
Interval order representation via shortest paths

Garth Isaak

Department of Mathematics, Lehigh University, Bethlehem, PA 18015
gisaak@lehigh.edu

Summary. Our goal in this paper is to illustrate how the representation theorems for finite interval orders and semiorders can be seen as special instances of existence results for potentials in digraphs. This viewpoint yields short proofs of the representation theorems and provides a framework for certain types of additional constraints on the intervals. We also use it to obtain a minimax theorem for the minimum number of endpoints in a representation. The techniques are based on techniques used by Peter Fishburn in proving results about bounded representations of interval orders.

1 Introduction

Recall that an interval order is an asymmetric binary relation \prec on a set U that satisfies ($a \prec x$ and $b \prec y$ implies $a \prec y$ or $b \prec x$) for all $a, b, x, y \in U$. These are (strict partial) orders as transitivity is implied by this definition. The name interval applies because these orders can be represented by a set of intervals in a linear order with the natural relation ‘less than’ for the intervals. A (closed real) interval representation of a strict order (U, \prec) is a set of closed real intervals $[l_x, r_x]$ for $x \in U$ such that $x \prec y$ if and only if $r_x < l_y$. A $\mathbf{2} + \mathbf{2}$ in an order is the disjoint union of two chains each with two elements. That is, a $\mathbf{2} + \mathbf{2}$ is a set of four elements x, y, a, b such that $a \prec x, b \prec y, a \sim y, b \sim x$ (with also $x \sim y, a \sim b$ implied by transitivity). Here we use the notation \sim to represent incomparability, $x \sim y$ if and only if $x \not\prec y$ and $y \not\prec x$. A $\mathbf{2} + \mathbf{2}$ corresponds to a violation of the interval order condition above so we can take as our definition that an interval order is a strict partial order with no $\mathbf{2} + \mathbf{2}$.

This result that an order can be represented by intervals if and only if it has no $\mathbf{2} + \mathbf{2}$ was anticipated by Weiner in 1914 (see Fishburn

and Monjardet [5]) and shown by Fishburn [2], [3]. Following Bogart [1] we will refer to this as Fishburn's Theorem. When the ground set is finite we can use either open or closed intervals in the real numbers for the interval representation. When the ground set is infinite and in particular uncountable, things are a little more complicated. A linear order other than the reals may be required. For more on this see Fishburn's book [4]. For this paper we will stick to the finite version and consider real number representations with the goal of seeing a connection to potentials, shortest paths and negative cycles in digraphs. We will assume that all orders considered in this paper are finite. We will not attempt to survey the many different proofs that have been given for these theorems nor the various related results. See Fishburn's book [4] for a very nice description of results on interval orders.

Fishburn's Theorem (for finite interval orders): A finite strict partial order has a closed real interval representation if and only if it has no $\mathbf{2} + \mathbf{2}$.

A finite unit interval order (also called a semiorder) is an interval order that has a real representation in which all of the intervals have the same length (which by scaling we can assume to be 1). Such a representation will be called a unit interval representation. Again, see [4] for a discussion of the infinite case. A $\mathbf{3} + \mathbf{1}$ is the disjoint union of a chain with 3 elements and a chain with 1 element. That is, r, s, t, u such that $r < s < t$ and $u \sim r, u \sim t$ (with $u \sim s$ implied by transitivity). This is the extra condition needed for unit interval representations as shown by Scott and Suppes [10].

Scott-Suppes' Theorem (for finite unit interval orders): A finite strict partial order has a closed real unit interval representation if and only if it has no $\mathbf{2} + \mathbf{2}$ or $\mathbf{3} + \mathbf{1}$.

We will give short proofs of these results using shortest paths in an associated digraph and also show how this framework can be used to get results about representations with with certain side constraints on the intervals. This technique was first used by Fishburn [6] for representations with bounds on the interval lengths and later by the author [7], [8] for representations with bounds and integral endpoints. A similar technique was used by Pirlot [9] for semiorder representation questions although with a different set of inequalities. The representation theorems are implicit in these results. However, without the bounds we can

get short proofs of the representation theorems. Finally we will use the framework to obtain a minimax theorem for the minimum number of endpoints in a representation.

2 Shortest Paths and Potentials

We first give a brief review of basic results on shortest paths and potentials which will provide the framework for our proofs. This can be found, for example in Schrijver, page 108 [11].

A weighted digraph is a set V of vertices along with a set A of ordered pairs (x, y) of vertices called arcs. Each arc (x, y) has an associated length $w(xy)$. A potential function $p(x)$ defined on the vertices is a function satisfying $p(y) - p(x) \leq w(xy)$ for all arcs (x, y) . That is, the potential value at y is at most the potential at x plus the length of arc (x, y) . An $x - y$ path in a digraph is a sequence of distinct vertices $x = x_1, x_2, \dots, x_t = y$ such that (x_i, x_{i+1}) is an arc for $i = 1, 2, \dots, t-1$. A cycle is the same as a path except that (x_t, x_1) is also an arc. The length of a path (cycle) is the sum of the arc lengths along the path (cycle). Let $\bar{p}(y)$ denote the shortest length of a path ending at y . It is a basic result that these shortest path lengths are defined for all vertices if and only if the graph has no negative cycle (i.e., a cycle with negative length). It is easy to see that if $\bar{p}(x)$ is the length of a shortest path ending at x for each x (assuming that these are well defined) then this is a potential. If a digraph has a negative cycle with length $-c < 0$ then ‘adding’ the inequalities for the arcs on the cycle produces the inconsistency $0 \leq -c$ showing that there is no potential.

Hence a digraph has a potential function if and only if it has no negative cycle. Furthermore, if there is no negative cycle then the lengths of shortest paths ending at each vertex yield a potential.

3 Interval Orders

That an order with an interval representation has no **2 + 2** is easy to check. Our goal is to show the converse: a finite order with no **2 + 2** has a closed interval representation. For each element $x \in U$ create two variables $p(r_x)$ and $p(l_x)$ (which will correspond to left and right endpoints of intervals in a representation). Consider the p as representing placement of the endpoints and in what follows (with a slight abuse of notation) a potential function.

Let γ be a positive number. Consider the following inequalities:

- (C) For all $x \prec y$, $p(r_x) \leq p(l_y) - \gamma$, equivalently $p(r_x) - p(l_y) \leq -\gamma$
(I) For all $x \sim y$, $p(r_x) \geq p(l_y)$, equivalently $p(l_y) - p(r_x) \leq 0$.

The intervals would be $[p(l_x), p(r_x)]$. The inequalities C enforce ‘less than’ for intervals of comparable elements and the inequalities I enforce ‘not less than’ for incomparable elements. It is not difficult to check that (U, \prec) has an interval representation using intervals $[p(l_x), p(r_x)]$ if and only if the $p(r_x), p(l_x)$ are a solution to the the system of inequalities above. The γ is just a convenience to avoid writing strict inequalities. Note that when $x = y$ in I we have $p(r_x) \geq p(l_x)$, ensuring that these really are intervals, with the right endpoints at least as large as the left endpoints. Also, when $x \sim y$, switching the roles of x and y we see that we have both $p(l_y) - p(r_x) \leq 0$ and $p(l_x) - p(r_y) \leq 0$

Each inequality has two variables, one with coefficient $+1$ and one with coefficient -1 . We then recognize the inequalities as those for a potential function on a particular digraph.

For a given order (U, \prec) we define (with a slight abuse of notation) a digraph D_U with vertices $\{l_x, r_x | x \in U\}$ and arcs $C \cup I$ where $C = \{(l_y, r_x) | x \prec y\}$ with length $-\gamma$ for some positive number γ and $I = \{(r_x, l_y) | x \sim y\}$ with length 0. Then, from the preceding section, (U, \prec) has an interval representation if and only if D_U has no negative cycles. Furthermore the length of a shortest path ending at r_x can be used for the right endpoints $p(r_x)$ and similarly the the length of a shortest path ending at l_y can be used for the left endpoints $p(l_y)$. We note that directly writing down D_U one can fairly easily show this claim without going through the idea of potentials. Looking at shortest paths, the length 0 on an arc (r_x, l_y) for $x \sim y$ forces a shortest path ending at l_y to have length no more than a shortest path ending at r_x . That is $p(l_y)$, would be at most $p(r_x)$ and so $x \not\prec y$. Similarly the length $-\gamma$ on arc (l_y, r_x) for $x \prec y$ forces a shortest path ending at r_x to have length strictly less a shortest path ending at l_y . That is, $p(r_x)$ would be less than $p(l_y)$ and so $x \prec y$. The framework of potentials is useful as motivation for why we construct the digraph in this manner and for easily yielding a proof that the technique works.

Proof of Fishburn’s Theorem for finite orders:

From the discussion above we need to show that (U, \prec) has a **2 + 2** if and only if D_U contains a negative cycle.

We have already noted that it is easy to check directly that an order with a $\mathbf{2} + \mathbf{2}$ has no interval representation. We can also show this using the digraph: If $a \prec x, b \prec y, a \sim y, b \sim x$ is a $\mathbf{2} + \mathbf{2}$ then l_x, r_a, l_y, r_b is a cycle with length $-2\gamma < 0$.

We now need to show that if D_U contains a negative cycle then (U, \prec) has a $\mathbf{2} + \mathbf{2}$. Observe that D_U is bipartite. Partition the vertex set into $R = \{r_x | x \in U\}$ and $L = \{l_y | y \in U\}$. Then there are two types of arcs, each with one end in R and the other in L . Arcs from L to R have length $-\gamma$ and arcs from R to L have length 0. Any cycle must alternate between these two types of arcs and thus any cycle is negative. So it is enough to show that if D_U contains a cycle then (U, \prec) has a $\mathbf{2} + \mathbf{2}$.

Consider a cycle with the minimum number of vertices. It is easy to see that it cannot have exactly two vertices. Since the digraph is bipartite the cycle contains l_x, r_a, l_y, r_b for some x, a, y, b (not necessarily distinct). The arcs imply that $a \prec x, b \prec y$ and $a \sim y$. If $b \prec x$ then (l_x, r_b) is an arc and replace the segment with this arc for a negative cycle with fewer vertices. If $x \prec b$ then in the order $a \prec x \prec b \prec y$ and by transitivity $a \prec y$ contradicting $a \sim y$. So $x \sim b$. Thus we have $y \sim a, a \prec x, b \prec y$ and $x \sim b$. Using transitivity it is easy to see that x, y, a, b are distinct and so x, y, a, b induce a $\mathbf{2} + \mathbf{2}$. \square

To prove Fishburn's Theorem with open intervals in the representation we would give an almost identical proof except that the arcs in I would have length $-\epsilon$ for some positive ϵ and the arcs in C would have length 0 (and the corresponding changes in the inequalities C and I).

4 Unit Interval Orders

That an order with a unit interval representation has no $\mathbf{2} + \mathbf{2}$ and no $\mathbf{3} + \mathbf{1}$ is easy to check. Our goal is to show the converse: a finite order with no $\mathbf{2} + \mathbf{2}$ and no $\mathbf{3} + \mathbf{1}$ has a closed unit interval representation. One approach is to use the same model as for interval orders and replace the constraints $p(r_x) \geq p(l_x)$ that right endpoints are at least as large as left endpoints with constraints $p(r_x) = p(l_x) + 1$ written as $p(r_x) - p(l_x) \leq 1$ and $p(l_x) - p(r_x) \leq -1$. Then appropriately adjust the corresponding digraph. Instead we will use a single variable for left endpoints, setting the right endpoints to be 1 more than the left.

We will use the same notation for D_U as the previous section however the construction here is different.

For each element $x \in U$ create a variable $p(l_x)$ (which will correspond to left endpoint of intervals in a representation). Consider the p as representing placement of the left endpoint and in what follows a potential function.

Let γ be a positive number. Consider the following inequalities:

- (C) For all $x \prec y$, $p(l_x) + 1 \leq p(l_y) - \gamma$,
 equivalently $p(l_x) - p(l_y) \leq -(1 + \gamma)$
 (I) For all $x \sim y$ with $x \neq y$, $p(l_x) + 1 \geq p(l_y)$,
 equivalently $p(l_y) - p(l_x) \leq 1$.

The intervals would be $[p(l_x), p(l_x) + 1]$. The inequalities C enforce ‘less than’ for intervals of comparable elements and the inequalities I enforce ‘not less than’ for incomparable elements. It is not difficult to check that (U, \prec) has an unit interval representation using intervals $[p(l_x), p(l_x) + 1]$ if and only if the $p(l_x)$ are a solution to the the system of inequalities above. The γ is just a convenience to avoid writing strict inequalities.

As with the interval order case we recognize the system of inequalities as corresponding to those for a potential in a particular digraph.

For a given order (U, \prec) we define (with a slight abuse of notation) a digraph D_U with vertices $\{l_x | x \in U\}$ and arcs $C \cup I$ where $C = \{(l_y, l_x) | x \prec y\}$ with length $-(1 + \gamma)$ for some positive number γ and $I = \{(l_x, l_y) | x \sim y\}$ with length 1. Observe that if $x \sim y$ we have both arcs (l_x, l_y) and (l_y, l_x) with length 1. From the connection to potentials (U, \prec) has a unit interval representation if and only if D_U has no negative cycles. Furthermore length of a shortest path ending at l_y can be used for the left endpoints $p(l_x)$. We note that directly writing down D_U one can fairly easily show this claim without going through the idea of potentials. Looking at shortest paths, the length 1 on an arc (l_x, l_y) for $x \sim y$ forces a shortest path ending at l_y to have length at most 1 more than a shortest path ending at l_x . That is the right endpoint $p(l_x) + 1$, would be at most $p(l_y)$ and so $x \not\prec y$. Similarly the length $-(1 + \gamma)$ on arc (l_y, l_x) for $x \prec y$ forces a shortest path ending at l_x to have length more than 1 less a shortest path ending at l_y . That is, the right endpoint $p(l_x) + 1$ would be less than $p(l_y)$ and so $x \prec y$.

Proof of Scott-Suppes' Theorem for finite orders:

From the discussion on potentials we need to show that (U, \prec) has a $\mathbf{2} + \mathbf{2}$ or a $\mathbf{3} + \mathbf{1}$ if and only if D_U contains a negative cycle.

We have already noted that it is easy to check directly that an order with a $\mathbf{2} + \mathbf{2}$ or a $\mathbf{3} + \mathbf{1}$ has no unit interval representation. We can also show this using the digraph: If $a \prec x, b \prec y, a \sim y, b \sim x$ is a $\mathbf{2} + \mathbf{2}$ then l_x, l_a, l_y, l_b is a cycle with length $-2\gamma < 0$. If $r \prec s \prec t, u \sim r, u \sim t$ is a $\mathbf{3} + \mathbf{1}$ then l_t, l_s, l_r, l_u is a cycle with length $-2\gamma < 0$.

We now need to show that if D_U contains a negative cycle then (U, \prec) has a $\mathbf{2} + \mathbf{2}$ or a $\mathbf{3} + \mathbf{1}$.

Observe that since arcs in C have length $-(1 + \gamma)$ and arcs in I have length $+1$ there are at least as many arcs from C as from I in any negative cycle. Consider a negative cycle with the minimum number of vertices. Using transitivity in the order, it is easy to see that it cannot have 3 vertices. Since there are at least as many C arcs as I arcs, the cycle either alternates between C arcs and I arcs or contains 2 consecutive arcs from C .

In the first case consider l_x, l_a, l_y, l_b along the cycle with $(l_x, l_a), (l_y, l_b)$ in C and (l_a, l_y) in I . The arcs imply that $a \prec x, b \prec y$ and $a \sim y$. If $b \prec x$ then (l_x, l_b) is an arc in C and replace the segment with this arc for a cycle with fewer vertices. If $x \prec b$ then in the order $a \prec x \prec b \prec y$ and by transitivity $a \prec y$ contradicting $a \sim y$. So $x \sim b$. Thus we have $y \sim a, a \prec x, b \prec y$ and $x \sim b$. So x, y, a, b induce a $\mathbf{2} + \mathbf{2}$.

In the second case some pair of consecutive arcs from C is followed by an arc from I . (If all arcs are from C we get a violation of transitivity in the order.) Consider l_t, l_s, l_r, l_u along the cycle with (l_t, l_s) and (l_s, l_r) in C and (l_r, l_u) in I . The arcs imply that $s \prec t, r \prec s$ and $r \sim u$. If $u \prec t$ then (l_t, l_u) is an arc from C and replace the segment by this arc for a cycle with fewer vertices. If $t \prec u$ then in the order $r \prec s \prec t \prec u$ and by transitivity $r \prec u$ contradicting $r \sim u$. So $t \sim u$. Thus we have $s \prec t, r \prec s$ and $r \sim u$ and $t \sim u$. So r, s, t, u induce a $\mathbf{3} + \mathbf{1}$. \square

To prove the Scott-Suppes' Theorem with open intervals in the representation we would give an almost identical proof except that the arcs in I would have length $1 - \epsilon$ for some positive ϵ and the arcs in C would have length -1 (and the corresponding changes in the inequalities C and I).

5 Side Constraints

We next briefly note how the framework of inequalities and shortest paths can be used in a more general setting. We can, for example, place lower and upper bounds on the interval lengths in a representation. That is, given numbers $0 \leq \alpha(x) \leq \beta(x)$ for each element $x \in U$ we add the constraints $p(r_x) - p(l_x) \geq \alpha(x)$ and $p(r_x) - p(l_x) \leq \beta(x)$ for all x . These specify lower and upper bounds on each interval length. The additional constraints add additional arcs to the digraph D_U . These representations have been examined by Fishburn, [4], [6] and with the additional requirement that the endpoints be integral in [7], [8]. The use of shortest paths to construct an interval representation does imply an efficient algorithm for determining when an order has a representation subject to the additional constraints on interval length. However, while the result that there is a representation if and only if D_U contains no negative cycle is in a sense a characterization theorem it would be good to translate this to minimal forbidden suborders as negative cycles implied a $\mathbf{2} + \mathbf{2}$ in the proof of Fishburn's Theorem. Unfortunately this seems to be fairly complicated. The minimal forbidden suborders for a representation with integral endpoints and interval lengths between 0 and some positive integer α is given in [7] but it is very messy. The same situation with a lower bound of 1 on interval lengths appears even messier although a possible infinite list is suggested in [7].

Looking at the inequalities and not using the digraph model one can consider interval representations that are 'optimal' by some other measure. For example, when there are lower bound on interval lengths, one could specify a utility for each element and seek an interval representation that minimizes the sums of the interval lengths weighted by the utilities. Linear programming algorithms and linear programming duality immediately give an efficient algorithm to find such a 'weighted least length' interval representation as well as minimax theorem. It may be interesting to investigate if this can be translated to a more direct statement in terms of the order.

6 Magnitude

In this section we will show that the representation for interval orders constructed with interval endpoints determined by the values of shortest paths in D_U uses the minimum number of distinct endpoint values among all interval representations. This minimum value is called the magnitude and discussed in section 2.3 of [4]. Magnitude is presented

from a slightly different perspective here. In particular we obtain a min-max result equating the minimum number of endpoints to the maximum ‘size’ of a certain suborder.

Since the arc lengths in D_U are 0 and $-\gamma$, the shortest path lengths $p(r_x), p(l_x)$ take on values $0, -\gamma, -2\gamma, \dots, -m\gamma$ for some nonnegative integer m . The interval representation constructed from D_U uses $m + 1$ distinct endpoints. We will show that the magnitude is $m + 1$.

In this section we will assume that all orders are interval orders. That is, D_U does not have a negative cycle.

We use a particular class of orders, called sequences of linked chains to obtain a lower bound on the magnitude and show that this bound matches the number of endpoints used in the shortest paths construction. These orders are special cases of the picycles introduced by Fishburn [6] and what are called sequences of linked chains in [7].

A sequence of linked chains in an interval order (U, \prec) is a sequence of chains $C_i = u_{i1} \succ u_{i2} \succ \dots \succ u_{in_i}$ for $i = 1, 2, \dots, k$ such that the chains are nontrivial ($n_i \geq 2$) and for $i = 1, 2, \dots, k - 1$ and $u_{in_i} \sim u_{(i+1)1}$. In addition, elements of the chains are distinct except that possibly an element can appear on two chains as $u_{in_i} = u_{j1}$ for some $j > i + 1$. The **size** of the sequence of linked chains is $1 + \sum_{i=1}^k (n_i - 1)$.

Observe that we have not specified all the relations between the elements so that different orders on the same set of elements can be sequences of linked chains. The relations that have been specified by the chains are enough for our proof.

Theorem: For a finite interval order (U, \prec) , the minimum number of distinct endpoints in an interval representation is equal to the maximum size of a sequence of linked chains in (U, \prec) . Furthermore, the representation constructed from shortest paths in D_U uses this many endpoints.

Proof: Consider any interval representation of an order that contains a sequence of linked chains with size $m + 1$. Use the notation $[l(u_{ij}), r(u_{ij})]$ for the intervals. From $u_{i(n_i-1)} \succ u_{in_i}$ and $u_{in_i} \sim u_{(i+1)1}$ we get $l(u_{i(n_i-1)}) > r(u_{in_i}) \geq l(u_{(i+1)1})$. Also, from the chains, the left endpoints satisfy $l(u_{i1}) > \dots > l(u_{in_i})$. Thus the endpoints $l(u_{ij})$ for $i = 1, 2, \dots, k$ and $j = 1, 2, \dots, n_i - 1$ are distinct. (It is possible that $l(u_{in_i}) = l(u_{(i+1)1})$.) Finally, $l(u_{kn_k})$ is distinct from any of these end-

points. Thus any representation requires at least $1 + \sum_{i=1}^k (n_i - 1) = m + 1$ endpoints.

As noted above, the length of a shortest path in D_U will be $-m\gamma$ for some nonnegative integer m . Partition the I arcs in D_U into $I_1 = \{(r_x, l_x) | x \in U\}$ and $I_2 = \{(r_x, l_y) | x \sim y, x \neq y\}$. Since arcs in I have length 0 we can assume that a shortest path is of the form $l_{u_1}, r_{u_2}, \dots, r_{u_i}$. Deleting the arcs from I_2 leaves subpaths which alternate between C arcs and I_1 arcs. That is, subpaths $l_{u_{i_1}}, r_{u_{i_2}}, l_{u_{i_3}}, \dots, r_{u_{i(n_i-1)}}, l_{u_{i(n_i-1)}}, r_{u_{i n_i}}$ for some $i = 1, 2, \dots, k$ (where we assume the i^{th} subpath appears before the $(i + 1)^{\text{st}}$). These correspond to chains $u_{i_1} \succ u_{i_2} \succ \dots \succ u_{i n_i}$ in the order. From the deleted I_2 arcs we have $u_{i n_i} \sim u_{(i+1)_1}$ and hence the order contains a sequence of linked chains. For each of the chains there are $n_i - 1$ corresponding arcs from C in the path in D_U and these are the only arcs from C . Hence the path length satisfies $m = \sum_{i=1}^k (n_i - 1)$ and the sequence of linked chains has size $m + 1$.

The first paragraph shows that the minimum number of endpoints in an interval representation of (U, \prec) is at least the maximum size of a sequence of linked chains. If the shortest path length in D_U is $-m\gamma$ then the construction in the proof of Fishburn's theorem yields a representation with exactly $m + 1$ distinct endpoints. The second paragraph shows that (U, \prec) contains a sequence of linked chains of size $m + 1$. \square

Acknowledgement: I would like to thank Peter Fishburn for his valuable support and encouragement early in my career.

References

1. K.P. Bogart (1993) An obvious proof of Fishburn's interval order theorem. *Discrete Math.* 118:239–242
2. P.C. Fishburn (1970) Intransitive indifference with unequal indifference intervals. *J. Math. Psychology* 7:144 – 149
3. P.C. Fishburn (1970) *Utility theory for decision making*. Wiley, New York
4. P.C. Fishburn (1985) *Interval orders and interval graphs*. Wiley, New York
5. P.C. Fishburn, B. Monjardet (1992) Norbert Wiener on the theory of measurement (1914, 1915, 1921). *J. Math. Psych.* 36:165–184
6. P.C. Fishburn (1983) Threshold-bounded interval orders and a theory of picycles. *SIAM J. Alg. Disc. Meth.* 4:290–305

7. G. Isaak (1990) Odd forests, reversing numbers and discrete representations of interval orders. Ph.D. Thesis, Rutgers University, New brunswick, New Jersey
8. G. Isaak (1993) Bounded discrete representations of interval orders. *Discrete Appl. Math.* 44:157–183
9. M. Pirlot (1990) Minimal representation of a semiorder. *Theory and Decision* 28:109–141
10. D. Scott, P. Suppes (1958) Foundational aspects of theories of measurement. *J. Symb. Logic* 23:113–128
11. A. Schrijver (2003) *Combinatorial Optimization: Polyhedra and Efficiency*. Springer, Berlin Heidelberg New York