VAPNIK–CHERVONENKIS CLASSES OF DEFINABLE SETS

MICHAEL C. LASKOWSKI

ABSTRACT

We show that a class of subsets of a structure uniformly definable by a first-order formula is a Vapnik–Chervonenkis class if and only if the formula does not have the independence property. Via this connection we obtain several new examples of Vapnik–Chervonenkis classes, including sets of positivity of finitely subanalytic functions.

Introduction

Bernoulli's theorem states that the relative frequency of an event A in a sequence of independent trials converges to the probability of the event. Vapnik and Chervonenkis [**19**] gave a sufficient condition on a class $\mathscr{C}$ of events on the same space so that the maximum difference over the class between the relative frequency of an event in the class and its probability still converges to zero.

Given a set X and a finite subset F, a collection $\mathscr{C}$ of subsets of X *shatters* F if for every subset F' of F there is $C \in \mathscr{C}$ with $F' = F \cap C$. We call $\mathscr{C}$ a *Vapnik–Chervonenkis* class (or VC class) if there is an n such that no subset of X of size n is shattered by $\mathscr{C}$. If $\mathscr{C}$ is a VC class, its Vapnik–Chervonenkis dimension $\mathscr{V}(\mathscr{C})$ is the least such n.

Suppose that X is a space, P_X is a probability measure on X, and $\mathscr{C}$ is a family of events (measurable subsets of X). For each k endow X^k with the measure P_X^k of k independent trials with underlying distribution P_X. For each event $A \in \mathscr{C}$ let the relative frequency of A, RF_A^k, be a map from X^k to $\mathbb{R}$ defined by $(1/k)\sum_{i=1}^{k}\chi_A(x_i)$, where χ_A is the characteristic function of A. Vapnik and Chervonenkis showed that if $\mathscr{C}$ is a VC class then, for each $\varepsilon > 0$,

$$\lim_{k\to\infty} P_X^k(\sup\{|RF_A^k - P_X(A)| : A \in \mathscr{C}\} > \varepsilon) = 0.$$

Since their introduction, many properties of VC classes and further applications to multivariate non-parametric statistics have been found by Dudley [**5, 6, 7**], Dudley and Philipp [**8**], Pollard [**14**], and Assouad [**1**].

In light of their utility, a considerable amount of research has involved identifying natural VC classes of subsets of spaces. Many of the early examples of VC classes are due to Dudley [**5, 6**], who showed that the set of rectangles and the sets of positivity of polynomials of bounded degree are VC classes of subsets of $\mathbb{R}^n$. In this paper we expand on the work of Stengle and Yukich [**17**] and investigate whether families of subsets of a structure uniformly definable by a first-order formula form a VC class (see Section 1 for definitions).

Received 19 June 1990; revised 21 November 1990.

1991 *Mathematics Subject Classification* 60D05.

Research supported by NSF grant DMS #8505550.

Let $\mathscr{A}$ be a structure in a language L and let

$$\phi(\bar{x},\bar{y}) := \phi(x_1,\ldots,x_m,y_1,\ldots,y_k)$$

be a first-order formula (see Section 1 for definitions). For $\bar{b}\in A^k$ let

$$\phi(A^m,\bar{b}) = \{\bar{c}\in A^m : \mathscr{A}\models\phi[\bar{c},\bar{b}]\}$$

and let $\mathscr{C}_\phi = \{\phi(A^m,\bar{b}):\bar{b}\in A^k\}$.

The key observation of this paper (Proposition 1.3) is that $\mathscr{C}_\phi$ is a VC class of subsets of A^m if and only if the formula ϕ fails to have the independence property (see Definition 1.1). We also give a combinatorial proof of a theorem of Shelah (Corollary 2.5) that a structure has the independence property if and only if some formula has the independence property. Consequently, for a given structure $\mathscr{A}$, every $\mathscr{C}_\phi$ forms a VC class if and only if $\mathscr{A}$ fails to have the independence property. The independence property for a formula or a structure was defined by Shelah [**15**] and has become one of the central notions of model theory. The abstract properties of the independence property have been extensively investigated by Shelah [**15, 16**] and Poizat [**12, 13**].

Using the connection above as a bridge, we are able to exhibit numerous examples of VC classes of definable subsets. Most notably, if a structure is either o-minimal (see Definition 3.1) or is stable then every $\mathscr{C}_\phi$ is a VC class. As an example,

$$\{\{\bar{x}\in I^m : f(\bar{x},\bar{a}) > 0\}:\bar{a}\in I^n\}$$

is a VC class for any real analytic function $f: I^m\times I^n\to\mathbb{R}$, where I denotes $[0,1]$. This observation answers a question of Stengle and Yukich [**17**].

Also, it follows from Corollary 2.5 that every $\mathscr{C}_{\phi(\bar{x},\bar{y})}$ is a VC class if and only if every $\mathscr{C}_{\psi(x,\bar{y})}$ (with a single x-variable) is a VC class and this is the case if and only if every $\mathscr{C}_{\theta(\bar{x},y)}$ (with a single y-variable) is a VC class.

1. *The correspondence*

We begin this section with a very rough outline of the basic definitions of first-order logic. Any reader unfamiliar with the definitions given below is referred to [**4**, Sections 1.3 and 1.4] for examples and a more thorough account.

A language L is a set of function, relation and constant symbols. Given a language L, the set of (first-order) L-formulas is defined inductively. The set of L-terms consists of the constants in L and an infinite set of variables, and is closed under applications of the L-function symbols. The L-atomic formulas are simply $t_1 = t_2$ and $R(t_1,\ldots,t_n)$, where each t_i is a term and R is a relation symbol in L. The set of L-formulas is the closure of the set of L-atomic formulas under conjunction, negation and quantification over *elements* (as opposed to subsets or the natural numbers).

An L-structure is a non-empty set A, together with interpretations of every function, relation and constant symbol in L. Given an L-structure $\mathscr{A}$, an L-formula $\phi(x_1,\ldots,x_n)$ (that is, all of the free variables of ϕ are among $x_1,\ldots,x_n$), and $\langle a_1,\ldots,a_n\rangle\in A^n$, the notion '$\phi[a_1,\ldots,a_n]$ is true in $\mathscr{A}$', denoted by $\mathscr{A}\models\phi[a_1,\ldots,a_n]$, is defined by induction on the complexity of ϕ in the obvious way.

Thus, to each formula $\phi(x_1,\ldots,x_n)$ there is a corresponding subset

$$\phi(A^n) = \{\langle a_1,\ldots,a_n\rangle : \mathscr{A}\models\phi[a_1,\ldots,a_n]\}.$$

The collection of all such subsets forms the set of *definable* subsets of $\mathscr{A}$. Under this identification, the operations of conjunction, negation and existential quantification of formulas correspond to intersection, complementation and projection of the respective definable subsets.

We shall adhere to the standard practice among logicians of identifying every non-negative integer n with the set $\{0, \ldots, n-1\}$.

DEFINITION 1.1. A formula $\phi(x;y) := \phi(x_1, \ldots, x_m; y_1, \ldots, y_k)$ has the independence property with respect to $\mathscr{A}$ if for every n there is a sequence $\langle a_i : i \in n \rangle$ of elements from A^k so that for every $w \subseteq n$ there is a $c_w \in A^m$ with

$$\mathscr{A} \models \phi[c_w, a_i] \quad \text{if and only if } i \in w.$$

If ϕ does not have the independence property, its *independence dimension* $\mathscr{I}(\phi)$ is the least n so that the above fails. If ϕ has the independence property we say $\mathscr{I}(\phi) = \infty$.

As the definition of a formula having the independence property is dependent on the given partition of the free variables we use a semicolon (;) to denote the partition. For $\phi(x;y)$ a formula let $\psi(y;x) := \phi(x;y)$ be its dual. That is, ϕ and ψ are identical as formulas, but the roles of x and y are reversed in the definition of the independence property.

DEFINITION 1.2. A structure $\mathscr{A}$ has the independence property if there is a formula $\phi(x;y)$ with only a single x-variable having the independence property with respect to $\mathscr{A}$.

The goal of Section 2 will be to prove that this definition is equivalent to the more natural statement that some formula (with no restriction on the number of x-variables) has the independence property with respect to $\mathscr{A}$.

The following proposition establishes the connection between the independence property and VC classes.

PROPOSITION 1.3. *Let $\mathscr{A}$ be a structure and $\phi(x;y)$ a formula. Then $\mathscr{C}_\phi$ is a VC class if and only if ϕ does not have the independence property. If $\mathscr{V}(\mathscr{C}_\phi) = d$ and $\mathscr{I}(\phi) = n$ then $n \leqslant 2^d$ and $d \leqslant 2^n$.*

The proposition follows immediately from the two lemmas below.

LEMMA 1.4. *Let $\mathscr{A}$ be a structure, $\phi(x;y)$ a formula and $\psi(y;x)$ its dual. Then $\mathscr{V}(\mathscr{C}_\phi) \leqslant d$ if and only if $\mathscr{I}(\psi) \leqslant d$.*

Proof. Clearly $\mathscr{V}(\mathscr{C}_\phi) > d$ if and only if there are $c_0, \ldots, c_{d-1}$ from A^m such that for every $F \subseteq d$ there is $a_F \in A^k$ such that

$$\mathscr{A} \models \phi[c_i, a_F] \quad \text{if and only if } i \in F.$$

However, this is exactly the definition of $\mathscr{I}(\psi) > d$.

The following lemma as stated is due to Poizat [**13**, Theorem 12.16]. However, the analogous statement for arbitrary VC classes is the essence of [**6**, Theorem 9.3.2].

LEMMA 1.5. *Let $\mathscr{A}$ be a structure, $\phi(x;y)$ a formula and $\psi(y;x)$ its dual. Then $\mathscr{I}(\phi) \leqslant n$ implies that $\mathscr{I}(\psi) \leqslant 2^n$.*

Proof. Assume that $\mathscr{I}(\psi) > 2^n$. Then we can find a sequence $\langle \bar{c}_l : l \in \mathscr{P}(n) \rangle$ such that for any $s \subseteq \mathscr{P}(n)$ there is a tuple $\bar{a}_s$ so that

$$\mathscr{A} \models \phi[\bar{c}_l, \bar{a}_s] \quad \text{if and only if } l \in s.$$

Now, for each $i \in n$, let $s_i = \{l \subseteq n : i \in l\}$ and let $\bar{b}_i = \bar{a}_{s_i}$. It is now easy to check that the sequence $\langle \bar{b}_i : i \in n \rangle$ demonstrates that $\mathscr{I}(\phi) > n$.

We remark that the bounds on the dimensions given in Proposition 1.3 are sharp. To see this, let L be a language consisting of a unary predicate U and a binary relation R. For any n, let $p = 2^n - 1$ and let $\mathscr{A}_p$ be an L-structure having p elements $\{c_i : i \in p\}$ for which the predicate U holds and 2^p elements $\{a_s : s \subseteq p\}$ for which U fails. Define the relation R by $R(c_i, a_s)$ if and only if $i \in s$. Let ϕ be the formula $R(x; y)$. It is easy to verify that $\mathscr{V}(\mathscr{C}_\phi) > p$. However, as $p < 2^n$, for any choice $\langle a_i : i \in n \rangle$ from the elements where U fails there must be $w \subseteq n$ such that

$$\{i \in n : \mathscr{A} \models R(c, a_i)\} \neq w$$

for any $c \in A$ where U holds. Thus $\mathscr{I}(\phi) \leqslant n$, so $2^{\mathscr{I}(\phi)} \leqslant \mathscr{V}(\mathscr{C}_\phi)$. Dually, Lemma 1.4 states that $2^{\mathscr{V}(\mathscr{C}_\psi)} \leqslant \mathscr{I}(\psi)$, where ψ is the formula $R(x, y)$, but with the roles of x and y exchanged.

2. *Sufficiency of a single variable*

The equivalence of a structure having the independence property and some formula (with an arbitrary number of x-variables) having the independence property is due to Shelah **[15]**. Shelah's original proof required a theorem of Baumgartner **[2]** that it is consistent with the ZFC axioms of set theory that there is an infinite cardinal κ such that no tree with κ nodes has 2^κ branches. He then proves the equivalence in such a model of set theory. However, from the syntactic form of the equivalence it follows that it is *absolute*, that is it holds in every model of ZFC if and only if it holds in some model of ZFC. One of the drawbacks of Shelah's proof is that if a formula $\phi(\bar{x}, \bar{y})$ has the independence property, it gives no information about *which* formula having only a single x-variable has the independence property. This section is devoted to giving a strictly combinatorial proof of this equivalence.

For all of this section assume that we are working inside a fixed L-structure $\mathscr{A}$ whose underlying set is denoted by A. We shall tacitly assume that any formula mentioned is an L-formula.

Suppose that $\tau(\bar{x}_1, \ldots, \bar{x}_r)$ is a formula with $l(\bar{x}_j) = k$ (that is, $\bar{x}_j$ is a k-tuple) for all $1 \leqslant j \leqslant r$, and $\langle \bar{a}_n : n \in N \rangle$ is a finite sequence of k-tuples from A. We call the sequence *τ-indiscernible* if, for all sequences $i_1 < \ldots < i_r < N$ and $j_1 < \ldots < j_r < N$,

$$\mathscr{A} \models \tau[\bar{a}_{i_1}, \ldots, \bar{a}_{i_r}] \leftrightarrow \tau[\bar{a}_{j_1}, \ldots, \bar{a}_{j_r}].$$

If F is a set of such formulas we call $\langle \bar{a}_n : n \in N \rangle$ *F-indiscernible* if it is τ-indiscernible for all τ in F. The following lemma is simply a restatement of the finite Ramsey theorem (see, for example, Parsons **[10]**).

LEMMA 2.1. *Let $\tau(\bar{x}_1, \ldots, \bar{x}_r)$ be a formula with $l(\bar{x}_j) = k$ for $1 \leqslant j \leqslant r$. For every positive integer p there is an N (depending only on r and p) such that if $\langle \bar{a}_n : n \in N \rangle$ is a sequence of k-tuples from A^k then there is a τ-indiscernible subsequence $\langle \bar{a}_n : n \in J \rangle$ of length at least p.*

LEMMA 2.2. *Assume that a formula* $\phi(\bar{x};\bar{y})$ *has the independence property. Then for any finite set* $F = \{\tau_i(\bar{y}_1, \ldots, \bar{y}_{r_i}) : i \in I\}$ *of formulas with* $l(\bar{y}_j) = k$ *for all* $i \in I$ *and* $1 \leqslant j \leqslant r_i$, *and every p, there is an F-indiscernible sequence* $\langle \bar{a}_n : n \in p \rangle$ *from* A^k *and a tuple* $\bar{b} \in A^m$ *so that* $\mathscr{A} \models \phi[\bar{b}, \bar{a}_n]$ *if and only if n is even.*

Proof. By successively applying Lemma 2.1 for each τ in F there is a number N such that if $\langle \bar{a}_n : n \in N \rangle$ is any sequence of k-tuples from A then there is an F-indiscernible subsequence of length p. As ϕ has the independence property we can find a sequence $\langle \bar{a}_n : n \in N \rangle$ so that

$$\mathscr{A} \models \exists \bar{x}\Big(\bigwedge_{i \in w} \phi(\bar{x}, \bar{a}_i) \wedge \bigwedge_{i \in N \setminus w} \neg\phi(\bar{x}, \bar{a}_i)\Big) \tag{1}$$

for any subset w of N. Now fix an F-indiscernible subsequence $\langle \bar{a}_{i_j} : j \in p \rangle$ and take $\bar{b}$ to be a witness to (1), where $w = \{i_{2j} : 2j \in p\}$.

Given a formula $\phi(x_1, \ldots, x_m, \bar{y})$, a positive integer N and a subset of $w \subseteq N$, let

$$\theta^{\phi}_{w}(\bar{x}, \bar{y}_1, \ldots, \bar{y}_N) \quad \text{be} \quad \bigwedge_{i \in w} \phi(\bar{x}, \bar{y}_i) \wedge \bigwedge_{i \in N \setminus w} \neg\phi(\bar{x}, \bar{y}_i),$$

let

$$\Gamma^{\phi}_{w}(x_1, \bar{y}_1, \ldots, \bar{y}_N) \quad \text{be} \quad \exists x_2 \ldots \exists x_m\, \theta^{\phi}_{w}(x_1, x_2, \ldots, x_m, \bar{y}_1, \ldots, \bar{y}_N),$$

let

$$\rho^{\phi}_{w}(\bar{y}_1, \ldots, \bar{y}_N) \quad \text{be} \quad \exists \bar{x}\theta^{\phi}_{w}(\bar{x}, \bar{y}_1, \ldots, \bar{y}_N),$$

and let $\Delta_N(\phi) = \{\rho^{\phi}_{w} : w \subseteq N\}$.

The following is a converse to Lemma 2.2.

LEMMA 2.3. *Let* $\phi(\bar{x};\bar{y})$ *be a formula with* $\mathscr{I}(\phi) \leqslant N$. *If* $\langle \bar{a}_n : n \in M \rangle$ *is a* $\Delta_N(\phi)$*-indiscernible sequence from* A^k *and* $\bar{b} \in A^m$, *then there do not exist* $i_0 < \ldots < i_{2N-1}$ *so that* $\mathscr{A} \models \phi[\bar{b}, \bar{a}_{i_j}]$ *if and only if j is even.*

Proof. Assume that $\langle \bar{a}_n : n \in M \rangle$ is a $\Delta_N(\phi)$-indiscernible sequence from A^k and there is a sequence of integers $i_0 < \ldots < i_{2N-1}$ and $\bar{b} \in A^m$ so that $\mathscr{A} \models \phi[\bar{b}, \bar{a}_{i_j}]$ if and only if j is even. We claim that the sequence $\langle \bar{a}_i : i < N \rangle$ demonstrates that $\mathscr{I}(\phi) > N$. Indeed, for any $w \subseteq N$, define a function $k_w : \{0, \ldots, N-1\} \to \{i_0, \ldots, i_{2N-1}\}$ by

$$k_w(j) = \begin{cases} i_{2j} & \text{if } j \in w, \\ i_{2j+1} & \text{if } j \notin w. \end{cases}$$

Now $\mathscr{A} \models \theta^{\phi}_{w}[\bar{b}, \bar{a}_{k_w(1)}, \ldots, \bar{a}_{k_w(N)}]$, so $\mathscr{A} \models \rho^{\phi}_{w}[\bar{a}_{k_w(1)}, \ldots, \bar{a}_{k_w(N)}]$. However, as $k_w(1) < \ldots < k_w(N)$ and $\langle \bar{a}_n : n \in M \rangle$ is ρ^{ϕ}_{w}-indiscernible, this implies that $\mathscr{A} \models \rho^{\phi}_{w}[\bar{a}_1, \ldots, \bar{a}_N]$, as desired.

The following is the major theorem of this section. Note that as we identify an integer with the set of elements preceding it, Γ_l is just Γ_{w_l}, where $w_l = \{0, \ldots, l-1\}$ for each $0 \leqslant l \leqslant N$.

THEOREM 2.4. *Assume that* $\phi(x_1, \ldots, x_m;\bar{y})$ *has the independence property. Then either* $\phi^*(x_2, \ldots, x_m; x_1, \bar{y}) := \phi(x_1, \ldots, x_m; \bar{y})$ *has the independence property or, letting* $N = \mathscr{I}(\phi^*)$, *there is an l,* $0 \leqslant l \leqslant N$, *so that* $\Gamma^{\phi}_{l}(x_1; \bar{y}_1, \ldots, \bar{y}_N)$ *has the independence property.*

Proof. Assume that ϕ has the independence property, N is a positive integer and for each l, $0 \leqslant l \leqslant N$, Γ_l^ϕ does not have the independence property (hence no Γ_w^ϕ has the independence property for any $w \subseteq N$). We must show that $\mathscr{I}(\phi^*) > N$.

For each $w \subseteq N$ let $M_w = \mathscr{I}(\Gamma_w^\phi)$ and let M be the maximum of the M_w. Let

$$F_w = \{v_{u,w}(\bar{z}^1, \ldots, z^{M_w}) : u \subseteq M_w\},$$

where z^j is the (Nk)-tuple $\langle y_1^j, \ldots, y_N^j \rangle$ and

$$v_{u,w} \equiv \exists x_1 (\bigwedge_{j \in u} \Gamma_w^\phi(x_1, z^j) \wedge \bigwedge_{j \in M_w \setminus u} \neg \Gamma_w^\phi(x_1, z^j))$$

and let $F = \bigcup \{F_w : w \subseteq N\}$. Note that $v_{u,w} \equiv \rho_u^{\Gamma_w^\phi}$ for all $u \subseteq M_w$.

Now fix an integer $p \geqslant (2M)^{2^N} \cdot 2N$. As ϕ has the independence property, by Lemma 2.2 we can find an F-indiscernible sequence $\langle \bar{a}_n : n \in p \rangle$ from A^k and elements $b \in A$ and $c \in A^{m-1}$ so that $\mathscr{A} \models \phi[b, c, \bar{a}_n]$ if and only if n is even. Thus, $\mathscr{A} \models \phi^*[c, b, \bar{a}_n]$ if and only if n is even. By Lemma 2.3 to show that $\mathscr{I}(\phi^*) > N$ it suffices to show that there is an interval $i_0 \leqslant n < i_0 + 2N$ of length $2N$ such that the sequence $\langle ba_n : i_0 \leqslant n < i_0 + 2N \rangle$ of $(k+1)$-tuples is $\Delta_N(\phi^*)$-indiscernible.

CLAIM. *Suppose that R and Q are integers with $Q \geqslant 2MR$, $i_0 \leqslant n < i_0 + Q$ is an interval of p and $w \subseteq N$. Then there is $j_0, i_0 \leqslant j_0 < i_0 + Q - R$, such that $\langle ba_n : j_0 \leqslant n < j_0 + R \rangle$ is $\rho_w^{\phi^*}$-indiscernible.*

Proof. Fix integers R, Q and a subset w of N. As the element b remains fixed throughout the sequence, merely by reindexing the variables this is equivalent to asking for j_0 so that $\langle a_n : j_0 \leqslant n < j_0 + R \rangle$ is $\Gamma_w^\phi(b, y_1, \ldots, y_N)$-indiscernible. So, assume that there is no such interval. Then certainly for each $j < 2M$ there is an increasing sequence $i_1^j < \ldots < i_N^j$ with $Rj < i_1^j$ and $i_N^j < R(j+1)$ so that $\mathscr{A} \models \Gamma_w^\phi[b, a_{i_1^j}, \ldots, a_{i_N^j}]$ if and only if j is even, lest $Rj \leqslant n < R(j+1)$ be such an interval. Let $C_j = \langle a_{i_1^j}, \ldots, a_{i_N^j} \rangle$ for each $j < 2M$. Now as $\rho_u^{\Gamma_w^\phi} = v_{u,w}$ and $\langle \bar{a}_n : i_0 \leqslant n < i_0 + Q \rangle$ is an F-indiscernible sequence, the underlying order of the $\bar{a}_n$ within an increasing sequence of C_j implies that $\langle C_j : j < 2M \rangle$ is a $\Delta_{M_w}(\Gamma_w^\phi)$-indiscernible sequence. However, as $\mathscr{A} \models \Gamma_w^\phi[b, C_j]$ if and only if j is even, we obtain a contradiction to Lemma 2.3.

Now, from our bound on p, by successively iterating the claim for each subset w of N, we obtain an interval $i_0 \leqslant n < i_0 + 2N$ so that $\langle b\bar{a}_n : i_0 \leqslant n < i_0 + 2N \rangle$ is $\Delta_N(\phi^*)$-indiscernible, as desired.

COROLLARY 2.5. *Let $\mathscr{A}$ be a structure and assume that every formula $\phi(x; y)$ (where $l(x) = 1$) does not have the independence property. Then no formula $\psi(x; z)$ has the independence property.*

Proof. This follows by induction on m using Theorem 2.4.

3. *Examples of VC classes*

DEFINITION 3.1. Let $\mathscr{A}$ be a structure in a language that includes the symbol $<$ and satisfies axioms stating that $<$ is a dense linear order without end-points. $\mathscr{A}$ is called *order-minimal* (or o-minimal) if for any formula $\phi(x; y)$ with a single x-variable and any $b \in A^k$, $\phi(A, b)$ is a finite union of points and intervals with end-points in $A \cup \{-\infty, \infty\}$.

It was shown by Knight, Pillay and Steinhorn [9] that if $\mathscr{A}$ is o-minimal then for any formula $\phi(x;\bar{y})$ with a single x-variable there is a uniform bound on the number of components of $\phi(A,\bar{b})$ as $\bar{b}$ varies. Now if $\mathscr{A}$ is any structure where $<$ is a linear ordering then the formulas $x = y$ and $x < y$ cannot have the independence property. However, for any structure the set of formulas that do not have the independence property is closed under boolean combinations. (This can be seen directly or else one can quote Dudley [6, Theorem 9.2.3] and apply Proposition 1.3.) These three facts imply that no o-minimal structure can have the independence property. Consequently, if $\mathscr{A}$ is o-minimal then $\mathscr{C}_\phi$ is a VC class for every formula ϕ.

Because of this connection, it is of considerable interest here to know what structures are o-minimal. It follows from the Tarski–Seidenberg theorem that $\bar{\mathbb{R}} := \langle \mathbb{R}; +, \cdot, 0, 1, < \rangle$ is o-minimal. We ask what other functions and relations can be added to this structure that preserve o-minimality. One of the central open questions in model theory is whether expanding the language of $\bar{\mathbb{R}}$ by adding the function e^x preserves o-minimality.

As for positive results, the most general result to date on o-minimal expansions of $\bar{\mathbb{R}}$ is given by van den Dries [18]. A set $X \subseteq \mathbb{R}^m$ is *semianalytic at the point* $x \in \mathbb{R}^m$ if there is an open neighbourhood U of x such that $U \cap X$ can be written as a finite union of sets of the form

$$\{y \in U : f(y) = 0, g_1(y) > 0, \ldots, g_n(y) > 0\},$$

where $f, g_1, \ldots, g_n$ are real analytic functions on U. A set $X \subseteq \mathbb{R}^m$ is semianalytic if it is semianalytic at every $x \in \mathbb{R}^m$.

A set $X \subseteq \mathbb{R}^m$ is *subanalytic at the point* $x \in \mathbb{R}^m$ if there is an open neighbourhood U of x and a *bounded* semianalytic set $S \subseteq \mathbb{R}^{m+n}$ for some n, so that $U \cap X = \pi(S)$, where $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^m$ is the projection map onto the first m coordinates. Now $X \subseteq \mathbb{R}^m$ is subanalytic if it is subanalytic at every $x \in \mathbb{R}^m$. The reader is referred to Bierstone and Milman [3] for a thorough treatment of the theory of subanalytic sets.

Following van den Dries, we call a subset X of $\mathbb{R}^m$ *finitely subanalytic* if its image under the semialgebraic map

$$(x_1, \ldots, x_m) \longrightarrow (x_1/(1+x_1^2)^{\frac{1}{2}}, \ldots, x_m/(1+x_m^2)^{\frac{1}{2}})$$

from $\mathbb{R}^m$ to $\mathbb{R}^m$ is a subanalytic subset of $\mathbb{R}^m$.

As examples, it is easily seen that the restriction of a real analytic function (defined on an open subset of $\mathbb{R}^m$) to a compact, subanalytic subset has a finitely subanalytic graph. Further, as both sine and cosine restricted to $(-\pi/2, \pi/2)$ are finitely subanalytic, it follows that the (total) arctangent function is finitely subanalytic.

Van den Dries [18] proves that the expansion of the reals formed by adding an m-place predicate for every finitely subanalytic subset of $\mathbb{R}^m$ for every m remains o-minimal. Consequently, by simply piecing together the facts above, we have the following theorem.

THEOREM 3.2. *Let $\phi(\bar{x};\bar{y})$ be any formula in the language L containing $+$, $\cdot$, $<$, and predicates for all finitely subanalytic sets. Then $\mathscr{C}_\phi = \{\phi(\mathbb{R}^m, \bar{a}) : \bar{a} \in \mathbb{R}^k\}$ is a VC class of subsets of $\mathbb{R}^m$.*

Another large class of structures that do not have the independence property are the *stable* structures (see Shelah [15]). This class includes algebraically closed fields and modules over a ring R in the language containing $+$ and for each $r \in R$ a unary function symbol $\cdot_r$ corresponding to scalar multiplication by r. Thus, for example, given any variety V of $\mathbb{C}^{m+k}$, $\{V_a : \bar{a} \in \mathbb{C}^k\}$ is a VC class of subsets of $\mathbb{C}^m$, where $V_a = \{\bar{b} \in \mathbb{C}^m : (\bar{b}, \bar{a}) \in V\}$.

References

1. P. ASSOUAD, 'Denseté et dimension', *Ann. Inst. Fourier Grenoble* 33 (1983) 233–282.
2. J. BAUMGARTNER, 'Almost disjoint sets, the dense set problem and partition calculus', *Ann. of Math. Logic* 9 (1976) 401–439.
3. E. BIERSTONE and P. D. MILMAN, 'Semianalytic and subanalytic sets', Publications Mathématiques 67 (Institut des Hautes Etudes Scientifiques, Paris, 1988) 5–42.
4. C. C. CHANG and H. J. KEISLER, *Model theory* (North-Holland, Amsterdam, 1973).
5. R. M. DUDLEY, 'Central limit theorems for empirical measures', *Ann. Probab.* 6 (1978) 899 929.
6. R. M. DUDLEY, 'A course on empirical processes', *Ecole d'été de probabilités de Saint-Fleur* XII-1982, Lecture Notes in Mathematics 1097 (Springer, Berlin, 1984) 1–142.
7. R. M. DUDLEY, 'The structure of some Vapnik–Chervonenkis classes', *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer* 2 (ed. L. M. LeCam and R. A. Olshen; Wadsworth, Monterrey, 1985) 495–508.
8. R. M. DUDLEY and W. PHILIPP, 'Invariant principles for sums of Banach space valued random elements and empirical processes', *Z. Wahrsch. Verw. Gebeite* 62 (1983) 509–552.
9. J. KNIGHT, A. PILLAY and C. STEINHORN, 'Definable sets in ordered structures II', *Trans. Amer. Math. Soc.* 295 (1986) 593–605.
10. T. D. PARSONS, 'Ramsey graph theory', *Selected topics in graph theory* (ed. L. Beinecke and R. Wilson; Academic Press, New York, 1978) 361–384.
11. A. PILLAY and C. STEINHORN, 'Definable sets in ordered structures I', *Trans. Amer. Math. Soc.* 295 (1986) 565–605.
12. B. POIZAT, 'Théories instables', *J. Symbolic Logic* 46 (1981) 513–522.
13. B. POIZAT, *Cours de théorie des modèles* (Nur al-Mantiq wal-Ma'rifah, Villeurbanne, 1985).
14. D. POLLARD, *Convergence of stochastic processes* (Springer, Berlin, 1984).
15. S. SHELAH, 'Stability, the f.c.p. and superstability', *Ann. of Math. Logic* 3 (1971) 271–362.
16. S. SHELAH, *Classification theory* (North-Holland, Amsterdam, 1978).
17. G. STENGLE and J. E. YUKICH, 'Some new Vapnik–Chervonenkis classes', *Ann. Statist.* 17 (1989) 1441–1446.
18. L. VAN DEN DRIES, 'A generalization of the Tarski–Seidenberg theorem and some non-definability results', *Bull. Amer. Math. Soc.* 15 (1986) 189–193.
19. V. N. VAPNIK and A. YA. CHERVONENKIS, 'On the uniform convergence of relative frequencies of events to their probabilities', *Theory Probab. Appl.* 16 (1971) 264–280.

Mathematical Sciences Research Institute
1000 Centennial Drive
Berkeley
California 94720
USA

Current address:
Department of Mathematics
University of Maryland
College Park
Maryland 20742
USA