SIAM J. COMPUT.
Vol. 17, No. 4, August 1988

004

LOCALITY, COMMUNICATION, AND INTERCONNECT LENGTH IN MULTICOMPUTERS*

PAUL M. B. VITÁNYI†

Abstract. We derive a lower bound on the average interconnect (edge) length in d-dimensional embeddings of arbitrary graphs, expressed in terms of diameter and symmetry. It is optimal for all graph topologies we have examined, including complete graph, star, binary n-cube, cube-connected cycles, complete binary tree, and mesh with wraparound (e.g., torus, ring). The lower bound is technology independent, and shows that many interconnection topologies of today's multicomputers do not scale well in the physical world ($d = 3$). The new proof technique is simple, geometrical, and works for wires with zero volume, e.g., for optical (fibre) or photonic (fibreless, laser) communication networks. Apparently, while getting rid of the "von Neumann" bottleneck in the shift from sequential to nonsequential computation, a new communication bottleneck arises because of the interplay between locality of computation, communication, and the number of dimensions of physical space. As a consequence, realistic models for nonsequential computation should charge extra for communication, in terms of time and space.

Key words. multicomputers, complexity of computation, locality, communication, wire length, general communication network, edge-symmetric graph, binary n-cube, cube-connected cycles, tree, Euclidean embedding, scalability, optical computing

AMS(MOS) subject classifications. 68C05, 68C25, 68A05, 68B20, 94C99

1. The tyranny of physical space. In many areas of the theory of parallel computation we meet graph-structured computational models. These models encourage the design of parallel algorithms where the cost of communication is largely ignored. Yet it is well known that the cost of computation—in both time and space—vanishes with respect to the cost of communication in parallel or distributed computing. As multiprocessor systems with really large numbers of processors start to be constructed, this effect becomes more and more apparent. Thinking Machines Corporation of Cambridge, Massachusetts, has just marketed the "Connection Machine," a massive multiprocessor parallel computer. The prototype contains microscopically fine-grained processor/memory cells, 65,536 of them, each with 4,096 bits of memory and a simple arithmetical unit. The communication network connecting the processors is packet-switched and based on the binary 16-cube. (A binary n-cube network consists of 2^n nodes, each node identified by an n-bit name, and an edge between nodes which differ in a single bit.) This is implemented by packing a cluster of 16 processors and one router circuit on a single chip. The 4,096 routers (in casu chips) are connected by 24,576 bidirectional wires in the pattern of the binary 12-cube. The last chapter of [3], "New Computer Architectures and their Relationship to Physics or, Why Computer Science is No Good," expresses the dissatisfaction of the designers with traditional computer science, "which abstracts the wire away into a costless and volumeless idealized connection. [The] old models do not impose a locality of connection, even though the real world does.... In classical computation the wire is not even considered. In current engineering it may be the most important thing." Here we shall argue that,

* Received by the editors February 13, 1987; accepted for publication (in revised form) September 16, 1987. This work was supported in part by the Office of Naval Research under contract N00014-85-K-0168, by the Office of Army Research under contract DAAG29-84-K-0058, by the National Science Foundation under grant DCR-83-02391, and by the Defense Advanced Research Projects Agency under contract N00014-83-K-0125. Preliminary results were reported in VLSI Algorithms and Architectures, Lecture Notes in Computer Science 227, Springer-Verlag, Berlin, New York, 1986.

† Massachusetts Institute of Technology, Laboratory for Computer Science, Cambridge, Massachusetts 02139 and Centrum voor Wiskunde en Informatica, Kruislaan 413, 1098 SJ Amsterdam, the Netherlands.

while getting rid of the so-called "von Neumann" bottleneck,[1] in the shift from serial to nonserial computing, we run into a new *communication* bottleneck due to the three-dimensionality of physical space.

Models of parallel computation that allow processors to randomly access a large shared memory, such as PRAMs, or rapidly access a large number of processors, such as NC computations, can provide new insights in the inherent parallelizability of algorithms to solve certain problems.[2] For instance, in the form of distributing copies of the entire problem instance, or pieces of the problem instance, among an exponential number of processors in a linear number of steps (i.e., the number of steps in the longest causal chain is linear). Or, as in NC, among a polynomial number of processors in a polylogarithmic number of steps. This sometimes leads to the obscure thought that VLSI technology opens the way to implement tree machines which solve NP-complete problems in linear time. Now, the way a problem instance can be divided and partial answers put together may give genuine insight into its parallelizability. However, it *cannot* give a reduction from an asymptotic exponential time best algorithm in the sequential case to an asymptotic polynomial time algorithm in *any* parallel case. At least, if by "time" we mean time. This is a folklore fact dictated by the Laws of Nature. Namely, if the parallel algorithm uses 2^n processing elements, regardless of whether the computational model assumes bounded fan-in and fan-out or not, it cannot run in time polynomial in n, because *physical space* has us in its tyranny. Namely, if we use 2^n processing elements of, say, unit size each, then the tightest they can be packed is in a three-dimensional sphere of volume $N=2^n$. Assuming that the units have no "funny" shapes, e.g., are spherical themselves, no unit in the enveloping sphere can be closer to all other units than a distance of radius R (Fig. 1),

$$R=\left(\frac{3N}{4\pi}\right)^{1/3}. \tag{1.1}$$

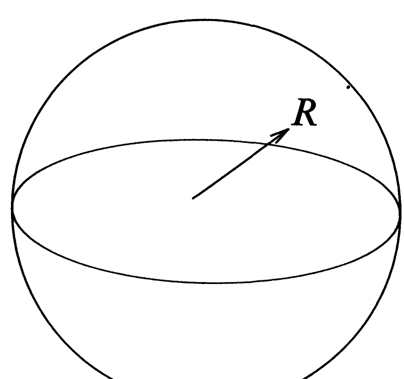

FIG. 1

[1] When the operations of a computation are executed serially in a single Central Processing Unit (CPU), each one entails a "fetch data from memory to CPU; execute operation in CPU; store data in memory" cycle. The cost of this cycle, and therefore of the total computation, is dominated by the cost of the memory accesses which are essentially operation-independent. This is called the "von Neumann" bottleneck, after the brilliant Hungarian mathematician John von Neumann.

[2] For example, in [10] it is demonstrated that any program that requires T steps on a CRCW PRAM with n processors and m shared variables (m polynomial in n) can be simulated by a bounded degree network of n processors such as the Ultracomputer [7] that runs in deterministic "time" $O(T(\log n)^2 \log\log n)$ steps.

Unless there is a major advance in physics, it is impossible to transport signals over $2^{\alpha n}$ ($\alpha > 0$) distance in polynomial $p(n)$ time. In fact, the assumption of the bounded speed of light says that the lower time bound on *any* computation using 2^n processing elements is $\Omega(2^{n/3})$ outright. Or, for the case of NC computations which use n^{α} processors, $\alpha > 0$, the lower bound on the computation time is $\Omega(n^{\alpha/3})$.[3] Science fiction buffs may want to keep open the option of embedding circuits in hyper dimensions. Counter to intuition, this does not help—at least, not all the way (see the Appendix). The situation is worse than it appears. At present, many popular multicomputer architectures are based on highly symmetric communication networks with small diameter. Like all networks with small diameter, such networks will suffer from the communication bottleneck above, i.e., they necessarily contain *some* long interconnects (embedded edges). However, the desirable fast permutation properties of symmetric networks do not come free, since they require that the average of *all* interconnects is long. (Note that "embedded edge," "wire," and "interconnect" are used synonymously.) This brings us to the main topic of this paper, the analysis of the *amount of wire required.* To prevent arguments that the results have little practical importance because they hold only asymptotically, or because processors are huge and wires thin, we calculate precisely without hidden constants[4] and assume that wires have length but no volume and can pass through everything. The key Theorem 2 in the next section gives a lower bound on the *average* edge length for *arbitrary* graphs that is arguably optimal.

Let us illustrate the novel approach with a popular architecture, say the *binary n-cube.* Recall that this is the network with $N = 2^n$ nodes, each of which is identified by an n-bit name. There is a two-way communication link between two nodes if their identifiers differ by a single bit. The network is represented by an undirected graph $C = (V, E)$, with V the set of nodes and $E \subseteq V \times V$ the set of edges, each edge corresponding with a communication link. There are $n2^{n-1}$ edges in C. Let C be embedded in three-dimensional Euclidean space, each node as a sphere with unit volume. The distance between two nodes is the Euclidean distance between their centers. Let x be any node of C. There are at most $2^n/8$ nodes within Euclidean distance $R/2$ of x, with R as in (1.1). Then, there are $\geqq 7 \cdot 2^n/8$ nodes at Euclidean distance $\geqq R/2$ from x (Fig. 2). Construct a spanning tree T_x in C of depth n with node x as the root. Since the binary n-cube has diameter n, such a shallow tree exists. There are N nodes in T_x, and $N-1$ paths from root x to another node in T_x. Let P be such a path, and let $|P|$ be the *number of edges* in P. Then $|P| \leqq n$. Let *length*(P) denote the Euclidean length of the embedding of P. Since 7/8th of all nodes are at Euclidean distance at least $R/2$ of root x, the average of *length*(P) satisfies

$$(N-1)^{-1} \sum_{P \in T_x} \mathit{length}(P) \geqq \frac{7R}{16}.$$

The average Euclidean length of an embedded edge *in a path P* is bounded below as follows:

$$(N-1)^{-1} \sum_{P \in T_x} \left(|P|^{-1} \sum_{e \in P} \mathit{length}(e) \right) \geqq \frac{7R}{16n}. \tag{1.2}$$

[3] It is sometimes argued that this effect is significant for large values of n only, and therefore can safely be ignored. However, in the theory of computation many results are of asymptotic nature, i.e., they hold only for large values of n, so the effect is especially relevant there.

[4] Ω is used sometimes to simplify notation. The constant of proportionality can be reconstructed easily in all cases, and is never very small.

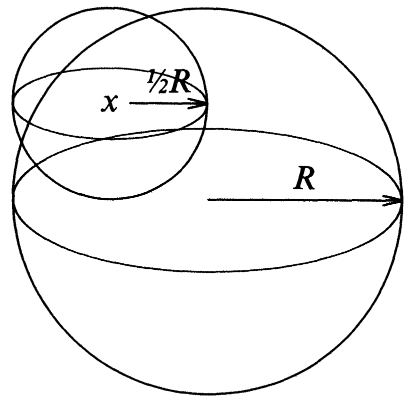

FIG. 2. *At most 1/8th of all nodes in the large sphere are also contained in the small sphere centered on x.*

This does *not* give a lower bound on the average Euclidean length of an edge, the average taken *over all edges* in T_x. To see this, note that if the edges incident with x have Euclidean length $7R/16$, then the average edge length *in each path* from the root x to a node in T_x is $\geqq 7R/16n$, even if all edges not incident with x have length 0. However, the average edge length *in the tree* is dominated by the many short edges near the leaves, rather than the few long edges near the root. In contrast, in the case of the binary n-cube, because of its symmetry, if we squeeze a subset of nodes together to decrease local edge length, then other nodes are pushed farther apart increasing edge length again. We can make this intuition precise.

LEMMA 1. *The average Euclidean length of the edges in the three-space embedding of C is at least $7R/(16n)$.*

Proof. Denote a node a in C by an n-bit string $a_1a_2 \cdots a_n$, and an edge (a, b) between nodes a and b differing in the kth bit by

$$(a_1 \cdots a_{k-1}a_ka_{k+1} \cdots a_n, a_1 \cdots a_{k-1}(a_k \oplus 1)a_{k+1} \cdots a_n)$$

where $\oplus$ denotes modulo 2 addition. Since C is an undirected graph, an edge $e=(a, b)$ has two representations, namely (a, b) and (b, a). Consider the set A of automorphisms $\alpha_{v,j}$ of C consisting of

(1) modulo 2 addition of a binary n-vector v to the node representation, followed by

(2) a cyclic rotation over distance j.

Formally, let $v = v_1v_2 \cdots v_n$, with $v_i = 0, 1$ $(1 \leqq i \leqq n)$, and let j be an integer $1 \leqq j \leqq n$. Then $\alpha_{v,j}$: $V \to V$ is defined by

$$\alpha_{v,j}(a) = b_{j+1} \cdots b_nb_1 \cdots b_j$$

with $b_i = a_i \oplus v_i$ for all i, $1 \leqq i \leqq n$.

Consider the spanning trees $\alpha(T_x)$ isomorphic to T_x, $\alpha \in A$. The argument used to obtain (1.2) implies that for *each* α in A separately, in each path $\alpha(P)$ from root $\alpha(x)$ to a node in $\alpha(T_x)$, the average of *length*$(\alpha(e))$ over all edges $\alpha(e)$ in $\alpha(P)$ is at least $7R/16n$. Averaging (1.2) additionally over all α in A, the same lower bound applies:

$$(1.3) \qquad (N \log N)^{-1} \sum_{\alpha \in A} \left[(N-1)^{-1} \sum_{P \in T_x} \left(|P|^{-1} \sum_{e \in P} \textit{length}\,(\alpha(e)) \right) \right] \geqq \frac{7R}{16n}.$$

Now fix a particular edge e in T_x. We sum *length* $(\alpha(e))$ over all α in A, and show that this sum equals twice the total edge length. Together with (1.3), this will yield the desired result. For each edge f in C there are $\alpha_1, \alpha_2 \in A$, $\alpha_1 \neq \alpha_2$, such that $\alpha_1(e) = \alpha_2(e) = f$, and for all $\alpha \in A - \{\alpha_1, \alpha_2\}$, $\alpha(e) \neq f$. (For $e = (a, b)$ and $f = (c, d)$ we have $\alpha_1(a) = c$, $\alpha_1(b) = d$, and $\alpha_2(a) = d$, $\alpha_2(b) = c$.) Therefore, for each $e \in E$,

$$\sum_{\alpha \in A} \textit{length}\,(\alpha(e)) = 2 \sum_{f \in E} \textit{length}\,(f).$$

Then, for *any* path P in C,

$$\sum_{e \in P} \sum_{\alpha \in A} \textit{length}\,(\alpha(e)) = 2|P| \sum_{f \in E} \textit{length}\,(f). \tag{1.4}$$

Rearranging the summation order of (1.3), and substituting (1.4), yields the lemma. □

2. Interconnect length in Euclidean space. Deriving the total required wire length for embeddings of networks in Euclidean space, I will not make *any* assumptions about the volume of a wire of unit length, or the way they are embedded in space. Compare this with previous VLSI-related arguments (see e.g., [9]) which are the only other ones on this issue known to me. It is consistent with our results that wires have *zero* volume, and that *infinitely* many wires pass through a unit area. Concretely, the problem is posed as follows. Let $G = (V, E)$ be a finite undirected graph, without loops or multiple edges, *embedded* in Euclidean d-space. (For the physical space in which we put our computers, $d = 3$.) Let each embedded node have unit *volume.* For convenience of the argument, each node is embedded as a sphere, and is *represented* by the single point in the center. The *distance* between a pair of nodes is the Euclidean distance between the points representing them. The *length* of the embedding of an edge between two nodes is the distance between the nodes. How large does the *average* edge length need to be?

Theorem 2 expresses a lower bound on this quantity for *any* graph, in terms of certain symmetries and diameter. The new argument is based on graph automorphism, graph topology, and Euclidean metric. For each graph topology I have examined, the resulting lower bound turned out to be sharp. This includes the binary n-cube, cube-connected cycles (CCC), complete graph, star, complete binary tree, and meshes with wraparound such as ring and torus. It could be that the lower bound is optimal in general. All mentioned graphs, except the cube-connected cycles and tree, exhibit a type of symmetry called edge-symmetry. Because of the significance of this class of graphs, in Corollary 4 we set off a lower bound on the average interconnect length for edge-symmetric graphs in general.

2.1. Lower bound based on symmetry and diameter. What symmetry of a graph yields large edge length? Not that of the complete binary tree. There the diameter is small, yet the average Euclidean length of an embedded edge is $O(1)$. This is borne out by the familiar H-tree layout [9], where the average edge length is less than 3 or 4. The symmetry property we are after is "edge-symmetry." We recall the definitions from [2]. Let $G = (V, E)$ be a simple undirected graph, and let Γ be the automorphism group of G. Two edges $e_1 = (u_1, v_1)$ and $e_2 = (u_2, v_2)$ of G are *similar* if there is an automorphism γ of G such that $\gamma(\{u_1, v_1\}) = \{u_2, v_2\}$. We consider only connected graphs. The relation "similar" is an equivalence relation, and partitions E into nonempty equivalence classes, called *orbits*, $E_1, \cdots, E_m$. We say that Γ *acts transitively* on each E_i, $i = 1, \cdots, m$. A graph is *edge-symmetric* if every pair of edges are similar ($m = 1$). The following property of orbits is obvious.

Property. For each pair of edges $e_1, e_2 \in E_i$, the set $\{\gamma \in \Gamma: \gamma(e_1) = e_2\}$ has $|\Gamma|/|E_i|$ elements, $i = 1, \cdots, m$. (Hint: Let $0 \in E_i$ and $\Gamma_0 = \{\gamma \in \Gamma: \gamma(0) = 0\}$. For $e, f \in E_i$, define $\gamma_{ef} \in \Gamma$ by $\gamma_{ef}(e) = f$. Fix e and f arbitrarily. Then $\gamma \in \gamma_{e0}\Gamma_0\gamma_{0f}$ if and only if $\gamma_{e0}^{-1}\gamma\gamma_{0f}^{-1} \in \Gamma_0$.)

We need the following notions. Let $D < \infty$ be the *diameter* of G. If x and y are nodes, then $d(x, y)$ denotes the number of edges in a *shortest path* between them. For $i = 1, \cdots, m$, define $d_i(x, y)$ as follows. If (x, y) is an edge in E_i then $d_i(x, y) = 1$, and if (x, y) is an edge not in E_i then $d_i(x, y) = 0$. Let Π be the set of shortest paths between x and y. If x and y are not incident with the same edge, then $d_i(x, y) = |\Pi|^{-1} \sum_{P \in \Pi} \sum_{e \in P} d_i(e)$. Clearly,

$$d_1(x, y) + \cdots + d_m(x, y) = d(x, y) \leqq D. \tag{2.1}$$

Denote $|V|$ by N. The ith *orbit frequency* is

$$\delta_i = N^{-2} \sum_{x,y \in V} \frac{d_i(x, y)}{d(x, y)},$$

$i = 1, \cdots, m$. Finally, define the *orbit skew coefficient* of G as $M = \min\{|E_i|/|E|: 1 \leqq i \leqq m\}$. Consider a d-space embedding of G, with embedded nodes, distance between nodes, and edge length as above. Let R be the *radius* of a d-space sphere with volume N, e.g., (1.1) for $d = 3$. We are now ready to state the main result. Just in case the reader does not notice, (i) is the most general form.

THEOREM 2. *Let graph G be embedded in d-space with the parameters above, and let $C = (2^d - 1)/2^{d+1}$*[5]

(i) *Let $l_i = |E_i|^{-1} \sum_{e \in E_i} l(e)$ be the average length of the edges in orbit E_i, $i = 1, \cdots, m$. Then, $\sum_{1 \leqq i \leqq m} l_i \geqq \sum_{1 \leqq i \leqq m} \delta_i l_i \geqq CRD^{-1}$.*

(ii) *Let $l = |E|^{-1} \sum_{e \in E} l(e)$ be the average length of an edge in E. Then, $l \geqq CRMD^{-1}$.*

Proof. Without loss of generality, we give the proof for the physically relevant case $d = 3$. If x and y are nodes, let $l(x, y)$ be the Euclidean *distance* between x and y in three-space. For $i = 1, \cdots, m$, define $l_i(x, y)$ as follows. If (x, y) is an edge in E_i, then $l_i(x, y) = l(x, y)$, and if (x, y) is an edge not in E_i, then $l_i(x, y) = 0$. If x and y are not incident with the same edge, then $l_i(x, y) = |\Pi|^{-1} \sum_{P \in \Pi} \sum_{e \in P} l_i(e)$, with Π as above. By the triangle inequality,

$$l(x, y) \leqq l_1(x, y) + \cdots + l_m(x, y). \tag{2.2}$$

Consider Fig. 2 again. Let x be any node of G. There are at most $N/8$ nodes within distance $R/2$ of x, with R given by (1.1). Therefore, there are $\geqq 7N/8$ nodes at distance $\geqq R/2$ from x, for N large enough. Thus, the sum of all $l(x, y)$, taken over all node pairs x, y, satisfies

$$\sum_{x,y \in V} l(x, y) \geqq \frac{7RN^2}{16}. \tag{2.3}$$

Using (2.1) and (2.2), we obtain from (2.3),

$$\sum_{x,y \in V} \sum_{i=1}^{m} \frac{l_i(x, y)}{d(x, y)} \geqq \sum_{x,y \in V} \frac{l(x, y)}{d(x, y)} \geqq \frac{7RN^2}{16D}. \tag{2.4}$$

[5] This constant C can be improved. For $d = 3$, $C = 7/16$ is the value of $c(1 - c^3)$ for $c = 2^{-1}$. This function reaches its optimum value $(3/4)2^{-2/3}$ for $c = 2^{-2/3}$. By refining the argument we can improve the constant to $\frac{3}{4}$. Namely, to obtain (2.3), sum $(c, c + dc]R\delta(x, y)$ with $\delta(x, y) = 1$ if $cR < l(x, y)$ and $\delta(x, y) = 0$ otherwise, with c ranging from 0 to 1, for each pair of nodes x, y. This replaces $C = 7/16$ in (2.3) by $C = \int_0^1 (1 - c^3) dc = \frac{3}{4}$. Similarly, in two dimensions we can improve C from 3/8 to 2/3.

Now fix a particular edge e in some E_i. We average $l(\gamma(e))$ over all γ in Γ. By the property above, there are precisely $|\Gamma|/|E_i|$ distinct automorphisms in Γ that map edge e onto edge f, for each pair $e, f \in E_i$. Therefore, the sum of $l(\gamma(e))$ over all γ in Γ equals *precisely* $|\Gamma|/|E_i|$ times the sum of the lengths of all edges in E_i. Formally,

$$|\Gamma|^{-1} \sum_{\gamma \in \Gamma} l(\gamma(e)) = |E_i|^{-1} \sum_{f \in E_i} l(f) \quad \text{for each } e \in E_i, \quad i = 1, \cdots, m,$$

and therefore, for all $x, y \in V$,

$$|\Gamma|^{-1} \sum_{\gamma \in \Gamma} l_i(\gamma(x), \gamma(y)) = |E_i|^{-1} d_i(x, y) \sum_{f \in E_i} l(f) \quad \text{for } i = 1, \cdots, m. \tag{2.5}$$

We now finish the argument. Averaging (2.4) additionally over all γ in Γ, leaves the lower bound invariant:

$$|\Gamma|^{-1} \sum_{\gamma \in \Gamma} \sum_{x,y \in V} \sum_{i=1}^{m} \frac{l_i(\gamma(x), \gamma(y))}{d(\gamma(x), \gamma(y))} \geqq \frac{7RN^2}{16D}. \tag{2.6}$$

By rearranging the summation order in (2.6), and substitution of (2.5), we obtain

$$\sum_{i=1}^{m} \sum_{x,y \in V} \frac{d_i(x, y)}{d(x, y)} |E_i|^{-1} \sum_{e \in E_i} l(e) \geqq \frac{7RN^2}{16D}.$$

That is, $\sum_{1 \leqq i \leqq m} \delta_i l_i \geqq 7R/(16D)$. Since $\delta_i \leqq 1$, $i = 1, \cdots, m$, this proves (i). For the *average* edge length l, this yields $l = \sum_{1 \leqq i \leqq m} (|E_i|/|E|) l_i \geqq M \sum_{1 \leqq i \leqq m} l_i$, which proves (ii). □

Example 1. Binary n-cube. Let Γ be an automorphism group of the binary n-cube, e.g., A in the proof of Lemma 1. Let $N = 2^n$. The orbit of each edge under Γ is E. Substituting R, D, $m = 1$, and $d = 3$ in Theorem 2(i) proves Lemma 1. Denote by L the *total* edge length $\sum_{f \in E} l(f)$ in the three-space embedding of C. Then

$$L \geqq \frac{7RN}{32}. \tag{2.7}$$

Recapitulating, the sum *total* of the lengths of the edges is $\Omega(N^{4/3})$, and the *average* length of an edge is $\Omega(N^{1/3} \log^{-1} N)$. (In two dimensions we obtain in a similar way $\Omega(N^{3/2})$ and $\Omega(N^{1/2} \log^{-1} N)$, respectively.)

Example 2. Cube-connected cycles. The binary n-cube has the drawback of unbounded node degree. Therefore, in the fixed-degree version of it, each node is replaced by a *cycle* of n trivalent nodes [9]; whence the name *cube-connected cycles* or CCC. If $N = n2^n$, then the CCC version, say CCC $= (V, E)$, of the binary n-cube has N nodes, $3N/2$ edges, and diameter $D < 2.5n$.

COROLLARY 3. *The average Euclidean length of edges in a three-space embedding of CCC is at least* $7R/(120n)$.

Proof. Denote a node a by an n-bit string with one marked bit, $a = a_1 \cdots a_{i-1}\overline{a_i}a_{i+1} \cdots a_n$. There is an edge (a, b) between nodes $a = a_1 \cdots a_{i-1}\overline{a_i}a_{i+1} \cdots a_n$ and $b = a_1 \cdots a_{j-1}\overline{b_j}\, a_{j+1} \cdots a_n$, if either $i \equiv j \pm 1 (\text{mod } n)$, $a_i = b_i$ and $a_j = b_j$ (edges in cycles), or $i = j$ and $a_i \neq b_i$ (edges between cycles). Consider the set A of automorphisms $\alpha_{v,j}$, with $v = v_1 \cdots v_n$ a binary n-vector and j an integer $1 \leqq j \leqq n$, such that

$$\alpha_{v,j}(a_1 \cdots a_{i-1}\overline{a_i}a_{i+1} \cdots a_n) = b_{j+1} \cdots b_n b_1 \cdots b_j,$$

with $b_i = \overline{a_i \oplus v_i}$ and $b_k = a_k \oplus v_k$ for $k \neq i$, $1 \leqq k \leqq n$. Clearly, A is a subgroup of the automorphism group of CCC. The similarity relation induced by A partitions E in two orbits: the set of *cycle* edges and the set of *noncycle* edges. Since there are $N/2$

noncycle edges, N cycle edges, and $3N/2$ edges altogether, the orbit skew coefficient M is $\frac{1}{3}$. Substitution of R, D, M, and $d=3$ in Theorem 2(ii) yields the corollary. □

That is, the *total* edge length is $\Omega(N^{4/3}\log^{-1} N)$ and the *average* edge length is $\Omega(N^{1/3}\log^{-1} N)$. (In two dimensions $\Omega(N^{3/2}\log^{-1} N)$ and $\Omega(N^{1/2}\log^{-1} N)$, respectively.) Similar lower bounds are expected to hold for other fast permutation networks like the *butterfly*, *shuffle-exchange*, and *de Bruijn* graphs.

Example 3. Edge-symmetric graphs. Recall that a graph $G=(V, E)$ is *edge-symmetric* if each edge is mapped to every other edge by an automorphism in Γ. We set off this case especially, since it covers an important class of graphs. (It includes the binary n-cube but excludes CCC.) Let $|V|=N$ and $D<\infty$ be the diameter of G. Substituting R, $m=1$, and $d=3$ in Theorem 2(i) we obtain the following.

COROLLARY 4. *The average Euclidean length of edges in a three-space embedding of an edge-symmetric graph is at least* $7R/(16D)$.

For the complete graph K_N, this results in an average wire length of $\geqq 7R/16$. That is, the average wire length is $\Omega(N^{1/3})$, and the total wire length is $\Omega(N^{7/3})$.

For the complete bigraph $K_{1,N-1}$ (the star graph on N nodes) we obtain an average wire length of $\geqq 7R/32$. That is, the average wire length is $\Omega(N^{1/3})$, and the total wire length is $\Omega(N^{4/3})$.

For an N-node δ-dimensional mesh with wraparound (e.g., a ring for $\delta=1$, and a torus for $\delta=2$; for a formal definition see Appendix), this results in an average wire length of $\geqq 7R/(8\delta N^{1/\delta})$. That is, the average wire length is $\Omega(\delta^{-1}N^{(\delta-3)/3\delta})$, and the total wire length is $\Omega(N^{(4\delta-3)/3\delta})$.

To give some indication of the scope of Corollary 4, we note that every edge-symmetric graph with no isolated nodes is node-symmetric or bipartite, by a theorem attributed to Elayne Dauber [2], and that every Cayley graph is symmetric [1]. (A graph is symmetric if it is both node-symmetric and edge-symmetric. A graph is node-symmetric if for each pair of nodes there is an automorphism that maps one to the other.)

Example 4. Complete binary tree. The complete binary tree T_n on $N-1$ nodes ($N=2^n$) has $n-1$ orbits $E_1, \cdots, E_{n-1}$. Here E_i is the set of edges at level i of the tree, with E_1 is the set of edges incident with the leaves, and E_{n-1} is the set of edges incident with the root. Let l_i and l be as in Theorem 2 with $m=n-1$. Then $|E_i|=2^{n-i}$, $i=1, \cdots, n-1$, the orbit skew coefficient $M=2/(2^n-2)$, and we conclude from Theorem 2(ii) that l is $\Omega(N^{-2/3}\log^{-1} N)$ for $d=3$. This is consistent with the known fact l is $O(1)$. However, we obtain significantly stronger bounds using the more general part (i) of Theorem 2. In fact, we can show that one-space embeddings of complete binary trees with $o(\log N)$ average edge length are impossible.[6]

COROLLARY 5. *The average Euclidean length of edges in a d-space embedding of a complete binary tree is* $\Omega(1)$ *for* $d=2, 3$, *and* $\Omega(\log N)$ *for* $d=1$.

Proof. Consider d-space embeddings of T_n, $d\in\{1, 2, 3\}$ and $n>1$. By Theorem 2,

$$\sum_{i=1}^{n-1} \delta_i l_i \geqq CRD^{-1}. \tag{2.8}$$

CLAIM. $\delta_i \leqq (n-1)^{-1}$, for $i=1, \cdots, n-1$.

Proof of claim. The proof is by induction on n. Denote by $\delta_i^{(n)}$ the ith orbit frequency of T_n, the complete binary tree with 2^n-1 nodes. Note that T_{n+1} consists of two copies of T_n, with the roots attached to a root node that is in neither

[6] Using $\sum_{i=1}^{n-1} l_i \geqq CRD^{-1}$ instead of (2.8), also yields l is $\Omega(1)$ for $d=2, 3$, but only l is $\Omega(\log\log N)$ for $d=1$.

of them. Set $\delta_i^{(j)}=0$ for $i \geqq j$. For $n=2$ the claim holds trivially. Assume the claim holds for $n \geqq 2$. Then we prove it holds for $n+1$, as follows. We obtain $(2^{n+1}-1)^2\delta_i^{(n+1)}$ by dividing $\sum_{x,y\in V} d_i(x,y)/d(x,y)$ in two parts: with both nodes x, y in the same T_n subtree and with nodes x, y in different subtrees. The first subsum equals $2(2^n-1)^2\delta_i^{(n)}$. To obtain the second subsum, we sum $d_i(x,y)/d(x,y)$ with x and y ranging over the consecutive levels of different T_n-subtrees (so the shortest path between x and y contains the root of T_{n+1}). This yields the following recurrences, for each $i=1,\cdots,n$ (with $\delta_n^{(n)}=0$):

$$(2^{n+1}-1)^2\delta_i^{(n+1)}=2(2^n-1)^2\delta_i^{(n)}+\sum_{j=0}^{n} 2^j \sum_{k=n-i+1}^{n} \frac{2^{k-1}}{k+j}.$$

Evaluating the double sum for $i=n$, and substituting $\delta_i^{(n)} \leqq (n-1)^{-1}$, we find after due computation, $\delta_i^{(n+1)} \leqq n^{-1}$. □

Substitution of $|E_i|$, $|E|$, m in the expression for l in Theorem 2 gives

$$l=(2^n-2)^{-1}\sum_{i=1}^{n-1} 2^{n-i} l_i. \tag{2.9}$$

Substitute in (2.8) the values of C, R (depending on d) and $D=2(n-1)$. Next, substitute $(n-1)^{-1}$ for δ_i and multiply both sides with $n-1$. Use the resulting expression to substitute in (2.9), after rearranging the summation, as follows:

$$l=(2^n-2)^{-1}\left(\sum_{j=1}^{n-1} 2^{j-1}\sum_{i=1}^{n-j} l_i+\sum_{i=1}^{n-1} l_i\right)\in\Omega\left(2^{-(1-1/d)n}\sum_{j=1}^{n-1} 2^{(1-1/d)j}\right).$$

Therefore, for $d=2, 3$, we obtain l is $\Omega(1)$. However, for $d=1$, l is $\Omega(\log N)$. □

2.2. Optimality conjecture. There is evidence that the lower bound of Theorem 2 is optimal. Namely, it is within a constant multiplicative factor of an upper bound for several example graphs of various diameters. Consider only three-dimensional Euclidean embeddings, and recall the assumption that wires have length but no volume, and can pass through nodes. For the *complete graph* K_N with diameter 1, the lower bound on the average wire length is $7R/16$, while $2R$ is a trivial upper bound. For the *star* graph on N nodes the bounds are $7R/32$ and $2R$, respectively. The upper bound on the total wire length to embed the *binary n-cube* requires more work. Let $N=2^n$.

The construction is straightforward. For convenience we assume now that each node is embedded as a three-space cube of volume 1. Recursively, embed the binary n-cube in a cube of three-dimensional Euclidean space with *sides* of length S_n. Use eight copies of binary $(n-3)$-cubes embedded in Euclidean $S_{n-3}\times S_{n-3}\times S_{n-3}$ cubes, with $S_{n-3}=S_n/2$. Place the eight small cubes into the large cube by fitting each small cube into an octant of the large cube. First connect the copies pairwise along the first coordinate to form four binary $(n-2)$ cubes. Connect these four pairwise along the second coordinate to form two binary $(n-1)$ cubes, which in turn are connected along the third coordinate into one binary n-cube. This requires no more than $4\cdot 2^{n-3}$ wires of length at most $\sqrt{3/2}\cdot S_n$, another $2\cdot 2^{n-2}$ wires of length at most $3S_n/2$ and 2^{n-1} wires of length at most $\sqrt{3}\cdot S_n$. Assume $S_1=1$ and $n-1$ is a multiple of 3. Since $S_n=2S_{n-3}$, we have $S_n=2^{(n-1)/3}$. The *total* wire length $L(n)$ required to embed the binary n-cube is

$$\begin{aligned} L(n) &\leqq 2^{n-1}(\sqrt{3/2}+3/2+\sqrt{3})S_n+8L(n-3)\\ &\leqq \sum_{i=1}^{(n-1)/3} 2^{4i}\cdot 2^{n-1-3i}(\sqrt{3/2}+3/2+\sqrt{3}). \end{aligned}$$

Substitute $i=-j+(n-1)/3$ and round off the bracketed sum to 5 to obtain

$$L(n)<5\cdot 2^{4(n-1)/3}\sum_{j=0}^{(n-4)/3}2^{-j}.$$

Summing the infinite series $\sum_{j=0}^{\infty}2^{-j}$ yields an upper bound $L(n)<4N^{4/3}$. Together with Lemma 1, the optimum of the average interconnect length for the binary n-cube is in between $7R/16n$ and $8N^{1/3}/n$.

For the *cube-connected cycles* with $N=n2^n$ nodes, we derive an upper bound by the same argument. Squeeze the n nodes of each cycle in a three-space cube of volume n in the obvious way. This takes, say, about $L_1<n2^n$ total interconnect length for the cycle edges. Recall that each such cycle corresponds to a particular node of the binary n-cube above. Apply the same construction as for the binary n-cube with $S_1=n^{1/3}$. Then obtain $L_2<4\cdot 2^{4n/3}n^{1/3}$ total interconnect length for the edges between cycles. Together with Corollary 3, we obtain that the optimum of the average interconnect length for the cube-connected cycles is in between $7R/120n$ and $8N^{1/3}/(3n)+2/3$. For *δ-dimensional meshes with wraparound*, with $\delta=1$, 2, 3 and diameter $2^{-1}\delta N^{1/\delta}$, a lower bound of $\Omega(1)$ follows from Corollary 4, and the upper bound is $O(1)$ by the obvious embedding. Note that $\delta=1$ is the ring and $\delta=2$ is the torus. For the *complete binary* tree, for $d=2$, 3, the H-tree construction gives an average edge length $O(1)$ [9], matching the $\Omega(1)$ lower bound. In the one-dimensional case, the obvious embedding gives $O(\log N)$ average edge length, matching the lower bound $\Omega(\log N)$ of Corollary 5.

2.3. Robustness. Theorem 2 is robust in the sense that if $G'=(V',E')$ is a subgraph of $G=(V,E)$, and the theorem holds for either one of them, then a related lower bound holds for the other. Essentially, this results from the relation between the orbit frequencies of G, G'. Let us look at some examples, with $d=3$.

Let a graph G have the binary n-cube C as a subgraph and $N=2^n$. Let G have $N'\leqq 8N$ nodes and at most $N'\log N'$ edges. The lower bound on the total wire length $L(G)$ of a three-space embedding of G follows trivially from $L(G)\geqq L(C)$, with $L(C)\geqq 7RN/32$ the total wire length of the binary n-cube. Therefore, expressing the lower bounds in N' and radius R' of a sphere with volume N' yields $L(G)\geqq 7R'N'/512$, and the average edge length of G is at least $7R'/(512\log N')$.

Let the binary n-cube C have a subgraph G with $n2^{n-1}-2^{n-5}$ edges. The lower bound on the total wire length $L(G)$ of a three-space embedding of G follows from the observation that each deleted edge of C has length at most twice the diameter R of (1.1). That is, $L(G)\geqq L(C)-2^{n-4}R$ with $L(C)$ as above. Note that G has $N'\geqq 2^n-(2^{n-6}/n)$ nodes. Therefore, expressing the lower bounds in N' and radius R' of a sphere with volume N' yields $L(G)\geqq 5RN/32\geqq 5R'N'/32$, and the average edge length of G is at least $5R/16n\sim 5R'/(16\log N')$.

3. Interconnect length and volume. An effect that becomes increasingly important at the present time is that most space in the device executing the computation is taken up by the wires. Under very conservative estimates that the unit length of a wire has a volume which is a constant fraction of that of a component it connects, we can see above that in three-dimensional layouts for binary n-cubes, the volume of the $N=2^n$ components performing the actual computation operations is an asymptotic fastly vanishing fraction of the volume of the wires needed for communication:

$$\frac{\text{volume computing components}}{\text{volume communication wires}}\in o(N^{-1/3})$$

If we charge a constant fraction of the unit volume for a unit wire length and add the volume of the wires to the volume of the nodes, then the volume necessary to embed the binary n-cube is $\Omega(N^{4/3})$. However, this lower bound ignores the fact that the added volume of the wires pushes the nodes further apart, thus necessitating longer wires again. How far does this go? A rigorous analysis is complicated and is not important here. The following intuitive argument indicates well enough what we can expect. Denote the volume taken by the nodes as V_n and the volume taken by the wires as V_w. The total volume taken by the embedding of the cube is $V_t = V_n + V_w$. The total wire length required to lay out a binary n-cube as a function of the volume taken by the embedding is, substituting $V_t = 4\pi R^3/3$ in (2.7),

$$L(V_t) \geqq \frac{7N}{32}\left(\frac{3V_t}{4\pi}\right)^{1/3}.$$

Since $\lim_{n\to\infty} V_n/V_w \to 0$, assuming unit wire length of unit volume, we set $L(V_t) \sim V_t$. This results in a better estimate of $\Omega(N^{3/2})$ for the volume needed to embed the binary n-cube. When we want to investigate an upper bound to embed the binary n-cube under the current assumption, we have a problem with the unbounded degree of unit volume nodes. There is no room for the wires to come together at a node. For comparison, therefore, consider the fixed-degree version of the binary n-cube, the CCC (see above), with $N = n2^n$ trivalent nodes and $3N/2$ edges. The same argument yields $\Omega(N^{3/2} \log^{-3/2} N)$ for the volume required to embed CCC with unit volume per unit length wire. It is known, that every small degree N-vertex graph, e.g., CCC, can be laid out in a three-dimensional grid with volume $O(N^{3/2})$ using a unit volume per unit wire length assumption [5]. This neatly matches the lower bound.

Because of current limitations to layered VLSI technology, previous investigations have focused on embeddings of graphs in two-space (with unit length wires of unit volume). We observe that the above analysis for two dimensions leads to $\Omega(N^2)$ and $\Omega(N^2 \log^{-2} N)$ volumes for the binary n-cube and the cube-connected cycles, respectively. These lower bounds have been obtained before, using *bisection-width* arguments and are known to be optimal [9]. It can be even worse, namely, in [6], [12] it is shown that we cannot always assume that a unit length of wire has $O(1)$ volume (for instance, if we want to drive the signals to very high speed on chip).

4. Conclusion. In contrast to other investigations, my goal here was to derive *hard* lower bounds on the total wire length *independent* of the ratio between the volume of a unit length wire and the volume of a processing element. Clearly this is desirable, since this ratio changes with different technologies and granularity of computing components. The arguments we have developed are purely geometrical, apply to any graph, and give optimal lower bounds in all cases we have examined.

Such technology-independent, but huge, lower bounds are a theoretical prelude to many *wiring problems* currently starting to plague computer designers and chip designers alike. Formerly, a wire had magical properties of transmitting data "instantly" from one place to another (or better, to many other places). A wire did not take room, did not dissipate heat, and did not cost anything—at least, not enough to worry about. This was the situation when the number of wires was low, somewhere in the hundreds. Current designs use many millions of wires (on chip), or possibly billions of wires (on wafers). In a computation of parallel nature, most of the time seems to be spent on communication—transporting signals over wires. Thus, thinking that the von Neumann bottleneck has been conquered by nonsequential computation, we are unaware that a non von Neumann communication bottleneck looms large. The following innominate quote covers this matter admirably:

> Without me they fly they think;
> But when they fly I am the wings.

It is clear that these communication mishaps will influence the architecture and the algorithms to be designed for the massive multiprocessors of the future, just like existing algorithms influenced (or were inspired by) the novel architectures of today. What is needed, therefore, are *realistic* formal models for nonsequential computation. In particular, we need to formulate the appropriate cost measures for multicomputer computations. Such costs must account for the communication overhead in (physical) time due to the computer aggregates used in the computation and the overhead in space due to the topology of those aggregates. That is beyond the scope of this paper.

Mesh-connected architectures may be the ultimate solution for interconnecting the extremely large (in numbers) computer complexes of the future. Mesh architectures have desirable properties of *scalability*, *modular extensibility*, *and uniformity*, when embedded in physical space. These notions are generally used in a very loose fashion, and with a great deal of intuition, so I do not try to define them here. Circuits with lower bound $f(N)$, $f(N)\to\infty$ for $N\to\infty$, on the average interconnect length do not scale well. (N is the number of nodes.) Namely, composing a larger such circuit from smaller ones, the average wire length needs to increase. Thus, embeddings of such circuits are not uniformly modular extensible. This positive dependency of the interconnect length on the number of nodes to be connected we call nonscalability.

Nonscalability. No edge-symmetric graph on N nodes with a diameter $o(N^{1/3})$ is scalable (i.e., uniformly modular extensible) when embedded in physical space.

Tomorrow, optical communication will be used in multicomputers, either wireless by means of lasers/infrared light or by using virtually unlimited bandwidth optical fiber or integrated waveguides [8]. In the current jargon: we can obtain three-dimensional mesh interconnect structures by stacking wafer circuit boards and providing optical interconnections vertically between wafers over the entire wafer in addition to planar connections. This may use hybrid mounting of optical components, combined with integrated optical waveguides and lenses on a large area silicon wafer-scale integrated (WSI) electronic circuit combining electronic and photonic functions [4]. However, it is unlikely that any clever scheme or technology will free us from practical communication problems forever. Even though Nature is not malicious, she is subtle.

Appendix A. What happens with embeddings in higher-dimensional spaces? Lest the reader conclude that I indulge in the same avoidance of reality that I decry in others, I have delegated this digression to the Appendix. These mathematical curiosities have no more bearing on realistic formal models for multicomputers than space warps have on the theory of propulsion of space vehicles.

A.1. Communication and interconnect length in higher dimensions. Assume that a node (processor) has unit volume, say spherical, in any number d of dimensions we care to consider. This is in order to obtain comparable reasoning to the physical relevant case of three dimensions. Our intuition about higher-dimensional Euclidean geometry turns out to be quite unreliable. The Euclidean volume V_d of a d-dimensional sphere of radius R_d is

$$V_d=\frac{(R_d)^d\pi^{d/2}}{\Gamma(1+d/2)},$$

with Γ the gamma function providing a natural generalization of the factorial function. With radius 1 this gives, for dimensions $d=1, 2, \cdots$, the volumes 2, 3.14, 4.18, 4.93, 5.26, 4.72, 4.06, $\cdots$. The volume of the unit radius sphere comes to a maximum for $d=5$ and falls off rather rapidly toward zero as d approaches infinity. On the other hand, d can be chosen to minimize the radius of a d-dimensional sphere of volume N. However, even with the optimal d (a function of N) the radius is $\Omega(\log^{1/2} N)$. Namely, setting $V_d = N$ and $d = 2k$, we have

$$N = \left(\frac{\pi^k}{k!}\right)(R_{2k})^{2k} = \frac{(\pi(R_{2k})^2)^k}{k!}.$$

By Stirling's approximation,

$$R_{2k} \sim \sqrt{\frac{k}{e\pi}(N\sqrt{2\pi k})^{1/k}}.$$

Observe that the lower bound in Theorem 2(i) is therefore $\Omega(N^{1/d} \cdot d^{1/2} D^{-1})$. Differentiating, we find that R_{2k} reaches its *minimum* $R_{2k}^{\min}$ for

$$k \sim \log N,$$

where log denotes the natural logarithm. Therefore, with $N^{1/\log N} = e$, and $(2\pi k)^{1/k} \downarrow 1$ for $k \to \infty$, we obtain

$$R_{2k}^{\min} \sim \sqrt{\frac{\log N}{\pi}}.$$

We may think that it is the unfortunate accident of having a physical space of only three dimensions that makes it hard to embed edge-symmetric graphs with small diameter. However, this is not the case. By this analysis and Theorem 2, to embed edge-symmetric graphs of diameter $o(\log^{1/2} N)$ requires the average length of an embedded edge to rise unbounded with N, independent of the number of dimensions. As another curiosity, the average edge length of the complete binary tree in $d > 1$ dimensions is not $O(1)$, but turns out to be $\Omega(d^{1/2})$. That is, in higher dimensions the H-tree construction increasingly loses efficiency.

A.2. Meshes in higher dimensions. Let $N = n^\delta$, n a positive integer. Define a δ-dimensional *mesh with wraparound* as a set of nodes $(i_1, \cdots, i_\delta)$, $i_j = 0, \cdots, N^{1/\delta} - 1$ $(1 \leqq j \leqq \delta)$. Node $(i_1, \cdots, i_\delta)$ is connected by an edge with node $(j_1, \cdots, j_\delta)$, if they are equal in all coordinates except one where they differ by 1 mod $N^{1/\delta}$.

Again assume that a node (processor) has unit volume in any number d of dimensions we care to consider. For d-dimensional embeddings of N-node, δ-dimensional meshes with wraparound we have an average interconnect length $\geqq (2^d - 1)R_d/(2^d \delta N^{1/\delta})$. This lower bound is a small positive constant for $d \geqq \delta$ and d is small (this is necessary because of the curious behavior of the ratio between volume and radius in higher dimensions). Since the lower bound can be matched by an upper bound, such meshes are feasible architectures for large N. However, since the average Euclidean interconnect length exceeds

$$\delta^{-1} N^{(\delta-d)/d\delta} \sqrt{\frac{d}{2\pi e}},$$

it rises unbounded with N for $\delta > d$. (It also rises unbounded with d for fixed N and δ.)

Acknowledgments. Remarks by Andries Brouwer, Evangelos Kranakis, F. Tom Leighton, Lambert Meertens, Yoram Moses, and the referees were helpful.

REFERENCES

[1] N. L. BIGGS AND A. T. WHITE, *Permutation Groups and Combinatorial Structures*, Cambridge University Press, Cambridge, 1979.

[2] F. HARARY, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.

[3] W. D. HILLIS, *The Connection Machine*, MIT Press, Cambridge, MA, 1985.

[4] L. A. HORNAK, S. K. TEWKSBURY, M. HATAMIAN, A. LIGTENBERG, B. SUGLA, AND P. FRANZON, *Through-wafer optical interconnects for multi-wafer wafer-scale integrated architectures*, Proc. SPIE Conference, San Diego, August, 1986.

[5] F. T. LEIGHTON AND A. ROSENBERG, *Three-dimensional circuit layouts*, Tech. Report TM-262, Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, MA, June, 1984.

[6] C. MEAD AND M. REM, *Minimum propagation delays in* VLSI, IEEE J. on Solid State Circuits, 17 (1982), pp. 773-775. (Correction: IEEE J. Solid State Circuits, 19 (1984), p. 162.)

[7] J. T. SCHWARTZ, *Ultracomputers*, ACM Trans. Programming Languages and Systems, 2 (1980), pp. 484-521.

[8] S. K. TEWKSBURY, L. A. HORNAK, AND A. LIGTENBERG, *The impact of component interconnects on future large scale systems*, Proc. IEEE, to appear.

[9] J. D. ULLMAN, *Computational Aspects of* VLSI, Computer Science Press, Rockville, MD, 1984.

[10] E. UPFAL AND A. WIGDERSON, *How to share memory in a distributed system*, J. Assoc. Comput. Mach., 34 (1987), pp. 116-127.

[11] P. M. B. VITÁNYI, *Non-sequential computation and laws of nature*, in VLSI Algorithms and Architectures, Lecture Notes in Computer Science 227, Springer-Verlag, New York, Berlin, 1986, pp. 108-120.

[12] ———, *Area penalty for sublinear signal propagation delay on chip*, Proc. 26th Annual IEEE Symposium on Foundations of Computer Science, 1985, pp. 197-207.