INTERVAL NUMBER OF SPECIAL POSETS AND RANDOM POSETS

Tom Madej
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana

Douglas B. West[1]
Department of Mathematics
University of Illinois
Urbana, Illinois 61801

Abstract

The *interval number* $i(P)$ of a poset P is the smallest t such that P is a containment of sets that are unions of at most t real intervals. For the special poset $B_n(k)$ consisting of the singletons and k-subsets of an n-element set, ordered by inclusion, $i(B_n(k)) = \min\{k, n-k+1\}$ if $|n/2 - k| \geq n/2 - (n/2)^{1/3}$. For bipartite posets with n elements or n minimal elements, $i(P) \leq \lceil \frac{n}{\lg n - \lg\lg n} \rceil + 1$. Finally, the fraction of the n-element posets having interval number between $(1-\varepsilon)\frac{n}{8\lg n}$ and $(3/2)(\lceil \frac{n}{\lg n - \lg\lg n} \rceil + 1)$ approaches 1 as $n \to \infty$ (i.e., this involves the Kleitman-Rothschild model of random posets).

Keywords: poset, representation parameter, interval number, containment representation, dimension

Running head: INTERVAL NUMBER OF POSETS

[1]Research supported in part by NSA/MSP Grant MDA904-90-H-4011.

1. INTRODUCTION

In the special issue devoted to the preceding meeting in this series, we introduced [10] the notion of the interval (inclusion) number of a poset P. A *containment representation* of P is an assigment of a set $f(x)$ to each $x \in P$ such that $x < y$ if and only if $f(x) \subset f(y)$. The *interval number* $i(P)$ is the minimum value t such that P has a containment representation f in which each $f(x)$ is the union of at most t intervals on the real line. The results in [10] include $i(P) \le \lceil \dim(P)/2 \rceil$, $i(B_{2k}) = k$ (where B_n denotes the Boolean algebra on n elements), bounds on the interval number for compositions and products of posets, and the fact that testing $i(P) \le k$ for fixed $k \ge 2$ is NP-complete.

In this paper we obtain bounds on the interval number for special subposets of the Boolean algebra and study the interval number for random posets. Both of these investigations were motivated by the question of how large $i(P)$ can be for a poset on n elements. A poset with interval number 3 must have dimension at least 5; the smallest known poset with interval number 3 is the subposet of B_7 consisting of the singletons and 5-sets. Hence in Section 2 we study $i(B_n(k))$, where $B_n(k)$ denotes the subposet of B_n consisting of the singletons and the k-sets. We use a counting argument to prove that if $|n/2 - k| > n/2 - (n/2)^{1/3}$, then a simple construction is optimal and yields $i(B_n(k)) = \min\{k, n-k+1\}$. However, when k is near $n/2$ there are better constructions. The dimension of these posets was studied by Dushnik [1] and by Spencer [14], see also [3] for a modern viewpoint. The results for $i(B_n(k))$ are very much different from those for $\dim(B_n(k))$.

In Section 3, we study the interval number of large posets in terms of their size, for bipartite posets and for random posets. *Bipartite posets* are those whose comparability graph is bipartite. This is equivalent to every element being maximal or minimal. This term avoids the disagreement in terminology over whether such a poset should be said to have height 1 or height 2. We prove that $i(P) \le \lceil n/(\lg n - \frac{1}{2}\lg\lg n) \rceil + 1$ for every n-element bipartite poset P. For random posets, we use the Kleitman-Rothschild model [8], in which all n-element posets are equally likely. We prove that in this model almost every poset has interval number between 1/8 and 3/2 times the bound above for bipartite posets, roughly speaking. Nevertheless, the asymptotic behavior of $\max_{|P| = n} i(P)$ remains open; in particular, we do not know whether it grows linearly or sublinearly with n. Section 4 contains a collection of open problems about interval number of posets.

2. SINGLETONS AND k-SETS

The "standard" example of an n-dimensional poset is the subposet of B_n induced by the singletons and $n-1$-element sets. However, this poset has interval number 2, which is a special case of the following theorem. We use the notation $[n] = \{1, \ldots, n\}$.

THEOREM 1. If $k \le (n/2)^{1/3}$ or $k \ge n - (n/2)^{1/3}$, then $i(B_n(k)) = \min\{k, n-k+1\}$.

Proof. First suppose $k \le n/2$ and put $P = B_n(n-k)$. For each $X \in P$ define $f(X) = \cup_{i \in X}(i-.5, i+.5]$. Clearly f is a containment representation for P. Each singleton receives one interval. Each set X of size $n-k$ has at most k gaps, which means $f(X)$ consists of at most $k+1$ intervals, and hence $i(P) \le k+1$.

To prove $i(P) > k$ if $k \le (n/2)^{1/3}$, we assume the contrary and fix a representation f for P that uses no more than k intervals per element. We call the intervals assigned to the singleton sets the *1-intervals*. Because containment representation is defined using strict inclusion, we may assume that all endpoints are distinct, and hence there is a unique leftmost 1-interval $[a, b]$ and unique rightmost 1-interval $[c, d]$, which by symmetry we may assume are assigned to $n-1$ and n, respectively. Let $I = [a, d]$. Now consider an $(n-k)$-set X such that $\{n-1, n\} \subset X \subset [n]$. Since $[a, b] \cup [c, d] \subset f(X)$, the set $I - f(X)$ is the union of at most $k-1$ intervals; we call these the *gaps* of X. Each of the k elements of $[n] - X$ is assigned a 1-interval intersecting a gap of X. By the pigeonhole principle, some gap of X intersects 1-intervals for distinct singletons.

We choose a label $l(X)$ for each such set X. Consider the first gap of X that intersects 1-intervals for distinct singletons. If some (first) point p within the gap belongs to 1-intervals for distinct singletons, we let $l(X) = \{j\colon\ p \in f(\{j\})\}$. If there is no such point p, we let $l(X)$ consist of the two singletons with leftmost 1-intervals intersecting this gap. In either case we have $l(X) \subseteq [n-2]$, $X \cap l(X) = \emptyset$, and $2 \le |l(X)| \le k$. For a given set $T \subseteq [n-2]$ of size t, at most $\binom{n-2-t}{k-t}$ of the $n-k$-subsets of $[n]$ can be assigned T as their label, since a set X for which $l(X) = T$ must omit T and an additional $k-t$ elements chosen from $[n-2] - T$. Since all labels have size at least 2, we have $\binom{n-2-t}{k-t} \le \binom{n-4}{k-2}$. Let r denote the number of distinct labels. Since we choose labels for the $\binom{n-2}{k}$ sets of size k containing $\{n-1, n\}$, we have $r\binom{n-4}{k-2} \ge \binom{n-2}{k}$, which simplifies to

$$r \ge \frac{(n-2)(n-3)}{k(k-1)}\,. \tag{1}$$

On the other hand, the possibilities for distinct labels are restricted by the configuration of 1-intervals. Consider a candidate point p traveling from left to right through the interval I. For any given position of p, if p belongs to distinct 1-intervals or lies outside all 1-intervals, then there is a potential label consisting of the singletons whose 1-intervals contain p or the singletons corresponding to the 1-intervals nearest to p on each side, respectively. Furthermore, these are the only possible labels. Therefore, from the current position of p, when moving to the right we can only obtain a new label when we encounter the left endpoint of a new 1-interval in $[n-2]$ or pass the right endpoint of an old 1-interval in $[n-2]$. There are at most $2k(n-2)$ such points, and in fact the first and last cannot generate labels, so we have $r < 2k(n-2)$. Together with (1), this becomes $k^2(k-1) > (n-3)/2$.

For the other case, suppose $P = B_n(k)$, where again $k \le n/2$. The representation defined above uses at most k intervals for each element of P. Again we prove optimality by associating labels with the large sets. Suppose $i(P) < k$ and fix an optimal representation. For each k-set $X \subset [n]$, some interval J assigned to X must contain 1-intervals for distinct singletons. As above, we assign to X a label $l(X) \subset [n]$ such that $2 \le l(X) \le k$, but this time with $l(X) \subset X$. It consists of the singletons whose 1-intervals contain the leftmost duplicated point in J or the two singletons corresponding to leftmost 1-intervals in J. If $|T| = t$, then at most $\binom{n-t}{k-t}$ of the k-sets can have T as label, and this is at most $\binom{n-2}{k-2}$ since $t \ge 2$. If r is the total number of labels, we have $r\binom{n-2}{k-2} \ge \binom{n}{k}$, or $r \ge n(n-1)/[k(k-1)]$. By counting the possible labels arising in a left-to-right scan of the 1-intervals, we obtain $r < 2kn$, or $k^2(k-1) > (n-1)/2$. ■

These bounds are much smaller than the dimension of these posets. Dushnik [1] proved that if $k \ge 2\sqrt{n}$, then $\dim(B_n(k)) = n - r + 1$, where r is the smallest value such that $k \ge \lfloor n/r \rfloor + r - 2$. Spencer [14] proved that $\dim(B_n(2)) \sim \lg\lg n$.

The range of Theorem 1 is somewhat limited, and we soon need to consider other arguments. We showed in [10] that $i(B_5(3)) = i(B_6(4)) = 2$, meaning that the conclusion of Theorem 1 does not extend to all (n, k). For large enough n, the theorem implies that $i(B_n(n-2)) = 3$, but when $n = 7$ the theorem does not yet apply. We consider the example of $B_7(5)$ in detail, because it is the smallest known example (28 elements) of a poset with interval number 3. Our lower bound proof for $i(B_7(5))$ uses a result about interval intersection representations of graphs. For an undirected graph G a *t-representation* is a function f on $V(G)$ such that for every $u \in V$, $f(u)$ is the union of at most t intervals from the real line, and such that $uv \in E$ if and only if $f(u) \cap f(v) \neq \emptyset$ (for distinct u, v). A t-representation has *depth* r if no point on the real line belongs to more than r of the image sets $f(v)$. The *depth-r interval number* of G, denoted $i_r(G)$, is the minimum t such that G has a t-representation of depth r. Scheinerman [11] and Maas [9] proved that $i_r(G) \geq (e + \binom{r}{2})/[n(r-1)]$ for a graph G with n vertices and e edges (see also [13]); this is proved by counting the intersections that can yield edges as the representation is traversed from left to right. The case $r = 2$ was used in the early papers [15] and [4]; $i_2(G) \geq (e+1)/n$. In particular, $i_2(K_5) > 2$. We will use this to prove $i(B_7(5)) > 2$.

THEOREM 2. $i(B_7(5)) = 3$.

Proof. The upper bound follows from the standard construction beginning the proof of Theorem 1. To prove optimality, assume that $i(B_7(5)) \leq 2$ and fix a 2-representation f in which 6 and 7 are the singletons assigned the leftmost and rightmost 1-intervals, in the terminology of the preceding proof. By modifying the 1-intervals for the other elements, we will obtain a depth-2 2-representation for the complete graph K_5, which is impossible. Consider a 5-set X containing $\{6, 7\}$. Since $f(X)$ has only two intervals, X has only one gap, and 1-intervals for the two elements of $[5] - X$ must intersect that gap. Extend these 1-intervals toward each other until they intersect (unless they already intersect). Since no other singleton can have a 1-interval intersecting this gap, this does not cause any point to be contained in more than two 1-intervals. We make this modification for each X containing $\{6, 7\}$. Since every 2-subset of [5] is the complement of such an X, the resulting intervals for [5] form a 2-representation of K_5.

In order to make this into a depth-2 2-representation, we make further modifications to limit the depth at points not belonging to any gap. Because all edges of K_5 have been represented within the gaps, it suffices to reduce the depth outside gaps to 2 by deleting portions of 1-intervals; in doing this, we must avoid increasing the number of intervals used for any singleton. Let J be a maximal interval not intersecting any gap. It suffices to show that at most two 1-intervals can contain J and extend on both sides, because any other portions of 1-intervals in J can be deleted without increasing the number of 1-intervals. In fact, if $f(\{i\})$ and $f(\{j\})$ both contain J (and extend on both sides), let $X = [7] - \{i, j\}$. Then these 1-intervals for i and j both intersect the gaps that occur on each side of J. The only 5-set for which these can be gaps is X, but if X has two gaps then $f(X)$ has at least three intervals. Hence in fact at most one 1-interval properly contains J in this way, and we can reduce the depth outside gaps to 1, forming the impossible representation of K_5. ■

In light of Theorem 2, in which the conclusion of Theorem 1 still holds, it is natural to try to extend the range of applicability of Theorem 1. The next theorem places a limit on how far the formula of Theorem 1 can hold. It would be interesting to know the asymptotics of the maximum value of k such that $i(B_n(n-k)) = k + 1$.

THEOREM 3. If $k > \lceil (\sqrt{2n+1} - 1)/2 \rceil$, then $i(B_n(n-k)) \le k$.

Proof. Suppose $n = 2r(r+1)$ and $k = r+1$; we provide a k-representation for $B_n(n-k)$. Treat the elements of $[n]$ as pairs (i, j) indexing the rows and columns of an r by $2r+2$ array. We create an ordering of 1-intervals so that any pair of singletons in a single row of the array have adjoining 1-intervals. Such adjacencies correspond to a complete graph of order $2r+2$; add one dummy element to obtain a complete graph in which the vertices have even degree. This graph is Eulerian; order the 1-intervals by their appearances on an Eulerian circuit, starting with the dummy element, then delete the dummy intervals. Do this successively for the singletons in each row. Note that we use $k = r+1$ intervals for each singleton.

To complete the containment representation, we must add k intervals for each $(n-k)$-set X that include all 1-intervals for the elements in X but omit at least one 1-interval for each element of the complement. Since $|\bar{X}| = r+1$, the set X omits some pair of elements in the same row. These have a consecutive pair of 1-intervals, so we need only allow r gaps to miss 1-intervals for the $r+1$ elements of $\bar{X}$.

For the general case, let $r = \lceil \frac{1}{2}\sqrt{2n} \rceil$, and fill out the grid of elements to an r by $2r+2$ array by adding dummy elements. With $k > r$, there are still two omitted elements in the same row, and the construction described above works. ■

When k is near $n/2$, or indeed is between αn and βn for $0 < \alpha < \beta < 1$ when n is large, much more efficient representations can be found. For such values of k, it follows from the results of the next section that a factor of $\lg n$ can be saved compared to $\min\{k, n-k+1\}$.

3. BOUNDS IN TERMS OF POSET SIZE

In this section we consider posets with n elements. Hiraguchi [5] proved that $\dim(P) \le n/2$ when $n \ge 4$; hence we know $i(P) \le \lceil n/4 \rceil$. Hiraguchi's inequality is tight for the posets $B_{n/2}(n/2 - 1)$, but these have interval number 2. We prove next that the interval number of a bipartite poset cannot be linear in the size, which suggests that the maximum over all n-element posets will be sublinear.

THEOREM 4. If P is an n-element bipartite poset, then

$$i(P) \le \left\lceil \frac{n}{\lg n - \lg\lg n} \right\rceil + 1.$$

Proof. We establish a universal structure that will accomodate representations for all bipartite posets with elements $[n]$. Let $m = \lg n - \lg\lg n$. Partition $[n]$ into $\lceil n/m \rceil$ disjoint subsets X_i with $|X_i| \le m$. For each i, create $2^{|X_i|}$ disjoint subintervals within the interval $[i-1, i]$. Assign each subset $S \subseteq X_i$ one of these intervals, arbitrarily, and assign each element of S a distinct point within the interval $g(S)$ assigned to S. Note that each element of X_i is assigned to at most 2^{m-1} distinct points in $[i-1, i]$. With $m = \lg n - \lg\lg n$, we have $2^{m-1} = n/(2\lg n) < \lceil n/m \rceil + 1$.

Now consider an arbitrary poset P on the set $[n]$. For each element $x \in P$, the order ideal $D(x) = \{y \in P\colon y \le x\}$ generated by x is a subset of $[n]$. If x is a minimal element in P, we represent x by the points assigned to it within $[i-1, i]$, where i is the unique index such that $x \in X_i$. If x is not a minimal element, then the partition $\{X_i\}$ of $[n]$ induces a partition of $P - D(x)$, with $Y_i = X_i \cap (P - D(x))$. We let $f(x) = [0, \lceil n/m \rceil] - \cup_i g(Y_i)$. Since we have deleted $\lceil n/m \rceil$ intervals, $f(x)$

consists of at most $\lceil n/m \rceil + 1$ intervals. Furthermore, we have provided a gap in $f(x)$ containing an interval for each minimal element outside $D(x)$ and a portion of intervals for each maximal element other than x, so this is a representation. ■

Note that this construction does not work for posets of arbitrary height. It works for bipartite posets because the inclusion relationships between the non-minimal elements are destroyed within the parts X_i of the partition. In fact, the construction applies to any class of bipartite posets having no more than n minimal elements, yielding a representation in which maximal elements are assigned $\lceil \frac{n}{\lg n - \lg\lg n} \rceil + 1$ intervals and minimal elements are assigned at most $\frac{1}{2}n/(2\lg n)$. In particular, we have the bound on $i(B_n(k))$ advertised in the preceding section, which is an improvement over the trivial construction when $\alpha n < k < \beta n$ and n is large.

COROLLARY 5. $i(B_n(k)) \le \lceil \frac{n}{\lg n - \lg\lg n} \rceil + 1$ for all $1 < k < n$. ■

Finally, we show that this is the right order of magnitude for the interval number of almost every poset. This result is based on the description of the random n-element poset due to Kleitman and Rothschild [8]. They obtained an asymptotic formula for the number of n-element posets, proving that its base-2 logarithm is asymptotic to $n^2/4$. Asymptotically, this many posets can be constructed using three levels L_0, L_1, L_2 of sizes $n/4, n/2, n/4$, respectively, and adding some subset of the cover relations $x < y$ such that $x \in L_0$, $y \in L_1$, or $x \in L_1$, $y \in L_2$. Hence almost every n-element poset has this structure, in that the fraction not having this structure goes to 0. Furthermore, if we view these cover relations as introduced independently, with probability 1/2, we see that in almost all posets we have $x < z$ for all $x \in L_0$, $z \in L_2$. (The expected number of unrelated pairs of this form is $(n^2/16)(3/4)^{n/2}$, which approaches 0, so the probability of having no bad pair approaches 1).

THEOREM 6. Suppose P is a random n-element poset, with each such poset being equally likely. For any $\varepsilon > 0$, the probability approaches 1 that

$$(1-\varepsilon)\frac{n}{8\lg n} \le i(P) \le \frac{3}{2}\lceil \frac{n}{\lg n - \lg\lg n} \rceil + 1.$$

Proof. For the upper bound, the discussion above allows us to assume that P has three levels and that every element of the bottom level is less than every element of the top level. We use the construction in Theorem 4 to represent the poset P_0 induced by the bottom two levels and the poset P_1 induced by the top two levels. Translate these representations to disjoint portions of the real line, and give each element of the top level an interval containing the full representation for P_0. The result is a representation of P. Letting $k = \lceil \frac{n}{\lg n - \lg\lg n} \rceil + 1$, it uses at most $k/2$ intervals for each element of the bottom level, $k+1$ for each element of the top level, and $k + k/2$ for each element of the middle level.

The lower bound follows from a counting argument that is standard for lower bounds on representation parameters for combinatorial structures where the representations use intervals (see [2], [12], for example). A representation is determined by the ordering of the endpoints of its intervals. The number of distinct orderings of $2k$ letters of each of n types is $(2kn)!/((2k)!)^n$. The base-2 logarithm of this is asymptotic to $2kn\lg n$. Since the base-2 logarithm of the number of n-element posets is asymptotic to $n^2/4$ [8], almost all n-element posets fail to be representable if we allow at most $k = (1 - \varepsilon)n/(8\lg n)$ intervals per vertex. ■

4. OPEN PROBLEMS

1. Determine the asymptotic behavior of the function $\Phi(n) = \max_{|P| = n} i(P)$; in particular, is it linear or sublinear? Presently we know $n/(8\lg n) \le \Phi(n) \le \lceil n/4 \rceil$. How tight is the bound of Theorem 4 for bipartite posets?

2. Determine the maximum value of k such that $i(B_n(n-k)) = k+1$. We know this value is at least $(n/2)^{1/3}$ and at most $\lceil (\sqrt{2n+1} - 1)/2 \rceil$. Determine the maximum over k of $i(B_n(k))$. In particular, is it $\Omega(n/\lg n)$, and for what k as a function of n does it occur?

3. In [10], we studied ‘‘removal theorems’’ analogous to those of dimension theory. We showed that $i(P - x) \ge i(P) - 1$ if x is a maximal or minimal element of P, but for an arbitrary element we could only prove $i(P - x) \ge i(P)/2$. However, we have no example with $i(P - x) < i(P) - 1$. What is the maximum by which deletion of an element can reduce the interval number?

4. Is the interval number bounded for useful special classes of posets? In [10], we proved $i(P) \le 2$ if P is an interval order. Is there a bound for angle orders, n-gon orders, or planar posets? We have no example of a planar poset with $i(P) > 2$. Kelly [7] proved that dimension can be arbitrarily large for planar posets.

5. What are the bounds on $i(P)$ in terms of other poset parameters? The Boolean algebras show that the bound $i(P) \le \lceil (\dim P)/2 \rceil$ is best possible. In [10], we proved that $i(P) \le \lceil (3/2)\mathrm{idim}(P) \rceil$, where idim denotes interval dimension. Is this best possible? We do not yet know of any poset with interval dimension 2 and interval number 3. Also, since $\dim P$ is bounded by the width of P (maximum antichain size) [6], it follows that $i(P) \le \lceil w(P)/2 \rceil$, but it seems likely that this bound could be substantially improved.

6. In [10], we proved for the lexicographic composition of posets that $i(P[Q]) \le \max\{i(P), \lceil (\dim Q)/2 \rceil\}$. Can this inequality be sharpened to $i(P[Q]) \le 1 + \max\{i(P), i(Q)\}$?

7. The proof in [10] of NP-completeness for recognizing posets with fixed interval number uses posets of unbounded height. Is the problem still NP-complete for posets of fixed height or for bipartite posets?

References

[1] B. Dushnik, Concerning a certain set of arrangements, *Proc. Amer. Math. Soc.* 1(1950), 788-796.

[2] P. Erdös and D.B. West, A note on the interval number of a graph, *Discrete Math.* 55(1985), 129-133.

[3] P.C. Fishburn and W.T. Trotter, Dimensions of hypergraphs, *J. Comb. Theory (B)*, to appear.

[4] J.R. Griggs and D.B. West, Extremal values of the interval number of a graph, *SIAM J. Algeb. Disc. Meth.* 1(1980), 1-7.

[5] T. Hiraguchi, On the dimension of partially ordered sets, *Sci. Rep. Kanazawa Univ.* 1(1951), 77-94.

[6] T. Hiraguchi, On the dimension of orders, *Sci. Rep. Kanazawa Univ.* 4(1955), 1-20.

[7] D. Kelly, On the dimension of partially ordered sets, *Discrete Math.* 35(1981), 135-156.

[8] D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.* 205(1975), 205-220.

[9] C. Maas, Some results about the interval number of a graph, *Discrete Appl. Math.* 6(1983), 99-102.

[10] T. Madej and D.B. West, The interval inclusion number of a partially ordered set, *Discrete Math.* 88(1991), 259-277.

[11] E.R. Scheinerman, Intersection graphs and multiple intersection parameters of graphs, Ph.D. Thesis (Princeton Univ., 1984).

[12] E.R. Scheinerman, On the interval number of random graphs, *Discrete Math.* 82(1990), 105-109.

[13] E.R. Scheinerman and D.B. West, Interval representations of cliques and of subset intersection graphs, *Combinatorial Mathematics - Proc. Third Intl. Conf. (1985)*, *Annals N.Y. Acad. Sci* 555(1989), 363-367.

[14] J. Spencer, Minimal scrambling sets of simple orders, *Acta Math. Acad. Sci. Hung.* 22(1971), 349-353.

[15] W.T. Trotter and F. Harary, On double and multiple interval graphs, *J. Graph Theory* 2(1978), 137-142.