Enumerating Membrane Structures

Vincenzo Manca

Verona University, Computer Science Department,
Strada LeGrazie 15, 37134 Verona, Italy
vincenzo.manca@univr.it

Abstract. A recurrent formula for enumerating membrane structures is given. It is deduced by elementary combinatorial methods, by providing a simplification with respect to a classical formula for the enumeration of trees, which is based on the analytical method of generating functions.

1 Introduction

The computation of the number of different membrane structures constituted by n membranes was considered at early beginning of Membrane Computing [6], in a preliminary draft of [7]. It is a well known combinatorial problem equivalent to the enumeration of unlabeled unordered trees [4]. Therefore, it is related to Catalan numbers and to a lot of combinatorial problems [2] which recently were proved to be investigated even by Greek mathematicians (*e. g.*, Hypparcus' problem and its modern variant known as Schröder's problem [8]).

For the enumeration of (this kind of) trees, no exact analytical formula is available, but a recurrent formula, based on integer partitions, was given in [5], which was deduced by means of generating functions. In the same paper also a complex asymptotic formula is presented.

In this note, we provide a new recurrent formula related to a simple combinatorial analysis of membrane structures.

Finite cumulative multisets are an extension of finite sets, such as $[a, a, b, c]$, or $[a, a, b], [a, a, b], [b]]$, where elements are put inside bracket pairs ("membranes"), and any element can occur more than one time. In a membrane structure all elements are built by starting from the empty multiset $[\,]$ ("elementary membrane"). The most external pair of brackets is called "skin", and the elements inside the skin are called components of the structure.

The set $\mathbb{M}$ of finite membrane structures can be defined by induction by means of the multiset sum $+$ (summing the element occurrences of two multisets), which is a binary commutative and associative operation, and by means of the multiset singleton operation which, given a multiset S, provides the multiset $[S]$ having S as its unique element.

$$
\begin{array}{ll}
[\,] \in \mathbb{M} & \textit{Base step} \\
S,\ S_1,\ S_2 \in \mathbb{M} \implies [S],\ S_1 + S_2 \in \mathbb{M} & \textit{Inductive step}
\end{array}
$$

For example,
$[[\,]] \in \mathbb{M}$ and $[[\,]] + [[\,]] = [[\,], [\,]] \in \mathbb{M}$.

D. Corne et al. (Eds.): WMC9 2008, LNCS 5391, pp. 292–298, 2009.

2 Proto-Agglomerates, Neo-Agglomerates and Conglomerates

Given n elements, the number of multisets of k elements over the n elements, according to [1], is given by:

$$\binom{n+k-1}{k} \tag{1}$$

By using formula (1), the following recurrent formula is given in Knuth's book [4] (pag. 386), which provides the number $T(n)$ unlabeled unordered trees of n nodes (membrane structures with n membranes), where $\mathbb{N}$ is the set of natural numbers, $n > 0 \in \mathbb{N}$, $T(0) = 1$, and $j, n_1, n_2, \ldots, n_j, k_1, k_2, \ldots, k_j \in \mathbb{N}$:

$$T(n+1) = \sum_{k_1 \cdot n_1 + k_2 \cdot n_2, \ldots k_j \cdot n_j = n} \quad \prod_{i=1,\ldots,\, j} \binom{T(n_i) + k_i - 1}{k_i} \tag{2}$$

Unfortunately, formula (2) is not manageable for an effective computation, because it is based on integer partitions, which grow, according to Ramanujan's exponential

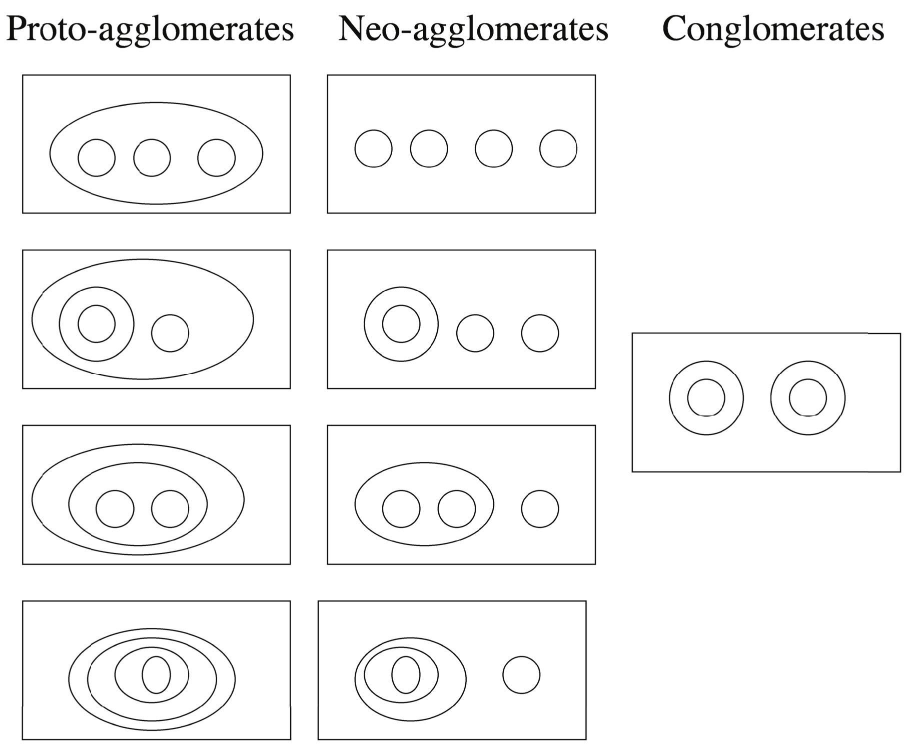

Fig. 1. A representation of the membrane structures with four membranes (the skin is represented by a rectangle)

asympthotic formula [3], making the evaluation of the sum in formula (2) very complex. Now, we adopt an alternative enumeration strategy, by counting all the possible different membrane structures without considering the skin. For this reason we make the following partition of membrane structures, based on the structure of membranes inside the skin: i) **Proto-agglomerates**, ii) **Neo-agglomerates**, and iii) **Conglomerates**. Proto-agglomerates are structure where the skin contains only a singleton multiset. Neo-agglomerates are structure where the skin contains empty multisets. Conglomerates are structures different from proto-agglomerates and neo-agglomerates, that is, inside the skin there is not only a singleton and there are not empty multisets. In Fig. 1 structures on the left side are proto-agglomerates, structures in the middle are neo-agglomerates, and the structure on the right is a conglomerate (the skin is represented by a rectangle). Proto-agglomerates are essentially rooted unlabeled unordered trees, while conglomerates and neo-agglomerates represent unlabeled unordered forests. We denote by $M(n) = T(n+1)$ the number of membrane structures having n membranes (pairs of matching brackets) inside the skin, and by $P(n), N(n), C(n)$, the number of proto-agglomerates, neo-agglomerates, and conglomerates, respectively, having n membranes inside the skin (the skin will not mentioned anymore and it is not counted in the number of membranes). It easy to realize that a membrane structure with n membranes, when it is put inside a further membrane, provides a proto-agglomerate with $n+1$ membranes, while united with the multiset $[[\ ]]$ provides a neo-conglomerate with $n+1$ membranes. The following lemmas are simple consequences of this partition of membrane structures.

Lemma 1. *For $n > 0$ the following equations hold: $M(n) = N(n+1) = P(n+1)$.*

Lemma 2. *For $n > 0$, $M(n+1) = 2M(n) + C(n+1)$.*

Lemma 3. $C(1) = C(2) = C(3) = 0$. *For $n > 0$, $C(n+1) \leq M(n)$.*

Proof. Removing the external membrane in a component of a conglomerate with $n+1$ membranes, provides a membrane structure with n membranes. Therefore conglomerates with $n+1$ membranes are at most $M(n)$.

Lemma 4. *Let $C_i(n)$ denote the number of conglomerates having n membranes and exactly i components, then for $n > 2$*

$$C(n) = \sum_{i=2,\lfloor n/2 \rfloor} C_i(n).$$

Proof. At most $\lfloor n/2 \rfloor$ components can be present in a conglomerate with n membranes. Putting together lemmas 2 and 3 we get the following lemma.

Lemma 5. *For $n > 1$*

$$2M(n) \leq M(n+1) \leq 3M(n).$$

$$2^n \leq M(n+1) \leq 3^n.$$

According to the previous lemmas, we see that in the number $M(n+1)$ the part $2M(n)$ refers to proto-agglomerates plus neo-agglomerate. Therefore, if $M(n)$ is known, the real problem for the computation of $M(n+1)$ is the evaluation of $M(n+1)-2M(n) = C(n+1)$, that is, the number of conglomerates with $n+1$ membranes.

In the case of 1, 2, and 3 membranes we have $M(0) = 1$, $M(1) = 1$, $M(2) = 2$, $M(3) = 4$, as it is indicated in the following schema (skin brackets are not indicated).

1	[ ]			
2	[ ], [ ]	[[ ]]		
3	[ ], [ ], [ ]	[[[ ]]]	[ ], [[ ]]	[[ ], [ ]]

From lemma 6 we evaluate immediately $M(4) = 2M(3)+1 = 9$. In fact, $C(4) = 1$, because there i s only a conglomerate with 4 membranes: [[ ]], [[ ]]. Analogously, $M(5) = 2M(4) + 2 = 18 + 2 = 20$, because there are two conglomerates with 5 membranes: [[ ]], [[ ], [ ]], and [[ ]], [[[ ]]]. The sequence from $M(1)$ up to $M(12)$ (sequence A000081 of The On-Line Encyclopedia of Integer Sequences [8]) provides the following values:

n	1	2	3	4	5	6	7	8	9	10	11	12
M(n)	1	2	4	9	20	48	115	286	719	1842	4766	12486

Let $\mathbb{N}^*$ be the set of finite sequences over the set $\mathbb{N}$ of natural numbers. If $X \in \mathbb{N}^*$, and $j \in \mathbb{N}$ we denote by by $X(j)$ the number which occurs in X at position j. Let $\Pi_{n,k}$ be the set of partitions of the the integer n as sum of k summands. A partition μ of integers is a multiset of integers, let us denote by $\mu(j)$ the number of occurrences of the integer j in μ.

The following operation associates, for any $i \in \mathbb{N}$, a natural number to any sequence $X \in \mathbb{N}^*$ of length n.

$$\bigotimes^i X \;=\; \sum_{\mu \,\in\, \Pi_{\mathbf{n-i+1,\ i}}} \;\prod_{j\in\mu} \binom{X(j)+\mu(j)-1}{\mu(j)} \tag{3}$$

For $i, j \in \mathbb{N}$, let $M(1,..,j)$ denote the sequence $(M(1),\ldots,M(j))$, then the main lemma follows.

Lemma 6. *For* $n > 2$

$$C(n+1) = \sum_{i=2,\lfloor (n+1)/2 \rfloor} \bigotimes^i M(1,..,n). \tag{4}$$

Proof Outline. Conglomerates with n+1 membranes may have 2, 3, ..., but at most a number $\lfloor (n+1)/2 \rfloor$ of components. If we fix a number i of components, then i membranes, of the $n+1$ membranes, must be used for wrapping these i components, therefore the remaining $n+1-i$ are partitioned among these components in all the possible ways. In conglomerates with 2 components $n+1-2$ membranes can be distributed in 2 components. In conglomerates with 3 components $n+1-3$ membranes can be distributed in 3 components, and so on, up to $n+1-\lfloor(n+1)/2\rfloor$ membranes in $\lfloor(n+1)/2\rfloor$ components. In order to compute the number of all possible membrane arrangements, each partition μ of $n+1$ into i summands must be "read", according to the formula $\prod_{j\in\mu}\binom{M(j)+\mu(j)-1}{\mu(j)}$, on the sequence $M(1,..,n)$ of membrane structure numbers. For example, if a partition has three parts, with two equal parts, say $n+1-3=p+p+q$, then in a corresponding conglomerate of three components q membranes can be arranged in $M(q)$ ways in a component, and p membranes can be arranged in $M(p)$ ways in the other two components. However, in the two components with p membranes the repetitions of the same configurations must be avoided. For this reason, the product $\prod_{j\in\mu}\binom{M(j)+\mu(j)-1}{\mu(j)}$ is used. When $\mu(j)=1$, then this formula provides the value $M(j)$, but, when $\mu(j)>1$, the number of different multisets of $M(j)$ elements with multiplicity $\mu(j)$ is provided. In conclusion, the number of all possible colonies is the sum of $\bigotimes^i M(1,\ldots,n)$ for all possible number i of components.

This lemma suggests an algorithm for computing $C(n+1)$. From lemmas 2, 4, and 6 the final proposition follows. The application of the formula of lemma 6, tested for $n=0,\ldots,11$, provided the same values, previously given, of the sequence A000081.

Proposition 1. *For* $n>2$

$$M(n+1)=2M(n)+\sum_{i=2,\lfloor(n+1)/2\rfloor}\bigotimes{}^i M(1,..,n).$$

As an example we provide the computation of $C(11)$. According to lemma 6 the value $C(11)$ is given by:

$$C(11)=\sum_{i=2,5}\bigotimes{}^i M(1,\ldots,10)$$

Now, according to formula (3), we need to compute the right member of this equation for the values $\Pi_{9,2},\Pi_{8,3},\Pi_{7,4},\Pi_{6,5}$ corresponding to the values $2,3,4,5$ of i.

The integer partitions of 9 in two summands yield the following set:

$$\Pi_{9,2}=\{\{8,1\},\{7,2\},\{6,3\},\{5,4\}\}$$

therefore:

$$\bigotimes{}^2 M(1,\ldots,10)=\sum_{\mu\in\Pi_{9,2}}\prod_{j\in\mu}\binom{M(j)+\mu(j)-1}{\mu(j)}=$$

$$[\binom{286+1-1}{1}\binom{1+1-1}{1}]+[\binom{115+1-1}{1}\binom{2+1-1}{1}]+[\binom{48+1-1}{1}\binom{4+1-1}{1}]+[\binom{20+1-1}{1}\binom{9+1-1}{1}]=$$

$286+115\cdot 2+48\cdot 4+20\cdot 9=888.$

The integer partitions of 8 in three summands yield the following set:

$$\Pi_{8,3} = \{\{6,1,1\},\{5,2,1\},\{4,3,1\},\{4,2,2\},\{3,3,2\}\}$$

therefore:

$$\bigotimes^{3} M(1,\ldots,10) = \sum_{\mu\in\Pi_{8,3}} \prod_{j\in\mu} \binom{M(j)+\mu(j)-1}{\mu(j)}$$

$\left[\binom{48+1-1}{1}\binom{1+2-1}{2}\right] + \left[\binom{20+1-1}{1}\binom{2+1-1}{1}\binom{1+1-1}{1}\right] + \left[\binom{9+1-1}{1}\binom{4+1-1}{1}\binom{1+1-1}{1}\right] + \left[\binom{9+1-1}{1}\binom{2+2-1}{2}\right] + \left[\binom{4+2-1}{2}\binom{2+1-1}{1}\right] = 48 + 20\cdot 2 + 9\cdot 4 + 9\cdot 3 + 10\cdot 2 = 171.$

The integer partitions of 7 in four summands yield the following set:

$$\Pi_{7,4} = \{\{4,1,1,1\},\{3,2,1,1\},\{2,2,2,1\}\}$$

therefore:

$$\bigotimes^{4} M(1,\ldots,10) = \sum_{\mu\in\Pi_{7,4}} \prod_{j\in\mu} \binom{M(j)+\mu(j)-1}{\mu(j)} =$$

$\left[\binom{9+1-1}{1}\binom{1+3-1}{3}\right] + \left[\binom{4+1-1}{1}\binom{2+1-1}{1}\binom{1+2-1}{2}\right] + \left[\binom{2+3-1}{3}\binom{1+1-1}{1}\right] =$

$9 + 4\cdot 2 + 4 = 21.$

The integer partitions of 6 in five summands yield the following set:

$$\Pi_{6,5} = \{\{2,1,1,1,1\}\}$$

therefore:

$$\bigotimes^{5} M(1,\ldots,10) = \sum_{\mu\in\Pi_{6,5}} \prod_{j\in\mu} \binom{M(j)+\mu(j)-1}{\mu(j)} =$$

$\left[\binom{2+1-1}{1}\binom{1+4-1}{4}\right] = 2.$

In conclusion, $C(11) = 888 + 171 + 21 + 2 = 1082$, therefore:

$$M(11) = 2M(10) + 1082 = 4766.$$

References

1. Aigner, M.: Discrete Mathematics. American Mathematical Society (2007)
2. Cayley, A.: On the analytical forms called trees, with application to the theory of chemical combinations. Mathematical Papers 9, 427–460 (1875)
3. Conway, J.H., Guy, R.K.: The Book of Numbers. Springer, Heidelberg (1996)
4. Knuth, D.: The Art of Computer Programming. Fundamental Algorithms, vol. 1. Addison Wesley, Reading (1968)

5. Otter, R.: The number of trees. The Annals of Mathematics 2nd Ser., 49, 583–599 (1948)
6. Păun, G.: Personal Communication (October 1998)
7. Păun, G.: Computing with membranes. J. Comput. System Sci. 61, 108–143 (2000)
8. Sloane, N.J.A.: The on-line encyclopedia of integer sequences. Notices of The American Mathematical Society 59, 912–915 (2003)