 61, 41, 1, 1, 1, 127, 1, 43, 1, 131, 1, 1, 67, 1, 1, 137, 1, 139, 1, 47, 71, 1, 1, 29, 73, 1, 1, 149, 1, 151, 1, 1, 1, 1, 1, 157, 1, 53, 1, 1, 1, 163, 1, 1, 83, ...

Conjecturally, every prime appears except 3 and 7.

No proof yet!