Enumerating Minimal Hypotheses and Dualizing Monotone Boolean Functions on Lattices

Mikhail A. Babin and Sergei O. Kuznetsov

State University Higher School of Economics,
School for Applied Mathematics and Information Science Pokrovskii bd. 11, 109028
Moscow, Russia
mikleb@yandex.ru, skuznetsov@hse.ru

Abstract. Any monotone Boolean function on a lattice can be described by the set of its minimal 1 values. If a lattice is given as a concept lattice, this set can be represented by the set of minimal hypotheses of a classification context. Enumeration of minimal hypotheses in output polynomial time is shown to be impossible unless P = NP, which shows that dualization of monotone functions on lattices with quasipolynomial delay is hardly possible.

1 Introduction

One of the first models of machine learning that used lattices (closure systems) was the JSM-method[1] of automated hypothesis generation [1,2]. In this model positive hypotheses are sought among intersections of positive example descriptions (object intents), same for negative hypotheses. For classification purposes, it suffices to have hypotheses minimal by inclusion, so-called minimal hypotheses. It is well-known that hypotheses may be generated with polynomial delay [7]. However, the problem of generating minimal hypotheses with polynomial delay remained an open one. In our paper we prove that minimal hypotheses cannot be generated with polynomial delay unless P=NP. This finding has an important implication for the theory of monotone Boolean functions.

The rest of the paper is organized as follows: In the second section we give most important definitions, in the third section we prove the main result about minimal hypotheses, and in the fourth section we discuss the implication of this result for the problem of dualizing monotone Boolean functions.

2 Main Definitions

We use standard definitions from [4]. Let G and M be sets, called the set of objects and attributes, respectively. Let I be a relation $I \subseteq G \times M$ between objects and attributes: for $g \in G, m \in M, gIm$ holds iff the object g has the

[1] Called so in honor of the English philosopher John Stuart Mill, who introduced methods of inductive reasoning in 19th century.

P. Valtchev and R. Jäschke (Eds.): ICFCA 2011, LNAI 6628, pp. 42–48, 2011.

attribute m. The triple $\mathbb{K} = (G, M, I)$ is called a *(formal) context*. If $A \subseteq G, B \subseteq M$ are arbitrary subsets, then the *Galois connection* is given by the following *derivation operators*:

$$A' = \{m \in M \mid gIm \; \forall g \in A\}$$

$$B' = \{g \in G \mid gIm \; \forall m \in B\}$$

The pair (A, B), where $A \subseteq G$, $B \subseteq M$, $A' = B$, and $B' = A$ is called a *(formal) concept (of the context* $\mathbb{K}$*)* with *extent* A and *intent* B (in this case we have also $A'' = A$ and $B'' = B$). The set of attributes B is *implied by the set of attributes* A, or implication $A \to B$ holds, if all objects from G that have all attributes from the set A also have all attributes from the set B, i.e. $A' \subseteq B'$.

Now we present a learning model from [1,2] in terms of FCA [8]. This model complies with the common paradigm of learning from positive and negative examples (see, e.g. [8], [7]): given a positive and negative examples of a "target attribute", construct a generalization of the positive examples that would not cover any negative example.

Assume that w is a target (*functional*) attribute, different from attributes from the set M, which correspond to *structural* attributes of objects. For example, in pharmacological applications the structural attributes can correspond to particular subgraphs of molecular graphs of chemical compounds.

Input data for learning can be represented by sets of positive, negative, and undetermined examples. *Positive examples* (or (+)-examples) are objects that are known to have the attribute w and *negative examples* (or (−)-examples) are objects that are known not have this attribute.

Definition 1. *Consider positive context* $\mathbb{K}_+ = (G_+, M, \mathcal{I}_+)$ *and negative context* $\mathbb{K}_- = (G_-, M, \mathcal{I}_-)$*. The context* $\mathbb{K}_\pm = (G_+ \cup G_-, M \cup \{w\}, \mathcal{I}_+ \cup \mathcal{I}_- \cup G_+ \times \{w\})$ *is called a* learning context*. The derivation operator in this context is denoted by superscript* $\pm$*.*

Definition 2. *The subset* $H \subseteq M$ *is called a positive (or (+)-)-hypothesis of learning context* $\mathbb{K}_\pm$ *if* H *is intent of* $\mathbb{K}_+$ *and* H *is not a subset of any intent of* $\mathbb{K}_-$*.*

In the same way negative (or (-)-) hypotheses are defined.

Besides classified objects (positive and negative examples), one usually has objects for which the value of the target attribute is unknown. These examples are usually called undetermined examples, they can be given by a context $\mathbb{K}_\tau := (G_\tau, M, I_\tau)$, where the corresponding derivation operator is denoted by $(\cdot)^\tau$.

Hypotheses can be used to classify the undetermined examples: If the intent

$$g^\tau := \{m \in M \mid (g, m) \in I_\tau\}$$

of an object $g \in G_\tau$ contains a positive, but no negative hypothesis, then g^τ is *classified positively*. Negative classifications are defined similarly. If g^τ contains hypotheses of both kinds, or if g^τ contains no hypothesis at all, then the classification is contradictory or undetermined, respectively. In this case one can apply probabilistic techniques.

In [6], [7] we argued that one can restrict to *minimal* (w.r.t. inclusion $\subseteq$) hypotheses, positive as well as negative, since an object intent obviously contains a positive hypothesis if and only if it contains a minimal positive hypothesis.

Definition 3. *Let* $G = \{g_1, \ldots, g_n\}$ *and* $M = \{m_1, \ldots, m_n\}$ *be sets with same cardinality. Then the context* $\mathbb{K} = (G, M, \mathcal{I}_{\neq})$ *is called* contranominal scale*, where* $\mathcal{I}_{\neq} = G \times M - \{(g_1, m_1), \ldots, (g_n, m_n)\}$.

The contranominal scale has the following property, which we will use later: for any $H \subseteq M$ one has $H'' = H$ and $H' = \{g_i \mid m_i \notin H, 1 \leq i \leq n\}$.

3 Enumeration of Minimal Hypotheses

Here we discuss algorithmic complexity of enumerating all minimal hypotheses. Note that there is an obvious algorithm for enumerating all hypotheses (not necessary minimal) with polynomial delay [7]. This algorithm is an adaptation of an algorithm for computing the set of all concepts, where the branching condition is changed.

Problem: Minimal hypotheses enumeration (MHE)
INPUT: Positive and negative contexts $\mathbb{K}_+ = (G_+, M, \mathcal{I}_+), \mathbb{K}_- = (G_-, M, \mathcal{I}_-)$
OUTPUT: All minimal hypotheses of $\mathbb{K}_\pm$.

Unfortunately, this problem cannot be solved in output polynomial time unless $P = NP$. In order to prove this result we study complexity of the following decision problem.

Problem: Additional minimal hypothesis (AMH)
INPUT: Positive and negative contexts $\mathbb{K}_+ = (G_+, M, \mathcal{I}_+), \mathbb{K}_- = (G_-, M, \mathcal{I}_-)$ and a set of minimal hypotheses $\mathcal{H} = \{H_1, \ldots, H_k\}$.
QUESTION: Is there an *additional* minimal hypothesis H of $\mathbb{K}_\pm$ i.e. minimal hypothesis H that is $H \notin \mathcal{H}$.

We reduce the most known NP-complete problem satisfiability of CNF to AMH.

Problem: CNF satisfability (SAT)
INPUT: A Boolean CNF formula $f(x_1, \ldots, x_n) = C_1 \wedge \ldots \wedge C_k$
QUESTION: Is f satisfable?

Consider an arbitrary CNF instance $C_1, \ldots, C_k$ with variables $x_1, \ldots, x_n$, where $C_i = (li_1 \vee \ldots \vee l_{ir_i}), 1 \leq i \leq k$ and $l_{ij} \in \{x_1, \ldots, x_n\} \cup \{\neg x_1, \ldots, \neg x_n\}$ $(1 \leq i \leq k$, $1 \leq j \leq r_i)$ are some variables or their negations called literals. From this instance we construct a positive context $\mathbb{K}_+ = (G_+, M, \mathcal{I}_+)$ and a negative context

$\mathbb{K}_- = (G_-, M, \mathcal{I}_-)$. Define

$$M = \{C_1, \ldots, C_k\} \cup \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$$

$$G_+ = \{g_{x_1}, g_{\neg x_1}, \ldots, g_{x_n}, g_{\neg x_n}\} \cup \{g_{C_1}, \ldots, g_{C_k}\}$$
$$G_- = \{g_{l_1}, \ldots, g_{l_n}\}$$

The incidence relation of the positive context is defined by $\mathcal{I}_+ = \mathcal{I}_\mathcal{C} \cup \mathcal{I}_{\neq} \cup \mathcal{I}_=$, where

$$\mathcal{I}_\mathcal{C} = \{(g_{x_i}, C_j) \mid x_i \notin C_j, 1 \leq i \leq n, 1 \leq j \leq k\}$$
$$\cup \{(g_{\neg x_i}, C_j) \mid \neg x_i \notin C_j, 1 \leq i \leq n, 1 \leq j \leq k\}$$

$$\mathcal{I}_{\neq} = \{g_{x_1}, g_{\neg x_1}, \ldots, g_{x_n}, g_{\neg x_n}\} \times \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$$
$$- \{(g_{x_1}, x_1), (g_{\neg x_1}, \neg x_1), \ldots, (g_{x_n}, x_n), (g_{\neg x_n}, \neg x_n)\}$$
$$\mathcal{I}_= = \{(g_{C_1}, C_1), \ldots, (g_{C_k}, C_k)\}$$

that is for i-th clause $C_i^+ \cap \{g_{x_1}, g_{\neg x_1}, \ldots, g_{x_n}, g_{\neg x_n}\}$ is the set of literals not included in C_i, $\mathcal{I}_{\neq}$ is relation of contranominal scale.

The incidence relation of the negative context is given by $\mathcal{I}_- = \mathcal{I}_\mathcal{L}$ where

$$\mathcal{I}_\mathcal{L} = G_- \times \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$$
$$- \{(g_{l_1}, x_1), (g_{l_1}, \neg x_1), \ldots, (g_{l_n}, x_n), (g_{l_n}, \neg x_n)\}.$$

		$C_1\ C_2\ \cdots\ C_k$	$x_1\ \neg x_1\ \cdots\ x_n\ \neg x_n$
$\mathbb{K}_+$	g_{x_1}, $g_{\neg x_1}$, $\vdots$, g_{x_n}, $g_{\neg x_n}$	$\mathcal{I}_\mathcal{C}$	$\mathcal{I}_{\neq}$
	g_{C_1}, $\vdots$, g_{C_k}	$\mathcal{I}_=$	
$\mathbb{K}_-$	g_{l_1}, $\vdots$, g_{l_n}		$\mathcal{I}_\mathcal{L}$

As the set of minimal hypotheses we take $\mathcal{H} = \{\{C_1\}, \{C_2\}, \ldots, \{C_k\}\}$. It is easy to see that $\mathbb{K}_\pm$ with $\mathcal{H}$ is a correct instance of AMH.

If a hypothesis (not necessary minimal) is not included in $\mathcal{H}$ we will call it *additional*.

Proposition 4. *If H is an additional minimal hypothesis of $\mathbb{K}_\pm$ then $H \subseteq \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$.*

Proof. Suppose $H \nsubseteq \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$ then since H is not empty there is some $C_i \in H$, $1 \le i \le k$. But H is a minimal hypothesis and thus it does not contain any hypothesis. Hence $H = C_i$ and this contradicts that H is an *additional* minimal hypothesis. □

For any $H \subseteq \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$ that for an $1 \le i \le n$ satisfies $\{x_i, \neg x_i\} \nsubseteq H$ we define the truth assignment ϕ_H in a natural way:

$$\phi_H(x_i) = \begin{cases} true, & \text{if } x_i \in H; \\ false, & \text{if } x_i \notin H; \end{cases}$$

In the case $\{x_i, \neg x_i\} \subseteq H$ for some $1 \le i \le n$, ϕ_H is not defined.

Symmetrically, for a truth assignment ϕ define the set $H_\phi = \{x_i \mid \phi(x_i) = true\} \cup \{\neg x_i \mid \phi(x_i) = false\}$.

Below, for the sake of convenience, if $H \subseteq \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$ we will denote the complement of H in $\{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$ by $\overline{H}$.

Proposition 5. *If a subset $H \subseteq \{x_1, \neg x_1, \ldots, x_n, \neg x_n\}$ is not contained in the intent of any negative concept (i.e $\forall g \in G_-, H \nsubseteq g^-$), then $\phi_{\overline{H}}$ is correctly defined. Conversely, for a truth assignment ϕ the set $\overline{H_\phi}$ is not contained in the intent of any negative concept.*

Proof. The proof is straightforward. □

The following theorem proves NP-hardness of AMH.

Theorem 6. *AMH has a solution if and only if SAT has a solution.*

Proof.($\Rightarrow$) Let H be an additional minimal hypothesis of $\mathbb{K}_\pm$. First note that by Proposition 4 and Proposition 5 the truth assignment $\phi_{\overline{H}}$ is correctly defined. Since H is a nonempty concept intent of $\mathbb{K}_+$, Proposition 4 together with the fact that $I_{\neq}$ is the relation of contranominal scale implies $H^+ = \{g_{x_i} \mid x_i \in \overline{H}\} \cup \{g_{\neg x_i} \mid \neg x_i \in \overline{H}\}$. Now $H^{++} \cap \{C_1, C_2, \ldots, C_k\} = \emptyset$, hence for any C_i $(1 \le i \le k)$ there is some $g_l \in H^+$ such that $g_l \notin C_i^+$. According to the definition of $\mathcal{I}_\mathcal{C}$ the letter means that literal l belongs to clause C_i. Thus $f(\phi_{\overline{H}}) = true$.

($\Leftarrow$) Let ϕ be a truth assignment and $f(\phi) = true$. Define $H = \overline{H_\phi}$. Note that $H^+ = \{g_{x_i} \mid x_i \in H_\phi\} \cup \{g_{\neg x_i} \mid \neg x_i \in H_\phi\}$, because $\mathcal{I}_{\neq}$ is the relation of contranominal scale and $H \cap g_{C_j}^+ = \emptyset, 1 \le i \le k$. Suppose that $C_i \in H^{++}$ for some $1 \le i \le k$. This is equivalent to $H^+ \subseteq C_i^+$. Hence, by definition of $\mathcal{I}_\mathcal{C}$, there is no literal $l \in H_\phi$ such that $l \in C_i$. Therefore, the clause C_i does not hold and this contradicts that ϕ satisfies CNF f. Thus $H^{++} = H$ and H is a hypothesis. Since H does not contain any $\{C_i\}$, it must contain additional minimal hypothesis. □

Corollary 1. *MHE cannot be solved in output polynomial time, unless $P = NP$.*

Proof. Assume there is an output polynomial algorithm $\mathcal{A}$ that generates all minimal hypotheses in time $pol(|G_+|, |M|, |\mathcal{I}_+|, |G_-|, |\mathcal{I}_-|, N)$, where N is the number of minimal hypotheses. Use this algorithm to construct $\mathcal{A}'$ that makes first $p(|G_+|, |M|, |\mathcal{I}_+|, |G_-|, |\mathcal{I}_-|, k+1)$ steps of $\mathcal{A}$. Clearly, if there is more than k minimal hypotheses, then $\mathcal{A}'$ generates $k+1$ minimal hypotheses, hence we can solve AMH in polynomial time. □

4 Dualizing Monotone Boolean Functions on Lattices

Let $\mathfrak{B}$ be a complete lattice and f be a monotone Boolean function on it. Without loss of generality we can assume that $\mathfrak{B}$ is a concept lattice $\mathfrak{B}(G, M, I)$ from the corresponding formal context $\mathbb{K}(G, M, I)$. Then $A \subseteq B \Rightarrow f((A, A')) \leq f((B, B'))$. It is known that any monotone Boolean function on a lattice is uniquely given by its minimal 1 values, i.e. by the set $\{(A, A') \mid (A, A') \in \mathfrak{B}, f((A, A')) = 1, f((B, B')) = 0 \; \forall B \subset A\}$. We can represent the set of minimal 1 values of a monotone Boolean function as the set of minimal hypotheses of the learning context defined by $\mathbb{K}_+$ and $\mathbb{K}_-$, where $\mathbb{K}_+ = \mathbb{K}$ and object intents of $\mathbb{K}_-$ are precisely maximal 0 values of f. Symmetrically, a learning context $\mathbb{K}_\pm$ specifies a monotone Boolean function f on concept lattice of $\mathbb{K}_+$ such that maximal 0 values of f are (inclusion) maximal object intents of $\mathbb{K}_-$. Consider the following

Problem: Minimal true values enumeration (MTE)
INPUT: A formal context $\mathbb{K}$ and a set of maximal 0 values of monotone Boolean function f on the concept lattice of $\mathbb{K}$
OUTPUT: Set of minimal 1 values of f.

From Corollary 1 it follows that MTE cannot be solved in output polynomial time unless $P = NP$. Note that in the case of Boolean lattice this problem is polynomially equivalent to Monotone Boolean Dualism (see [3]) and the minimal hypotheses in this case can be enumerated with quasi-polynomial delay $O(n^{o(\log n)})$, where n is the input size.

5 Conclusions

The enumeration of minimal hypotheses of a learning context in output polynomial time was shown to be impossible unless $P = NP$. This implies that dualizing monotone Boolean functions on lattices given by their contexts is not possible in quasi-polynomial time unless $P = NP$.

Acknowledgments

The authors were supported by the project of the Russian Foundation for Basic Research, grant no. 08-07-92497-NTsNIL_a.

References

1. Finn, V.K.: On Machine-Oriented Formalization of Plausible Reasoning in the Style of F. Backon–J. S. Mill. Semiotika Informatika 20, 35–101 (1983) (in Russian)
2. Finn, V.K.: Plausible Reasoning in Systems of JSM Type. Itogi Nauki i Tekhniki, Seriya Informatika 15, 54–101 (1991) (in Russian)
3. Fredman, M.L., Khachiyan, L.: On the complexity of dualization of monotone disjunctive normal forms. Journal of Algorithms 21, 618–628 (1996)
4. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, Berlin (1999)
5. Garey, M., Johnson, D.: Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, San Francisco (1979)
6. Ganter, B., Kuznetsov, S.: Formalizing Hypotheses with Concepts. In: Ganter, B., Mineau, G.W. (eds.) ICCS 2000. LNCS (LNAI), vol. 1867, pp. 342–356. Springer, Heidelberg (2000)
7. Kuznetsov, S.O.: Complexity of Learning in Concept Lattices from Positive and Negative Examples. Discrete Applied Mathematics (142), 111–125 (2004)
8. Ganter, B., Kuznetsov, S.O.: Hypotheses and Version Spaces. In: Ganter, B., de Moor, A., Lex, W. (eds.) ICCS 2003. LNCS (LNAI), vol. 2746, pp. 83–95. Springer, Heidelberg (2003)