How To Guess a Generating Function

Simon Plouffe
CECM

Most of people knows the following problem,

here is a sequence: 1, 2, 4, 7, ...

What is the next term ?
What expression would give us the general term ?

It would be nice to have a formula to
generate all the terms in the sequence.

how do we find the formula?

here is the way to become a 'genius' at MENSA tests.

We took a source of integer sequences

A Handbook of Integer Sequences.
by
N.J.A. Sloane

Academic Press, 1973

4568 sequences. (1991)

for example the sequence #391 of HIS is

1,2,4,7,11,16,22,29,37,46,56,67,79,92...

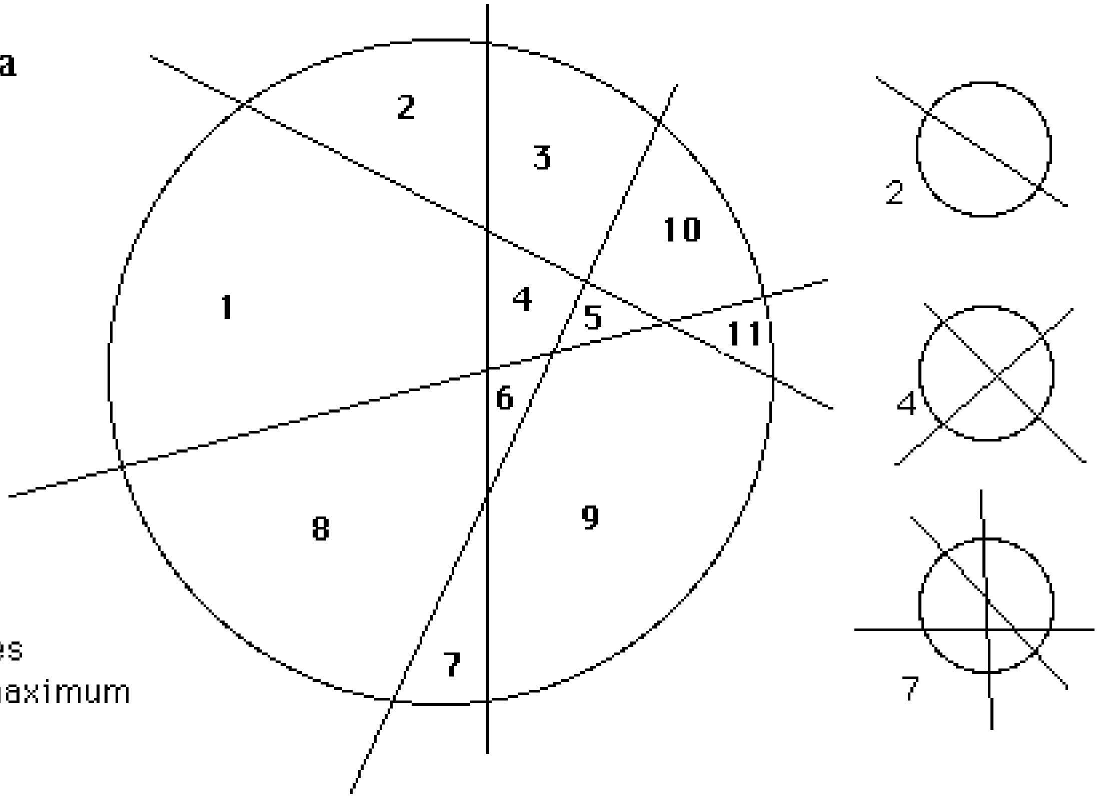
"slicing a cake with n slices"
or the "lazy caterer sequence"
references are : Mathematical Magazine vol. 30 page 150 (1946)
and Fibonacci Quartely vol. 30 page 296 (1961).

One may find by hand,

$$\downarrow$$

$$n^2 - \frac{1}{2}n + 1$$

slices of a cake...



with 4 slices
there is a maximum
of 11 parts.

a_0 , a_1 , a_2 , a_3 , ... , a_n

$a_i \in \mathbb{Z}$ i.e. the integers.

Let's look at power series representation of a sequence.

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_2 x^3 + a_2 x^4 \ldots + a_n x^n$$

There is always a way to put ,

$$f(x) = P(x)/Q(x)$$

A rational fraction.

$Q(x) \neq 0$ et $\deg(P)+\deg(Q) < n$

More generally speaking will are looking at
a general formula for the n'th term of a sequence.

A polynomial is ALWAYS a rational fraction.
The converse is not true.

Other methods of attack are leading to the same result.

i.e. Finite Differences,
Montmort Formula,
A base of the polynomials 1, n , n^2, n^3, …

This gives back the terms in the sequence once developped
into Taylor series.

```
                 2
       1 - x + x
f := - ----------
              3
       (1 - x)
```

```
> (1-x+x**2)/(1-x)**3;
```

$$\frac{1-x+x^2}{(1-x)^3}$$

```
> series(",x,12);
```

$$1 + 2\,x + 4\,x^2 + 7\,x^3 + 11\,x^4 + 16\,x^5 + 22\,x^6 + 29\,x^7 + 37\,x^8 + 46\,x^9 + 56\,x^{10} + 67\,x^{11} + O(x^{12})$$

A simple example.

$$\frac{1}{1 - x - x^2}$$

By developping this fraction
(long division) we generate Fibonacci Numbers.

$$1 + x + 2 x^2 + 3 x^3 + 5 x^4 + 8 x^5 + 13 x^6 + 21 x^7 + 34 x^8 + 55 x^9 + 89 x^{10} + \ldots$$

It is not a polynomial.
The n'th term is given by Binet's formula.

$$\frac{1}{\sqrt{5}} (\phi_1^n + \phi_2^n)$$

This gives us also information on the behaviour of F(n) when n gets large.

Puzzle #1

try to find a closed expression of f(n) if
the rational polynomial is

$$\frac{1}{1 - 6 x + 25 x^2}$$

Puzzle #2

here is a exponential g.f.
guess what is f(n) ?

$$- \frac{2}{5} \frac{\sinh(1/2\ 5^{1/2}\ x)\ 5^{1/2}}{\exp(x)^{1/2}}$$

answer f(n) are the Fibonacci Numbers !

To go from
sequence ------> rational fraction,
is mechanical…
We put the series into a system of linear equations.

$$a_0 + a_1X + a_2X\hat{}2 + a_2X\hat{}3 + a_2X\hat{}4 \ldots + a_nX\hat{}n = P(X)/Q(X)$$

We simply solve that system, by using the method of Undetermined Coefficients.

There are a FINITE number of steps to do,
in principle this problem is SOLVED, so to speak.

Well (if we are lazy) we use a program for doing such.

The 'routine' CONVERT/RATPOLY of MapleV
does it so well.
(a jewell of program : 800 lines of Maple) : 2 Ph. D. thesis in it.

Why not using it on all the sequences in the table ??!!

The first experimental result found

$$\frac{1 + x + 2x^2 + 4x^3}{(1 + x)(1 + x^2)(x - 1)^2}$$

Hurwitz-Radon function at powers of 2.
Theory of binary quadratic forms, T.Y. Lam p.131

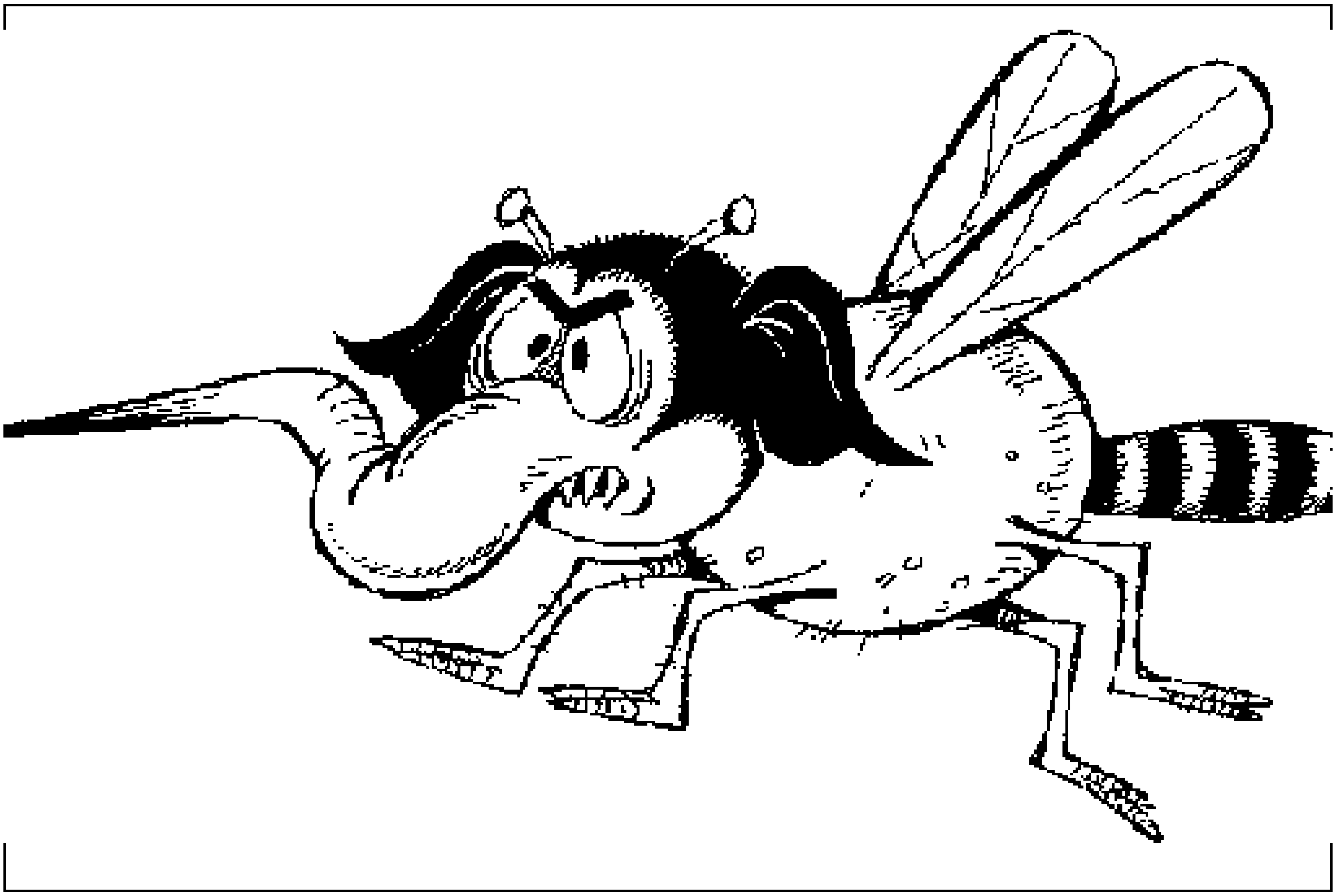
For example if we take the sequence of the first prime numbers…

$$2 + 3x + 5x^2 + 7x^3 + 11x^4 + 13x^5 + 17x^6 + 19x^7 + 23x^8 + 29x^9 + 31x^{10} + 37x^{11} + 41x^{12} + 43x^{13} + 47x^{14} + 53x^{15} + 59x^{16} + 61x^{17} + 67x^{18} + 71x^{19}$$

The program returns this formula (Yuk !)

$$\left(2 + \frac{19979}{4767}x + \frac{22105}{4767}x^2 + \frac{58889}{9534}x^3 + \frac{12844}{1589}x^4 + \frac{106381}{9534}x^5 + \frac{41033}{4767}x^6 + \frac{51467}{9534}x^7 + \frac{4261}{4767}x^8 + \frac{12856}{1589}x^9 + \frac{10337}{4767}x^{10}\right) \Big/ \left(1 + \frac{2839}{4767}x - \frac{10247}{9534}x^2 - \frac{2749}{9534}x^3 - \frac{4037}{9534}x^4 + \frac{8783}{9534}x^5 - \frac{14015}{9534}x^6 - \frac{6061}{3178}x^7 - \frac{289}{4767}x^8 + \frac{4842}{1589}x^9\right)$$

The expression with 48 terms is a monster…



BUT ,
how to decide if the rational fraction found is a good one?

The DEGREE of the expression found tells
if we are in the presence of a good candidate.

P(X) : initial conditions
Q(X) : recurrence relation itself.

The LENGTH of the expression is a
simple criteria.

So, why is it good ?

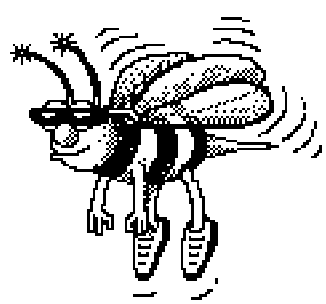
The algorithm of Geddes/Padé/Choi/Cabay
is one of the best known.

Results are reduced (in size) to a maximum.

The C.A.S. always returns an answer,
even a monster, so we can always decide.

Some examples ------------->

# of hits : 614



a bug

IF,
from a given sequence we can decide wheter or not we have
a simple P(x)/Q(x),

THEN why not use it to find,

ln(sequence),
exp(sequence),
derivative(sequence),
logarithmic derivative(sequence),
inverse(sequence)
and ***then***
try IF P(x)/Q(x) ? *

*F.B.

By doing such , we would be able to find formulas like,

$$\ln\left(\frac{P(x)}{Q(x)}\right) \qquad e^{\left(\frac{P(x)}{Q(x)}\right)} \qquad e^{\int\left(\frac{P(x)}{Q(x)}\right)}$$

$$y(x) = \left(\frac{P(x)}{Q(x)}\right)^{<-1>}$$

…and combinations of these.

It is not necessary to combine an arbitrary number of these…
only 3 are enough for most of the simple cases.
(experimental evidence).
The first 2, are equivalent to F.O.D.E.

That is ,

the derivative,
the logarithmic derivative
and
the functional inverse.

guessgf( ) of GFUN

The functional INVERSE an example, with Motzkin Numbers.

```
> read A1006;
```

[1, 1, 2, 4, 9, 21, 51, 127, 323, 835, 2188, 5798, 15511, 41835, 113634, 310572, 853467, 2356779, 6536382, 18199284, 50852019, 142547559, 400763223, 1129760415]

```
> listtoseries(",x,revogf);
```

$$x - x^2 + x^4 - x^5 + x^7 - x^8 + x^{10} - x^{11} + x^{13} - x^{14} + x^{16} - x^{17} + x^{19} - x^{20} + x^{22} - x^{23} + O(x^{25})$$

```
> convert(",ratpoly);
```

$$\frac{x}{1 + x + x^2}$$

```
> "=y;
```

$$\frac{x}{1 + x + x^2} = y$$

```
> solve(",x);
bytes used=2001524, alloc=1310480, time=2.80
                                 2 1/2                               2 1/2
         - 1 + y + (1 - 2 y - 3 y )             - 1 + y - (1 - 2 y - 3 y )
  - 1/2 ----------------------------, - 1/2 ----------------------------
                    y                                    y

> op(1,["]);

                                                2 1/2
                         - 1 + y + (1 - 2 y - 3 y )
                 - 1/2 ----------------------------
                                    y

> series(",y,35);
```

$$y + y^{2} + 2y^{3} + 4y^{4} + 9y^{5} + 21y^{6} + 51y^{7} + 127y^{8} + 323y^{9} + 835y^{10} + 2188y^{11} + 5798y^{12} + 15511y^{13} + 41835y^{14} + 113634y^{15} + 310572y^{16} + 853467y^{17} + 2356779y^{18} + 6536382y^{19} + 18199284y^{20} + 50852019y^{21} + 142547559y^{22} + 400763223y^{23} + 1129760415y^{24} + 3192727797y^{25} + 9043402501y^{26} + 25669818476y^{27} + 73007772802y^{28} + 208023278209y^{29} + 593742784829y^{30} + 1697385471211y^{31} + 4859761676391y^{32}$$

21,483,6468,66066,570570,4390386,31039008,205633428,1293938646,
7808250450,45510945480...

Rooted genus-2 maps with n edges.
Réf : JCT A13 page 215 1972.

$$21 \frac{1 + x}{(1 - 4x)^{11/2}}$$

In this case,
we could *extend* the sequence,
correct errors from the original paper
and verify typing errors.

Some nice formulas,...

$$\frac{1}{(1 - 6x + x^2)^{1/2}}$$

$$\frac{1 + x}{(1 - 2x)^{5/2}}$$

Bessel polynimials

$$\frac{4x^2 + 10x + 1}{(1 - 4x)^{7/2}} \qquad \frac{1 + 2x}{(1 - 4x)^{5/2}} \qquad \frac{1}{(1 - 4x)^{3/2}}$$

Apéry's Numbers

In some cases,
we have found families of formulas
, in order to deduce a general formula...

example : Delauney Numbers,
ref. Comtet, page 81.

$$\frac{(1 + x)^2}{(x - 1)^3} \qquad \frac{(1 + x)^3}{(x - 1)^4} \qquad \frac{(1 + x)^5}{(x - 1)^6} \qquad \frac{(1 + x)^6}{(x - 1)^7}$$

$$\frac{-1-6x^2+4x^3}{(-1+2x^2)(1-5x+2x^2)}$$

1,5,31,141,659,3005,13739,62669,285931,1304285

Worst case of a Jacobi symbol algorithm.

$$\frac{\exp(-1/4\,x\,(x+2))}{(1-x)^{1/2}}$$

Clouds with n points.

Comtet page 277.

$$\frac{1}{(x-1)^2} - \frac{2}{x-1} - \ln(x-1)$$

Associated Stirling numbers.

$$\frac{(1+3x)}{(1-2x)^{**}(7/2)}$$

$$\frac{(2x^3+57x^2+64x+6)}{(1-2x)^{11/2}}$$

$$\frac{\exp\left(1/4\,\frac{x\,(x+1)\,(x-2)}{x-1}\right)}{(x-1)^{1/2}}$$

Stochastic matrices of integers.

The nicest…

Extreme points of set of n x n symmetric doubly-substochastic matrices.

European Journal of Combinatorics
vol. 1 page 180 1980.
Richard P. Stanley

$$\frac{(x+1)^{1/4}\exp\left(\frac{1}{2}\,\frac{x\,(x^3 - x + 2x^2 - 3)}{(x-1)(x+1)}\right)}{(x-1)^{1/4}}$$

The wonderful Demlo Numbers, seq. #2339.
1,121,12321,1234321,123454321,12345654321,1234567654321,
123456787654321,12345678987654321,1234567900987654321,…
are generated by the following trick.

1*1
11*11
111*111
1111*1111
11111*11111
111111*111111
1111111*1111111
…
see HIS for references.

Are in fact simply generated by

$$\frac{1 + 10x}{(1 - x)(1 - 10x)(1 - 100x)}$$

Finally , we had 1031 formulas ,

out of the 4568 sequences in the table.

22% success !

Then ?, why not apply this methodology
to sequences of polynomials !
For example with the Chebyshev polynomials…

$$1$$

$$x$$

$$2x^2 - 1$$

$$4x^3 - 3x$$

$$8x^4 - 8x^2 + 1$$

$$16x^5 - 20x^3 + 5x$$

$$32x^6 - 48x^4 + 18x^2 - 1$$

$$64x^7 - 112x^5 + 56x^3 - 7x$$

nice christmas tree

After a few seconds of hesitation,
the program finds,

$$\frac{t\,x - 1}{2\,t\,x - 1 - x^{2}}$$

By developping this into a series in regard to x,
we obtain the diagonals of the tree.

We can recuperate most of the orthogonal polynomials in the A. & S. book (the bible of numerical analysis).

an example...

$$\frac{1}{\exp(x)} + \frac{(1 - 2\exp(x) + 2\exp(x)^{2})^{1/2}}{\exp(x)}$$

$$1 - (3 - 2\exp(x))^{1/2}$$

Planted binary phylogenetic trees with n labels.
Lecture Notes in Mathematics #884 page 196 1981.

We could obtain a large class of expressions.

$$e^{t(e^x - 1)}$$

$$e^{\left(\frac{1-\sqrt{1-2xt}}{x} - t\right)}$$

$$\frac{1}{1 - xe^{A(x)}}$$

A(x) is the solution of the
functional equation , A(x) = x exp(A(x))

What is happening here is the application
of a model to a sequence.

Most of the formulas (80%)
are from the D-finite (or P-recurrent)
world : R.P. Stanley , EJC 1980 vol. 1 page 80.

$$a_0 P_0(n) = a_1 P_1(n) + a_2 P_2(n) + a_3 P_3(n) \; \ldots \; a_k P_k(n)$$

D-finite is equivalent to P-recurrent.
We “know” that.
We “can” go from one to the other by
using GFUN --> listtorec, listtodiffeq, rectodiffeq,diffectorec.

All algebraic generating functions are D-finite !

Comtet, 1964 describes a procedure to go
from algebraic g.f. ---> diff. eq. (P-recurrence).

The CONVERSE is <u>quite</u> difficult to do in general… (?).

there is a pseudo-algorithm with LLL (S.P.)

To detect if a sequence is D-finite (?),

we may use a method based on Und. Coefficients,
GFUN does it, for most of the simple cases.

The method is called : brute force.

Not so intelligent : but it works.

Just plug the numbers and wait.

In somes cases we can go
from Diff. eq. ---> explicit solution (Ei(x), W(x), exp(x), Bessel…)
using GFUN/Maple (with Bruno Salvy around).

P-recurrences ---> Hypergeometrics

Well, are there other models ?

What about all those sequences related to partitions ?
all these infinite products…

Euler found this trick (apparently) : Andrews book on partitions

Let's suppose that a(n) is given by

$$\prod_{n \geq 1} \frac{1}{(1-x^n)^{c(n)}}$$

Given c(n) , it is easy to find a(n),
just by expanding the product.

But, if we have a(n) and we are looking for c(n),

then , well : we know that for a(n) = ord. partitions
c(n) = 1,1,1,1,1….

There are several ways to find c(n) , we could expand and 'collect' the coefficients of the same degree… (by hand).

or use the logarithmic derivative.
and the Mobius function,
log(product) --> sum of terms of same degree : collect the terms.
each degree can be expressed simply with Mobius function
on all n<= degree. It works.

IT IS a good model : # of hits : 125

Raymond strings, Theta series (some of them).
sequences related to lattices of all kind…

some surprises also.

Beatty Sequences.

If a and b are irrational then

if 1/a + 1/b = 1 then

the sequences :

[a], [2*a], [3*a], [4*a], …

and

[b], [2*b], [3*b], [4*b], …

COVERS N with no overlap and holes.

Let's suppose that the sequence is given simply by

[x*n]

[ ] = floor function
x irrational.
or maybe a polynomial in n ?

How can we detect that ?

Simple !

By using Least Squares there is a way o decide if the sequence is a Beatty sequence or not.
If YES then x can be calculated with usually 4 digits.
# of hits : 45 (errors corrected also).

In average finally we had,

around 25 % of success.

Bosonic string states

CU86.

A5308

Euler

Produit infini

$$\prod_{n\geq 1} \frac{1}{(1 - Z^{n})^{c(n)}}$$

c(n) = 0,0,0,1,1,2,2,3,3,4,4,...

1, 0, 0, 0, 1, 1, 2, 2, 4, 4, 7, 8, 14, 16, 25, 31

$$n!\frac{\left[\frac{[n/2]!\,2^{[n/2]}}{\sqrt{e}} + \frac{1}{2}\right]}{2^{[n/2]}[n/2]!}$$

is the EXACT expression of

the n'th term of

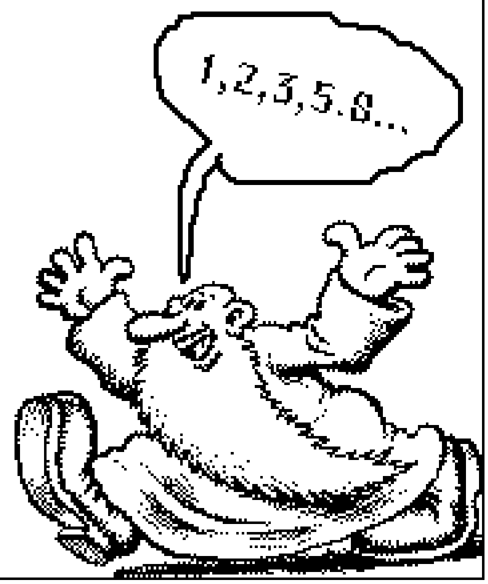
$$\frac{1}{\exp(x**2/2)(1-x)}$$

Classical asymptotic analysis gives (n-1)!/sqrt(e) **ONLY**.

sequences are from

The Encyclopedia of Integer Sequences
N.J.A. Sloane
and Simon Plouffe

24068 sequences,
+10000 references
with formulas.



or the SEQUENCE SERVER,

sequences@research.att.com
superseeker@research.att.com

gfun program in the share library of Maple.

REFERENCES

-Comtet, Louis , Advance Combinatorics, Reidel 1974.
-Routine convert/ratpoly of Maple

THE SEQUENCE SERVER

sequence@research.att.com
superseeker@research.att.com

lookup 1 2 5 14 42
from 10 to 100 requests / day