

VC density was introduced in [1] by Aschenbrenner, Dolich, Haskell, MacPherson, and Starchenko as a natural notion of dimension for NIP theories. In a NIP theory we can define a VC function

$$vc : \mathbb{N} \longrightarrow \mathbb{N}$$

Where $vc(n)$ measures complexity of definable sets n -dimensional space. Simplest possible behavior is $vc(n) = n$ for all n . Theories with that property are known to be dp-minimal, i.e. having the smallest possible dp-rank. In general, it is not known whether there can be a dp-minimal theory which doesn't satisfy $vc(n) = n$.

In this paper we work with trees viewed as posets. Parigot in [3] showed that such theory is NIP. This result was strengthened by Simon in [2] showing that trees are dp-minimal. [1] has the following problem

Problem 0.1. ([1] p. 47) Determine the VC density function of each (infinite) tree.

Here we settle this question by showing that theory of trees has $vc(n) = n$.

1. PRELIMINARIES

We use notation $a \in T^n$ for tuples of size n . For variable x or tuple a we denote their arity by $|x|$ and $|a|$ respectively.

We work with finite relational languages. Given a formula we can define its complexity n as the depth of quantifiers used to build up the formula. More precisely

Definition 1.1. Define complexity of a formula by induction:

$$\begin{aligned} \text{Complexity}(\text{q.f. formula}) &= 0 \\ \text{Complexity}(\exists x \phi(x)) &= \text{Complexity}(\phi(x)) + 1 \\ \text{Complexity}(\phi \cap \psi) &= \max(\text{Complexity}(\phi), \text{Complexity}(\psi)) \\ \text{Complexity}(\neg \phi) &= \text{Complexity}(\phi) \end{aligned}$$

A simple inductive argument verifies that there are (up to equivalence) only finitely many formulas with fixed complexity and the number of free variables. We will use the following notation for types and Stone Space:

Definition 1.2. Let n, m be naturals, \mathbf{B} a structure, A a parameter set and a, b tuples in \mathbf{B} .

- $\text{tp}_{\mathbf{B}}^n(a/A)$ will stand for all the A -formulas of complexity $\leq n$ that are true of a in \mathbf{B} . Note that if A is finite there are finitely many of them (up to equivalence). Conjunction of those formulas would still have complexity $\leq n$ so we can just associate a single formula describing this type. If $A = \emptyset$ we may write $\text{tp}_{\mathbf{B}}^n(a)$. \mathbf{B} will be omitted as well if it is clear from context.
- $\mathbf{B} \models a \equiv_A^n b$ means that a, b have the same type of complexity n over A in \mathbf{B} , i.e. $\text{tp}_{\mathbf{B}}^n(a/A) = \text{tp}_{\mathbf{B}}^n(b/A)$
- $S_{\mathbf{B}, m}^n(A)$ will stand for all m -types of complexity n over A .

Language for the trees consists of a single binary predicate $\{\leq\}$. Theory of trees states that \leq defines a partial order and for every element a we have $\{x \mid x < a\}$ a linear order. For visualization purposes we assume trees grow upwards, with smaller elements on the bottom and larger elements on the top. If $a \leq b$ we will say that a is below b and b is above a .

Definition 1.3. Work in a tree T . For $x \in T$ let $I(x) = \{t \in T \mid t \leq x\}$ denote all elements below x . *Meet* of two tree elements a, b is the greatest element of $I(a) \cap I(b)$ (if one exists).

Theory of meet trees requires that any two elements in the same connected component have a meet. Colored trees are trees with a finite number of colors added via unary predicates.

For the entire paper when we say tree we mean colored tree. We allow our trees to be disconnected or finite unless otherwise stated.

VC-DENSITY FOR TREES

ANTON BOBKOV

ABSTRACT. We show that for the theory of infinite trees we have $\text{vc}(n) = n$ for all n .

VC density was introduced in [1] by Aschenbrenner, Dolich, Haskell, MacPherson, and Starchenko as a natural notion of dimension for NIP theories. In a NIP theory we can define a VC function

$$\text{vc} : \mathbb{N} \longrightarrow \mathbb{N}$$

Where $\text{vc}(n)$ measures complexity of definable sets n -dimensional space. Simplest possible behavior is $\text{vc}(n) = n$ for all n . Theories with that property are known to be dp-minimal, i.e. having the smallest possible dp-rank. In general, it is not known whether there can be a dp-minimal theory which doesn't satisfy $\text{vc}(n) = n$.

In this paper we work with trees viewed as posets. Parigot in [3] showed that such theory is NIP. This result was strengthened by Simon in [2] showing that trees are dp-minimal. [1] has the following problem

Problem 1.4. ([1] p. 47) Determine the VC density function of each (infinite) tree.

Here we settle this question by showing that theory of trees has $\text{vc}(n) = n$.

2. PRELIMINARIES

We use notation $a \in T^n$ for tuples of size n . For variable x or tuple a we denote their arity by $|x|$ and $|a|$ respectively.

We work with finite relational languages. Given a formula we can define its complexity n as the depth of quantifiers used to build up the formula. More precisely

Definition 2.1. Define complexity of a formula by induction:

$$\begin{aligned} \text{Complexity}(\text{q.f. formula}) &= 0 \\ \text{Complexity}(\exists x \phi(x)) &= \text{Complexity}(\phi(x)) + 1 \\ \text{Complexity}(\phi \cap \psi) &= \max(\text{Complexity}(\phi), \text{Complexity}(\psi)) \\ \text{Complexity}(\neg \phi) &= \text{Complexity}(\phi) \end{aligned}$$

A simple inductive argument verifies that there are (up to equivalence) only finitely many formulas with fixed complexity and the number of free variables. We will use the following notation for types and Stone Space:

Definition 2.2. Let n, m be naturals, \mathbf{B} a structure, A a parameter set and a, b tuples in \mathbf{B} .

- $\text{tp}_{\mathbf{B}}^n(a/A)$ will stand for all the A -formulas of complexity $\leq n$ that are true of a in \mathbf{B} . Note that if A is finite there are finitely many of them (up to equivalence). Conjunction of those formulas would still have complexity $\leq n$ so we can just associate a single formula describing this type. If $A = \emptyset$ we may write $\text{tp}_{\mathbf{B}}^n(a)$. \mathbf{B} will be omitted as well if it is clear from context.
- $\mathbf{B} \models a \equiv_A^n b$ means that a, b have the same type of complexity n over A in \mathbf{B} , i.e. $\text{tp}_{\mathbf{B}}^n(a/A) = \text{tp}_{\mathbf{B}}^n(b/A)$
- $S_{\mathbf{B},m}^n(A)$ will stand for all m -types of complexity n over A .

Language for the trees consists of a single binary predicate $\{\leq\}$. Theory of trees states that \leq defines a partial order and for every element a we have $\{x \mid x < a\}$ a linear order. For visualization purposes we assume trees grow upwards, with smaller elements on the bottom and larger elements on the top. If $a \leq b$ we will say that a is below b and b is above a .

Definition 2.3. Work in a tree T . For $x \in T$ let $I(x) = \{t \in T \mid t \leq x\}$ denote all elements below x . Meet of two tree elements a, b is the greatest element of $I(a) \cap I(b)$ (if one exists).

Theory of meet trees requires that any two elements in the same connected component have a meet. Colored trees are trees with a finite number of colors added via unary predicates.

For the entire paper when we say tree we mean colored tree. We allow our trees to be disconnected or finite unless otherwise stated. For completeness we also present definition of VC function. One should refer to [1] for more details. Suppose we have a collection \mathcal{S} of subsets of X . We define a shatter function $\pi_{\mathcal{S}}(n)$

$$\pi_{\mathcal{S}}(n) = \max\{|A \cap \mathcal{S}| : A \subset X \text{ and } |A| = n\}$$

Sauer's Lemma asserts that asymptotically $\pi_{\mathcal{S}}$ is either 2^n or polynomial. In the polynomial case we define VC density of \mathcal{S} to be power of polynomial that bounds $\pi_{\mathcal{S}}$. More formally

$$\text{vc}(\mathcal{S}) = \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{S}}}{\log n}$$

Given a model $M \models T$ and a formula $\phi(x, y)$ we define

$$\mathcal{S}_\phi = \{\phi(M^{|n|}, b) : b \in M^{|y|}\}$$

$$\text{vc}(\phi) = \text{vc}(\mathcal{S}_\phi)$$

One has to check that this definition is independent of realization of T , see [1], Lemma 3.2. For a theory T we define the VC function

$$\text{vc}(n) = \sup\{\text{vc}(\phi(x, y)) : |x| = n\}$$

3. PROPER SUBDIVISIONS

Definition 3.1. Let $\mathbf{A}, \mathbf{B}, \mathbf{T}$ be models in some (possibly different) finite relational languages. If A, B partition T (i.e. $T = A \sqcup B$) we say that (\mathbf{A}, \mathbf{B}) is a *subdivision* of \mathbf{T} .

Definition 3.2. (\mathbf{A}, \mathbf{B}) subdivision of \mathbf{T} is called *n-proper* if for all $p, q \in \mathbb{N}$, for all $a_1, a_2 \in A^p$ and $b_1, b_2 \in B^q$ we have

$$\mathbf{A} \models a_1 \equiv_n a_2$$

$$\mathbf{B} \models b_1 \equiv_n b_2$$

then

$$\mathbf{T} \models a_1 b_1 \equiv_n a_2 b_2$$

Definition 3.3. (\mathbf{A}, \mathbf{B}) subdivision of \mathbf{T} is called *proper* if it is *n-proper* for all $n \in \mathbb{N}$.

Lemma 3.4. Consider a subdivision (\mathbf{A}, \mathbf{B}) of \mathbf{T} . If it is 0-proper then it is proper.

Proof. Prove the subdivision is *n-proper* for all n by induction. Case $n = 0$ is given by the assumption. Suppose $n = k + 1$ and we have $\mathbf{T} \models \exists x \phi^k(x, a_1, b_1)$ where ϕ^k is some formula of complexity k . Let $a \in T$ witness the existential claim i.e. $\mathbf{T} \models \phi^k(a, a_1, b_1)$. $a \in A$ or $a \in B$. Without loss of generality assume $a \in A$. Let $\mathbf{p} = \text{tp}_{\mathbf{A}}^k(a, a_1)$. Then we have

$$\mathbf{A} \models \exists x \text{tp}_{\mathbf{A}}^k(x, a_1) = \mathbf{p}$$

Formula $\text{tp}_{\mathbf{A}}^k(x, a_1) = \mathbf{p}$ is of complexity k so $\exists x \text{tp}_{\mathbf{A}}^k(x, a_1) = \mathbf{p}$ is of complexity $k + 1$ by inductive hypothesis we have

$$\mathbf{A} \models \exists x \text{tp}_{\mathbf{A}}^k(x, a_2) = \mathbf{p}$$

Let a' witness this existential claim so that

$$\text{tp}_{\mathbf{A}}^k(a', a_2) = \mathbf{p}$$

$$\text{tp}_{\mathbf{A}}^k(a', a_2) = \text{tp}_{\mathbf{A}}^k(a, a_1)$$

$$\mathbf{A} \models a' a_2 \equiv_k a a_1$$

by inductive assumption we have

$$\mathbf{T} \models a a_1 b_1 \equiv_k a' a_2 b_2$$

$$\mathbf{T} \models \phi^k(a', a_2, b_2) \quad \text{as } \mathbf{T} \models \phi^k(a, a_1, b_1)$$

$$\mathbf{T} \models \exists x \phi^k(x, a_2, b_2)$$

□

We use this lemma for colored trees. Suppose we have T to be a model of a colored tree in some language $\mathcal{L} = \{\leq, \dots\}$ and \mathbf{A}, \mathbf{B} be in some languages $\mathcal{L}_A, \mathcal{L}_B$ which will be expands of \mathcal{L} , with \mathbf{A}, \mathbf{B} substructures of T as reducts to \mathcal{L} . In this case we'll refer to (\mathbf{A}, \mathbf{B}) as a proper subdivision (of T).

Example 3.5. Suppose a tree consists of two connected components C_1, C_2 . Then (C_1, \leq) and (C_2, \leq) form a proper subdivision.

Example 3.6. Work with a (colored) tree T in language $\mathcal{L} = \{\leq, C_1, \dots, C_n\}$ with colors interpreted by sets S_1, \dots, S_n . Fix $a \in T$. Let $B = \{t \in T \mid a < t\}$, $S = \{t \in T \mid t \leq a\}$, $A = T - B$. Consider structures

$$\mathbf{A} = (A, \leq, S_1 \cap A, \dots, S_n \cap A, S) \text{ in language with an extra color } \mathcal{L}_A = \mathcal{L} \cup \{C\}$$

$$\mathbf{B} = (B, \leq, S_1 \cap B, \dots, S_n \cap B) \text{ in language } \mathcal{L}$$

Then (\mathbf{A}, \mathbf{B}) is a proper subdivision of T .

Proof. By Lemma 2.4 we only need to check that it is 0-proper. Suppose we have

$$a = (a_1, a_2, \dots, a_p) \in A^p$$

$$a' = (a'_1, a'_2, \dots, a'_p) \in A^p$$

$$b = (b_1, b_2, \dots, b_q) \in B^q$$

$$b' = (b'_1, b'_2, \dots, b'_q) \in B^q$$

with $(\mathbf{A}, a) \equiv_0 (\mathbf{A}, a')$ and $(\mathbf{B}, b) \equiv_0 (\mathbf{B}, b')$. We need to make sure that ab has the same quantifier free type as $a'b'$. Any two elements in T can be related in the four following ways

$$x = y$$

$$x < y$$

$$x > y$$

$$x, y \text{ are incomparable}$$

We need to check that the same relations hold for pairs of $(a_i, b_j), (a'_i, b'_j)$ for all i, j .

- It is impossible that $a_i = b_j$ as they come from disjoint sets.
- It is impossible $a_i > b_j$
- Suppose $a_i < b_j$. This forces $a_i \in S$ thus $a'_i \in S$ and $a'_i < b'_j$
- Suppose a_i and b_j are incomparable. This forces $a_i \notin S$ so $a'_i \notin S$ making a'_i, b'_j incomparable.

Also we need to check that ab has the same colors as $a'b'$. That follows directly from definition. \square

4. INTERVALS

Definition 4.1. For $c_1 > c_2$ in a tree T we define an (open) interval

$$(c_1, c_2) = \{t \in T \mid c_1 > t \text{ and } c_2 \not\geq t\}$$

We also allow endpoints to be infinite, $c_1 = -\infty$ or $c_2 = \infty$.

Lemma 4.2. Suppose we have $I = (c_1, c_2)$, $a, b \in I$ and $a \equiv_{c_1 c_2}^n b$. Denote $D = T - I$. Then $a \equiv_D^n b$.

To prove this, first we do the following two lemmas:

Lemma 4.3. *Suppose we have $I = (-\infty, c_2)$, $a, b \in I$ and $a \equiv_{c_2}^n b$. Denote $U = T - I$. Then $a \equiv_U^n b$.*

Proof. Consider subdivision as in the example 2.6 and fix a tuple $c \in U^{|c|}$. Note that $B = U$. Universe and relations in \mathbf{A} are c_2 -definable in \mathbf{T} . Then $\mathbf{T} \models a \equiv_{c_2}^n b$ implies that $\mathbf{A} \models a \equiv^n b$. We automatically have that $\mathbf{B} \models c \equiv^n c$. Thus by properness we get $\mathbf{T} \models ac \equiv^n bc$ or $\mathbf{T} \models a \equiv_c^n b$. As choice of c was arbitrary we are done. \square

Lemma 4.4. *Suppose we have $I = (c_1, \infty)$, $a, b \in I$ and $a \equiv_{c_1}^n b$. Denote $L = T - I$. Then $a \equiv_L^n b$.*

Proof. Similar to the proof above. \square

Proof. (of lemma 3.2) Easy combination of the lemmas above. Pick an arbitrary tuple $d \in D^{|d|}$ and partition it as $d = (l, u)$ such that $l \in L^{|l|}$ and $u \in U^{|u|}$ (where L and U are as in the lemmas above). Then

$$\begin{aligned}
 a &\equiv_{c_1}^n b \\
 a &\equiv_{c_2 l}^n b && \text{by lemma 3.4} \\
 la &\equiv_{c_2}^n lb \\
 la &\equiv_{c_2 u}^n lb && \text{by lemma 3} \\
 a &\equiv_{c_2 l u}^n b \\
 a &\equiv_{c_2 l u}^n b \\
 a &\equiv_d^n b
 \end{aligned}$$

\square

Note 4.5. If we weren't restricting complexity of type, this lemma could be proven in an easier way by modifying automorphism sending a to b . However to have a uniform bound on complexity we need to use a finer analysis involving proper subdivisions.

Next we state a generalization of this result for a collection of disjoint intervals. This will only required for computing vc -density for formulas $\phi(x, y)$ with $|y| > 1$. For technical reasons we will allow some of the intervals to be single points $\{c\}$.

Corollary 4.6. *Work in a fixed tree. Let I_i with $i \in [1..m]$ be a collection of disjoint sets where each set is either interval (l_i, u_i) or a point $\{c_i\}$. We denote E_i denote endpoints of I_i . If $I = (l_i, u_i)$ then $E_i = \{l_i, u_i\}$ and if $I = \{c_i\}$ then $E_i = \{c_i\}$. Also we have tuples*

$$\begin{aligned}
 a &= (a_1, a_2, \dots, a_m) \\
 b &= (b_1, b_2, \dots, b_m)
 \end{aligned}$$

s.t. $a_i \in I_i^{|a_i|}$ and $b_i \in I_i^{|b_i|}$ for all $i \in [1..m]$ (i.e every a_i and b_i are also tuples). Moreover we require

$$a_i \equiv_{E_i}^n b_i \text{ for all } i \in [1..m]$$

Let $E = \bigcup_i E_i$ be a collection of all endpoints. Then we $a \equiv_E^n b$.

Proof. We apply lemma m times - one for each tuple. For example assume $I_1 = (l_1, u_1)$ and denote $C_1 = T - I_1$ (case for when I_i is singleton is trivial). Then

$$\begin{aligned} a_1 &\equiv_{l_1 u_1}^n b_1 \\ a_1 &\equiv_{C_1}^n b_1 \\ a_1 &\equiv_{E a_2 a_3 \dots a_m}^n b_1 \\ a_1 a_2 a_3 \dots a_m &\equiv_E^n b_1 a_2 a_3 \dots a_m \end{aligned}$$

Denote $C_2 = T - I_2$ and assume $I_2 = (l_2, u_2)$. Then

$$\begin{aligned} a_2 &\equiv_{l_2 u_2}^n b_2 \\ a_2 &\equiv_{C_2}^n b_2 \\ a_2 &\equiv_{E b_1 a_3 \dots a_m}^n b_2 \\ b_1 a_2 a_3 \dots a_m &\equiv_E^n b_1 b_2 a_3 \dots a_m \end{aligned}$$

Continuing in such fashion we get

$$\begin{aligned} a &= a_1 a_2 a_3 \dots a_m \equiv_E^n \\ b_1 a_2 a_3 \dots a_m &\equiv_E^n \\ b_1 b_2 b_3 \dots a_m &\equiv_E^n \\ &\dots \\ b_1 b_2 b_3 \dots b_m &= b \end{aligned}$$

i.e. $a \equiv_E^n b$ as needed. \square

5. PARTITION INTO INTERVALS

Here we show that every tree can be partitioned into intervals with prescribed endpoints. To do this we must restrict our attention to meet trees (closed under meets). Given a finite subset B of a meet tree T let $\text{cl}(B)$ denote its closure under meets.

Lemma 5.1. *Suppose $S \subseteq T$ is a non-empty finite subset of size n in a meet tree and $S' = \text{cl}(S)$ its closure under meets. Then $|S'| \leq 2n$.*

Proof. We can partition S into connected components and prove the result separately for each component. Thus we may assume elements of S lie in the same connected component. We prove the claim by induction on n . Base case $n = 1$ is clear. Suppose we have S of size k with closure of size at most $2k - 1$. Take a new point s , and look at its meets with all the elements of S . Pick the smallest one, s' . Then $S \cup \{s, s'\}$ is closed under meets. \square

Lemma 5.2. *Let C be a non-empty finite set closed under meets in a tree T . Then there is a collection of disjoint intervals $\{I_i = (a_i, b_i)\}_{1..n}$ such that*

- (1) *endpoints are either in C or infinite.*
- (2) *$n \leq 2|C|$.*
- (3) *I_i partition $T - C$.*

Proof. For every $c \in C$ define c^+ to be supremum of all elements of C less than c . If there aren't any, let $c^+ = -\infty$. Separate C into connected components \mathcal{C} . For every $D \in \mathcal{C}$ let c_D be supremum of all the elements in that component. Our desired collection of intervals is $(c, c^+)_{c \in C} \cup (c_D, \infty)_{D \in \mathcal{C}}$. \square

6. TYPE COUNTING

Definition 6.1. For $\phi(x, y)$, $A \subseteq T^{|x|}$ and $B \subseteq T^{|y|}$

- Let $\phi(A, b) = \{a \in A \mid \phi(a, b)\} \subseteq A$
- Let $\phi(A, B) = \{\phi(A, b) \mid b \in B\} \subseteq \mathcal{P}(A)$

That is $\phi(A, B)$ is a collection of subsets of A definable by ϕ with parameters from B .

Note 6.2. For $A_1 \subseteq A_2$ we have $|\phi(A_1, B)| \leq |\phi(A_2, B)|$.

Definition 6.3. Fix a (finite relational) language \mathcal{L}_B , and $n, |y|$. Let $N = N(n, |y|, \mathcal{L}_B)$ be smallest number such that for any structure \mathbf{B} in \mathcal{L}_B we have $|S_{\mathbf{B}, |y|}^n| \leq N$. This number is well-defined as there are a finite number (up to equivalence) of possible formulas of complexity $\leq n$ with $|y|$ free variables. Note the following easy inequalities

$$\begin{aligned} n \leq m &\Rightarrow N(n, |y|, \mathcal{L}_B) \leq N(m, |y|, \mathcal{L}_B) \\ |y| \leq |z| &\Rightarrow N(n, |y|, \mathcal{L}_B) \leq N(n, |z|, \mathcal{L}_B) \\ \mathcal{L}_A \subseteq \mathcal{L}_B &\Rightarrow N(n, |y|, \mathcal{L}_A) \leq N(n, |y|, \mathcal{L}_B) \\ N(n, |y|, \mathcal{L}_B) \cdot N(n, |z|, \mathcal{L}_B) &\leq N(n, |y| + |z|, \mathcal{L}_B) \end{aligned}$$

Lemma 6.4. Suppose $\phi(x, y)$ is a formula of complexity n in some language \mathcal{L} of a (colored) tree. Let $I = (c_1, c_2)$ an interval and let its compliment be D . Then

$$|\phi(D^{|x|}, I^{|y|})| < N(n, |y|, \mathcal{L}')$$

where \mathcal{L}' is \mathcal{L} with two extra constant symbols.

Note that the bound provided is uniform - it doesn't depend on the model or the interval.

Proof. Fix $a, b \in I^{|y|}$. Then if $a \equiv_{c_1 c_2}^n b$ then $\phi(D^{|x|}, a) = \phi(D^{|x|}, b)$ by lemma 3.2. Thus $|\phi(D^{|x|}, I^{|y|})|$ is bounded by $|S^n(y/c_1 c_2)|$. \square

We extract a similar result from Corollary 3.6.

Lemma 6.5. Suppose $\phi(x, y)$ is a formula of complexity n . Let I_i with $i \in [1..m]$ be a disjoint collection of set that are either intervals or points. Denote collection of all endpoints by E . Moreover fix naturals k_1, \dots, k_m such that $|y| = k_1 + k_2 + \dots + k_m$. Then

$$|\phi(E^{|x|}, I_1^{k_1} \times I_2^{k_2} \times \dots \times I_m^{k_m})| \leq N(n, |y|, \mathcal{L}')$$

where \mathcal{L}' is \mathcal{L} with two extra constant symbols.

Proof. Fix $a_i, b_i \in I_i^{k_i}$ for each i such that $a_i \equiv_{E_i}^n b_i$ where E_i is endpoints of I_i . Then by Corollary 3.6 $\phi(E^{|x|}, a_1, a_2, \dots, a_m) = \phi(E^{|x|}, b_1, b_2, \dots, b_m)$. This implies that

$$\begin{aligned} |\phi(E^{|x|}, I_1^{k_1} \times I_2^{k_2} \times \dots \times I_m^{k_m})| &\leq |S_{k_1}^n(E_1)| \cdot |S_{k_2}^n(E_2)| \cdot \dots \cdot |S_{k_m}^n(E_m)| \leq \\ &\leq N(n, k_1, \mathcal{L}') \cdot N(n, k_2, \mathcal{L}') \cdot \dots \cdot N(n, k_m, \mathcal{L}') \leq \\ &\leq N(n, k_1 + k_2 + \dots + k_m, \mathcal{L}') = N(n, |y|, \mathcal{L}') \end{aligned}$$

\square

7. PROOF FOR ONE-DIMENSIONAL CASE

To demonstrate our method in a simple setting we present the computation of vc -density for formulas $\phi(x, y)$ with $|x| = |y| = 1$ in an infinite (colored) meet tree \mathbf{T} . Let B be an arbitrary finite set in T . We would like to show that $|\phi(B, T)| = O(|B|)$. This would show that vc -density of ϕ is 1. First let $C = \text{cl}(B)$. By Lemma 4.1 we have $|C| \leq 2|B|$. Using Lemma 4.2 we can partition T into intervals with endpoints exactly C . More precisely we get a collection of intervals $\{I_i\}_{i < m}$ such that $T - C = \bigsqcup I_i$ with $m \leq 2|C|$. Observe

$$\phi(C, T) = \phi(C, \bigsqcup I_i \sqcup C) = \bigcup \phi(C, I_i) \cup \phi(C, C)$$

Thus

$$|\phi(C, T)| = \left| \bigcup \phi(C, I_i) \cup \phi(C, C) \right| \leq \sum |\phi(C, I_i)| + |\phi(C, C)| = \sum_{1..|C|} |\phi(C, I_i)| + |C|$$

Let n_ϕ denote complexity of ϕ and let \mathcal{L}' be language of \mathbf{T} with two extra constant symbols. Now observe that by Lemma 5.4 $|\phi(C, I_i)| \leq (n_\phi, 1, \mathcal{L}')$ for any i . Denote $N = N(n_\phi, 1, \mathcal{L}')$. We have

$$|\phi(C, T)| \leq \sum_{i < m} |\phi(C, I_i)| + |C| \leq \sum_{i < m} N + |C| = mN + |C|$$

Finally, as $B \subseteq C$ we have $|\phi(B, T)| \leq |\phi(C, T)|$ so

$$|\phi(B, T)| \leq |\phi(C, T)| \leq mN + |C| \leq 4N|B| + 2|B|$$

As N is independent from B , this finishes the proof.

8. MAIN PROOF

Basic idea for the proof is that we are able to divide our parameter space into $O(n)$ many intervals. Each of q parameters can come from any of those $O(n)$ intervals giving us $O(n)^q = O(n^q)$ many choices for parameter configuration. With every parameter coming from a fixed partition the number of definable sets is constant and in fact is uniformly bounded by some N . This gives us $NO(n^q) = O(n^q)$ possibilities for different definable sets.

Theorem 8.1. *Let \mathbf{T} be an infinite (colored) meet tree and $\phi(x, y)$ a formula with $|x| = p$ and $|y| = q$. Then $vc(\phi) \leq q$.*

Proof. Pick a finite subset of $S_0 \subset T^p$ of size n . Let $S_1 \subset T$ consist of coordinates of S_0 . Let $S \subset T$ be a closure of S_1 under meets. Using Lemma 4.1 we have $|S| \leq 2|S_1| \leq 2p|S_0| = 2pn = O(n)$. We have $S_0 \subseteq S^p$, so $|\phi(S_0, T^q)| \leq |\phi(S^p, T^q)|$. Thus it is enough to show $|\phi(S^p, T^q)| = O(n^q)$. Using Lemma 4.2 construct collection of intervals $\{I_i\}_{i \in \mathcal{I}}$ for the set S . We have $|\mathcal{I}| = 2|S| + 1$. Extend this collection of intervals by adding singletons $\{s\}$ for each $s \in S$. Denote new collection $\{I_j\}_{j \in \mathcal{J}}$. We have $|\mathcal{J}| = |\mathcal{I}| + |S| = 3|S| + 1 \leq 6pn + 1 = O(n)$. We now have that $\{I_j\}_{j \in \mathcal{J}}$ partitions T i.e. $T = \bigsqcup_{j \in \mathcal{J}} I_j$. For $(j_1, j_2, \dots, j_q) = \vec{j} \in \mathcal{J}^q$ denote

$$I_{\vec{j}} = I_{j_1} \times I_{j_2} \times \dots \times I_{j_q}$$

Then we have the following identity

$$T^q = \left(\bigsqcup_{j \in \mathcal{J}} I_j \right)^q = \bigsqcup_{\vec{j} \in \mathcal{J}^q} I_{\vec{j}}$$

Thus we have that $\{I_{\vec{j}}\}_{\vec{j} \in \mathcal{J}^q}$ partition T^q . Compute

$$\begin{aligned} |\phi(S^p, T^q)| &= \left| \bigcup_{\vec{j} \in \mathcal{J}^q} \phi(S^p, I_{\vec{j}}) \right| \leq \\ &\leq \sum_{\vec{j} \in \mathcal{J}^q} |\phi(S^p, I_{\vec{j}})| \end{aligned}$$

We can bound $|\phi(S^p, I_{\vec{j}})|$ uniformly using Lemma 5.5. $\{I_j\}_{j \in \mathcal{J}}$ satisfies the requirements of the lemma and $I_{\vec{j}}$ looks like the argument in the lemma after possibly permuting some variables in ϕ . Applying the lemma we get

$$|\phi(S^p, I_{\vec{j}})| \leq N (= N(\text{complexity of } \phi, q, \mathcal{L}'))$$

with N only depending on q and complexity of ϕ . We complete our computation

$$\begin{aligned} |\phi(S^p, T^q)| &\leq \sum_{\vec{j} \in \mathcal{J}^q} |\phi(S^p, I_{\vec{j}})| \leq \\ &\leq \sum_{\vec{j} \in \mathcal{J}^q} N \leq \\ &\leq |\mathcal{J}^q| N \leq \\ &\leq (6pn + 1)^q N = O(n^q) \end{aligned}$$

□

Corollary 8.2. *In the theory of infinite (colored) meet trees we have $vc(n) = n$ for all n .*

Corollary 8.3. *In the theory of infinite (colored) trees we have $vc(n) = n$ for all n .*

Proof. Let \mathbf{T}' be a tree. We can embed it in a larger tree that is closed under meets $\mathbf{T}' \subset \mathbf{T}$. Expand \mathbf{T} by an extra color and interpret it by coloring the subset \mathbf{T}' . Thus we can interpret \mathbf{T}' in T^1 . By Corollary 3.17 in [1] we get that $vc^{\mathbf{T}'}(n) \leq vc^T(1 \cdot n) = n$ thus $vc^{\mathbf{T}'}(n) = n$ as well. □

REFERENCES

- [1] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, S. Starchenko, *Vapnik-Chervonenkis density in some theories without the independence property*, I, preprint (2011)
- [2] P. Simon, *On dp-minimal ordered structures*, J. Symbolic Logic 76 (2011), no. 2, 448-460
- [3] Michel Parigot. Théories d'arbres. *Journal of Symbolic Logic*, 47, 1982.
E-mail address: bobkov@math.ucla.edu