Classification Using Information*

William Gasarch[†]
University of Maryland

Mark G. Pleszkoch[‡]
IBM Corporation

Frank Stephan[§]
University of Heidelberg

Mahendran Velauthapillai[¶]
Georgetown University

Abstract

Let $\mathcal{A}$ be a set of functions. A *classifier for* $\mathcal{A}$ is a way of telling, given a function f, if f is in $\mathcal{A}$. We will define this notion formally. We will then modify our definition in three ways: (1) Allow the classifier to ask questions to an oracle A (thus increasing the classifiers *computational* power). (2) Allow the classifier to ask questions about f (thus increasing the classifiers *information access*). (3) Restrict the number of times the classifier can change its mind (thus decreasing the classifiers *information access*). By varying these parameters we will gain a better understanding of the contrast between computational power and informational access.

We have determined exactly (1) which sets are classifiable (Theorem 3.6), (2) which sets are classifiable with queries to some oracle (Theorem 3.2), (3) which sets are classifiable with queries to some oracle and queries about f, (Theorem 5.2), and (4) which sets are classifiable with queries to some oracle, queries about f, and a bounded number of mindchanges, (Theorem 5.2). The last two items involve the Borel hierarchy.

*The first, second and fourth author presented a preliminary version of this paper on the Conference on Algorithmic Learning Theory 1994 in Reinhardsbrunn [5].

[†]Dept. of C.S. and Inst. for Adv. Stud., University of Maryland, College Park, MD 20742, U.S.A. Supported in part by NSF grants CCR-8803641 and CCR-9020079 (email: gasarch@cs.umd.edu).

[‡]IBM Corporation, Gaithersburg, MD 20879, U.S.A., (email: markp@vnet.ibm.com).

[§]Mathematical Institute of the University of Heidelberg, Im Neuenheimer Feld 294, 69121 Heidelberg, Germany, EU, Supported by the Deutsche Forschungsgemeinschaft (DFG) grants Me 672/4-2 and AM 60/9-1 (email: fstephan@math.uni-Heidelberg.de).

[¶]Department of Computer Science, Georgetown University, Washington, D.C., 20057, U.S.A., (email: mahe@cs.georgetown.edu).

1 Introduction

Let $\mathcal{FS} = \{f \mid (\exists x)(\forall y)[y \geq x \Rightarrow f(x) = 0]\}$. ($\mathcal{FS}$ stands for 'finite support'.) If you were given $g(0), g(1), \ldots$ you could never classify g with respect to $\mathcal{FS}$, even in the limit. Even if you had access to K (or some other oracle) you could not classify g. The barrier to your classification is not computational, but is instead informational. By contrast, assume you could ask existential questions about g. Initially guess NO ($g \notin \mathcal{FS}$). Then ask the following questions until you get an answer of YES (this might never happen).

$$\begin{array}{l}(\forall x \geq 0)[g(x) = 0]?\\(\forall x \geq 1)[g(x) = 0]?\\(\forall x \geq 2)[g(x) = 0]?\\\vdots\end{array}$$

If an answer of YES is given then guess YES ($g \in \mathcal{FS}$). You have successfully classified g in the limit. We will see later that $\mathcal{FS}$ is dense and co-dense in the standard topology of the function space and that no dense and co-dense set is classifiable. Hence you really needed that additional information.

Let $\mathcal{UK} = \{f \mid (\forall x)[f(x) \leq K(x)]\}$. (The 'UK' stands for 'Under K'.) If you were given $g(0), g(1), \ldots$ you could never classify g with respect to $\mathcal{UK}$, even in the limit. However, if you had access to K, then you could classify g with respect to UK in the limit as follows: Guess YES until an x is spotted such that $g(x) > K(x)$, at which point change the guess to NO (and never change your mind thereafter). Hence the barrier to classification is computational.

When a class of functions cannot be classified it may be for either computational or information-theoretic reasons. Information-theoretic means that not enough information is available to classify. This is pinned down by topology; for the rest of the paper we will use the mathematically precise word 'topological' rather than the intuitive word 'information-theoretical.'

In the next section we define classification formally. We vary the amount of information the learner can access. To increase the model's ability to access information, we give it the ability to ask questions about the function. We also regulate the type of question by both restricting the query language and restricting the number of alternations of quantifiers a question can have. To decrease the models ability to access information, we will bound the number of mindchanges it may make.

Carl Smith and Rolf Wiehagen [13] introduced a model of classification that is similar to the Gold model of learning [7]. The classifier M sees longer and longer initial segments of the graph of a function f. At each segment it guesses either YES (for $f \in \mathcal{A}$) or NO (for $f \notin \mathcal{A}$). The guesses converge for each function f to the value $M(f) \in \{\text{YES}, \text{NO}\}$. $M(f) = \text{YES}$ means that $f \in \mathcal{A}$ and $M(f) = \text{NO}$ that

$f \notin \mathcal{A}$. In this model the classifier is limited in *both* computing power and access to information. In particular the learner is limited to Turing computability and initial segments of the function to be classified. Shai Ben-David [1] and Kevin Kelly [8] studied the same topic, but did not consider limitations of computational power.

2 Definitions and Notations

In this section we formalize our notions. N denotes the set $\{0, 1, 2, \ldots\}$ of natural numbers, Σ will denote a fixed set of symbols such that $\{0, 1\} \subseteq \Sigma \subseteq \mathsf{N}$. Σ^* denotes the set of all finite sequences of symbols in Σ. Σ^ω denotes the set of all countably infinite sequences of symbols in Σ. If $\sigma \in \Sigma^*$ and $f \in \Sigma^* \cup \Sigma^\omega$ then $\sigma \preceq f$ means that σ is a prefix of f. If $\sigma, \tau \in \Sigma^*$ then $\sigma\tau$ denotes their concatenation. We may use $\sigma \cdot \tau$ for clarity.

Throughout this section $\mathcal{A}$ denotes a subset of Σ^ω and $\overline{\mathcal{A}}$ denotes its complement; it can't be confused with the topological closure operation since the closure operation is not used in this paper. $\#A$ denotes the cardinality of a set A.

Definition 2.1 A *classifier* is a recursive function $M : \Sigma^* \to \{\text{YES, NO}, \text{DK}\}$ (DK stands for DON'T KNOW). Our intention is that M is fed initial segments of some f and eventually decides if it is in $\mathcal{A}$ or not. Let $f : \mathsf{N} \to \Sigma$ be a function. *M classifies f with respect to $\mathcal{A}$* if (1) when M is given initial segments of f as input, the resultant sequence of answers converges (after some point there are no more mindchanges) (2) if $f \in \mathcal{A}$ then the sequence converges to YES, and (3) if $f \notin \mathcal{A}$ then the sequence converges to NO.

Note 2.2 In the above definition we restrict a classifier to be a recursive function that only has access to the function via initial segments. We will later allow classifiers to have access to oracles and/or be able to ask questions about the function. The type of classifier will be clear from context.

Definition 2.3 *M classifies $\mathcal{A}$* if, for every function f, M classifies f with respect to $\mathcal{A}$. The class DE is the collection of all sets $\mathcal{A}$ such that there exists a classifier M that classifies $\mathcal{A}$ (DE stands for DEcision). We denote this by saying "$\mathcal{A} \in$ DE via M." Formally M is a function; however, we will often describe it as a process that continually receives values of f (in order) and outputs conjectures. Such a description can clearly be restated in terms of M being a function. The class DE$[A]$ denotes decision relative to an oracle A and DE$[all]$ is the collection of all classes DE$[A]$:

$$\mathcal{A} \in \text{DE}[all] \Leftrightarrow \mathcal{A} \in \text{DE}[A] \text{ for some set } A \subseteq \mathsf{N}.$$

The class DE_c is the collection of all sets in DE that have classifiers that change their mind about each f at most c times. The initial change from DK to either YES or NO is not counted as a mindchange. We will mostly be concerned with $\mathrm{DE}[all]$ since we wish to study how much information is needed independent of computational resources.

We now define classifiers that can make queries. This is analogous to the query inference machines defined by Gasarch and Smith [6].

Definition 2.4 A *query language* consists of the usual logical symbols (and equality), symbols for first order variables, symbols for every element of N, symbols for some functions and relations on N, and a special symbol f. A query language is denoted by the symbols for these functions and relations, A well-formed formula over L is defined in the usual way.

Convention 2.5 Small letters are used for first order variables which range over N. All questions are assumed to be sentences in prenex normal form (quantifiers followed by a quantifier-free formula, called the matrix of the formula) and questions containing quantifiers are assumed to begin with an existential quantifier. This convention entails no loss of generality. The special symbol f will represent the function we are trying to classify.

Definition 2.6 Let L be a query language. A *query over* L is a formula $\phi(f)$ such that the following hold.

(I) $\phi(f)$ uses symbols from L.

(II) f is a free function variable and is the only free variable.

We think of a query $\phi(f)$ as asking a question about an as yet unspecified function f. If f is a function then $\phi(f)$ will be either true or false.

Definition 2.7 Let L be a query language. Informally, a *classifier over* L (usually just 'classifier') is a total Turing machine that can ask questions about the recursive function f in the language L and by using the answers to these questions, eventually outputs 0 or 1 in the limit. Formally a *classifier* is is a total Turing machine M, which takes as input a string of bits σ (the empty string is allowed), corresponding to the answers to previous queries about f, outputs first one value $M(\sigma) \in \{\mathrm{YES}, \mathrm{NO}, \mathrm{DK}\}$ in order to indicate whether it at the moment guesses $f \in \mathcal{A}$ and second a new question $\phi(\sigma)$ in the language L. Our intention is that M is conjecturing whether f is in $\mathcal{A}$ or not and also generating the next question to ask about f. The definition of when M *classifies* f *with respect to* $\mathcal{A}$ is straightforward but tedious (it is analogous to the definition in [6]).

Definition 2.8 Let L be a query language. The class QDE$[L]$ is the collection of all sets $\mathcal{A}$ such that there exists a classifier that classifies $\mathcal{A}$ and only asks queries that use the symbols in L. We denote this by saying "$\mathcal{A} \in \mathrm{QDE}[L]$ via M." The class $\mathrm{QDE}_a[L]$ is the collection of all sets in QDE$[L]$ that have classifiers that change there mind about each f at most a times. The initial change from DK to either YES or NO is not counted as a mindchange. Furthermore QDE$[all]$ is the union of all classes QDE$[L]$ as L goes over all possible query languages.

All the query languages that we will consider allow the use of quantifiers. Restricting the applications of quantifiers is a technique that we will use to regulate the expressive power of a language. Of concern to us is the alternations between blocks of existential and universal quantifiers.

Definition 2.9 Suppose that $f \in \mathrm{QDE}[L](M)$ for some M and L. If M only asks quantifier-free questions, then we will say that $f \in \mathrm{Q_0DE}[L](M)$. If M only asks questions with existential quantifiers, then we will say that $f \in \mathrm{Q_1DE}[L](M)$. In general, if M's questions begin with an existential quantifier and involve a alternations between blocks of universal and existential quantifiers, then we say that $f \in \mathrm{Q}_{a+1}\mathrm{DE}[L](M)$. The classes $\mathrm{Q}_c\mathrm{DE}[L]$ and $Q_c\mathrm{DE}_b[L]$ are defined analogously.

Note 2.10 We use the notations DE$[A]$ and QDE$[L]$. In the first case the A is a set which we ask question to and in the second the language L is a language we express questions in. Note that in DE$[A]$ we are allowing more computational power to the inference device and in QDE$[L]$ we are allowing greater access of information. One of the points of this paper will be to compare computational to information.

3 Classification with Oracles

The class DE$[all]$ has various topological characterizations. In this section we present the main ones.

Definition 3.1 The following two topological spaces are useful.

(i) $\mathcal{F}$ is the set of all functions from N to Σ. We place a topology on it by letting the basic open sets be $\mathcal{F}_\sigma = \{f \mid \sigma \preceq f\}$ where σ ranges over Σ^*.

(ii) $\mathcal{N}$ is the set N. We place a topology on it by letting the basic open sets be $\mathcal{N}$, $\emptyset$, and and all sets of the form $\{y \in \mathsf{N} \mid y \geq x\}$ with $x \in \mathsf{N}$.

Theorem 3.2 *$\mathcal{A}$ is in* DE[*all*] *iff there is a continuous function $F : \mathcal{F} \to \mathcal{N}$ such that $\mathcal{A} = \{f \mid F(f)$ is odd $\}$.*

Proof: Recall that F is continuous iff the inverse image of every open subset of $\mathcal{N}$ is an open set in $\mathcal{F}$."

Let M be an classifier which witnesses $\mathcal{A} \in$ DE[A] for some oracle A. We can assume that $M(\emptyset) =$ NO. Now let $F(\sigma)$ denote the number of mindchanges on input σ; note that $F(\sigma)$ is even iff $M(\sigma) =$ NO. Classifying each function f, M makes only finitely many mindchanges and thus $F(f) = \lim_{\sigma \preceq f} F(\sigma)$ exists for each function f. Now $f \in \mathcal{A}$ iff M converges on f to YES iff M makes an odd number of mindchanges on f iff $F(f)$ is odd. It remains to show that F is a continuous function from $\mathcal{F}$ to $\mathcal{N}$.

Let $y \in \mathsf{N}$, $\mathcal{U}_y = \{f \mid F(f) \geq y\}$ and $f \in \mathcal{U}_y$. There is a $\sigma \preceq f$ such that $F(\sigma) \geq F(f)$. By the definition of F, $F(\tau) \geq F(\sigma)$ for all $\tau \succeq \sigma$ and $F(g) \geq F(\sigma) \geq y$ for all $g \succeq \sigma$. Thus the basic open set $\mathcal{F}_\sigma$ is contained in $\mathcal{U}_y$; so $\mathcal{U}_y$ is the union of basic open sets; therefore $\mathcal{U}_y$ is open and F is continuous.

For the other way round, let $F : \mathcal{F} \to \mathcal{N}$ be a continuous function and $\mathcal{A} = \{f \mid F(f)$ is odd$\}$. Now for each σ let $F(\sigma) = \min\{F(f) \mid \sigma \preceq f\}$. $F(\sigma)$ is defined since the natural numbers are well-ordered. Let $y = F(f)$. Since F is continuous there is a string $\sigma \preceq f$ such that $F(g) \geq y$ for all $g \succeq \sigma$. Therefore $F(\tau) = y$ for all τ with $\sigma \preceq \tau \preceq f$ and the classifier

$$M(\sigma) = \begin{cases} \text{YES} & \text{if } F(\sigma) \text{ is odd;} \\ \text{NO} & \text{if } F(\sigma) \text{ is even;} \end{cases}$$

decides $\mathcal{A}$: If $F(f)$ is even then M converges on f to NO and if $F(f)$ is odd then M converges on f to YES. ∎

Corollary 3.3 *Let $\mathcal{A} \subseteq \Sigma^\omega$. Assume $\mathcal{A} \in$ DE[all]. Then* (1) *there is a σ such that either $\mathcal{F}_\sigma \subseteq \mathcal{A}$ or $\mathcal{F}_\sigma \subseteq \overline{\mathcal{A}}$, and* (2) *the topological boundary $\partial\mathcal{A}$ is nowhere dense. Hence $\mathcal{FS} = \{f \mid (\forall^\infty x)[f(x) = 0]\} \notin$ DE[all].*

Proof: Assume that $\mathcal{A} \in$ DE[*all*] witnessed by a continuous $F : \mathcal{F} \to \mathcal{N}$. Again extend F onto the finite strings $\sigma \in \Sigma^*$ by $F(\sigma) = \min\{F(f) \mid \sigma \preceq f\}$. Let $\sigma_0 = \lambda$. As long as possible find an extension $\sigma_{n+1} \succ \sigma_n$ with $F(\sigma_{n+1}) > F(\sigma_n)$. If this process never terminates, then $F(f) \geq F(\sigma_n) \geq n$ for the limit f of all σ_n; but this contradicts the fact that $F(f) \in \mathsf{N}$. Therefore the process stops for some σ_n. Now $F(\tau) = F(\sigma_n)$ for all $\tau \geq \sigma_n$ and therefore $F(g) = F(\sigma_n)$ for all $g \succeq \sigma_n$. The basic open set $\mathcal{F}_{\sigma_n}$ either belongs to $\mathcal{A}$ or to $\overline{\mathcal{A}}$.

This construction indeed provides such a basic open set above any given string. Thus each string τ is extended by some σ with either $\mathcal{F}_\sigma \subseteq \mathcal{A}$ or $\mathcal{F}_\sigma \subseteq \overline{\mathcal{A}}$. Thus $f \notin \partial\mathcal{A}$ for all $f \succeq \sigma$ and $\partial\mathcal{A}$ is nowhere dense.

Since $\partial\mathcal{FS} = \mathcal{F}$, $\mathcal{FS} \notin \mathcal{A}$. To see that every $f : \mathsf{N} \to \Sigma$ is in $\partial\mathcal{FS}$, note that for each $\sigma \preceq f$, $\sigma 0^\omega \in \mathcal{FS}$ and $\sigma 1^\omega \notin \mathcal{FS}$, thus f is approximated by a sequence inside $\mathcal{FS}$ and an other sequence outside $\mathcal{FS}$. So f is in the boundary of $\mathcal{FS}$. ∎

Similarly one can show that $\mathcal{B} = \{\sigma 0^\omega \,|\, \sigma \in \{0,1\}^*\} \cup \{f \,|\, (\exists x)[f(x) \geq 2]\}$ is not in DE[*all*] for $\Sigma = \{0,1,2\}$. $\partial\mathcal{B}$ is nowhere dense since $\mathcal{F}_{\sigma\cdot 2} \subseteq \mathcal{B}$ for all σ. So the first two statements of the corollary are not "if and only if". Another topological characterization is based on the following observation:

Theorem 3.4 *If $\mathcal{A}$ is open in $\mathcal{F}$ then $\mathcal{A} \in \mathrm{DE}_1[all]$.*

Proof: Since $\mathcal{A}$ is open, $\mathcal{A} = \bigcup_{\sigma\in W} \mathcal{F}_\sigma$ for some set W. Without loss of generality we can assume that $W = \{\sigma \,|\, \mathcal{F}_\sigma \subseteq \mathcal{A}\}$. Now the classifier M given by

$$M(\sigma) = \begin{cases} \text{YES} & \text{if } \sigma \in W; \\ \text{NO} & \text{if } \sigma \notin W; \end{cases}$$

$\mathrm{DE}_1[W]$ classifies $\mathcal{A}$. If W is r.e., then even $\mathcal{A} \in \mathrm{DE}_1$.

An alternative proof — which only shows $\mathcal{A} \in \mathrm{DE}[all]$ — uses the topological characterization of Theorem 3.2: Let F be the characteristic function of $\mathcal{A}$, i.e., $F(f) = 1$ if $f \in \mathcal{A}$ and $F(f) = 0$ if $f \notin \mathcal{A}$. Then the inverse images of the open set N is $\mathcal{F}$, of the open set $\{x \,|\, x \geq 1\}$ is the open set $\mathcal{A}$ and of all other open sets is $\emptyset$. Since $\emptyset$ and $\mathcal{F}$ are also open, F is continuous. ∎

So one might ask, how the class of all sets in DE[*all*] can be generated from the open sets. The answer follows from the following definition:

Definition 3.5 Let $\mathcal{C}$ be a collection of subsets of Σ^ω. $\mathcal{A}$ is a the *well-defined symmetric difference of* $\mathcal{C}$ (denoted $\mathcal{A} = \mathrm{WDSD}(\mathcal{C})$) if $\mathcal{A}$ consists of all f such that (1) f is contained only in finitely many sets $\mathcal{B} \in \mathcal{C}$, and (2) $\{\mathcal{B} \in \mathcal{C} \,|\, f \in \mathcal{B}\}$ has an odd number of elements.

Note that if $\mathcal{A}$ is a Boolean combination of open sets, then it is also the WDSD of a finite collection of open sets. Further if $\mathcal{A}$ is a WDSD of a collection of open sets, then $\mathcal{A}$ is also a Borel set. But none of these two implications have a reverse: $\{f \,|\, \min(f)$ is odd$\}$ is a WDSD of a collection of open sets but not the Boolean combination of finitely many open sets; $\mathcal{FS}$ is a Borel-set since $\mathcal{FS}$ is countable, but $\mathcal{FS}$ is not the WDSD of some collection of open sets. Now DE[*all*] has the following characterization:

Theorem 3.6 *$\mathcal{A} \in$ DE[all] iff $\mathcal{A}$ is the well-defined symmetric difference of some collection of open sets.*

Proof: Let $\mathcal{A} \in \mathrm{DE}[all]$ be given and $F : \mathcal{F} \to \mathcal{N}$ be the continuous function from Theorem 3.2 such that $f \in \mathcal{A} \Leftrightarrow F(f)$ is odd. Now let $\mathcal{C} = \{\mathcal{U}_y \mid y \in \mathsf{N}\}$ with $\mathcal{U}_y = \{f \mid F(f) \geq y\}$ for $y \geq 1$. All sets $\mathcal{U}_y$ are open and each f is in the finitely many sets $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_{F(f)}$. So $\mathcal{A} = \mathrm{WDSD}(\mathcal{C})$ and the "only if" direction holds.

Now let $\mathcal{A} = \mathrm{WDSD}(\mathcal{C})$ for some collection $\mathcal{C}$ of open sets. Further let $F(f)$ denote the cardinality of the set $\{\mathcal{B} \in \mathcal{C} \mid f \in \mathcal{B}\}$. By definition, $F(f)$ is odd iff $f \in \mathcal{A}$. It remains to show, that F is continuous. Let $\mathcal{C}_y$ be the collection of all sets which are the intersection of at least y different sets from $\mathcal{C}$. Then $F(f) \geq y$ iff there is some $\mathcal{B} \in \mathcal{C}_y$ with $f \in \mathcal{B}$. It follows that $\mathcal{U}_y = \{f \mid F(f) \geq y\}$ is just the union of all sets in $\mathcal{C}_y$ and so each set $\mathcal{U}_y$ is open. Therefore F is a continuous mapping from $\mathcal{F}$ to $\mathcal{N}$. ▮

There is an effective version of this theorem. This version works with *basic open sets* instead of *open sets*. This is needed since open sets can be highly nonrecursive, whereas basic open sets are recursive.

Theorem 3.7 *$\mathcal{A} \in$ DE iff $\mathcal{A} = \mathrm{WDSD}\{\mathcal{F}_\sigma \mid \sigma \in W\}$ for some r.e. set W, i.e., iff $\mathcal{A}$ is the well-defined symmetric difference of an r.e. collection of basic open sets.*

Proof: We establish the "only if" direction. If $\mathcal{A} \in \mathrm{DE}$ via some classifier M then let $W = \{\sigma a \mid M(\sigma) \neq M(\sigma a)\}$. We can assume, without loss of generality, that $M(\lambda) = \mathrm{NO}$ (if $M(\lambda) = \mathrm{YES}$ one has to add λ to W). W is even recursive. Since $f \in \mathcal{A}$ iff M makes an odd number of mind changes, $f \in \mathcal{A}$ iff there is an odd number of strings $\sigma \in W$ with $\sigma \preceq f$, i.e., iff $f \in \mathrm{WDSD}\{\mathcal{F}_\sigma \mid \sigma \in W\}$.

We now establish the "if" direction. Let $\mathcal{A} = \mathrm{WDSD}\{\mathcal{F}_\sigma \mid \sigma \in W\}$ for an r.e. set W. First W has to be replaced by a recursive set which is sufficiently similar to W. Let $\sigma_0, \sigma_1, \ldots$ be a recursive 1-1 enumeration of W such that $|\sigma_n| \leq n$ for all n; in order to achieve this condition, $\sigma_n = \#$ is allowed. Now the set

$$V = \{\tau \mid \sigma_{|\tau|} \neq \# \wedge \sigma_{|\tau|} \preceq \tau\}$$

is recursive and for each f the sets $\{\sigma \in W \mid \sigma \preceq f\}$ and $\{\tau \in V \mid \tau \preceq f\}$ have the same finite cardinality. Thus $\mathcal{A} = \mathrm{WDSD}\{\mathcal{F}_\tau \mid \tau \in V\}$. Now N given by

$$N(\eta) = \begin{cases} \mathrm{YES} & \text{if } \#\{\tau \in V \mid \tau \preceq \eta\} \text{ is odd;} \\ \mathrm{NO} & \text{otherwise } (\#\{\tau \in V \mid \tau \preceq \eta\} \text{ is even}); \end{cases}$$

is a recursive classifier which classifies $\mathcal{A}$. ▮

Shai Ben-David [1] found a further topological characterization based on the notion of countable unions of closed sets, the so called "F_σ-sets". The next theorem is his. We proof it for completeness.

Theorem 3.8 *$\mathcal{A} \in \mathrm{DE}[all]$ iff $\mathcal{A}$ and $\overline{\mathcal{A}}$ both are countable unions of closed sets.*

Proof: Let $\mathcal{A} = \{f \mid F(f) \text{ is odd}\}$ for a continuous function $F : \mathcal{F} \to \mathcal{N}$. The sets $\mathcal{U}_y = \{f \mid F(f) \geq y\}$ are open and each satisfies $\mathcal{U}_y = \bigcup_{\sigma \in W_y} \mathcal{F}_\sigma$ for certain sets W_y. The basic open sets are not only open, but also closed. The sets $\mathcal{F}_\sigma - \mathcal{U}_y$ are also closed. From the relation

$$\mathcal{A} \;=\; \bigcup_y \{f \mid F(f) = 2y+1\} \;=\; \bigcup_y \bigcup_{\sigma \in W_{2y+1}} (\mathcal{F}_\sigma - \mathcal{U}_{2y+2})$$

follows that $\mathcal{A}$ and similarly $\overline{\mathcal{A}}$ are the union of countably many closed sets.

For the other way assume that $\mathcal{A} = \mathcal{C}_1 \cup \mathcal{C}_3 \cup \mathcal{C}_5 \ldots$ and $\overline{\mathcal{A}} = \mathcal{C}_0 \cup \mathcal{C}_2 \cup \mathcal{C}_4 \ldots$ are the countable unions of the closed sets $\mathcal{C}_0, \mathcal{C}_1, \ldots$; further let $F(f)$ be the first y such that $f \in \mathcal{C}_y$. Since the $\mathcal{C}_y$ cover the whole set $\mathcal{F}$, $F(f)$ is defined and $F(f)$ is odd iff $f \in \mathcal{A}$. The function F is continuous since the sets

$$\{f \mid F(f) \geq y\} = \overline{\mathcal{C}_0 \cup \mathcal{C}_1 \cup \ldots \cup \mathcal{C}_{y-1}}$$

are open: each of them is the complement of a finite union of closed sets. ∎

The next theorem shows that $\mathrm{DE}[A]$ and $\mathrm{DE}[B]$ have a simple relationship.

Theorem 3.9 *$\mathrm{DE}[A] \subseteq \mathrm{DE}[B]$ iff $A \leq_T B$.*

Proof: The "if"-direction is clear, for the "only-if"-direction let $\mathrm{DE}[A] \subseteq \mathrm{DE}[B]$, $\Sigma = \{0,1\}$, $f = \chi_A$ and $\mathcal{A} = \{f\}$. Now $\{f\} \in \mathrm{DE}[B]$ via some classifier M^B. For all sufficiently long input $\sigma \preceq f$ the classifier outputs YES, by a finite modification one can obtain $M^B(\sigma) = \text{YES}$ for all $\sigma \preceq f$. A further modification gives, that M^B makes at most one mind change: if $M^B(\sigma) = \text{NO}$, then one can set $M^B(\tau) = \text{NO}$ for all $\tau \succeq \sigma$ since no $g \succeq \sigma$ is in $\{f\}$. Thus $T = \{\sigma \in \Sigma^* \mid M^B(\sigma) = \text{YES}\}$ is a B-recursive tree with f being its only infinite branch. Therefore $f \leq_T B$ and $A \leq_T B$. ∎

4 Arbitrary Query-Languages

This section looks for relations between the number of quantifiers (allowed in queries) and bounds on mindchanges. Queries allow one to extract more information than just looking at initial segments. For example $\mathcal{FS} \in \mathrm{Q}_2\mathrm{DE}_0[\emptyset] - \mathrm{DE}[all]$. Most results in this section do not depend on a specific query language.

Theorem 4.1 $\mathrm{Q}_1\mathrm{DE}_0[<,+,\times] \subseteq \mathrm{DE}$.

Proof: Let M be a classifier which $\mathrm{Q}_1\mathrm{DE}_0[<,+,\times]$ infers a family $\mathcal{A}$. Classifying any function f the new DE classifier N simulates M but replaces every query of the form $(\exists x_1,\ldots,x_n)[\phi(x_1,\ldots,x_n)]$ by the query $(\exists x_1,\ldots,x_n < m)$ $[\phi(x_1,\ldots,x_n)]$ which can be recursively decided using the initial segment $f(0), f(1), \ldots, f(m\text{–}1)$ of f. Since

$$(\exists x_1,\ldots,x_n)[\phi(x_1,\ldots,x_n)] \Leftrightarrow (\forall^\infty m)(\exists x_1,\ldots,x_n < m)[\phi(x_1,\ldots,x_n)]$$

and since M makes only finitely many queries until M makes its first and only guess $c(f) \in \{\mathrm{YES},\mathrm{NO}\}$, the guess $c_m(f)$ of the emulation equals $c(f)$ for almost all m. ▮

The above theorem can be generalized and strengthened.

Theorem 4.2 $\mathrm{Q}_{a+1}\mathrm{DE}_0[L] \subseteq \mathrm{Q}_a\mathrm{DE}[L]$ *for all* $a \in \mathsf{N}$ *and all languages* L.

Proof: Let $\mathcal{A} \in \mathrm{Q}_{a+1}\mathrm{DE}_0[L]$ via M. It is possible to find out the value of a query $(\exists x_1,\ldots,x_n)[\phi(x_1,\ldots,x_n)]$ with $\phi(x_1,\ldots,x_n)$ having only a alternations of the quantifiers by inserting each n-tuple $(x_1,\ldots,x_n)$ one after an other. As long as none of this n-tuples $(x_1,\ldots,x_n)$ has evaluated $\phi(x_1,\ldots,x_n)$ to YES the classifier assumes that the answer of $(\exists x_1,\ldots,x_n)[\phi(x_1,\ldots,x_n)]$ is NO and emulates M on the queries following this NO. It is easy to see that the algorithm converges to the correct index. ▮

One might look for an inverse of Theorem 4.2, i.e., whether $\mathrm{Q}_a\mathrm{DE}[L] \subseteq \mathrm{Q}_{a+1}\mathrm{DE}_0[H]$ for some sufficiently powerful language H depending on L. This reverse holds for $a \geq 1$:

Theorem 4.3 $\mathrm{Q}_a\mathrm{DE}[all] = \mathrm{Q}_{a+1}\mathrm{DE}_0[all]$ *for all* $a \geq 1$.

Proof: Let $\mathcal{A} \in \mathrm{Q}_a\mathrm{DE}[L]$ via classifier M. W.l.o.g. L contains the ability to unpack tuples, therefore all queries are of the form $(\exists x)[\phi(x)]$ where ϕ contains $a-1$ alternating quantifiers starting with $\forall$. The formula ϕ depends of the previous answers, M received. We make those previous answers a parameter by letting $(\exists x)[\phi(\sigma,k,x)]$ denote the query which would be asked after receiving the answers $\sigma(0),\sigma(1),\ldots,\sigma(k\text{–}1)$, $k \leq |\sigma|$. Using this parametrization one can define a $\mathrm{Q}_{a+2}\mathrm{DE}_0[H]$ classifier N which searches for a number n such that one of the following conditions hold:

(1) $$(\forall m \geq n)(\forall \sigma \in \{\mathrm{NO},\mathrm{YES}\}^m)(\exists k < m)(\forall x)(\exists y)\\ [M(\sigma) = \mathrm{YES} \vee (\sigma(k) = \mathrm{YES} \wedge \neg\phi(\sigma,k,x)) \vee (\sigma(k) = \mathrm{NO} \wedge \phi(\sigma,k,y))\,];$$

(2) $$(\forall m \geq n)(\forall \sigma \in \{\mathrm{NO},\mathrm{YES}\}^m)(\exists k < m)(\forall x)(\exists y)\\ [M(\sigma) = \mathrm{NO} \vee (\sigma(k) = \mathrm{YES} \wedge \neg\phi(\sigma,k,x)) \vee (\sigma(k) = \mathrm{NO} \wedge \phi(\sigma,k,y))\,].$$

Formula (1) says that either $M(\sigma)$ outputs YES or σ is not a string of answers obtained by M's sequent queries to f. Since M converges there must be some n such that either (1) or (2) holds. In the first case N outputs YES and in the second case N outputs NO.

The only problem is that the queries have too many alternations of quantifiers. It is necessary to swap the $(\exists k < m)$ and $(\forall x)$. This can be done by first replacing x by a finite function — coded as a string — with domain $\{0, 1, \ldots, m-1\}$. After swapping the quantifiers x depends from k and is replaced by $\tau(k)$. The new queries (1) and (2) of N are:

$$
\begin{aligned}
(1)\qquad & (\forall m \geq n)(\forall \sigma \in \{\mathrm{NO}, \mathrm{YES}\}^m)(\forall \tau \in \Sigma^m)(\exists k < m)(\exists y) \\
& \quad [M(\sigma) = \mathrm{YES} \vee (\sigma(k) = \mathrm{YES} \wedge \neg\phi(\sigma, k, \tau(k))) \\
& \quad\quad \vee (\sigma(k) = \mathrm{NO} \wedge \phi(\sigma, k, y))\,];
\end{aligned}
$$

$$
\begin{aligned}
(2)\qquad & (\forall m \geq n)(\forall \sigma \in \{\mathrm{NO}, \mathrm{YES}\}^m)(\forall \tau \in \Sigma^m)(\exists k < m)(\exists y) \\
& \quad [M(\sigma) = \mathrm{NO} \vee (\sigma(k) = \mathrm{YES} \wedge \neg\phi(\sigma, k, \tau(k))) \\
& \quad\quad \vee (\sigma(k) = \mathrm{NO} \wedge \phi(\sigma, k, y))\,].
\end{aligned}
$$

Now N witnesses $\mathcal{A} \in \mathrm{Q}_{a+1}\mathrm{DE}_0[all]$. ∎

This proof does not work for $a = 0$ since the first existential quantifier of the formula is used to cover the bounded quantifier $(\exists k < m)$. The next result shows that it is impossible to overcome this gap:

Theorem 4.4 $\mathrm{DE}_1 \not\subseteq \mathrm{Q}_1\mathrm{DE}_0[all]$.

Proof: There are several variants of this proof, they all look for properties which can be discovered by examing larger and larger parts of the graph but not by finitely many $\exists$-queries. One such property is that f has a "loop" [4, 5]. Easier is the following:

$$\mathcal{A} = \{f \mid (\exists x)[x \notin \{f(2^x), f(2^x+1), \ldots, f(2^{x+1}-1)\}\,]\,\}.$$

Such functions are said to have a gap at x. A finite string σ has a gap iff there is an x with $2^{x+1} \leq |\sigma|$ and $x \notin \{\sigma(2^x), \sigma(2^x+1), \ldots, \sigma(2^{x+1}-1)\}$. Otherwise σ is said to have no gap; note that a string σ can be extended to a function without a gap iff σ does not have a gap itself. Since

$$\mathcal{A} = \bigcup\nolimits_{\sigma \text{ has a gap}} \mathcal{F}_\sigma,$$

$\mathcal{A}$ is open and in $\mathrm{DE}_1[all]$. Since the set of all strings σ having a gap is recursive, $\mathcal{A} \in \mathrm{DE}_1$.

So it remains to show that $\mathcal{A} \notin \mathrm{Q}_1\mathrm{DE}_0[all]$. Assume, by way of contradiction, that there exists a classifier M such that $\mathcal{A} \in \mathrm{Q}_1\mathrm{DE}_0[all]$ via M. We assume, without loss of generality, that M asks queries $\phi_0, \phi_1, \phi_2, \ldots$, independent of the answers given. A string σ is said to satisfy a formula ψ iff all $f \succeq \sigma$ satisfy ψ. Now a sequence σ_n of strings is constructed as follows

$$\begin{aligned}
\sigma_0 &= \lambda; \\
b_n &= \begin{cases} 1 & \text{if there is some } \eta_n \succeq \sigma_n \text{ such that} \\ & \quad \eta_n \text{ has no gap and } \eta_n \text{ satisfies } \phi_n; \\ 0 & \text{otherwise, i.e., there is no such } \eta_n; \end{cases} \\
\sigma_{n+1} &= \begin{cases} \eta_n & \text{for the } \eta_n \text{ from above if } b_n = 1; \\ \sigma_n & \text{otherwise, i.e., if } b_n = 0. \end{cases}
\end{aligned}$$

We envision running M and answering ϕ_i by b_i. After finitely many queries, say after the m queries $\phi_0, \phi_1, \ldots, \phi_{m-1}$, M outputs its only guess. Since there is a function extending σ_m which has no gap, this guess is YES.

Let k be greater than the number of occurrences of f in all formulae $\phi_0, \phi_1, ..., \phi_{m-1}$. Furthermore, without loss of generality, we can assume $2^k = |\sigma_m|$. Assume now $f \succeq \sigma_m$ has a gap. Then f satisfies some ϕ_i, $i < m$, which σ_m does not satisfy, since otherwise the classifier would also classify f by YES. So there is a set X of k elements such that g satisfies ϕ_i whenever $(\forall x \in X)[g(x) = f(x)]$. Now let

$$\eta_i(x) = \begin{cases} f(x) & \text{if } x \in X \text{ or } x < |\sigma_m|; \\ y & \text{if } 2^y \leq x < 2^{y+1} \text{ and } x \notin X \\ & \quad \text{and } x \geq |\sigma_m| \text{ and } y \leq \max(X); \\ \uparrow & \text{otherwise } (2^{1+\max(X)} \leq x). \end{cases}$$

Obviously η_i satisfies ϕ_i, η_i extends σ_i and η_i has no gaps since for each $y \geq k$, some value $x \in \{2^y, 2^y+1, \ldots, 2^{y+1}-1\}$ satisfies $\eta_i(x) = y$ because of cardinality-reasons. This contradicts the construction which demands that $b_i = 1$ and that ϕ_i has to be satisfied, e.g. via η_i, whenever this is possible. So $\mathcal{A} \notin \mathrm{Q_1DE_0}[all]$. ∎

The inclusion $\mathrm{Q_1DE_0} \subseteq \mathrm{DE}[all]$ is proper. But it cannot be extended to an inclusion $\mathrm{Q_1DE_0} \subseteq \mathrm{DE}_a[all]$ for any bound a on the number of mindchanges:

Theorem 4.5 $\mathrm{Q_1DE_0}[\emptyset] \not\subseteq \mathrm{DE}_a[all]$ *for all* $a \in \mathsf{N}$.

Proof: Let $\mathcal{A} = \{f \mid \min(f) \text{ is odd}\}$. $\mathcal{A}$ is in $Q_1\mathrm{DE}_0[\emptyset]$ as follows. To classify $\mathcal{A}$ ask the queries $(\exists x)[f(x) = y]$ for $y = 0, 1, 2, \ldots$ until the answer is YES. Let z denote the first y with positive answer. z is obviously the minimum of f and since every function has a minimum, the search terminates. Output YES if z is odd and NO otherwise.

Let M be a classifier which DE classifies $\mathcal{A}$. Let $a \in \mathsf{N}$. We show that there are functions that M takes more than a mindchanges to classify. Consider the following decreasing function f which is defined inductively:

$$f(x) = \begin{cases} a+1 & \text{if } x = 0; \\ f(x-1) & \text{if } x > 0 \text{ and } [f(x-1) = 0 \text{ or} \\ & (M(f(0)\dots f(x-1)) \neq \text{YES} \wedge f(x-1) \text{ is odd) or} \\ & (M(f(0)\dots f(x-1)) \neq \text{NO} \wedge f(x-1) \text{ is even)]}; \\ f(x-1)-1 & \text{otherwise.} \end{cases}$$

In other words f begins constantly with $a+1$ and whenever M classifies that f has an odd minimum, then f takes an new even value below all its previous values, and whenever M classifies f to have an even minimum, then f takes an odd value below all its previous values. This is iterated a+1 times until f reaches the level 0. Therefore M needs $a+1$ mindchanges to classify f. ∎

Kevin Kelly pointed out to us that there is an affirmative answer to the following question, originally posed in [5]: Is $\text{Q}_2\text{DE}_0[\emptyset] \not\subseteq \text{DE}[all]$? (This would extend our previous result $\text{Q}_2\text{DE}_0[\text{Succ}] \not\subseteq \text{DE}[all]$ [5, Theorem 5] where "Succ" denotes the successor-function $\text{Succ}(x) = x+1$.) We present his proof.

Theorem 4.6 $\text{Q}_2\text{DE}_0[\emptyset] \not\subseteq \text{DE}[all]$ *and* $\text{Q}_1\text{DE}_1[\emptyset] \not\subseteq \text{DE}[all]$.

Proof: The class $\mathcal{A} = \{f \mid f \text{ is surjective}\}$ can be identified via a single query: $(\forall y)(\exists x)[f(x) = y]$. Further $\mathcal{A}$ is $\text{Q}_1\text{DE}_1[\emptyset]$ classifiable by first suggesting YES, i.e., that $f \in \mathcal{A}$, and then asking whether $(\exists x)[f(x) = y]$ for all constants $y = 0, 1, 2, \dots$; if once such a query receives an negative answer, the classifier makes a mindchange to NO.

So it remains to show that $\mathcal{A} \notin \text{DE}[A]$ for any oracle A. Assume, by way of contradiction, that M^A classifies $\mathcal{A}$. f is inductively defined by

$$f(x) = \begin{cases} 0 & \text{if } x = 0; \\ f(x-1) & \text{if } x > 0 \text{ and } M^A(f(0)\dots f(x-1)) \neq \text{NO}; \\ f(x-1)+1 & \text{otherwise } (x > 0 \wedge M^A(f(0)\dots f(x-1)) = \text{NO}). \end{cases}$$

So if M^A converges on f to YES then $f(x) = f(x-1)$ for almost all x and f has finite range, i.e., f is not surjective. If M^A converges to NO on f then $f(x) = f(x-1) + 1$ for almost all x; thus f has infinite range and is even surjective. Therefore M^A does not classify f correctly and $\mathcal{A} \notin \text{DE}[all]$. ∎

5 The Query and Borel Hierarchies

DE[all] had a topological characterization. It is possible to extend this characterization to $\text{QDE}_0[all]$ and QDE[all] using the notion of Σ_a Borel sets where a is an ordinal:

Definition 5.1 The Σ_1 Borel sets are just the open sets.
A set is Π_a Borel iff it is the complement of a Σ_a Borel set.
If $a > 1$ is any ordinal, then a set $\mathcal{A}$ is Σ_a Borel iff $\mathcal{A}$ is the union of countably many sets $\mathcal{A}_n$ where each $\mathcal{A}_n$ is a Π_{b_n} Borel set for some $b_n < a$.

The next theorem shows the connections between the query- and Borel hierarchy; in addition it shows that the query hierarchy does not collapse.

Theorem 5.2 *Query and Borel Hierarchy:*
(a) $\mathcal{A} \in \mathrm{Q}_a\mathrm{DE}[all] \Rightarrow \mathcal{A}$ *is* Σ_{a+2} *Borel and* Π_{a+2} *Borel.*
(b) $\mathcal{A} \in \mathrm{QDE}_0[all] \Leftrightarrow \mathcal{A}$ *is* Σ_ω *Borel and* Π_ω *Borel.*
(c) $\mathcal{A} \in \mathrm{QDE}[all] \Leftrightarrow \mathcal{A}$ *is* $\Sigma_{\omega+1}$ *Borel and* $\Pi_{\omega+1}$ *Borel.*
(d) *There is some Borel set* $\mathcal{A} \notin \mathrm{QDE}[all]$.
(e) $\mathrm{Q}_0\mathrm{DE}[all] \subset \mathrm{Q}_1\mathrm{DE}[all] \subset \mathrm{Q}_2\mathrm{DE}[all] \subset \ldots \subset \mathrm{QDE}_0[all] \subset \mathrm{QDE}[all]$,
i.e. no two levels of this hierarchy collapse.

Proof: **(a)**: The first statement is shown via induction on $a \in \mathsf{N}$. Since $\mathrm{Q}_0\mathrm{DE}[all] = DE[all]$, Shai Ben-David's result covers the case $a = 0$. Assume that $\mathcal{A} \in \mathrm{Q}_a\mathrm{DE}[all]$ via M, $0 < a < \mathsf{N}$ and that the statement (a) holds for $a-1$. Now let

$$\mathcal{A}_n = \{f \mid \text{ the } n\text{-th guess of } M \text{ is YES}\}.$$

Since $\mathcal{A}_n \in \mathrm{Q}_a\mathrm{DE}_0[all] \subseteq \mathrm{Q}_{a-1}\mathrm{DE}[all]$ for all n, each set $\mathcal{A}_n$ is Π_{a+1} Borel by induction hypothesis. Now the sets $\mathcal{B}_n = \bigcap_{m \geq n} \mathcal{A}_m$ are Π_{a+1} Borel and therefore their union is Σ_{a+2} Borel. Each $f \in \mathcal{A}$ is in some $\mathcal{B}_n$ since M converges on f to YES at some stage n and then $f \in \mathcal{A}_m$ for all $m \geq n$. On the other hand, if $f \notin \mathcal{A}$, then $f \notin \mathcal{A}_m$ for all arbitrary large m and therefore $f \notin \mathcal{B}_n$ for all n. Thus $\mathcal{A} = \bigcup_n \mathcal{B}_n$ and $\mathcal{A}$ is a Σ_{a+2} Borel set. In the same way it follows that $\overline{\mathcal{A}}$ is a Σ_{a+2} Borel set and $\mathcal{A}$ is also a Π_{a+2} Borel set.

(b): First the direction "$\Rightarrow$" is shown. Now let $\mathcal{A} \in \mathrm{QDE}_0[all]$ via M. On each input f, M asks only the first $n(f)$ questions $\phi_0, \phi_1, \ldots, \phi_{n(f)-1}$ for some $n(f)$ depending on f. We can assume, without loss of generality, that (1) the queries do not depend on f and (2) each query ϕ_i has at most i quantifiers. For each f let

$$\mathcal{A}_f = \{g \mid (\forall i < n(f))[\phi_i(f) \Leftrightarrow \phi_i(g)]\}.$$

Each such set $\mathcal{A}_f$ is uniquely determined by the answers $b_{f,0}, b_{f,1}, \ldots, b_{f,n(f)-1}$ given to the $n(f)$ queries of the classifier, therefore the sets $\mathcal{A}_f$ are indexable via strings in $\{\mathrm{YES}, \mathrm{NO}\}^*$ and there are only countably many different sets $\mathcal{A}_f$. If $f \in \mathcal{A}$ then $\mathcal{A}_f \subseteq \mathcal{A}$; if $f \notin \mathcal{A}$ then $\mathcal{A}_f \subseteq \overline{\mathcal{A}}$. $\mathcal{A}_f$ is a $\Sigma_{n(f)}$ Borel and a $\Pi_{n(f)}$ Borel set. Now

$\mathcal{A} = \bigcup_{f \in \mathcal{A}} \mathcal{A}_f$ and $\overline{\mathcal{A}} = \bigcup_{f \notin \mathcal{A}} \mathcal{A}_f$; both unions are countable and therefore $\mathcal{A}$ is Σ_ω Borel and Π_ω Borel.

The direction "$\Leftarrow$" needs also some claim on the cases $a \in \mathsf{N}$, it would be sufficient to show that every Σ_a Borel set is in $\mathrm{Q}_b\mathrm{DE}[all]$ for some $b \in \mathsf{N}$, but it is even possible to give an upper bound for this b:

Claim: If $a \in \mathsf{N} - \{0\}$ then every Σ_a Borel set is in $\mathrm{Q}_{a+1}\mathrm{DE}_0[all]$.

For $a = 1$ this follows already from Shai Ben-David's result and from $\mathrm{DE}[all] \subseteq \mathrm{Q}_2\mathrm{DE}_0[all]$. We show the claim just at the example of Σ_3 Borel sets, but the proof easily generalizes to all $a \in \mathsf{N}$. Any Σ_3 Borel set $\mathcal{A}$ is of the form

$$\bigcup_i \bigcap_j \bigcup_k \mathcal{F}_{\sigma_{i,j,k}}$$

where the $\mathcal{F}_{\sigma_{i,j,k}}$ are the basic open sets generated by a — not necessarily recursive — family $\sigma_{i,j,k}$ of strings. Now the formula

$$f \in \mathcal{A} \Leftrightarrow (\exists i)(\forall j)(\exists k)(\forall h)[h < |\sigma_{i,j,k}| \Rightarrow f(h) = \sigma_{i,j,k}(h)]$$

witnesses that $\mathcal{A} \in \mathrm{Q}_4\mathrm{DE}_0[all]$. This finishes the proof of the claim.

If $\mathcal{A}$ is Σ_ω Borel and Π_ω Borel, then $\mathcal{A} = \mathcal{A}_0 \cup \mathcal{A}_2 \cup \mathcal{A}_4 \cup \ldots$ and $\overline{\mathcal{A}} = \mathcal{A}_1 \cup \mathcal{A}_3 \cup \mathcal{A}_5 \cup \ldots$ for some Σ_i Borel sets $\mathcal{A}_i$. Now a classifier for $\mathcal{A}$ searches via queries for the first i such that $f \in \mathcal{A}_i$ and then outputs YES if i is even and NO if i is odd.

(c): The proof that all sets in $\mathrm{QDE}[all]$ are $\Sigma_{\omega+1}$ Borel and $\Pi_{\omega+1}$ Borel is similar to that for the induction hypothesis in (a); therefore we leave it to the reader and deal only with the other direction.

The idea to show that every set which is $\Sigma_{\omega+1}$ Borel and $\Pi_{\omega+1}$ Borel is in $\mathrm{QDE}[all]$ is also similar to that of (b), but a bit more complicated: There are Σ_j Borel sets $\mathcal{A}_{i,j}$ such that $\mathcal{A} = \mathcal{A}_0 \cup \mathcal{A}_2 \cup \mathcal{A}_4 \ldots$, $\overline{\mathcal{A}} = \mathcal{A}_1 \cup \mathcal{A}_3 \cup \mathcal{A}_5 \ldots$ and $\mathcal{A}_i = \bigcap_j \mathcal{A}_{i,j}$. The sets $\mathcal{A}_{i,j}$ are in $\mathrm{Q}_{j+1}\mathrm{DE}_0[L_{i,j}]$ via one query $\phi_{i,j}$ in the language $L_{i,j}$; let L denote the union of all these languages $L_{i,j}$. Now the classifier N to witness $\mathcal{A} \in \mathrm{QDE}[L]$ works as follows:

Initialization: Let $i = 0$ and goto Stage 0.

Stage i: Let $j = 0$ and Goto Substage 0.

Stage i, Substage j: If i is even then output YES else output NO.
  Find out if $f \in \mathcal{A}_{i,j}$, i.e., ask whether $\phi_{i,j}(f)$ holds.
  If $f \in \mathcal{A}_{i,j}$ then goto substage $j+1$ of stage i else goto stage $i+1$.

Assume that the algorithm is in stage i. If $f \in \mathcal{A}_i$ then the algorithm remains for ever in stage i since $f \in \mathcal{A}_{i,j}$ for all j. If $f \notin \mathcal{A}_i$ then the algorithm finds some j

with $f \notin \mathcal{A}_{i,j}$ and goes to stage $i+1$ after finite time. Since each f is in some $\mathcal{A}_i$, the algorithm reaches after finite time some stage i which it does not leave again and from now on it outputs the correct guess.

(d): Engelking, Holsztyński and Sikorski [2] showed that there is a $\Sigma_{\omega+2}$ Borel set which is not $\Sigma_{\omega+1}$ Borel. In particular this Borel set is not in QDE[*all*] by part (c).

(e): The inclusions of the hierarchy $\mathrm{Q}_0\mathrm{DE}[all] \subset \mathrm{Q}_1\mathrm{DE}[all] \subset \mathrm{Q}_2\mathrm{DE}[all] \subset \ldots \subset \mathrm{QDE}_0[all] \subset \mathrm{QDE}[all]$ are clear; the only difficult one $\mathrm{Q}_a\mathrm{DE}[all] \subseteq \mathrm{QDE}_0[all]$ follows from $\mathrm{Q}_a\mathrm{DE}[all] \subseteq \mathrm{Q}_{a+1}\mathrm{DE}_0[all]$ for all a with $1 \leq a < \mathsf{N}$.

There is a set $\mathcal{A}$ which is Σ_ω Borel but not Π_ω Borel — otherwise every $\Pi_{\omega+1}$ Borel set would be the countable intersection of Π_ω Borel sets and the hierarchy of Borel sets would collapse in contradiction to the results in [2]. Thus $\mathcal{A}$ is QDE[*all*] classifiable but not $\mathrm{QDE}_0[all]$ classifiable.

Further if $\mathrm{Q}_{a+1}\mathrm{DE}[all] = \mathrm{Q}_a\mathrm{DE}[all]$ for some $a < \mathsf{N}$, then $a > 0$ since $\mathcal{FS} \in \mathrm{Q}_1\mathrm{DE}[all] - \mathrm{Q}_0\mathrm{DE}[all]$. Now $\mathrm{Q}_{a+1}\mathrm{DE}_0[all] = \mathrm{Q}_{a+2}\mathrm{DE}_0[all]$ follows and $\mathrm{Q}_{a+1}\mathrm{DE}[all] = \mathrm{Q}_{a+2}\mathrm{DE}[all]$: Let $\mathcal{A} \in \mathrm{Q}_{a+2}\mathrm{DE}[L]$. The n-th guess of a $\mathrm{Q}_{a+2}\mathrm{DE}[L]$ classifier M can be obtained via a $\mathrm{Q}_{a+2}\mathrm{DE}_0[L']$ classifier for some query language L'. Thus $\mathcal{A}_n = \{f \mid$ the n-th guess of M is YES$\}$ is in $\mathrm{Q}_{a+1}\mathrm{DE}_0[H]$ for some H and $\mathcal{A} \in \mathrm{Q}_{a+1}\mathrm{DE}[H] \subseteq \mathrm{Q}_a\mathrm{DE}[all]$. It follows that the whole hierarchy would collapse to $\mathrm{QDE}_a[all]$. Thus $\mathrm{Q}_a\mathrm{DE}[all] = \mathrm{Q}_{a+3}\mathrm{DE}[all]$. By the claim, every Σ_{a+3} Borel set is $\mathrm{Q}_{a+4}\mathrm{DE}_0[all]$ classifiable and therefore $\mathrm{Q}_{a+3}\mathrm{DE}[all]$ classifiable. But since every $\mathrm{Q}_a\mathrm{DE}[all]$ classifiable set is Σ_{a+2} Borel, the assumption $\mathrm{Q}_{a+1}\mathrm{DE}[all] = \mathrm{Q}_a\mathrm{DE}[all]$ contradicts the fact that there is a Σ_{a+3} Borel set which is not Σ_{a+2} Borel. ▮

6 Fixed Languages and Arbitrary Oracles

The inclusion $\mathrm{Q}_a\mathrm{DE}[all] \subseteq \mathrm{Q}_{a+1}\mathrm{DE}[all]$ from Theorem 4.3 inherently needs to increase the power of the query language. This section now looks at the case where this is prohibited — on the other hand the classifier still should have access to nonrecursive information. Therefore it now may either ask a query in L about the function to be learned or a membership query to an oracle A. $\mathrm{QDE}[L; A]$ denotes this new class and $\mathrm{QDE}[L; all]$ denotes the union of all these classes with fixed language L but A running over all possible oracles. Indeed Theorem 4.3 does not hold in this context:

Theorem 6.1 $\mathrm{DE}_1[all] \not\subseteq \mathrm{QDE}_0[L; all]$ *for every language* L.

Proof: The main idea of this theorem is that given a fixed language L it is not possible to identify each singleton languages $\mathcal{A} = \{g\}$ with finitely many queries: If $\mathcal{A} \in \mathrm{QDE}_0[L; all]$ then there must be finitely many queries $\phi_0, \phi_1, \ldots, \phi_{n-1}$ in the language L to a function f such that $f = g$ iff all these queries receive the answer YES. Since there are only countably many such queries, there are also only countably many finite combination of such queries and there are only countably many singletons $\{g\} \in \mathrm{QDE}_0[L; all]$. In particular there is some singleton $\{g\} \notin \mathrm{QDE}_0[L; all]$.

On the other hand it is possible to $\mathrm{DE}_1[A]$ classify each singleton $\{g\}$ with A being the graph of g: First the classifier conjectures YES, i.e., that $f = g$. If it then discovers at some point a difference between f and g, i.e., if it discovers $(x, f(x)) \notin A$ for some x, then the classifier makes a mind change to NO. ∎

To generalize this theorem, the following result is necessary, which is a nonrecursive variant of the Tree Method [9, Section V.5].

Theorem 6.2 *For every language L there is a tree T such that the following holds:*

- $T : \Sigma^* \to \Sigma^*$ *is total and* $(\forall \sigma, \tau \in \Sigma^*)[T(\sigma) \preceq T(\tau) \Leftrightarrow \sigma \preceq \tau]$.
- $\phi_n(T(f)) \equiv \phi_n(T(g))$ *for the n-th formula $\phi_n \in L$, $\sigma \in \Sigma^n$ and all functions* $f, g \succeq \sigma$.

Proof: This function T can be produced by a transfinite computer program with a variable k ranging over all ordinals below ω^2. Since T is not required to be recursive, termination of the program in the "ordinary" sense is not necessary. Let $\phi_0, \phi_1, \ldots$ be a listing of all formulae in L; they are normed in the way that they all start with an existential quantifies - since it is equivalent if an algorithm asks ϕ_n or $\neg\phi_n$. Further the function variable g is the only free variable of a formula ϕ_n and all bound variables in ϕ_n range over N. Now the algorithm runs as follows:

(1) Initialization: $T(\sigma) = \sigma$ for all $\sigma \in \Sigma^*$.
(2) For all $k < \omega^2$, $k = m\omega + n$ with $m, n \in \mathsf{N}$:
(3) If ϕ_n has m bound variables then do for all $\sigma \in \Sigma^n$:
(4) If $m = 0$ then find $\eta \in \Sigma^*$ with
$(\forall f \succeq \sigma\eta)[\psi(c, T(\sigma\eta 0^\omega)) = \psi(c, T(f))]$
And replace $T(\sigma\tau)$ by $T(\sigma\eta\tau)$ for all $\tau \in \Sigma^*$.
(5) If $m > 0$, $\phi_n(g) = (\exists x)[\psi(x, g)]$ and there are $c \in \mathsf{N}$, $\eta \in \Sigma^*$ with
$(\forall f \succeq \sigma\eta)[\psi(c, T(f))]$
Then replace $T(\sigma\tau)$ by $T(\sigma\eta\tau)$ for all $\tau \in \Sigma^*$.

First one has to verify that each time the search in (4) terminates. But since a statement without quantifiers only looks at each function $T(f)$ at finitely many places, the equivalence always holds if $T(\sigma)$ is sufficiently long, i.e., $\phi_n(g)$ takes the same value for all functions $g \succeq T(\sigma)$ if $|T(\sigma)| > c$ for any numerical constant c appearing in ϕ_n. Further for all $k \geq \omega$, the basis case of the following inductive hypothesis holds:

If $k \geq m\omega$, ϕ_n has m bound variables and $\sigma \in \Sigma^n$ then

$$(*) \quad \phi_n(T(f)) = \phi_n(T(\sigma 0^\omega)) \text{ for all } f \succeq \sigma.$$

It remains to show, that also (5) goes through and preserves the inductive hypothesis. Assume that $m > 0$ and $\phi_n(g) = (\exists x)[\psi(x, g)]$. Either for all $f \succeq \sigma$ and all c, $\psi(c, T(f))$ does not hold. Then there is nothing to do and $(*)$ is satisfied. Or there is some $f \succeq \sigma$ and some c such that $\psi(c, T(f))$ holds. Now $\psi(c, g) \in \{\phi_l(g), \neg\psi_l(g)\}$ for some l and ϕ_l has $m-1$ bound variables. Let $\eta = (f(n), f(n+1), \ldots, f(l))$ if $l > n$ and $\eta = \lambda$ otherwise. By the inductive hypothesis, $\phi_l(T(f')) \equiv \phi_l(T(f))$ for all $f' \succeq \sigma\eta$ and therefore $\phi_n(T(f))$ holds for all $f \succeq \sigma\eta$. After replacing all values $T(\sigma\tau)$ by $T(\sigma\eta\tau)$, $(*)$ is also satisfied for ϕ_n and σ.

Each step only replaces the set $\{T(f) \mid f : \mathsf{N} \to \Sigma\}$ by a proper subset. Therefore if $(*)$ once hold for some n, it will never be destroyed again. Thus at the end, $(*)$ holds for every formula. Further only the stages $k = m\omega + n$ with $m + n \leq |\sigma|$ can change the value of $T(\sigma)$. Therefore $T(\sigma)$ is always defined and converges "transfinitely" to a fixed value. The resulting tree T has the desired properties. ▮

Theorem 6.3 $\mathrm{DE}_{a+1}[all] \not\subseteq \mathrm{QDE}_a[L; all]$ *for every language* L *and* $a \in \mathsf{N}$.

Proof: Let $\Sigma = \{0, 1, \ldots, a+1\}$, T be the mapping from Theorem 6.2 and $\phi_0, \phi_1, \ldots$ the enumeration of all formula in L. Now define the oracle A as follows:

$$\chi_A(\sigma) = \begin{cases} 1 & \text{if } \phi_{|\sigma|}(T(f)) \text{ holds for all } f \succeq \sigma; \\ 0 & \text{if } \phi_{|\sigma|}(T(f)) \text{ holds for no } f \succeq \sigma. \end{cases}$$

By the definition of T always one of these two cases holds, thus χ_A is never undefined. Further let

$$\begin{aligned} \mathcal{A} &= \{T(f) \mid f \in \mathcal{B}\} \quad \text{where} \\ \mathcal{B} &= \{f : \mathsf{N} \to \Sigma \mid \min(f) \text{ is odd}\}. \end{aligned}$$

By Theorem 6.1 $\mathcal{B} \notin \mathrm{DE}_a[all]$. Assume that $\mathcal{A} \in \mathrm{QDE}_a[L; B]$. Now it is easy to translate any $\mathrm{QDE}_a[L; B]$ into a $\mathrm{DE}_a[A \oplus B]$ algorithm for $\mathcal{B}$ by replacing every query ϕ_n to $T(f)$ by the query whether $\sigma = (f(0), f(1), \ldots, f(n-1)) \in A$. Now the classifier classifies $\mathcal{B}$ only by membership-queries to the oracles A and B. Thus if $\mathcal{A} \in \mathrm{QDE}_a[L; B]$ then $\mathcal{B} \in \mathrm{DE}_a[A \oplus B]$ which is known to be wrong. So $\mathcal{A} \notin \mathrm{QDE}_a[L; B]$.

The other direction needs only the oracle T. Let $T^{-1}(\tau)$ be the unique string $\sigma \in \Sigma^*$ such that $T(\sigma) \preceq \tau \prec T(\sigma c)$ for some $c \in \Sigma$ and $T^{-1}(\tau)\uparrow$ iff there is no such σ. Now let

$$N^T(\tau) = \begin{cases} \text{YES} & \text{if } T^{-1}(\tau)\downarrow = \sigma \text{ and } \min(\{\sigma(x) \mid x < |\sigma|\}) \text{ is odd}; \\ \text{NO} & \text{if } T^{-1}(\tau)\downarrow = \sigma \text{ and } \min(\{\sigma(x) \mid x < |\sigma|\}) \text{ is even} \\ & \text{or } T^{-1}(\tau)\uparrow. \end{cases}$$

If $T^{-1}(\tau)\uparrow$ then also $T^{-1}(\eta)\uparrow$ for all $\eta \succeq \tau$. N^T makes at most $a+1$ mindchanges: At every mindchange it computes a new minimum for $T^{-1}(\tau)$ or changes from $T^{-1}(\tau)\downarrow$ to $T^{-1}(\tau c)\uparrow$. The latter causes only a mindchange if $N^T(\tau)$ has been on the value YES and therefore at most a mindchanges have occurred before. Thus $\mathcal{A} \in \mathrm{DE}_{a+1}[all]$. ▮

The proof of Theorem 3.1 in [12] shows, that for the language with the extra symbols "<" and "+" it is possible to choose T and A recursive. This gives the following corollary:

Corollary 6.4 $\mathrm{DE}_{a+1} \not\subseteq \mathrm{QDE}_a[+,<;all]$ *for all* $a \in \mathsf{N}$.

7 Classification with Anomalies

In this section we establish exactly when $DE_a^b \subseteq DE_c^d$. We will show that if $d \neq *$ then $DE_a^b \subseteq DE_c^d$. iff $a \leq c$ and $b \leq d$. The class DE_0^* is surprisingly powerful: we will show $\mathrm{DE}[all] \subseteq \mathrm{DE}_0^*$.

Definition 7.1 Let f and g be functions. If $\#\{x \mid g(x) \neq f(x)\} \leq a$, then we say that g is an *a-variant of* f and denote this by $f =^a g$. If $\{x \mid g(x) \neq f(x)\}$ has finite cardinality then we say g is a *finite variant* or $*$*-variant of* f and denote this by $f =^* g$.

Definition 7.2 Let $\mathcal{A}$ be a set of functions, M a classifier, $a \in \mathsf{N} \cup \{*\}$ and f be any function. We say $\mathcal{A} \in DE^a$ via M iff for every function f the following holds:

- M converges on f to some value $M(f) \in \{\mathrm{NO}, \mathrm{YES}\}$;
- $M(f) = \mathrm{YES} \quad \Rightarrow \quad$ some a-variant of f is in $\mathcal{A}$;
- $M(f) = \mathrm{NO} \quad \Rightarrow \quad$ some a-variant of f is in $\overline{\mathcal{A}}$.

Note that for a function having a-variants in both, $\mathcal{A}$ and $\overline{\mathcal{A}}$, M can converge to YES or to NO as it wants.

A set $\mathcal{A}$ is called *closed under* $=^*$ iff $f \in \mathcal{A} \Leftrightarrow g \in \mathcal{A}$ for all f, g with $f =^* g$. If $\mathcal{A}$ is closed under $=^*$ then $\mathcal{A} \in \mathrm{QDE}^*[all]$ iff $\mathcal{A} \in \mathrm{QDE}[all]$. In particular every set $\mathcal{A} \in \mathrm{QDE}^*[all]$ which is closed under $=^*$ is a Borel set.

Theorem 7.3 *There is a set* $\mathcal{A} \notin \mathrm{QDE}^*[all]$.

Proof: Let $\mathcal{B}$ be not a Borel set. Then also

$$\mathcal{A} = \{f \mid (\exists g \in \mathcal{B})(\forall^\infty \langle x, y\rangle)[f(\langle x, y\rangle) = g(x)]\}$$

is not a Borel set, but closed under $=^*$. Thus $\mathcal{A} \notin \mathrm{QDE}^*[all]$. ▮

Using this result, it is easy to establish the hierarchy:

Theorem 7.4 *There is a set* $\mathcal{A} \in \mathrm{DE}_0^{n+1} - \mathrm{QDE}^n[all]$.

Proof: Let $\mathcal{B} \notin \mathrm{QDE}^*[all]$ be $=^*$-closed and

$$\mathcal{A} = \{f \in \mathcal{B} \mid (\exists x \leq n)[f(x) = 0]\}.$$

$\mathcal{A}$ is QDE_0^{n+1} classifiable via always guessing NO since for each function f the $n+1$-variant f_1 is not in $\mathcal{A}$ where

$$f_b(x) = \begin{cases} b & \text{if } x \leq n; \\ f(x) & \text{otherwise } (x > n); \end{cases}$$

for $b = 0, 1$. On the other hand, assume that M $\mathrm{QDE}^n[L; B]$ classifies $\mathcal{A}$. f_0 is an $n+1$-variant of f. Every n-variant g of f_0 takes a 0 on one of the first $n+1$ places. Therefore $g \in \mathcal{A}$ iff $g \in \mathcal{B}$ iff $f \in \mathcal{B}$. So the relation

$$f \in \mathcal{B} \Leftrightarrow M(f_0) = \mathrm{YES}$$

holds and $\mathcal{B} \in \mathrm{QDE}^*[L; B]$ in contrary to the assumption to $\mathcal{B}$. ∎

On the other hand a direct corollary from Corollary 3.3 is that there is some σ with $\mathcal{F}_\sigma \subseteq \mathcal{A}$ or $\mathcal{F}_\sigma \subseteq \overline{\mathcal{A}}$. Since every function f is a $*$-variant of some function extending σ, either every function f is a $*$-variant of a function inside $\mathcal{A}$, or every function is a $*$-variant of a function outside $\mathcal{A}$. So the following theorem follows:

Theorem 7.5 $\mathrm{DE}[all] \subset \mathrm{DE}_0^*$.

But it is impossible to improve this result to QDE. $\mathcal{FS}$ is closed under $=^*$, therefore a classifier M DE^* classifies $\mathcal{FS}$ iff M DE classifies $\mathcal{FS}$. Since $\mathcal{FS} \notin \mathrm{DE}[all]$, $\mathcal{FS} \notin \mathrm{DE}^*[all]$ and $\mathrm{Q_1DE} \not\subseteq \mathrm{DE}^*[all]$.

Theorem 7.6 $\mathrm{DE}_a^m \subseteq \mathrm{DE}_b^n$ *iff* $a \leq b$ *and* $m \leq n$.

Proof: Let $\mathrm{DE}_a^m \subseteq \mathrm{DE}_b^n$. The condition $m \leq n$ directly follows from Theorem 7.4. Let a be odd and consider the set

$$\mathcal{A} = \{f \mid (\exists c < a)[c \text{ is even and } 4nc \leq \#\{x \mid f(x) > 0\} < 4nc + 4n]\}$$

If $c \leq a$ and $\#\{x \mid f(x) > 0\} = 4nc + n$ then there is some $g =^n f$ with $g \in \mathcal{A}$ iff c is even and there is some $g =^n f$ with $g \notin \mathcal{A}$ iff c is odd.

Assume that the classifier M classifies $\mathcal{A}$ under the requirement DE^n. M first has to guess YES and then M has to make a mindchange each time after reading $4n$ new arguments x with $f(x) > 0$ until $4na$ arguments x with $f(x) > 0$ are found. So a mindchanges are necessary to classify $\mathcal{A}$. It is easy to see that they are also sufficient; even for $m = 0$. ∎

8 Classification with Teams

Team inference was introduced by Smith [11]. Next we define team classifiers similar to team inference defined by Smith.

Definition 8.1 For $m, n \in \mathsf{N}$ such that $1 \leq m \leq n$, $a \in \mathsf{N}$ and for any f, $[m, n]DE_a$ denotes a team of n classifiers out of which at least m of them correctly classify f after at most a mindchanges.

Every set is $[1, 2]\mathrm{DE}_0$ classifiable since one classifier always guesses NO while the other always guesses YES. Thus only $[a, b]\mathrm{DE}_c$ teams with $\frac{a}{b} < \frac{1}{2}$ are interesting. While $[a, b]\mathrm{DE} = \mathrm{DE}$ the connection is more complicated if $c < *$:

Theorem 8.2 *$[b + 1, 2b + 1]DE_c = DE_{(2b+1)c}$ for all $b, c \in \mathsf{N}$.*

Proof: Let $M_1, \ldots, M_{2b+1}$ be a team which $[b+1, 2b+1]DE_c$ classifies $\mathcal{A}$; without loss of generality we can assume that no of them makes more than c mindchanges. Now a single classifier M_0 emulates the team and waits always until at least b members of the team either output YES or output NO. Then M_0 makes its first guess. Now M_0 always outputs the guess of the majority of the team — since each mindchange of the team means that at least one of its members changes the mind from YES to NO or vice versa, M_0 makes at most $(2b + 1)c$ mindchanges.

For the other way around assume let M_0 be a given classifier which makes at most $(2b+1)c$ mindchanges. The team $M_1, \ldots, M_{2b+1}$ waits until M_0 makes its first guess, say YES. Then $M_1, \ldots, M_{b+1}$ guess YES and $M_{b+2}, \ldots, M_{2b+1}$ guess NO. Further two markers ODD and EVEN are placed on $b + 1$ and $2b + 1$. If M_0 makes an odd mindchange from YES to NO, then M_{ODD} makes also a mindchange from YES to NO and the marker moves from position ODD to ODD -1 if ODD > 1 or to $2b + 1$ if ODD $= 1$. Similar if M_0 makes an even mindchange from NO to YES then M_{EVEN} makes a mindchange from NO to YES and the marker moves either to EVEN -1 or to $2b+1$ depending whether the old value of EVEN is greater than or equal to 1. The following example illustrates this for $b = 5$ and $c = 2$:

M_0	ODD	EVEN	M_1	M_2	M_3	M_4	M_5
YES	3	5	YES	YES	YES	NO	NO
NO	2	5	YES	YES	NO	NO	NO
YES	2	4	YES	YES	NO	NO	YES
NO	1	4	YES	NO	NO	NO	YES
YES	1	3	YES	NO	NO	YES	YES
NO	5	3	NO	NO	NO	YES	YES
YES	5	2	NO	NO	YES	YES	YES

NO	4	2	NO	NO	YES	YES	NO
YES	4	1	NO	YES	YES	YES	NO
NO	3	1	NO	YES	YES	NO	NO
YES	3	5	YES	YES	YES	NO	NO

So while M_0 makes 10 mindchanges, each member of the team makes only two but nevertheless the majority of the team always agrees with M_0. ▮

This result can be generalized to $2a - b > 1$. Then the same algorithm works by moving the markers always $2a - b$ positions and making always $2a - b$ classifiers to change their mind. So the outcome is:

Corollary 8.3 $[a, b]\mathrm{DE}_c = \mathrm{DE}_d$ *iff* $\frac{1}{2} < \frac{a}{b} \leq 1$ *and* $d \leq \frac{bc}{2a-b} < d + 1$.

Acknowledgments

We would like to thank Shai Ben-David, Kevin Kelly and Oliver Schulte, for helpful discussions. We would also like to thank two anonymous referees for helpful comments.

References

[1] Shai Ben-David. Can Finite Samples Detect Singularities of Real-Valued Functions? *Proceedings of the 24th Annual ACM Symposium on the Theory of Computer Science, Victoria, B.C.*, 1992, pages 390–399.

[2] R. Engelking, W. Holsztyński and R. Sikorski. Some examples of Borel sets. *Colloquium Mathematicum*, 15:271–274, 1966.

[3] William Gasarch, Efim Kinber, Mark Pleszkoch, Carl Smith, and Thomas Zeugmann. Learning via queries, teams, and anomalies. *Fundamenta Informaticae*, 23(1):67–89, May, 1995. Shorter version in Third Annual Conference on Computational Learning Theory, 1990, pages 327-337, published by Morgan Kaufman.

[4] William Gasarch, Mark Pleszkoch and Robert Solovay. Learning via queries to $[+, <]$. *Journal of Symbolic Logic*, 57(1):53–81, Mar. 1992.

[5] William Gasarch, Mark Pleszkoch and Mahendran Velauthapillai. Classification Using Information. *Proceedings of the Fifth International Workshop on Algorithmic Learning Theory, Reinhardsbrunn*, 1994, pages 290-300, published as Notes in Artificial Intelligence 872 by Springer Verlag.

[6] William Gasarch and Carl Smith. Learning via queries. *Journal of the ACM*, 39(3):649–675, July 1992. A shorter version is in 29th FOCS conference, 1988, pp. 130-137.

[7] Mark Gold. Language identification in the limit. *Information and Control*, 10(10):447–474, 1967.

[8] Kevin Kelly. *The Logic of Reliable Inquiry.* Oxford University Press, Oxford. Forthcoming.

[9] Piergiorgio Odifreddi. *Classical Recursion Theory.* North-Holland, Amsterdam 1989.

[10] Daniel Osherson, Micheal Stob, and Scott Weinstein. Aggregating inductive expertise. *Information and Computation*, 70:69–95, 1986.

[11] Carl Smith. The power of pluralism for automatic program synthesis. *Journal of the ACM*, 29(4):1144–1165, October 1982.

[12] Frank Stephan. Learning via queries and oracles. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 162–169. ACM Press, New York, NY, 1995.

[13] Rolf Wiehagen and Carl Smith. Generalization versus classification. *Journal of Experimental and Theoretical Artificial Intelligence*, 7, 1995. Shorter version in Fifth Annual Conference on Computational Learning Theory, 1992, pages 224-230, Published by Morgan Kaufman.