Probability in the Engineering and Informational Sciences, **30**, 2016, 244–260.
doi:10.1017/S0269964815000364

ON THE COMBINATORICS OF BINARY SERIES-PARALLEL GRAPHS

MICHA HOFRI
Department of Computer Science,
Worcester Polytechnic Institute,
Worcester, MA, USA

CHAO LI
Department of Computer Science,
Worcester Polytechnic Institute,
Worcester, MA, USA

HOSAM MAHMOUD
Department of Statistics,
The George Washington University,
Washington, D.C., USA
E-mail: hosam@guru.edu

Binary series-parallel (BSP) graphs have applications in transportation modeling, and exhibit interesting combinatorial properties. This work is limited to the second aspect. The BSP graphs of a given size – measured in edges – are enumerated. Under a uniform probability model, we investigate the asymptotic distribution of the order (number of vertices) and the asymptotic average length of a random walk (under different strategies) of large graphs of the same size. The systematic method throughout is to define the graphs, and the features we evaluate by a structural equation (equivalent to a regular expression). The structural equation is translated into an equation for a generating function, which is then analyzed.

All (distinct) BSP graphs are created equal

1. INTRODUCTION

Complex networks are spreading all around us. Their importance is manifested in our use of social networks, communication systems and road maps. We need to propose and analyze models for many types of emerging networks. The class of series-parallel (SP) graphs is one model of networks representing flow of commodities and merchandize from a source (producer) to a sink (market).

There are two definitions of families of SP graphs. The complete graph K_n on n vertices is at the core of these definitions. In the first definition, an SP graph is an undirected connected graph that does not contain K_4 as a minor (i.e., the graph cannot be reduced to K_4 after any sequence of edge deletions, node deletions or node identifications), [1].

© Cambridge University Press 2016 0269-9648/16 $25.00

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

This definition does not distinguish two vertices as a source and a sink. The other popular definition is recursive; it constructs SP graphs from smaller members of the family. In this variant, the smallest SP graph is K_2. The two vertices are called *poles* (one *North* and one *South*). Larger SP graphs are obtained from smaller ones by one of two edge operations: A *series operation*, which converts an edge into a two-edge path, and a *parallel operation*, which creates an edge parallel to an existing one; both are detailed in the following section.

The *size* of an SP graph is the number of edges in it, and its number of nodes is the graph *order*. For the network flow application we have in mind, we think of SP graphs as directed, with orientation assigned to the edges to allow the flow to move from the North pole to the South pole.

Applications often impose restrictions on the growth of random graphs. For example, in hardwired communication circuit models there can be a limit on the fan-out (outdegree) of a node in order to avoid the flow of large electrical current through the wire feeding into their shared northern node. In a road network between two trading cities (one producing and one consuming), each town along the trading routes from the producer to the market may have a limited budget to build at most two roads toward the market.

We study some combinatorics of directed binary series-parallel (BSP) graphs under the natural probability assumption: Distinct graphs of the same size are assigned the same likelihood. This provides an alternative model of randomness on BSP graphs to that in the incremental model of [3], where the author uses the same graph creation process, and assigns the same likelihood to each sequence of edge selections, even though many such different sequences lead to the same BSP graph. Resulting graphs of the same size have then a range of different probabilities. Equal likelihood to all distinct random BSP graphs of a given size is the measure used for the rest of this note. The arbitrariness of this distributional assumption is illusory: The BSP graphs can arise independently of the insertion process we describe in Section 2; they can be generated from the structure formulas described throughout, and properly used, those formulas would only generate distinct graphs.

We shall address a number of combinatorial questions on random BSP graphs. We organize the discussion as follows. Section 2 describes the generation process of the graphs systematically. We then proceed to look at some of their properties: In Section 3, the BSP graphs of a given size are enumerated. Section 4 establishes structural properties of one side of a BSP graph. In Section 5, we pursue the distribution of the order of a random BSP graph. The properties of Section 4 are used in Section 6 to aid in obtaining results on the length of a random walk from the North pole to the South pole under different walking strategies.

2. GENERATING ALL BSP GRAPHS

The starting point in creating the family of BSP graphs is K_2, which we consider as generation 1 (the initial graph) of the family. As usual, the initial graph K_2 is diagrammed as ⁝. All edges in the graphs are viewed as oriented, from north to south (from top to bottom on the page), and the two nodes in this BSP graph are called the North pole and the South pole; they are never changed.

The graphs evolve in steps; each graph created in step n is a member of the nth generation. Graphs created at step n have n edges, and are collected in the set G_n. In creating the next generation we use the algorithm below. It calls for two possible edge-doubling operations:

Parallelization: The edge uv (directed from a vertex u to vertex v), is doubled by adding another edge, also directed from u to v.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

Serialization: The edge uv, is replaced with two edges: ux (directed from u to x), and xv (directed from x to v), adding a new node x.

This is how the graphs in the set G_n are processed to create the BSP graphs of the next step, the set G_{n+1}, which is initially empty.

```
1.    For each g ∈ G_n
2.        for each e an edge of g
3.            double (e):
4.            if the North pole of e is of degree one
5.                then parallelize e;
6.                else serialize e;
7.            if the created graph is distinct from all previous G_{n+1} graphs,
8.                then add it to G_{n+1};
9.                else discard it;
10.   end
```

Parallelization is impossible for an edge which shares its northern vertex with another edge (southern vertices can be shared indefinitely). For such an edge we use the alternative method of doubling. Some, possibly many of these graphs may not be distinct; all duplicates but one are discarded.

How many distinct such graphs of size n exist? This is a question of enumeration that we take up in Section 3. The sequel deals with further properties of graphs in the family so created.

The idea behind giving priority to parallelization is to increase the outdegree of the northern side of the selected edge as quickly as possible to the allotted quota (a limit of 2 in the binary case), perhaps to spend the road building budget, always a desired budgetary strategy in city planning.

As an illustration, consider the BSP graphs in generation 2, as created from K_2, the single BSP graph in generation 1. The lone edge in K_2 can be parallelized, and we get , the only BSP in generation 2, denoted by G_2.

What happens in generation 3? Neither edge in G_2 can be parallelized. We select an edge and serialize it, creating thus one graph, then go back to G_2 and serialize its second edge. This creates the two BSP graphs (which we denote, from left to right, by $G_3^{(1)}$ and $G_3^{(2)}$):  and .

The next section continues the discussion of this evolution, and connects it with the combinatorics of the generated graphs. When describing the graphs, we refer to a pair of parallel edges as a *bubble*, whether it is as simple as the two-edge BSP G_2, or more complex, such as the two graphs of size 3; they are both bubbles. As seen below, bubbles can appear within bubbles, and all BSP graphs of more than a single edge are bubbles. A bubble has two *sides*, left (or West) and right (East). We use the pairs of terms interchangeably. Each side of a bubble can consist of any number of edges and bubbles interleaved, except that the first (top-most) is always a simple edge, since it could not be parallelized into a bubble in its own right (that would require the North pole of the bubble to have a forbidden outdegree, of 3 or more).

The generation process creates certain BSP graphs multiple times, just as several permutations of a set of keys, when inserted into a binary search tree, can end up creating identical trees. For example, in generating BSP graphs of size 4, we get the same graph if

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

we choose the left edge in $G_3^{(1)}$ or the right edge of $G_3^{(2)}$. Both choices give the same graph, which is . Throughout the discussion we are interested in distinct BSP graphs, and not in the generation sequences that produced them.

3. COUNTING BSP GRAPHS

We use combinatorial structure expressions to describe all possible BSP graphs and their components. Let the class of all BSP graphs be denoted by B, with an ordinary generating function $\varphi_B(x)$, where x marks the number of edges in the graph: $\varphi_B(x) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \tau_n x^n$; which we interpret as saying that τ_n is the total count of BSP graphs of size n.

The following notation guides us in the construction of BSP graphs. The initial BSP graph, a single edge between two nodes (the graph K_2), , is denoted by a'. The same edge carrying a single node (in the North pole), , is denoted by a. We found it convenient to see the addition of an edge by serialization as replacing an edge by two a graphs in series. The node of the top a merges with the North pole of the replaced edge, the node of the bottom a marks the serialization, and provides the top edge with a South pole; the bottom a inherits its South pole from the serialized edge.

All other BSP graphs are bubbles that are two-sided structures; this set is denoted by b. Such a structure can be simply two edges from the North pole to the South pole, such as the graph G_2 (our second generation graph), or more complex, such as $G_3^{(1)}$ and $G_3^{(2)}$ (the two BSP graphs in the third generation). Observe that since all edge choices are used in the generation algorithm, then however large a size we consider, there are BSP graphs where one of the sides, originally from , remains pristine, unchanged since infancy, while all subsequent development occurs in the other side. In the next section, we show how many such lopsided graphs exist.

There are $(n-1)!$ evolutionary paths leading from the initial K_2 to the set of all BSP graphs of size n; however, there are far fewer such distinct graphs. The phenomenon of distinct evolutionary paths leading to the same BSP graph is seen first with graphs of size 4, where the graph  can be generated twice (yet we count both occurrences as one BSP graph) from the two distinct BSP graphs $G_3^{(1)}$ and $G_3^{(2)}$. We mentioned the similarity to constructing an n-node binary search tree by inserting an n-long permutation of distinct values into an initially empty tree. We shall see the similarity going even further.

We proceed to count BSP graphs using the elements defined above. By defining them as elements in a class captured by regular expressions, with operators guaranteed to generate distinct structures, we shall obtain a count of distinct BSP graphs. The enumeration result is in terms of the celebrated sequence of Catalan numbers, known to count several other combinatorial classes, such as binary trees, plane trees, Dyck languages, polygon triangulations and more (Exercise 6.19 in [6] lists 66 combinatorial varieties counted by Catalan numbers). We denote the nth Catalan number, $\binom{2n}{n}/(n+1)$, by C_n. A good asymptotic estimate of these numbers for a large argument is

$$C_n = \frac{4^n}{\sqrt{\pi n^3}} \left(1 - \frac{9}{8n} + O\Big(\frac{1}{n^2}\Big)\right). \tag{1}$$

THEOREM 1: *The number of distinct* BSP *graphs of size* n *is given by* $\tau_n = C_{n-1} = \binom{2n-2}{n-1}/n$.

PROOF: The key observation is that we can describe each of the two sides of b by the expression $a(a+b)^\star$. Only an a structure can appear to "start a side" at the North pole (and

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

whenever it is selected for propagation, it manufactures another a edge by serialization). The choice of any other a leads to create a basic b bubble, which continues to evolve in the same way. The resulting structural relation is

$$B \models a' + b; \quad b \models (a(a+b)^\star)^2, \tag{2}$$

where the notation $\models$ can be read as "has a structure as," or "is modeled by." Similarly to the definition of $\varphi_B(x)$ at the beginning of the section, we shall use $\varphi_b(x)$ which is its main part, and differs in not containing τ_1; in other words, $\varphi_b(x) = \sum_{n=2}^{\infty} \tau_n x^n$.

Relation (2) corresponds to equations for the corresponding ordinary generating functions:

$$\varphi_B(x) = x + \varphi_b(x); \quad \varphi_b(x) = \left(x \frac{1}{1 - x - \varphi_b(x)} \right)^2.$$

The equation for the bubble counting generating function is then

$$\varphi_b(x)\big(1 - x - \varphi_b(x)\big)^2 = x^2,$$

and this cubic equation obliges by factoring into

$$\big(1 - \varphi_b(x)\big)\big(\varphi_b(x)\big(1 - \varphi_b(x)\big) - 2x\varphi_b(x) - x^2\big) = 0.$$

The root $\varphi_b(x) = 1$ is spurious, and the quadratic equation has the two solutions

$$\varphi_b(x) = \frac{1 \pm \sqrt{1 - 4x}}{2} - x.$$

Again, one of the solutions (the one with the "+" sign) is spurious, giving rise to negative coefficients. The second option provides us with the needed counts:

$$\varphi_b(x) = \frac{1 - \sqrt{1 - 4x}}{2} - x \quad \Longrightarrow \quad \varphi_B(x) = \frac{1 - \sqrt{1 - 4x}}{2}, \tag{3}$$

which expands as $\tau_n = [x^n]\varphi_B(x) = \frac{1}{n}\binom{2n-2}{n-1} = C_{n-1}$. ■

Figure 1 shows all the $C_4 = 14$ distinct BSP graphs having five edges:

In addition to providing a count of the BSP graphs, we shall see that the Catalan numbers permeate all the calculations that relate to these graphs.

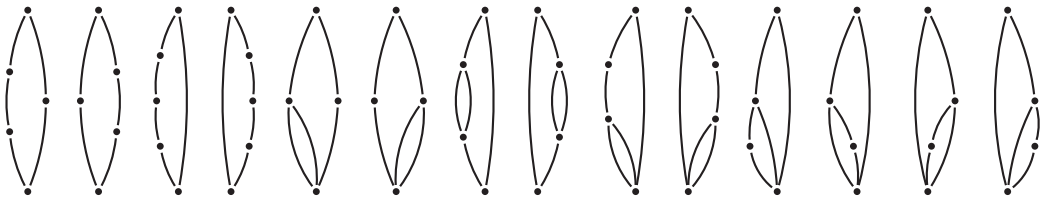


FIGURE 1. All $C_4 = 14$ BSPs of five edges.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

We noted that the evolution process of BSP graphs is very similar to what we find with binary search trees carrying n labels: They can be created using $n!$ insertion sequences, and end up having only C_n distinct configurations. We can even tighten the similarity by observing that the n-node binary search trees have also $n-1$ edges.

The practice of evaluating the moments of several structural properties of binary search trees under two probability measures is long established [5, Chapter 6]. One measure is seen when we consider each *insertion sequence* as equally likely. Some sequences generate identical trees; the second measure is based on seeing each *distinct tree* as equally likely. Some properties, such as height, are strikingly different under the two probability measures. Here we use a similar approach: One of us, in [3], computed some parameters in such structures, when each BSP graph edge-addition sequence is considered equally likely. In the sections below we take the complementary point of view: Each of the C_{n-1} BSP graphs of size n is equally likely.

4. CONSTRUCTIVE COMBINATORICS OF BSP GRAPH SIDES

The versatility of the basic recurrence satisfied by the Catalan numbers,

$$C_{k-1} = \sum_{j=1}^{k-1} C_{j-1} C_{k-j-1}, \quad k \geqslant 1, \tag{4}$$

appears in several aspects of BSP graphs, just as it does for binary trees.

THEOREM 2:

(1) *Let the random variable W_n be the number of edges "allocated" to the West side of a* BSP *graph of size n. Then, its probability distribution is given by*

$$\mathbb{P}(W_n = j) = \frac{C_{j-1} C_{n-j-1}}{C_{n-1}}, \quad j \geqslant 1, \tag{5}$$

(2) *Let the random variable T_m be the number of edges from the North pole down to the first bubble in a side of size m (having m edges) of a* BSP *graph. Its probability distribution is given by*

$$\mathbb{P}(T_m = k) = \frac{1}{C_{m-1}} \times \begin{cases} 3\frac{m-k-1}{m-k+1} C_{m-k-1}, & \text{when } 1 \leqslant k < m; \\ 1, & \text{when } k = m. \end{cases} \tag{6}$$

PROOF: **1.** Consider a BSP graph of size $n \geqslant 2$, and denote by w the number of edges it has on the West side. If we denote by R_k the number of configurations that a side with k edges can have, we require $C_{n-1} = \sum_{w=1}^{n-1} R_w R_{n-w}$. The solution of this recurrence is given by (4) as $R_w = C_{w-1}$, and the fraction of all BSP graphs of size n which have w edges on the left is as claimed in (5); the initial conditions agree as well ($R_1 = C_0 = 1$).
2. When a side of m edges begins at the North pole with a string of k edges, till the first bubble is encountered, all of its C_{m-1} possible configurations occur in the lower $m-k$ edges. Let us denote this structure by V; Equation (2) tells us to represent it as $V \models b(a+b)^\star$, hence the generating function for the number of V structures with a given number of edges

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

is provided as

$$\varphi_V(x) = \frac{\varphi_b(x)}{1 - x - \varphi_b(x)} = \frac{1 - s(x) - 2x}{1 + s(x)} = \frac{1 - 3x - s(x)(1-x)}{2x},$$

where $s(x) = \sqrt{1-4x}$ as before. The expansion of this function informs us that the number of such structures of size j is given by

$$V_j = 3C_{j-1}\frac{j-1}{j+1}, \quad j \geqslant 2. \tag{7}$$

Also $V_0 = V_1 = 0$, as is evident. Observing all possible configurations of the side, we see that when $T_m = k$ the other $m - k$ edges can be used in V_{m-k} ways, hence

$$\mathbb{P}(T_m = k) = \frac{V_{m-k}}{C_{m-1}},$$

which is the same as the statement of the theorem. ■

The asymptotics of the Catalan numbers provide us with local limit theorems.

COROLLARY 1: *As $n \to \infty$, we have*

$$\mathbb{P}(W_n = j) = \frac{C_{j-1}}{4^j}\left(1 + \frac{3j}{n} + O\left(j^2\right)\right), \quad j = o(n),$$

and as $m \to \infty$, we have

$$\mathbb{P}(T_m = k) = \frac{3}{4^k}\left(1 + \frac{3k}{m} + O\left(k^2\right)\right), \quad k = o(m).$$

As we can write a convergence relation for the probability distribution of T_m in the form

$$\mathbb{P}(T_m = k) \to \frac{3}{4}\left(\frac{1}{4}\right)^{k-1},$$

we surmise that T_m converges in distribution to a geometric random variable with rate of success $3/4$.

5. THE ORDER OF BSP GRAPHS

Binary trees and BSP graphs share many aspects, beyond being counted by the Catalan number sequence, but they differ when we consider both nodes and edges. A binary tree of n edges has exactly $n + 1$ nodes, deterministically, but a BSP graph has more variability: n edges can never accommodate more than n nodes, but they can live in a BSP graph with many fewer, down to barely more than $n/2$ nodes; the reason was seen above: Whenever the step in the evolution of a BSP graph is parallelization, an edge is added, but not a node. Consider the 14 BSP graphs of size 5 above; the first four graphs are of order 5, and the rest are of order 4: As we shall see, larger graphs exhibit a richer choice.

For the counting, we return to the structural relation (2), and fit it with a *bivariate* ordinary generating function, in x and u, to mark edges and nodes, respectively. Let $h_{n,m}$ denote the number of distinct BSP graphs of size n having m nodes. Accordingly, $\psi_B(x,u) = \sum_{n=1}^{\infty}\sum_{m=1}^{n-1} h_{n,m}x^n u^m$. Naturally, $\psi_B(x,1) = \varphi_B(x)$.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

The accounting for both features requires some refinement in the components we use to describe the structures. The symbol a' stands as above for an edge with two end nodes (K_2), and the generating function xu^2. When a' drops its South pole, we have the simpler a, and a generating function xu, and $\overline{a}$ is a bare edge (shorn of both end nodes), with the generating function x. Similarly b' is a complete bubble structure, and b denotes a b' "missing" its South pole. Here is how we use these symbols to assemble all possible BSP graphs:

$$B \models a' + b'; \quad b \models a(a+b)^\star \times \overline{a}(a+b)^\star; \quad b' \models b \times \text{South pole}.$$

Notice how the closure gives rise to complete structure: Both a and b are without a South pole, and when an a, say, follows a b, the North pole of the a is the needed completion of the b above it, to a meaningful b'.

This translates to the following equation:

$$\begin{aligned} \psi_B(x,u) &= xu^2 + \psi_{b'}(x,u); \quad \psi_{b'}(x,u) = u\psi_b(x,u); \\ \psi_b(x,u) &= \frac{x^2 u}{(1 - xu - \psi_b(x,u))^2}. \end{aligned} \tag{8}$$

Again a cubic equation for $\psi_b(x,u)$, but unlike its single-variable version, no reasonable-to-handle solution is evident. The computer algebra system MAPLE could be coaxed to present the solutions, ungainly expressions that they be, and one among the three was found to provide the following prefix of its expansion at $x = 0$: $\varphi_B(x,u) = xu^2 + x^2u^2 + 2x^3u^3 + x^4\left(3u^4 + 2u^3\right) + x^5\left(4u^5 + 10u^4\right) + x^6\left(5u^6 + 30u^5 + 7u^4\right) + x^7\left(6u^7 + 70u^6 + 56u^5\right) + \cdots$, where the coefficient of x^5 corresponds to the diagrams in Figure 1.

While the distribution of the order of a random BSP graph is not available explicitly, moments are interesting, and computing them poses no difficulty. First we compute the mean number of nodes in a BSP graph of size n. We get this value by differentiating $\psi_B(x,u)$ with respect to u and evaluating at $u = 1$, since $[x^n](\partial/\partial u)\psi_B(x,u)|_{u=1}$ is the total number of nodes in all the BSP graphs of size n. Unlike the value of $\varphi_B(x,u)$ itself, the derivatives with respect to u can be calculated simply, solving a linear equation for each order; our needs are satisfied by the first two derivatives.

THEOREM 3: *Let ω_n be the order of a random* BSP *graph of size n. The mean and variance of this random variable are given by the asymptotic values*

$$\mathbb{E}[\omega_n] = \frac{2}{3}n + \frac{26}{27} + o(1); \quad \mathbb{V}\mathrm{ar}[\omega_n] = \frac{4}{81}n + o(1).$$

PROOF: We use the values from (8) and find $(\partial/\partial u)\psi_B(x,u) = 2xu + \psi_b(x,u) + (\partial/\partial u)\psi_b(x,u)$. For the last part we differentiate the last relation in (8), denoting $\psi_b(x,u)$ by ψ_b, and $(\partial/\partial u)\psi_b(x,u)$ by ψ_u for conciseness, and get

$$\psi_u = \frac{x^2(1 - xu - \psi_b) + 2x^2u(x + \psi_u)}{(1 - xu - \psi_b)^3} = x^2\frac{1 + xu - \psi_b + 2u\psi_u}{(1 - xu - \psi_b)^3}.$$

This is a linear equation in ψ_u. Using (3) and $s(x) = \sqrt{1 - 4x}$, we find

$$\psi_u\big|_{u=1} = \frac{x(3s(x) - (2x+3)(1-4x))}{2(2+x)(1-4x)} \implies (\psi_B)_u(x,1) = \frac{1+x}{2+x} - \frac{1 - 5x - 2x^2}{s(x)(2+x)}. \tag{9}$$

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

Obtaining the coefficient of x^n in the above is routine – and for $n \geqslant 1$ we have

$$[x^n]\,(\psi_B)_u\,(x,1) = \left(-\frac{1}{2}\right)^{n+1} - \frac{1}{2}\binom{2n}{n} + \frac{11}{4}\binom{2n-2}{n-1} - \frac{3}{2}\sum_{k=2}^{n}\left(-\frac{1}{2}\right)^{k}\binom{2(n-k)}{n-k}.$$

No closed form is likely, but the second, main term in (9) is suitable for estimation via, for example, Theorem 3 of [2, Section 6.2.2]; the dominant pole is that of $s(x)$, at $x = 1/4$; this is the ratio of successive terms, and evaluating the factor $(1-5x-2x^2)/(2+x)$ at $x = 1/4$ provides $1/6$. Since $1/s(x)$ expands as $\sum_{j=0}^{\infty}\binom{2j}{j}x^j$, our estimate of the total number of nodes in all C_{n-1} BSP graphs of size n is $\frac{1}{6}\binom{2n}{n}$. Since the number of such BSP graphs is C_{n-1} we have our mean, by dividing the two, giving the leading term.

A better estimate is possible by rearranging the terms of the above series. The coefficients of the main term in (9) can be written

$$-[x^n]\frac{1-5x-2x^2}{s(x)} = -\left(\binom{2n}{n} - 5\binom{2n-2}{n-1} - 2\binom{2n-4}{n-2}\right) = \frac{3}{2}\binom{2n}{n}\frac{n^2-2}{(2n-3)(2n-1)},$$

and when convolved with contributions from the factor $1/(2+x)$, (excluding terms which make no contribution to large n) the total number of nodes in BSP graphs of size n is given by

$$\frac{1}{2}\sum_{k=0}^{n}\frac{3}{2}\binom{2k}{k}\frac{k^2-2}{(2k-3)(2k-1)}\left(-\frac{1}{2}\right)^{n-k}.$$

Still, no closed form is available, but by letting k range on the last few terms only (to get the main contribution), we get a value that when divided by C_{n-1} is asymptotic to $2n/3 + 26/27$.

One more derivative with respect to u is needed to compute the variance, and it provides at, $u = 1$,

$$\psi_{b_{uu}} \stackrel{\text{def}}{=} \psi_{b_{uu}}(x,1) = \frac{x}{(2+x)^3}\left(1 - x^2 + \frac{(1+2x)(11x^2+8x-1)}{s^3(x)}\right).$$

We have $\psi_{B_{uu}} = 2x + 2\psi_u + \psi_{b_{uu}}$, and adding the number of nodes, $\psi_{B_u}(x,1)$, we get the sum of squares of node counts in all graphs. The main term here is $\psi_{b_{uu}}$, and the leading term there is

$$\frac{x(1+2x)(11x^2+8x-1)}{s^3(x)(2+x)^3}.$$

Using the approach as above, we again see that $x/s^3(x)$ has the dominant singularity. The rest of the term, evaluated at $x = 1/2$, equals $2/9$. Since $2x/s^3(x) = \sum_{j=0}^{\infty}\binom{2j}{j}jx^j$, our leading term estimate for $\mathbb{E}[\omega_n^2]$ is $\frac{n}{9}\binom{2n}{n}$; dividing by C_{n-1}, we find its contribution to $\mathbb{E}[\omega_n^2]$ is $\frac{4}{9}n^2$. When computing the variance, this $\frac{4}{9}n^2$ cancels with the main term of $\mathbb{E}^2[\omega_n]$. The pole at $x = -2$ provides several exponentially smaller corrections (to what are integer coefficients); they can be disregarded. The lower-order terms available from the singularities of $s(x)$ are similar to the series above for the first moment. They are at most linear in n, and we have not computed them, going instead for a direct numerical evaluation, by observing the coefficients obtained from expanding the functions above. This led to the very tight estimate $\mathbb{V}\mathrm{ar}[\omega_n] = \frac{4}{81}n\,(1+o(1))$. ■

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

The relatively small variance is indicative of sharp concentration.

COROLLARY 2: *As $n \to \infty$, we have*

$$\frac{\omega_n}{n} \xrightarrow{\text{P}} \frac{2}{3}.$$

PROOF: By Chebyshev's inequality, we have

$$\mathbb{P}\big(\big|X_n - \mathbb{E}[\omega_n]\big| > \varepsilon \mathbb{E}[\omega_n]\big) \leqslant \frac{\mathbb{V}\text{ar}[\omega_n]}{\varepsilon^2 \mathbb{E}^2[\omega_n]} \sim \frac{4n/81}{4\varepsilon^2 n^2/9} = \frac{1}{9\varepsilon^2 n} \to 0.$$

Therefore, $\frac{\omega_n}{\mathbb{E}[\omega_n]} \xrightarrow{\text{P}} 1$. According to Theorem 3, we also have $\mathbb{E}[\omega_n]/n \to 2/3$. Taking the product of the two convergence relations, the corollary follows. ■

The order of a graph lumps all nodes: Those that have outdegree one with those that are the North poles of bubbles, and have an outdegree of two. The analysis in the next section can be adapted to count separately each type.

6. PATH LENGTH IN BSP GRAPHS

One can take a random walk on a given BSP graph from the North pole to the South one following different strategies. No decision needs to be made when a node of outdegree 1 is reached: Follow the edge coming out of the node. But at each node of outdegree 2 a decision is needed. For example, in the absence of any knowledge of the structure of the graph, the walker can use a Bernoulli random variable to reach a decision, with or without dependence on previous decisions. A line of research that we do not follow here is the range of strategies available to the walker with advance information about the graph structure.

Let L'_n be the length of a path from the North pole to the South in a BSP graph of size n. The length is counted by edges traversed, and all edges are taken *down*. We want its distribution, assuming all graphs of a given size are equally likely and the implied distribution of edges between the East and West sides is as given by Theorem 2, part 1.

The calculations below use the symbol L_n for the path length through a bubble-BSP graph, the structures denoted in Section 3 by the symbol b. The symbol L_n is only defined for $n \geqslant 2$.

6.1. Going Down a Random Path

An agent traveling down the path makes a probabilistic choice at each possible decision point. We can posit probabilities $p_w,\ p_e = 1 - p_w$ to denote the preference of pursuing each direction. The uniformity of our distribution over the set of each equal-sized graphs also induces symmetry with respect to the left and right sides of each graph. As a result, all random variables evaluated on complete sets of graphs (all BSP graphs of a given size), are invariant with respect to this navigation probability. Our computations, however, are more natural with the assumption $p = 1/2$. We note in passing that this invariance means that the same results would be obtained, if the agent always picked the West branch (or always the East one).

We refer to L_n as a random variable using the same language we used for ω_n, but note that these two variables live on quite different spaces. The order ω_n gets its values from counting nodes on each of the BSP graphs, and we compute moments by accumulating those

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

counts and dividing by C_{n-1}. The path length is not yet determined when a specific BSP graph is chosen; we have the random decision taken each time a node with outgoing degree two is reached.

THEOREM 4: *The bivariate, unnormalized probability generating function of the path lengths in* BSP *graphs is given by*

$$\Lambda(x,z) = \frac{1}{2}\left(1 - xz - \sqrt{\big(s(x)+xz\big)^2 + 4x(1-z)}\right).$$

The prefix of expanding this function reads:

$$\begin{aligned}\Lambda(x,z) =& x^2z + x^3(z^2+z) + x^4(z^3+2z^2+2z) + x^5(z^4+4z^3+4z^2+5z)\\ &+ x^6(z^5+7z^4+10z^3+10z^2+14z)\\ &+ x^7(z^6+11z^5+24z^4+26z^3+28z^2+42z) + \cdots .\end{aligned}$$

The coefficient of x^5 can be read off the diagrams in Figure 1. To normalize the distributions generated by this function we need to divide the polynomial accompanying x^n by C_{n-1}.

PROOF: A few more variables need to be defined. In a BSP graph of size n, once a side of size m is entered, the path will have the length K_m. This path is initiated with a sequence of T_m edges, with the distribution seen in Theorem 2. Once it reaches a node with two outgoing edges it enters the part of the side we denoted by V, which has now the size $m - T_m$. The path length through a V structure of size j is denoted by Q_j. The structure of this variable is interesting – it echoes that of V itself. Just as the characterization of V was $V \models b(a+b)^\star$, we have $Q \models L(1+L)^\star$, where L is the path length through a complete bubble-BSP graph (of size two or more), the variable we started with.

One further notation is needed: The convolution of k independent identically distributed random variables $X_1, X_2, \ldots, X_k$ is written as

$$X^{(k)} \stackrel{\text{def}}{=} X_1 + X_2 + \cdots + X_k.$$

It is not evident how to write a relation for the distribution of L directly. We pursue a strategy to connect L with Q, then Q with K, and finally K with L again. We derive the three relations one at a time.

First, we derive the connection between L and Q. The representation of Q as $L(1+L)^*$ is used, to describe the possible paths down a structure V. We remark that this representation can be read as leading directly to the result in (13), yet we show it in some detail. The path consists of $r+1$ parts, for some $r \geqslant 0$: an initial L, and r that are either single edges (when "1" is selected from $1+L$), or additional instances of L. We use the letter i here to denote the number of times (out of r) that the single edge is selected. While r is not formally bounded above, it is clearly at most the size of V. The total length traversed in the bubble is the convolution of i single edges and $L_{m-i}^{(r-i+1)}$, the length of path through $r-i+1$ BSP bubbles on the way down the chosen side, using a total $m-i$ edges. Since we count each graph equally likely, we multiply the path length probability of each of the $r-i+1$ bubble paths by its population size, as given in Theorem 1. This requires an elaboration of the convolution notation: We write the multiplier as $C_{m-i}^{(n-r+1)}$, when attached to the above convolution of L; finally, we divide by V_m, the total number of configurations of

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

the structure that Q_m navigates. We write

$$\mathbb{P}(Q_m = t) = \frac{1}{V_m} \sum_{r=0}^{\infty} \sum_{i=0}^{r} \binom{r}{i} \sum_{(m-i)|[r-i+1]} C_{m-i}^{(r-i+1)} \mathbb{P}(L_{m-i}^{(r-i+1)} = t - i).$$

The innermost summation accounts for all the ways we can allocate $m - i$ edges among the $r - i + 1$ bubbles. A more explicit expression for this allocation, where the jth bubble is allocated $s_j \geqslant 2$ edges, yields

$$V_m \mathbb{P}(Q_m = t) = \sum_{r=0}^{\infty} \sum_{i=0}^{r} \binom{r}{i} \sum_{\substack{s_1, s_2, \ldots, s_{r-i+1} \geqslant 2 \\ s_1 + s_2 + \cdots + s_{r-i+1} = m-i}} \left(\prod_{j=1}^{r-i+1} C_{s_j - 1} \right) \mathbb{P} \left(\sum_{k=1}^{r-i+1} L_{s_k} = t - i \right). \tag{10}$$

Much simplification accrues from the following natural abbreviations:

$$\lambda_{m,t} \stackrel{\text{def}}{=} C_{m-1}\, \mathbb{P}(L_m = t), \qquad \eta_{m,t} \stackrel{\text{def}}{=} V_m\, \mathbb{P}(Q_m = t), \tag{11}$$

and the related generating functions:

$$\Lambda(x, z) = \sum_{m=1}^{\infty} \sum_{t=1}^{\infty} \lambda_{m,t} x^m z^t = \sum_{m=1}^{\infty} \lambda_m(z) x^m, \quad H(x, z) = \sum_{m=2}^{\infty} \sum_{t=1}^{\infty} \eta_{m,t} x^m z^t. \tag{12}$$

Note that the definition leads to $\Lambda(x, 1) = \varphi_b(x)$.

Multiplying (10) by $x^m z^t$ and summing for $1 \leqslant t \leqslant m$, we obtain on the left-hand side $H(x, z)$, and the right-hand side needs rearranging:

$$\sum_{r=0}^{\infty} \sum_{i=0}^{r} \binom{r}{i} (xz)^i \sum_{\substack{s_1, s_2, \ldots, s_{r-i+1} \geqslant 2 \\ s_1 + s_2 + \cdots + s_{r-i+1} = m-i}} \sum_{\substack{u_1, u_2, \ldots, u_{r-i+1} \geqslant 1 \\ u_1 + u_2 + \cdots + u_{r-i+1} = t-i}} \left(\prod_{k=1}^{r-i+1} \lambda_{s_k, u_k}\, x^{s_k} z^{u_k} \right)$$

$$= \sum_{r=0}^{\infty} \sum_{i=0}^{r} \binom{r}{i} (xz)^i\, \Lambda^{r-i+1}(x, z).$$

The summation on t and its fragments u_k is an ordinary convolution of independent random variables. The summation over m and its fragments s_k does not have the same deep interpretation, but shares, as the definition of $\Lambda(x, z)$ in (12) shows, the same happy computational mechanism.

Summing over i, we find for this right-hand side

$$\Lambda(x, z) \sum_{r=0}^{\infty} (xz + \Lambda(x, z))^r = \frac{\Lambda(x, z)}{1 - xz - \Lambda(x, z)}.$$

Hence, we have the relation

$$H(x, z) = \frac{\Lambda(x, z)}{1 - xz - \Lambda(x, z)}. \tag{13}$$

So far, we have obtained the relation between L and Q in terms of their generating functions.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

Next we connect Q and K. We write the distribution of K_m in the form

$$\mathbb{P}(K_m = k) = \mathbb{P}(T_m + Q_{m-T_m} = k) = \sum_r \mathbb{P}(T_m = r)\,\mathbb{P}(Q_{m-r} = k - r).$$

Using the distribution of T_m as given in (6) we write

$$\mathbb{P}(K_m = k) = \frac{\delta_{k,m}}{C_{m-1}} + \frac{3}{C_{m-1}} \sum_{r=1}^{m-1} C_{m-r-1} \frac{m-r-1}{m-r+1}\,\mathbb{P}(Q_{m-r} = k - r).$$

Multiplying by C_{m-1}, and replacing $m - r$ by j leads to

$$C_{m-1}\,\mathbb{P}(K_m = k) = \delta_{k,m} + 3 \sum_{j=1}^{m-1} C_{j-1} \frac{j-1}{j+1}\,\mathbb{P}(Q_j = k - m + j).$$

Using (11) and (7), the sum is over $\eta_{j,k-m+j}$. For further simplification, we add the notation

$$\rho_{m,t} \stackrel{\text{def}}{=} C_{m-1}\,\mathbb{P}(K_m = t), \quad R(x,z) = \sum_{m=1}^{\infty} \sum_{t=1}^{m} x^m z^t \rho_{m,t}.$$

The abbreviations allow us to write the last relation as $\rho_{m,k} = \delta_{m,k} + \sum_{j=1}^{n-1} \eta_{j,k-m+j}$, which we use to obtain a relation between $K(x,z)$ and $H(x,z)$, multiplying $\rho_{m,k}$ by $x^m z^k$ and summing over $1 \leqslant k \leqslant m$. The Kronecker delta contributes $xz/(1-xz)$, and the rest contribute as follows:

$$\begin{aligned}
R(x,z) &\leftarrow \sum_{m=1}^{\infty} \sum_{k=1}^{m} x^m z^k \sum_{j=1}^{m-1} \eta_{j,k-m+j}, \quad \text{now replacing } m-k \text{ by } i \\
&= \sum_{m=1}^{\infty} \sum_{i=0}^{n-1} x^m z^{m-i} \sum_{j=1}^{m-1} \eta_{j,j-i}, \quad \text{now replacing } j-i \text{ by } r \\
&= \sum_{m=1}^{\infty} x^m \sum_{j=1}^{m-1} z^{m-j} \sum_{r=0}^{j} z^r \eta_{j,r}, \quad \text{we denote the inner sum by } \eta_j(z) \\
&= \sum_{j=1}^{\infty} \eta_j(z) z^{-j} \sum_{m=j+1}^{\infty} (xz)^m \\
&= \frac{xz}{1-xz} \sum_{j=1}^{\infty} x^j H_j(z) \\
&= \frac{xz}{1-xz}\,H(x,z).
\end{aligned}$$

Hence, we have

$$R(x,z) = \frac{xz}{1-xz}\big(1 + H(x,z)\big), \tag{14}$$

relating K and Q in terms of their generating functions.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

To complete the derivation we need to define a relation between L, and the path length K, once a side is chosen. Let W be the number of edges in the left edge, with distribution given in (5), and write

$$\mathbb{P}(L_m = t) = \frac{1}{2}\sum_w \mathbb{P}(W = w)\big(\mathbb{P}(K_w = t) + \mathbb{P}(K_{m-w} = t)\big),$$

that leads to

$$2C_{m-1}\,\mathbb{P}(L_m = t) = \sum_{w=1}^{n-1}\left[C_{n-w-1}\left(C_{w-1}\,\mathbb{P}(K_w = t)\right) + C_{w-1}\left(C_{m-w-1}\,\mathbb{P}(K_{m-w} = t)\right)\right],$$

and with the above abbreviations, we can write

$$2\lambda_{m,t} = \sum_{w=1}^{m-1}\big(C_{m-w-1}\rho_{w,t} + C_{w-1}\rho_{m-w,t}\big).$$

These are convolutions on the right-hand side. Let $\mathcal{L}_t(x) = \sum_{m=1}^{\infty} x^m\lambda_{m,t}$, $\mathcal{R}_t(x) = \sum_{m=1}^{\infty} x^m\rho_{m,t}$, and we have found that

$$\mathcal{L}_t(x) = \varphi_B(x)\mathcal{R}(x) \implies \Lambda(x,z) = \varphi_B(x)\rho(x,z), \tag{15}$$

where $\varphi_B(x)$ is as given in (3).

Combining (14) and (15) we find

$$\Lambda(x,z) = \varphi_B(x)\frac{xz}{1-xz}\big(1+\eta(x,z)\big).$$

Substituting relation (13) leads to a quadratic equation for the weighted probability generating function $\Lambda(x,z)$:

$$\begin{aligned}&\Lambda(x,z)^2 - \Lambda(x,z)(1-xz) + xz\varphi_B(x) = 0\\ &\implies \quad \Lambda(x,z) = \frac{1}{2}\left(1 - xz - \sqrt{\big(s(x)+xz\big)^2 + 4x(1-z)}\right).\end{aligned}$$

■

As expected, $\Lambda(x,1) = \varphi_b(x)$. The corresponding generating function of L', which we denote by $\Lambda'(x,z)$, over all BSP graphs, has the additional term xz.

6.2. Moments of the Path Length

Turning to moment calculation: The first derivative of $\Lambda'(x,z)$ is

$$\Lambda'_z(x,z) = \frac{x}{2} - \frac{x(xz+s(x)) - 2x}{2\sqrt{(s(x)+xz)^2 + 4x(1-z)}}.$$

At the point $z = 1$, we find $\Lambda'_z(x,1) = x/(x+s(x))$. No explicit representation is available for the expansion of this function. It is listed in the on-line encyclopedia of integer

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

sequences [4] as sequence A081696. The first few, with the coefficient of x^j divided by C_{j-1}, provide the prefix, for small graphs, of the generating function for the expected path lengths:

$$x + x^2 + \frac{3}{2}x^3 + \frac{9}{5}x^4 + \frac{29}{14}x^5 + \frac{97}{42}x^6 + \frac{333}{132}x^7 + \frac{1,165}{429}x^8 + \frac{4,135}{1,430}x^9 + \frac{14,845}{4,862}x^{10} + \frac{53,791}{16,796}x^{11} + \cdots .$$

The same series may be more transparent as follows: $x + x^2 + 1.5x^3 + 1.8x^4 + 2.071x^5 + 2.3095x^6 + 2.5227x^7 + 2.7156x^8 + 2.8916x^9 + 3.0532x^{10} + 3.2026x^{11} + \cdots$.

The coefficient of x^5, $29/14 \approx 2.071$, can be read-off the diagrammed graphs in Figure 1.

For larger graphs we look at an asymptotic expansion of the function by writing

$$\frac{x}{x + s(x)} = \frac{1}{1 + s(x)/x} = 1 - \frac{s(x)}{x} + \frac{s^2(x)}{x^2} - \frac{s^3(x)}{x^3} + \frac{s^4(x)}{x^4} + \cdots .$$

The fragment $[x^n](-s(x)/x)$ provides $2C_n$. The even-powers fragments provide only few terms, and none contributes beyond the constant term. The odd-powers fragments contribute terms with decreasing powers of n. In general, for $j = 2k - 1$, we find that for small j the convenient form is

$$[x^n]s^j(x) = \binom{j/2}{n}(-4)^n = (-1)^k\binom{2n}{n}\frac{\binom{n}{k}}{\binom{2n}{2k}}, \quad j = 2k - 1.$$

The first $-s^3(x)/x^3$ provides the main error term $-12C_{n+1}/(n + 3)$. We need yet to divide the term $[x^n]x/(x + s(x))$ by C_{n-1}. From (1) we can write $2C_n/C_{n-1} = 8(1 - 3/2n + o(1/n))$, and we finally obtain

$$\mathbb{E}[L_n] = \frac{[x^n]\frac{x}{x+s(x)}}{C_{n-1}} = 8\left(1 - \frac{51}{2n} + o\left(n^{-1}\right)\right).$$

The average path length approaches a constant! A few values obtained with the help of MAPLE are shown in the following table.

n	50	100	200	400	800	1600	3200
$\mathbb{E}[L_n]$	5.6446	6.5303	7.1519	7.5392	7.7578	7.8758	7.9371

This is indeed a surprise, since the calculations made in [3] under the assumptions that each of the $(n - 1)!$ distinct generation sequences is equally likely, found the expected path length grows without bound, at the asymptotic rate of $n^{\phi-1} = n^{0.6180339880\ldots}$ (here ϕ is the golden ratio $(1 + \sqrt{5})/2$). Obviously, many random BSP graphs have large expected path lengths, yet under "our" measure their contribution is limited, and the average length of a random walk is bounded by constant, even in arbitrarily large BSP graphs. Observe that for large n, $\mathbb{P}(L_n = 1) = 1/4$, and $\mathbb{P}(L_n = 2)$ is close to $1/8$ under this measure. This furnishes some insight into this quite different behavior. On the other hand, these probabilities are in $O(n^{-1})$ under the measure used in [3]. In the running analogy to binary trees, under the uniform model about a fourth of them have an empty left subtree, whereas under the binary search tree model the number of nodes in the left subtree grows to infinity almost surely.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

Evaluating the second derivative, directly at $z = 1$ we find

$$\Lambda'_{zz}(x,1) = \frac{2x^2(1-x-s(x))}{(x+s(x))^3} = 2x^2\left(\frac{1}{(x+s(x))^3} - \frac{1}{(x+s(x))^2}\right).$$

Adding the generating function of $C_{n-1}\,\mathbb{E}[L_n]$, found as $x/(x+s(x))$ we find that the generating function of $C_{n-1}\mathbb{E}[L_n^2]$ is

$$\begin{aligned} v(z) &\stackrel{\text{def}}{=} \Lambda'_{zz}(x,1) + \Lambda'_z(x,1) = 2x^2\left(\frac{1}{(x+s(x))^3} - \frac{1}{(x+s(x))^2}\right) + \frac{x}{x+s(x)} \\ &= \frac{x-2x^2-x^3}{(x+s(x))^3}. \end{aligned}$$

Extraction of coefficients is done by writing

$$[x^n]\big(x+s(x)\big)^{-3} = \sum_{j=0}^{\infty}\binom{-3}{j}[x^{n+3+j}]s^j(x).$$

As for the first moment, only odd j values matter, and since the ratio between successive terms is n, two were used (writing $j = 2k-1$, we used the values 1 and 2 for k). The explicit expression is not very presentable, but for large n we used MAPLE to expand them at $n = \infty$, and find

$$\mathbb{E}[L_n^2] = 168\left(1 - \frac{1213}{14n} + o\left(n^{-1}\right)\right), \quad \mathbb{V}\mathrm{ar}[L_n] = 104\left(1 - \frac{2823}{26n} + o\left(n^{-1}\right)\right).$$

As for the first moment, the graph size has to increase well into the hundreds, before the last relations give a close estimate of the actual values. The so-called "variance ratio," $\sqrt{\mathbb{V}\mathrm{ar}[L_n]}/\mathbb{E}[L_n] \to 1.274755$ confirms intuitive impressions of rare but extravagant deviations from values close to the mean.

7. NOTES AND OPEN PROBLEMS

The calculations in Section 6 were made under the assumption that the path is taken with no advance knowledge of the detailed structure of the graph. One can ask about the experience of an enlightened traveler, who approaches such a graph with a complete map (or can visualize its entire geometry on sight). What path distribution would such a traveler find, if he wants the shortest path? What if the agent needs a dilatory strategy, traversing the longest possible route? How would he fare? For $n = 5$ we found above that the "mapless" agent has an average route of 29/14 edges. With complete information, the average shortest path is 18/14 edges long; in particular, in 10 out of the 14 graphs there is a path of size one! The traveler that wishes to maximize his travel will traverse an average of three edges, more than double the minimal average, and nearly 45% higher than the uninformed explorer. How do these value change with n? And in particular, how do those ratios evolve?

At the end of Section 5 we observe that the calculation of the path-length can be extended with little effort – by adding a variable in the equation following (12) to mark the counter i used in (10) – and provide separate counts of the nodes according to their outdegree, which is limited to the values one and two. The corresponding question about indegrees, when we look at nodes in their roles as South poles in the graphs, requires a different approach.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press

A different type of questions arise when the BSP graphs are considered as a realistic model for transportation. Such an application requires two constraints on the graphs, respecting practical limitations: The indegree of South pole nodes need to be capped, due to physical road layout constraints. In addition, the allocation of edges to the two sides may not be extreme, possibly limiting the difference to a small fraction of the graph size.

References

1. Bodirsky, M., Giménez, O., Kang, M., & Noy, M. (2007). Enumeration and limit laws of series-parallel graphs. *European Journal of Combinatorics* 28: 2091–2105.
2. Hofri, M. (1995). *Analysis of algorithms: computational methods & mathematical tools*. New York: Oxford University Press.
3. Mahmoud, H.M. (2014). Some properties of binary series-parallel graphs. *Probability in the Engineering and Informational Sciences* 28: 565–572. doi:10.1017/S0269964814000138.
4. On-line Encyclopedia of Integer sequences. Accessed at www.oeis.org
5. Sedgewick, R. & Flajolet, P. (2013). *An introduction to the analysis of algorithms*, 2nd edn., Upper Saddle Valley, New Jersey: Addison-Wesley.
6. Stanley, R. (1999). *Enumerative combinatorics*, Vol. 2. New York: Cambridge University Press.

https://doi.org/10.1017/S0269964815000364 Published online by Cambridge University Press