

## ON THE ENUMERATION OF INTERVAL GRAPHS

JOYCE C. YANG AND NICHOLAS PIPPENGER

(Communicated by Patricia Hersh)

ABSTRACT. We present upper and lower bounds for the number  $i_n$  of interval graphs on  $n$  vertices. Answering a question posed by Hanlon, we show that the ordinary generating function  $I(x) = \sum_{n \geq 0} i_n x^n$  for the number  $i_n$  of  $n$ -vertex interval graphs has radius of convergence zero. We also show that the exponential generating function  $J(x) = \sum_{n \geq 0} i_n x^n / n!$  has radius of convergence at least  $1/2$ .

### 1. INTRODUCTION

An undirected graph is an *interval graph* if there is a one-to-one correspondence between its vertices and a set of intervals of the real numbers such that two vertices are adjacent if and only if their corresponding intervals overlap (that is, have a non-empty intersection). The concept of an interval graph appears first to have been formulated in 1957 by Hajós [2], who posed the problem of determining whether a given graph is an interval graph. In 1959, the biologist Benzer [1] independently reformulated the problem in connection with the question of whether the observed overlaps among gene fragments were compatible with the hypothesis that the structure of the gene is linear. The first intrinsic characterization of interval graphs was given in 1962 by Lekkerkerker and Boland [4], who showed that a graph is an interval graph if and only if it is *chordal* (that is, every cycle of length at least four has a chord) and *asteroidal* (that is, among every three vertices, there is one that is adjacent to every path between the other two). They also gave a characterization in terms of an infinite family of forbidden induced subgraphs.

In 1982, Hanlon [3] addressed the problem of enumerating interval graphs (that is, determining the number  $i_n$  of (isomorphism classes of) interval graphs with  $n$  vertices). He solved this problem by giving a system of equations defining a set of formal power series including the (ordinary) generating function  $I(x) = \sum_{n \geq 0} i_n x^n$  for interval graphs. This implicit enumeration allowed him to tabulate  $i_n$  for  $n$  up to 30 ( $i_{30}$  has 27 decimal digits), but it did not give an explicit formula for  $i_n$  or even allow determination of its asymptotic behavior. Indeed, Hanlon posed the question of whether the radius of convergence of  $I(x)$  is positive (that is, whether  $i_n$  is eventually bounded by  $C^n$  for some constant  $C$ ). (In Hanlon’s tabulation, the entry  $i_4 = 10$  is missing, and the entry  $i_5 = 27$  is erroneously given as  $i_5 = 10$ . The correct values are given as sequence A005975 by Sloane [5].)

In Section 2, we shall give a lower bound to  $i_n$  that grows “factorially” and therefore shows that the radius of convergence of  $I(x)$  is zero. In Section 3, we

---

Received by the editors September 24, 2016 and, in revised form, October 24, 2016.

2010 *Mathematics Subject Classification*. Primary 05C30.

*Key words and phrases*. Interval graphs, enumeration.

shall give an upper bound to  $i_n$  that shows that the exponential generating function  $J(x) = \sum_{n \geq 0} i_n x^n / n!$  has a radius of convergence that is at least  $1/2$ .

## 2. LOWER BOUND

In this section, we shall prove the lower bound

$$(2.1) \quad i_{3k} \geq k! / 3^{3k}.$$

Since  $i_n$  is non-decreasing, (2.1) implies that the coefficients in the formal power series  $I(x)$  are bounded below by those of  $(1 + x + x^2) \sum_{k \geq 0} k! x^{3k} / 3^{3k}$ . Since this last power series has radius of convergence zero, so does  $I(x)$ , thus answering Hanlon's question.

To prove (2.1), we shall associate with each permutation  $\pi$  of the set  $\{1, \dots, k\}$  a 3-colored  $(3k)$ -vertex interval graph  $G_\pi$ . This association will be one-to-one (that is, the permutation  $\pi$  can be recovered from the colored graph  $G_\pi$ ). Since there are only  $i_{3k} 3^{3k}$  distinct 3-colored  $(3k)$ -vertex interval graphs, we have  $i_{3k} 3^{3k} \geq k!$ , which is equivalent to (2.1).

Let  $\pi$  be a permutation of  $\{1, \dots, k\}$ . We shall construct  $3k$  intervals, with endpoints  $1, \dots, 6k$ . First we construct  $k$  red intervals,  $R_j = [3j - 2, 3j]$  for  $1 \leq j \leq k$ . Then we construct  $k$  blue intervals,  $B_j = [3k + 3j - 2, 3k + 3j]$  for  $1 \leq j \leq k$ . Finally, we construct  $k$  white intervals,  $W_j = [3j - 1, 3k + 3\pi(j) - 1]$  for  $1 \leq j \leq k$ . Let  $G_\pi$  be the interval graph whose colored vertices  $r_1, \dots, r_k, b_1, \dots, b_k, w_1, \dots, w_k$  correspond to the colored intervals just constructed. It remains to show that the permutation  $\pi$  can be recovered from the colored graph  $G_\pi$ . Let us define the *red degree*  $\deg_R(w)$  of a white vertex  $w$  to be the number of red vertices adjacent to  $w$  and the *blue degree*  $\deg_B(w)$  to be the number of blue vertices adjacent to  $w$ . The interval  $W_j$  intersects  $R_j, R_{j+1}, \dots, R_k$  and  $B_1, B_2, \dots, B_{\pi(j)}$ . The corresponding vertex  $w_j$  thus has  $\deg_R(w_j) = k - j + 1$  and  $\deg_B(w_j) = \pi(j)$ . Thus,  $\pi(j) = \deg_B(w_j)$ , which shows that  $\pi$  can be recovered from  $G_\pi$ .

It is clear that (2.1) could be improved slightly by using only  $k - 1$  red and  $k - 1$  blue intervals, for example, and by using a sharper upper bound to the number of colorings used. But we have not pursued these improvements, as none of them improve the factor 3 in the relation between  $i_{3k}$  and  $k!$ .

## 3. UPPER BOUND

In this section, we shall prove the upper bound

$$(3.1) \quad i_n \leq (2n - 1)!!,$$

where  $(2n - 1)!! = (2n - 1) \cdot (2n - 3) \cdots 3 \cdot 1$ . We have  $(2n - 1)!! \leq 2^n n!$ . The last inequality yields

$$\begin{aligned} J(x) &= \sum_{n \geq 0} i_n x^n / n! \\ &\leq \sum_{n \geq 0} 2^n x^n \\ &= 1 / (1 - 2x), \end{aligned}$$

which implies, because the coefficients of  $J(x)$  are non-negative, that the radius of convergence of  $J(x)$  is at least  $1/2$ . It remains an open question to determine if  $J(x)$  has a larger radius of convergence or is perhaps even an entire function.

To prove (3.1), we observe that the  $2n$  endpoints of the  $n$  intervals in the representation of an  $n$ -vertex interval graph can, without loss of generality, be taken to be the  $2n$  positive integers  $1, \dots, 2n$ . The  $n$  intervals then correspond to a partition of these  $2n$  integers into  $n$  blocks, each containing two integers that are the endpoints of an interval. This partition can be specified by first choosing the mate of 1 from among the  $2n - 1$  integers  $2, \dots, 2n$  (which can be done in  $2n - 1$  ways), then choosing the mate of the smallest as-yet-unpaired integer from among the  $2n - 3$  greater as-yet-unpaired integers (which can be done in  $2n - 3$  ways), and continuing in this way until all  $2n$  integers have been partitioned into  $n$  mated pairs. This can thus be done in  $(2n - 1) \cdot (2n - 3) \cdots 3 \cdot 1 = (2n - 1)!!$  ways. Since every interval graph corresponds to at least one partition, we have established (3.1).

#### 4. CONCLUSION

In view of Stirling's formula, one way of roughly stating our results is

$$\frac{1}{3}n \log n + O(n) \leq \log i_n \leq n \log n + O(n).$$

It remains an open problem to bring the coefficients  $1/3$  and  $1$  in the lower and upper bounds closer together, perhaps even to obtain an estimate of the form

$$\log i_n = C n \log n + O(n)$$

for some constant  $C$ .

#### REFERENCES

- [1] S. Benzer, *On the topology of the genetic fine structure*, Proc. Nat. Acad. Sci. **45** (1959), 1607–1620.
- [2] G. Hajós, *Über eine Art von Graphen*, Int. Math. Nachr. **11** (1957), 65 pp.
- [3] Phil Hanlon, *Counting interval graphs*, Trans. Amer. Math. Soc. **272** (1982), no. 2, 383–426, DOI 10.2307/1998705. MR662044
- [4] C. G. Lekkerkerker and J. Ch. Boland, *Representation of a finite graph by a set of intervals on the real line*, Fund. Math. **51** (1962/1963), 45–64. MR0139159
- [5] N. J. A. Sloane (editor), *The Online Encyclopedia of Integer Sequences*, published electronically at <https://oeis.org> (2016).

DEPARTMENT OF MATHEMATICS, HARVEY MUDD COLLEGE, 301 PLATT BOULEVARD, CLAREMONT, CALIFORNIA 91711

*E-mail address*: [jcyang@hmc.edu](mailto:jcyang@hmc.edu)

DEPARTMENT OF MATHEMATICS, HARVEY MUDD COLLEGE, 301 PLATT BOULEVARD, CLAREMONT, CALIFORNIA 91711

*E-mail address*: [njp@math.hmc.edu](mailto:njp@math.hmc.edu)