Notes on A299041

Peter Bala, Feb 04 2018

NAME: Irregular table: $T(n,k)$ equals the number of alignments of length k of n strings each of length 3.

An alignment of n strings of various lengths is a way of inserting blank characters into the n strings so that the resulting strings all have the same length. This common length is called the length of the alignment. Insertion of a blank character into the same position in each of the n strings is not allowed. By writing the strings one under another we can consider an alignment of length L of n strings as an $n \times L$ matrix.

Example. Listed below are the 12 alignments of length 4 of two strings ABC and DEF.

$$\begin{matrix} A & B & C & - \\ - & D & E & F \end{matrix} \quad \begin{matrix} A & B & - & C \\ - & D & E & F \end{matrix} \quad \begin{matrix} A & - & B & C \\ - & D & E & F \end{matrix} \quad \begin{matrix} - & A & B & C \\ D & - & E & F \end{matrix}$$

$$\begin{matrix} A & B & C & - \\ D & - & E & F \end{matrix} \quad \begin{matrix} A & B & - & C \\ D & - & E & F \end{matrix} \quad \begin{matrix} A & - & B & C \\ D & E & - & F \end{matrix} \quad \begin{matrix} - & A & B & C \\ D & E & - & F \end{matrix}$$

$$\begin{matrix} A & B & C & - \\ D & E & - & F \end{matrix} \quad \begin{matrix} A & B & - & C \\ D & E & F & - \end{matrix} \quad \begin{matrix} A & - & B & C \\ D & E & F & - \end{matrix} \quad \begin{matrix} - & A & B & C \\ D & E & F & - \end{matrix}$$

Sequence comparison and alignment is a central tool in computational molecular biology and computational linguistics. For example, in molecular biology a DNA sequence may be considered as a mathematical string $x = (x_1, x_2, \ldots, x_n)$, where $x_i \in \{A, C, G, T\}$, $i = 1, 2, \ldots, n$, is one of the four nucleotides, adenine, cytosine, guanine and thymine. A typical problem is to align and compare the sequence x with another DNA sequence $y = (y_1, y_2, \ldots, y_m)$, to measure the similarity between both sequences. Regions of similarity may indicate functional, structural and/or evolutionary relationships between the two sequences. For the number of alignments of length k of n strings of length 1 (resp. length 2) see A131689 (resp. A122193). Here we are considering the number of alignments of length k of n strings where each string has length 3. The unique shortest possible alignment of n such strings has length 3 and the longest possible alignments of n strings have length $3n$.

1) Explicit formula for table entries.

$$T(n,k) = \sum_{i=3}^{k} (-1)^{k-i} \binom{k}{i} \binom{i}{3}^n \qquad (1)$$

Proof. Slowinski [Sl'98] proved the following general result on the enumeration of alignments of strings: the number of alignments of length L of n strings $s_1, ..., s_n$ of lengths $l_1, ..., l_n$ is given by the formula

$$\sum_{i=0}^{L}(-1)^i\binom{L}{i}\prod_{j=1}^{n}\binom{L-i}{L-l_j-i}$$

or equivalently,

$$\sum_{i=0}^{L}(-1)^{L-i}\binom{L}{i}\prod_{j=1}^{n}\binom{i}{l_j}.$$

Applying Slowinski's result in the present case where $l_1 = l_2 = \cdots = l_n = 3$ produces (1). □

2) A series expansion for the row polynomials.

Let $R_n(x)$ denote the nth row polynomial of A299041.

$$R_n(x) \quad = \quad \sum_{i=3}^{\infty}\binom{i}{3}^n\frac{x^i}{(1+x)^{i+1}}, \qquad [n \geq 1] \qquad (2)$$

Proof. Using the binomial expansion

$$\frac{1}{(1-x)^{i+1}} \quad = \quad \sum_j\binom{i+j}{i}x^j$$

we obtain

$$\sum_{i=3}^{\infty}\binom{i}{3}^n\frac{x^i}{(1+x)^{i+1}} \quad = \quad \sum_{i\geq 3}\sum_{j\geq 0}(-1)^j\binom{i+j}{i}\binom{i}{3}^n x^{i+j}.$$

The coefficient of x^k on the right-hand side is

$$\begin{aligned}\sum_{i+j=k}(-1)^j\binom{i+j}{i}\binom{i}{3}^n \quad &= \quad \sum_{i=3}^{k}(-1)^{k-i}\binom{k}{i}\binom{i}{3}^n \\ &= \quad T(n,k).\end{aligned}$$

Hence

$$\begin{aligned}\sum_{i=3}^{\infty}\binom{i}{3}^n\frac{x^i}{(1+x)^{i+1}} \quad &= \quad \sum_k T(n,k)x^k \\ &= \quad R(n,x). \ \square\end{aligned}$$

3) Recurrence equation for the row polynomials.

$$R_{n+1}(x) = \tfrac{1}{3!}x^3\tfrac{d^3}{dx^3}\left((1+x)^3R_n(x)\right) \tag{3}$$

Proof. By (2) we see that

$$(1+x)^3R_n(x) = \sum_{i=3}^{\infty}\binom{i}{3}^n\frac{x^i}{(1+x)^{i-2}} \quad [n\geq 1]. \tag{4}$$

Now one easily checks that

$$\frac{1}{3!}x^3\frac{d^3}{dx^3}\left(\frac{x^i}{(1+x)^{i-2}}\right) = \frac{\binom{i}{3}x^i}{(1+x)^{i+1}}. \tag{5}$$

Hence applying the operator $\frac{1}{3!}x^3\frac{d^3}{dx^3}$ to (4) we obtain (for $n\geq 1$)

$$\begin{aligned}\frac{1}{3!}x^3\frac{d^3}{dx^3}\left((1+x)^3R_n(x)\right) &= \sum_{i=3}^{\infty}\binom{i}{3}^{n+1}\frac{x^i}{(1+x)^{i+1}}\\ &= R_{n+1}(x)\end{aligned}$$

by (2). The result also holds for $n=0$ if we take $R_0(x)=1$. □

4) Recurrence equation for table entries.

$$T(n+1,k) = \binom{k}{3}\left(T(n,k)+3T(n,k-1)+3T(n,k-2)+T(n,k-3)\right) \tag{6}$$

Proof. Follows easily by equating the coefficient of the x^k term on either side of (3). □

Remark. It is an easy consequence of (3) that for $n\geq 1$, the row polynomials $R_n(x)$ have the form

$$R_n(x) = x^3+\cdots+T(n,3n)x^{3n}.$$

Therefore the boundary conditions for the recurrence (6) are $T(0,0)=1$ and for $n\geq 1$, $T(n,3)=1$ while $T(n,k)=0$ if $(k<3)$ or $(k>3n)$.

5) Double exponential generating function.

For comparison purposes we also include the double exponential generating functions for A131689 and A122193.

A131689

$$\exp(-x)\sum_{i=0}^{\infty}\exp\left(\binom{i}{1}y\right)\frac{x^i}{i!} = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\left(A131689(n,k)\frac{x^k}{k!}\right)\frac{y^n}{n!} \tag{7}$$

A122193

$$\exp(-x)\sum_{i=0}^{\infty}\exp\left(\binom{i}{2}y\right)\frac{x^i}{i!} = \sum_{n=0}^{\infty}\sum_{k=0}^{2n}\left(A122193(n,k)\frac{x^k}{k!}\right)\frac{y^n}{n!} \tag{8}$$

A299041

$$\exp(-x)\sum_{i=0}^{\infty}\exp\left(\binom{i}{3}y\right)\frac{x^i}{i!} = \sum_{n=0}^{\infty}\sum_{k=0}^{3n}\left(T(n,k)\frac{x^k}{k!}\right)\frac{y^n}{n!} \tag{9}$$

Proof of (9).

The expansion of the left side of (9) is

$$\begin{aligned}\exp(-x)\sum_{i=0}^{\infty}\exp\left(\binom{i}{3}y\right)\frac{x^i}{i!} &= \sum_{j=0}^{\infty}(-1)^j\frac{x^j}{j!}\sum_{i=0}^{\infty}\left(\sum_{n=0}^{\infty}\binom{i}{3}^n\frac{y^n}{n!}\right)\frac{x^i}{i!}\\ &= \sum_{i,j,n\geq 0}(-1)^j\binom{i}{3}^n\frac{x^{i+j}}{j!i!}\frac{y^n}{n!}.\end{aligned}$$

The coefficient of $\frac{x^k}{k!}\frac{y^n}{n!}$ in the expression on the right-hand side equals $\sum_i(-1)^{k-i}\binom{i}{3}^n\binom{k}{i}$, which equals $T(n,k)$ by (1). □

The expansion of the double e.g.f. for A299041 begins

$$\begin{aligned}\exp(-x)\sum_{n=0}^{\infty}\exp\left(\binom{n}{3}y\right)\frac{x^n}{n!} &= 1+\left(\frac{x^3}{3!}\right)\frac{y}{1!}+\left(\frac{x^3}{3!}+\frac{12x^4}{4!}+\frac{30x^5}{5!}+\frac{20x^6}{6!}\right)\frac{y^2}{2!}\\ &+(\frac{x^3}{3!}+\frac{60x^4}{4!}+\frac{690x^5}{5!}+\frac{2940x^6}{6!}+\frac{5670x^7}{7!}\\ &+\frac{5040x^8}{8!}+\frac{1680x^9}{9!})\frac{y^3}{3!}+\cdots.\end{aligned}$$

Exercise. Show the double exponential generating function $A(x,y) = \exp(-x)\sum_{i=0}^{\infty}\exp\left(\binom{i}{3}y\right)\frac{x^i}{i!}$ satisfies the partial differential equation

$$\frac{\partial A}{\partial y} = \frac{x^3}{3!}\left(A + 3\frac{\partial A}{\partial x} + 3\frac{\partial^2 A}{\partial x^2} + \frac{\partial^3 A}{\partial x^3}\right)$$

6) Row polynomials as a black diamond product.

$$R_n(x) = x^3 \blacklozenge \cdots \blacklozenge x^3 \ (n \text{ factors}) \tag{10}$$

Dukes and White [DuWh'16], in their study of the combinatorics of web diagrams and web matrices, introduced a commutative and associative $\mathbb{C}$-bilinear product of power series, which they named the black diamond product and denoted by the symbol $\blacklozenge$. The black diamond product of monomial polynomials is given by the formula

$$x^m \blacklozenge x^n = \sum_{k=0}^{m}\binom{n+k}{k}\binom{n}{m-k}x^{n+k}. \tag{11}$$

The stated expression for the row polynomial $R_n(x)$ as a black diamond product may be easily proved by simple induction argument, making use of the following particular case of (11):

$$x^3 \blacklozenge x^n = \binom{n}{3}x^n + 3\binom{n+1}{3}x^{n+1} + 3\binom{n+2}{3}x^{n+2} + \binom{n+3}{3}x^{n+3}.$$

REFERENCES

[Ba'18] P. Bala, Notes on A122193, uploaded to A122193

[DuWh'16] M. Dukes and C. D. White, Web matrices: structural properties and generating combinatorial identities, Electronic Journal Of Combinatorics, 23(1) (2016), #P1.45.

[Sl'98] J. B. Slowinski, The Number of Multiple Alignments, Molecular Phylogenetics and Evolution 10:2 (1998), 264-266, doi: 10.1006/mpev.1998.0522