Discrete Applied Mathematics 40 (1992) 19–28
North-Holland

The communication complexity of interval orders

Ulrich Faigle

Faculty of Applied Mathematics, University of Twente, NL-7500 AE Enschede, Netherlands

Rainer Schrader

Institut für Operations Research, Universität Bonn, D-5300 Bonn 1, Germany

György Turán

Department of Mathematics and Computer Science, University of Illinois, Chicago, Il 60637, USA; and Automata Theory Research Group, Hungarian Academy of Sciences, Szeged, Hungary

Received 11 January 1989
Revised 15 March 1990

Abstract

Faigle, U., R. Schrader and G. Turán, The communication complexity of interval orders, Discrete Applied Mathematics 40 (1992) 19–28.

The communication complexity of interval orders is studied within the following model. Two players choose two elements x and y and want to determine whether $x < y$ holds by exchanging as few bits of information as possible. It is shown that an optimal one-way protocol exists by first establishing a rank-optimality result for a subclass of generalized interval orders. It turns out that the deterministic and nondeterministic communication complexities coincide for generalized interval orders. The analogous statement for the complementary relation is true for interval orders in the strict sense while it need not hold for generalized interval orders.

Keywords. Communication complexity, interval order.

1. Introduction

Interest in the communication complexity arises from the desire to establish lower bounds for the complexity of VLSI-computations. The model proposed by Yao [11] views a chip as a device to compute the value of a Boolean (0,1)-valued function

Correspondence to: Professor U. Faigle, Faculty of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, Netherlands.

$f(x, y)$ whose input is represented by the pair (x, y) of variables. Thinking of x as being associated with a player I and y with a player II, the communication complexity measures the minimum amount of information bits the two players have to exchange in a cooperative effort in order to evaluate $f(x, y)$.

The theory of communication complexity allows many analogies with the complexity theory of Turing machines (viz. Babai et al. [2]). In particular, deterministic as well as nondeterministic complexity are meaningful notions. As it turns out, the study of these notions amounts to the study of combinatorial properties of (0,1)-matrices, namely the tables associated with the functions $f(x, y)$. Moreover, these combinatorial properties often reveal themselves as properties of underlying (partially) ordered sets (viz. Lovász and Saks [9]).

In the present investigation, we are concerned with the communication complexity of Boolean functions $f(x, y)$ that are the incidence functions of (strict) interval order relations. In other words, we study the case where player I chooses an element x and player II an element y of an interval order. It is to be determined whether the relation "$x<y$" holds. We show in Section 4 that a deterministic one-way protocol for deciding this question exists whose complexity achieves the rank lower bound of Mehlhorn and Schmidt [10] and thus is seen to be optimal.

Furthermore, we prove that also the nondeterministic communication complexity respects the rank lower bound. Our analysis is based on the close connection between the nondeterministic complexity and the so-called setup number for interval orders and especially the defect lower bound studied by Gierz and Poguntke [6]. Generalizing the results of Faigle and Schrader [3], we exhibit in Section 3 a large class of generalized interval orders to be rank (defect respectively) optimal, which may be interesting in its own right. This fact is then used in Section 4 to establish the bound. Let us just remark that there is a close relationship between the communication complexity and "classical" parameters of more general classes of ordered sets. For more details, see Faigle and Turán [4].

It follows from the definition that the deterministic communication complexity of an order is the same as that of its complementary relation. The analogous statement about the nondeterministic complexity may be false for generalized interval orders. Therefore somewhat surprisingly, it is seen to be true for the class of interval orders in the strict sense (Corollary 4.4). It would be interesting to know if the nondeterministic complexity of a generalized interval order is at least always within factor 2, say, of the nondeterministic complexity of its complementary relation.

2. Communication complexity of ordered sets

In this section, we describe the computational model we use for the communication complexity. Rather than working with Yao's [11] original model, we will take over the covenient model proposed by Lovász and Saks [9].

Let E_1 and E_2 be two finite sets and $f: E_1 \times E_2 \to \{0,1\}$ a Boolean function. We are interested in the following decision problem associated with the binary relation $Q = f^{-1}(1)$. Determine whether a given input (x, y) satisfies $f(x, y) = 1$, i.e., $(x, y) \in Q$.

A *deterministic communication protocol* for recognizing Q is a decision tree T whose nodes are of two types. An internal node of type i ($i = 1, 2$) is labeled by a function from E_i to the set of children of that node. A leaf of type i is labeled by a function from E_i to the set {YES, NO}. Each input $(x, y) \in E_1 \times E_2$ specifies a unique path from the root to a leaf of T in such a way that Q consists exactly of those inputs (x, y) yielding the outcome YES.

T is a *one-way* protocol if T has depth 1. The *cost* of an internal node in T equals the (rounded) logarithm (here always assumed relative to base 2) of the number of its children, i.e., the number of bits needed to specify a child. The cost $c(P)$ of a path P from the root to a leaf in T is the sum of the costs of its internal nodes. Thus the *complexity* $c(T)$ of the protocol T can be defined as

$$c(T) = \max\{c(P) \colon P \text{ path in } T\}.$$

The *deterministic* (*communication*) *complexity* $\mathrm{cc}(Q)$ of the binary relation Q is given by

$$\mathrm{cc}(Q) = \min\{c(T) \colon T \text{ protocol for } Q\}.$$

We say that Q is *one-way optimal* if there exists an optimal protocol for Q which is one-way. Note that the one-way complexity of Q equals $\lceil \log d \rceil$, where d is the smaller of the number of distinct rows and the number of distinct columns in the incidence matrix of Q.

A *proof scheme* for the relation $Q \subseteq E_1 \times E_2$ consists of a set P of *proofs* together with two *verification relations* $V_1 \subseteq E_1 \times P$ and $V_2 \subseteq E_2 \times P$ such that $(x, y) \in Q$ if and only if there exists a proof $p \in P$ with the property $(x, p) \in V_1$ and $(y, p) \in V_2$.

For all $p \in P$, consider now the sets

$$R(p) = \{(x, y) \in Q \colon (x, p) \in V_1 \text{ and } (y, p) \in V_2\}.$$

$R(p)$ is a *rectangle* of Q, i.e., there are subsets $F_1 \subseteq E_1$ and $F_2 \subseteq E_2$ such that $R(p) = F_1 \times F_2$. As Lipton and Sedgewick [8] have observed, a proof scheme of Q may therefore equivalently be defined as a set $\mathcal{R}$ of rectangles whose union equals Q. The *nondeterministic* (*communication*) *complexity* $\mathrm{cc}^*(Q)$ is the number

$$\mathrm{cc}^*(Q) = \min\{\log|\mathcal{R}| : \mathcal{R} \text{ proof scheme for } Q\}.$$

The lower bound $\mathrm{cc}^*(Q) \le \mathrm{cc}(Q)$ is not hard to see (each path in a protocol T for Q leading to a YES-leaf specifies a rectangle of Q).

Thinking of the function $f: E_1 \times E_2 \to \{0,1\}$ as a (0, 1)-matrix with set E_1 of rows and set E_2 of columns, the *rank* $r(Q)$ of the relation $Q = f^{-1}(1)$ is well defined as the rank of its (0, 1)-incidence matrix. This leads to another important lower bound for the deterministic complexity, due to Mehlhorn and Schmidt [10]:

Theorem 2.1. *Relative to any field K or, more generally, the ring of integers,*

$$\log(r(Q)) \le \mathrm{cc}(Q).$$

We will be concerned with the following situation. P is a (partial) order on the set E and the Boolean function f is defined relative to $E_1 = E_2 = E$ via

$$f(x,y) = \begin{cases} 1, & \text{if } x<y \text{ in } P, \\ 0, & \text{otherwise.} \end{cases}$$

It is convenient to write the incidence matrix associated with P via f in triangular form by listing both the rows and columns according to a linear extension of P. Recall that a listing $L = x_1x_2\cdots x_i\cdots x_n$ of the ground set E is a *linear extension* (also called a *topological sorting*) of P if for all i, j,

$$x_i < x_j \quad \text{implies} \quad i<j.$$

Define the *lineality* of the linear extension L as

$$l(L) = |\{(x_i, x_{i+1}) : x_i < x_{i+1}\}|.$$

Gierz and Poguntke [6] have observed:

Theorem 2.2. *If L is a linear extension of the order P, then*

$$l(L) \le r(P).$$

We call the order P *rank-optimal* if there exists a linear extension L of P such that $l(L) = r(P)$. (Note that our rank-optimal orders are exactly the *defect-optimal* orders of Gierz and Poguntke [6]. The different terminology arises from their consideration of the *setup number* $s(L) = (n-1) - l(L)$ of the linear extension L.)

Theorem 2.3. *If L is a linear extension and $\mathscr{R}$ a proof scheme of P, then*

$$l(L) \le |\mathscr{R}|.$$

Proof. Assume $L = x_1x_2\cdots x_i\cdots x_n$ and consider two distinct pairs (x_i, x_{i+1}) and (x_j, x_{j+1}) such that $f(x_i, x_{i+1}) = f(x_j, x_{j+1}) = 1$.

If $i<j$ then $f(x_j, x_{i+1}) = 0$ by the definition of a linear extension. Hence the two pairs cannot be contained in the same rectangle $R \in \mathscr{R}$. Thus $\mathscr{R}$ comprises at least $l(L)$ rectangles. □

Corollary 2.4. *If P is rank-optimal, then*

$$\log(r(P)) \le \mathrm{cc}^*(P) \le \mathrm{cc}(P).$$

With every Boolean function $f: E_1 \times E_2 \to \{0,1\}$, we may associate the complementary function $\bar{f}: E_1 \times E_2 \to \{0,1\}$ via

$$\bar{f}(x,y) = 1 + f(x,y) \pmod 2.$$

Thus $Q = f^{-1}(1)$ gives rise to the binary relation $\bar{Q} = \bar{f}^{-1}(1) = (E_1 \times E_2) \setminus Q$. Apparently,

$$\mathrm{cc}(Q) = \mathrm{cc}(\bar{Q}),$$

while equality may not hold for the nondeterministic complexities. We set

$$\overline{\mathrm{cc}}^*(Q) := \mathrm{cc}(\bar{Q}).$$

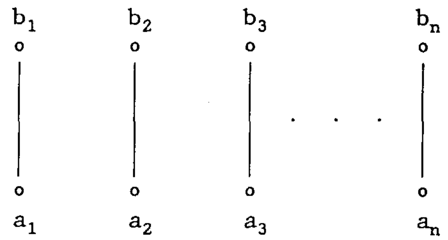
The following upper bound on the deterministic complexity in terms of the nondeterministic complexities has been derived by Aho et al. [1] (see also Halstenberg and Reischuk [7]).

Theorem 2.5. *If Q is an arbitrary binary relation, then*

$$\max\{\mathrm{cc}^*(Q), \overline{\mathrm{cc}}^*(Q)\} \le \mathrm{cc}(Q) \le \mathrm{O}(\mathrm{cc}^*(Q) \cdot \overline{\mathrm{cc}}^*(Q)).$$

The question naturally arises whether an analog of Theorem 2.3 also holds for the complementary relation $\bar{P}$ associated with the order P. The following example shows that this is generally not the case. In Section 4, however, we will exhibit the analogous property for interval orders (Corollary 4.4).

Example 2.6. Let the order P be represented by the following Hasse diagram.



Then $L = a_1 b_1 a_2 b_2 \cdots a_n b_n$ satisfies $l(L) = n$. The complementary relation $\bar{P}$ can be covered with $\mathrm{O}(\log n)$ rectangles.

3. A class of rank-optimal orders

Let P be an order on E. With each $x \in E$ we associate the open lower ideal

$$N(x) = \{u \in E: u < x\}.$$

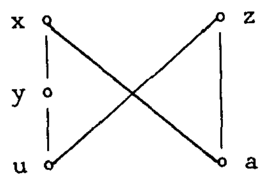
Recall that P is an *interval order*, i.e., can be represented by a set of intervals with the natural ordering of pairs of disjoint intervals, if and only if for all $x, y \in E$,

$$N(x) \subseteq N(y) \quad \text{or} \quad N(y) \subseteq N(x)$$

(see Fishburn [5]). We say that P is a *generalized interval order* if for all $x, y \in E$,

$$N(x) \cap N(y) \neq \emptyset \quad \text{implies} \quad N(x) \subseteq N(y) \text{ or } N(y) \subseteq N(x).$$

Denoting by $\mathcal{P}$ the class of all generalized interval orders, observe that $\mathcal{P}$, in particular, contains all rooted trees (with the understanding that the root is the maximal element) and hence properly contains the interval orders in the stricter sense above. The members of $\mathcal{P}$ are not necessarily rank-optimal as demonstrated by the order $OC_5 \in \mathcal{P}$ with Hasse diagram



OC_5

The purpose of this section is to generalize the results of Faigle and Schrader [3] to yield the following theorem.

Theorem 3.1. *Let P be a generalized interval order such that P contains no induced suborder isomorphic to OC_5. Then P is rank-optimal.*

We need more terminology. The *lineality* of the order P is the number

$$l(P) = \max\{l(L) \colon L \text{ linear extension of } P\}.$$

An *N-maximal* element of P is an element $x \in E$ such that for all $y \in E$, either $N(y) \subseteq N(x)$ or $N(y) \cap N(x) = \emptyset$.

Lemma 3.2. *Let $P \in \mathcal{P}$. Then for each N-maximal element x, there is an optimal linear extension of P ending with x.*

Proof. Consider an arbitrary N-maximal element x of P. If $L = ab \cdots xy \cdots z$ is an optimal linear extension (with respect to lineality) then we may move y forward and insert it immediately after the last element of $N(y)$. Since x and y are not comparable in P, the resulting linear extension has the same lineality. Hence we may assume that for each element y occurring after x in L, also $N(y)$ occurs after x. This means that $L' = y \cdots za \cdots x$ also is an optimal linear extension of P. $\square$

We now return to Theorem 3.1.

Proof of Theorem 3.1. Recall from Theorem 2.2 that the inequality $l(P) \leq r(P)$ always holds. Let x be N-maximal in P. By induction, we may assume $l(P \setminus x) = r(P \setminus x)$.

If $l(P) \geq l(P \setminus x) + 1$ then we have

$$l(P) \geq l(P \setminus x) + 1 = r(P \setminus x) + 1 \geq r(P)$$

since the incidence matrices of P and $P \setminus x$ only differ in the nonzero column corresponding to the maximal element x. Hence we may furthermore assume that $l(P) = l(P \setminus x)$ and it suffices to show that $r(P) = r(P \setminus x)$.

Suppose now that $P \setminus x$ decomposes into at least two nonempty connected components $P_1, P_2, \ldots$. (Recall that a *connected component* of an order is a maximal suborder whose Hasse diagram is connected.) Note that

$$l(P) = l(P_1) + l(P_2) + \cdots$$

since an optimal linear extension of $P \setminus x$ is obtained by concatenating optimal linear extensions of the components of $P \setminus x$. For the same reason, each component P_i of $P \setminus x$ must satisfy $l(P_i) = l(P_i \cup x)$ (otherwise we could concatenate an optimal extension of $P_i \cup x$ after optimal extensions of all other components of $P \setminus x$ and conclude $l(P) \geq l(P \setminus x) + 1$). Thus by induction also $r(P_i) = r(P_i \cup x)$ must hold. So each restriction of the column vector x to P_i is a linear combination of the column vectors in P_i and, consequently, x is in the span of the other vectors, i.e., $r(P) = r(P \setminus x)$.

It remains to deal with the case where $P \setminus x$ is connected. If some element y of $N(x)$ is N-maximal in $P \setminus x$, then Lemma 3.2 implies the existence of an optimal linear extension of $P \setminus x$ ending with y. Concatenation with x then yields $l(P) \geq l(P \setminus x) + 1$, which contradicts the assumption $l(P) = l(P \setminus x)$.

So no element of $N(x)$ is N-maximal in $P \setminus x$. Let z be N-maximal in $P \setminus x$. If $N(z) = N(x)$, then clearly $r(P \setminus x) = r(P)$. We may, therefore, assume that there exists an element y in $N(x) \setminus N(z)$ which is maximal in $P \setminus x$.

Because y is not N-maximal in $P \setminus x$, $N(z) \setminus N(y)$ must be nonempty. Let a be a minimal element of $N(z) \setminus N(y)$. Note that $N(a)$ is contained in $N(y)$. We claim $N(a) = N(y)$. Indeed, if there existed an element u in $N(y) \setminus N(a)$, the connectedness of $P \setminus x$ would imply $u < z$ and so $\{u, y, x, a, z\}$ would give rise to the forbidden suborder OC_5.

Now $N(a) = N(y)$ clearly gives $r(P \setminus x) = r((P \setminus x) \setminus y)$. On the other hand, adjoining yx at the end of an optimal linear extension for $(P \setminus x) \setminus y$ shows that

$$l(P) \geq l((P \setminus x) \setminus y) + 1 = r((P \setminus x) \setminus y) + 1 = r(P \setminus x) + 1 \geq r(P)$$

as required. $\square$

4. One-way optimality of generalized interval orders

Again let P be an arbitrary generalized interval order on the set E. Our main result in this section states that the patterns of 1's of the distinct rows in an incidence matrix of P provide an optimal proof scheme for P.

We first show that P can be assumed to have a bipartite graph as its Hasse diagram. To this end, let E' be a disjoint copy of E and match the corresponding elements $x \leftrightarrow x'$.

On $S = E \cup E'$, we define an order Q via

$$s<t \text{ in } Q \quad \text{iff} \quad s \in E, t \in E' \text{ and}$$
$$t = x' \text{ for some } x \in E \text{ with } s<x \text{ in } P.$$

Apparently, Q is also a generalized interval order with $r(Q)=r(P)$. Indeed, the incidence matrix of Q reduces to that of P by deleting the zero rows corresponding to E' and the zero columns corresponding to E. Hence each proof scheme for P can be viewed as a proof scheme for Q and conversely.

Theorem 4.1. *Let P be a generalized interval order and denote by $\bar{r}(P)$ the number of distinct nonzero rows in the incidence matrix of P. Then*

$$\bar{r}(P) = r(P).$$

Proof. By the foregoing discussion, it suffices to prove the theorem for the class of those generalized interval orders P having no three-element suborder of the form $x<y<z$. We will do this by induction on $|P|$. Note that each order in this class contains no OC_5 and hence is rank-optimal (Theorem 3.1).

Without loss of generality, we assume that P is connected. Hence the N-maximal elements z of P satisfy $N(z) = \min(P)$, i.e., every nonzero row vector has "1" in the position corresponding to the column of z. This implies $\bar{r}(P \setminus z) \leq \bar{r}(P) \leq r(P \setminus z) + 1$.

If $\bar{r}(P \setminus z) = \bar{r}(P)$, then the relation

$$\bar{r}(P) \geq r(P) \geq r(P \setminus z) = \bar{r}(P \setminus z)$$

finishes the proof.

We are left to deal with the case $\bar{r}(P) = \bar{r}(P \setminus z) + 1$. Then there exists an $y \in \min(P) \cap \max(P \setminus z)$. Thus y occurs as last element of some optimal linear extension of $P \setminus z$,

$$r(P) = l(P) > l(P \setminus z) = r(P \setminus z) = \bar{r}(P \setminus z).$$

Consequently, $r(P) = \bar{r}(P)$. □

Corollary 4.2. *For every generalized interval order P,*

$$\log(r(P)) \leq \mathrm{cc}^*(P) \leq \mathrm{cc}(P) \leq \log(r(P)+1).$$

Proof. The inequality $\mathrm{cc}(P) \leq \log(r(P)+1)$ is implied by the following one-way protocol for the decision problem "$?x<y?$":

The rows of the incidence matrix are partitioned according to identical row vectors. The protocol specifies the block of the partition which contains the x-row, which can be done with $\log(\bar{r}(P)+1)$ bits.

The inequality $\log(r(P)) \leq \mathrm{cc}^*(P)$ follows from Corollary 2.4 by replacing, if necessary, P with an equivalent rank-optimal order. □

We finally turn to the discussion of the nondeterministic complexity $\overline{\mathrm{cc}}^*(P)$. We will henceforth assume that P is an interval order *in the strict sense*, i.e., P satisfies for all $x, y \in P$,

$$N(x) \subseteq N(y) \quad \text{or} \quad N(y) \subseteq N(x).$$

Let $f: E \times E \to \{0,1\}$ be the Boolean function describing the incidence matrix of P and set

$$\bar{f}(x,y) = 1 + f(x,y) \pmod 2.$$

With $\bar{Q} = \bar{f}^{-1}(1) = f^{-1}(0)$, we are interested in the nondeterministic complexity $\mathrm{cc}^*(\bar{Q})$.

Take a disjoint copy E' of E and define the order $\bar{P}$ on $S = E \cup E'$ as follows:

$$s < t \text{ in } \bar{P} \quad \text{iff } s \in E, t \in E' \text{ and } \bar{f}(s,x) = 1, \quad \text{where } t = x'.$$

Apparently, $\mathrm{cc}^*(\bar{P}) = \mathrm{cc}^*(\bar{Q})$. Morever,

$$\bar{r}(P) \leq \bar{r}(\bar{P}),$$

where $\bar{r}$ denotes the number of distinct nonzero row vectors in the associated incidence matrices. Looking at the collections of column vectors, we observe:

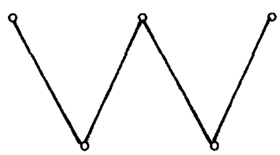
Lemma 4.3. *P is an interval order in the strict sense if and only if $\bar{P}$ is an interval order in the strict sense.*

An application of Corollary 4.2 to $\bar{P}$ therefore yields:

Corollary 4.4. *Let P be an interval order in the strict sense. Then*

$$\log(r(P)) \leq \overline{\mathrm{cc}}^*(P) \leq \mathrm{cc}(P) \leq \log(r(P)+1).$$

Note that our argument does not apply to generalized interval orders. For example, the analogue of Lemma 4.3 is false for the order



In fact, the order P of Example 2.6 is a generalized interval order and thus shows that Corollary 4.4 may not be true for generalized interval orders.

Acknowledgement

The authors thank the referees for their careful reading and detailed remarks, which helped to enhance the presentation.

References

[1] A.V. Aho, J.D. Ullman and M. Yannakakis, On notions of information transfer in VLSI circuits, in: Proceedings 15th ACM Symposium on Theory of Computing (1983) 133–139.

[2] L. Babai, P. Frankl and J. Simon, Complexity classes in communication complexity theory, in: Proceedings 27th IEEE Symposium on Foundations of Computer Science (1986) 337–347.

[3] U. Faigle and R. Schrader, Interval orders without odd crowns are defect optimal, Computing 38 (1987) 59–69.

[4] U. Faigle and G. Turán, Communication complexity, Computing Supplementum 7 (Springer, Vienna, 1990) 141–153.

[5] P.C. Fishburn, Intransitive indifference with inequal indifference intervals, J. Math. Psych. 7 (1970) 144–149.

[6] G. Gierz and W. Poguntke, Minimizing setups for ordered sets: a linear algebraic approach, SIAM J. Algebraic Discrete Methods 4 (1983) 132–144.

[7] B. Halstenberg and R. Reischuk, On different modes of communication, in: Proceedings 20th ACM Symposium on Theory of Computing (1988) 162–172.

[8] R.J. Lipton and R. Sedgewick, Lower bounds for VLSI, in: Proceedings 13th ACM Symposium on Theory of Computing (1981) 300–307.

[9] L. Lovász and M. Saks, Lattices, Möbius functions and communication complexity, in Proceedings 29th IEEE FOCS (1988) 81–90.

[10] K. Mehlhorn and E.M. Schmidt, Las Vegas is better than determinism in VLSI and distributed computing, in: Proceedings 14th ACM Symposium on Theory of Computing (1982) 330–337.

[11] A.C. Yao, Some complexity questions related to distributive computing, in: Proceedings 11th ACM Symposium on Theory of Computing (1979) 209–213.