ELSEVIER

Discrete Mathematics 256 (2002) 129–136

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Containment orders for similar ellipses with a common center

P.C. Fishburn[a,*], W.T. Trotter[b]

[a] Information Sciences Research Centre, AT&T Labs-Research, Room C227, 180 Park Avenue, Florham Park, NJ 07932, USA

[b] Arizona State University, Tempe, AZ 85287, USA

Received 13 July 1999; received in revised form 18 May 2000; accepted 30 May 2000

Abstract

Every finite two-dimensional partially ordered set has an inclusion representation by similar noncircular ellipses centered at the origin. The representing ellipses have the same ratio r of minor axis length to major axis length, and any $r \in (0,1)$ can be used for the representation.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Partial orders; Containment orders; Similar ellipses

1. Introduction

A finite irreflexive poset (partially ordered set) $P = (X, \prec)$ has an inclusion representation by circular disks centered at the origin $(0,0) \in \mathbb{R}^2$ if and only if $\prec$ on X is a weak order, i.e., if and only if the symmetric complement $\sim$ of $\prec$, defined by $u \sim v$ if neither $u \prec v$ nor $v \prec u$, is an equivalence relation. The representation maps equivalent points in X into the same circular disk, and the disk for u has a smaller radius than the disk for v if and only if $u \prec v$.

Weak orders constitute a small subset of the family of posets of order dimension [2] 1 or 2, where $\dim(P) \leqslant 2$ if $\prec$ equals the intersection of not necessarily different linear orders $<_1$ and $<_2$ on X. Our purpose is to show that the circular result of the preceding paragraph changes dramatically when we allow inclusion representations by similar noncircular ellipses centered at the origin. In particular, every finite two-dimensional poset has such an inclusion representation, and this is true for every family of similar ellipses that are not circular disks. A precise statement of our main result follows shortly. A survey of other geometric inclusion representations of posets is given in [4].

* Corresponding author. Tel.: +1-973-360-8413; fax: +1-973-360-8178.
E-mail address: fish@research.att.com (P.C. Fishburn).

0012-365X/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0012-365X(00)00366-6

Two ellipses in $\mathbb{R}^2$ are *similar* if one can be obtained from the other by translation, rotation, and uniform rescaling. In other words, ellipses E and E' are similar if both are circular disks, or if neither is circular and they have the same ratio $r<1$ of minor axis length to major axis length. For each $0<r<1$ let $\mathscr{E}(r)$ denote the set of all planar ellipses with minor to major axis length ratio r, and let $\mathscr{E}_0(r)$ be the subset of $\mathscr{E}(r)$ whose members are centered at the origin.

Theorem 1. *Suppose $P=(X,\prec)$ is a finite poset with* $\dim(P)\leqslant 2$, *and* $0<r<1$. *Then there is a map $f: X\to\mathscr{E}_0(r)$ such that, for all $u,v\in X$,*

$$u\prec v \Leftrightarrow f(u)\subset f(v). \tag{1}$$

Theorem 1 is proved in the next two sections. It is remarkable that the conclusion obviously fails for circular disks, but holds for every other origin-centered family of similar ellipses, even those with r arbitrarily near 1. We also remark that, when $r<1$ is fixed, each member of $\mathscr{E}_0(r)$ is fully determined by the length and slope of its major axis. It then follows from the degrees-of-freedom theorem in Alon and Scheinerman [1] that for every $r\in(0,1)$ there are finite three-dimensional posets that are not inclusion-representable by ellipses in $\mathscr{E}_0(r)$.

In contrast to the situation for $\mathscr{E}_0(r)$, the picture for $\mathscr{E}(r)$, where centers are unrestricted, is unclear. We know [3] that some finite three-dimensional posets are not inclusion-representable by circular disks, but do not know if this is true also for similar noncircular ellipses. Our ignorance on the matter along with Theorem 1 suggests the following two-part question for further research.

Question: Is every finite three-dimensional poset inclusion-representable by ellipses in $\mathscr{E}(r)$ for some r? If so, does the existence of such a representation depend on r in any way other than $0<r<1$?

2. Key lemma

The next section describes a scheme for positioning right-hand major axis endpoints of ellipses in $\mathscr{E}_0(r)$ for any finite two-dimensional poset that will confirm (1) when a special angle θ of the scheme is suitably small. The present section proves a key lemma that will be used to analyze the scheme. We assume throughout that $0<r<1$ and let

$$\varepsilon=r^2.$$

Lemma 2. *Suppose E_1 and E_2 are ellipses in $\mathscr{E}_0(r)$ with major semi-axis lengths of l_1 and l_2, respectively, and with angle $\tau\in(0,\pi/4)$ between their major axes. If $E_2\subset E_1$ and the two touch at boundary points, then*

$$\left(\frac{l_2}{l_1}\right)^2=1-\frac{(1-\varepsilon)\sin\tau}{2\varepsilon}\left[\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}-(1-\varepsilon)\sin\tau\right].$$

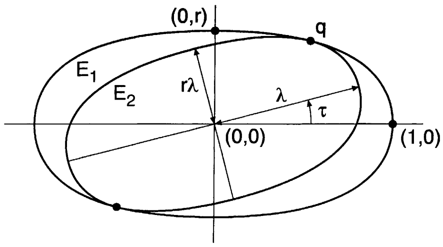

Fig. 1. Containment with touching points.

In proving this, we let $\lambda = l_2/l_1$ or, with no loss of generality, let $l_1 = 1$ and $l_2 = \lambda$. We assume also, with no loss of generality, that E_1's major axis lies on the abscissa and E_2's major axis makes an angle τ counterclockwise from the positive abscissa. We assume further that $E_2 \subset E_1$ and that the two touch at their boundaries. The upper touching point, where the boundary tangents are equal, is denoted by q: see Fig. 1. Our task is to show that λ^2 equals the right-hand side of the equation in Lemma 2.

The boundary equation of E_1 is $x^2 + y^2/r^2 = 1$, or

$$\varepsilon x^2 + y^2 = \varepsilon,$$

and the boundary equation of E_2 is

$$\varepsilon(x\cos\tau + y\sin\tau)^2 + (y\cos\tau - x\sin\tau)^2 = \varepsilon\lambda^2$$

with $0 < \lambda < 1$. We solve these for y in the vicinity of q:

$$y = r\sqrt{1-x^2}, \tag{2}$$

$$y = [x(1-\varepsilon)\cos\tau\sin\tau + r\sqrt{\lambda^2 R - x^2}]/R, \tag{3}$$

where $R = \varepsilon + (1-\varepsilon)\cos^2\tau$.

Point q is determined by two equations. The first equation says that q lies on both boundaries, so we equate the right-hand sides of (2) and (3) to get

$$Rr\sqrt{1-x^2} = x(1-\varepsilon)\cos\tau\sin\tau + r\sqrt{\lambda^2 R - x^2}. \tag{4}$$

The second, for tangency, says that $\mathrm{d}y/\mathrm{d}x$ for (2) equals $\mathrm{d}y/\mathrm{d}x$ for (3). Differentiation gives

$$\mathrm{d}y/\mathrm{d}x = -rx/\sqrt{1-x^2}, \tag{2'}$$

$$\mathrm{d}y/\mathrm{d}x = [(1-\varepsilon)\cos\tau\sin\tau - rx/\sqrt{\lambda^2 R - x^2}]/R. \tag{3'}$$

So we equate the right-hand sides of (2′) and (3′) to get

$$Rrx/\sqrt{1-x^2} = rx/\sqrt{\lambda^2 R - x^2} - (1-\varepsilon)\cos\tau\sin\tau. \tag{5}$$

We now solve (4) and (5) for the λ terms. Let

$$S = (1-\varepsilon)\cos\tau\sin\tau/r.$$

Then (4) and (5) rewritten are

$$\sqrt{\lambda^2 R - x^2} = R\sqrt{1-x^2} - Sx \tag{4}$$

$$\sqrt{\lambda^2 R - x^2} = \frac{x\sqrt{1-x^2}}{Rx + S\sqrt{1-x^2}}. \tag{5}$$

By equating the right-hand sides here, we obtain

$$(R\sqrt{1-x^2} - Sx)(Rx + S\sqrt{1-x^2}) = x\sqrt{1-x^2}. \tag{6}$$

In addition, multiplication of (4) and (5) rewritten gives

$$\lambda^2 R - x^2 = \frac{(R\sqrt{1-x^2} - Sx)x\sqrt{1-x^2}}{Rx + S\sqrt{1-x^2}},$$

which reduces to $\lambda^2 = x/(Rx + S\sqrt{1-x^2})$ or, using (6), to

$$\lambda^2 = R - \frac{Sx}{\sqrt{1-x^2}} = \varepsilon + (1-\varepsilon)\cos^2\tau - \frac{Sx}{\sqrt{1-x^2}}. \tag{7}$$

We complete the derivation of λ^2 by assessing $x/\sqrt{1-x^2}$. By (6),

$$RS(1-2x^2) = x\sqrt{1-x^2}(1 + S^2 - R^2). \tag{8}$$

We solve this for x^2, but before doing this, we note that $1 > R^2 \Rightarrow 1 + S^2 - R^2 > 0$, so (8) requires $x^2 < \frac{1}{2}$, or $x < 0.707\ldots$ for point q. More specifically,

$$\begin{aligned}
1 + S^2 - R^2 &= S^2 + (1+R)(1-R) \\
&= S^2 + [(1+\varepsilon) + (1-\varepsilon)\cos^2\tau][(1-\varepsilon) - (1-\varepsilon)\cos^2\tau] \\
&= (1-\varepsilon)^2\cos^2\tau\sin^2\tau/\varepsilon + (1-\varepsilon)\sin^2\tau[(1+\varepsilon) + (1-\varepsilon)\cos^2\tau] \\
&= \left(\frac{1-\varepsilon}{\varepsilon}\right)[(1-\varepsilon)\cos^2\tau\sin^2\tau + \varepsilon(1+\varepsilon)\sin^2\tau + \varepsilon(1-\varepsilon)\cos^2\tau\sin^2\tau] \\
&= \frac{(1-\varepsilon)(1+\varepsilon)\sin^2\tau}{\varepsilon}[\varepsilon + (1-\varepsilon)\cos^2\tau] \\
&= \frac{(1-\varepsilon^2)(\sin^2\tau)R}{\varepsilon},
\end{aligned}$$

so, in addition,

$$\begin{aligned}
\frac{RS}{1+S^2-R^2} &= \left[\frac{\varepsilon}{(1-\varepsilon^2)\sin^2\tau}\right]\frac{(1-\varepsilon)\cos\tau\sin\tau}{r} \\
&= \frac{r\cos\tau}{(1+\varepsilon)\sin\tau}.
\end{aligned} \tag{9}$$

Let

$$v = x^2,$$
$$A = (RS)^2,$$
$$B = (1 + S^2 - R^2)^2$$

and square both sides of (8) to get

$$v^2(4A+B) - v(4A+B) + A = 0.$$

Because $x^2 < \frac{1}{2}$, we use the $-$sign of the quadratic solution for v to obtain

$$\begin{aligned} v = x^2 &= \frac{4A+B-\sqrt{(4A+B)^2-4A(4A+B)}}{2(4A+B)} \\ &= \frac{1}{2}\left(1-\frac{1+S^2-R^2}{\sqrt{4A+B}}\right). \end{aligned}$$

By (8), the result just derived, and (9), we have

$$\begin{aligned} \frac{\sqrt{1-x^2}}{x} &= \frac{RS}{1+S^2-R^2}\left(\frac{1}{x^2}-2\right) \\ &= \left(\frac{RS}{1+S^2-R^2}\right)\frac{2(1+S^2-R^2)}{\sqrt{4(RS)^2+(1+S^2-R^2)^2}-(1+S^2-R^2)} \\ &= \frac{2RS/(1+S^2-R^2)}{\sqrt{4(RS)^2/(1+S^2-R^2)^2+1}-1} \\ &= \frac{2r\cos\tau/[(1+\varepsilon)\sin\tau]}{\sqrt{4\varepsilon\cos^2\tau/(1+\varepsilon)^2\sin^2\tau+1}-1} \\ &= \frac{2r\cos\tau}{\sqrt{4\varepsilon\cos^2\tau+(1+\varepsilon)^2\sin^2\tau}-(1+\varepsilon)\sin\tau} \\ &= \frac{2r\cos\tau}{\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}-(1+\varepsilon)\sin\tau}. \end{aligned} \tag{10}$$

Using (10), it follows from (7) that

$$\begin{aligned} \lambda^2 &= \varepsilon+(1-\varepsilon)\cos^2\tau-\left[\frac{(1-\varepsilon)\cos\tau\sin\tau}{r}\right]\left[\frac{\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}-(1+\varepsilon)\sin\tau}{2r\cos\tau}\right] \\ &= \varepsilon+(1-\varepsilon)\cos^2\tau-\frac{(1-\varepsilon)\sin\tau}{2\varepsilon}[\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}-(1+\varepsilon)\sin\tau] \\ &= \frac{2\varepsilon^2+2\varepsilon(1-\varepsilon)\cos^2\tau+(1-\varepsilon^2)\sin^2\tau-(1-\varepsilon)\sin\tau\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}}{2\varepsilon} \\ &= \frac{2\varepsilon+(1-\varepsilon)^2\sin^2\tau-(1-\varepsilon)\sin\tau\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}}{2\varepsilon} \\ &= 1-\frac{(1-\varepsilon)\sin\tau}{2\varepsilon}[\sqrt{(1-\varepsilon)^2\sin^2\tau+4\varepsilon}-(1-\varepsilon)\sin\tau]. \end{aligned}$$

This completes the proof of Lemma 2.

3. Proof of Theorem 1

Let

$$P_m = (\{1, 2, \ldots, m\}^2, <_0),$$

with

$$(i,j) <_0 (k,l) \quad \text{if } (i \leqslant k, j \leqslant l, i+j < k+l).$$

It is easily seen that every finite $P=(X, \prec)$ with $\dim(P) \leqslant 2$ is isomorphic to an induced subset of P_m for some m, so it suffices to prove Theorem 1 for P_m with $m \in \{2, 3, \ldots\}$.

Our proof for P_m uses a partial nesting of ellipses in $\mathscr{E}_0(r)$ that are determined by their right-hand major axis endpoints, which lie on, above, or below the positive abscissa. The designated endpoint for the ellipse assigned to $(i,j) \in \{1, 2, \ldots, m\}^2$ is denoted by $E(i,j)$. The proof uses two basic rules for locating the $E(i,j)$. First, all $E(i,j)$ with the same value of $i+j$ lie on the same circle centered at the origin. The radii of these circles increase as $i+j$ increases, so the ellipse for $i+j$ is smaller than that for $k+l$ if $i+j<k+l$. Second, all $E(i,j)$ with the same value of $j-i$ lie on the same ray emanating from the origin. The angle between the positive abscissa and the ray for constant $j-i$ is $(j-i)\theta$, where θ is a small positive angle, so the $E(i,j)$ are clustered around the positive abscissa. Those with $i=j$ lie on the abscissa. When $j'-i'=j-i$ and $i<i'$, we have $j<j'$ and $(i,j) <_0 (i',j')$, and the ellipse for (i,j) lies inside the ellipse for (i',j'). The delicate part of the proof is to choose θ and the radii of the circles that contain the $E(i,j)$ so that $E(i,j)$ is properly included in $E(k,l)$ if and only if $(i,j) <_0 (k,l)$. We now consider the details of the construction.

Let θ be a small angle with $0 < m\theta < \pi/4$. Given θ and $\varepsilon = r^2$, define $\lambda > 0$ by the equation of Lemma 2 when $\tau = \theta$, i.e.,

$$\lambda^2 = 1 - \frac{(1-\varepsilon)\sin\theta}{2\varepsilon}\left[\sqrt{(1-\varepsilon)^2\sin^2\theta + 4\varepsilon} - (1-\varepsilon)\sin\theta\right]. \tag{11}$$

We then take $E(i,j)$ as the point at distance λ^{2m-i-j} from the origin on the ray from the origin at angle $(j-i)\theta$ with the positive abscissa. As noted above, all $E(i,j)$ with the same $j-i$ are on the same ray (above the abscissa if $j>i$, below if $j<i$), and all $E(i,j)$ with the same $i+j$ lie on the circle of radius λ^{2m-i-j} centered at the origin.

Let E_{ij} denote the ellipse assigned to (i,j). Fig. 2 illustrates the right halves of $E_{j-1,j+1}$, E_{jj}, and $E_{j-1,j}$. By the definition of λ and Lemma 2, $E_{j-1,j}$ is as large as possible within the intersection of $E_{j-1,j+1}$ and E_{jj}. The boundary of $E_{j-1,j}$ touches the upper boundary of E_{jj} at q and touches the lower boundary of $E_{j-1,j+1}$ at q'. A similar picture, unique up to rotation around the origin and uniform rescaling, applies to every triple $(E_{i-1,j+1}, E_{ij}, E_{i-1,j})$, with $E_{i-1,j}$ as large as possible within $E_{i-1,j+1} \cap E_{ij}$.

To satisfy the inclusion representation of Theorem 1 for P_m, we need

$$(i,j) <_0 (k,l) \Leftrightarrow E_{ij} \subset E_{kl} \tag{12}$$

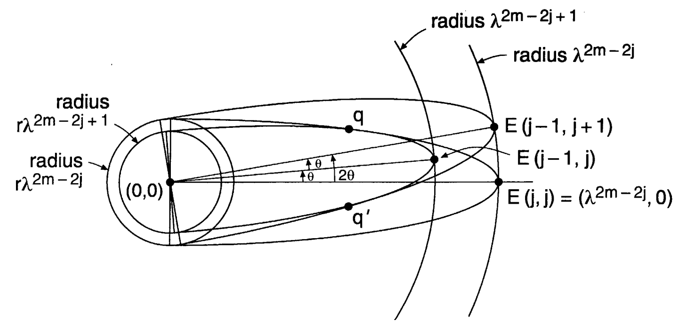

Fig. 2. $E_{j-1,j} \subset E_{j-1,j+1} \cap E_{jj}$, snugly.

for all distinct (i,j), $(k,l) \in \{1,2,\ldots,m\}^2$. The $\Rightarrow$ part of (12) follows from our construction and transitivity: if $k=i+a$, $l=j+b$, $a \geqslant 0$, $b \geqslant 0$ and $a+b \geqslant 1$, then a sequence of unit increases in the indices gives $E_{ij} \subset \cdots \subset E_{i+a,j+b}$. This part of (12) is insensitive to θ, but it may be necessary to make θ very small to ensure all desired noninclusions.

Some noninclusions demanded by (12) are obvious, as when $i+j=k+l$, or when $i+j=k+l-1$ and $|i-k| \geqslant 2$. The analytically most sensitive occur between $E_{i+1,1}$ and $E_{i,m}$, and between $E_{1,i+1}$ and $E_{m,i}$. In fact, if $E_{1,i+1} \not\subset E_{m,i}$ for $i=1,2,\ldots,m-1$, then $E_{i+1,1} \not\subset E_{i,m}$ by symmetry, and all other noninclusions required for (12) are valid. Moreover, the scaling aspects of our construction imply that if $E_{1,i+1} \not\subset E_{m,i}$ for one i, then it is true for every i. Hence to complete the proof, it suffices to show that if $\theta > 0$ is sufficiently small then

$$E_{12} \not\subset E_{m1}. \tag{13}$$

Let $n=m-2$. The angle between the major axes of E_{12} and E_{m1} is $\theta-(1-m)\theta=(n+2)\theta$, and the square of the ratio of the major semi-axis length of E_{12} to that of E_{m1} is

$$[\lambda^{2m-3}/\lambda^{2m-m-1}]^2 = (\lambda^2)^n.$$

Let E_0 denote the ellipse in $\mathscr{E}_0(r)$ whose major axis is collinear with the major axis of E_{12} and whose boundary touches the boundary of E_{m1} above the major axis of E_{m1} at point q_0 with $E_0 \subseteq E_{m1}$. By Lemma 2 with $\tau=(n+2)\theta$, the square of the ratio of the major semi-axis length of E_0 to that of E_{m1} is

$$\mu^2 = 1 - \frac{(1-\varepsilon)\sin[(n+2)\theta]}{2\varepsilon} \times \left\{\sqrt{(1-\varepsilon)^2\sin^2[(n+2)\theta]+4\varepsilon} - (1-\varepsilon)\sin[(n+2)\theta]\right\}.$$

It follows that $E_{12} \not\subset E_{m1}$ if and only if $\mu^2 < (\lambda^2)^n$, for then the major semi-axis length of E_{12} is greater than that of E_0, so a segment of E_{12} lies outside and above E_{m1} near

q_0. We conclude that (13) holds if and only iff

$$1-\frac{(1-\varepsilon)\sin[(n+2)\theta]}{2\varepsilon}\left\{\sqrt{(1-\varepsilon)^2\sin^2[(n+2)\theta]+4\varepsilon}-(1-\varepsilon)\sin[(n+2)\theta]\right\}$$
$$<\left\{1-\frac{(1-\varepsilon)\sin\theta}{2\varepsilon}[\sqrt{(1-\varepsilon)^2\sin^2\theta+4\varepsilon}-(1-\varepsilon)\sin\theta]\right\}^n. \tag{14}$$

We complete the proof of Theorem 1 by noting that (14) holds when θ is near 0.

Fix n and ε, $0<\varepsilon<1$, and let

$$t=\sin\theta.$$

When $(n+2)\theta/\varepsilon$ is near zero, we have

$$\sin[(n+2)\theta]=(n+2-\delta_0)t,$$

$$\sqrt{(1-\varepsilon)^2\sin^2[(n+2)\theta]+4\varepsilon}-(1-\varepsilon)\sin[(n+2)\theta]=2r-\delta_1,$$

$$\sqrt{(1-\varepsilon)^2\sin^2\theta+4\varepsilon}-(1-\varepsilon)\sin\theta=2r-\delta_2,$$

where the δ_i are positive and approach 0 as $\theta\to 0$. Because the δ_i can be made vanishingly small in comparison to the fixed terms they are subtracted from, i.e., $n+2$ and $2r$, we approximate the left-hand side of (14) by

$$1-\frac{(1-\varepsilon)(n+2)t}{r} \tag{15}$$

and the right-hand side of (14) by

$$\left\{1-\frac{(1-\varepsilon)t}{r}\right\}^n. \tag{16}$$

Binomial expansion of (16) gives

$$1-\frac{(1-\varepsilon)nt}{r}+\sum_{k=2}^{n}(-1)^k\binom{n}{k}\left[\frac{(1-\varepsilon)t}{r}\right]^k,$$

and, by taking $n(1-\varepsilon)t/r$ near 0 with small θ, we can ensure that $\sum_2^n$ is negligible in comparison to $(1-\varepsilon)nt/r$. In summary, (14) can be rewritten as

$$-\frac{(1-\varepsilon)(n+2)t}{r}+\delta_3<-\frac{(1-\varepsilon)nt}{r}+\delta_4,$$

where $\delta_i/[(1-\varepsilon)nt/r]\to 0$ as $\theta\to 0$ for $i=3,4$, and it follows that (14) holds for suitably small θ.

References

[1] N. Alon, E.R. Scheinerman, Degrees of freedom versus dimension for containment orders, Order 5 (1988) 11–16.

[2] B. Dushnik, E.W. Miller, Partially ordered sets, Amer. J. Math. 63 (1941) 600–610.

[3] S. Felsner, P.C. Fishburn, W.T. Trotter, Finite three-dimensional partial orders which are not sphere orders, Discrete Math. 201 (1999) 101–132.

[4] P.C. Fishburn, W.T. Trotter, Geometric containment orders: a survey, Order 15 (1999) 167–182.