Research Article

Bargraphs of combinations with repetition

Aubrey Blecher, Arnold Knopfmacher

The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of Witwatersrand, South Africa

(Received: 13 June 2023. Received in revised form: 21 June 2023. Accepted: 15 January 2024. Published online: 4 April 2024.)

© 2024 the authors. This is an open-access article under the CC BY (International 4.0) license (www.creativecommons.org/licenses/by/4.0/).

Abstract

Generating functions that track some geometrical features of combinations with repetition are developed; namely, the semi-perimeter, inner site-perimeter, and outer site-perimeter, each of whose meanings depends on the interpretation of the combination as a bargraph. The paper has three theorems, which respectively give the total number for each of these statistics based on separate generating functions tracking these statistics.

Keywords: combinations with repetition; generating functions; semi-perimeter; site-perimeter.

2020 Mathematics Subject Classification: 05A10, 05A15.

1. General introduction

In elementary mathematics, choosing (allowing replacement of) \( p \) objects from \( n \) different objects, regardless of the order of choice, is called a combination with repetition. The number of possible such combinations is a basic and elementary statistic. Indeed, choosing a combination of \( j \) objects with repetition from a set of \( n \) objects can be done in \( \binom{n+j-1}{j} \) ways. By definition, such a combination is of the form \( \{c_1,c_2,\cdots,c_j\} \) where

\[
1 \leq c_1 \leq c_2 \leq \cdots \leq c_j \leq n.
\]

In this paper, we track the number of ways of choosing \( j \) integers from \([n] \), where we allow repetitions but ignore the order in which we choose the \( j \) integers. What is different here is that we track certain geometrical statistics of the combinations, namely the semi-perimeter, inner site-perimeter, and outer site-perimeter, which are defined below. The idea for this paper follows that of Mansour and Shabani [8], where these same statistics are studied in combinations that do not allow repetition. The geometric parameters above are based on the interpretation of such a combination as a bargraph. For a survey of topics related to bargraphs, see [7]. A bargraph representation of a combination \( c_1,c_2,\ldots,c_j \) of \( j \) integers from \([n] \) is a left-to-right sequence of columns whose lower boundaries are at the same level and where the \( i \)th column has \( c_i \) vertical squares, each of which is called a cell. We think of the cells of a bargraph as lying within a larger grid of such cells, some of which are inside the bargraph and the rest being outside.

In this paper, we study three statistics. Firstly, the semi-perimeter of a bargraph is half the number of edges on the boundary of the bargraph. Secondly, the inner site-perimeter is the number of cells inside the bargraph that have at least one common edge with an outside cell. Thirdly, the outer site-perimeter is the number of cells outside the bargraph that have at least one common edge with a cell inside the bargraph.

In this paper, we track the number of elements or the cardinality of combinations. We develop a different generating function in each of these sections; in the next section, the semi-perimeter, then the outer site-perimeter, and finally in the last section, the inner site-perimeter (see [6]).

As an example of the inner site-perimeter and outer site-perimeter, we illustrate in Figure 1.1, the combination \( \{2,4,4,5\} \). The outer site-perimeter is the sum of the cells marked by “o”, i.e., 16, whereas the inner site-perimeter is the sum of the cells marked by “x”, i.e., 12. The semi-perimeter of the bargraph is 9. The number of parts is clearly 4 and this is a combination whose parts come from \([n] \), where \( n \) is any integer greater than or equal to the largest part of the combination, say \( n = 6 \).
2. Semi-perimeter

We denote by \(\text{card}(\pi)\) the cardinality of the combination \(\pi\) of \([n]\). Let \(P_n(p, q)\) be the generating function for the number of combinations \(\pi\) of \([n]\) with repetition according to the statistics \(\text{card}\) tracked by \(p\) and semi-perimeter \((sp)\) tracked by \(q\).

\[
P_n(p, q) := \sum_{\pi \in C_n} p^{\text{card}(\pi)} q^{sp(\pi)}.
\]

Since each combination of \([n]\) either contains \(n\) or not, we have

\[
P_n(p, q) = P_{n-1}(p, q) + P_n(p, q|n), \tag{1}
\]

where \(P_n(p, q|i)\) is defined to be the generating function for such combinations ending in \(i\). We consider the two cases of cardinality 1 or larger (where the last two columns are \(jn\)) and therefore obtain:

\[
P_n(p, q|n) = pq^{n+1} + p \sum_{j=1}^{n-1} q^{n+1-j} P_j(p, q|j) \tag{2}
\]

which implies

\[
qP_{n-1}(p, q|n-1) = pq^{n+1} + p \sum_{j=1}^{n-1} q^{n+1-j} P_j(p, q|j) \tag{3}
\]

with \(P_1(p, q|1) = P_1(p, q) = \frac{pq^2}{1-pq}\). Subtracting Equation (3) from (2), we obtain

\[
P_n(p, q|n) = \frac{qP_{n-1}(p, q|n-1)}{1-pq},
\]

which we iterate \(n-1\) times to obtain

\[
P_n(p, q|n) = \frac{pq^n}{(1-pq)^n}.
\]

So, from Equation (1), we have

\[
P_n(p, q) = P_{n-1}(p, q) + \frac{pq^n}{(1-pq)^n}.
\]

Again, we iterate the latter equation to obtain

\[
P_n(p, q) = pq \sum_{j=1}^{n} \frac{q^j}{(1-pq)^j} = \frac{pq^2 (1-q^n (1-pq)^{-n})}{1 - (1+p)q}.
\]

To find the generating function for the total semi-perimeter we differentiate the last equation with respect to \(q\) and set \(q = 1\) to obtain

\[
\frac{\partial}{\partial q} \left( \frac{pq^2 (1-q^n (1-pq)^{-n})}{1 - (1+p)q} \right) \bigg|_{q=1} = -\frac{(1-p)^{-1-n} (1 - (1-p)^n + (-2 - n + 2(1-p)^n) p - (-1 + (1-p)^n) p^2)}{p}. \tag{4}
\]

Now, we extract the coefficient of \(p^j\) and obtain

\[
\frac{(1+j(1+j+n)) \binom{j+n}{j+1}}{j+n}.
\]

Finally, dividing by \(\binom{j+n-1}{j}\) (which is the number of combinations of \([n]\) with \(j\) parts allowing repetition), we obtain the next result.
Theorem 2.1. For fixed \( n \) and \( j \), the average semi-perimeter for the class of size \( (j+n-1) \) of all combinations of \([n]\) with repetition and with \( j \) parts, is

\[
\frac{1 + j(1 + j + n)}{1 + j}.
\]

Example 2.1. When \( n = 3 \) and \( j = 4 \), we obtain the following list of the 15 combinations with repetition

\[
\{1,1,1,1\}, \{1,1,1,2\}, \{1,1,1,3\}, \{1,1,2,2\}, \{1,1,2,3\},
\{1,1,3,3\}, \{1,2,2,2\}, \{1,2,2,3\}, \{1,2,3,3\},
\{1,3,3,3\}, \{2,2,2,2\}, \{2,2,2,3\}, \{2,2,3,3\}, \{2,3,3,3\}, \{3,3,3,3\}\]

and average semi-perimeter of 33/5 as asserted by Theorem 2.1.

Next, we briefly consider the following problem. Suppose we are given a combination from \([n]\) with \( j \leq n \) parts and semi-perimeter \( k \), where \( k \geq j \). Can we find the generating function for the area of the different combinations having these properties?

The bargraph representation of the combination \( c_1, c_2, \ldots, c_j \) has area \( c_1 + c_2 + \cdots + c_j \) and fits within a rectangle of width \( j \) and height \( n \). The number of combination bargraphs that have exactly \( j \) columns and fit within such a rectangle is given by

\[
q^j \binom{n + j - 1}{j}_q,
\]

where

\[
\binom{n + j - 1}{j}_q
\]

is the \( q \)-binomial coefficient. This is in accordance with Equation (7.1) in [1]. Also \( q^j \binom{n+j-1}{j}_q \) is the generating function tracked by area for all combination bargraphs with \( j \) parts and semi-perimeter \( \leq n + j \).

Hence the generating function for all combinations with \( j \) parts and semi-perimeter \( k \) is given by

\[
q^j \binom{k-1}{j}_q - q^j \binom{k-2}{j}_q.
\]

Thus, by letting \( q \to 1 \), the number of bargraphs with semi-perimeter \( k \) is

\[
\binom{k-1}{j} - \binom{k-2}{j}.
\]

This gives the formula for the total semi-perimeter:

\[
\sum_{k=j+1}^{j+n} k \left( \binom{k-1}{j} - \binom{k-2}{j} \right) = \frac{n(j(n+1)+1) \left( \binom{j+n}{j} - \binom{j+n-1}{j} \right)}{j(j+1)}.
\]

Dividing this by \( (n+j-1) \) we re-derive the average given in Theorem 2.1.

3. Outer site-perimeter

Outer site-perimeter is defined as the number of nearest neighbour cells outside the bargraph. See for example, [4, 5], where the statistic is simply called the site-perimeter. Outer site-perimeter is illustrated in the left portion of Figure 1.1.

Let \( P_n(p, q) \) be the generating function for the number of combinations \( \pi \) of \([n]\) according to the statistics \( \text{card} \) tracked by \( p \) and \( \text{op} \) (outer site-perimeter) tracked by \( q \).

Also, we define \( P_0(p, q) = 1 \). First, let us write a recurrence relation for \( P_n(p, q) \). Since each combination of \([n]\) either contains \( n \) or not, we have

\[
P_n(p, q) = P_{n-1}(p, q) + P_n(p, q|n),
\]

where \( P_n(p, q|n) \) is the generating function for the number of combinations of \([n]\) that contain \( n \), again according to the statistics \( \text{card} \) and \( \text{op} \). Since each combination \( \pi \) of \([n]\) that contains \( n \) either has only one element, or the second maximal element in \( \pi \) is either \( j, 1 \leq j \leq n - 1 \) or \( n \), we obtain

\[
P_n(p, q|n) = pq^{2n+2} + \sum_{j=1}^{n-1} q^{2n+1-2j} P_j(p, q|j) + pq^2 P_n(p, q|n).
\]
Hence
\[ P_n(p, q|n) = \frac{pq^{2n+2}}{1 - pq^2} + \frac{p}{1 - pq^2} \sum_{j=1}^{n-1} q^{2n+1-2j} P_j(p, q|j), \]
from which we obtain
\[ q^2 P_{n-1}(p, q|n-1) = \frac{pq^{2n+2}}{1 - pq^2} + \frac{p}{1 - pq^2} - \sum_{j=1}^{n-2} q^{2n+1-2j} P_j(p, q|j). \]
Subtracting the latter equation from the previous one, we obtain the recursion
\[ P_n(p, q|n) = \left( q^2 + \frac{pq^3}{1 - pq^2} \right) P_{n-1}(p, q|n-1) \]
where
\[ P_1(p, q|1) = \frac{pq^4}{1 - pq^2}, \]
and iterating this yields
\[ P_n(p, q|n) = \frac{pq^2}{1 + pq - pq^2} \left( \frac{q^2(-1 + p(-1 + q)p)}{-1 + pq^2} \right)^n. \]
Now, substituting this into Equation (6) we obtain
\[ P_n(p, q) = 1 + \sum_{j=1}^{n} P_n(p, q|j) \]
\[ = \frac{1 - (1+p)q^2 - pq^3 - pq^4 \left( -2 + \left( q^2(-1 + p(-1 + q)p) \right) \right)^n}{1 - (1+p)q^2 - pq^3 + pq^4}. \] (7)
Differentiating with respect to \( q \) and putting \( q = 1 \), we find the generating function for the total outer site-perimeter to be
\[ \left. \frac{\partial P_n(p, q)}{\partial q} \right|_{q=1} = -3 + \frac{2}{p} + \left( 5 + 2n - \frac{2}{p} + (-3 - n)p + np^2 \right) \left( \frac{1}{1-p} \right)^{n+1}. \]
Next, we extract the coefficient of this partial derivative to obtain
\[ [p^j] \frac{\partial P_n(p, q)}{\partial q} \bigg|_{q=1} = (5 + 2n) \binom{n+j}{j} - 2 \binom{n+j+1}{j+1} - (3 + n) \binom{n+j-1}{j-1} + n \binom{n+j-2}{j-2} \]
\[ = -3 + j^2 + 2j^3 + 3 \left( 1 + j^2 \right) n + 2jn^2 \left( j + n \right) \frac{1}{(n+j)(n+j-1)}. \] (8)
Now, dividing by \( \binom{j+n-1}{j} \), we obtain the next result.

**Theorem 3.1.** For fixed \( n \) and \( j \), the average outer site-perimeter for the class of size \( \binom{j+n-1}{j} \) of all combinations of \([n]\) with repetition and with \( j \) parts, is
\[ \frac{3 - j^2 - 2j^3 - 3 \left( 1 + j^2 \right) n - 2jn^2}{(1+j)(1-j-n)}. \]

**Example 3.1.** When \( n = 3 \) and \( j = 4 \), the list of the 15 combinations with repetition given in (5) has average outer site-perimeter of 61/5 as asserted by Theorem 3.1.

## 4. Inner site-perimeter

The inner site-perimeter is defined as the number of cells inside the bargraph that have at least one-edge lying on the bargraph perimeter; for example, see [2, 3]. The inner site-perimeter is illustrated in the right-hand portion of Figure 1.1.

Let \( P_n(p, q) \) be the generating function for the number of combinations \( \pi \) of \([n]\) according to the statistics \( \text{card} \) tracked by \( p \) and \( \text{ip} \) (i.e., inner site-perimeter) tracked by \( q \).

Notationally, we use \( P_n(p, q|j_1, j_2, \ldots, j_s) \) to mean the generating function for combinations of \([n]\) that end with the parts \( j_1, j_2, \ldots, j_s \).

The strategy for obtaining the generating function in this section is to express \( P_n(p, q|jn) \) in terms of \( P_{j+1}(p, q|j(j+1)) \). This will enable us to solve for the generating function recursively as specified later, in Equation (17).
Also, we define \( P_0(p, q) = 1 \). First, let us write a recurrence relation for \( P_n(p, q) \). Since each combination of \([n]\) either contains \( n \) or not, we have

\[
P_n(p, q) = P_{n-1}(p, q) + P_n(p, q|n)
\]

(9)

where \( P_n(p, q|n) \) is the generating function for the number of combinations of \([n]\) that contain \( n \), again according to the statistics \( \text{card} \) and \( i_p \). Since each combination \( \pi \) of \([n]\) that contains \( n \) either has only one element, or the second maximal element in \( \pi \) is either \( j \) (\( 1 \leq j \leq n-1 \)) or \( n \), we obtain

\[
P_n(p, q|n) = pq^n + \sum_{j=1}^{n-1} P_n(p, q|jn) + P_n(p, q|nn)
\]

(10)

\[
= pq^n + \sum_{j=1}^{n-1} q^{n-1-j} P_{j+1}(p, q|j(j+1)) + P_n(p, q|nn),
\]

(11)

where the summand \( q^{n-1-j} P_{j+1}(p, q|j(j+1)) \) is a replacement for the summand of the previous line \( P_n(p, q|jn) \) by exchanging \( n \) with \( j+1 \) and compensating with \( q^{n-1-j} \). Now, focusing on \( P_n(p, q|nn) \), we obtain

\[
P_n(p, q|nn) = p^2q^{2n} + \sum_{j=1}^{n-1} P_n(p, q|jnn) + P_n(p, q|nnn)
\]

(12)

\[
= p^2q^{2n} + \sum_{j=1}^{n-1} pq^{n+1-j} P_n(p, q|jn) + pq^2 P_n(p, q|nn).
\]

Hence

\[
P_n(p, q|nn) = \frac{p^2q^{2n}}{1-pq^2} + \frac{p}{1-pq^2} \sum_{j=1}^{n-1} q^{n+1-j} P_n(p, q|jn).
\]

(13)

Now, we substitute the latter equation into Equation (10) to obtain

\[
P_n(p, q|n) = pq^n + \sum_{j=1}^{n-1} P_n(p, q|jn) + \frac{p^2q^{2n}}{1-pq^2} + \frac{p}{1-pq^2} \sum_{j=1}^{n-1} q^{n+1-j} P_n(p, q|jn)
\]

\[
= pq^n + \frac{p^2q^{2n}}{1-pq^2} + \sum_{j=1}^{n-1} \left(1 + \frac{pq^{n+1-j}}{1-pq^2}\right) P_n(p, q|jn)
\]

\[
= pq^n + \frac{p^2q^{2n}}{1-pq^2} + \sum_{j=1}^{n-1} \left(1 + \frac{pq^{n+1-j}}{1-pq^2}\right) q^{n-1-j} P_{j+1}(p, q|j(j+1)),
\]

(14)

where the last line follows using (11).

Next, for \( j > 1 \), we have

\[
P_{j+1}(p, q|j(j+1)) = p^2q^{2j+1} + \sum_{i=1}^{j-1} P_{j+1}(p, q|i(j+1))
\]

\[
= p^2q^{2j+1} + \sum_{i=1}^{j-1} q^{2(j-1-i)} P_{i+2}(p, q|i+1(i+2)) + P_{j+1}(p, q|j(j+1))
\]

\[
= p^2q^{2j+1} + p \sum_{i=1}^{j-1} q^{2j+1-2i} P_{i+1}(p, q|i+1) + pq^2 P_{j+1}(p, q|j(j+1)).
\]

Hence for \( j > 1 \), we have

\[
P_{j+1}(p, q|j(j+1)) = \frac{p^2q^{2j+1}}{1-pq^2} + \frac{p}{1-pq^2} \sum_{i=1}^{j-1} q^{2j+1-2i} P_{i+1}(p, q|i+1).
\]

(15)

Thus, for \( j > 2 \), we obtain

\[
P_{j+1}(p, q|j(j+1)) - q^2 P_j(p, q|(j-1)j) = \frac{p}{1-pq^2} q^2 P_j(p, q|(j-1)j).
\]

(16)

\[
P_{j+1}(p, q|j(j+1)) = q^2 \left( \frac{pq}{1-pq^2} + 1 \right) P_j(p, q|(j-1)j).
\]

(17)
Since $P_3(p, q|23) = \frac{1}{1-pq} \frac{p^2q^2}{1-pq^2}$, by iteration for $j \geq 2$, we have

$$P_{j+1}(p, q, j(j+1)) = P_j(p, q|23)q^{2j-4} \left(\frac{pq}{1-pq^2} + 1\right)^{j-2}$$

(18)

Substitute this into Equation (14) to obtain

$$P_n(p, q|n) = pq^n + \frac{p^2q^{2n}}{1-pq^2} + \left(1 + \frac{pq^n}{1-pq^2}\right) q^{n-2}P_2(p, q|12) + \sum_{j=2}^{n-1} \frac{p^2q^{2n-1-j}}{1-pq^2} q^{n-1-j}P_n(p, q|23)q^{2j-4} \left(1 + \frac{pq}{1-pq^2}\right)^{j-2} \left(1 + \frac{pq^{n+1-j}}{1-pq^2}\right).$$

(19)

Now, by substituting this into Equation (9), we obtain

$$P_n(p, q) = 1 + \sum_{j=1}^{n} P_n(p, q|j) = 1 + pq \left(1 - pq^2\right) \times \frac{(1 - (1 + 2p)q^2 - pq^3 + pq^4 - q^4(1 - (1 + p)q^2 + pq) + pq^{2+2n} \left(1 + \frac{pq}{1-pq^2}\right)^n)}{(1 - pq)(1 - q - 2pq^2 + pq^3)(1 - (1 + p)q^2 - pq^3 + pq^4)},$$

(20)

where the simplification is a result of summing a finite geometric series. Differentiating with respect to $q$ and putting $q = 1$, we find the generating function for the total inner site-perimeter to be

$$\left.\frac{\partial P_n(p, q)}{\partial q}\right|_{q=1} = \frac{3 - 4p + np + 2p^2 - np^2}{(1 - p)p} - \frac{3 - 4p - 2np + 2p^2 + np^2 - np^3}{(1 - p)p} \left(\frac{1}{1 - p}\right)^n.$$

(21)

Next, we extract the coefficient of this partial derivative to obtain

$$\left.\frac{[p^j]}{p} \frac{\partial P_n(p, q)}{\partial q}\right|_{q=1} = -3 \binom{n+j}{j+1} + (1+2n) \binom{n+j-1}{j} + (n-1) \binom{n+j-2}{j-1} + (2n-1) \sum_{i=2}^{j} \binom{n+j-1-i}{j-i} + 1$$

$$= (2n-1) \binom{j+n-2}{j-2} + (n-1) \binom{j+n-2}{j-1} + (2n+1) \binom{j+n-1}{j} - 3 \binom{j+n}{j+1} + 1.$$

(22)

Dividing by $\binom{j+n-1}{j}$, we obtain the next theorem.

**Theorem 4.1.** For fixed $n$ and $j$, the average inner site-perimeter for the class of size $(j+n-1)$ of all combinations of $[n]$ with repetition and with $j$ parts, is

$$\frac{(2j-1)n}{1+j} + j - 2 + \frac{3}{1+j} - \frac{j}{n} + \frac{(j-1)n}{j+n-1} + \frac{1}{(j+n-1)}.$$

**Example 4.1.** When $n = 3$ and $j = 4$, the list of the 15 combinations with repetition given in (5) has average inner site-perimeter of 113/15 as asserted by Theorem 4.1.
References