NONDETERMINISTIC LINEAR-TIME TASKS MAY REQUIRE SUBSTANTIALLY NONLINEAR DETERMINISTIC TIME IN THE CASE OF SUBLINEAR WORK SPACE[1]

Yuri Gurevich
Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109-2122

Saharon Shelah
Mathematics, Hebrew University, Jerusalem 91904, Israel
Mathematics, Rutgers University, New Brunswick, NJ 08903

Abstract. Log-size Parabolic Clique Problem is a version of Clique Problem solvable in linear time by a log-space nondeterministic Turing machine. However, no deterministic machine (in a very general sense of this term) with sequential-access read-only input tape and work space n^σ solves Log-size Parabolic Clique Problem within time $n^{1+\tau}$ if $\sigma + 2\tau < 1/2$.

§1 Main Theorem

Definition. A subset $\{v_1, ..., v_\lambda\}$ of an initial segment $\{0, ..., \nu-1\}$ of natural numbers is parabolic if there are integers a_0, a_1, a_2 with $a_0 + a_1 i + a_2 i^2 = v_i \mod \nu$ for all i.

Definition. Log-size Parabolic Clique Problem is the following version of Clique Problem:

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

Question: Is there a parabolic clique of size at least $\log_2\nu$ in the graph ?

Remark. How does w represent the adjacency matrix? It will be convenient to fix a specific form of representation. View w as a function from the initial segment $\{0, ..., n-1\}$ of natural numbers to $\{0,1\}$. Code a pair $\{u,v\}$ of distinct vertices by the number $Cd(u,v) = q(q-1)/2 + p$ where $p = \min\{u,v\}$ and $q = \max\{u,v\}$. (This corresponds to ordering pairs of distinct vertices first by the maximal member, and then by the minimal.) Set $w(Cd(u,v))$ = [If $\{u,v\}$ is an edge then 1 else 0].

Claim. There is a log-space nondeterministic Turing machine that solves Log-size Parabolic Clique Problem in linear time.

Proof. We describe an accepting computation of the desired nondeterministic machine. The machine computes $\lambda = \lceil \log_2\nu \rceil$ and guesses parameters a_0, a_1, a_2. For each positive integer $i \le \lambda$, let v_i be the vertex equal $a_0 + a_1 i + a_2 i^2 \mod \nu$. The machine initializes a variable

[1] The work of the first author was partially supported by NSF grant DCR 85-03275. The work of both authors was partially supported by a grant of Binational US-Israel Science Foundation. In the main, the work was done during a week in Fall 1987 when both authors visited Rutgers.

© 1988 ACM-0-89791-264-0/88/0005/0281 $1.50

s to $Cd(v_1,v_2)$, and the input head h_0 starts moving; s will always have the form $Cd(v_i,v_j)$. When h_0 reaches the cell number s, the machine checks that $w(s) = 1$. If $s = Cd(v_{\lambda-1},v_\lambda)$ then the machine halts and accepts; otherwise s gets the next value and h_0 resumes the motion.

There is only one difficulty. In order to be able to recognize the cell number s, we would like to compute in real time the binary notation for the current position c of h_0, but it takes more than a constant time to compute the binary notation for c+1 from that for c. Here is one way to overcome the difficulty. Let $l = \lceil \log_2 n \rceil$. An auxiliary tape T_1 has $l+1$ cells. A head h_1 of T_1 moves if and only if h_0 does. From the leftmost cell, h_1 moves to the rightmost cell, then back to leftmost cell, then again to rightmost cell, etc. It takes $2l$ moves of h_0 to move h_1 from the leftmost position back to leftmost positon. When h_1 is in the leftmost position, the binary notation for the whole number $c/(2l)$ is written on T_1. As h_1 makes one round, it adds 1 to the binary notation on the tape. Q.E.D.

Remark. Notice that the nondeterministic log-space machine

(1) guesses $\leq 3\lceil \log_2 \nu \rceil$ bits,

(2) moves the input-tape head only to the right,

(3) spends time n + polylog(n) on every accepting computation.

The input tape of the nondeterministic log-space machine is a sequential-access read-only tape. Turing machines with such input tape are called off-line. Notice that an off-line deterministic Turing machine T with a bound S(n) on the work space can be viewed as a very uniform sequence of finite Turing machine T_n with input tape of size n and work tape of size S(n).

Definition. In this paper, deterministic machines M are infinite sequences $\langle M_0, M_1, \ldots \rangle$, where each M_n is an extension of a finite machine by means of a built-in (unbounded) clock, which generalize off-line deterministic Turing machines in the following aspects:

(1) The number η of the input heads of M_n may grow with n. Somewhat arbitrarily, we suppose that η is a poly-log function of n. The current positions of input heads are described by a function Head: $\{0, \ldots, \eta-1\}$ --> $\{0, \ldots, n-1\}$. The composition of Head and the input w (viewed as a function from $\{0, \ldots, n-1\}$ to $\{0,1\}$) gives the currently scanned string.

(2) There is no restriction on the nature of the work space of M_n, but it can have only finite many states; the base-two logarithm of the number of states is the size of work space.

(3) A built-in clock counts the number of steps. A configuration of M_n comprises a state of the function Head and a state of the work space. The next configuration of M_n is an arbitrary function of the current configuration, the currently scanned string and the reading of the clock except that the input heads do not jump (they move as the heads of a multihead two-way finite automaton).

Remarks. The work space can be thought of as a set of cells with a read-write head in each cell. There is no uniformity requirement on the sequence $\langle M_0, M_1, \ldots \rangle$.

Main Theorem. Let σ and τ be positive reals such that $\sigma + 2\tau < 1/2$. No deterministic machine with work space $\leq n^\sigma$ solves Log-size Parabolic Clique Problem within time $n^{1+\tau}$.

Main Theorem will be proved in Sections 2 - 5. Similar results can be proved for some other NP problems with small witnesses.

Remarks. (1) Let ν range over positive integers, $n = \nu(\nu-1)/2$, $\sigma + 2\tau < 1/2$, X_n range over n-th constituents of (infinite) deterministic machines with work space $\leq n^\sigma$, and P_n be the restriction of Log-size Parabolic Clique Problem to graphs with ν vertices. Because of the possible nonuniformity of infinite deterministic machines, Main Theorem implies that, for some n, no X_n solves P_n within time $n^{1+\tau}$.

The proof gives a little more: For every sufficiently large prime ν, no X_n solves P_n within time $n^{1+\tau}$. It is not difficult to establish that there exists m such that, for every $n \geq m$, no X_n solves P_n within time $n^{1+\tau}$. Also, estimations of m can be given.

(2) The proof of Main Theorem also gives the following: There is no deterministic machine with work space n^σ which, given the representation of a graph G with ν vertices, outputs 0 within time $n^{1+\tau}$ if G has no clique of size $\geq 3 \cdot \log_2 \nu$, and outputs 1 within time $n^{1+\tau}$ if G has a parabolic clique of size $\geq 3 \cdot \log_2 \nu$. Why $3 \cdot \log_2 \nu$? Consider the uniform probability distribution on graphs with ν vertices. We use the fact that the probability that a random graph has a parabolic clique of size $\log_2 \nu$ converges to 0 when n grows to infinity; see Lemma 2.3 below. Instead we can use the fact [Bo, Theorem 4 in Chapter XI] that the probability that a random graph has any clique of size $3 \cdot \log_2 \nu$ converges to 0.

(3) Janos Simon [Si] noticed that Main Theorem may generalize to the case of reliable probabilistic acceptors. We are thankful to Janos for communicating to us his observation and hope to address the issue in the full paper.

Finally, we compare Main Theorem with related results in the literature. Duris and Galil [DG] have found a simple language whose time and space complexities T and S (on Turing machines) satisfy $T^2S = \Omega(n^3)$. However, nondeterminism is not of much help in their case: even nondeterministic Turing machines that decide the Duris-Galil language satisfy $T^2S = \Omega(n^3)$.

Another time-space tradeoff $T^2S = \Omega(n^3)$ has been proved by Borodin, Fich, Meyer auf der Heide, Upfal and Widgerson [BFMUW] for a different simple problem. Their 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, Pippinger, Szemeredi and Trotter have succeeded to prove that, for Turing machines with several linear tapes, nondeterministic linear tasks may require nonlinear deterministic time [PPST]. 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.

§2 First randomization

Suppose that Main Theorem is false and let $M = \langle M_0, M_1, \ldots \rangle$ be an infinite deterministic machine such that for all n,

(1) the work space of M_n is at most n^σ, and

(2) M_n solves each instance of size n of Log-size Parabolic Clique Problem within time $n^{1+\tau}$.

Let w an input of length n.

Definition. A <u>cell</u> (or a <u>tape square</u>) is a natural number < n. A <u>segment</u> is an interval of cells. A <u>moment</u> is a natural number $\leq n^{1+\tau}$. For each moment t, $\rho_w(t)$ is the t-th configuration in the run of M_n on w. (If M_n halts at some moment $t' < t$, then $\rho_w(t) = \rho_w(t')$.)

Definition. Let S be a set of cells, and t be a moment. $Act_w(S,t)$ is the number of heads residing in S at moment t with respect to $\rho_w(t)$, and $Act_w(S) = \sum_t Act_w(S,t)$. Any maximal time-interval I, with $Act_w(S,t) > 0$ for all t in I, is a w-<u>session</u> for S.

The subscripts may be omitted if input is clear from the context.

Lemma 2.1. There are positive reals α, β, γ, δ such that:

(1) $\tau < \alpha$,

(2) $3(\alpha + \sigma) < \beta < \gamma < 1/2 - \alpha$,

(3) $\delta < \alpha - \tau$, $\delta < \gamma - \beta$, $\delta < 1/2 - (\alpha + \gamma)$.

Proof. Since $2\tau < 1/2 - \sigma$, the open interval $(\tau, 1/4 - \sigma/2)$ is not empty; pick any α there. Then (1) is satisfied, $2\alpha < 1/2 - \sigma$, and therefore $\alpha + \sigma < 1/2 - \alpha$. Pick β and γ such that $\alpha + \sigma < \beta < \gamma < 1/2 - \alpha$. Then (2) is satisfied. Finally, choose δ to satify (3). Q.E.D.

Let α, β, γ, δ be as in Lemma 2.1.

Definition. A set S of cells is *superactive* (wrt a given input) if S has more than n^α sessions or $\mathrm{Act}(S) > |S| \cdot n^\alpha$. Otherwise S is *moderate*.

Definition. A *binary function* is a function with values in $\{0,1\}$. 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.

A binary string of length n may be viewed as a binary function on $\{0, \dots, n-1\}$.

Definition. Let S be a segment and x be a binary function on the complement $\overline{S}$ of S. A binary function y on S is x-*moderate* if S is moderate wrt $x \cup y$. Two x-moderate functions y and z are x-*equivalent* if :

(1) the $(x \cup y)$-sessions for S are exactly the $(x \cup z)$-sessions for S, and

(2) if t is an end-point of any $(x \cup y)$-session for S then $\rho_{x \cup y}(t) = \rho_{x \cup z}(t)$.

Lemma 2.2. The number of x-equivalence classes is $o(5^{**}n^{\alpha+\sigma})$.

Proof. An x-equivalence class is fully determined by the number of sessions, the end-points of sessions, and the configuration at each such end-points. Let $l = \log_2 n$. There are at most n^α sessions. For each number of sessions, there are at most $2n^\alpha$ end-points and therefore at most $(n^{1+\tau})^{**}2n^\alpha$ sets of end-points. There are at most $2^{**}n^\sigma$ states of the work space and n^η possible states of the function Head. Thus, the number of x-equivalence classes is at most

$$n^\alpha \cdot [(n^{1+\tau})^{**}2n^\alpha] \cdot [((2^{**}n^\sigma) \cdot n^\eta)^{**}(2n^\alpha)] =$$

$$2^{**}[l\alpha + l(1+\tau)2n^\alpha + 2n^{\alpha+\sigma} + 2l\eta n^\alpha] =$$

$$o(5^{**}n^{\alpha+\sigma}).$$ Q.E.D.

Definition. Let S be a segment, and x, y be binary functions on $\overline{S}$, S respectively. A cell s in S is *flexible* wrt $x \cup y$ and S if S is x-moderate and the x-equivalence class of y contains some z with $z(s) \neq y(s)$; otherwise s is *rigid* wrt $x \cup y$ and S. For brevity, we say that S has k cells rigid wrt w if it has k cells rigid wrt w and S.

Now suppose that w (or w_n) is a random input with respect to the uniform probability distribution on inputs of length n.

Theorem 2.1. The probability of the event

[There exists a w-moderate segment S such that $n^\gamma < |S| \leq 2n^\gamma$ and S has $\geq n^\beta$ cells rigid wrt w]

converges to 0 when n grows to infinity.

Proof. We will use the following obvious and well-known facts.

Claim 2.1. Let E and $H_0, \dots, H_{k-1}$ be events in an arbitrary probability space Ω. If $H_0, \dots, H_{k-1}$ partition Ω then :

(1) $\Pr[E] = \sum_{i<k} \Pr[E \mid H_i] \cdot \Pr[H_i]$.

(2) $\Pr[E] \leq k \cdot \max\{\Pr[E \mid H_i]: i < k\}$.

(3) If all events H_i are equally probable then $\Pr[E] \leq \max\{\Pr[E \mid H_i]$.

Let S be a w-modest segment with $n^\gamma < |S| \leq 2n^\gamma$, x be a binary function on $\overline{S}$, and y range over binary functions on S. If S has $\geq n^\beta$ cells rigid wrt $x \cup y$ then the x-equivalence class of y contains at most $2**(|S| - n^{\beta})$ members. By Lemma 2.2, there are at most

$$(5**n^{\alpha+\sigma}) \cdot 2**(|S| - n^{\beta})$$

y's such that S has $\geq n^{\beta}$ rigid cells wrt $x \cup y$. Hence

$$\Pr[S \text{ has } \geq n^{\beta} \text{ rigid cells } | \; w|\overline{S} = x] \leq$$

$$(5**n^{\alpha+\sigma}) \cdot 2**(|S| - n^{\beta}) \cdot 2**(-|S|) =$$

$$(5**n^{\alpha+\sigma}) \cdot 2**(- n^{\beta}).$$

By Claim 2.1(3),

$$\Pr[S \text{ has } \geq n^{\beta} \text{ rigid cells}] \leq$$

$$(5**n^{\alpha+\sigma}) \cdot 2**(- n^{\beta}).$$

There are at most $n^{1+\gamma}$ segments S of length $n^{\gamma} < |S| \leq 2n^{\gamma}$ because there are at most n choices for $\min(S)$ and, given $\min(S)$, at most n^{γ} choices for $\max(S)$. By Claim 2.1(2),

$$\Pr[\text{There is a } w\text{-moderate } S \text{ such that } n^{\gamma} < |S| \leq 2n^{\gamma} \text{ and } S \text{ has } \geq n^{\beta} \text{ cells rigid wrt } w] \leq$$

$$n^{1+\gamma} \cdot \Pr[S \text{ has } \geq n^{\beta} \text{ rigid cells}] \leq$$

$$n^{1+\gamma} \cdot (5**n^{\alpha+\sigma}) \cdot 2**(- n^{\beta}) = o(1).$$

Q.E.D.

Lemma 2.3. Suppose that $n = \nu(\nu-1)/2$ and G is the graph with vertices $\{0, \ldots, \nu-1\}$ whose adjacency matrix is represented by w_n. The probability that G has a parabolic clique of size $\log_2 \nu$ converges to 0 when n grows to infinity.

Proof. Let $\lambda = \log_2 \nu$, and a_0, a_1, a_2 range over the vertices of G, and $X(a_0,a_1,a_2)$ be the set of vertices v_i, $1 \leq i \leq \lambda$, such that $v_i = a_0 + a_1 i + a_2 i^2 \bmod \nu$ for all i. The probability that $X(a_0,a_1,a_2)$ forms a cliques is $2^{-\lambda(\lambda-1)/2} = \nu^{-(\lambda-1)/2}$. There are ν^3 different sets $X(a_0,a_1,a_2)$. Hence the probability that some $X(a_0,a_1,a_2)$ forms a cliques is at most

$$\nu^3 \cdot \nu^{-(\lambda-1)/2} = o(1).$$

Q.E.D.

§3 Wards

Let w be an input of length n; w is the default input in this section.

Definition. A cell s is <u>active</u> at a moment t if some head resides in s at t, i.e. if $Act(\{s\},t) > 0$); s is <u>superactive</u> if $\{s\}$ is superactive, i.e. $Act(\{s\}) > n^{\alpha}$; and s is <u>moderate</u> if $\{s\}$ is moderate.

Definition. Partition $\{0, \ldots, n-1\}$ into segments Seg_i such that $\min(Seg_0) = 0$ and each $\min(Seg_{i+1})$ is the least w-moderate cell s such that $c \geq \min(Seg_i) + n^{\gamma}$. These segments are w-<u>wards</u>. The w-ward that contains a cell s will be denoted $Ward_w(s)$.

Lemma 3.1. Each w-ward has at most n^{α} active sessions.

Proof. Let $W_0, \ldots, W_m$ be all w-wards in the natural order. We suppose that some W_i has $k > n^{\alpha}$ sessions, and prove that some $\min(W_j)$ is superactive.

Case $i = 0$. At the beginning of any but the first session for W_0 some head enters W_0 from the right, and at the end of any but the last session some head leaves W_0 to the right. Thus, $\min(W_1)$ is active at least

$$1 + 2(k-2) + 1 = 2k - 2 \geq 2(n^{\alpha} + 1) - 2 = 2n^{\alpha} > n^{\alpha}$$

times and therefore is superactive.

Case $i = m$ is similar; $\min(W_m)$ turns out to be superactive.

Case $0 < i < m$. Let $a = \min(W_i)$ and $b = \min(W_{i+1})$. At the beginning of each session for W_i some head enters W_i, and at the end of any but the last session for W_i some head leaves W_i. Hence,

$$Act_w(a) + Act_w(b) \geq 2(k-1) + 1 \geq 2n^{\alpha} + 1$$

and therefore either a or b is superactive. Q.E.D.

Definition. A ward W is w-*regular* if it is w-moderate and $n^\gamma \le |W| < 2n^\gamma$. A cell s is w-*flexible* if $\mathrm{Ward}_w(s)$ is w-regular and s is flexible with respect to w and $\mathrm{Ward}_w(s)$; otherwise s is w-*rigid*.

Theorem 3.1. Suppose that every w-regular ward has less than n^β w-rigid cells. Then the total number of w-rigid cells is $o(n^{1-\delta})$.

Proof. First we prove that the union of all nonregular wards has $o(n^{1-\delta})$ cells. Let S_0 be the set of superactive cells, S_1 be the union of superactive wards, and S_2 be the union of extra-long (of length at least $2n^\gamma$) wards, and S_3 be the union of extra-short (of length $< n^\gamma$) wards.

Obviously, $|S_0|\cdot n^\alpha < \mathrm{Act}(S_0)$. By Lemma 3.1, $|W|\cdot n^\alpha < \mathrm{Act}(W)$ for every superactive ward. Hence, for each $i < 2$,

$$|S_i|\cdot n^\alpha < \mathrm{Act}(S_i) \le \mathrm{Act}(\{0, ..., n-1\}) = \eta n^{1+\tau}$$

and therefore $|S_i| < \eta n^{1+\tau-\alpha} = o(n^{1-\delta})$. All but the first n^γ cells of any ward W are superactive; if W is extra-long then at least one half of W-cells are superactive. Thus, $|S_2| \le 2|S_0| = o(n^{1-\delta})$. Finally, only the last ward can be extra-short; hence $|S_3| < n^\gamma$. But $\gamma < 1/2 - \alpha < 1 + \tau - \alpha < 1 - \delta$.

It remains to prove that all regular wards together contain $o(n^{1-\delta})$ rigid cells. Since each regular ward contains at least n^γ cells, there are at most $n^{1-\gamma}$ regular wards. Each of them contains at most n^β rigid cells. Hence all regular wards contain at most $n^{1+\beta-\gamma} = o(n^{1-\delta})$ rigid cells. Q.E.D.

§4 Independence

Again, w is the default input of length n.

Definition. For each w-ward U, each session I for U and each input head h, let $\mathrm{PAZ}_w(U,I,h)$ be the segment $[s - |I|, s + |I|]$ where s is the position of h in $\rho_w(\min(I))$. Here PAZ abbreviates the phrase "potentially active zone". Let

$$\mathrm{PAZ}_w(U,I) = \bigcup_h \mathrm{PAZ}_w(U,I,h), \text{ and}$$

$$\mathrm{PAZ}_w(U) = \bigcup_I \mathrm{PAZ}_w(U,I).$$

Lemma 4.1. If U is a regular w-ward then $\mathrm{PAZ}_w(U)$ intersects $< 7\eta n^\alpha$ regular w-wards.

Proof. Since each PAZ(U,I,h) is a segment of length $\le 1+2|I|$ and regular wards are at least n^γ long, each PAZ(U,I,h) intersects at most $3 + 2|I|n^{-\gamma}$ wards. Hence each PAZ(U,I) intersects at most $3\eta + 2\eta|I|n^{-\gamma}$ regular wards. Since U is regular, it has at most n^α sessions, and

$$\sum\{|I| : I \text{ is a session for } U\} < \mathrm{Act}(U) \le |U|n^\alpha < 2n^{\alpha+\gamma}.$$

Hence PAZ(U) intersects $< 3\eta\cdot n^\alpha + 2\eta n^{-\gamma}\cdot 2n^{\alpha+\gamma} = 7\eta n^\alpha$ regular wards. Q.E.D.

Lemma 4.2. Let S be a regular w-ward, $I = [a,b]$ be a w-session for S, $x = w|\overline{S}$, and $y = w|S$. Let x' be a binary function on $\overline{S}$ which coincides with x on $\mathrm{PAZ}_w(S,I)$, y' be a binary function on S which is x-equivalent to y, and $v = x' \cup y'$. If $\rho_v(a) = \rho_w(a)$ then $\rho_v(b) = \rho_w(b)$.

Proof. Let $u = x \cup y'$. Since y and y' are x-equivalent, I is a u-session for S and $\rho_u(a) = \rho_w(a)$, $\rho_u(b) = \rho_w(b)$. Suppose $\rho_v(a) = \rho_w(a)$. Then all three configurations $\rho_u(a)$, $\rho_v(a)$ and $\rho_w(a)$ coincide. Notice that $\mathrm{PAZ}_w(S,I)$ is completely defined by $\rho_w(a)$, and the same holds for u and v. Thus, $\mathrm{PAZ}_u(S,I) = \mathrm{PAZ}_v(S,I) = \mathrm{PAZ}_w(S,I)$, and therefore x' coincide with x on $\mathrm{PAZ}_u(S,I)$.

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, $\rho_v(b) = \rho_u(b) = \rho_w(b)$. Q.E.D.

Definition. Two w-wards U an V are _independent_ if they are regular, V is disjoint from $PAZ_w(U)$, and U is disjoint from $PAZ_w(V)$. Three or more w-wards are _independent_ if every two of them are.

Theorem 4.1. Let u be an input, $W_0, \ldots, W_{l-1}$ be independent u-wards, $y_i = u|W_i$, and z_i be a binary function on W_i which is $(u|\overline{W_i})$-equivalent to y_i. Let v be the result of the simultaneous replacement of $y_0, \ldots, y_{l-1}$ by $z_0, \ldots, z_{l-1}$ in u. Then $\rho_u(t) = \rho_v(t)$ for all t, and therefore v is accepted if and only if u is.

Proof. Since the wards W_i are independent, their sessions are disjoint. Let $[a_0,b_0], \ldots, [a_{k-1},b_{k-1}]$ be all sessions for all u-wards $W_0, \ldots, W_{l-1}$ in the natural order (so that $b_i < a_j$ if $i < j$)).

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 :

(1) $\rho_u(t) = \rho_v(t)$ for all $t \leq a_0$,

(2) if $\rho_u(b_i) = \rho_v(b_i)$ then $\rho_u(a_{i+1}) = \rho_v(a_{i+1})$ for all $i < k-1$,

(3) if $\rho_u(b_{k-1}) = \rho_v(b_{k-1})$ then $\rho_u(t) = \rho_v(t)$ for all $t > b_{k-1}$.

It remains to prove that, for all $i < k$,

(4) if $\rho_u(a_i) = \rho_v(a_i)$ then $\rho_u(b_i) = \rho_v(b_i)$.

Without loss of generality, W_0 is active during $[a_i,b_i]$. Now use Lemma 4.2 with w = u, I = $[a_i,b_i]$, S = W_0, x' being the result of the simultaneous replacement of $y_1, \ldots, y_{l-1}$ by $z_1, \ldots, z_{l-1}$ in u, and $y' = z_0$. Q.E.D.

§5 Random cliques

Lemma 5.1. For every sufficiently large ν there is an input w of size $n = \nu(\nu-1)/2$ such that :

(1) every w-regular ward W has less than n^β rigid cells, and

(2) the graph represented by n has no parabolic cliques of size $\log_2\nu$.

Proof. Use Theorem 2.1 and Lemma 2.3. Q.E.D.

In the rest of this section, ν is a prime number, $n = \nu(\nu-1)/2$, $\lambda = \log_2\nu$ and $w = w_\nu$ is an input of size n satisfying clauses (1) and (2) of Lemma 5.1. We are interested in the graph with vertices $\{0, \ldots, \nu-1\}$ whose adjacency matrix is represented by w.

Consider a new sample space: sample points are triples (a_0,a_1,a_2) of numbers $< \nu$, and the probability distribution is uniform. Let (a_0,a_1,a_2) be a random sample point. The binary function Cd was defined in Section 1.

Definition. For each positive integer $i \leq \lambda$, $v_i = v_i(a_0,a_1,a_2)$ is the vertex such that $v_i = a + bi + ci^2 \bmod \nu$. If $v_i \neq v_j$ then $s_{ij} = Cd(v_i,v_j)$ and $W_{ij} = Ward(s_{ij})$; if $v_i = v_j$ then s_{ij} and W_{ij} are undefined.

Definition. If the probability Pr[E] of an event E is $o(\nu^{-\delta})$ then E (as well as Pr[E]) is _negligible_ and the complement $\overline{E}$ is _almost sure_.

We intend to prove that the event [The wards W_{ij} are independent] is almost sure.

Lemma 5.2. (1) $Pr[v_i = v_j]) = \nu^{-1}$ for $i \neq j$.

(2) $Pr[v_i = v] = \nu^{-1}$ for all i and all $v < \nu$.

(3) $Pr[s_{ij}$ is defined and equal to $s] = 2\nu^{-2}$ for all s_{ij} and all cells s.

(4) $\Pr[v_j = v | v_i = u] = \nu^{-1}$ for all $i \neq j$ and all $u \neq v$.

(5) Every event [s_{ij} is defined and flexible] is almost sure.

Proof. (1) $v_i = v_j$ <--> $a_0 + a_1 i + a_2 i^2 = a_0 + a_1 j + a_2 j^2 \mod \nu$ <--> $a_1(j-i) + a_2(j^2 - i^2) = 0 \mod \nu$ <--> $a_1 = - a_2(i+j) \mod \nu$. Hence the event $[v_i = v_j]$ contains ν^2 sample points.

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

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

(3) There are unique natural numbers $u < v$ such that $s = v(v-1)/2 + u$; hence $s_{ij} = s$ if and only if $\{v_i, v_j\} = \{u, v\}$. The system

$$a_0 + a_1 i + a_2 i^2 = u \mod \nu$$

$$a_0 + a_1 j + a_2 j^2 = v \mod \nu$$

has ν solutions, and the system

$$a_0 + a_1 i + a_2 i^2 = v \mod \nu$$

$$a_0 + a_1 j + a_2 j^2 = u \mod \nu$$

has ν solutions. The two sets of solutions are disjoint. Hence $\Pr[s_{ij} = s] = 2\nu/\nu^3$.

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

(5) We prove that the event [s_{ij} is undefined or rigid] is negligible. By (1), the event [s_{ij} is undefined] is negligible. It remains to prove that the event [s_{ij} is defined and rigid] is negligible. By Theorem 3.1, the number r of rigid cells is $o(\nu^{2-2\delta})$. By (3), $\Pr$[s_{ij} is defined and rigid] $\leq r \cdot 2\nu^{-2} = o(\nu^{-2\delta})$. Q.E.D.

Lemma 5.3. For all distinct i, j, k and l,

$\Pr$[s_{ij} and s_{kl} are flexible, and PAZ(W_{ij}) intersects W_{kl}]

is negligible.

Proof. Without loss of generality, we may restrict attention to the case i = 1, j = 2, k = 3 and l = 4. By Claim 2.1(3) and Lemma 5.2(3), it suffices to prove that

$\max_s \Pr$[s_{34} is flexible, and PAZ(W_{12}) intersects W_{34} | $s_{12} = s$],

where s ranges over flexible cells, is negligible. Fix a flexible s where the maximum is reached; let u, v be the distinct vertices with s = Cd(u,v). By virtue of symmetry, it suffices to prove that

$\Pr$[s_{34} is flexible, and PAZ(W_{12}) intersects W_{34} | $v_1 = u$ and $v_2 = v$]

is negligible. Let E = [$v_1 = u$ and $v_2 = v$], PAZ(s) = PAZ(Ward(s)) and U(s) be the union of all regular wards intersected by PAZ(s). It suffices to prove that

$\Pr$[s_{34} belongs to U(s) | E]

is negligible. By Lemma 4.1, PAZ(s) intersects at most $7\eta n^\alpha$ wards. Since each regular ward contains at most $2n^\gamma$ cells, U(s) contains at most $14\eta n^{\alpha+\gamma}$ cells; by the choice of δ, $|U(s)| = o(\nu^{1-\delta})$ cells. It is easy to see that the event [E and $s_{34} = s'$] contains either 0 or 2 sample points for every s'. Thus,

$\Pr$[s_{34} belongs to U(s) | E] =
$\Pr$[E and s_{34} belongs to U(s)] / $\Pr$[E] $\leq$
$(2 \cdot |U(s)|/\nu^3) / \nu^{-2} = o(\nu^{-\delta})$.

Q.E.D.

Definition. A vertex u is <u>bad</u> if there are $\geq \nu^{1-\delta}$ vertices v such that Cd(u,v) is rigid; otherwise u is <u>good</u>.

Lemma 5.4. Every $\Pr$[v_i is bad] is negligible.

Proof. Let k is the number of bad vertices. By Lemma 5.2(2), $\Pr$[v_i is bad] = $k\nu^{-1}$. It suffices to prove that $k = o(\nu^{1-\delta})$.

There are at least $k \cdot \nu^{1-\delta}/2$ rigid cells of the form Cd(u,v) where u or v is bad. By Theorem 3.1, the number of rigid cells is $o(\nu^{2-2\delta})$. Hence $k = o(\nu^{1-\delta})$. Q.E.D.

Lemma 5.5. For all distinct i, j and k,

> Pr[s_{ij} and s_{ik} are flexible, and PAZ(W_{ij}) intersects W_{ik}],

is negligible.

Proof. Without loss of generality, we may restrict attention to the case i = 1, j = 2 and k = 3. Since Pr[v_1 is bad] is negligible, it suffices to prove that

> Pr[v_1 is good, and s_{12}, s_{13} are flexible, and PAZ(W_{12}) intersects W_{13}]

is negligible. By Claim 2.1(3) and Lemma 5.2(2), it suffices to prove that

> $\max_u$ Pr[s_{12}, s_{13} are flexible, and PAZ(W_{12}) intersects W_{13} $|$ $v_1 = u$],

where u ranges over good vertices, is negligible. Fix a good u where the maximum is reached; let F_u be the set of vertices v such that $v \neq u$ and the cell Cd(u,v) is flexible. By Claim 2.1(3) and Lemma 5.2(3), it suffices to prove that

> $\max_v$ Pr[s_{13} is flexible, and PAZ(W_{12}) intersects W_{13} $|$ $v_1 = u$ and $v_2 = v$],

where v ranges over F_u, is negligible. Fix some v in F_u where the maximum is reached; let s = Cd(u,v), PAZ(s) = PAZ(Ward(s), U(s) be the union of all regular wards intersected by PAZ(s), and E be the event [$v_1 = u$ and $v_2 = v$]. It suffices to prove that

> Pr[s_{13} belongs to U(s) $|$ E]

is negligible. As in the proof of Lemma 5.2, $|U(s)| = o(\nu^{1-\delta})$ cells. Different sample points of E give different value to v_3 and therefore to s_{13}. Thus,

> Pr[s_{13} belongs to U(s) $|$ E] =
> Pr[E and s_{13} belongs to U(s)] / Pr[E] $\leq$
> $(|U(s)|/\nu^3) / \nu^{-2} = o(\nu^{-\delta})$. Q.E.D.

Theorem 5.1. The event [All vertices v_i are different, and all cells s_{ij} are flexible, and the wards W_{ij} are independent] is almost sure.

Proof. Use Lemmas 5.3 and 5.5. Q.E.D.

Finally, we are ready to finish the proof of Main Theorem. By virtue of Theorem 5.1, we can choose a parabolic subset $v_1, \ldots, v_\lambda$ of different vertices such that all cells s_{ij} are flexible and the wards W_{ij} are independent. For every W_{ij}, let $y_{ij} = w|W_{ij}$ and z_{ij} be a binary function on W_{ij} such that z_{ij} is $(w|\overline{W_{ij}})$-equivalent to y_{ij} and $z_{ij}(s_{ij}) = 1$. Let w' be the result of simultaneous replacements of every y_{ij} by z_{ij}. By Theorem 4.1, the machine M_n accepts w', but the graph represented by w' contains a parabolic clique $v_1, \ldots, v_\lambda$ which is impossible. Q.E.D.

References

Bo B. Bollobas, "Random graphs", Academic Press, London, 1985.

BFMUW A. Borodin, F. Fich, F. Meyer auf der Heide, E. Upfal and A. Widgerson, "A time-space tradeoff for element distinctness", SIAM J. Computing 16, nu. 1, Feb. 1987.

DG P. Duris and Z. Galil, "A time-space tradeoff for language recognition", Math. Systems Theory 17 (1984), 3-12.

PPST W. J. Paul, N. Pippenger, E. Szemeredi and W. T. Trotter, "On determinism versus non-determinism and related problems", Proc. 24th Symposium on Foundation of Computer Science, Nov. 1983, Tucson, Arizona, 429 - 438.

Si J. Simon, Private Communication, Feb. 1988.