ON THE COMMUNICATION COMPLEXITY OF GRAPH PROPERTIES

ANDRÁS HAJNAL[1,2], WOLFGANG MAASS[1,3], GYÖRGY TURÁN[1,4]

Abstract.

We prove $\theta(n \log n)$ bounds for the deterministic 2-way communication complexity of the graph properties CONNECTIVITY, s-t-CONNECTIVITY and BIPARTITENESS (for arbitrary partitions of the variables into two sets of equal size). The proofs are based on combinatorial results of Dowling-Wilson and Lovász-Saks about partition matrices using the Möbius function, and the Regularity Lemma of Szemerédi. The bounds imply improved lower bounds for the VLSI complexity of these decision problems and sharp bounds for a generalized decision tree model which is related to the notion of evasiveness.

[1]Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago.

[2]Mathematical Institute of the Hungarian Academy of Sciences, Budapest.

[3]Supported in part by NSF Grant CCR8703889.

[4]Automata Theory Research Group of the Hungarian Academy of Sciences, Szeged.

1. Introduction.

The communication complexity of Boolean functions is a complexity measure corresponding to the amount of information transfer necessary to compute the function. Let f be a Boolean function and (X_1, X_2) be a partition of the variables. Assume that I and II are two processors such that I knows the values of variables in X_1 and II knows the values of variables in X_2. Then COMM(f, X_1, X_2) is the number of bits that have to be exchanged by the processors to compute f. COMM(f), the (deterministic, 2-way) communication complexity of f is the minimal communication complexity taken over all equal size partitions of the variables.

The model for a fixed partition was defined by Yao [14] (a related model was considered by Abelson [1]). The variable partition model is discussed in Yao [15].

This paper is concerned with the problem of proving lower bounds for COMM(f) for particular functions. Yao [15] proved a $\Omega(n^2)$ lower bound for graph isomorphism and raised the problem of proving lower bounds for other graph properties, noting that this "seems to be a difficult problem in general." Ja'Ja' [6] proved a $\theta(n \log n)$ bound for the CONNECTED COMPONENTS function (given G, output the connected components of G). As it is remarked in [6] his methods do not seem to work for decision problems. Other lower bounds are obtained in Lipton-Sedgewick [7], Mehlhorn-Schmidt [9], Papadimitriou-Sipser [10] and Yao [15].

In this paper we consider the communication complexity of the decision problem: given G, decide whether G is connected. For the $\binom{n}{2}$ variable function CONNECTIVITY$_n$ that decides connectivity for graphs on n vertices we prove a $\theta(n \log n)$ bound for COMM(CONNECTIVITY$_n$).

An important aspect of lower bounds on communication complexity (for arbitrary equal size partitions of the input variables) is that they provide lower bounds for VLSI complexity. If A is the area and T is the time required by a VLSI circuit to compute f then $AT^2 = \Omega(\text{COMM}(f)^2)$ (Thompson [12], Lipton-Sedgewick [7], Yao [15]). Our lower bound of $\Omega(n \log n)$ for COMM(CONNECTIVITY$_n$) (Theorem 10) implies an $\Omega(n^2 \log^2 n)$ lower bound on AT^2 for the decision problem CONNECTIVITY$_n$. Apparently the best known upper bound (for the VLSI bit-model) is $O(n^2 \log^9 n)$ (Hambrusch [4], see also [5]).

The proof of the lower bound on COMM(CONNECTIVITY$_n$) is based on two important combinatorial results: a theorem of Dowling and Wilson [3] (see also [13]) which uses Möbius functions to determine the rank of partition matrices, and the fundamental Regularity Lemma of Szemerédi [11] about the structure of dense graphs.

Recently, Lovász and Saks [8] proved a result analogous to the Dowling-Wilson theorem about another kind of partition matrix (also using Möbius functions). Using the proof method of the lower bound for CONNECTIVITY this result implies $\theta(n \log n)$ bounds for s-t-CONNECTIVITY (given G and vertices s and t, decide whether there is a path from s to t in G) and for BIPARTITENESS (given G, decide whether it is bipartite).

Thus, in conclusion, our approach seems to provide a general method to prove lower bounds for the communication complexity of graph properties.

In section 5 we will consider applications of these lower bounds to decision tree complexity. There are several well-known results about the decision tree complexity of graph properties (see Bollobás [2]), centered around the concept of evasiveness and the Aanderaa-Rosenberg conjecture. CONNECTIVITY is a standard example of a graph property that is evasive, i.e., in any decision tree where one can only test whether a certain edge is present one has to test in the worst case each of the $\binom{n}{2}$ possible edges. In section 5 we will consider a more powerful decision tree model which allows tests of the form "Is any edge in X present?" where X is an arbitrary set of edge slots. It is easy to see that only $O(n \log n)$ tests are needed to decide CONNECTIVITY$_n$ in this generalized model. The lower bound on the communication complexity of CONNECTIVITY$_n$ implies that this bound is in fact optimal.

2. Preliminaries.

In this section we summarize the definitions and the combinatorial results that are used. Furthermore, we outline the strategy for the proof of the main theorem (Theorem 10).

Let $f : \{0,1\}^m \to \{0,1\}$ be a Boolean function and let (X_1, X_2) be a partition of the variable set X into two parts of equal size (i.e., $\big||X_1| - |X_2|\big| \leq 1$). By $M^f_{X_1,X_2}$ we denote the $2^{|X_1|} \times 2^{|X_2|}$ matrix describing f.

A deterministic two-way communication protocol π to compute f given the partition (X_1, X_2) is of the form

$$\pi = \{g^1_k, g^2_k, h^1_k, h^2_k : \begin{aligned}[t] & g^1_k : \{0,1\}^{|X_1|} \times \{0,1\}^{k-1} \to \{0,1\} \\ & g^2_k : \{0,1\}^{|X_2|} \times \{0,1\}^{k} \to \{0,1\} \\ & h^1_k : \{0,1\}^{|X_1|} \times \{0,1\}^{k-1} \hookrightarrow \{0,1\} \\ & h^2_k : \{0,1\}^{|X_2|} \times \{0,1\}^{k} \hookrightarrow \{0,1\}, \\ & k \geq 1\} \end{aligned}$$

where g^1_k, g^2_k give the k-th message of the processors computed from their input and the messages received; and h^1_k, h^2_k are partial functions which whenever defined give the value of f. We do not formalize further the requirements guaranteeing that π is a correct protocol (see e.g., [14]).

$C(\pi, \underline{x})$ is the number of bits exchanged by π on input $\underline{x}$. $C(\pi) = \max\{C(\pi, \underline{x}) : \underline{x} \in \{0,1\}^m\}$ is the complexity of π. COMM $(f, X_1, X_2) = \min\{C(\pi)$: π computes f given the partition $(X_1, X_2)\}$ is the complexity of f given the partition (X_1, X_2) and COMM $(f) = \min\{\text{COMM}(f, X_1, X_2) : (X_1, X_2)$ is a partition of X into two parts of equal size$\}$ is the deterministic 2-way communication complexity of f.

Theorem 1. *(Mehlhorn-Schmidt [9]).*

$$COMM(f, X_1, X_2) \geq \log_2(rank(M^f_{X_1,X_2})) - 1.$$

□

Let $P_1, \ldots, P_t$ be an enumeration of all partitions of $S = \{1, \ldots, s\}$ (into arbitrarily many sets), where $t = B_s$ is the s-th Bell number.

$P_i \vee P_j$ is the finest partition P such that both P_i and P_j are refinements of P. By M^s we denote the $t \times t$ $0-1$ matrix given by $M^s_{ij} = 1$ if and only if $P_i \vee P_j$ is the trivial partition $\underline{1}$ consisting of the single set $\{1, \ldots, s\}$. By N^s we denote the $t \times t$ $0-1$ matrix given by $N^s_{ij} = 1$ if and only if elements 1 and 2 are in the same subset in the partition $P_i \vee P_j$.

Theorem 2. *(Dowling-Wilson [3], [13]).*

$$rank(M^s) = B_s.$$

□

Theorem 3. *(Lovász-Saks [8]).*

$$rank(N^s) = B_s - B_{s-1}.$$

□

Corollary 4. $\log(rank(M^s)) = \Omega(s \log s)$ *and* $\log(rank(N^s)) = \Omega(s \log s)$. □

Undirected graphs $G = (V, E)$ with $V = \{1, \ldots, n\}$ can be described by $\underline{x}_G \in \{0,1\}^{\binom{n}{2}}$ corresponding to the adjacency matrix of G. CONNECTIVITY$_n$ is the $\binom{n}{2}$-variable Boolean function for which CONNECTIVITY$_n(\underline{x}_G) = 1 \Leftrightarrow G$ is connected. Variables correspond to edges of the complete graph K^n on $\{1, \ldots, n\}$ and a partition of the variables will be given in the form (R, B) ("red" and "blue" edges), where $\big||R| - |B|\big| \leq 1$.

In order to motivate the proof of the lower bound on COMM (CONNECTIVITY$_n$) we first consider the special case where the $\binom{n}{2}$ variables have been partitioned in the following way.

Assume that the vertex set is partitioned into 3 disjoint sets A, B, C of size $n/3$. Consider the special partition of the edges where processor I has all edges from $A \times B$ and processor II has all edges from $B \times C$ (other edges are distributed arbitrarily). Let $P = (S_1, \ldots, S_k)$ be a partition of B.

Define the bipartite graph $G^1_P = (A, B, E_P)$ as follows: for every subset S_i choose a vertex $v_i \in A$ and connect it to all vertices in S_i (if $i \neq j$ then $v_i \neq v_j$); connect all vertices in $A - \{v_1, \ldots, v_k\}$ to some vertex in B.

Define G^2_P similarly with A replaced by C.

Then for two partitions P and P' of B, $G^1_P \cup G^2_{P'}$ is connected if and only if $P \vee P' = \underline{1}$, thus the matrix of the connectivity function corresponding to this partition of the variables contains $M^{n/3}$ as a submatrix and therefore in this case the communication complexity is $\Omega(n \log n)$ (by Corollary 4 together with Theorem 1).

What remains is to prove that *every* partition of the variables contains a configuration similar to the above and thus M^s can always be embedded into the matrix corresponding to the partition (for some $s = \Omega(n)$). We will show in the next section that this can be derived from the Regularity Lemma of Szemerédi, which is stated below. In fact it appears that the Regularity Lemma provides a general method to deduce lower bounds for arbitrary partitions from a lower bound for a special partition.

If $G = (V, E)$ is a graph and V_1, V_2 are disjoint subsets of V then $e(V_1, V_2)$ denotes the number of edges of G between V_1 and V_2. $G|(V_1, V_2)$ is the bipartite graph induced by G on (V_1, V_2). If $H = (A, B, E)$ is a bipartite graph, $A' \subseteq A$ and $B' \subseteq B$ then $\delta(A', B') = e(A', B')/(|A'| \cdot |B'|)$ is the density of (A', B'). H is ϵ-regular if for every $A' \subseteq A$, $B' \subseteq B$ it holds that $|A'| \geq \epsilon|A|$, $|B'| \geq \epsilon|B|$ imply $|\delta(A', B') - \delta(A, B)| < \epsilon$.

Theorem 5. *(Szemerédi Regularity Lemma [11]).*

For every $\epsilon > 0$ and $m \in \mathbf{N}$ there are numbers $M = M(\epsilon, m)$ and $N = N(\epsilon, m)$ such that for all graphs $G = (V, E)$ on $n \geq N$ vertices the following holds: there is a partition $V = C_0 \cup C_1 \cup \ldots \cup C_k$ such that

(1) $m \leq k \leq M$,

(2) $|C_0| < \epsilon N$,

(3) $|C_1| = \ldots = |C_k|$,

(4) with the exception of $< \epsilon k^2$ pairs, for all pairs (C_i, C_j) the bipartite graph $G|(C_i, C_j)$ is ϵ-regular.

3. The Application of the Regularity Lemma.

Let $H = (A, B, E)$ be an ϵ-regular graph with $|A| = |B| = r$ and assume that every degree is ar least δr, for some $\delta > \epsilon$. Two paths connecting vertices v and w are disjoint if they have no common inner vertices.

Lemma 6. *If $v \in A$ and $w \in B$ then there are at least $\frac{1}{2}(\delta - \epsilon)\delta r$ pairwise disjoint paths of length at most 3 connecting v and w.*

Proof: Let N_v (resp. N_w) be a set of δr neighbors of v (resp. w). As $|E| \geq \delta r^2$ and H is ϵ-regular, $H|(N_v, N_w)$ has density $> \delta - \epsilon$ and therefore $> (\delta - \epsilon)|N_v|\,|N_w| = (\delta - \epsilon)\delta^2 r^2$ edges.

Now $H|(N_v, N_w)$ contains $\geq \frac{1}{2}(\delta - \epsilon)\delta r$ independent edges. To see this, select edges one by one. After i edges are selected, these exclude $\leq 2i\delta r$ edges sharing an endpoint with these edges. Hence we can continue as long as $2i\delta r \leq (\delta - \epsilon)\delta^2 r^2$.

These edges can be completed to disjoint paths of length ≤ 3 connecting v and w. □

Lemma 7. *If v and w are different vertices of A then there are at least $\frac{1}{6}(\delta - \epsilon)\delta r$ pairwise disjoint paths of length ≤ 4 connecting v and w.*

Proof: Let N_v be any set of δr neighbors of v.

Then there are $\geq \frac{1}{6}(\delta - \epsilon)\delta r$ paths of length ≤ 3 connecting N_v and w such that any two of these paths have only w as a common vertex and none of them contains v. Indeed, after selecting i such paths, these exclude $\leq 3i + 1$ paths between w and any vertex of N_v not contained in these paths. Lemma 6 implies that we can continue as long as $3i + 1 \leq \frac{1}{2}(\delta - \epsilon)\delta r$.

These paths can be completed to pairwise disjoint paths of length ≤ 4 connecting v and w. □

If $F \subseteq E$ then it induces a partition P_F on A obtained by intersecting the connected components of $H' = (A, B, F)$ with A.

Lemma 8. *If $S \subseteq A$ and $|S| \leq \frac{1}{24}(\delta - \epsilon)\delta r$ then every partition of S is induced by some $F \subseteq E$.*

Proof: Let $P = (S_1, \ldots, S_k)$ be a partition of S, where $S_i = \{v_{i,1}, \ldots, v_{i,r_i}\}$, $r_i = |S_i|$, $i = 1, \ldots, k$.

We construct $F = \bigcup_{i=1}^{k} E_i \subseteq E$, where the E_i's are the connected components of F and $S_i = V(E_i) \cap S$.

In order to construct E_i we select for every $j \in \{1, \ldots, r_i - 1\}$ a path $P_{i,j}$ of length ≤ 4 between $v_{i,j}$ and $v_{i,j+1}$ so that the inner vertices of $P_{i,j}$ are disjoint from $V(E_1) \cup \ldots \cup V(E_{i-1}) \cup S$. Such path $P_{i,j}$ exists for every j by Lemma 7, because

$$|V(E_1) \cup \ldots \cup V(E_{i-1}) \cup S| \leq 4|S| - 1 < \frac{1}{6}(\delta - \epsilon)\delta r.$$

We define E_i to be the union of these paths $P_{i,j}$. □

Lemma 8 requires H to have large minimal degree besides ϵ-regularity. The next lemma, which is a direct consequence of the Szemerédi Regularity Lemma, shows that for graphs with $\frac{1}{2}\binom{n}{2}$ edges such bipartite graphs can be found, simultaneously in the graph and its complement. The constructed sets V_1, $V_0^1 \cap V_0^2$, V_2 will play a similar role in the proof of the main theorem as the sets A, B, C in the sketch of the proof (for a special case) in section 2.

Lemma 9. *If ϵ is sufficiently small then there is an α $(0 < \alpha < 1)$ and an $n_0 \in \mathbf{N}$ such that for every graph $G = (V, E)$ with $n \geq n_0$ vertices and $\frac{1}{2}\binom{n}{2}$ edges, and for every $\gamma (2\epsilon < \gamma \leq \frac{1}{2})$ the following holds: there are subsets $V_0^1, V_0^2, V_1, V_2 \subseteq V$ such that*

a) each set has size $r \geq \alpha n$,

b) V_0^1, V_0^2, V_1 and V_2 are pairwise disjoint,

c) $(\gamma - 2\epsilon)r \leq |V_0^1 \cap V_0^2| \leq 2\gamma r$,

d) $G|(V_0^1, V_1)$ is ϵ-regular with all degrees $\geq \frac{1}{3}r$,

e) $G|(V_0^2, V_2)$ is ϵ-regular with all degrees $\leq \frac{2}{3}r$.

Proof: Apply the Szemerédi Regularity Lemma for some ϵ' and m to be specified later. If G has $n \geq N(\epsilon', m)$ vertices, we get a partition $(C_0, C_1, \ldots, C_k)$ of V, where $m \leq k \leq M(\epsilon', m)$. Put $t = |C_1| = \ldots = |C_k|$.

Case 1. There is a non-exceptional pair (C_i, C_j) with density δ such that $\frac{1}{3} + 6\epsilon' \leq \delta \leq \frac{2}{3} - 6\epsilon'$.

Let $C_i^1, C_i^2 \subseteq C_i$, $C_j^1, C_j^2 \subseteq C_j$ be sets of size $q := \lfloor t/2 \rfloor$ such that $C_i^1 \cap C_i^2 = \emptyset$ and $|C_j^1 \cap C_j^2| = \lceil \frac{\gamma}{2} t \rceil$.

Consider the bipartite graphs $G_1 := G|(C_i^1, C_j^1)$ and $G_2 := G|(C_i^2, C_j^2)$.

The sets required by the lemma are obtained by excluding vertices of small and large degree. In particular we will remove the sets D_i^1 and D_j^1 of vertices that have too small degree in G_1, where

$$D_i^1 := \{v \in C_i^1 : d_{G_1}(v) \le (\frac{1}{3} + 5\epsilon')q\} \text{ and}$$
$$D_j^1 := \{v \in C_j^1 : d_{G_1}(v) \le (\frac{1}{3} + 5\epsilon')q\}.$$

We will also remove the sets D_i^2 and D_j^2 of vertices that have too large degree in G_2, where

$$D_i^2 := \{v \in C_i^2 : d_{G_2}(v) \ge (\frac{2}{3} - 5\epsilon')q\} \text{ and}$$
$$D_j^2 := \{v \in C_j^2 : d_{G_2}(v) \ge (\frac{2}{3} - 5\epsilon')q\}.$$

Each of these sets has size $< \epsilon' \cdot t$. Assume that e.g., $|D_i^1| \ge \epsilon' \cdot t$. Then applying the ϵ'-regularity of (C_i, C_j) to (D_i^1, C_j^1) we get

$$\frac{1}{3} + 5\epsilon' \le \delta - \epsilon < \frac{e(D_i^1, C_j^1)}{|D_i^1||C_j^1|} \le \frac{(\frac{1}{3}+5\epsilon')|C_i^1||C_j^1|}{|C_i^1||C_j^1|} = \frac{1}{3} + 5\epsilon',$$

a contradiction. The other inequalities follow in the same way.

Now choose $V_0^1 \subseteq C_j^1 \backslash (D_j^1 \cup D_j^2)$, $V_0^2 \subseteq C_j^2 \backslash (D_j^1 \cup D_j^2)$, $V_1 \subseteq C_i^1 \backslash D_i^1$, $V_2 \subseteq C_i^2 \backslash D_i^2$ to be sets of equal size r such that $q - 2\epsilon' t \le r \le q$ and

$$\lceil \frac{\gamma}{2} t \rceil - 2\epsilon' t \le |V_0^1 \cap V_0^2| \le \lceil \frac{\gamma}{2} t \rceil.$$

Such a choice is possible by the bounds on the sizes of the sets deleted.

We claim that these sets satisfy the conditions of the lemma. Straightforward calculations give that if t is sufficiently large then $r \ge (\frac{1}{2} - 3\epsilon')t$, $|V_0^1 \cap V_0^2| \ge (\gamma - 4\epsilon')r$, all degrees in $G|(V_0^1, V_1)$ are $\ge \frac{1}{3} r$ and all degrees in $G|(V_0^2, V_2)$ are $\le \frac{2}{3} r$. Furthermore, if $\epsilon' = \frac{\epsilon}{2(1+3\epsilon)}$ then $G|(V_0^1, V_1)$ and $G|(V_0^2, V_2)$ are ϵ-regular. Observing $\epsilon' < \frac{\epsilon}{2}$ for c) and putting $\alpha := (\frac{1}{2} - 3\epsilon')(1 - \epsilon')\frac{1}{M(\epsilon', m)}$ completes the proof for Case 1.

Case 2. Every non-exceptional pair is either sparse (has density $< \frac{1}{3} + 6\epsilon'$) or dense (has density $> \frac{2}{3} - 6\epsilon'$).

We show that there are classes C_i, C_j and C_k such that (C_i, C_j) is a dense non-exceptional pair and (C_j, C_k) is a sparse non-exceptional pair. If this is true then for these classes the construction of Case 1 can be repeated and the same bounds hold.

Let s (resp. d) denote the number of sparse (resp. dense) non-exceptional pairs. Then

$$d\left(\frac{2}{3} - 6\epsilon'\right)\left(\frac{(1-\epsilon')n}{k}\right)^2 \le \frac{1}{2}\binom{n}{2}$$

hence

$$s \ge \binom{k}{2} - \left(\frac{1}{4(\frac{2}{3} - 6\epsilon')(1-\epsilon')^2} + \epsilon'\right)k^2.$$

This means that if ϵ' is sufficiently small and m is sufficiently large then there are $\ge \frac{1}{9}k^2$ sparse non-exceptional pairs and $\ge \frac{1}{9}k^2$ dense non-exceptional pairs (since the coefficient of k^2 in the second inequality tends to $\frac{3}{8}$ for $\epsilon' \to 0$). But for $\mathcal{C} = \{C_i : C_i \text{ is in } \ge 6\epsilon' k \text{ exceptional pairs}\}$ it holds that $|\mathcal{C}| \le k/3$.

Assume there are no classes C_i, C_j, C_k as claimed. Consider a class $C_i \notin \mathcal{C}$ contained in a sparse non-exceptional pair (as $\binom{|\mathcal{C}|}{2} < \frac{1}{9}k^2$ there is such a class). There are $> (1 - 6\epsilon')k - 1$ classes C_j such that (C_i, C_j) is a sparse non-exceptional pair. Considering a class $C_k \notin \mathcal{C}$ contained in a dense non-exceptional pair, we obtain a contradiction (since we can find a class C_j that works for both C_i and C_k). □

4. The bounds for communication complexity.

We consider the $\binom{n}{2}$-variable Boolean functions CONNECTIVITY$_n$, s-t-CONNECTIVITY$_n$ and BIPARTITENESS$_n$. Partitions of the variables are denoted by (R, B).

Theorem 10. *COMM(CONNECTIVITY$_n$)* $= \theta(n \log n)$.

Proof: The upper bound is easily seen even using 1-way communication only (I transmits to II a list of the vertices in each connected component of (V, R)). To prove the lower bound, let (R, B) be a partition of the edges into two sets of equal size.

Let $V^* \subseteq V$ be a set and let $P_1, \ldots, P_t$ be an enumeration of all partitions of V^*. Assume that there are graphs $G_i^R = (V, E_i^R)$, $G_i^B = (V, E_i^B)$ with $E_i^R \subseteq R$, $E_i^B \subseteq B$ for $i = 1, \ldots, t$ such that for every i and j, $G_i^R \cup G_j^B$ is connected if and only if $P_i \vee P_j = \underline{1}$. Then the partition matrix M^s is a submatrix of $M_{R,B}^{CONN_n}$ for $s = |V^*|$. If $|V^*| = \Omega(n)$ then Theorems 1, 2, and Corollary 4 can be applied to prove the theorem. Hence what remains is to find V^* and graphs G_i^R, G_i^B with the above properties.

Let ϵ, α, n_0 be numbers suitable for Lemma 9, and assume $n \ge n_0$. Apply Lemma 9 to $G = (V, R)$ with $\gamma = \frac{1}{148}(\frac{1}{3} - \epsilon)$ to obtain sets V_0^1, V_0^2, V_1, V_2 (we assume ϵ is small enough so that $\gamma > 2\epsilon$ holds).

Put $V^* = V_0^1 \cap V_0^2$ and let $P_1, \ldots, P_t$ be the partitions of V^*. Note that by Lemma 9c $|V^*| = \Omega(n)$.

As $|V^*| \le 2\gamma r = \frac{1}{24}(\frac{1}{3} - \epsilon)\frac{1}{3} r$ we can apply Lemma 8 to $V^* \subseteq V_0^1$ and $H := G|(V_0^1, V_1)$ with $\delta := \frac{1}{3}$ to get for each partition P_i of V^* a set $F_i^R \subseteq R$ of edges between V_0^1 and V_1 that induces P_i. Similarly, Lemma 8 can be applied to $G|(V_0^2, V_2)$ and an arbitrary partition P_j of V^* to obtain $F_j^B \subseteq B$ inducing P_j.

Now consider all vertices outside of V^* that occur as endpoint of an edge in some F_i^R or F_j^B:

$$\begin{aligned} V_1^* =& \{v \in V_1 \cup (V_0^1 \backslash V^*) : v \text{ is an endpoint of an edge of} \\ & F_i^R \text{ for some } i, 1 \le i \le t\} \\ V_2^* =& \{v \in V_2 \cup (V_0^2 \backslash V^*) : v \text{ is an endpoint of an edge of} \\ & F_j^B \text{ for some } j, 1 \le j \le t\}. \end{aligned}$$

Let v^* be a fixed vertex in V^*.

When we construct G_i^R and G_j^B, all vertices outside $V^* \cup V_1^* \cup V_2^*$ (these are "superfluous" vertices) will be connected to v^* by an edge which is always present. We set $\widetilde{E} = \{(v, v^*) : v \in V \backslash (V^* \cup V_1^* \cup V_2^*)\}$, $\widetilde{E}^R = \widetilde{E} \cap R$, $\widetilde{E}^B = \widetilde{E} \cap B$.

We also need "helping" edges from vertices in $V_1^* \cup V_2^*$ which guarantee that these vertices are never isolated. For $v \in V_1^*$, let

$e_v \in R$ be any edge joining v to V_1^* is $v \in V_0^1 \backslash V^*$ and any edge joining v to V^* if $v \in V_1$ (there exists such an edge by the definition of V_1^*). Similarly, for $v \in V_2^*$, let $e_v \in B$ be any edge joining v to V_2^* if $v \in V_0^2 \backslash V^*$ and any edge joining v to V^* if $v \in V_2$.

Finally: define $G_i^R = (V, E_i^R)$, $G_j^B = (V, E_j^B)$ by

$$E_i^R = F_i^R \cup \widetilde{E}^R \cup \{e_v : v \in V_1^* \text{ is isolated in } F_i^R\} \text{ and}$$
$$E_j^B = F_j^B \cup \widetilde{E}^B \cup \{e_v : v \in V_2^* \text{ is isolated in } F_j^B\}.$$

Then $G_i^R \cup G_j^B$ is connected if and only if $P_i \vee P_j = \underline{1}$. Indeed, if $P_i \vee P_j = \underline{1}$ then $F_i^R \cup F_j^B$ is a connected edge set and all other vertices are connected to it using $\widetilde{E}$ and the edges e_v. If $P_i \vee P_j \neq \underline{1}$ then $F_i^R \cup F_j^B$ is a disconnected edge set and the other edges will not connect it as they form paths of length 1 or 2 "hanging" from $F_i^R \cup F_i^B$. This completes the proof. □

Theorem 11. *a) COMM(s-t-CONNECTIVITY$_n$) = $\theta(n \log n)$,*
b) COMM(BIPARTITENESS$_n$) = $\theta(n \log n)$.

Proof: a) Similar to the proof above, using Theorem 3 instead of Theorem 2.

b) This is another consequence of Theorem 3. In order to show for some $s = \Omega(n)$ that N^s is a submatrix of NONBIPARTITENESS$_n$ one defines edges sets E_i^R, E_j^B similarly as in the proof of Theorem 10, but with the additional property that connected vertices of V^* are only connected by paths of *even* length in $(V, E_i^R \cup E_j^B)$. This can easily be accomplished with the help of the fact that one can get in Lemma 7 at least $\frac{1}{6}(\delta - \epsilon)\delta r - 1$ paths of length *exactly* 4 connecting v and w. Finally one adjusts E_i^R, E_j^B by adding the edge $\{s,t\}$ for the special vertices s,t in V^* (corresponding to the special points 1,2 in the definition of matrix N^s). In this way one gets edge sets $\widetilde{E}_i^R, \widetilde{E}_j^B$ so that 1,2 are in the same connected component of $P_i \vee P_j$ if and only if $(V, \widetilde{E}_i^R \cup \widetilde{E}_j^B)$ has a cycle of odd length (note that by construction such cycle of odd length would have to go through edge $\{s,t\}$).□

5. Application to generalized decision trees.

Let f be a Boolean function. Assume we want to compute f by asking questions of the form "does any variable in Y have value 1?", where Y is any subset of the variable set X. Thus the model we are considering is a binary decision tree with tests of the above form. This is a generalization of the standard decision tree model here a test is of the form "is the value of x_i equal to 1?", where x_i is any variable, i.e., $|Y| = 1$. Let $GD(f)$ be the complexity of f in the generalized model. E.g., a simple adversary argument shows that GD(PARITY$_n$) = n.

Theorem 12. *a) GD(CONNECTIVITY$_n$) = $\theta(n \log n)$,*
b) GD(s-t-CONNECTIVITY$_n$) = $\theta(n \log n)$,
c) GD(BIPARTITENESS$_n$) = $\theta(n \log n)$.

Proof: The upper bounds are easy. E.g. for CONNECTIVITY$_n$ one uses the fact that for any set of vertices $U \subseteq V$ one can find via binary search with $O(\log n)$ tests a vertex $v \in V - U$ that is adjacent to some vertex in U, provided there exists such v: One divides $V - U$ into two sets W_1, W_2 of equal size and carries out tests for $Y_1 := U \times W_1$ and $Y_2 := U \times W_2$. If i is minimal so that one gets a positive answer for Y_i one repeats this process with V replaced by W_i.

Once one has found a vertex v that is adjacent to U one repeats the whole procedure with U replaced by $U \cup \{v\}$ (in order to locate another vertex in the connected component of $U \cup \{v\}$). Initially one sets $U = \{s\}$ for some arbitrary vertex s.

Note that with the same upper bound (up to constant factors) one can in fact *construct* a spanning tree for each connected component of the considered graph (one just has to find out to *which* vertex of U the new vertex v is adjacent in the procedure above; this can be done in the same way).

The lower bounds can be derived from the lower bounds on communication complexity (Theorems 10., 11.) with the help of the following observation.

Lemma 13. *For every Boolean function f,*

$$COMM(f) \leq 2\, GD(f).$$

Proof: Every generalized decision tree gives rise to a communication protocol for any partition (X_1, X_2) of the input variables: for each test set Y processor I tells processor II whether some variable in $Y \cap X_1$ has value 1 and processor II tells I whether some variable in $Y \cap X_2$ has value 1. □

Remark 14. In the same fashion one can derive somewhat weaker lower bounds for more powerful versions of the generalized decision tree model. In particular one gets a lower bound of $\Omega(n)$ for the case where one allows tests of the form "how many edges in Y are present?", and for the case where one allows linear tests on the $\binom{n}{2}$ input variables with coefficients of polynomial size.

The following result separates generalized decision trees from oblivious generalized decision trees. We write $OGD(f)$ for the complexity of f on oblivious generalized decision trees (with tests of the form "does any variable in Y have value 1?") where all nodes on the same level of the tree ask this question for the same set Y (in other words: the sequence of test sets Y is specified in advance).

Theorem 15. *OGD(CONNECTIVITY$_n$) = $\Omega(n^2/\log^2 n)$.*

Proof: Consider an oblivious generalized decision tree T for CONNECTIVITY$_n$ with test sets $Y_1, \ldots, Y_t$. We can assume that $t \leq n^2$ (otherwise we are done).

It is easy to see that for every partition V_1, V_2 of V (where $|V| = n$) one has for

$$E(V_1, V_2) := \{\{x,y\} : x \in V_1 \text{ and } y \in V_2\}$$

that

$$E(V_1, V_2) = \bigcup \{Y_i : Y_i \subseteq E(V_1, V_2)\}$$

(consider graphs where all edges inside V_1 and inside V_2 are present and potentially one edge of $E(V_1, V_2)$; note that we identify edge variables with the corresponding edges).

Thus it is sufficient to show that there is some partition V_1, V_2 of V so that $\frac{n}{3} \leq |V_1| \leq \frac{2}{3}n$ and $Y_i \not\subseteq E(V_1, V_2)$ for all test sets Y_i in T with at least $6 \log n$ vertices incident with edges in Y_i.

Let Y be any set of edges with $s \geq 6 \log n$ incident vertices. Then there are at most $2^{n-s/2-1}$ partitions of V into sets V_1, V_2 with $Y \subseteq E(V_1, V_2)$ (there are 2^{n-s} choices for distributing those vertices that are not incident with Y, furthermore there are up to $2^{s/2}$ ways of distributing the remaining vertices to V_1, V_2 so that $Y \subseteq E(V_1, V_2)$). This implies that there are at most $n^2 \cdot 2^{n-3\log n-1} = 2^{n-1}/n$ partitions V_1, V_2 so that $Y_i \subseteq E(V_1, V_2)$ for some test set Y_i of T with $\geq 6 \log n$ incident vertices. On the other hand there are $\geq \left(1 - \frac{24}{n}\right) \cdot 2^n$ partitions V_1, V_2 of V with $\frac{n}{3} \leq |V_1| \leq \frac{2}{3}n$. □

6. Open Problems.

There are some interesting decision problems on graphs for which the communication complexity in the here considered model (with arbitrary input partitions) is still unknown, e.g. planarity, bipartite matching and Hamiltonian circuits. Furthermore almost nothing is known about the complexity of graph properties in the probabilistic version of this model (in particular the complexity of CONNECTIVITY_n is open for this model).

Finally the complexity of other graph properties in the generalized decision tree model of section 5 remains to be determined. Further open questions arise if one wants to determine the complexity of graph properties in the more powerful decision tree model that allows linear tests (in particular the complexity of CONNECTIVITY_n is not known, not even if the weights in the linear tests are required to have polynomial size).

Acknowledgements.

We would like to thank László Babai, Susanne Hambrusch, László Lovász, Michael Saks, Endre Szemerédi and Richard Wilson for valuable discussions.

References.

[1] H. Abelson, Lower bounds on information transfer in distributed computing, *19th FOCS*, (1978), 151-158.

[2] B. Bollobás, Extremal Graph Theory, Academic Press (1978).

[3] T.A. Dowling, R.M. Wilson, Whitney number inequalities for geometric lattices, *Proc. Amer. Math. Soc.*, **47**(1975), 504-512.

[4] S.E. Hambrusch, The complexity of graph problems in VLSI, Ph.D. Dissertation, Pennsylvania State Univ. (1982).

[5] S.E. Hambrusch, J. Simon, Solving undirected graph problems on VLSI, *SIAM J. Comp.*, **14**(1985), 527-544.

[6] J. Ja'Ja', The VLSI complexity of selected graph problems, *JACM*, **31**(1984), 377-391.

[7] R. Lipton, R. Sedgewick, Lower bounds for VLSI, *13th STOC*, (1981), 300-307.

[8] L. Lovász, M. Saks, to be published.

[9] K. Mehlhorn, E.M. Schmidt, Las Vegas is better than determinism in VLSI and distributed computing, *14th STOC*, (1982), 330-337.

[10] C.H. Papadimitriou, M. Sipser, Communication Complexity, *JCSS*, **28**(1984), 260-269.

[11] E. Szemerédi, Regular partitions of graphs, *Problémes combinatoires et théorie des graphs, Coll. CNRS*, (1976), 399-401.

[12] C. Thompson, Area-time complexity for VLSI, *11th STOC*, (1979), 81-88.

[13] D.J.A. Welsh, Matroid Theory, Academic Press, (1976).

[14] A. Yao, Some complexity questions related to distributed computing, *11th STOC*, (1979), 209-213.

[15] A. Yao, The entropic limitations of VLSI computations, *13th STOC*, (1981), 308-311.