Nondeterministic Linear-Time Tasks May Require Substantially Nonlinear Deterministic Time in the Case of Sublinear Work Space

YURI GUREVICH

University of Michigan, Ann Arbor, Michigan

AND

SAHARON SHELAH

Hebrew University, Jerusalem, Israel, and Rutgers University, New Brunswick, New Jersey

Abstract. A technique is developed for establishing lower bounds on the computational complexity of certain natural problems. The results have the form of time–space trade-off and exhibit the power of nondeterminism. In particular, a form of the clique problem is defined, and it is proved that:

—a nondeterministic log-space Turing machine solves the problem in linear time, but

—no deterministic machine (in a very general use of this term) with sequential-access input tape and work space n^σ solves the problem in time $n^{1+\tau}$ if $\sigma + 2\tau < \frac{1}{2}$.

Categories and Subject Descriptors: F.1.1 [**Computation by Abstract Devices**]: Models of Computations—*automata*; *bounded-action devices*; F.1.3 [**Computation by Abstract Devices**]: Complexity Classes—*relations among complexity classes*; *relations among complexity measures*; F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*computations on discrete structures*; F.2.3 [**Analysis of Algorithms and Problem Complexity**]: Trade-offs among Complexity Measures; G.2.1 [**Discrete Mathematics**]: Combinatorics—*combinatorial algorithms*; *counting problems*; G.2.2 [**Discrete Mathematics**]: Graph Theory—*graph algorithms*; *cliques*

General Terms: Algorithms, Performance

Additional Key Words and Phrases: Clique problem, lower bounds, nondeterminism, nonlinear time, time–space trade-off

The work of Y. Gurevich was partially supported by the National Science Foundation grant DCR 85-03275. The work of both authors was partially supported by a grant from Binational US–Israel Science Foundation. In the main, the work was done during a week in Fall 1987 when both authors visited Rutgers University. During the revision (in 1988/89), the first author was with Stanford University and IBM Almaden Center (on a sabbatical leave from the University of Michigan).

An extended abstract of this paper was published in *Proceedings of the 20th ACM Symposium on the Theory of Computing* (Chicago, Ill., May 2–4). ACM, New York, 1988, pp. 281–289.

Authors' addresses: Y. Gurevich, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122; S. Shelah, Department of Mathematics, Hebrew University, Jerusalem 91904, Israel.

1. *Introduction*

One of the problems amenable to our lower-bound technique is:

LOG-SIZE CLIQUE PROBLEM:

Instance: A graph G on an initial segment $[0, \nu)$ of natural numbers given by a binary string of length $n = \nu(\nu - 1)/2$ representing the adjacency matrix of G.

Question: Is there a clique of size $\mathrm{lh}(n)$ in G?

Here and throughout the paper, $\mathrm{lh}(m)$ is the length of the binary notation for a natural number m, and a clique is simply a complete subgraph (not necessarily a maximal complete subgraph). We reserve the letter n to denote the input size. If the input string w represents the adjacency matrix of a graph, the number of vertices of the graph will be always denoted ν, so that $|w| = n = \nu(\nu - 1)/2$. The question remains how does w represent the adjacency matrix. Any reasonable form of representation will do, but it will be convenient to fix a particular one. View w as a binary function on $[0, n)$ and, given natural numbers $i < j$, define $\mathrm{Cd}(i, j) = \mathrm{Cd}(j, i) = j(j - 1)/2 + i$; the function Cd orders pairs first by the maximal member, and then by the minimal. Set $w(\mathrm{Cd}(u, v)) = 1$ (resp., $w(\mathrm{Cd}(u, v)) = 0$) if $\{u, v\}$ is (resp., is not) an edge.

Log-Size Clique Problem is not known to be decidable in polynomial time; we proceed to define an easier subproblem of it. Call a subset $\{v_1, \ldots, v_\lambda\} \subseteq [0, \nu)$ *parabolic* if there are integers a_0, a_1, a_2 such that every $v_i = a_0 + a_1 i + a_2 i^2$ modulo ν. If a parabolic set happens to be a clique in a graph G on $[0, \nu)$, then it is a *parabolic clique* of G.

LOG-SIZE PARABOLIC CLIQUE PROBLEM:

Instance: A graph G on an initial segment $[0, \nu)$ of natural numbers given by a binary string of length $n = \nu(\nu - 1)/2$ representing the adjacency matrix of G.

Question: Is there a parabolic clique of size $\mathrm{lh}(\nu)$ in G?

Log-Size Parabolic Clique Problem is especially easy for nondeterministic machines.

CLAIM 1.1. *Some log-space nondeterministic Turing machine M solves Log-Size Parabolic Clique Problem in time n + polylog(n). Moreover, M starts by guessing $\leq 4 \cdot lh(\nu)$ bits and then proceeds in a deterministic fashion, and the input head of M moves only to the right.*

Claim 1.1 will be proved in the appendix, but the idea is obvious: M guesses ν and parameters a_0, a_1, a_2 and then traverses the input to verify that there are 1's in all relevant places.

We turn now to our notion of deterministic machines. In this paper, a *generalized automaton* is a two-way multihead automaton with a clock for counting computation steps. The current positions of input heads are given by a mapping Heads from $[0, \eta)$ to $[0, n)$ where η is the number of input heads. The composition of Heads and the input gives the *currently scanned string.* A *configuration* of a generalized automaton comprises a state and a particular mapping Head. The next configuration is a function of the current configuration, the currently scanned string and the reading of the clock. A *deterministic machine M* with work space $S(n)$ is an infinite sequence (with no uniformity requirement) of generalized automata M_n such that M_n takes inputs of length n, has $2^{S(n)}$ states and polylog many input heads.

THEOREM 1.1. *If $\sigma + 2\tau < \frac{1}{2}$, then no deterministic machine with work space n^σ solves Log-Size Parabolic Clique Problem in time $n^{1+\tau}$.*

One consequence is that no deterministic log-space Turing machine can solve Log-Size Parabolic Clique Problem in time $n^{1+\tau}$ if $\tau < \frac{1}{4}$.

Theorem 1.1 will be proved in Sections 2–5. Here is a sketch of the proof. We suppose, by contradiction, that a deterministic machine M with work space n^σ solves Log-Size Parabolic Clique Problem in time $n^{1+\tau}$ and analyze the run of M on an input w. Segments $S \subseteq [0, n)$ of medium (in some technical sense) length are divided into moderate and superactive in a way that ensures a moderate majority. A moderate segment S is called flexible at a cell $c \in S$ if the restriction $y = w \mid S$ of w to S can be altered to some y' in such a way that $y'(c) \neq y(c)$ and the machine does not notice the alteration. The main result of Section 2 is that, with probability $1 - o(1)$, all moderate segments of a random input are flexible at most of their cells. In addition, we check in Section 2 that, with probability $1 - o(1)$, a random graph with ν vertices has no parabolic cliques of size $\mathrm{lh}(\nu)$.

In Section 3, the interval $[0, n)$ is sliced into standard segments called wards. Two wards are called independent if they never host input heads at the same time. In Section 4, we prove that independent wards can be altered simultaneously without changing the result of the computation.

In Section 5, we fix a sufficiently random graph G with sufficiently large number ν of vertices, so that, in particular, G has no parabolic cliques of size $\mathrm{lh}(\nu)$ and therefore M does not accept the presentation w of G. For technical reasons, it is convenient to assume that ν is prime. Then, we introduce a new probability space: Sample points are parabolic subsets X of $[0, \nu)$ and the probability distribution is uniform. We check that, with great probability, cells $c = \mathrm{Cd}(u, v)$, where u, v are different vertices in X, belong to different wards $W(c)$, and, for each c, the ward $W(c)$ is flexible at c, and the wards $W(c)$ are independent. Choosing a sufficiently random X, we alter the input in the relevant wards $W(c)$ making sure that X is a clique in the altered graph. M_n fails to notice the alteration.

By the way, the proof of Theorem 1.1 gives a little more than is stated in the theorem. Namely, let $\eta(n)$ be any polylog function of n. There exists m such that for no prime $\nu \geq m$, is there a generalized automaton with $\eta(n)$ input heads and work space n^σ which solves Log-Size Parabolic Clique problem for graphs with ν vertices in time $n^{1+\tau}$. The primality restriction can be removed.

THEOREM 1.2. *If $\sigma + 2\tau < \frac{1}{2}$, then no deterministic machine with work space n^σ solves Log-Size Clique Problem in time $n^{1+\tau}$.*

The proof of Theorem 1.2 is similar to that of Theorem 1.1; the necessary changes are given in Section 6. In particular, we use there the known fact [2, 7] that, with probability $1 - o(1)$, a random (with respect to the uniform distribution) graph G_ν with ν vertices has no cliques of size $\mathrm{lh}(n)$. It is however probable that G_ν has cliques of size $\mathrm{lh}(\nu)$. This is way we speak about cliques of size $\mathrm{lh}(n)$ (rather than $\mathrm{lh}(\nu)$) in the definition of Log-Size Clique Problem. One may work with nonuniform probability distributions (to strengthen Theorem 1.2 for example); computations become messier though.

One obvious generalization of Theorems 1.1 and 1.2 is obtained by noticing that cliques may be replaced with subgraphs of a different form, for example, circles (where every element has exactly two neighbors). A less obvious generalization (requiring some work) is obtained by allowing multidimensional Euclidean input tapes. In the new situation, wards will be cubes of appropriate dimension. It is important that the border of a cube is much less than its interior. But the topology of tape matters. It is easy to see that Theorem 1.1 fails for input tapes in the tree form even though subtrees have one-point borders. (For each triple (a_1, a_2, a_3) in turn, the desired algorithm checks whether the corresponding parabolic set is a clique.) It is not enough that wards have relatively small borders. It is also important that a great majority of them have moderate borders.

We may allow writing on the input tape (say, two-dimensional input tape) if there is only one reading head. Define a *2D automaton* as a finite automaton with a clock, a two-dimensional tape and a read–write head. Depending on the current state, the currently observed symbol and the reading of the clock, the automaton prints a new symbol in the observed cell, goes to a new state, moves the head to one of the four neighboring positions and advances the clock by 1. Initially, cells $(0, 0), \ldots, (n - 1, 0)$ hold the input; the input size is fixed for a given 2D automaton. Define a *2D machine* with work space $S(n)$ as a sequence (with no uniformity requirement) of 2D automata M_n such that M_n takes inputs of length n and has $2^{S(n)}$ states. For simplicity, we suppose that the alphabet of tape symbols does not depend on n.

THEOREM 1.3. *If $\sigma + 2\tau < \frac{1}{2}$, then no 2D machine with work space n^σ solves Log-Size Clique Problem (or Log-Size Parabolic Clique Problem) in time $n^{1+\tau}$.*

Theorem 1.3 is proved in Section 7.

Janos Simon (private communication) asked whether our trade-off results can be generalized to probabilistic acceptors. This seems to be a good research problem.

Now let us briefly discuss related results in the literature. Kannan [5] proves that there exists a universal constant k such that for all "nice" time bounds $t(n)$, the class of languages that can be accepted simultaneously in deterministic time $O(t(n))$ and space $o((t(n))^{1/k})$ is strictly contained in NTIME$(t(n))$. Thus, in the case of restricted space, deterministic machines may require more time. We have shown that they may require substantially more time and that the gap may contain natural problems. The methods (of Kannan and ours) are different: Our method is purely combinatorial, Kannan relies heavily on diagonalization.

Duris and Galil [3] have found a simple language whose time and space complexities T and S (on Turing machines) satisfy a condition $T^2S = \Omega(n^3)$. Nondeterminism is of no help in their case: Even nondeterministic Turing machines that decide the Duris–Galil language satisfy $T^2S = \Omega(n^3)$.

A series of time–space trade-offs has been proved for comparison-based branching programs. The latest paper in the series is that of Yao [8] who mentions the previous papers in the series as well as other papers on time–space trade-offs. The comparison-based model allows random access to input but is restricted in the sense that the basic operation is comparison. Proving lower bounds, one has to deal with a possibility that, instead of behaving rationally, the machine does some black magic and then comes up with a correct result. This poses a greater problem in the case of Turing machines.

Paul et al. have proved that, for Turing machines with several linear tapes, nondeterministic linear-time tasks may require nonlinear deterministic time [6].

Our result is somewhat similar but the work space is restricted. On the positive side, the nonlinearity of deterministic time is more substantial in our case, nondeterministic machines are more restricted and deterministic machines are more general.

Grandjean [4] exhibited natural NP-complete problems which are not in DTIME(n), but are solvable in linear time by alternating Turing machines using only one alternation.

Beame [1] showed that any CRCW PRAM that recognizes k-cliques in ν-node graphs in time T requires $\nu^{\Omega(k/T^2)}$ processors independent of its memory size.

2. *Moderate Segments of Random Inputs*

Let M_n be a generalized automaton with n-bit inputs, 2^{n^σ} states and time bound $n^{1+\tau}$. Let $\eta(n)$ be the number of inputs heads of M_n. Natural numbers $< n$ will be called *cells*, nonempty intervals of cells will be called *segments*, and natural numbers $\leq n^{1+\tau}$ will be called *moments*. If w is an input and t is a moment, let $\rho_w(t)$ be the configuration number t in the run of M_n on input w. (If M_n halts at some moment $t' < t$, then $\rho_w(t) = \rho_w(t')$.)

A cell c is *active* at moment t (with respect to a given input) if at least one head resides in c at moment t. The number of heads residing in c at moment t is the *activity* $A(c, t)$ of c at t. The *total activity* of c is $A(c) = \sum_t A(c, t)$. A set S of cells is *active* at a moment t if at least one cell $c \in S$ is active at t. The *activity* $A(S, t)$ of S at t is the sum of the activities of the cells of S. The *total activity* of S is $A(S) = \sum_t A(S, t)$; the *average activity* of S is $A(S)/|S|$. Any maximum time-interval I such that S is active at all $t \in I$ will be called a *session* for S.

Recall that $\sigma + 2\tau < \frac{1}{2}$. Choose real numbers $\alpha, \beta, \gamma, \delta$ such that:

- $\tau < \alpha$ and $\sigma + 2\alpha < \frac{1}{2}$.
- $\alpha + \sigma < \beta < \gamma < \frac{1}{2} - \alpha$.
- δ is less than $\alpha - \tau$, $\beta - \alpha - \sigma$, $\gamma - \beta$, $\frac{1}{2} - \alpha - \gamma$.

Call a segment S (of cells) *short*, *medium*, or *long* if $|S| \leq n^\gamma$, $n^\gamma < |S| \leq 2n^\gamma$ or $|S| > 2n^\gamma$, respectively. Call a segment *moderate* if it is medium, has $\leq n^\alpha$ sessions and its average activity $\leq n^\alpha$. A segment with $> n^\alpha$ sessions or with average activity $> n^\alpha$ will be called *S superactive*.

Functions with values in $\{0, 1\}$ will be called *binary*. Recall that we view input as a binary function. If x and y are binary functions with disjoint domains, let $x \cup y$ be the extension of x and y to the union of their domains.

Definition. Let $S = [c, d]$ be a medium segment, x be a binary function on the complement $\bar{S}$ of segment S, and y, z be binary functions on S such that S is moderate with respect to both inputs $u = x \cup y$ and $v = x \cup z$. The functions y, z are *x-equivalent* if

- $y(c) = z(c)$ and $y(d) = z(d)$,
- the sessions of S with respect to input u are exactly the sessions of S with respect to input v, and
- if $I = [s, t]$ is any of those sessions, then, $\rho_u(t) = \rho_v(t)$.

It follows that, if $I = [s, t]$ is one of the sessions for S with respect to u or v, then $\rho_u(s) = \rho_v(s)$. It follows also that if M_n accepts $x \cup y$ then it accepts $x \cup z$. The following lemma is crucial from the point of view of possible generalizations of Theorems 1.1 and 1.2. In that lemma and later, $\exp(m) = 2^m$.

LEMMA 2.1. *Let x be a binary function on the complement of a medium segment S. The number of x-equivalence classes is $\exp[o(n^\beta)]$.*

PROOF. Let y range over binary functions on S such that S is moderate with respect to $x \cup y$. The x-equivalence class of any y is determined by the values of y at the end-points of S, by the sessions of S with respect to $x \cup y$ and by the configurations of M_n at the end of each session.

There are only four possibilities for the values of y at the end-points of S.

The total number of possible collections of sessions is $\exp[o(n^\beta)]$. For, the number of sessions is at most n^α and, for each number k of sessions, the number of possible collections of sessions is bounded by

$$(n^{1+\tau})^{2k} \leq \exp[2n^\alpha \cdot (1 + \tau)\log n] = \exp[o(n^\beta)].$$

Choose $\sigma' > \sigma$ such that $\alpha + \sigma' < \beta$ and consider sufficiently large n. The number N of configurations of M_n equals the number $\exp[n^\sigma]$ of states times the number n^η of possible mappings Heads. Hence, $N < \exp[n^{\sigma'}]$. Therefore, the number of functions that assign configurations of M_n to (the final points of) any collection of $\leq n^\alpha$ disjoint time intervals is bounded by

$$N^{n^\alpha} < \exp[n^{\sigma'+\alpha}] = \exp[o(n^\beta)].$$

This finishes the proof of the lemma. □

Notice the use of inequality $\alpha + \sigma < \beta$.

Definition. Let S be a moderate segment (with respect to a given input). Let x be the restriction of the input to $\bar{S}$, and y be the restriction of the input to S. The segment S is *flexible at* a cell $c \in S$ if the x-equivalence class of y contains some z with $z(c) \neq y(c)$; otherwise S is *rigid at c*. Further, S itself is *flexible* (resp., *rigid*) if it is rigid at $<n^\beta$ (resp., $\geq n^\beta$ cells).

We interrupt the main flow of this section to recall an easy and well-known fact about probability spaces that will be used in Theorem 2.1 and later.

CLAIM 2.1. *Let E and $H_1, \ldots, H_k$ be events in an arbitrary probability space. If the events H_i are pairwise disjoint and cover E, then $Pr[E] \leq max_i Pr[E \mid H_i]$.*

THEOREM 2.1. *Consider a random (with respect to the uniform probability distribution) input. The probability that at least one moderate segment is rigid is $o(1)$.*

PROOF. First, consider a fixed medium segment S, a fixed binary function x on $\bar{S}$ and an auxiliary probability space where sample points are binary functions y on S and the probability distribution is uniform. Let

$$\text{Bad}(S, x) = \{y : S \text{ is moderate and rigid with respect to input } x \cup y\},$$

and, for every $y \in \text{Bad}(S, x)$, let $C(y)$ be the x-equivalence class of y. The rigidity requirement implies that

$$|C(y)| \leq \exp(|S| - n^\beta).$$

By Lemma 2.1,

$$|\{C(y) : S \text{ is moderate with respect to } x \cup y\}| = \exp[o(n^\beta)].$$

Hence, $|\text{Bad}(S, x)| \leq \exp[|S| - n^\beta + o(n^\beta)]$ and therefore

$$\Pr[\text{Bad}(S, x)] \leq \exp[-n^\beta + o(n^\beta)].$$

Second, consider the probability space indicated in the theorem: Sample points are binary functions on $[0, n)$ and the probability distribution is uniform. For each medium segment S, let

$$\mathrm{Bad}(S) = \{w : S \text{ is moderate and rigid with respect to } w\}.$$

Let x range over binary functions on $\bar{S}$. We have shown that there exists a uniform bound $\exp[-n^{\beta} + o(n^{\beta})]$ on the conditional probabilities $\Pr[\mathrm{Bad}(S) \mid w \mid \bar{S} = x]$. By Claim 2.1, $\Pr[\mathrm{Bad}(S)] \leq exp[-n^{\beta} + o(n^{\beta})]$. Since there are at most $n^{1+\gamma}$ medium segments,

$$\Pr[(\exists S)\mathrm{Bad}(S)] \leq n^{1+\gamma} \cdot \exp[-n^{\beta} + o(n^{\beta})] = o(1). \qquad \square$$

PROPOSITION 2.1. *Let G_ν be a random (with respect to the uniform probability distribution) graph with ν vertices. The probability that G_ν has a parabolic clique of size $lh(\nu)$ converges to* 0 *when ν grows to infinity.*

PROOF. Let a_0, a_1, a_2 range over $[0, \nu)$, and $X(a_0, a_1, a_2)$ be the set of vertices $a_0 + a_1 i + a_2 i^2 \mathrm{mod}\ \nu$ where $1 \leq i \leq \lambda = \mathrm{lh}(\nu)$. The probability that $X(a_0, a_1, a_2)$ forms a clique is $2^{-\lambda(\lambda-1)/2} \leq \nu^{-(\lambda-1)/2}$. There are ν^3 different sets $X(a_0, a_1, a_2)$. Hence, the probability that at least one of these sets forms a clique is at most $\nu^3\nu^{-(\lambda-1)/2} = o(1)$. $\square$

3. *Moderate Wards of Input*

Call a cell c *moderate* if its total activity is bounded by $n^{\alpha}/2$; otherwise, call c superactive. A segment with moderate end-points has at most n^{α} sessions.

Definition. Partition the interval $[0, n)$ of natural numbers into segments $[c_i, c_{i+1})$ such that $c_0 = 0$ and, if $c_i < n$, then either c_{i+1} is the minimal number in the set

$$\{c : c > c_i + n^{\gamma} \text{ and } c \text{ is moderate}\}$$

or else this set is empty and $c_{i+1} = n$. If is easy to see that there is only one such partition. The segments $[c_i, c_{i+1})$ will be called *wards*. The ward that contains a cell c will be denoted $W(c)$.

Call a cell c *flexible* (with respect to a given input) if $W(c)$ is flexible at c; otherwise, call c *rigid*.

THEOREM 3.1. *Suppose that every moderate ward is flexible. Then the total number of rigid cells is $o(n^{1-\delta})$.*

PROOF. The union S of superactive wards contains $o(n^{1-\delta})$ cells. For, let W be a superactive ward. By the definition of superactivity, either W has $>n^{\alpha}$ sessions or else its total activity $A(W)$ exceeds $|W| \cdot n^{\alpha}$. The first alternative is impossible because the end-points of W are moderate. Hence, $A(W) > |W| \cdot n^{\alpha}$. Hence, $A(S) > |S| \cdot n^{\alpha}$. But $A(S) \leq A[0, n) = \eta n^{1+\tau}$. Thus, $|S| < \eta n^{1+\tau-\alpha} = o(n^{1-\delta})$.

A similar argument shows that the total number of superactive cells is $o(n^{1-\delta})$. This fact implies that the union of all long wards is $o(n^{1-\delta})$ because if W is a long ward, then, by the definition of wards, $A(c) > n^{\alpha}/2$ for at least one half of cells $c \in W$. There is at most one short ward; it contains at most $n^{\gamma} = o(n^{1-\delta})$ cells. ($\delta < \frac{1}{2} - \alpha - \gamma < 1 - \gamma$ and therefore $\gamma < 1 - \delta$.) It remains to prove that all moderate wards together contain $o(n^{1-\delta})$ rigid cells.

Since each moderate ward contains more than n^{γ} cells, there are at most $n^{1-\gamma}$ moderate wards. By the assumption, each moderate ward is flexible, that is, contains less than n^{β} rigid cells. Then all moderate wards together contain less than $n^{1+\beta-\gamma} = o(n^{1-\delta})$ rigid cells. □

Notice the explicit use of inequalities $\delta < \alpha - \tau$, $\delta < \gamma - \beta$ and $\delta < \frac{1}{2} - \alpha - \gamma$ and the implicit use of inequalities $\tau < \alpha$ and $\beta < \gamma$.

4. *Independent Wards of Input*

Definition. Let $I = [t, t + l]$ be a session (of length $l + 1$) for a ward W and c be the position of a head h at moment t (all this with respect to a given input). The *potentially active zone* $Z(W, I, h)$ of the triple (W, I, h) is the segment $(c - l, c + l)$ of length $2l + 1 = 2|I| - 1$; it is easy to see that h cannot leave this segment during the session I. Further,

$$Z(W, I) = \bigcup_{h} Z(W, I, h),$$

$$Z(W, h) = \bigcup_{I} Z(W, I, h),$$

$$Z(W) = \bigcup_{I} Z(W, I) = \bigcup_{h} Z(W, h).$$

If there is a need to show the input explicitly, we write Z_w instead of Z.

LEMMA 4.1. *If W is a moderate ward, then $Z(W)$ intersects $O(\eta n^{\alpha})$ moderate wards.*

PROOF. We prove that each $Z(W, h)$ intersects $\leq 6n^{\alpha}$ wards. Since each zone $Z(W, I, h)$ is of length $< 2|I|$ and moderate wards are at least $n^{\gamma} + 1$ long, each $Z(W, I, h)$ intersects at most $2|I|/n^{\gamma} + 2$ moderate wards. Since W is moderate, it has at most n^{α} sessions and the total length of all sessions of W is at most $A(W) \leq |W| \cdot n^{\alpha} \leq 2n^{\gamma}n^{\alpha}$. Hence, $Z(W, h)$ intersects at most

$$\sum_{I} \left(\frac{2|I|}{n^{\gamma}} + 2 \right) \leq \frac{2A(W)}{n^{\gamma}} + 2n^{\alpha} \leq 6n^{\alpha}$$

moderate wards. □

LEMMA 4.2. *Let $I = [s, t]$ be a session for a moderate ward S with respect to an input w. Let $x = w \mid \bar{S}$, $y = w \mid S$, x' be a binary function on $\bar{S}$, which coincides with x on $Z_w(S, I)$, y' be a binary function on S, which is x-equivalent to y, and $v = x' \cup y'$. If $\rho_v(s) = \rho_w(s)$, then $\rho_v(t) = \rho_w(t)$.*

PROOF. By the definition of x-equivalence, M_n will not notice if we substitute y' for y without altering x. But x' is identical with x at every cell that can be possibly examined during the session I. Therefore, if M_n does not distinguish between w and v at moment s, it will not distinguish between them at moment t. A more formal version of this argument follows.

Let $u = x \cup y'$. Since y and y' are x-equivalent, y' coincides with y on the endpoints of S, S is moderate with respect to u, I is a u-session for S, $\rho_u(s) = \rho_w(s)$ and $\rho_u(t) = \rho_w(t)$. Suppose that $\rho_v(s) = \rho_w(s)$. Then, all three configurations $\rho_u(s)$, $\rho_v(s)$, and $\rho_w(s)$ coincide. Notice that $\rho_w(s)$ uniquely defines $Z_w(S, I)$, and the same holds for u and v. Thus, $Z_u(S, I) = Z_v(S, I) = Z_w(S, I)$, and therefore x' coincides with x on $Z_u(S, I)$. Recall that the next configuration of the machine is completely defined by the current configuration, the currently scanned string and the current reading of the clock. By obvious induction, $\rho_v(t') = \rho_u(t')$ for all t' in I. Hence, $\rho_v(t) = \rho_u(t) = \rho_w(t)$. □

Two distinct wards U and V (with respect to the same input) will be called *independent* if V is disjoint from $Z(U)$, and U is disjoint from $Z(V)$. Three or more wards (with respect to the same input) are *independent* if every two of them are.

THEOREM 4.1. *Suppose that $W_0, \ldots, W_l$ are moderate independent wards with respect to an input u, and $x_j = u \mid \bar{W}_j$, $y_j = u \mid W_j$. Let v be obtained from u by simultaneous replacement of each y_j with a binary function z_j on W_j which is x_j-equivalent to y_j. Then M_n accepts v if and only if it accepts u.*

PROOF. Since the u-wards W_j are independent, their sessions are disjoint. List all sessions $[c_0, d_0], \ldots, [c_k, d_k]$ for the wards $W_0, \ldots, W_l$ in the natural order (so that $d_i < c_{i+1}$). Recall that the next configuration of the machine is completely defined by the current configuration, the currently scanned string and the current reading of the clock. It follows that:

- $\rho_u(t) = \rho_v(t)$ for all $t < c_0$,
- if $\rho_u(d_i) = \rho_v(d_i)$, then $\rho_u(c_{i+1}) = \rho_v(c_{i+1})$ for all $i < k$, and
- if $\rho_u(d_k) = \rho_v(d_k)$, then $\rho_u(t) = \rho_v(t)$ for every moment $t > d_k$.

It remains to prove that, for all $i \leq k$,

- if $\rho_u(c_i) = \rho_v(c_i)$, then $\rho_u(d_i) = \rho_v(d_i)$.

Without loss of generality, $[c_i, d_i]$ is a session for W_0. Use Lemma 4.2 with $w = u$, x' being the result of the simultaneous replacement of $y_1, \ldots, y_l$ by $z_1, \ldots, z_l$ in u, and $y' = z_0$. □

5. *Random Parabolic Subgraphs*

For technical reasons, it will be convenient to suppose that the number ν of vertices of the given graph is prime. Let w be an n-bit string. Call a pair (ν, w) *appropriate* if ν is prime and sufficiently large and w is sufficiently random, so that with respect to Theorem 2.1 and Proposition 2.1 we have that:

- every w-moderate segment has less than n^β rigid cells, and
- the graph G represented by w has no parabolic cliques of size $\mathrm{lh}(\nu)$.

Given an appropriate pair (ν, w), consider a new probability space where sample points are triples of natural numbers $< \nu$, and the probability distribution is uniform. Abbreviate $\mathrm{lh}(\nu)$ to λ.

Let (a_0, a_1, a_2) be a random sample point. For each positive integer $i \leq \lambda$, let $v_i = v_i(a_0, a_1, a_2)$ be the vertex such that

$$v_i = a_0 + a_1 i + a_2 i^2 \mathrm{mod}\ \nu.$$

The binary function Cd was defined in Section 1. If $v_i \neq v_j$, let $c_{i,j} = \mathrm{Cd}(v_i, v_j)$ and $W_{i,j} = W(c_{i,j})$; if $v_i = v_j$, then $c_{i,j}$ and $W_{i,j}$ are undefined.

If the probability $\Pr[E]$ of an event E is $o(\nu^{-\delta})$, we say that E (as well as $\Pr[E]$) is *negligible* and the complement $\bar{E}$ of E (as well as $\Pr[\bar{E}]$) is *almost sure.* We intend to prove that the event [The wards $W_{i,j}$ are (defined and) independent] is almost sure.

LEMMA 5.1

(1) $\Pr[v_i = v_j] = 1/\nu$ *for all* $i \neq j$.
(2) $\Pr[v_i = v] = 1/\nu$ *for all* i *and all* $v < \nu$.
(3) $\Pr[v_i = u$ *and* $v_j = v] = 1/\nu^2$ *for all* $i \neq j$ *and all* u, v.
(4) $\Pr[c_{i,j} = c] = 2/\nu^2$ *for all* $i \neq j$ *and all cells* c.
(5) $\Pr[v_j = v \mid v_i = u] = 1/\nu$ *for all* $i \neq j$ *and* $u \neq v$.
(6) *Every event* $[c_{i,j}$ *is flexible*$]$ *is almost sure.*

PROOF

(1) The total number of sample points is ν^3, and there are exactly ν^2 sample points that solve the equation

$$a_0 + a_1 i + a_2 i^2 = a_0 + a_1 j + a_2 j^2 \text{mod } \nu.$$

(2) There are exactly ν^2 sample points that solve the equation

$$a_0 + a_1 i + a_2 i^2 = v \text{ mod } \nu.$$

(3) There are exactly ν sample points solving the system

$$a_0 + a_1 i + a_2 i^2 = u \text{ mod } \nu \qquad \text{and} \qquad a_0 + a_1 j + a_2 j^2 = v \text{ mod } \nu.$$

(4) Use the third assertion of this lemma.

(5) $\Pr[v_j = v \mid v_i = u] = \Pr[v_i = u \text{ and } v_j = v]/\Pr[v_i = u] = \nu^{-2}/\nu^{-1}$.

(6) We prove that the event $[c_{i,j}$ undefined or rigid$]$ is negligible. By (1), the event $[c_{i,j}$ is undefined$]$ is negligible. By Theorem 3.1, the total number of rigid cells is $o(\nu^{2-2\delta})$. By (4), the probability that $c_{i,j}$ is (defined and) rigid is $o(\nu^{-2\delta})$. □

LEMMA 5.2. *For all distinct* i, j, k *and* l, *the probability that* $c_{i,j}$ *and* $c_{k,l}$ *are flexible and* $Z(W_{i,j})$ *intersects* $W_{k,l}$ *is negligible.*

PROOF. Without loss of generality, we may restrict attention to the case $i = 1$, $j = 2$, $k = 3$, and $l = 4$. Choose an event $E = [v_1 = u$ and $v_2 = v]$ such that the cell $c = \text{Cd}(u, v)$ is flexible and the conditional probability

$$\Pr[c_{3,4} \text{ is flexible and } Z(W_{1,2}) \text{ intersects } W_{3,4} \mid E]$$

is maximal possible. By Claim 2.1, it suffices to prove that this conditional probability is negligible. Recall that flexible cells belong to moderate wards. Let U be the union of all moderate wards intersected by $Z(W(c))$. It suffices to prove that the conditional probability $\Pr[c_{3,4} \in U \mid E]$ is negligible.

By Lemma 4.1, U contains $O(\eta n^{\alpha})$ moderate wards. Since each moderate ward contains at most $2n^{\gamma}$ cells, U contains

$$O(\eta n^{\alpha+\gamma}) = o(n^{1/2-\delta})$$

cells. Since any event $[c_{3,4} = d$ and $E]$ contains <2 sample point, the event $[c_{3,4} \in U$ and $E]$ contains $o(n^{1/2-\delta})$ sample points. Hence

$$\Pr[c_{3,4} \in U \mid E] = \frac{\Pr[c_{3,4} \in U \text{ and } E]}{\Pr[E]} = o(\nu^{-\delta}). \qquad \square$$

Notice the explicit use of inequality $\delta < \frac{1}{2} - \alpha - \gamma$ and the implicit use of inequality $\gamma < \frac{1}{2} - \alpha$.

Call a vertex u *bad* if there are $\geq \nu^{1-\delta}$ vertices v such that $\text{Cd}(u, v)$ is rigid; otherwise, call u *good*.

LEMMA 5.3. *Every event* $[v_i$ *is bad*$]$ *is negligible.*

PROOF. Let k be the number of bad vertices. By Lemma 5.1(1), the probability that v_i is bad equals k/ν. Hence, it suffices to prove that $k = o(\nu^{1-\delta})$.

For every bad u, there are at least $\nu^{1-\delta}$ rigid cells of the form $\mathrm{Cd}(u, v)$. Hence, there are at least $k\nu^{1-\delta}/2$ rigid cells $\mathrm{Cd}(u, v)$ where u or v is bad. By Theorem 3.1, the total number of rigid cells is $o(\nu^{2-2\delta})$. Hence, $k = o(\nu^{1-\delta})$. □

LEMMA 5.4. *For all distinct* i, j, k, *the probability that* $c_{i,j}$ *and* $c_{i,k}$ *are flexible and* $Z(W_{i,j})$ *intersects* $W_{i,k}$ *is negligible.*

PROOF. Without loss of generality, we may restrict attention to the case $i = 1$, $j = 2$, $k = 3$. By the previous lemma, it suffices to prove that the event

$$[v_1 \text{ is good, and } c_{1,2}, c_{1,3} \text{ are flexible, and } Z(W_{1,2}) \text{ intersects } W_{1,3}]$$

is negligible. Choose a good vertex u such that the conditional probability

$$\Pr[c_{1,2}, c_{1,3} \text{ are flexible, and } Z(W_{1,2}) \text{ intersects } W_{1,3} \mid v_1 = u]$$

is maximal possible. By Claim 2.1, it suffices to prove that this conditional probability is negligible. Let F_u be the set of vertices v such that $v \neq u$ and the cell $\mathrm{Cd}(u, v)$ is flexible. Choose a vertex $v \in F_u$ such that the conditional probability

$$\Pr[c_{1,2}, c_{1,3} \text{ are flexible, and } Z(W_{1,2}) \text{ intersects } W_{1,3} \mid v_1 = u \text{ and } v_2 = v]$$

is maximal possible. Let $c = \mathrm{Cd}(u, v)$, $Z' = Z(W(c)$, U be the union of all moderate wards intersected by Z', and E be the event $[v_1 = u$ and $v_2 = v]$. It suffices to prove that $\Pr[c_{1,3} \in U \mid E]$ is negligible.

As in the proof of Lemma 5.2, $|U| = o(n^{1/2-\delta})$. It is easy to check that different sample points of E give different values to v_3 and therefore to $c_{1,3}$. Hence, $\Pr[E$ and $c_{1,3} \in U] \leq \Pr[U]$. Thus,

$$\Pr[c_{1,3} \in U \mid E] \leq \frac{\Pr[U]}{\Pr[E]} = \frac{|U|}{|E|} = o(\nu^{-\delta}). \qquad \square$$

THEOREM 5.1. *The event* [*All vertices* v_i *are different, and all cells* $c_{i,j}$ *are flexible, and the wards* $W_{i,j}$ *are independent*] *is almost sure.*

PROOF. Use Lemmas 5.2 and 5.4. □

Finally, we are ready to finish the proof of Theorem 1.1. Let (ν, w) be an appropriate pair. By contradiction, suppose that an automaton M_n solves the case of Log-Size Parabolic Clique Problem with inputs of length n. Since the graph represented by w has no parabolic cliques of size $\mathrm{lh}(\nu)$, M_n does not accept w. By virtue of Theorem 5.1, we can choose a parabolic subset $v_1, \ldots, v_\lambda$ of different vertices such that all cells $c_{i,j}$ are flexible and the wards $W_{i,j}$ are independent. For every $W_{i,j}$, let $y_{i,j}$ (resp., $x_{i,j}$) be the restriction of the input w to $W_{i,j}$ (resp., to the complement of $W_{i,j}$), and let $z_{i,j}$ be a binary function on $W_{i,j}$ such that $z_{i,j}$ is $x_{i,j}$-equivalent to $y_{i,j}$ and $z_{i,j}(c_{i,j}) = 1$. Let w' be the result of simultaneous replacements of every $y_{i,j}$ by $z_{i,j}$. By Theorem 4.1, M_n does not accept w', but the graph represented by w' contains the parabolic clique $v_1, \ldots, v_\lambda$, which gives us the desired contradiction. □

6. *Log-Size Clique Problem*

This section is devoted to the proof of Theorem 1.2.

PROPOSITION 6.1. *Let G_v be a random (with respect to the uniform probability distribution) graph with v vertices. The probability that G_v has a clique of size $lh(n)$ converges to* 0 *when v grows to infinity.*

PROOF. See [2] or [9]. □

Choose a sufficiently large prime number v and a sufficiently random binary string w of length n, so that (with respect to Theorem 2.1 and Proposition 6.1):

- every w-moderate segment has less than n^β rigid cells, and
- the graph G represented by w has no cliques of size $\mathrm{lh}(n)$.

Let $l = \mathrm{lh}(n)$. Consider a new probability space where sample points are functions from $[1, l]$ to $[0, v)$, and the probability distribution is uniform. Let f be a random sample point. In order to make the notation closer to that of Section 5, abbreviate $f(i)$ to v_i. As in Section 5, define $c_{i,j} = \mathrm{Cd}(v_i, v_j)$ and $W_{i,j} = W(c_{i,j})$. Define negligible (resp., almost sure) events and probabilities as in Section 5.

LEMMA 6.1. *Lemma* 5.1 *remains true in the new setting.*

PROOF. The first five assertions are obvious. The proof of the last assertion is the same as in Section 5. □

The rest of the proof of Theorem 1.2 is the same as in the case of Theorem 1.1. □

7. *Proof of Theorem* 1.3

Suppose that M_n is a 2D automaton with n-bit inputs, exp (n^σ) states and a time bound $n^{1+\tau}$. Tape cells of M_n will be called *little squares* in this section. Only n little squares $(0, 0), \ldots, (n-1, 0)$ will be called cells. These n cells will be identified with numbers $0, \ldots, n-1$, respectively. (The goal is to be closer to the terminology in Sections 2–5.)

A *pattern* of a column L of the tape is a function from the little squares of L to tape symbols. A *configuration* of M_n comprises a state and a little square (the head position). Let an *extended configuration* of M_n comprise a configuration and a column pattern (the pattern of the column where the head is located).

Define segments, moments, activity, and sessions as in Section 2. Let $R_w(t)$ be the extended configuration of M_n at moment t in the run on input w.

Recall that $\sigma + 2\tau < \frac{1}{2}$. Choose reals $\alpha, \beta, \gamma, \delta$ such that

- $\tau < \alpha$ and $\sigma + 3\alpha < \frac{1}{2}$.
- $(2\alpha + \sigma) < \beta < \gamma < \frac{1}{2}$.
- δ is less than $\alpha - \tau, \beta - 2\alpha - \sigma, \gamma - \beta, \frac{1}{2} - \gamma$.

(These are not exactly the inequalities of Section 2.) Define short, medium, long, moderate, and superactive segments as in Section 2. Say that a column pattern is *moderate* if the number of nonblank symbols is bounded by n^α. An extended configuration is *moderate* if its column pattern is so.

Definition. Let $S = [c, d]$ be a medium segment, x be a binary function on the complement $\bar{S} = [0, n) - S$ of segment S, and y, z be binary functions on S such that S is moderate with respect to both inputs $u = x \cup y$ and $v = x \cup z$. The functions y, z are *x-equivalent* if

- the sessions of S with respect to input u are exactly the sessions of S with respect to input v, and
- if $I = [s, t]$ is any of those sessions, then $R_u(t) = R_v(t)$.

LEMMA 7.1. *Let $S = [c, d]$ be a medium segment and x be a binary function on $[0, n) - S$. The number of x-equivalence classes is exp$[o(n^\beta)]$.*

PROOF. The equivalence class of any relevant y is defined by the collection of sessions and the extended configurations of M_n at the final moments of the sessions; all those extended configurations are moderate. As in the proof of Lemma 2.1, the number of relevant collections of sessions is bounded by $\exp[o(n^\beta)]$. It suffices to show that, for each set J of $\leq n^\alpha$ moments, the number of functions from J to moderate extended configurations of M_n has a bound of the form $\exp[o(n^\beta)]$ which is independent of J.

Choose $\alpha' > \alpha$ and $\sigma' > \sigma$ such that $2\alpha' + \sigma' < \beta$ and consider sufficiently large n. M_n has $\exp(n^\sigma)$ states and polynomial number of possible positions of the head. Hence, the number of configurations of M_n is $<\exp[n^{\sigma'}]$. To simplify notation, we suppose that M_n has only three tape symbols: 0, 1 and the blank. Then a moderate column pattern is described by the set of $\leq n^\alpha$ zeroes in the column and the set of $\leq n^\alpha$ ones in the column. Hence, the number of possible moderate column patterns is bounded by

$$\exp[2n^\alpha(1 + \tau)\log n] < \exp[n^{\alpha'}]$$

and the number of moderate extended configurations is bounded by $\exp[n^{\sigma'+\alpha'}]$ and the number of functions from J to moderate extended configurations is bounded by

$$\exp[n^{\sigma'+\alpha'+\alpha}] = \exp[o(n^\beta)]. \qquad \square$$

The rest of the proof of Theorem 1.3 is similar (and simpler) than that of Theorem 1.1. We explain only how we can get away with weaker inequalities $\gamma < \frac{1}{2}$ and $\delta < \frac{1}{2} - \gamma$ (rather than $\gamma < \frac{1}{2} - \alpha$ and $\delta < \frac{1}{2} - \alpha - \gamma$). Since M_n has only one head, the notion of independent wards becomes trivial. Define $Z(W) = W$. Then $|U| \leq 2n^\gamma$ in the proofs of the analogs of Lemma 5.2 and Lemma 5.4, which gives us the desired $|U| = o(n^{1/2-\delta})$. $\square$

Appendix

PROOF OF CLAIM 1.1. We describe an accepting computation of the desired nondeterministic machine M. The machine starts by guessing ν and parameters a_0, a_1, a_2. The rest of the computation is deterministic. For each positive integer $i \leq \mathrm{lh}(\nu)$, let v_i be the vertex equal to $a_0 + a_1 i + a_2 i^2 \bmod \nu$.

Notice that the binary notation for the cell

$$F(d) = \min\{\mathrm{Cd}(v_i, v_j) : \mathrm{Cd}(v_i, v_j) \geq d\}$$

is computable from the binary notation for a given cell d in space $O(\mathrm{lh}(n))$ and in time that is polylog in n. The strategy of M is obvious. It computes $F(0)$ and walks to $F(0)$, then it computes $F(F(0) + 1)$ and walks to $F(F(0) + 1)$, and so on until it reaches $cd(v_{\mathrm{lh}(\nu-1)}, v_{\mathrm{lh}(\nu)})$. Let c be the current head position. All we need to show is that, having the binary notation for a cell $d \geq c$ on one of its work tapes, M is able to determine whether $c < d$ or $c = d$.

It is natural to use a counter for the binary notation for c. The difficulty is that too often the updates will require more than one step. (Recall that the whole computation should take n + polylog(n) time. To overcome the difficulty, we use a slightly unusual counter on a special auxiliary work tape T of length lh(n) + 1. The head h of T moves if and only if the input head does. From the leftmost position, h moves to the rightmost position, then back to leftmost position, then again to rightmost position, etc.; it takes 2lh(n) steps to make a full circle. When h is in the leftmost position, the binary notation for the whole number c/(2lh(n)) is written on T. As h makes one round, it adds 1 to the binary notation on the tape. □

ACKNOWLEDGMENT. We are thankful to the referee for constructive criticism.

REFERENCES

1. BEAME, P. Lower bounds for recognizing small cliques on CRCW PRAM's. *Disc. Appl. Math.*, to appear.
2. BOLLOBAS, B. *Random Graphs.* Academic Press, London, 1985.
3. DURIS, P., AND GALIL, Z. A time-space tradeoff for language recognition. *Math. Syst. Theory 17* (1984), 3–12.
4. GRANDJEAN, E. A natural NP-complete problem with a nontrivial lower bound. *SIAM J. Comput. 17* (1988), 786–809.
5. KANNAN, R. Towards separating nondeterministic time from deterministic time. In *Proceedings of the 22nd IEEE Conference on Foundations of Computer Science.* IEEE, New York, 1981, pp. 235–243.
6. PAUL, W. J., PIPPENGER, N., SZEMEREDI, E., AND TROTTER, W. T. On determinism versus non-determinism and related problems. In *Proceedings of the 24th IEEE Symposium on Foundation of Computer Science* (Tucson, Az., Nov.). IEEE, New York, 1983, pp. 429–438.
7. SPENCER, J. Ten lectures on the probabilistic method. CBMS-NSF Regional Conference Series in Applied Math., number 52, SIAM.
8. YAO, A. C. Near-optimal time-space tradeoff for element distinctness. In *Proceedings of the 29th Symposium on Foundations of Computer Science.* IEEE, Computer Society Press, Las Alamitos, Calif., 1988, pp. 91–97.

RECEIVED FEBRUARY 1988; REVISED JANUARY 1989 AND JUNE 1989; ACCEPTED AUGUST 1989