Monotone versus Positive

MIKLOS AJTAI

IBM Research, San Jose, California

AND

YURI GUREVICH

University of Michigan, Ann Arbor, Michigan

Abstract. In connection with the least fixed point operator the following question was raised: Suppose that a first-order formula $\varphi(P)$ is (semantically) monotone in a predicate symbol P on finite structures. Is $\varphi(P)$ necessarily equivalent on finite structures to a first-order formula with only positive occurrences of P? In this paper, this question is answered negatively. Moreover, the counterexample naturally gives a uniform sequence of constant-depth, polynomial-size, monotone Boolean circuits that is not equivalent to any (however nonuniform) sequence of constant-depth, polynomial-size, positive Boolean circuits.

Categories and Subject Descriptors: F1.1 [**Computation by Abstract Devices**]: Models of Computation; F.4.1 [**Mathematical Logic and Formal Languages**]: Mathematical Logic; G.2.1 [**Discrete Mathematics**]: Combinatorics; G.3 [**Mathematics of Computing**]: Probability and Statistics

General Terms: Algorithms, Languages, Theory, Verification

Additional Key Words and Phrases: Boolean circuits, constant-depth, first-order formulas, monotone, polynomial size, positive, probabilistic

1. *Introduction*

Let $\varphi(P)$ be a first-order formula where P is an l-ary predicate symbol. Let σ be the rest of the signature of φ, and $x_1, \ldots, x_m$ be the free individual variables of $\varphi(P)$. View the symbols in σ as constants and P as a predicate variable. Then $\varphi(P)$ represents an operator assigning the m-ary predicate $P' = \lambda(x_1, \ldots, x_m).\varphi(P)$ to each l-ary predicate P.

$\varphi(P)$ is called monotone (in P) if $P \subseteq Q$ logically implies $P' \subseteq Q'$. Here Q is a new l-ary predicate variable and $P \subseteq P'$ abbreviates $\forall x_1 \cdots x_l[P(x_1, \ldots, x_l) \rightarrow Q(x_1, \ldots, x_l)]$. Recognizing monotonicity is an undecidable problem: A σ-sentence α is valid if and only if the sentence $\exists x_1 \cdots x_l P(x_1, \ldots, x_l) \rightarrow \alpha$ is monotone. A sufficient condition for the monotonicity of $\varphi(P)$ is that $\varphi(P)$ is logically

A preliminary version of this paper appeared in Monotone versus positive, IBM Tech. Rep. RJ 4697 (50035). IBM Research Center, San Jose, Calif., May 1985.

The work of M. Ajtai was partially supported by National Science Foundation grant DMS 84-01281. The work of Y. Gurevich was partially supported by National Science Foundation grants MCS 83-01022 and DCR 85-03275.

Authors' addresses: M. Ajtai, 16712 Chirco Drive, Los Gatos, CA 95030; Y. Gurevich, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122.

equivalent to a first-order formula that is positive in P (has only positive occurrences of P); and the positivity is, of course, easily recognizable. Lyndon proved that this sufficient condition is necessary [11]. (This desired $\psi(P)$ is a Lyndon interpolant [4] for the implication $(P' \subseteq P)$ & $\varphi(P') \rightarrow \varphi(P)$.)

In the rest of this paper we consider only finite structures. Thus, $\varphi(P)$ will be called monotone if $P \subseteq P'$ implies $\varphi(P) \subseteq \varphi(P')$ on finite structures. Again, recognizing monotonicity is an undecidable problem. And again, positivity is a sufficient condition for monotonicity: $\varphi(P)$ is monotone in P if it is equivalent on finite structures to a first-order formula positive in P. It was conjectured that this sufficient condition is not necessary [6]. The conjecture is proved here.

Actually, we prove a stronger result. To formulate it nicely, we use a generalization of the classical Borel hierarchy to finite topological spaces [7]. The original goal of the generalization was to understand the paper [1]. This paper is, in a sense, an offspring of [1], which shares some results with [5]. Connections between circuit complexity and the classical Borel hierarchy were explored in [12].

Subsets of a set S will be called *points* over S, and sets of points over S will be called *point-sets* over S.

Definition. Let M be a point-set over a set S. M is *positive Borel of level* 0 if M is empty or M contains all points over S or $M = \{X \subseteq S : a \in X\}$ for some $a \in S$. M is *positive Borel of level* $i+1$ if it is the union or the intersection of at most $|S|$ positive Borel point-sets over S of level i.

CLAIM. *Let σ be a signature comprising only predicate symbols and individual constants, P be an additional unary predicate symbol, and $\psi(P)$ be a sentence of the signature $\sigma \cup \{P\}$. Suppose that $\varphi(P)$ is positive in P, d is the logical depth of $\varphi(P)$, and S is a σ-structure. Then the point-set $\{P \subseteq S : S \vDash \psi(P)\}$ over S is positive Borel of level d.*

PROOF. First extend σ by means of individual constants corresponding to elements of S. Then prove the claim (for sentences in the extended signature) by induction on the logical depth. □

MAIN THEOREM. *There exists a sequence $S_1, S_2, \ldots$ of structures of some signature σ, a first-order σ-sentence φ_0, and a first-order sentence $\varphi_1(P)$ in the signature σ plus an additional unary predicate symbol P such that σ contains only predicate symbols and*

(*i*) *an arbitrary σ-structure satisfies φ_0 if and only if it is isomorphic to some S_n,*
(*ii*) *the sentence φ_0 & $\varphi_1(P)$ is monotone in P, and*
(*iii*) *for every i there is n_i such that for every $n \geq n_i$, the point-set $\{P \subseteq S_n : S_n \vDash \varphi_1(P)\}$ over S_n is not positive Borel of level i.*

COROLLARY 1. *The sentence φ_0 & $\varphi_1(P)$ is monotone in P but not equivalent on finite structures to any first-order sentence positive in P.*

PROOF. By contradiction, suppose that φ_0 & $\varphi_1(P)$ is equivalent on finite structures to a first-order sentence $\psi(P)$ that is positive in P.

Without loss of generality, the signature of $\psi(P)$ contains only predicate symbols and even is included into $\sigma \cup \{P\}$. The extra predicates can be made identically true; the equivalence $(\varphi_0$ & $\varphi_1(P)) \leftrightarrow \psi(P)$ will survive.

Let d be the logical depth of $\psi(P)$ and n be the number n_d of (iii). By the Claim, the point-set $\{P \subseteq S_n : S_n \vDash \psi(P)\}$ is positive Borel of level d, which contradicts the clause (iii). □

We consider Boolean circuits with AND, OR, and NOT gates. A circuit will be called *monotone* if it computes a Boolean function $f(x_1, \ldots, x_l)$ that is monotone: If $x_i \le y_i$ for all i, then $f(x_1, \ldots, x_l) \le f(y_1, \ldots, y_l)$. A circuit will be called *positive* if it has no NOT gates.

COROLLARY 2. *There is a constant-depth polynomial-size sequence $C_1, C_2 \ldots$ of monotone Boolean circuits such that for all natural numbers d, s and any sufficiently large natural number n, no positive Boolean circuit of depth $\le d$ and size $\le n^s$ computes the Boolean function of C_n.*

PROOF. Let structures S_n and formulas φ_0, $\varphi_1(P)$ be as in the Main Theorem. Let σ_n be the extension of the signature σ by $|S_n|$ individual constants $C_1, C_2, \ldots$ naming the elements of S_n. Turn every first-order sentence ψ in the signature $\sigma_n \cup \{P\}$ into a Boolean formula ψ^*, as follows: If ψ is an atomic σ_n-sentence that holds (respectively, fails) in S_n, then ψ^* is an identically true (respectively, false) Boolean formula. If $\psi = P(c_i)$, then $\psi^* = \psi$. If ψ is the conjunction (respectively, disjunction) of α and β, then ψ^* is the conjunction (respectively, disjunction) of α^* and β^*. If $\psi = \neg\alpha$, then $\psi^* = \neg\alpha^*$. If ψ is $(\forall x)\alpha(x)$ (respectively, $(\exists x)\alpha(x)$), then ψ^* is the conjunction (respectively, disjunction) of formulas $(\alpha(c_i))^*$. The Boolean formula $(\varphi_0 \mathbin{\&} \varphi_1(P))^*$ gives the desired circuit C_n. □

Remark. In our construction, $|S_n| = n \times \lfloor \log_2 n \rfloor$ (for $n > 1$) and C_n has $|S_n|$ inputs. It is not difficult to achieve $|S_n| = n$ as one may desire.

Corollary 2 is a kind of lower bound on the complexity of positive Boolean circuits. Boppana [2] gives lower bounds of a different type on the complexity of positive Boolean circuits (called *monotone* in [2]).

Remark. The sequence of structures S_n is uniform in our construction. The corresponding sequence of circuits C_n is uniform, of constant depth and polynomially bounded size. This sequence of circuits is not equivalent to any (whatever nonuniform) sequence of constant-depth, polynomial-size, positive circuits.

Note that every monotone Boolean circuit C is equivalent to some positive Boolean circuit. Consider for example a minimal Boolean formula ψ in the disjunctive normal form which is equivalent to the Boolean formula of C. If ψ has a disjunct $\alpha \mathbin{\&} \neg y$, then, by the monotonicity, every assignment satisfying α satisfies ψ. Hence $\alpha \mathbin{\&} \neg y$ can be replaced by α, which contradicts the minimality of ψ.

Corollary 2 indicates that the conversion of a monotone circuit into an equivalent positive one may not be easy. Since recognizing the monotonicity of a circuit is co-NP-complete [7], there is no polynomial-time algorithm—unless P = NP—which transforms an arbitrary circuit C into an equivalent circuit C' in such a way that C' is positive whenever C is monotone.

The Main Theorem is proved in Sections 2–4. Structures S_n and sentences φ_0, $\varphi_1(P)$ are defined in Section 2, and statements (i) and (ii) are proved there too. The universe of S_n (for $n > 1$) consists of pairs (x, y), of natural numbers such that $x < l = \lfloor \log_2 n \rfloor$ and $y < n$. With every subset P of S_n, we associate a function

$$P^*(x) = \max(\{0\} \cup \{y : S_n \models P(x, y)\}), \qquad 0 \le x < l.$$

The desired $\varphi_1(P)$ says that $\sum P^*(x) \ge l(n-1)/2$. Obviously, $\varphi_1(P)$ is monotone in P. In Section 3, we prove that $\{P^* : \varphi_1(P)\}$ has many minimal elements with respect to the componentwise ordering. In Section 4, we prove that sets of functions

P^*, definable by positive Borel conditions, do not have many minimal elements with respect to the componentwise ordering. This will establish statement (iii). The main difficulty is to formulate an appropriate notion of "many"; our notion of "many" has a probabilistic character.

The question about the status of Lyndon's theorem in the case of finite structures was raised by Chandra and Harel [3] in relation to the extension FO + LFP of first-order logic by means of the following least fixed point formation rule. A formula $\varphi(P, \bar{x})$, where the arity of a predicate variable P equals the length of the tuple $\bar{x}$ of individual variables, represents an operator $P \mapsto \lambda\bar{x}.\varphi(P, \bar{x})$ which can be iterated; if φ is positive in P, then the operator is monotone in P and therefore has a least fixed point $\mathrm{LFP}_{P;\bar{x}}(\varphi)$. If Lyndon's theorem were true in the case of finite structures, it would be an indication that FO + LFP loses no expressive power by sticking to positive rather than arbitrary monotone formulas. Fortunately, no expressive power is lost anyway: the two extensions coincide by their expressive power [8].

2. *The Monotone Formula*

In this section we prove the Main Theorem, except for statement (iii).

Definition. $\sigma 1$ is the signature $\{\leq, \mathrm{Sum}, \mathrm{Prod}, \mathrm{Exp}\}$, where $\leq$ is a binary predicate symbol, and Sum, Prod, Exp are ternary predicate symbols. A $\sigma 1$-structure S of cardinality n is *standard* if

(i) the universe of S is the interval $[0, n)$ of natural numbers,
(ii) $\leq$ is the standard ordering of the universe of S,

$$\begin{array}{lll} S \models \mathrm{Sum}(x, y, z) & \text{if and only if} & x + y = z \quad \text{modulo } n, \\ S \models \mathrm{Prod}(x, y, z) & \text{if and only if} & x \cdot y = z \quad \text{modulo } n, \\ S \models \mathrm{Exp}(x, y) & \text{if and only if} & 2^x = y \quad \text{modulo } n. \end{array}$$

THEOREM 2.1. *There is a first-order $\sigma 1$-sentence $\psi 1$ such that an arbitrary $\sigma 1$-structure is a model of $\psi 1$ if and only if it is isomorphic to a standard structure.*

PROOF. $\psi 1$ says that $\leq$ is a linear order, and Sum, Prod, Exp satisfy the usual recursive definitions [10]. □

LEMMA 2.1. *For any positive integer m, let $U(m)$ be the least common multiple of all positive integers up to m. For any integer $n \geq 2$, $U(2\lceil \log_2 n \rceil) \geq n$.*

PROOF. See [9]. □

LEMMA 2.2. *For every integer $n \geq 6$, the number of functions from the interval $[0, \lfloor (\log_2 n)^{1/4} \rfloor)$ of natural numbers to the interval $[0, 2\lceil \log_2 n \rceil)$ of natural numbers is at most n.*

PROOF. Omitted. □

Definition. $\sigma 2 = \sigma 1 \cup \{Q, R\}$ where Q is a ternary predicate symbol and R is a quaternary predicate symbol. Let $l = \log_2 n$ and $f_0, f_1, \ldots, f_{m-1}$ be the list of functions from $[0, \lfloor l^{1/4} \rfloor)$ to $[0, 2\lceil l \rceil)$ in the lexicographical order. A $\sigma 2$-structure S of cardinality n is *standard* if $n \geq 6$ and

(i) the $\sigma 1$-reduct of S is standard,
(ii) $S \models Q(i, j, k)$ iff $i < m$, $j < \lfloor l^{1/4} \rfloor$, and $f_i(j) = k$, and
(iii) $S \models R(i, x, y, p)$ iff $i < m$, $x \leq \lfloor l^{1/4} \rfloor$, $2 \leq p \leq 2\lceil l \rceil$, and $\sum \{f_i(j) : j < x\} = y$ modulo p. (The empty sum by definition is 0.)

THEOREM 2.2. *There is a first-order $\sigma 2$-sentence $\psi 2$ such that an arbitrary $\sigma 2$-structure is a model of $\psi 2$ if and only if it is isomorphic to a standard $\sigma 2$-structure.*

PROOF. $\psi 2$ is a conjunction of four sentences. The first conjunct is $\psi 1$. The second conjunct says that the universe contains at least six elements. The third conjunct describes Q by induction on the first argument. The forth conjunct describes R by induction on the second argument. □

THEOREM 2.3. *Let $\sigma 3 = \sigma 2 \cup \{f\}$ where f is a unary function symbol. There is a $\sigma 3$-formula $\psi 3(y)$ satisfying the following condition. Let S be a $\sigma 3$-structure with a standard $\sigma 2$-reduct, and let $n = |S|$, $l = log_2 n$. Suppose that $\sum \{f(j): j < \lfloor l^{1/4} \rfloor\} < n$. Then $S \models \psi 3(y)$ if and only if $\sum \{f(j): j < \lfloor l^{1/4} \rfloor\} = y$.*

PROOF. Let $f_0, f_1, \ldots, f_{m-1}$ be as in the definition of standard $\sigma 2$-structures, and let $r = \lfloor l^{1/4} \rfloor$. The desired formula $\psi 3(y)$ uses the predicates Q and R to say that for every $p \leq 2\lceil l \rceil$ there is $i < m$ such that

(a) for all $j < r$, $f(j) = f_i(j)$ modulo p, and
(b) $\sum \{f_i(j): j < r\} = y$ modulo p.

In virtue of Lemma 2.2, the equality $\sum \{f(j): j < r\} = y$ implies $S \models \psi 3(y)$. Suppose that $S \models \psi 3(y)$. Then $\sum \{f(j): j < r\} = y$ modulo every positive $p \leq 2\lceil l \rceil$. Hence $\sum \{f(j): j < r\} = y$ modulo the least common multiple of all positive numbers $p \leq 2\lceil l \rceil$. By Lemma 1, $\sum \{f(j): j < r\} = y$. □

THEOREM 2.4. *There is a $\sigma 3$-formula $\psi 4(y)$ satisfying the following condition. Let S be a $\sigma 3$-structure with a standard $\sigma 2$-reduct, and let $n = |S|$, $l = \lfloor log_2 n \rfloor$. Suppose that $\sum \{f(i): i < l\} < n$. Then $S \models \psi 4(y)$ if and only if $\sum \{f(i): i < l\} = y$.*

PROOF. Let $\alpha(f, y) = \psi 3(y)$. Let $r = \lceil l^{1/4} \rceil$, so that $r^4 \geq l$.

The formula $\alpha(\lambda j.f(ru + j), y)$ says that $\sum \{f(j): ru \leq j < ru + r\} = y$. Let $g(u) = \sum \{f(j): ru \leq j < ru + r\}$. The formula $\alpha(g, y)$ says that $y = \sum \{f(j): j < r^2\}$.

Using the same trick again, we arrive at a formula $\beta(f, y)$ saying that $\sum \{f(j): j < r^4\} = y$. Let $h(j) = f(j)$ if $j < l$, and $h(j) = 0$, otherwise. $\beta(h, y)$ is the desired $\psi 4(y)$.

The above proof assumes implicitly that $r^4 \leq n$. We ignore the modification needed to cover the case $r^4 > n$. □

Now we are ready to describe the desired structures S_n.

Definition. $\sigma = \sigma 2 \cup \{D, T\}$ where D is a unary predicate symbol and T is a ternary predicate symbol.

Definition. Let $n \geq 4$ and $l = \lfloor \log_2 n \rfloor$. S_n is the σ-structure such that

(i) the universe of S consists of pairs (x, y) of natural numbers where $0 \leq x < l$ and $0 \leq y < n$,
(ii) the map $(x, y) \to nx + y$ is an isomorphism of the $\sigma 2$-reduct of S onto the $\sigma 2$-standard structure of cardinality ln,
(iii) the interpretation of T in S is $\{(x, 0), (0, y), (x, y): 0 \leq x < l, 0 \leq y < n\}$,
(iv) the interpretation of D in S is $\{(x, x): 0 \leq x < l\}$.

Clause (ii) requires $ln \geq 6$ because a standard $S2$-structure contains at least six elements. This explains the restriction $n \geq 4$. Structures S_1, S_2, S_3 should be defined separately; we ignore them and suppose $n \geq 4$ in the rest of the proof of the Main Theorem. Let $l = \lfloor \log_2 n \rfloor$.

It is easy to write down the desired first-order σ-sentence φ_0 such that every model of φ_0 is isomorphic to some S_n. In the following description of φ_0 we suppose that S is a model of φ_0. Our φ_0 is a conjunction. One conjunct is the sentence $\psi 2$ of Theorem 2.2; this guarantees that the $\sigma 2$-reduct of S is standard. Let A, B be the first and the second projections of T in S, respectively; the projections inherit orderings from S. The second conjunct of φ_0 says that T is an isomorphism of the lexicographically ordered set $A \times B$ onto the $\{<\}$-reduct of S. The third conjunct says (using the arithmetic built into S) that there are $a, b \in S$ such that $\forall y(y < b \leftrightarrow y \in B)$, $a = \lfloor \log_2 b \rfloor$, $a \cdot b = |S|$ and $D = \{x \cdot b + x : x < a\}$.

Definition. With each subset P of S_n we associate the function $P^*(x) = \max(\{0\} \cup \{y : S \models P(x, y)\})$ where $0 \leq x < l$ (the *associate function* of P).

Let P be a unary predicate symbol. Using the formula $\psi 4$ of Theorem 2.4, write down a $\sigma \cup \{P\}$-sentence saying in each S_n that $\sum \{P^*(x) : 0 \leq x < l\} \geq l(n-1)/2$; this is the desired sentence $\varphi_1(P)$.

The sentence φ_0 & $\varphi_1(P)$ obviously is monotone in P.

3. *Abundance of Minimal Elements*

Consider any structure S_n constructed in Section 2. A function P^* was associated with each subset P of S_n. In this section we prove that the set $\{P : S_n \models \varphi_1(P)\}$ has many minimal elements with respect to the componentwise ordering.

Definition. If X is a set and Y is a linearly ordered set, then $\{X \to Y\}$ is the lattice of functions from X to Y ordered componentwise: $f \leq g$ if and only if $f(x) \leq g(x)$ for all $x \in X$.

Definition. A lattice L is called *regular* if there are finite nonempty intervals X, Y of natural numbers, functions f and g from X to Y, and a positive integer n such that $g(x) = f(x) + n - 1$ for all $x \in X$ (so that each interval $[f(x), g(x)]$ contains n natural numbers), and L is the interval $[f, g] = \{h \in L : f \leq h \leq g\}$ of $[X \to Y]$. The interval X is called the *source* of L. The function f is called the *bottom* of L and denoted bottom_L, the function g is called the *top* of L and denoted top_L. The number n is called the *height* of L.

Every sublattice of a regular lattice L is an interval $[h_1, h_2]$ for some functions $h_1 \leq h_2$. An interval of a regular lattice is called regular if it forms a regular sublattice.

Let $l = \lfloor \log_2 n \rfloor$, and $L = [X \to Y]$, where X, Y are the intervals $[0, l)$ and $[0, n)$ of natural numbers, respectively.

Proviso. The probability distribution of any random variable is supposed to be uniform (when all possible values are equally probable) unless the contrary is clear from the context.

THEOREM 3.1. *Suppose that m is an integer such that $3 \leq m < n$ and $n - m$ is even. Let J be a random regular interval of L of height m. Let E be the event that there is $f \in J$ such that $f >$ bottom$_J$ and $\sum f(x) = \lceil l(n-1)/2 \rceil$. Then $Pr[E] > (ln)^{-1}$.*

Remark. The restriction "$n - m$ is even" is not essential, but it simplifies somewhat the proof and suffices for our purposes.

We prove Theorem 3.1 using the following lemma.

LEMMA 3.1. *Let $Y_K = y_1 + \cdots + y_k$ where $y_1, \ldots, y_k$ are independent random variables on a fixed interval $[-r, r]$ of integers. Then $Pr[Y_k = a] \geq Pr[Y_k = b]$ for all integers a, b with $|a| \leq |b|$.*

PROOF. The case $k = 1$ is obvious. Suppose Lemma 3 is proved for k, $Y = Y_k$, $z = y_{k+1}$, $Z = Y_{k+1}$. We prove $\Pr[Z = a] \geq \Pr[Z = b]$ assuming $|b| = |a| + 1$. In virtue of symmetry, $\Pr[Z = c] = \Pr[Z = -c]$ for any c; hence we may suppose that $a \geq 0$ and $b = a + 1$.

$$\begin{aligned} P[Z = a] &= \textstyle\sum \{P[z = i] \times P[Y = a - i]: -r \leq i \leq r\} \\ &= (1/(2r + 1)) \times \textstyle\sum \{P[Y = a - i]: -r \leq i \leq r\}. \end{aligned}$$

Similarly, $P[Z = a + 1] = (1/(2r + 1)) \times \sum \{P[Y = a + 1 - i]: -r \leq i \leq r\}$. Then $(2r + 1) - (P[Z = a] \times P[Z = a + 1])$ equals $P[Y = a - r] - P[Y = a + 1 + r]$, which is nonnegative since $|a - r| \leq \max\{a, r\} < a + 1 + r$. $\square$

PROOF OF THEOREM 3.1. Let $f(x) = [\text{bottom}_J(x) + \text{top}_J(x)]/2$. Note that $\lfloor f(x) \rfloor > \text{bottom}_J(x)$ for all x. Since $n - m$ is even, every $f'(x) = f(x) - (n - 1)/2$ is integer. Every $f(x)$ has less than n possible values; hence $\sum f'(x)$ has fewer than ln possible values. By Lemma 3.1,

$$\Pr[\textstyle\sum f(x) = l(n - 1)/2] = \Pr[\textstyle\sum f'(x) = 0] > (ln)^{-1}.$$

If m is odd, then f belongs to J, and we have finished. Suppose that m is even. Set $g(x) = \lfloor f(x) \rfloor$ for $x < \lfloor l/2 \rfloor$, and $g(x) = \lceil f(x) \rceil$ for $\lfloor l/2 \rfloor \leq x < l$. Then g belongs to J. If l is even, then $\sum g(x) = \sum f(x)$; hence,

$$\Pr[\textstyle\sum f(x) = l(n - 1)/2] > (ln)^{-1}.$$

If l is odd, then

$$\sum g(x) = \frac{1}{2} + \sum f(x),$$

$$\left\lceil \frac{l(n-1)}{2} \right\rceil = \frac{1}{2} + \frac{l(n-1)}{2},$$

$$\Pr\left[\sum g(x) = \left\lceil \frac{l(n-1)}{2} \right\rceil\right] = \Pr\left[\sum f(x) = \frac{l(n-1)}{2}\right] > (ln)^{-1}. \qquad \square$$

4. *A Probabilistic Property of Positive Borel Sets*

In this section we complete the proof of the Main Theorem. We start with some notation and terminology.

Recall that a subset S of an arbitrary lattice L is called a *semifilter* if $S \ni x < y \rightarrow y \in S$. The union as well as the intersection of a family of semifilters is again a semifilter. If S is a semifilter in L and I is an interval of L, then $I \cap S$ is a semifilter in the sublattice I.

Every semifilter S of a regular lattice L is the union of intervals $[h, \text{top}_L]$, where h is a minimal member of S. Intuitively, S is rich if it has many minimal elements. The richness is related to the number of elements of the source of L in which some minimal member of S exceeds bottom_L. This motivates the following definition.

Definition. Let $S \subseteq L$ be a semifilter in a regular lattice L with a source W. The set $\{u \in W: \text{there is a minimal } h \in S \text{ with } h(u) > \text{bottom}_L(u)\}$ is the *base* of S in

L, written $\text{base}_L(S)$. The cardinality of $\text{base}_L(S)$ is the *dimension* of S in L, written $\dim_L(S)$.

Note that $\dim_L(S)$ equals 0 if and only if $S = \emptyset$ or $S = L$.

LEMMA 4.1. *Let F be $\bigcup$ or $\bigcap$ (so that in each lattice, F assigns a semifilter to each family of semifilters). Let L be a regular lattice with a source W, and N be a family of semifilters in L. Then*

(i) *for every interval J of L, $J \cap F(N) = F(\{J \cap S : S \in N\})$;*
(ii) *if $\dim_L(S) \le d \le |N|$ for all $S \in N$, then $F(N) = F\{(N_U : U \subseteq W$ and $|U| = d\}$, where $N_U = F\{S \in N : \text{base}_L(S) \subseteq U\}$; and*
(iii) *if $\text{base}_L(S) \subseteq U \subseteq W$ for all $S \in N$, then $\text{base}_L(F(N)) \subseteq U$.*

Comment. This simple lemma will allow us later (in the proof of Theorem 4.3) to treat the two operations $\bigcup$ and $\bigcap$ simultaneously.

PROOF. We only prove statement (iii) for the case of $\bigcap$; the rest is obvious. Let g be a minimal function in $\bigcap N$. Consider the collection X of functions $f \le g$ such that f is minimal in some $S \in N$, and let h be the joint of X. Obviously, $h \in \bigcap N$ and $h \le g$. By the minimality of g, $h = g$. Every $f \in X$ coincides with the bottom of I on $W - U$. Hence g coincides with the bottom of L on $W - U$. □

Definition. Let L be a regular lattice of height n, and $0 < \alpha < 1$. An *α-interval* of L is a regular interval of L of height $\lfloor n^\alpha \rfloor$.

LEMMA 4.2. *Let L be a regular lattice with a source W and the height $n > (2\,|W|)^4$. Let t be a positive integer, and α, β, γ be positive reals such that $\alpha \le 1/(4t)$, and β, $\gamma < 1$, and $\lfloor \lfloor n^\alpha \rfloor^\beta \rfloor = \lfloor n^\gamma \rfloor$. There is a set S of γ-intervals of L satisfying the following condition. Suppose that A is a random α-interval of L, B is a random β-interval of A, and C is a random γ-interval of L. Then*

(*i*) *$Pr[C \notin S] < n^{-t}$, and*
(*ii*) *for every $D \in S$, $Pr[C = D] < \sqrt{n} \times Pr[B = D]$.*

Comment. B is a random γ-interval of L with respect to a complicated probability distribution. Think about t as a big enough number. The lemma says that almost everywhere the complicated distribution is bounded by the uniform distribution times $\sqrt{n}$.

PROOF. Without loss of generality, the bottom and the top of L are $\lambda u.0$ and $\lambda u.(n-1)$, respectively. Let $a = \lfloor n^\alpha \rfloor$, $b = \lfloor a^\beta \rfloor$, $f = \text{bottom}_A$, $g = \text{bottom}_B$, and $h = \text{bottom}_C$. Both B and C have the height b.

For each u in W, $h(u)$ is a random element of $[0, n-b]$; $\Pr[h(u) = k] = 1/(n - b + 1)$ for each $0 \le k \le n - b$. The probability distribution of $g(u)$ is different. Let $0 \le k \le n - b$. $\Pr[g(u) = k] = p(k) \times [1/(a - b + 1)]$ where $p(k) = \Pr[f(u) \le k \le f(u) + a - b]$. If $a - b \le k \le n - a$, then $p(k) = (a - b + 1)/(n - a + 1)$ and $\Pr[g(u) = k] = 1/(n - a + 1) \ge \Pr[h(u) = k]$. In any case $p(k) \ge 1/(n - a + 1)$ and

$$\Pr[g(u) = k] \ge \frac{1}{(n-a+1)(a-b+1)} \ge a^{-1} \times \Pr[h(u) = k].$$

Let S be the set of γ-intervals D of L such that

$$|\{u : \text{either bottom}_D(u) < a - b \text{ or bottom}_D(u) > n - a\}| < 2t.$$

To prove (i), set $V = \{u : \text{either } h(u) < a - b \text{ or } h(u) > n - a\}$. For every $u \in W$, $\Pr[u \in V] < 2a/n$. If U is a subset of W of cardinality $2t$, then $\Pr[U \subseteq V] <$

$(2a/n)^{2t}$. Thus $\Pr[C \notin S] = \Pr[\text{There is } U \subseteq V, |U| = 2t] < |W|^{2t} \times (2a/n)^{2t} \le [|W|^2 \times 2^2 \times n^{2\alpha} \times n^{-2}]^t < [n^{1/2+1/2-2}]^t < n^{-t}$.

To prove (ii), let $D \in S$, $r = \text{bottom}_D$ and $V = \{u : r(u) < a - b \text{ or } r(u) > n - a\}$. Then $\Pr[C = D] = \Pr[h = r] = \prod \{\Pr[h(u) = r(u)] : u \notin V\} \times \prod \{\Pr[h(u) = r(u)] : u \in V\} < \prod \{\Pr[g(u) = r(u)] : u \notin V\} \times (a^{2t} \prod \{\Pr[g(u) = r(u)] : u \in V\}) < \Pr[g = r] \times n^{1/2}$. $\square$

THEOREM 4.1. *Suppose that L is a regular lattice with a source W and the height n, S is a semifilter in L of dimension $d > 0$, $b \in B = \text{base}_L(S)$, V is a subset of W such that $B - V = \{b\}$, ϵ is a positive real with $n^\epsilon < n/2$, F is a member of L such that $F(v) + n^\epsilon \le \text{top}_L(v)$ for all $v \in V$, and J is a random ϵ-interval of L with $\text{bottom}_J(v) = F(v)$ for all $v \in V$. Then $\Pr[b \in \text{base}_J(J \cap S)] < 2n^{\epsilon d - 1}$.*

Comment. The intuitive meaning of Theorem 4.1 is as follows: If the dimension d of the base of S is fixed and the height ϵ of J diminishes, then the probability of $b \in \text{base}_J(J \cap S)$ is rapidly diminishing. Why is this true? Consider the most interesting case $V = W - \{b\}$. Then the only random part of J is the interval $[\text{bottom}_J(b), \text{top}_J(b)]$ of integers; the rest of J is fixed. If $b \in \text{base}_J(J \cap S)$, then there is a minimal function g of $J \cap S$ with $g(b) > \text{bottom}_J(b)$. There are not too many choices for the restriction of g to $A = B - \{b\} \subseteq V$, so consider the case in which $g \mid A$ is fixed. How many choices do we have for $g(b)$ now? The answer is: just one.

PROOF. We may assume that $V = W - \{b\}$. For, let $U = (W - \{b\}) - V$ and r range over functions from U to the set ω of natural numbers such that $r(u) + n^\epsilon \le \text{top}_L(u)$ for $u \in U$. If every conditional probability

$$\Pr[b \in \text{base}_J(J \cap S) \mid \text{bottom}_J \text{ coincides with } r \text{ on } U]$$

is less than $2n^{\epsilon d - 1}$, then $\Pr[b \in \text{base}_J(J \cap S)] < 2n^{\epsilon d - 1}$.

We prove that the probability of a consequence of the event $b \in \text{base}_J(J \cap S)$ is less than $2n^{\epsilon d - 1}$. Let $A = B - \{b\}$,

$$H = \{h \in [A \to \omega] : F(a) \le h(a) \le F(a) + n^\epsilon \text{ for all } a \in A\},$$
$$H' = \{h \in H : \text{there is } f \in S \text{ with } f(a) \le h(a) \text{ for all } a \in A\},$$
$$h^* = \min\{f(b) : f \text{ is a minimal member of } S \text{ with } f(a) \le h(a) \text{ for } a \in A\},$$

for each $h \in H'$, and let E be the event "There is $h \in H'$ with $\text{bottom}_J(b) < h^* \le \text{top}_J(b)$."

E is a consequence of the event $b \in \text{base}_J(J \cap S)$. For, suppose $b \in \text{base}_J(J \cap S)$. Then there is a minimal member g of $J \cap S$ with $g(b) > \text{bottom}_J(b)$. Let h be the restriction of the function g to A; obviously, $h \in H'$. It suffices to prove that $h^* = g(b)$. There is a minimal member f of S such that $f(a) \le h(a)$ for $a \in A$ and $f(b) = h^*$. By the definition of h^*, $f(b) \le g(b)$. Since B is the base of S in L, we have $f(u) = \text{bottom}_L(u)$ for $u \in W - B$. Thus $f \le g$. Let f^J be the join of f and bottom_J; obviously $f^J \le g$ and $f^J \in J \cap S$. By the minimality of g, $g \le f^J$; hence $g = f^J$ and $g(b) = f^J(b) = h^*$.

It remains to estimate $\Pr[E]$. The number of values for J is $n - \lfloor n^\epsilon \rfloor + 1 > n - n^\epsilon > n/2$. $|H'| \le |H| = \lfloor n^\epsilon \rfloor^{d-1} \le n^{\epsilon(d-1)}$. For each $h \in H'$, the number of values for J with $\text{bottom}_J(b) < h^* \le \text{top}_J(b)$ is less than n^ϵ. Thus $\Pr[E] < n^{\epsilon(d-1)} \times n^\epsilon \div (n/2) = 2n^{\epsilon d - 1}$. $\square$

THEOREM 4.2. *Suppose that A is a regular lattice with a source W; $l = |W| \ge 3$, $a = \text{height}(A)$, for each $U \subseteq W$ of cardinality d, $S(U)$ is a semifilter in A with*

$base_A(S(U)) \subseteq U$; $\beta \le 1/(2d)$; B is a random β-interval of A; $M = \bigcup \{base_B(B \cap S(U)) : U \subseteq W, |U| = d\}$; and m is a positive integer. Then $Pr(|M| \ge dm) < l^{2dm}a^{-m/2}$.

Comment. The interesting case is when d, l, m are fixed and a increases. The intuitive meaning of the theorem is that the dimension of $B \cap S(U)$ in B is, in general, small.

PROOF. We evaluate the probability of a consequence of $|M| \ge dm$.

Let $Q = \{(U_1, \ldots, U_m, u_1, \ldots, u_m) : u_i \in U_i \subseteq W \text{ and } |U_i| = d \text{ for all } i\}$. Then $|Q| \le (l^d)^m \times d^m$. Since $(2d)^{1/d} < \max\{(2x)^{1/x} : 1 \le x < \infty\} < 3 \le l$, we have $d < l^d/2$ and $|Q| < (l^{2d}/2)^m$.

Let $R = \{(U_1, \ldots, U_m, u_1, \ldots, u_m) \in Q$: each $u_i \in \text{base}_B(B \cap S(U_i)) - \bigcup \{U_j : j < i\}\}$. The event $R \ne \varnothing$ is a consequence of the event $|M| \ge dm$. For, assume $|M| \ge dm$ and pick any $u_1 \in M$; there is $U_1 \subseteq W$ of cardinality d such that $u_1 \in \text{base}_B(B \cap S(U_1))$. Suppose that $i \le m$, and $U_1, \ldots, U_{i-1}$, $u_1, \ldots, u_{i-1}$ have been chosen. Pick any $u_i \in M - \bigcup \{U_j : j < i\}$; there is $U_i \subseteq W$ of power d such that $u_i \in \text{base}_B(B \cap S(U_i))$.

For each sequence $s = (U_1, \ldots, U_m, u_1, \ldots, u_m)$ in Q, let $p(s)$ be the probability that $s \in R$. If there are $j < i$ with $u_i \in U_j$, then $p(s) = 0$. We suppose that there is no pair $j < i$ with $u_i \in U_j$ and give an upper bound for $p(s)$.

Let $f = \text{bottom}_A$, $g = \text{bottom}_B$ and $b = \lfloor a^\beta \rfloor$. For each $u \in W$, $g(u)$ belongs to the interval $[f(u), f(u) + a - b]$. Let $W_0 = W - \bigcup \{U_i : 1 \le i \le m\}$; and for each $i = 1, \ldots, m$, let $W_i = (U_i - \bigcup_{j<i} U_j) - \{u_i\}$. Without loss of generality, the function g is produced by the following $1 + 2m$ independent choices: first choose the restriction of g to W_0, then choose the restriction of g to W_1, then choose $g(u_1)$, then choose the restriction of g to W_2, then choose $g(u_2)$, etc. By Theorem 4.1 (with $b = u_i$), $p(s) \le (2a^{\beta d - 1})^m \le (2a^{-1/2})^m$.

Clearly, $\Pr[R \ne \varnothing] \le |Q| \times \max\{p(s) : s \in Q\}$. But $|Q| - \max p(s) < (l^{2d}/2)^m - (2a^{-1/2})^m = l^{2dm}a^{-m/2}$. □

THEOREM 4.3. *Suppose that L is a regular lattice with a source W; $l = |W| \ge 3$; $n = height(L)$; N is a family of ln semifilters in L; Y is either $\bigcup N$ or $\bigcap N$; t is a positive integer; $\alpha \le 1/(4t)$; A is a random α-interval of L; for every $X \in N$ and some positive integer d, $Pr[dim_A(A \cap X) > d) < n^{-t}$; $\gamma \le \alpha/(3d)$; C is a random γ-interval of L; $m \ge (2t + 2)/\alpha$ is an integer; $n > 16l^4$; $n \ge l^{2dm}$. Then $Pr[dim_C(C \cap Y) \ge dm) < n^{-t+2}$.*

Comment. The intuitive meaning of the theorem is as follows. A relatively large family of semifilters is given. With a great probability, the dimension of each of these semifilters is bounded (by a given number) in a random α-interval. Then, with a great probability, the dimension of the union, as well as the intersection of all these semifilters, is bounded in a random γ-interval, provided γ is small enough.

PROOF. Let $a = \lfloor n^\alpha \rfloor$, $c = \lfloor n^\gamma \rfloor$, $\beta = \log_a n^\gamma$, and B be a random β-interval of A. Then $\beta \le 1/(2d)$ since $x^{3/2} > x + 1$ for $x \ge 3$, and, therefore, $a^{3/2} > a + 1 \ge n^\alpha$, $a \ge n^{2\alpha/3} \ge n^{2\gamma d} = a^{2\beta d}$. The proof uses two claims.

CLAIM 1. *The conditional probability*

$$Pr[dim_B(B \cap Y) \ge dm \mid \text{every } dim_A(A \cap X) \le d]$$

is less than n^{-t}.

PROOF OF CLAIM 1. Let F be $\bigcup$ or $\bigcap$, so that $Y = F(N)$. Let $N' = \{A \cap X: X \in N\}$ and $Y' = F(N')$. By Lemma 4.1, $Y' = A \cap Y$; hence, $B \cap Y = B \cap Y'$.

We are going to use Theorem 4.2. For every $U \subseteq W$, let $S(U) = F\{X: X \in N'$ and $\text{base}_A X \subseteq U\}$. Let $M = \bigcup \{\text{base}_B(B \cap S(U)): U \subseteq W, |U| = d\}$. If $\dim_A X \leq d$ for $X \in N'$, then, by Lemma 4.1, $Y' = F\{S(U): U \subseteq W, |U| = d\}$ and M includes $\text{base}_B(B \cap Y')$. But, by Theorem 4.2, $\Pr[|M| \geq dm] < l^{2dm} a^{-m/2} \leq n^{1-\alpha m/2} \leq n^{-t}$. Claim 1 is proved. □

CLAIM 2. *$Pr[dim_B(B \cap Y) \geq dm] < ln^{-t+1} + n^{-t}$.*

PROOF OF CLAIM 2. Let E be the event "There is $X \in N$ with $\dim_A(A \cap X) > d$," and E' be the complement of E. $\Pr[E] \leq ln \times \max\{\Pr[\dim_A(A \cap X) > d]: X \in N\} < ln \times n^{-t} = ln^{-t+1}$. Hence

$$\begin{aligned}\Pr[\dim_B(B \cap Y) \geq dm] &\leq \Pr[E] + \Pr[E' \text{ and } \dim_B(B \cap Y) \geq dm] \\ &\leq \Pr[E] + \Pr[E'] \times \Pr[\dim_B(B \cap Y) \geq dm \mid E'] \\ &= ln^{-t+1} + (1 - ln^{-t+1}) \times n^{-t} < ln^{-t+1} + n^{-t} \quad \text{(by Claim 1).}\end{aligned}$$

Claim 2 is proved. □

Now we are ready to prove Theorem 4.3. Note that the height of B equals that of C. By Lemma 4.2, there is a set S of γ-intervals of L such that $\Pr[C \notin S] < n^{-t}$ and $\Pr[C = D] < \sqrt{n} \times \Pr[B = D]$ for all $D \in S$. Thus,

$$\begin{aligned}&\Pr[\dim_C(C \cap Y) \geq dm] \\ &\quad = \sum \{\Pr[C = D]: D \text{ is a } \gamma\text{-interval of } L \text{ and } \dim_D(D \cap Y) \geq dm\} \\ &\quad \leq \sum \{\Pr[C = D]: D \in S \text{ and } \dim_D(D \cap Y) \geq dm\} + \sum \{\Pr[C = D]: D \notin S\} \\ &\quad < \sum \{\sqrt{n} \times \Pr[B = D]: D \in S \text{ and } \dim_D(D \cap Y) \geq dm\} + \Pr[C \notin S] \\ &\quad < n^{1/2} \times \Pr[\dim_B(B \cap Y) \geq dm] + n^{-t} \\ &\quad < ln^{-t+3/2} + n^{-t+1/2} + n^{-t} < n^{-t+2}.\end{aligned}$$

□

We turn now to some specific regular lattices. Recall the structures S_n constructed in Section 2. The universe of S_n consists of pairs (x, y) of natural numbers such that $0 \leq x < l = \lfloor \log_2 n \rfloor$ and $0 \leq y < n$. The subsets of S_n (i.e., the points over S_n) form a lattice with respect to the inclusion relation. With each subset P of S_n we associate a function $P^*(x) = \max(\{0\} \cup \{y: (x, y) \in P\})$ where $0 \leq x < l$. With each point-set M over S we associate $M^* = \{P^*: P \in M\}$. The associated functions form a regular lattice L_n with the source $W = [0, l)$; the height of L_n equals n. The mapping $P \mapsto P^*$ preserves the order: If $P \subseteq Q$, then $P^* \leq Q^*$. If M is a semifilter in the lattice of points over S, then M^* is a semifilter in L_n.

LEMMA 4.3. *Let M and N be semifilters over S_n. Then*

$$(M \cup N)^* = M^* \cup N^* \qquad \textit{and} \qquad (M \cap N)^* = M^* \cap N^*.$$

PROOF. We prove only the inclusion $M^* \cap N^* \subseteq (M \cap N)^*$; the rest is obvious. Suppose that $f \in M^* \cap N^*$, and let P be the subset $\{(x, y): y \leq f(x)\}$. Then $P^* = f$ and $P \in M \cap N$. □

THEOREM 4.4. *Let M be a positive Borel point-set over S_n of level i, and t be a positive integer. There exist a positive integer d and a positive real $\alpha < 1$, both dependent only on i and t, such that if $n \geq 16l^4$, $n \geq l^{2d}$ and if J is a random β-interval of L_n with $\beta \leq \alpha$, then $Pr[dim_J(J \cap M^*) > d] < n^{-t}$.*

PROOF (by induction on i). The case $i = 0$ is trivial. Suppose that $i > 0$ and the lemma is proved for $i - 1$. By Lemma 4.3, M^* is either the union or the intersection of a family $\{X_j: j = 0, \ldots, ln - 1\}$ of positive Borel point-set of level i.

Let $t' = t + 2$. By the induction hypothesis there are d' and α' such that, if $n \geq 16l^4$, $n \geq l^{2d'}$, and J is a random β-interval of L_n with $\beta \leq \alpha'$, then $\Pr[\dim_J(J \cap M_j^*) > d'] < n^{-t'}$ for each j.

Without loss of generality, $\alpha' \leq 1/(4t')$. Set $m = (2t' + 2)/\alpha'$, $d = d'm + 1$, and $\alpha = \alpha'/(3d)$. If $n \geq 16l^4$, $n \geq l^{2d}$ and if J is a random β-interval of L_n with $\beta \leq \alpha$, then (by Theorem 4.3) $\Pr[\dim_J(J \cap M^*) > d] < n^{-t}$. □

Now we are ready to finish the proof of the Main Theorem. The desired structures S_n and formulas φ_0, $\varphi_1(P)$ were constructed in Section 2. Statements (i) and (ii) of the Main Theorem were proved in Section 2. It remains to prove statement (iii). By contradiction, suppose that statement (iii) is false. Pick a natural number i that witnesses the failure of (iii). Let $t = 2$, and let d and α be as in Theorem 4.4. Since statement (iii) fails, there is a positive integer n such that $l = \lfloor \log_2 n \rfloor > d$, $n \geq 16l$, $n \geq l^{2d}$ and the point set $M = \{P \subseteq S_n : S_n \models \varphi_1(P)\}$ over S_n is positive Borel of level i. Let L_n, M^*, and J be as in Theorem 4.4. By Theorem 4.4, $\Pr[\dim_J(J \cap M^*) > d] < n^{-2}$. But, by the Theorem 3.1, $\Pr[\dim_J(J \cap M^*) = l] > (ln)^{-1}$. The two estimations contradict each other. The Main Theorem is proved. □

ACKNOWLEDGMENT. The authors would like to thank the referees for their very thorough review of this paper.

REFERENCES

1. AJTAI, M. Σ_1-formulae on finite structures. *Ann. Pure Appl. Logic 24* (1983), 1–48.
2. BOPPANA, R. H. Threshold functions and bounded depth monotone circuits. In *Proceedings of the 16th ACM Symposium on Theory of Computing* (Washington, D.C., Apr. 30–May 2). ACM, New York, 1984, pp. 475–479.
3. CHANDRA, A. K., AND HAREL, D. Structure and complexity of relational queries. *J. Comput. Syst. Sci. 25* (1980), 156–178.
4. CHANG, C. C., AND KEISLER, H. J. *Model Theory.* North-Holland, New York, 1977.
5. FURST, M., SAXE, J., AND SIPSER, M. Parity, circuits, and the polynomial time hierarchy. *Math. Syst. Theory 17* (1984), 13–27.
6. GUREVICH, Y. Toward logic tailored for computational complexity. In *Computation and Proof Theory,* M. M. Richter, E. Börger, W. Oberschelp, B. Schinzel, and W. Thomas, Eds. Lecture Notes in Mathematics, vol. 1104. Springer-Verlag, New York, 1984, pp. 175–216.
7. GUREVICH, Y. Logic and the challenge of computer science. In *Current Trends in Theoretical Computer Science,* E. Börger, Ed. Computer Science Press, Rockville, Md., 1987, pp. 1–57.
8. GUREVICH, Y., AND SHELAH, S. Fixed point extensions of first order logic. In *Proceedings of the 26th IEEE Symposium on Foundations of Computer Science.* IEEE, New York, 1985, pp. 346–353.
9. HARDY, G. H., AND WRIGHT, E. M. *An Introduction to the Theory of Numbers.* Clarendon Press, Oxford, England, 1971.
10. KLEENE, S. C. *Introduction to Metamathematics.* Van Nostrand, New York, 1952.
11. LYNDON, R. C. An interpolation theorem in the predicate calculus. *Pacific J. Math. 9* (1959), 155–164.
12. SIPSER, M. Borel sets and circuit complexity. In *Proceedings of the 15th ACM Symposium on Theory of Computing.* ACM New York, 1983, pp. 61–69.

RECEIVED FEBRUARY 1985; REVISED JUNE 1986; ACCEPTED JULY 1986