Discrete Mathematics 103 (1992) 25–40
North-Holland

Linear extensions of semiorders: A maximization problem

Peter C. Fishburn

AT&T Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974, USA

W.T. Trotter

Bellcore, 445 South St., Morristown, NJ 07962, USA
Arizona State University, Tempe, AZ 85287, USA

Received 21 July 1989
Revised 2 August 1990

Abstract

Fishburn, P.C. and W.T. Trotter, Linear extensions of semiorders: A maximization problem, Discrete Mathematics 103 (1992) 25–40.

We consider the problem of determining which partially ordered sets on n points with k pairs in their ordering relations have the greatest number of linear extensions. The posets that maximize the number of linear extensions for each fixed (n, k), $0 \leqslant k \leqslant \binom{n}{2}$, are semiorders. However, except for special cases, it appears difficult to say precisely which semiorders solve the problem. We give a complete solution for $k \leqslant n$, a nearly complete solution for $k = n + 1$, and comment on a few other cases.

1. Introduction

Let $(\boldsymbol{n}, >_0)$ denote the set $\boldsymbol{n} = \{1, 2, \ldots, n\}$ partially ordered by an irreflexive and transitive relation $>_0 \subseteq \boldsymbol{n}^2$, and let

$$\begin{aligned} e(\boldsymbol{n}, >_0) = |\{(\boldsymbol{n}, >_*)\colon >_* &\text{ is an irreflexive, transitive and complete} \\ &(a \neq b \Rightarrow a >_* b \text{ or } b >_* a) \text{ relation} \\ &\text{in } \boldsymbol{n}^2 \text{ that includes } >_0\}|, \end{aligned}$$

be the number of linear extensions of $(\boldsymbol{n}, >_0)$. We consider the problem of determining the posets that maximize $e(\boldsymbol{n}, >_0)$ when $>_0$ has exactly k ordered pairs in $\boldsymbol{n}^2$. That is, given $0 \leqslant k \leqslant \binom{n}{2}$ and letting

$$P(n, k) = \{(\boldsymbol{n}, >_0)\colon |>_0| = k\},$$

$$e(n, k) = \max_{P(n,k)} e(\boldsymbol{n}, >_0),$$

Elsevier Science Publishers B.V.

our aim is to characterize the members of $P(n, k)$ for which $e(\boldsymbol{n}, >_0) = e(n, k)$. We are also interested in the values of the $e(n, k)$.

The extreme cases for k are sufficiently restricted to make their answers obvious:

$$e(n, 0) = n!, \qquad e(n, 1) = n!/2,$$

$$e(n, \tbinom{n}{2} - 1) = 2, \qquad e(n, \tbinom{n}{2}) = 1,$$

with poset diagrams in Fig. 1(a)–(d) respectively. A little more effort shows that $e(n, 2) = n!/3$ and $e(n, \binom{n}{2} - 2) = 4$, see Fig. 1(e)–(f). Other cases tend to be far from obvious, and we resolve only a small number of them. They are summarized at the end of this introduction.

We refer to a poset $(\boldsymbol{n}, >_0)$ in $P(n, k)$ as a *realizer* of $e(n, k)$ if $e(\boldsymbol{n}, >_0) = e(n, k)$. Our search for realizers is greatly aided by a theorem of Trotter [5] which says that *every realizer is a semiorder*. Recall that $(\boldsymbol{n}, >_0)$ is a *semiorder* (Luce [2]) if, for all $a, b, x, y \in \boldsymbol{n}$,

$$a >_0 x \text{ and } b >_0 y \Rightarrow a >_0 y \text{ or } b >_0 x,$$

$$a >_0 x >_0 b \Rightarrow a >_0 y \text{ or } y >_0 b.$$

The only poset of Fig. 1 that is not a semiorder is the suboptimal poset in (e).

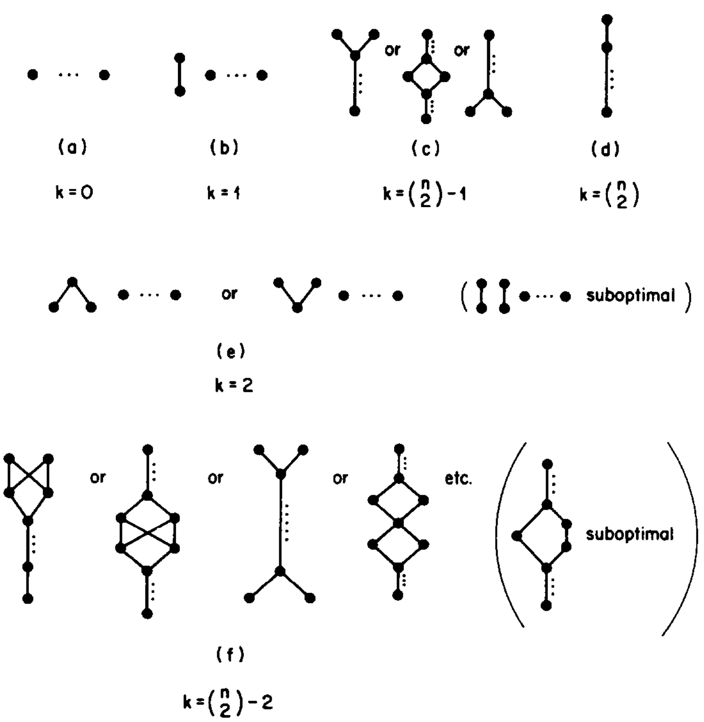

Fig. 1. Realizers of $e(n, k)$ for extreme k.

Semiorders have played a key role in the theory of ordered sets. We know, for example, that if $(\boldsymbol{n}, >_0)$ is a semiorder that is not a linear order or chain, then:

(1) it equals the intersection of either two or three of its linear extensions (Rabinovitch [3]);

(2) there exist $x, y \in \boldsymbol{n}$ such that the proportion of its linear extensions in which $x >_* y$ is between $\frac{1}{3}$ and $\frac{2}{3}$ (Brightwell [1]);

(3) there exists $f: \boldsymbol{n} \to \mathbb{R}$ such that, for all $x, y \in \boldsymbol{n}$, $x >_0 y \Leftrightarrow f(x) > f(y) + 1$ (Scott and Suppes [4]).

Trotter's addition to these seminal results is proved in the next section. We then describe a standard format for semiorders that is used thereafter. This format represents a semiorder for which $|>_0| = k$ by an integer vector $r = (r_1, \ldots, r_{n-1})$ in which $r_1 \geq \cdots \geq r_{n-1} \geq 0$, $r_i \leq n - i$ for each i, and $\sum r_i = k$. Our interpretation is that, for all $1 \leq i < j \leq n$, $i >_0 j \Leftrightarrow j \geq n - r_i + 1$. Up to relabeling, all semiorders on $\boldsymbol{n}$ can be thus represented.

Section 3 is devoted to the special but important case of $1 \leq k < n$. We prove first that, up to relabeling and inversion, $e(n, k)$ is uniquely realized by the height-1 semiorder $r = (k, 0, \ldots, 0)$ that has $e(n, k) = n!/(k + 1)$. The rest of the section considers the runner-up to $r = (k)$, where we omit the 0's in r for convenience. When (k) and its inverse are excluded, the number of linear extensions is uniquely maximized (up to $\cdots$) by $r = (k - 1, 1)$, with $n!(k + 2)/[2k(k + 1)]$ linear extensions, provided that $k \geq 7$. It follows that, when $k < n$ and k is large, the realizer of $e(n, k)$ has nearly twice as many linear extensions as the runner-up.

The restriction of $k \geq 7$ for the preceding runner-up result illustrates a problem endemic to our study in that certain uniform results apply only when we get beyond the first several integers. Consider, for example, the semiorders in Fig. 2 that have $k = 3n - 6$. Without claiming optimality in either case, we note that

$$r = (n - 1, n - 2, n - 3) \text{ has } (n - 3)! \text{ linear extensions,}$$

$$r = (n - 4, n - 4, n - 4, 6) \text{ has } 18(n - 4)! + 6(n - 3)!/7 \text{ linear extensions,}$$

so the second has more linear extensions when $n < 129$ while the first has more for all $n > 129$.

Section 4 concludes our present analysis for realizers with a few cases of $k \geq n$.

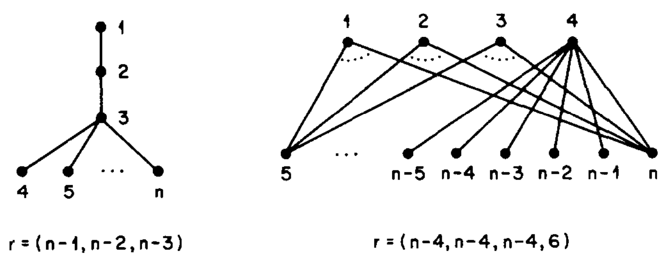

Fig. 2. $k = 3n - 6$.

We let H denote the *height* of a poset, so $H = c$ when the longest chain (linearly ordered subset) has $c + 1$ points. The first two of the three main results in the section further illustrate the theme of the preceding paragraph; the third comments on the smallest k that admits a height-3 semiorder. Uniqueness applies up to relabeling and inversion.

(1) Suppose $k = n \geqslant 4$. Then $r = (\lceil n/2 \rceil, \lfloor n/2 \rfloor)$ uniquely realizes $e(n, n)$ when $n \leqslant 8$, and the $H = 2$ vector $r = (n - 1, 1)$ uniquely realizes $e(n, n)$ when $n \geqslant 9$.

(2) Suppose $k = n + 1$, $n \geqslant 5$. For $n = 14$ and $n \geqslant 16$, $e(n, n + 1)$ is realized by the $H = 2$ semiorders $(n - 2, 2)$ and $(n - 1, 1, 1)$, and by no others. For $n \leqslant 13$, $e(n, n + 1)$ is realized by an $H = 1$ semiorder and never by an $H = 2$ semiorder. For $n = 15$, $e(n, n + 1)$ is realized by the vectors $r = (8, 8)$, $r = (14, 2)$, $r = (14, 1, 1)$ and by no others.

(3) For $n \geqslant 5$, no $H = 3$ semiorder realizes $e(n, 2n - 2)$.

While fragmentary, these results indicate the challenge posed by the realizer problem, and we hope they will provoke further research. Additional comments on open problems conclude the paper.

2. Semiorders

For a poset $(\boldsymbol{n}, >_0)$ let $\sim$, $U(x)$, $D(x)$, and $E(\boldsymbol{n}, >_0)$ denote, respectively, the symmetric complement of $>_0$, x's up set, x's down set, and the family of linear extensions of $(\boldsymbol{n}, >_0)$. Hence $x \sim y$ if neither $x >_0 y$ nor $y >_0 x$, $U(x) = \{y : y >_0 x\}$, $D(x) = \{y : x >_0 y\}$, and $e(\boldsymbol{n}, >_0) = |E(\boldsymbol{n}, >_0)|$. $A \setminus B$ denotes set subtraction.

Theorem 1. *For all $n \geqslant 1$ and all $0 \leqslant k \leqslant \binom{n}{2}$, every realizer of $e(n, k)$ is a semiorder.*

Proof. Assume that $(\boldsymbol{n}, >_0) \in P(n, k)$ realizes $e(n, k)$. Suppose the first semiorder condition is violated. Among all 4-sets in $\boldsymbol{n}$ that have $\{a >_0 x, b >_0 y, a \sim y, b \sim x\}$, choose one that minimizes $|U(a)| + |U(b)|$. It follows that $U(a) \subseteq U(b)$ or $U(b) \subseteq U(a)$. Assume $U(a) \subseteq U(b)$ for definiteness. Let $(\boldsymbol{n}, >')$ be the poset obtained from $(\boldsymbol{n}, >_0)$ by replacing $b >_0 z$ by $a >' z$ for all $z \in D(b) \setminus D(a)$. It is easily seen that $(\boldsymbol{n}, >') \in P(n, k)$.

We claim that $e(\boldsymbol{n}, >') > e(\boldsymbol{n}, >_0)$, thus contradicting our supposition. Let $(\boldsymbol{n}, >_*^{ab})$ denote the linear order obtained by interchanging a and b in $(\boldsymbol{n}, >_*)$. Suppose

$$(\boldsymbol{n}, >_*) \in E(\boldsymbol{n}, >_0) \setminus E(\boldsymbol{n}, >')$$

so that $b >_* z >_* a$ for some $z \in D(b) \setminus D(a)$. Then

$$(\boldsymbol{n}, >_*^{ab}) \in E(\boldsymbol{n}, >') \setminus E(\boldsymbol{n}, >_0).$$

Moreover, there are linear extensions $(\boldsymbol{n}, >_*)$ in $E(\boldsymbol{n}, >')\backslash E(\boldsymbol{n}, >_0)$, including those with $a >_* y >_* x >_* b$, for which $(\boldsymbol{n}, >_*^{ab}) \notin E(\boldsymbol{n}, >_0)\backslash E(\boldsymbol{n}, >')$. Therefore $e(\boldsymbol{n}, >') > e(\boldsymbol{n}, >_0)$.

It follows that the realizer $(\boldsymbol{n}, >_0)$ satisfies the first semiorder condition. Suppose it violates the second semiorder condition, say with $a >_0 x >_0 b$, $y \sim a$ and $y \sim b$. Form $(\boldsymbol{n}, >')$ from $(\boldsymbol{n}, >_0)$ by replacing $u >_0 x$ by $u >' y$ for all $u \in U(x)\backslash U(y)$. Then, with the first semiorder condition holding for $(\boldsymbol{n}, >_0)$, it follows that $(\boldsymbol{n}, >') \in P(n, k)$. And, with interchanges of x and y in this case, we get $e(\boldsymbol{n}, >') > e(\boldsymbol{n}, >_0)$, for a contradiction.

Hence $(\boldsymbol{n}, >_0)$ satisfies both semiorder conditions. □

We assume henceforth that $k \geqslant 1$.

To develop our standard format for semiorders, let $(\boldsymbol{n}, >_0)$ be a semiorder with unit interval representation $x >_0 y \Leftrightarrow f(x) > f(y) + 1$, assume with no loss of generality that the left ends of the unit intervals $[f(x), f(x) + 1]$ are distinct, and relabel the points so that

$$f(1) > f(2) > \cdots > f(n).$$

The $n \times n$ $>_0$-matrix for $(\boldsymbol{n}, >_0)$ thus labeled is the 0–1 matrix with 1 in cell (i, j) if $i >_0 j$, and 0 otherwise. Its 1's in row i (if any) are contiguous, are to the right of the main diagonal, and end in cell (i, n). Moreover, with r_i the number of 1's in row i, $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_n = 0$, $r_i \leqslant n - i$, and $\sum r_i = k$ when $|>_0| = k$. Conversely, if $(\boldsymbol{n}, >_0)$ is defined in the natural way from a 0–1 matrix with these properties, it is a semiorder.

Up to relabeling, each semiorder on n points is uniquely representable as a vector $r = (r_1, \ldots, r_{n-1})$ that has the preceding properties. We refer to r itself as a semiorder. Its *inverse* is the semiorder $r' = (r_1', \ldots, r_{n-1}')$ in which r_i' is the number of 1's in column $n + 1 - i$ of r's $>_0$-matrix. The diagram of r' is obtained by inverting r's diagram and relabeling point i by $n + 1 - i$ for $i = 1, \ldots, n$. The inverse of $r = (n - 1, n - 2, n - 3, 0, \ldots, 0)$ in Fig. 2 is $r' = (3, \ldots, 3, 2, 1)$ with points 1 through $n - 3$ above $n - 2$, $n - 1$ and n in order.

We say that r is in *standard format* if $r_i > r_i'$ for the smallest i at which $r_i \neq r_i'$. When $r = r'$, the two semiorders are identical, but not otherwise. Since r and r' always have the same number of linear extensions, we generally work with the one in standard format since its expression tends to be simpler. Whenever a claim of uniqueness is made for r, or a set of r's, it denotes uniqueness up to relabeling and inversion.

As a further convenience, we usually omit the 0's from r in standard format, i.e., as in $r = (r_1, \ldots, r_m)$ with $r_1 \geqslant \cdots \geqslant r_m \geqslant 1$. Since this can disguise the value of n for height-1 semiorders if there are isolated points, we denote by $e_n(r)$ the number of linear extensions of a semiorder $(\boldsymbol{n}, >_0)$ with vector r.

It is usually quite difficult to compute $e(\boldsymbol{n}, >_0)$ for a poset $(\boldsymbol{n}, >_0)$, and this is often true as well for the computation of $e_n(r)$ for a semiorder r. Our subsequent counts use two simple rules: an a-point antichain has $a!$ linear extensions; the union of disjoint a-point and b-point chains has $\binom{a+b}{a}$ linear extensions. We use these repeatedly and in various sequences. For example, if a poset has t points in a connected component along with $n-t$ isolated points, and if the t-point component has c linear extensions by itself, then $e = c(n!/t!)$ for the whole.

We conclude our preliminaries with two basic lemmas on heights of semiorders. Throughout, $k = |>_0|$ with $k \geq 1$. As before, H denotes height.

Lemma 1. *Every n-point semiorder with $k<n$ has $H=1$. If a semiorder has $H \geq 2$, its diagram is connected, i.e., it has no isolated points.*

We omit the simple proof. The complete bipartite diagram with $\lfloor n/2 \rfloor$ top points and $\lceil n/2 \rceil$ bottom points shows that k for $H=1$ can be as large as about $n^2/4$. The other lemma goes the opposite way.

Lemma 2. *Given $2 \leq h \leq n-1$, the smallest k for an n-point semiorder with height h is $k = n(h-1) - (h-2)(h+1)/2$.*

Proof. Let n and $2 \leq h \leq n-1$ be given. We are to minimize the number of 1's in an $n \times n$ $>_0$-matrix associated with a standard format r so that there are integers

$$1 < a_1 < a_2 < \cdots < a_{h-1} < n \tag{$*$}$$

for which the matrix has 1's in cells $(1, a_1)$, (a_1, a_2), $\ldots$, (a_{h-1}, n) and in all cells northeast of these. This produces the height-h chain $1 >_0 a_1 >_0 \cdots >_0 a_{h-1} >_0 n$. The number of 1's needed when $h=2$ is $(n-a_1) + a_1 = n$, with $a_1 \leq (n+1)/2$ for the standard format.

When $h \geq 3$, our matrix has

$$K = (a_2 - a_1) + (a_3 - a_2)a_1 + (a_4 - a_3)a_2 + \cdots + (n - a_{h-1})a_{h-2} + a_{h-1}$$

1's. To minimize K subject to $(*)$, observe that its contribution involving a_1 is $a_1(-1 + a_3 - a_2)$ with $-1 + a_3 - a_2 \geq 0$, so regardless of a_2 through a_{h-1} we can do no better than to minimize a_1 at $a_1 = 2$. Then

$$K = -2 + a_2(-1 + a_4 - a_3) + a_3(2 + a_5 - a_4) + a_4(a_6 - a_5) + \cdots,$$

so we minimize a_2 at $a_2 = 3$ and continue as indicated to conclude that K is minimized when $(a_1, a_2, \ldots, a_{h-2}) = (2, 3, \ldots, h-1)$. The value of a_{h-1} is immaterial, subject to $h - 1 < a_{h-1} \leq n-1$, since we already have

$$\begin{aligned} K &= 1 + 2 + 3 + \cdots + (h-3) + [a_{h-1} - (h-1)](h-2) \\ &\quad + (n - a_{h-1})(h-1) + a_{h-1} \\ &= (h-1)n - (h-2)(h+1)/2. \end{aligned}$$

This minimum K is the value of k given in the lemma. $\square$

3. Analysis for $k < n$

We begin with the realizer of $e(n, k)$, then consider the second-best linear extension maximizer, which also turns out to be a semiorder.

Theorem 2. *If $k < n$ then $r = (k)$ with $e_n(r) = n!/(k+1)$ is the unique realizer of $e(n, k)$.*

Proof. The number of linear extensions of $r = (k)$ is

$$k!(n-k-1)!\binom{n}{k+1} = n!/(k+1).$$

To show that this uniquely maximizes linear extensions for semiorders r in standard format, consider such an $r = (r_1, \ldots, r_m)$ with $m \geq 2$ nonzero components. We have $r_1 \geq m$, $r_1 \geq \cdots \geq r_m \geq 1$ and $\sum r_i = k < n$. Fig. 3(a) shows its diagram. As in Lemma 2, it has $H = 1$.

Suppose $r_2 = \cdots = r_m = 1$. Then $r_1 = k - m + 1$. By considering the possible positions of point n in the chain down from point 1 after we linearize the points under 1 except for n, and linearize points 2 through m above point n, as shown in Fig. 3(b), we get

$$\begin{aligned} e_n(r) &= n!\left[\sum_{j=0}^{k-m}\binom{m+j}{m-1}\right](k-m)!\,(m-1)!/(k+1)! \\ &= n!\left[\binom{k+1}{m} - 1\right](k-m)!\,(m-1)!/(k+1)! \\ &< n!/[m(k+1-m)] \leq n!/(k+1), \end{aligned}$$

where $k+1 \leq m(k+1-m)$ follows from $r_1 = k-m+1 \geq m$. Hence

$$e_n(k+1-m, 1, \ldots, 1) < e_n(k).$$

For general $r = (r_1, \ldots, r_m)$, $m \geq 2$, let

$$Z = \frac{(r_1 - r_2 + 2)(r_1 - r_3 + 3)\cdots(r_1 - r_m + m)}{(r_1+1)(r_1+2)(r_1+3)\cdots(r_1+m)},$$

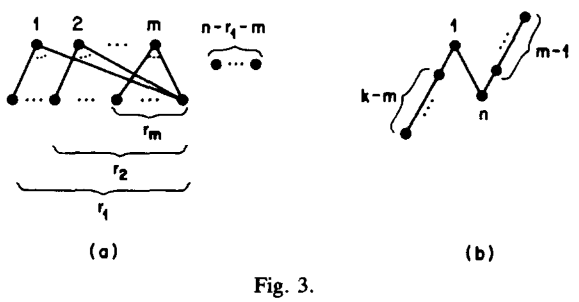

Fig. 3.

and note that $e_n(r) \leqslant n!\,Z$. For if the connected (r_1+m)-point component has C linear extensions, then $e_n(r) = n!C/(r_1+m)!$, and

$$C \leqslant r_1!(r_1-r_2+2)(r_1-r_3+3)\cdots(r_1-r_m+m)$$

as seen by first linearizing the r_1 points under 1 and then adding in points $2, 3, \ldots, m$ in sequence in the linear order.

Suppose $r_i \geqslant 2$ for some $i \geqslant 2$. Let j denote the largest i for which $r_i \geqslant 2$, and modify r by decreasing r_j by 1 and increasing r_1 by 1. A little algebra shows that this change increases Z. It follows that Z is maximized when $r_2 = r_3 = \cdots = r_m = 1$. Since $Z = 1/(k+1)$ in this case, we conclude that if $r_i \geqslant 2$ for some $i \geqslant 2$, then $e_n(r) \leqslant n!\,Z(r) < n!\,Z_{\max} = e_n(k)$. □

We now consider the second-best posets for $k<n$. Exhaustive computations for small k show that the semiorders of Fig. 4 maximize the number of linear extensions when $r=(k)$ is excluded. The pattern for $k \in \{7, 8\}$ persists for all larger k. Its r is $r=(k-1, 1)$ with

$$e_n(k-1, 1) = n!(k+2)/[2k(k+1)].$$

Since $e_n(k)/e_n(k-1, 1) = 2[k/(k+2)]$, the realizer has nearly twice as many linear extensions as the runner-up when k is large.

To prove the second-best optimality of $(k-1, 1)$, we begin with a lemma which says that the second-best poset, after (k) and its inverse, must be a semiorder.

Lemma 3. *Suppose $k<n$, $X \in P(n, k)$, and $e(X) = \max\{e(\boldsymbol{n}, >_0)$: $(\boldsymbol{n}, >_0) \in P(n, k)$ and $e(\boldsymbol{n}, >_0) < e_n(k) = n!/(k+1)\}$. Then X is a semiorder.*

Proof. Suppose $X \in P(n, k)$ is second-best as asserted in the hypotheses but is not a semiorder. By the proof of Theorem 1, a simple change in X (from $b >_0 z$ to $a >' z$ for $z \in D(b)\backslash D(a)$ in the first case; from $u >_0 x$ to $u >' y$ for $u \in U(x)\backslash U(y)$ in the second case) yields another poset $X' \in P(n, k)$ that has more linear extensions than X. Our supposition therefore implies that X' is the realizer (k), hence that X is such that the change yields (k) from X. Up to inversion, the only way that this can happen is for X to have height 1 with p points under point 1,

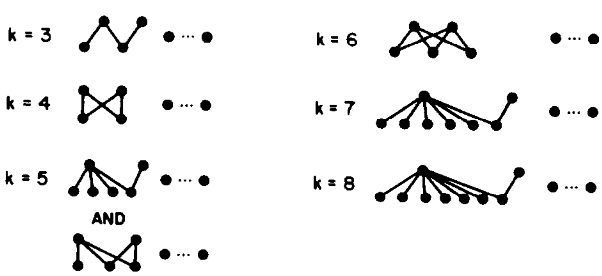

Fig. 4. Second-best posets, $k<n$.

another q points under point 2, $p+q=k$, and $n-(k+2)$ isolated points. This requires $k \leqslant n-2$.

Suppose X is as described. Then

$$e(X)=p!\,q!\,(n-k-2)!\binom{p+q+2}{p+1}\binom{n}{k+2}=n!/[(p+1)(q+1)].$$

With $p \geqslant 1$, $q \geqslant 1$ and $p+q=k$, $e(X)$ is maximized with $p=1$:

$$\max e(X)=n!/(2k).$$

However, we have a semiorder $(k-1, 1)$ that is not a realizer and has

$$e_n(k-1, 1)=n!\,(k+2)/[2k(k+1)].$$

Since $e_n(k-1, 1)>\max e(X)$, we contradict the supposition that the second-best poset is not a semiorder. □

Theorem 3. *If $7 \leqslant k<n$, then semiorder $(k-1, 1)$ uniquely maximizes the number of linear extensions over $P(n, k)$ when (k) and its inverse are excluded.*

Proof. Since our complete proof of the theorem is rather long, we omit some routine details. By Lemma 3, we need only consider semiorders. The proof is organized around the number $m \geqslant 2$ of positive components of r in standard format. For convenience let

$$e^*=e_n(k-1, 1)=n!\,(k+2)/[2k(k+1)].$$

$m=2$: For $1 \leqslant t \leqslant k/2$,

$$e_n(k-t, t)=n!\,(k+2)/[(t+1)(k-t+2)(k-t+1)].$$

The second derivative of $e_n(k-t, t)$ with respect to t is positive, so $e_n(k-t, t)$ is maximum at either $t=1(e^*)$ or $t=\lfloor k/2 \rfloor$. When k is even, $e^*>e_n(k/2, k/2)$ for $k \geqslant 8$; when k is odd, $e^*>e_n((k+1)/2, (k-1)/2)$ for $k \geqslant 7$. Hence the uniquely best semiorder for $m=2$ is $(k-1, 1)$ when $k \geqslant 7$.

$m=3$: Assume $k \geqslant 9$ henceforth. For $r=(r_1, r_2, r_3)$ with $r_1 \geqslant r_2 \geqslant r_3 \geqslant 1$ and $\sum r_i=k$, we have

$$e_n(r_1, r_2, r_3)=n!\,g(r_1, r_2, r_3)/[(r_1+1)(r_1+2)(r_1+3)],$$

$$g(r_1, r_2, r_3)=2+\frac{(r_1+1)(r_1+r_2+4)(r_2+r_3+2)}{(r_2+1)(r_2+2)(r_3+1)}.$$

To prove that $e^*>e_n(r_1, r_2, r_3)$, it is enough to show that

$$g(r_1, r_2, r_3)/[(r_1+1)(r_1+2)(r_1+3)]<1/(2k),$$

i.e.,

$$2<(r_1+1)\left[\frac{(r_1+2)(r_1+3)}{2(r_1+r_2+r_3)}-\frac{(r_1+r_2+4)(r_2+r_3+2)}{(r_2+1)(r_2+2)(r_3+1)}\right]. \qquad (*)$$

Differentiation shows that the right side of this inequality is minimized at either $r_3 = 1$ or $r_3 = r_2$.

Suppose $r_3 = r_2$. Substitution in (*) and differentiation shows that its right side is minimized when $r_2 = 1$ or $r_2 = r_1$. If $r_2 = r_1$ then (*) is

$$2 < (r_1 + 1)(r_1 + 2)(r_1 + 3)/(6r_1) - 4,$$

which is true when $r_1 \geqslant 3$. When $r_2 = 1$, the desired result follows directly from the following lemma.

Lemma 4. *$e_n(k + 1 - m, 1, \ldots, 1) < e^*$ if $k \geqslant 7$ and $3 \leqslant m \leqslant (k + 1)/2$.*

Lemma 4 is an easy consequence of the expression for $e_n(r)$ in the second paragraph of the proof of Theorem 2. We also consider $r_2 = 2$ because $k \geqslant 9$ allows smaller r_1 values when $r_2 + r_3$ is large. When $r_2 = r_3 = 2$, (*) is easily verified for $r_1 \geqslant 3$.

We complete verification of (*) by considering small values of r_3. When $r_3 = 1$, (*) is

$$4 < (r_1 + 1)\left[\frac{(r_1 + 2)(r_1 + 3)}{r_1 + r_2 + 1} - \frac{(r_1 + r_2 + 4)(r_2 + 3)}{(r_2 + 1)(r_2 + 2)}\right].$$

Differentiation again shows that the right side is minimized at either $r_2 = r_1$ or else when r_2 is small, and Lemma 4 and substitution give the desired result when $r_1 \geqslant 3$. When $r_3 = 2$, or $r_3 = 3$, we arrive at the same conclusion by similar means.

$m \geqslant 4$: For $m \geqslant 4$ we begin with the precise number of linear extensions computed with $m = 3$ for the first three top points in the connected component and then merge top points 4 through m in a greedy manner to get

$$\begin{aligned} e_n(r_1, \ldots, r_m) &\leqslant r_1!\, g(r_1, r_2, r_3)(r_1 - r_4 + 4) \cdots (r_1 - r_m + m)n!/(r_1 + m)! \\ &= n!\, \frac{g(r_1, r_2, r_3)R}{(r_1 + 1)(r_1 + 2)(r_1 + 3)}, \end{aligned}$$

$$R = \left(\frac{r_1 - r_4 + 4}{r_1 + 4}\right)\left(\frac{r_1 - r_5 + 5}{r_1 + 5}\right) \cdots \left(\frac{r_1 - r_m + m}{r_1 + m}\right).$$

To verify $e_n(r) < e^*$ here, it suffices to show that

$$\frac{g(r_1, r_2, r_3)R}{(r_1 + 1)(r_1 + 2)(r_1 + 3)} < \frac{1}{2k}, \tag{**}$$

given $4 \leqslant m \leqslant r_1$ and $9 \leqslant k = \sum r_i < n$.

Let $s = r_1 + r_2 + r_3$, so $r_4 + \cdots + r_m = k - s$. Consider maximization of R first since it is the only part of (**) that involves r_4 through r_m. It is easily seen that R is maximized by taking $r_4 = \cdots = r_m = 1$ so long as we arrive at $k - s$ for their total and maintain $m \leqslant r_1$, i.e., so long as $k - s + 3 \leqslant r_1$. Since this might be violated by the given values for r, k and m, we consider two cases.

Case 1 ($m \geqslant 4$): $k - s + 3 \leqslant r_1$.

Then R is maximized by taking $r_4 = \cdots = r_m = 1$ and $m = k - s + 3$. This gives $R = (r_1 + 3)/(r_1 + m)$. Substitution in ($**$) reduces it to

$$\frac{g(r_1, r_2, r_3)}{(r_1+1)(r_1+2)} < \frac{r_1 + m}{2(m+s-3)}.$$

Since the derivative of the right side with respect to m is negative if and only if $r_2 = r_3 = 1$, a case already covered by Lemma 4, we need only consider our present inequality at $m = 4$. When $m = 4$ is used therein, comparison with ($*$) and the present restriction of $r_1 \geqslant 4$ shows that the analysis for ($*$) covers the present situation, i.e., the preceding inequality holds for all possible cases when $m = 4$.

Case 2 ($m \geqslant 4$): $r_1 < k - s + 3$.

In this case it is easily seen that R is maximized by taking m equal to r_1 and making the r_i for $i \geqslant 4$ as equal as possible. Ignoring the restriction to integer values, we show that ($**$) holds when $r_4 = \cdots = r_m = (k-s)/(m-3)$. Then it must also hold when the r_i are integers. With $m = r_1$ and the equal values of r_4 through r_m, ($**$) is

$$2g(r_1, r_2, r_3)k \prod_{j=4}^{r_1} [(r_1 + j)(r_1 - 3) + s - k] < (r_1 - 3)^{r_1 - 3}(2r_1)!/r_1!,$$

subject to $s + (r_1 - 3) < k \leqslant s + r_3(r_1 - 3)$, which requires $r_3 \geqslant 2$. With r_1, r_2 and r_3 fixed, the left side of this inequality is maximized when k is as small as possible. Hence it suffices to show that the inequality holds when $k = s + r_1 - 3$. Substitution for this value of k and comparison with ($*$) as in Case 1 shows that our present inequality holds when $r_1 \geqslant 4$ and $r_2 \geqslant r_3 \geqslant 2$. □

4. Special cases for $k \geqslant n$

We consider the three special cases for $k \geqslant n$ outlined at the end of the introduction.

Theorem 4. *Suppose $k = n \geqslant 4$. If $n \leqslant 8$, $r = (\lceil n/2 \rceil, \lfloor n/2 \rfloor)$ is the unique realizer of $e(n, n)$; if $n \geqslant 9$, the $H = 2$ semiorder $r = (n-1, 1)$ is the unique realizer of $e(n, n)$.*

Proof. The possible $H = 2$ semiorders for $k = n \geqslant 4$ are shown in Fig. 5. We have $e_n(n-1, 1) = (n-1)!/2$ and, for $2 \leqslant t \leqslant (n-1)/2$,

$$e_n(n-t, 1, 1, \ldots, 1) < [(t+1)!/2](n-t-2)!\binom{n-1}{n-t-2} = (n-1)!/2,$$

where $<$ follows from merging the $t-1$ with the 3-point chain and then overcounting the merger of the $n-t-2$ with the others. Hence the unique $H = 2$ maximizer is $(n-1, 1)$.

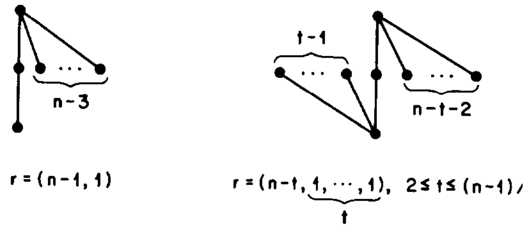

Fig. 5. Height-2 semiorders for $k = n$.

Consider next the $H = 1$ semiorders with $m = 2$ positive components in r. They are $(n-t, t)$, $2 \leqslant t \leqslant n/2$, with

$$e_n(n-t, t) = \frac{n!\,(n+2)}{(t+1)(n+1-t)(n+2-t)}.$$

Since the denominator of the ratio is concave over the range of t, $e_n(n-t, t)$ is maximized at either $t = 2$ or $t = \lfloor n/2 \rfloor$. For even n we get

$$e_n(n/2, n/2) > e_n(n-2, 2) \quad \text{for } n \leqslant 12,$$

$$e_n(n-2, 2) > e_n(n/2, n/2) \quad \text{for } n \geqslant 14;$$

for odd n,

$$e_n((n+1)/2, (n-1)/2) > e_n(n-2, 2) \quad \text{for } n \leqslant 9,$$

$$e_n(n-2, 2) > e_n((n+1)/2, (n-1)/2) \quad \text{for } n \geqslant 11.$$

Thus, given $m = 2$ and $H = 1$, the unique maximizer is

$$(\lceil n/2 \rceil, \lfloor n/2 \rfloor) \quad \text{for } n \leqslant 10 \text{ and } n = 12,$$

$$(n-2, 2) \qquad \text{for } n = 11 \text{ and } n \geqslant 13.$$

Since it is easily checked that $e_n(n-1, 1) > e_n(n-2, 2)$ for all $n \geqslant 9$, we ignore $(n-2, 2)$ henceforth. In addition,

$$e_n(\lceil n/2 \rceil, \lfloor n/2 \rfloor) > e_n(n-1, 1) \quad \text{for } n \leqslant 8,$$

$$e_n(n-1, 1) > e_n(\lceil n/2 \rceil, \lfloor n/2 \rfloor) \quad \text{for } n \geqslant 9.$$

Complete enumeration for $n \leqslant 8$ shows that $(\lceil n/2 \rceil, \lfloor n/2 \rfloor)$ is the unique realizer of $e(n, n)$ when $4 \leqslant n \leqslant 8$. To conclude the proof of the theorem, we show that $e_n(n-1, 1) > e_n(r)$ for all $H = 1$ semiorders r with $m \geqslant 3$ positive components when $n \geqslant 9$. To ensure $H = 1$ we require $r_2 \geqslant 2$.

Suppose $m = 3$. The $m = 3$ proof for Theorem 3 gives

$$e_n(r) = n!\,g(r_1, r_2, r_3)/[(r_1+1)(r_1+2)(r_1+3)].$$

Therefore $e_n(n-1, 1) > e_n(r)$ if and only if $g < (r_1+1)(r_1+2)(r_1+3)/(2n)$, or

$$2 < (r_1+1)\left[\frac{(r_1+2)(r_1+3)}{2(r_1+r_2+r_3)} - \frac{(r_1+r_2+4)(r_2+r_3+2)}{(r_2+1)(r_2+2)(r_3+1)}\right].$$

This is (*), and we have already verified it when $r_1+r_2+r_3 \geqslant 9$.

It remains to compare $e_n(n-1, 1)$ to $e_n(r)$ when $m \geqslant 4$ for $n \geqslant 9$. This has already been done in the $m \geqslant 4$ part of the proof of Theorem 3 by our use of $1/(2k)$ instead of $[1/(2k)](k+2)/(k+1)$ on the right side of (**). It follows that the $H=2$ semiorder $(n-1, 1)$ is the unique realizer of $e(n, n)$ for all $n \geqslant 9$. □

Theorem 5. *Suppose $k = n+1$, $n \geqslant 5$. For $n \leqslant 13$, $e(n, n+1)$ is realized by an $H = 1$ semiorder and no $H = 2$ semiorder. For $n = 14$ and $n \geqslant 16$, the $H = 2$ semiorders $(n-1, 2)$ and $(n-1, 1, 1)$ are the only realizers of $e(n, n+1)$. For $n = 15$, $e(15, 16)$ is realized by $(8, 8)$, $(14, 2)$ and $(14, 1, 1)$, and by no other semiorders.*

Proof. With $k = n+1 \geqslant 6$, Lemma 2 tells us that no semiorder has $H = 3$. The possible $H = 2$ semiorders are shown in Fig. 6. For the top two,

$$e_n(n-1, 2) = e_n(n-1, 1, 1) = (n-1)!/3.$$

When $t \geqslant 2$ and $r_2 = 2$ (bottom left), there are eight ways to merge a and b with the 3-point chain. The maximum number of ways to merge the remaining $n-5$ with the 5-point chain occurs when a and b go between 1 and n. Therefore

$$e_n(n-t, 2, 1, \ldots, 1) < 8(t-2)!\binom{t+2}{t-2}(n-t-3)!\binom{n-1}{n-t-3} = \frac{(n-1)!}{3}.$$

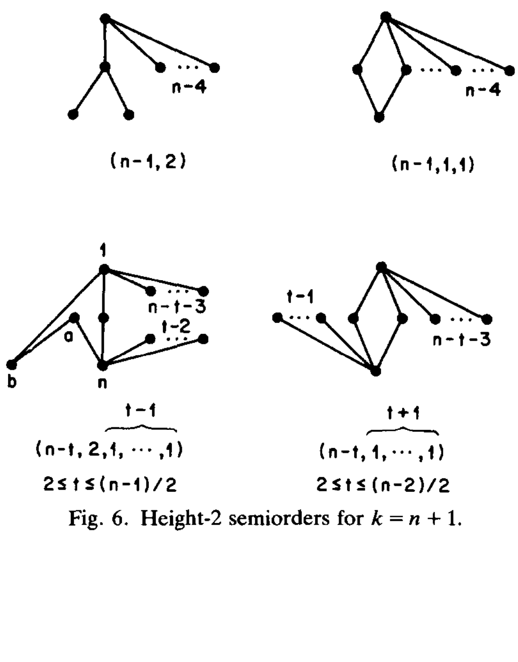

Fig. 6. Height-2 semiorders for $k = n+1$.

When $t \geqslant 2$ and $r_2 = 1$ (bottom right), we get

$$e_n(n-1, 1, \ldots, 1) < 2(t-1)!\binom{t+2}{t-1}(n-t-3)!\binom{n-1}{n-t-3} = \frac{(n-1)!}{3}.$$

Therefore $(n-1, 2)$ and $(n-1, 1, 1)$ are the only linear extension maximizers among the $H = 2$ semiorders.

Each $H = 1$ semiorder with $m = 2$ has $r = (n-t, t+1)$, $2 \leqslant t \leqslant (n-1)/2$, and

$$e_n(n-t, t+1) = \frac{n!\,(n+3)}{(t+2)(n-t+1)(n-t+2)}.$$

This is maximized at $t = 2$ or $t = \lfloor (n-1)/2 \rfloor$. Calculations show that the unique maximizer for $(H, m) = (1, 2)$ is

$$\begin{array}{ll} (\lceil (n+1)/2 \rceil, \lfloor (n+1)/2 \rfloor) & \text{for } n \leqslant 15, \\ (n-2, 3) & \text{for } n \geqslant 16. \end{array}$$

Comparisons to the $H = 2$ maximizers give: $(\lceil (n+1)/2 \rceil, \lfloor (n+1)/2 \rfloor)$ beats $H = 2$ if $n \leqslant 13$; the $H = 2$ best beat the $(H, m) = (1, 2)$ best if $n = 14$ or $n \geqslant 16$; the $H = 2$ best tie $(8, 8)$ at $n = 15$.

To complete the proof it suffices to note that $n \geqslant 14$ implies $(n-1)!/3 > e_n(r)$ whenever $m \geqslant 3$, $r_1 \leqslant n-2$ and $(r_2, r_3) \notin \{(1, 1), (2, 1)\}$. The $H = 1$ semiorders for $m = 3$ have the expression used earlier for this case (with $r_1 + r_2 + r_3 = n + 1$ here), and when this $e_n(r)$ is compared to $(n-1)!/3$ we conclude that the latter is larger if and only if

$$2 < (r_1 + 1)\left[\frac{(r_1+2)(r_1+3)}{3(r_1+r_2+r_3-1)} - \frac{(r_1+r_2+4)(r_2+r_3+2)}{(r_2+1)(r_2+2)(r_3+1)}\right].$$

This is like (*) except that $3(r_1 + r_2 + r_3 - 1)$ replaces $2(r_1 + r_2 + r_3)$ in the first denominator. Analysis similar to that for $m = 3$ in the proof of Theorem 3 shows that it holds for all $n \geqslant 14$. The analysis for $m \geqslant 4$ also mimics that for Theorem 3. In the present context we replace the right side of (**) by $1/(3n)$, or the slightly smaller $1/[3(r_1 + \cdots + r_m)]$, and find that (**) thus modified holds for $n \geqslant 14$. We omit further details. □

By Lemma 2, $k = n$ is the smallest k that admits an $H = 2$ semiorder, and Theorem 4 shows that an $H = 2$ semiorder uniquely realizes $e(n, n)$ when $n \geqslant 9$. Lemma 2 also says that $k = 2n - 2$ is the smallest k that admits an $H = 3$ semiorder. However, no such semiorder realizes $e(n, 2n-2)$, i.e., all realizers of $e(n, 2n-2)$ have height 1 or 2.

Theorem 6. *If $n \geqslant 5$ then no semiorder that is a realizer of $e(n, 2n-2)$ has $H = 3$.*

Proof. Assume that $n \geqslant 5$ and $k = 2n - 2$. By following the logic in the proof of Lemma 2 we find that the standard format semiorders with $H = 3$ are those shown

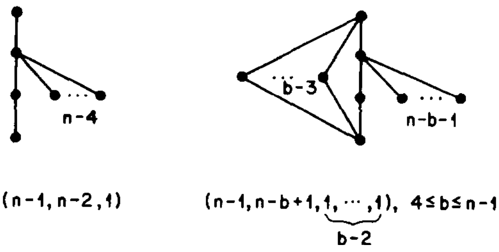

Fig. 7. Height-3 semiorders for $k = 2n - 2$.

in Fig. 7. The first of these has $e_n(n-1, n-2, 1) = (n-2)!/2$, and those for $b \geqslant 4$ have

$$e_n(n-1, n-b+1, 1, \ldots, 1) < \frac{(b-1)!\,(n-b-1)!}{2}\binom{n-2}{n-b-1} = (n-2)!/2.$$

Hence $(n-1, n-2, 1)$ is the unique height-3 maximizer.

We claim that $(n-1, n-2, 1)$ never realizes $e(n, 2n-2)$ when $n \geqslant 5$. For $n = 5$, $e_n(n-1, n-2, 1) = e_5(4, 3, 1) = 3$, but the $H = 2$ semiorder $(3, 3, 2)$ has $e_5(3, 3, 2) = 4$. For $n \geqslant 6$ the $H = 2$ semiorder $(n-2, n-3, 3)$ shown in Fig. 8 has $6(n-2)!/4!$ linear extensions for each of the two ways that point 2 can be above point 3 in a chain on $\{1, 2, 3\}$, and has $(n-3)!$ linear extensions when 2 is below 3. The total is

$$e_n(n-2, n-3, 3) = (n-3)! + (n-2)!/2,$$

which exceeds $(n-2)!/2$ for $n \geqslant 6$. □

5. Discussion

Although Theorem 1 identifies semiorders as the only posets that maximize the number of linear extensions of an n-point poset with k pairs in its ordering relation, characterization of the specific semiorders that accomplish the maxi-

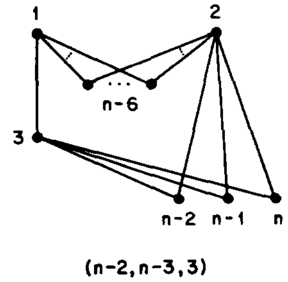

Fig. 8. A height-2 semiorder with $k = 2n - 2$.

mization appears difficult. We have done this only for $k \leqslant n$, and for $k = n + 1$ and $n \geqslant 14$, with further comments on a few other cases. As pointed out in the introduction, there is a 'small integers' problem that complicates complete characterizations for all (n, k) pairs in which k relates to n in a particular way, such as $k = n + 1$.

There are several avenues for further research. One is to focus on specific cases such as $k = n + 2$ or on all cases that do not admit semiorders of height 3 or more, i.e., those with $k < 2n - 2$.

Another avenue considers large n behavior to avoid the 'small integers' problem. For example, we have seen that, for a few cases, the optimal semiorder patterns that realize $e(n, k)$ when $k = an + b$ for fixed a and b are all similar when n is large. Is this true in general and, if so, can these patterns be described in a simple way?

References

[1] G.R. Brightwell, Semiorders and the $\frac{1}{3}$–$\frac{2}{3}$ conjecture, Order 5 (1989) 369–380.
[2] R.D. Luce, Semiorders and a theory of utility discrimination, Econometrica 24 (1956) 178–191.
[3] I. Rabinovitch, The dimension of semiorders, J. Combin. Theory, Ser. A 25 (1978) 50–61.
[4] D. Scott and P. Suppes, Foundational aspects of theories of measurement, J. Symbolic Logic 23 (1958) 113–128.
[5] W.T. Trotter, Partially ordered sets, in: R.L. Graham, M. Grötschel and L. Lovász, eds., Handbook of Combinatorics (1991).