Alcom-FT Technical Report Series
ALCOMFT-TR-01-21

A General Dimension for Exact Learning

José L. Balcázar, Jorge Castro*
Dept. LSI, Universitat Politècnica de Catalunya,
Campus Nord, 08034 Barcelona, Spain, {balqui,castro}@lsi.upc.es

David Guijarro†
Mannes Technology Consulting
Pl. Tirant lo Blanc 7, 08005 Barcelona, Spain, david@mannes-tech.com

1 Introduction

This paper is about query complexity, i.e. counting the number of queries needed to learn in any exact query learning model such as those defined in [1]. We do not deal with time complexity issues but, instead, concentrate our efforts in giving a *unique* combinatorial notion for *any* set of queries that characterizes the query complexity.

In the exact learning model there is a special type of queries whose answers can be always modeled with sets of examples: the example-based queries. In this case, there is a lot of work done for specific cases: certificates [11] and consistency dimension [4] for membership and equivalence queries, approximate fingerprints [2, 7] and strong consistency dimension [4] for equivalence queries and extended teaching dimension [10] for membership queries. Furthermore, there is a unifying concept, the abstract identification dimension (AIdim) [3], that gives a unique characterization for any set of example-based queries.

However, there are queries that do not fit in the notion of example-based queries, for instance restricted equivalence [1] (equivalence queries that do not supply counterexample) and unspecified attribute value queries [9] (UAVs). There are works on lower bounds for UAVs [5, 12] but, as far as we know, there is no general learning dimension for UAVs comparable to AIdim in the case of example-based queries.

In this work we introduce a new combinatorial notion, Gdim, that gives a good approximation of the number of queries needed to learn with any set of queries in the exact learning model. We show that AIdim is the natural derivation of Gdim in the case of example-based queries. And finally, we present two applications of Gdim for UAVs: (a) the first characterizations of the UAV protocols with membership and equivalence queries and membership queries alone, which are a straightforward but nice generalization of their standard counterparts and (b) the number of standard membership and equivalence queries needed to learn DNF formulas, which we prove to be polynomially bounded by the number of UAV membership and equivalence queries. This shows that, in a query complexity sense, UAVs do not help in learning DNF formulas.

*Work supported in part by the IST Programme of the EU under contract number IST-1999-14186 (ALCOM-FT) and the spanish government grant PB98-0937-C04-04.

†Part of this work was done while this author was still in LSI, UPC.

2 Notation and the abstract setting for exact learning

We focus on exact learning of Boolean functions and we fix all along the paper n as the number of variables. A Boolean function of arity n is a function from $\{0,1\}^n \to \{0,1\}$. The set of all Boolean functions is denoted by B_n.

An element x of $\{0,1\}^n$ is called an *assignment*. A pair (x,b), where $b \in \{0,1\}$ is a binary label, is called *example for function* $f \in B_n$ if $f(x) = b$. A *sample* is a collection of examples for some function $f \in B_n$. We denote by $\|X\|$ the cardinality of set X.

A partial assignment α is a word from $\{0,1,\star\}^n$. A complete assignment $x \in \{0,1\}^n$ satisfies a partial assignment α if they coincide in the positions where α is not $\star$. The hypercube H_α of a partial assignment α is the set of all complete assignments that satisfy α. Given $b \in \{0,1\}$, the expression $\alpha \cup v \leftarrow b$ denotes the partial assignment such that for any variable u it holds $u(\alpha \cup v \leftarrow b) = u(\alpha)$ when $u \neq v$ and $u(\alpha \cup v \leftarrow b) = b$ if $u = v$. We also denote by t_α the term that, when applied to a complete assignment x, evaluates to 1 if x satisfies α and to 0 otherwise and by c_α the clause such that $c_\alpha = \overline{t}_\alpha$.

Each function $f \in B_n$ can be formalized as an assignment in $\{0,1\}^{2^n}$. In this way, we identify the set of Boolean functions from $B_n \to \{0,1\}$ with B_{2^n}. Functions in B_{2^n} have a variable for each assignment y in $\{0,1\}^n$, which we denote by λ_y. We will make the following natural interpretation: $f \in B_n$ satisfies the positive (resp. negated) literal with variable λ_y when $(y,1)$ (resp. $(y,0)$) is an example of f; this is naturally denoted $\lambda_y(f)$. For $S \subseteq B_{2^n}$ and $f \in B_n$, the expression $S(f)$ is $\bigwedge_{\Lambda \in S} \Lambda(f)$.

In our abstract setting, queries are atomic objects. Answers provide some partial knowledge of the target. Since our target concepts are always Boolean functions of arity n, we assume that such partial knowledge is always modeled as a function in B_{2^n} that is interpreted as a property that the target function satisfies. That is, answers correspond to functions in B_{2^n}.

Thus, in any abstract learning protocol we have two participants: the set Q of queries and the set of all possible answers. Since the set of answers will be specifically defined by each learning protocol, we only write explicitly the dependence of the protocol on Q. A protocol *Protocol*(Q) is a subset of $\{\langle q, \Lambda\rangle \mid q \in Q, \Lambda \in B_{2^n}\}$.

For instance, if we want to talk about learning with the usual equivalence queries with hypothesis coming from a subset $H \subseteq B_n$, we define *Protocol*$_{\equiv}(H)$ as the set

$$\{\langle h, \lambda_y\rangle \mid h \in H, y \in \{0,1\}^n, h(y) = 0\} \cup \{\langle h, \overline{\lambda}_y\rangle \mid h \in H, y \in \{0,1\}^n, h(y) = 1\}$$

$$\cup$$

$$\{\langle h, (\bigwedge_y \lambda_y) \wedge (\bigwedge_z \overline{\lambda}_z)\rangle \mid h \in H,\ y, z \in \{0,1\}^n, h(y) = 1, h(z) = 0\}$$

where the first two sets correspond to counterexamples and the last set to YES answers.

Angluin defined in [1] equivalence queries without counterexamples. These queries, called *restricted* equivalence queries, have as possible answers YES or NO. We can consider this protocol in our setting as follows. For restricted equivalence queries with hypothesis $H \subseteq B_n$ the set *Protocol*$_{\equiv_r}(H)$ is

$$\{\langle h, (\bigwedge_y \lambda_y) \wedge (\bigwedge_z \overline{\lambda}_z)\rangle \mid h \in H,\ y, z \in \{0,1\}^n, h(y) = 1, h(z) = 0\}$$

$$\cup$$

$$\{\langle h, (\bigvee_y \overline{\lambda}_y) \vee (\bigvee_z \lambda_z)\rangle \mid h \in H,\ y, z \in \{0,1\}^n, h(y) = 1, h(z) = 0\}$$

where the first set corresponds to a YES answer and the second one to a negative answer. Observe that this protocol cannot be formalized in the framework presented in [3].

For a protocol P and a query q, P_q denotes the set of all answers associated, in P, to the query q; formally, $P_q = \{\Lambda \mid \langle q, \Lambda \rangle \in P\}$. For $f \in B_n$ we use P_q^f to denote the set of answers to the query q in P that are satisfied by f; formally, $P_q^f = \{\Lambda \mid \Lambda \in P_q \wedge \Lambda(f)\}$. We require all protocols P to be *complete* for the set of queries Q, which means that for all $q \in Q$ and for all $f \in B_n$, $P_q^f \neq \emptyset$. We also use P^f to denote the set $\{\langle q, \Lambda \rangle \mid q \in Q \wedge \langle q, \Lambda \rangle \in P \wedge \Lambda(f)\}$. The notation is extended to apply it to subsets of P.

An *answering scheme* T of a protocol P is a subset of P that for all $q \in Q$ the set of answers T_q is non empty. The set of all answering schemes of P is denoted as $\mathcal{T}(P)$. The notation is also extended to apply it to answering schemes of answering schemes. For any answering scheme T, T_Q denotes the set $\{\Lambda \mid q \in Q \wedge \langle q, \Lambda \rangle \in T\}$, i.e. the set of potential answers to all queries.

A protocol whose answers are always Boolean functions that can be expressed as terms, i.e. conjunctions of literals, is a *conjunctive* protocol. The answers of conjunctive protocols provide a partial knowledge of the target concept that can be modeled as the values of the target function on a subdomain. Thus, each answer is just a sample of the target. We note that queries that give this kind of answers are sometimes called example-based queries (see [8], for example). The most popular protocols as membership queries, equivalence queries and subset queries are conjunctive.

3 Exact learning and the general dimension

We use a generalization of the exact learning model via queries of Angluin [1]. A teacher answers with respect to $f \in B_n$ and using $P = \mathit{Protocol}(Q)$ if for each query $q \in Q$, it outputs some $\Lambda \in P_q^f$. A function class $C \subseteq B_n$ is learnable with d queries under $P = \mathit{Protocol}(Q)$ if there exists an algorithm A such that for any $f \in C$ and for any teacher B that answers with respect to f using P, the only remaining function in C that satisfies the answers received after at most d interactions is f. For a class $C \subseteq B_n$ and a protocol $P = \mathit{Protocol}(Q)$ we define the *learning complexity*, $\mathrm{LC}(C, P)$, as the smallest $d \geq 0$ such that C is learnable with d queries under P. If no such integer exists then $\mathrm{LC}(C, P) = \infty$.

A combinatorial parameter that captures the learning complexity can be defined using a chain of alternating quantifiers of queries and answers. We describe it here, as a way of introducing the idea, and also for the sake of comparison with the much nicer "flat" version we will describe below.

Given a protocol P with query set Q and a target class $C \subseteq B_n$ we define the *alternating dimension* as the smallest integer $d \geq 0$ such that

$$(\exists q_1 \in Q)(\forall \Lambda_1 \in P_{q_1}) \ldots (\exists q_d \in Q)(\forall \Lambda_d \in P_{q_d})(\|\{c \in C \mid \bigwedge \Lambda_i(c)\}\| \leq 1)$$

If no such integer exists then $\mathrm{Adim}(C, P) = \infty$.

Now, the following theorem can be easily proved by a fully standard adversary argument:

Theorem 1 *For any protocol P and any class $C \subseteq B_n$,*

$$LC(C, P) = Adim(C, P)$$

In the next part of the section we present a nicer dimension without alternating quantifiers that gives an approximation, up to a logarithm of the cardinality of the target class, of the learning complexity.

Given a target class $C \subseteq B_n$ and a protocol P on a query set Q, we define the *general dimension*, $\mathrm{Gdim}(C, P)$, as the minimum integer $d \geq 0$ such that

$$(\forall T \in \mathcal{T}(P))\,(\exists S \subseteq T_Q)\,(\|S\| \leq d \wedge \|\{c \in C \mid S(c)\}\| \leq 1)$$

If no such integer exists then $\mathrm{Gdim}(C, P) = \infty$.

That is, no matter what answering scheme is chosen there exists some set of at most d answers such that at most one function in the target class satisfies all those answers.

The following lemma will be central in the proof of our main result in this section and is interesting in its own right.

Lemma 2 *Let $D \subseteq C \subseteq B_n$ such that $\|D\| > 1$, $P = Protocol(Q)$ and $Gdim(C, P) = d$. There exists $q \in Q$ such that for any $\Lambda \in P_q$, at least $\frac{\|D\|-1}{d}$ functions from D do not satisfy Λ.*

Proof (sketch) Suppose that for each $q \in Q$ there exists some $\Lambda_q \in P_q$ such that less than $\frac{\|D\|-1}{d}$ functions do not satisfy Λ_q. A contradiction with $\mathrm{Gdim}(C, P) = d$ can be easily obtained simply considering the answering scheme T with $T_q = \{\Lambda_q\}$. □

Our main contribution of this section is the following characterization:

Theorem 3 *For any concept class $C \subseteq B_n$ and any protocol P on a query set Q,*

$$Gdim(C, P) \leq LC(C, P) \leq Gdim(C, P)\lceil \ln \|C\| \rceil$$

Proof (sketch) We will start showing that if $\mathrm{Gdim}(C, P) > k$ then any learning algorithm must ask more than k queries. Let T be the answering scheme such that

$$(\forall S \subseteq T_Q)(\|S\| \leq k \Rightarrow \|\{c \in C \mid S(c)\}\| > 1)$$

obtained by negation of the definition of Gdim. We answer all queries from a learner using T. After k interactions, the learner knows a set of given answers $S \subseteq T_Q$, and by the choice of T, there exist two different functions in C that satisfy all answers in S.

Now we show the upper bound. Let $\mathrm{Gdim}(C, P) = k$. The case $k = 0$ is easy. If $k = 1$ the theorem follows easily because Lemma 2 guarantees that $\mathrm{LC}(C, P) = 1$. Otherwise, let $\mathcal{V}$ be the version space consisting of functions in C that satisfy all the answers received so far (initially $\mathcal{V} = C$). Now we make the query whose existence is guaranteed by Lemma 2 when $D = \mathcal{V}$ such that no matter what the answer is, at least $\frac{\|\mathcal{V}\|-1}{k}$ functions from $\mathcal{V}$ do not satisfy that answer. A standard calculus shows that the number of such queries needed to reduce the number of surviving candidates to 1 is bounded by $k\lceil \ln \|C\| \rceil$. □

Observe that the algorithm implicit in the previous proof is simpler than its counterpart in [3] because it relies on a more powerful lemma. The key difference is that while the algorithm of [3] needed to make a majority vote strategy the present algorithm does not.

Next we show a necessary and sufficient condition for $\mathrm{Gdim}(C, P)$ being ∞ and a trivial upper bound when the dimension is finite (proof omitted).

Theorem 4 *For any $C \subseteq B_n$ and for any $P = Protocol(Q)$,*

1. *$Gdim(C, P) \neq \infty \Leftrightarrow \forall f, g \in C$, such that $f \neq g, \exists q \in Q$ such that $P_q^f \cap P_q^g = \emptyset$.*
2. *If $Gdim(C, P) \neq \infty$ and $C \neq \emptyset$ then $Gdim(C, P) \leq \|C\| - 1$.*

4 Abstract identification dimension revisited

We present now how to relate Gdim with the abstract identification dimension (AIdim) [3]. This concept is defined now by weakening the universal quantifier over answering schemes in the definition of Gdim and relaxing it to cover only answering schemes for which the conjunction of all of its answers is satisfiable by some function in B_n.

Formally, given a class $C \subseteq B_n$ and a protocol P on the query set Q, $\mathrm{AIdim}(C, P)$ is the minimum integer $d \geq 0$ such that

$$(\forall f \in B_n)\ (\forall T \in \mathcal{T}(P^f))\ (\exists S \subseteq T_Q)\ (\|S\| \leq d \wedge \|\{h \in C \mid S(h)\}\| \leq 1)$$

If no such integer exists then $\mathrm{AIdim}(C, P) = \infty$.

In [3], we showed that AIdim is a crucial notion in the sense that it unifies all learning dimensions of example-based query models. However, AIdim can not replace Gdim as a general notion of learning dimension. There are protocols where AIdim does not characterize learning complexity. For instance, consider Example 15 in the appendix.

The next theorem shows how AIdim corresponds with Gdim in conjunctive protocols. We recall that conjunctive protocols are the formalization in our abstract framework of the example-based query models.

Theorem 5 *Let P be a conjunctive protocol and $C \subseteq B_n$. Then*

$$\mathit{Gdim}(C, P) \leq \max(2, \mathit{AIdim}(C, P))$$

Proof Let T be any answering scheme. If for some function $f \in B_n$ it holds $T \in \mathcal{T}(P^f)$, we simply choose as set $S \subseteq T_Q$ in the definition of Gdim the set S promised by $\mathrm{AIdim}(C, P)$. If T cannot be satisfied by any function in B_n then there are two answers that already are not satisfied by any function. This is so because (a) a conjunction cannot be satisfied if and only if it contains two complemented literals and (b) the conjunction of conjunctive answers is still a conjunction. So this proves that $\mathrm{Gdim}(C, P) \leq \max(2, \mathrm{AIdim}(C, P))$. □

Observe that it is always the case that $\mathrm{AIdim}(C, P) \leq \mathrm{Gdim}(C, P)$ which is tantamount to say that they are equal if $\mathrm{AIdim}(C, P) \geq 2$. When $\mathrm{AIdim}(C, P) \leq 1$, Gdim and AIdim can be different. Consider the target class of all Boolean functions B_n and the protocol of equivalence queries with counterexamples where the set of hypothesis is also B_n. It is easy to prove that in this case AIdim is 1 and Gdim is 2.

The main difference between AIdim and Gdim relies in the fact that AIdim only considers answering schemes that contain answers that all together are satisfiable by, at least, one function. Theorem 5 shows that for conjunctive protocols we only need to take into account adversary strategies that give answers according to some function and discard all other possibilities because they can be discovered to be cheating with just two answers. This contrasts with the results shown in Example 15 (in the appendix) where the learner has to make too many queries (more than a constant) to discover a cheating teacher and therefore we cannot rule out the answering schemes that are not satisfiable.

5 Unspecified Attribute Value protocols

In this section we present our results for a concrete query set: the Unspecified Attribute Value (UAV) queries from [9, 5, 6]. Following [9], given a function $f \in B_n$ and a partial assignment

$\alpha \in \{0,1,\star\}^n$, we define $f(\alpha) = 1$ if and only if f is the constant 1 in H_α. Similarly, we define $f(\alpha) = 0$ if and only if f is the constant 0 in H_α. Otherwise, we say that $f(\alpha) =?$ (for unknown).

For a target concept f and a partial assignment α, the UAV membership query α returns YES when $f(\alpha) = 1$, NO if $f(\alpha) = 0$, and ? otherwise. In response to a UAV equivalence query with hypothesis h, the learner either is told that $h \equiv f$ or else is given a counterexample α, along with its classification, for which $f(\alpha) \neq h(\alpha)$. Note that the target function is a ternary function and all three values are distinct. So, for example, if $h(\alpha) = 1$ and $f(\alpha) =?$ then α could serve as a counterexample for h.

Below we formalize these queries in our framework and we show several results using our machinery.

5.1 Unspecified Attribute Value membership queries

Let H_α be the hypercube associated to the partial assignment α. The three possible answers to the UAV membership query α are modeled as follows

YES	NO	?
$\bigwedge_{y \in H_\alpha} \lambda_y$	$\bigwedge_{y \in H_\alpha} \overline{\lambda_y}$	$(\bigvee_{y \in H_\alpha} \lambda_y) \wedge (\bigvee_{y \in H_\alpha} \overline{\lambda_y})$

In a similar way as we did in the previous section we can get rid of the answering schemes that are not satisfiable. Let $PUAV_\in$ be the protocol for UAV membership queries.

Theorem 6 *For any $C \subseteq B_n$,*

$$Gdim(C, PUAV_\in) \leq \max(3, AIdim(C, PUAV_\in))$$

Proof As in the proof of Theorem 5 it is enough to show that any answering scheme T can be discovered to be unsatisfiable with at most 3 answers. Since T must be complete it must contain answers for all UAV queries on complete assignments (that do not contain $\star$s). If some of those answers is the corresponding function for ? then this single answer shows already the unsatisfiability (it will contain $\lambda_y \wedge \overline{\lambda_y}$ for some complete assignment y).

Otherwise we have a unique function that satisfies all the answers of the UAV membership queries with complete assignments. Call this function f. Now we proceed in a bottom up manner. We consider the answers to the UAV queries that contain $\star$s both from T and from f and we consider them by levels: first the queries with only one $\star$, then with two, and so on and so forth. We know, by the assumption made on the unsatisfiability of T, that at some point and for some query α the answer from T and the answer from f are different while all the answers to queries inside H_α coincide. Now let us consider some cases:

1. If T answers YES or NO to the query α and f YES or NO or ? (but not the same as T), then two answers suffice: the answer to α and some answer to a complete assignment in H_α.
2. If T answers ? and f answers YES or NO, then three answers suffice: the answer to α and the two answers to $\alpha \cup v \leftarrow 1$ and $\alpha \cup v \leftarrow 0$, for any variable v such that $v(\alpha) = \star$. □

The *extended teaching dimension* [10](see also [11] and [3]) characterizes the number of queries needed to learn in the standard membership queries learning protocol. We present below the counterpart of the extended teaching dimension for the $PUAV_\in$ protocol.

We need some new notation first. For a function $f \in B_n$ and a partial assignment $\alpha \in \{0,1,\star\}^n$ we say that f is single valued on α if f is constant on H_α. We denote by SV_f the set of all pairs (α, b) such that f is single valued on α and $f(\alpha) = b$. Given $S \subseteq SV_f$ and $g \in B_n$ we define $S(g)$ as

$$((\bigwedge_{y \in H_\alpha \wedge (\alpha,1) \in S} \lambda_y) \wedge (\bigwedge_{z \in H_\beta \wedge (\beta,0) \in S} \overline{\lambda}_z))(g).$$

Let $C \subseteq B_n$. We define the *UAV extended teaching dimension*, uavetdim(C), as the smallest integer $d \geq 0$ such that

$$(\forall f \in B_n)\,(\exists S \subseteq SV_f)\,((\|S\| \leq d) \wedge (\|\{c \in C \mid S(c)\}\| \leq 1)).$$

Note that this is a straightforward generalization of the extended teaching dimension where complete assignments have been replaced by single valued hypercubes.

Theorem 7 *For any $C \subseteq B_n$,*

$$AIdim(C, PUAV_\in) \leq uavetdim(C) \leq 2AIdim(C, PUAV_\in)$$

Proof First observe that any choice of S in the definition of uavetdim(C) is a valid choice for AIdim$(C, PUAV_\in)$ which implies directly the first inequality.

To prove the second inequality consider a function $f \in B_n$, an answering scheme T in $PUAV_\in^f$ and $\langle \alpha, ? \rangle \in T$. Let A be the set of target concepts that are not single valued on H_α. Let x and y two assignments that satisfy α and such that $f(x) \neq f(y)$. Clearly, the set of target concepts which agree with funcion f on x and y is a subset of A. This implies that we can substitute any ? answer by two conjunctive answers (YES or NO) to two UAV queries and the set of surviving functions from C does not grow. So if the set S of answers that shows AIdim$(C, PUAV_\in) = d$ contains some ? answers, we can replace each of them by two YES and NO answers corresponding to two examples in SV_f, which concludes the proof. □

Note that all this trick has consisted in removing the non conjunctive answers. The following corollary is immediate.

Corollary 8 *For any class $C \subseteq B_n$,*

$$\frac{uavetdim(C)}{2} \leq LC(C, PUAV_\in) \leq (uavetdim(C) + 2)\lceil \ln \|C\| \rceil$$

Our point is that with the abstract dimension uavetdim the proofs of many results are simplified a lot as we show in Example 16 and Example 19 that we have placed in the appendix due to space limitations.

5.2 Unspecified Attribute Value equivalence queries

We can formalize the UAV equivalence queries with hypothesis coming from $H \subseteq B_n$ as the protocol $PUAV_\equiv(H)$ whose answers to query $h \in H$ are:

Positive counter.	Negative counter.	? counter.	YES
$\bigwedge_{y \in H_\alpha} \lambda_y$ such that $h(\alpha) \neq 1$	$\bigwedge_{y \in H_\alpha} \overline{\lambda}_y$ such that $h(\alpha) \neq 0$	$(\bigvee_{y \in H_\alpha} \lambda_y) \wedge (\bigvee_{y \in H_\alpha} \overline{\lambda}_y)$ such that $h(\alpha) \neq ?$	$(\bigwedge_{h(y)=1} \lambda_y) \wedge (\bigwedge_{h(z)=0} \overline{\lambda}_z)$

With respect to AIdim this protocol has a different behavior from $PUAV_\in$. Example 21 (in the appendix) shows that AIdim is not a proper notion of dimension in $PUAV_\equiv(H)$.

5.3 UAV queries and the learnability of DNF Boolean formulas

Finally, we study the protocol $PUAV_{\in,\equiv}(H)$ that considers UAV membership and UAV equivalence queries with hypothesis in $H \subseteq B_n$. First, we observe that once again AIdim is a proper notion of dimension for $PUAV_{\in,\equiv}(H)$. As in this protocol we can do UAV membership queries, minor changes on the proof of Theorem 6 also shows

Theorem 9 *Let P be the protocol $PUAV_{\in,\equiv}(H)$. For any $C \subseteq B_n$,*

$$Gdim(C,P) \leq \max(3, AIdim(C,P))$$

The *certificate size* or *consistency dimension* [11] [4], formally defined in Example 16 in the appendix, characterizes the query complexity in the protocol of standard membership and equivalence queries. Theorem 9 is the key that allows to extend this notion in the protocol $PUAV_{\in,\equiv}(H)$. Similarly as we did in 5.1 introducing the UAV extended teaching dimension, we define the *UAV consistency dimension* for a target class $C \subseteq B_n$ and hypothesis class H, $C \subseteq H \subseteq B_n$, as the smallest integer $d \geq 0$ such that

$$(\forall f \in B_n)(f \notin H \Rightarrow (\exists S \subseteq SV_f)(\|S\| \leq d \wedge \{c \in C \mid S(c)\} = \emptyset)).$$

We denote by $\mathrm{uavcdim}(C,H)$ this number. Again, this is a direct generalization of the consistency dimension.

Theorem 10 *Let P be the protocol $PUAV_{\in,\equiv}(H)$. For any target class $C \subseteq B_n$ and hypothesis class $H \subseteq B_n$ such that $C \subseteq H$, if $uavcdim(C,H) > 0$ then*

$$AIdim(C,P) \leq uavcdim(C,H) \leq 2AIdim(C,P) + 1$$

Proof (sketch) We consider the first inequality. Let $f \in B_n$ be any function. If $f \in H$, the only answer of P^f to the equivalence query with hypothesis f discards all functions but at most one in C. Otherwise we choose the set of answers promised by the definition of $\mathrm{uavcdim}(C,H)$ to rule out all functions in C.

We prove now the second inequality. Let $f \notin H$ be a Boolean function and let S be the set of answers that shows $\mathrm{AIdim}(C,P) = d$. Using the trick present in the proof of Theorem 7, we can replace each ? answer in S by two answers YES and NO, and in such way that the set of functions ruled out does not decrease. Finally, if there is still one function c in C such that $S(c)$, we can get rid of c adding to S a labelled example (literal) that shows $f \neq c$. □

The learnability of DNF formulas using membership and equivalence queries is an important open problem. In [9, 5] is shown that $PUAV_{\in,\equiv}$ is stronger than the standard protocol for membership and equivalence queries. One may think that replacing the standard protocol by $PUAV_{\in,\equiv}$ some progress could be obtained. Below we show that for DNF the consistency and UAV consistency dimensions are close. So, as the learnability problem is equivalent in both protocols (recall that we are talking about query complexity only), we do not get any essential advantage of using UAV membership queries. We fix C as the class of m-term DNF where $m \geq 4$.

Lemma 11 *Let $\alpha_1, \ldots, \alpha_l$ and $\beta_1, \ldots, \beta_k$ partial assignments such that no function in C can evaluate 1 on $\cup_i H_{\alpha_i}$ and 0 on $\cup_j H_{\beta_j}$. Then, there exists a set S of complete assignments, $S \subseteq \cup_i H_{\alpha_i}$ such that*

$$(\forall \gamma \in \{0,1,\star\}^n)(\|H_\gamma \cap S\| > \|S\| n/m \Rightarrow H_\gamma \cap (\cup_j H_{\beta_j}) \neq \emptyset)$$

Proof For the sake of contradiction assume that for all $S \subseteq \cup_i H_{\alpha_i}$ there exists a partial assignment γ such that $\|H_\gamma \cap S\| > \|S\| n/m$ and $H_\gamma \cap (\cup_j H_{\beta_j}) = \emptyset$.

Let $S_1 = \cup_i H_{\alpha_i}$ and let γ_1 be the the partial assignment promised above. Then, the term t_{γ_1} (see definition in Section 2) evaluates 1 in H_{γ_1} and 0 in $\cup_j H_{\beta_j}$. Now, we consider $S_2 = S_1 - (H_{\gamma_1} \cap S_1)$ and we repeat the process. After m iterations we have a m-term DNF $t_{\gamma_1} \vee \ldots \vee t_{\gamma_m}$ that evaluates 1 on $\cup_i H_{\alpha_i}$ and 0 on $\cup_j H_{\beta_j}$. □

Lemma 12 *Let T be a set of complete assignments and let $\beta_1, \ldots, \beta_k$ be partial assignments such that*

$$(\forall \gamma \in \{0,1,\star\}^n)(\|H_\gamma \cap T\| \geq \|T\|/(m^2 n) \Rightarrow H_\gamma \cap (\cup_j H_{\beta_j}) \neq \emptyset),$$

then there exists a set $S \subseteq T$ of cardinality $\|S\| \leq m^2 n/3$ such that

$$(\forall \gamma \in \{0,1,\star\}^n)(\|H_\gamma \cap S\| \geq \|S\|/m \Rightarrow H_\gamma \cap (\cup_j H_{\beta_j}) \neq \emptyset).$$

Proof We choose $S \subseteq T$ of size $\lfloor m^2 n/3 \rfloor$ at random and let p be the probability of the event

$$(\exists \gamma \in \{0,1,\star\}^n)(\|H_\gamma \cap S\| \geq \|S\|/m \wedge H_\gamma \cap (\cup_j H_{\beta_j}) = \emptyset).$$

By the union bound p is at most $3^n q$, where q is the probability of the event $\|H_\gamma \cap S\| \geq \|S\|/m \wedge H_\gamma \cap (\cup_j H_{\beta_j}) = \emptyset$ for a fixed $\gamma \in \{0,1,\star\}^n$. By hypothesis,

$$H_\gamma \cap (\cup_j H_{\beta_j}) = \emptyset \Rightarrow \|H_\gamma \cap T\| < \|T\|/(m^2 n).$$

The bound $\binom{n}{k} p^k$ for the probability of obtaining at least k heads in n trials of a p biassed coin shows that q is at most

$$\binom{\|S\|}{\lceil \|S\|/m \rceil} \left(\frac{1}{m^2 n}\right)^{\lceil \|S\|/m \rceil}.$$

This bound for q shows that $p < 1$. □

Lemma 13 *Given $y \in \{0,1\}^n$ and $\beta \in \{0,1,\star\}^n$ there exists $y^\beta \in H_\beta$ such that:*

$$(\forall \gamma \in \{0,1,\star\}^n)(y \in H_\gamma \wedge H_\gamma \cap H_\beta \neq \emptyset \Rightarrow y^\beta \in H_\gamma).$$

Proof Consider the translation $\sigma_y : \{0,1\}^n \to \{0,1\}^n$ defined by $\sigma_y(x) = x \oplus y \oplus 1^n$. Note that $\sigma_y \circ \sigma_y$ is the identity function. Let z^β the maximum of the hypercube $\sigma_y(H_\beta)$ and let $y^\beta = \sigma_y(z^\beta)$. Note that y^β belongs to H_β.

Let γ be a partial assignment such that $y \in H_\gamma$ and $H_\gamma \cap H_\beta \neq \emptyset$. As we see below, these assumptions force $y^\beta \in H_\gamma$:

$$y^\beta \in H_\gamma \Leftrightarrow t_\gamma(y^\beta) = 1 \Leftrightarrow t_\gamma \circ \sigma_y(z^\beta) = 1.$$

We note that $t_\gamma \circ \sigma_y(z^\beta) = 1$ must be true because of: (1) $y \in H_\gamma$ implies $t_\gamma \circ \sigma_y$ is a monotone term, (2) z^β is the maximum of the hypercube $\sigma_y(H_\beta)$ and (3) $H_\gamma \cap H_\beta \neq \emptyset$ implies that $t_\gamma \circ \sigma_y$ is not the null function in $\sigma_y(H_\beta)$. □

Theorem 14 *Let C be the class of m-term DNF and let D be the class of s-term DNF where $m > s^2 n^2$ and $s \geq 4$. Let H be any hypothesis class such that $C \subseteq H$. It holds*

$$cdim(D, H) \leq (s^2 n/3)\, uavcdim(C, H)$$

Proof (sketch) Let $f \in B_n$ be a Boolean function and let $\{(\alpha_i, 1) \mid 1 \leq i \leq k\} \cup \{(\beta_j, 0) \mid 1 \leq j \leq l\}$ be the set of pairs in SV_f promised by the UAV consistency dimension. Thus, no function in C can evaluate 1 on $\cup_i H_{\alpha_i}$ and 0 on $\cup_j H_{\beta_j}$.

Apply Lemma 11 on $\alpha_1, \ldots, \alpha_k$ and $\beta_1, \ldots, \beta_l$, and afterwards, Lemma 12 replacing m by s (note that fraction n/m from Lemma 11 is less than $1/(s^2 n)$). We obtain a set $S \subseteq \cup_i H_{\alpha_i}$ of complete assignments of cardinality bounded by $s^2 n/3$ such that no term can evaluate 1 on a fraction $1/s$ of S and 0 on $\cup_j H_{\beta_j}$. Finally, for each pair $y \in S$ and $\beta_j, 1 \leq j \leq l$, we choose the assignment y^{β_j} promised by Lemma 13. It is easy to see that no function in D can evaluate 1 on S and 0 on $\{y^{\beta_j} \mid y \in S \wedge 1 \leq j \leq l\}$. □

6 Acknowledgments

The last subsection of the paper strongly benefits from conversations with Vijay Raghavan during his sabbatical in Barcelona.

References

[1] D. Angluin. Queries and concept learning. *Machine Learning*, 2:319–342, 1988.

[2] D. Angluin. Negative results for equivalence queries. *Machine Learning*, 5:121–150, 1990.

[3] J. L. Balcázar, J. Castro, and D. Guijarro. A new abstract combinatorial dimension of exact learning via queries. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 248–254. Morgan Kaufmann, San Francisco, 2000.

[4] J. L. Balcázar, J. Castro, D. Guijarro, and H. U. Simon. The consistency dimension and distribution-dependent learning from queries. In *ALT'99*, volume 1720, pages 77–92. LNAI. Springer, 1999.

[5] A. Birkendorf, N. Klasner, C. Kuhlmann, and H. U. Simon. Structural results about exact learning with unspecified attribute values. *J. of Comput. Syst. Sci.*, 60(2):258–277, 2000. Special Issue for COLT '98.

[6] N. H. Bshouty and D. K. Wilson. On learning in the presence of unspecified attribute values. In *Proc. 12th Annu. Conf. on Comput. Learning Theory*, pages 81–87. ACM Press, New York, NY, 1999.

[7] R. Gavaldà. On the power of equivalence queries. In *Computational Learning Theory: Eurocolt '93*, pages 193–203. Oxford University Press, 1994.

[8] S. Goldman and M. Kearns. On the complexity of teaching. *Journal of Computer and System Sciences*, 50:20–31, 1995.

[9] S. A. Goldman, S. S. Kwek, and S. D. Scott. Learning from examples with unspecified attribute values. In *Proc. 10th Annu. Conf. on Comput. Learning Theory*, pages 231–242. ACM Press, New York, NY, 1997.

[10] T. Hegedüs. Generalized teaching dimension and the query complexity of learning. In *8th Annu. Conf. on Comput. Learning Theory*, pages 108–117, New York, 1995. ACM Press.

[11] L. Hellerstein, K. Pillaipakkamnatt, V. Raghavan, and D. Wilkins. How many queries are needed to learn? *Journal of the ACM*, 43(5):840–862, 1996.

[12] N. Klasner and H. U. Simon. General lower bounds on the query complexity within the exact learning model. *Discrete Applied Mathematics*, 107:61–81, 2000.

A Examples

Example 15

Let us consider P as the protocol of restricted equivalence queries with hypothesis in B_n. For any class $C \subseteq B_n$, it holds $\mathrm{AIdim}(C,P) \leq 1$. We note that any answering scheme T in $\mathcal{T}(P^f)$ has, as answer to the query f, the conjunction $(\bigwedge_{f(y)=1} \lambda_y) \wedge (\bigwedge_{f(y)=0} \overline{\lambda}_y)$ that represents a YES answer. Exactly one function satisfies this answer, f itself.

However, it is well known that if C has N functions and $N > 0$, the learning complexity is $N-1$. In fact, the following answering scheme T shows that $\mathrm{Gdim}(C,P) > N-2$: each query is answered in T negatively (with the clause). Note that each of those answers is unsatisfied by one function from B_n, namely the function being queried. Therefore each of those answers is falsified by at most one function from C which implies that with $N-2$ answers there will be at least two satisfying functions from C.

Example 16

Let P be the protocol of standard membership and equivalence queries with hypothesis in H (see [3]). The *certificate size* or *consistency dimension* [11] [4] of a target class C ($C \subseteq H \subseteq B_n$), denoted by $\mathrm{cdim}(C,H)$, is the smallest integer $d \geq 0$ such that

$$(\forall f \in B_n)(f \notin H \Rightarrow (\exists S \subseteq P_Q^f)(\|S\| \leq d \wedge \{c \in C \mid S(c)\} = \emptyset)).$$

It is known thet $\mathrm{cdim}(C,H)$ is a good approximation of $\mathrm{LC}(C,P)$.

The CDNF size of a Boolean function is the maximun among their CNF and DNF sizes. We fix the hypothesis class H as the set of functions with CDNF size at most m. Let $C \subseteq H$ be any target class. The next theorem states that $\mathrm{cdim}(C,H)$ is related to $\mathrm{uavetdim}(C)$.

Theorem 17 *For classes C and H just defined, it holds*

$$\mathit{uavetdim}(C) \leq \max(2m, \mathit{cdim}(C,H))$$

Proof Let f be any function in B_n. If $f \in H$ we choose as set S in the definition of uavetdim a set of cardinality at most $2m$ whose elements are

$$\{(\alpha,1) \mid f = \bigvee t_\alpha\} \cup \{(\beta,0) \mid f = \bigwedge c_\beta\}.$$

Here, t_α and c_β are the term and the clause defined in section 2.

If $f \notin H$ we can choose as S the set of literals (examples) promised by the definition of cdim.
□

As a consequence of this theorem we have,

Corollary 18 *If C is polynomially query learnable by membership and equivalence queries in the CDNF size, then C is polinomially UAV-MQ query learnable.*

A stronger version of this Corollary is shown in [6], where a time efficient transformation is also proved.

Example 19

We consider $C \subseteq B_n$ as the set of functions with CDNF size at most m. In [6] is shown that the target class C is polynomially query learnable in the protocol $PUAV_{\in}$. The proof is based on Corollary 18 and in a result in [11] that shows that C is polynomially query learnable by membership and equivalence queries. We give here an direct proof of this fact using our abstract machinery. First, we introduce some new notation.

Given a partial assignment α and a function $f \in B_n$ we denote by $f_{[\alpha]}$ the projection of $f \in B_n$ with respect to α. On input x, $f_{[\alpha]}(x) = f(x_\alpha)$ where x_α is the assignment in H_α of minimum Hamming distance with respect to x.

Theorem 20 *Let C be the class of functions in B_n with CDNF size at most m, it holds*

$$uavetdim(C) \leq 4m$$

Proof (sketch) Let f be a function in B_n. If f has CDNF size at most $2m$, we proceed as in the first part of the proof of Theorem 17.

If f has CDNF size greater than $2m$ we use a projection trick. We choose a partial assignment α such that the CDNF size of $f_{[\alpha]}$, denoted by $|f_{[\alpha]}|_{\mathrm{cdnf}}$, is greater than $2m$ and for some variable v both projections $f_{[\alpha \cup v \leftarrow 0]}$ and $f_{[\alpha \cup v \leftarrow 1]}$ have CDNF size bounded by $2m$. Because of $f_{[\alpha]} = (\overline{v} \wedge f_{[\alpha \cup v \leftarrow 0]}) \vee (v \wedge f_{[\alpha \cup v \leftarrow 1]})$, and $f_{[\alpha]} = (\overline{v} \vee f_{[\alpha \cup v \leftarrow 1]}) \wedge (v \vee f_{[\alpha \cup v \leftarrow 0]})$, at least one of the two projections must have CDNF size greater than m. Therefore there exists $b \in \{0,1\}$ such that $|f_{[\alpha \cup v \leftarrow b]}|_{\mathrm{cdnf}} > m$. Let β be the partial assignment $\alpha \cup v \leftarrow b$, let $t_{\gamma_1}, \ldots, t_{\gamma_i}$ be the terms of $t_\beta \wedge f_{[\beta]}$ and let $c_{\delta_1}, \ldots, c_{\delta_j}$ the clauses of $c_\beta \vee f_{[\beta]}$. As $|f_{[\beta]}|_{\mathrm{cdnf}} \leq 2m$ we can assume that $i + j \leq 4m$. Now we choose as set S in the definition of uavetdim the set whose elements are $(\gamma_1, 1), \ldots, (\gamma_i, 1)$ and $(\delta_1, 0), \ldots, (\delta_j, 0)$. These labelled examples reveal the values of f in H_β and rule out all functions in C becouse of: (1) $f_{[\beta]} \notin C$, (2) The projection $g_{[\beta]}$ of any function g that agrees with f on H_β is $f_{[\beta]}$ and (3) C is projection closed. $\square$

Example 21

We consider as target class $C = SING_n$, as hypothesis class $H = B_n$ and we denote by P the protocol $PUAV_{\equiv}(H)$. It is not difficult to see that $\mathrm{AIdim}(C, P) = 1$. Each answering scheme T in $\mathcal{T}(P^f)$ has the formula $(\bigwedge_{f(y)=1} \lambda_y) \wedge (\bigwedge_{f(z)=0} \overline{\lambda}_z)$ as anwer to the hypothesis f. At most one function in C satisfies this answer, f itself.

However, the following adversary shows that $\mathrm{LC}(C, P)$ is $2^n - 1$. Each UAV-EQ query with hypothesis $h \neq 0$ is answered $(x, 0)$ being x a complete assignment such that $h(x) = 1$. If h is the constant 0 the adversary returns $(\star^n, ?)$. Note that each of these answers discard at most one function in $SING_n$.