[14] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, pp. 338–353, 1965.
[15] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems," *IEEE Syst., Man, Cybern*, vol. SMC-3, pp. 404–435, 1990.

Complexity of the Consistency Problem for Certain Post Classes

Ilya Shmulevich, Moncef Gabbouj, and Jaakko Astola

***Abstract*—The complexity of the consistency problem for several important classes of Boolean functions is analyzed. The classes of functions under investigation are those which are closed under function composition or superposition. Several of these so-called Post classes are considered within the context of machine learning with an application to breast cancer diagnosis. The considered Post classes furnish a user-selectable measure of reliability. It is shown that for realistic situations which may arise in practice, the consistency problem for these classes of functions is polynomial-time solvable.**

***Index Terms*—Computational complexity, computational learning theory, consistency problem, monotone Boolean function, Post class.**

I. INTRODUCTION

The consistency problem is an important problem in computational learning theory [2], [12] and can be thought of as a search of a rule from examples. That is, given some sets T and F of "true" and "false" vectors, respectively, we aim to discover a Boolean function f which takes on the value 1 for all vectors in T and the value 0 for all vectors in F. We may also assume that the target function f is chosen from some class of possible target functions. One important reason for studying the complexity of the consistency problem is its relation to the PAC approximate learning model of Valiant [12]. If the consistency problem for a given class is NP-hard, then this class is not PAC-learnable. Moreover, this would also imply that this class cannot be learned with equivalence queries [1].

In practice, it is often reasonable to require f to possess certain properties, known *a priori*, or equivalently, to belong to some specific class of functions. For example, the class of monotone Boolean functions is often representative of real-life phenomena and has been studied in the context of learning in such fields as medical diagnosis [5], manufacturing and reliability analysis [6], as well as signal processing [14]. In [3], the consistency problem was considered for various classes of Boolean functions, such as positive (monotone), k-DNF, h-term-DNF, self-dual, and many others. Specifically, the question of the existence of polynomial-time algorithms for various classes of functions was investigated. In this paper, we consider the consistency problem for several important classes of Boolean functions, known as Post classes.

In 1921, the American mathematician Emil Post described the structure of all classes of Boolean functions that are closed under superposition [10]. The results of Post have been applied outside the field of algebra of logic, for example, in the synthesis [9] and reliability [8] of control systems. For example, the class M of monotone Boolean functions is a Post class, since a superposition of monotone Boolean functions results in a monotone Boolean function. The consistency problem for M was shown to be polynomial-time solvable in [3]. In fact, this was shown to be the case for all transitive classes. In this paper, the classes under consideration are F_8^μ, F_4^μ, F_7^μ, F_3^μ and F_8^∞, F_4^∞, F_7^∞, F_3^∞, using the notation of Yablonsky [13]. This and other notation is established in Section II. The above classes have practical relevance in the context of learning. This is discussed in Section III where an example of breast cancer diagnosis is considered. Finally, Section IV contains the main results concerning the complexity of the consistency problem for these classes of Boolean functions.

Manuscript received October 29, 1999; revised November 22, 2000. This paper was recommended by Associate Editor S. Lakshmivarahan.

The authors are with the Tampere International Center for Signal Processing, Tampere University of Technology, Tampere, Finland (e-mail: ilya@cs.tut.fi).

Publisher Item Identifier S 1083-4419(01)02508-0.

II. DEFINITIONS AND NOTATION

Let E^n denote the n-cube $\{0, 1\}^n$. A Boolean function is a mapping $f\colon E^n \rightarrow E^1$. The elements $v \in E^n$ will be referred to as vectors. The relation $\alpha \preccurlyeq \beta$ holds for two vectors $\alpha = (\alpha_1, \cdots, \alpha_n)$ and $\beta = (\beta_1, \cdots, \beta_n)$ if $\alpha_1 \le \beta_1, \cdots, \alpha_n \le \beta_n$. A Boolean function f is monotone (also called positive) if for any two vectors α and β such that $\alpha \preccurlyeq \beta$, the inequality $f(\alpha) \le f(\beta)$ holds. The notation for various Post classes chosen here coincides with that used by Post as well as that in [13].

We now proceed to define several important Post classes.

Definition 1: A function f satisfies condition $\langle A^\mu \rangle$ (resp. $\langle a^\mu \rangle$), $\mu \ge 2$, if for any collection of μ vectors $v^{(1)}, v^{(2)}, \cdots, v^{(\mu)} \in E^n$ for which $f(v^{(1)}) = \cdots = f(v^{(\mu)}) = 1$ (resp. $f(v^{(1)}) = \cdots = f(v^{(\mu)}) = 0$), there exists an $i \in \{1, \cdots, n\}$ such that $v_i^{(1)} = v_i^{(2)} = \cdots = v_i^{(\mu)} = 1$ (resp. $v_i^{(1)} = v_i^{(2)} = \cdots = v_i^{(\mu)} = 0$).

Let $T(f) = \{v \in E^n : f(v) = 1\}$ be called the on-set of function f. Similarly, let $F(f) = \{v \in E^n : f(v) = 0\}$ be the off-set of f.

Definition 2: A function f satisfies condition $\langle A^\infty \rangle$ (resp. $\langle a^\infty \rangle$) if there exists an $i \in \{1, \cdots, n\}$ such that $v_i = 1$ (resp. $v_i = 0$) for all vectors $v \in T(f)$, (resp. $v \in F(f)$).

It is clear that if a function satisfies $\langle A^{\mu_1} \rangle$, then it also satisfies $\langle A^{\mu_2} \rangle$ for $2 \le \mu_2 \le \mu_1$. Also, the condition $\langle A^\infty \rangle$ implies $\langle A^\mu \rangle$ for any $\mu \ge 2$. Similar statements can be made for $\langle a^\mu \rangle$ and $\langle a^\infty \rangle$. Let F_8^μ and F_4^μ be the sets of all functions satisfying conditions $\langle A^\mu \rangle$ and $\langle a^\mu \rangle$, respectively. Similarly, let F_8^∞ and F_4^∞ be the set of functions satisfying $\langle A^\infty \rangle$ and $\langle a^\infty \rangle$. Finally, let F_7^μ (resp. F_3^μ) and F_7^∞ (resp. F_3^∞) be the classes of monotone functions satisfying conditions $\langle A^\mu \rangle$ (resp. $\langle a^\mu \rangle$) and $\langle A^\infty \rangle$(resp. $\langle a^\infty \rangle$).

Given sets $T \subseteq E^n$ and $F \subseteq E^n$, such that $T \cap F = \emptyset$, the partially defined function $g_{T,F}$ is defined as

$$g_{T,F}(v) = \begin{cases} 1, & v \in T \\ 0, & v \in F \\ *, & \text{otherwise.} \end{cases}$$

A function f is called an *extension* of $g_{T,F}$ if $T \subseteq T(f)$ and $F \subseteq F(f)$ [3]. The consistency problem (also called the extension problem) can be posed as: given some class C of functions and two sets T and F, is there an extension $f \in C$ of $g_{T,F}$? The problem for class C will be denoted as EXTENSION(C) [3] and the question of its polynomial-time solvability is central in the PAC theory of learning.

III. POST CLASSES AND BREAST CANCER DIAGNOSIS

In [7], a breast cancer diagnosis problem is considered within a learning framework. The goal is to find a set of rules, using some specific set of features, to determine whether or not an individual case is highly suspicious for malignancy. Another goal may be to determine whether or not a biopsy or short term followup is necessary. A radiologist studying a breast tumor may use a number of different features or attributes to make the decision regarding malignancy or biopsy. Some such attributes might be: number of calcifications per cm^2, total number of calcifications, variation in the shape of the calcifications, approximate volume of the lesion, and many others [15].

Each such attribute can be encoded by an appropriate number of binary valued variables. For example, suppose the total number of calcifications can be A: <10, B: 10 to 30, or C: >30. Thus, two binary variables would be needed to represent each case, e.g., (0,0) for A, (0,1) for B, and (1,0) for C. Accordingly, a binary vector $x = (x_1, \cdots, x_n)$ of the necessary length to encode all attributes represents any given breast tumor case. The goal then becomes to find a Boolean function f that would classify each example vector x according to the rules that are used by the expert, or radiologist. For example, $f(x) = 1$, if the tumor encoded by x is believed to be malignant, and $f(x) = 0$, otherwise. This can be considered to be a form of knowledge discovery from examples. In [7], it was assumed that the functions underlying the patterns are monotone, or in other words, $f \in M$. One reason for this seemingly restrictive assumption is rooted in the fact that any Boolean function can be expressed in terms of several increasing and decreasing monotone Boolean functions [11].

With the intention of adding a certain degree of reliability, it is natural to expect, or rather, demand, that a given number of positive examples, that is, those which are classified as being malignant, should exhibit some agreement among their attributes. In this sense, reliability does not refer to the process of learning itself, but rather to the constraints imposed on the hypothesis space. The weakest such requirement is that any two malignant cases must necessarily agree on at least one attribute, be it the total number of calcifications, approximate volume of the lesion, or any other. Of course, this attribute can be different depending on the two examples selected. For instance, malignant cases M1 and M2 might agree on attribute A, whereas malignant cases M3 and M4 might agree on attribute B, and so on. One way to strengthen this requirement is to increase the number μ of positive examples all of which must possess one or more common attributes. That is, any μ malignant cases must share at least one common attribute. Such a condition is equivalent to restricting the function f to be in the class F_8^μ or F_7^μ, if f is additionally assumed to be monotone. The strongest, and perhaps unreasonably restrictive, versions of this notion of reliability are captured by the classes F_8^∞ and F_7^∞, which call for *all* malignant cases to agree on at least one attribute.

The restriction to such classes, however, necessitates the use of more bits for the encoding of the cases. That is, one bit is required for every value of an attribute. So, for example, an attribute that can take three values A: <10, B: 10 to 30, or C: >30 should be coded as (1,0,0), (0,1,0), and (0,0,1), respectively. Such an encoding scheme would then allow one to check whether or not two or more malignant cases agree on the value of any given attribute. Having discussed the possible practical aspects of the above mentioned Post classes in the context of learning theory, we now move on to the main results in which we consider the complexity of the consistency problem for these classes.

IV. Extensions in F_8, F_7, F_4, and F_3

We first consider the problems EXTENSION(F_8^μ) and EXTENSION(F_7^μ) for a *fixed* constant μ. The results for F_4^μ and F_3^μ are identical and can be obtained using the duality principle. The question of polynomial-time solvability is posed in terms of n, $|T|$, and $|F|$.

Theorem 3: EXTENSION(F_8^μ) and EXTENSION(F_7^μ) are polynomial-time solvable for a fixed constant μ.

Proof: For the case of F_8^μ, we first must show that checking whether there exists an $i \in \{1, \cdots, n\}$ such that $v_i^{(1)} = v_i^{(2)} = \cdots = v_i^{(\mu)} = 1$ for a collection of μ vectors, takes time $O(\mu \cdot n)$. This is because we require n passes through all μ vectors, once for each coordinate, keeping a flag set if the coordinate is equal to 1. This condition is equivalent to saying that the conjunction of all μ vectors has nonzero Hamming weight.

For a fixed constant μ, checking the above condition for all possible μ-collections takes $O(\mu \cdot n \cdot |T|^\mu)$ time, which is polynomial in $|T|$ and n. If this condition is satisfied for all μ-collections of vectors from T, then as an extension in F_8^μ, we provide

$$f(v) = \begin{cases} 1, & v \in T \\ 0, & \text{otherwise.} \end{cases} \qquad (1)$$

For the case of F_7^μ, the above steps must be preceded by a check for monotonicity. This entails checking whether there is no pair of vectors $v \in T$ and $u \in F$ such that $v \preccurlyeq u$, which takes $O(n \cdot |T| \cdot |F|)$ time. Only if this condition is satisfied can an extension in F_7^μ possibly exist. Thus, EXTENSION(F_8^μ) and EXTENSION(F_7^μ) are both polynomial-time solvable in n, $|T|$, and $|F|$. ∎

We now examine the case in general, when $\mu \leq |T|$ is part of the input to the problems EXTENSION(F_8^μ) and EXTENSION(F_7^μ).

Theorem 4: EXTENSION(F_8^μ) and EXTENSION(F_7^μ) are co-NP-complete.

Proof: The problem EXTENSION(F_8^μ) has a negative answer if there exists an η-collection of vectors $V \subseteq T, (|V| = \eta$ and $\eta \leq \mu)$, such that for all coordinates $i \in \{1, \cdots, n\}$, there is at least one vector $v \in V$ for which $v_i = 0$. Thus, it is easy to see that EXTENSION(F_8^μ) ∈ co-NP, since a nondeterministic algorithm need only guess the correct collection of η vectors $v^{(1)}, \cdots, v^{(\eta)} \in T$ that satisfy the above property, and check this fact in polynomial time as in Theorem 3.

We show that EXTENSION(F_8^μ) is co-NP-hard by a reduction from the well-known MINIMUM-COVER problem [4], which is known to be NP-complete. The instance of the MINIMUM-COVER problem is a collection C of subsets of a finite set S and a positive integer $\mu \leq |C|$. The question is: does C contain a cover for S of size μ or less, that is, a subset $C' \subseteq C$ with $|C'| \leq \mu$ such that every element of S belongs to at least one member of C'?

Consider a finite set $S = \{x_1, \cdots, x_n\}$ and an arbitrary collection of subsets $C \subseteq 2^S$. Let us associate a set T to C by defining

$$T = \{\overline{(v^c)}: c \in C\}$$

where $\overline{v}$ denotes the bit-wise complement of vector v and v^c is the characteristic vector of set $c \in C$, that is, $x_i \in c \iff v_i^c = 1$. This set can be constructed from C in polynomial time in n and $|C|$. It is easy to see that C contains a cover for S if and only if EXTENSION(F_8^μ) has a negative answer.

Turning to F_7^μ, we note that the added constraint of monotonicity does not make the problem any easier. We have seen that the "hardness" of the problem lies entirely within the set T of true vectors, whereas monotonicity is concerned with the relationship between the sets T and F. So, even if we were furnished with a guarantee of monotonicity *a priori*, we would still have to solve the same exact problem as above, namely, to check whether condition $\langle A^\mu \rangle$ is satisfied within set T. Thus, we can conclude that EXTENSION(F_7^μ) is also co-NP-complete. ∎

Now, let us turn to the problem of finding an extension in F_8^∞ and F_7^∞, with F_4^∞ and F_3^∞ again being identical by duality. It is easy to see that checking whether all vectors $v \in T$ have at least one common unit component takes $O(n \cdot |T|)$ time. The question of monotonicity can be answered beforehand in $O(n \cdot |T| \cdot |F|)$ time, as in Theorem 3. In that case, an extension in F_8^∞ or F_7^∞ can be constructed as in (1), in polynomial time, giving us the following result.

Theorem 5: EXTENSION(F_8^∞) and EXTENSION(F_7^∞) are polynomial-time solvable.

V. Conclusion

We have analyzed the complexity of the consistency problem for several important Post classes, which have shown to be relevant in the study of reliability of control systems. Additionally, we discussed a possible application of the considered classes to diagnosis of breast cancer. The positive result obtained here is that if μ is known *a priori* and fixed, which would reflect the case in practice, the consistency problem in classes F_8^μ, F_4^μ, F_7^μ and F_3^μ is solvable in polynomial time. In the context of breast cancer diagnosis, the parameter μ would reflect the user-settable degree of required reliability for the rules being inferred. The higher the value of μ, the more strict this requirement becomes.

Additionally, the consistency problem for classes F_8^∞, F_4^∞, F_7^∞ and F_3^∞ is also polynomial time solvable. As part of future work, it would be worthwhile to consider best-fit extensions in the above classes of Boolean functions, especially for high values of μ, when a fewer number of possible rules is expected.

References

[1] D. Angluin, "Learning regular sets from queries and counterexamples," *Inform. Comput.*, vol. 75, no. 2, pp. 87–106, 1987.
[2] M. Anthony and N. Biggs, *Computational Learning Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[3] E. Boros, T. Ibaraki, and K. Makino, "Error-free and best-fit extensions of partially defined Boolean functions," *Inform. Comput.*, vol. 140, pp. 254–283, 1998.
[4] M. R. Garey and D. S. Johnson, *Computers and Intractability*. San Francisco, CA: Freeman, 1979.
[5] B. Kovalerchuk, E. Triantaphyllou, and J. F. Ruiz, "Monotonicity and logical analysis of data: A mechanism for evaluation of mammographic and clinical data," in *Proc. 13th Symp. Computer Applications in Radiology (SCAR)*, Denver, CO, June 6–9, 1996.
[6] B. Kovalerchuk, E. Vityaev, and E. Triantaphyllou, "How can AI procedures become more effective for manufacturing?," in *Proc. Artificial Intelligence and Manufacturing Research Planning Workshop*. Albuquerque, NM, 1996.
[7] B. Kovalerchuk, E. Triantaphyllou, and A. S. Deshpande, "Interactive learning of monotone Boolean functions," *Inform. Sci.*, vol. 94, no. 1–4, pp. 87–118, 1996.
[8] A. A. Muchnik and S. G. Gindikin, "On the completeness of systems of unreliable elements which realize functions of the algebra of logic" (in Russian), *Doclady AN SSSR*, vol. 144, no. 5, 1962.
[9] E. I. Nechiporuk, "On the complexity of circuits in some bases, containing nontrivial elements with zero weights" (in Russian), *Prob. Cybern.*, no. 8, 1962.
[10] E. Post, "Introduction to a general theory of elementary propositions," *Amer. J. Math.*, vol. 43, 1921.
[11] E. Triantaphyllou and J. Lu, "The knowledge acquisition problem in monotone Boolean systems," in *Encyclopedia of Computer Science and Technology*, A. Kent and J. G. Williams, Eds. New York: Marcel Dekker, 1998, vol. 39, pp. 89–106.
[12] L. G. Valiant, "A theory of the learnable," *Commun. Assoc. Comput. Mach.*, vol. 27, pp. 1134–1142, 1984.
[13] S. V. Yablonsky, G. P. Gavrilov, and V. B. Kudryavtsev, *Functions of the Algebra of Logic and Post Classes* (in Russian). Moscow, Russia: Nauka, 1966.
[14] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo, "Optimal weighted median filtering under structural constraints," *IEEE Trans. Signal Processing*, vol. 43, Mar. 1995.
[15] E. Triantaphyllou. "Discovery of decision rules for breast cancer diagnosis". [Online]. Available: http://www.imse.lsu.edu/vangelis/mining/cancer/cancer1.html-ssi.

An Efficient Blocking–Matching Algorithm Based on Fuzzy Reasoning

Pei-Yin Chen and Jer Min Jou

***Abstract*—Due to the temporal and spatial correlation of image sequence, the motion vector of a reference block is highly related to the motion vectors of its adjacent blocks in the same image frame. By using that idea, we propose a novel efficient fuzzy search (EFS) algorithm for block motion estimation. The experimental results show that the EFS performs better than other fast search algorithms, such as TSS, CS, NTSS, FSS, BBGDS, SES, and PSA in terms of picture quality, accuracy, computational complexity, and coding efficiency.**

***Index Terms*—Block motion estimation, fuzzy reasoning, motion vector.**

I. Introduction

Emerging information technologies such as video conferencing and high definition television (HDTV) require real time compression of video signals, which require a huge amount of bandwidth compared to audio or text information. The key to a successful video coding scheme is to exploit the spatial and temporal redundancies existing in video image sequences effectively. One common feature of most coding schemes is that they use motion estimation to reduce the temporal redundancy, and use block transform coding, such as the DCT or wavelet transform, to reduce spatial redundancy. The most popular technique for motion estimation is the block-matching algorithm (BMA) [1]–[7]. BMA finds the best match for a block in the current frame within a search area in the previous frame. Because of its simplicity and robustness, BMA has recently been adopted by various video coding standards, such as H263 for video conference and Moving Picture Experts Group (MPEG) for video communication. Besides, BMA is also used as a tool for computing the image flow [8], [9].

In the block-matching schemes, the coding performance depends heavily on accuracy, speed, and effectiveness of motion estimation. However, the characteristics of various image sequences bear a lot of uncertainty and are hard to extract. Therefore, it is not easy to devise a good estimation method which always provides accurate and fast motion estimation for different types of image sequences. Consequently fuzzy reasoning is proposed to develop an efficient algorithm for block motion estimation.

During the nearly thirty years since Zadeh first proposed the idea [10], a variety of applications of fuzzy logic [11], [12] have been implemented in various fields. By using the linguistic rules that capture the approximate and qualitative aspects of human knowledge, the fuzzy logic controller can infer desired output actions properly. In this paper, an efficient fuzzy search (EFS) algorithm for block motion estimation is proposed. With the help of fuzzy reasoning, the motion vectors of the adjacent blocks in the current frame are used to find the predicted motion vector of the current block. Using the predicted vector, the EFS can determine the initial center of the search area for the current block. Then, the EFS searches the whole search area with a 3 $\times$ 3 movable

Manuscript received April 1, 2000; revised April 16, 2000 and November 22, 2000. This work was supported in part by the National Science Council, R.O.C., under Grant NSC-89-2213-E-218-025. This paper was recommended by Associate Editor P. Willett.

P.-Y. Chen is with the Department of Electronic Engineering, Southern Taiwan University of Technology, Tainan, Taiwan, R.O.C.

J. M. Jou is with the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C. (e-mail: jou@92a21.ee.ncku.edu.tw).

Publisher Item Identifier S 1083-4419(01)02509-2.