Nonlinear Analysis: Modelling and Control, 2001, v. 6, No. 2, 71-86

Boolean Discriminant Functions in Symbolic Learning with Subclasses

A. J. Raškinis
Vytautas Magnus University
arrask@rc.lrs.lt

Received: 15.10.2001
Accepted: 23.10.2001

Abstract

Finding methods to increase the complexity of the Boolean discriminant functions and to stay within the limits of tractability set by combinatorics is an important task in the field of symbolic machine learning. The original formalism based on meta-features is introduced. Meta-features are predicates that describe relations between the features of the investigated objects and the subclasses (clusters inside classes) of the training set. The formalism facilitates finding Boolean discriminant functions of three variables. These are more complecated than simple conjunctions if the partition of the original training set into subclasses is given.. The structure of meta-feature predicates is close to the structure of statements used by domain experts to describe their knowledge. Consequently, the formalism can be applied in hybrid learning systems, which incorporate information obtained from domain experts.

***Keywords*:** symbolic machine learning, necessary and sufficient discriminant functions, expert knowledge, hierarchical learning

1 Introduction

Symbolic learning methods constitute an important subclass of machine learning methods, and can train functions of propositional logic from training examples. Typical symbolic learner outputs a set of k-DNF formulas, one formula per training class. For any particular class k-DNF formula represents the disjunction of conjuncts, i.e. the disjunction of sufficient conditions for that class [1, 2, 3, 4]. Conjuncts constituting the k-DNF formula are typically found through the search process exploring the space of conjunctions of feature values and testing each "candidate" conjunction against the set of training examples.

Different search strategies and various heuristics are used to guide this search process.

There is one well-known drawback of the abovementioned k-DNF approach. Sufficient conditions appear to be weak "building blocks" of the overall function. There are too many ways to generalize over training examples and consequently too many ways to create a k-DNF formula for any given class. Even if there were attempts to construct k-DNF rules from logical expressions more complex than conjunctions [5], most studies assumed sufficient conditions to be a basic building block in the process of the k-DNF construction. Thus, the reliability of machine-induced k-DNF functions remains questionable.

We proposed a method [6] previously, for building Boolean functions named *stable discriminant functions* (SDF) that represent both sufficient and necessary conditions for every class of the training set and discriminate objects from different classes without classification errors.

This method had three major particularities. First, every training class was approximated by one hyper-rectangle in the feature space that contained all of its examples. Second, the set of conjunctions was extended to contain 52 Boolean functions of greater complexity. Third, the new formalism based on the concept of the first-order meta-features [6] was introduced.

Meta-features are binary predicates, describing relations between features and training classes. The conditions of existence of SDF based on meta-features formalism were formulated and *conditions checking procedure* instead of example testing was elaborated.

The formalism of meta-features was implemented in the algorithms of hierarchical learning [7,8] and its efficiency was experimentally investigated [9]. Investigations compared the average execution time required by the condition checking procedure and by example testing. The experiments demonstrated that the execution time using conditions checking procedure is independent from the number of classes and is linearly related to the number of features. In the case of example-based testing the execution time was linearly related to the number of classes and demonstrated an exponential growth with the number of dataset features.

Approximation of classes by the hyper rectangle is a crude one and sometimes discriminant rules will fail to be discovered even if classes can be perfectly separated if the test is applied on the example-by-example basis .The current investigation is an attempt to propose a solution to this shortcoming. It no longer ignores the information about the layout of training examples within specified class hyper-rectangle. It is assumed that some layout information is known and stated as set of subclasses for each class. Mathematical formalism of meta-features of the second order describing relations between features and subclasses (not classes, as in the case of first order) are introduced in this paper The formalism allows selection of necessary and sufficient (instead of only sufficient) discriminant functions in the task of symbolic learning with subclasses. Instead of conjunctions, 40 Boolean functions of various complexities are allowed in the new formalism

It should be noted, that formalism of meta-features is oriented towards hybrid (in the sense of the method by which information is presented) learning systems, equipped with specific training algorithms processing both the training sample information, and expert knowledge of the subject field. Meta-features form a natural background for expression of expert information about the subject field. Introduction of meta-features of the second order allows perform a natural expert's knowledge acquisition about subclasses, e.g. patients are usually differentiated according to their sex, age, etc. when describing symptoms of some disease.

2 Concept of a meta-feature

Let us take an abstract machine-learning task described in standard form:

$\Sigma = \{\sigma^k\}$, $k = 1, \ldots, K$ *-the set of training classes,*

$\sigma^k = \{s_v^k\}$, $v = 1, \ldots, N_k$ - the set of training examples assigned to the class k,

$X = \{X_l\}$, $l = 1, \ldots, L$ - the set of parameters that describe training examples, where every training example s_v

$s_v^k = (x_{1v}^k, x_{2v}^k, \ldots, x_{Lv}^k)$ - can be represented as a L-dimensional vector in the parameter space X,

$Q = \{Q_r\}$, $r = 1, \ldots, R$ *- the set of logical features.*

Let the parameters X_l take real values. Then the elementary feature Q_r will be defined as a predicate, relating some parameter to some real-valued threshold:

$Q_r = \Pr_{X_l}(s_v^k) \geq \xi_r$ - "projection of an example s_v^k towards the axis X_l is not less than the threshold ξ_r", or "value of the parameter X_l for the object s_v^k is greater or equal to ξ_r", where ξ_r – value of the threshold determined in the training phase.

Stable discriminant function (SDF) $f(Q_1, \ldots, Q_R)$ is a Boolean (logical) function that combines one or more features Q_r. SDF must be stable, i.e. must assign the same value for all the examples of the same class, and it should be discriminant i.e. there should exist at least one class which is differentiated from the others by this value.

First-order *meta-features* – predicates defining relations between features Q_r and training classes σ^k. Two types of meta-features $\{P_r(k)\}$ ir $\{\Pi_r(k)\}$ are analyzed in [6]:.

Meta-feature $P_r(k)$ denotes the stability of the feature Q_r in respect to class k:

$$P_r(k) = \begin{cases} 1, \text{ if } Q_r(s_u^k) = Q_r(s_v^k) & \text{for all } u, v = 1, \ldots, N_k; \\ 0, \text{ if } Q_r(s_u^k) = Q_r(s_v^k) & \text{not for all } u, v = 1, \ldots, N_k; \end{cases} \quad (1)$$

Meta-feature $\Pi_r(k)$ indicates the predominant value of feature Q_r within class k:

$$\Pi_r(k) = \begin{cases} 1, \text{ if } N_r^1(k) > \dfrac{N_k}{2}, \\ 0, \text{ if } N_r^1(k) \le \dfrac{N_k}{2}, \end{cases} \tag{2}$$

where $N_r^1(k)$ is a number of examples s_v^k, for which $Q_r(s_v^k)=1$, $v=1, \ldots, N_k$.

The usage of meta-features for determining the conditions of existence of SDF is illustrated by the following example.

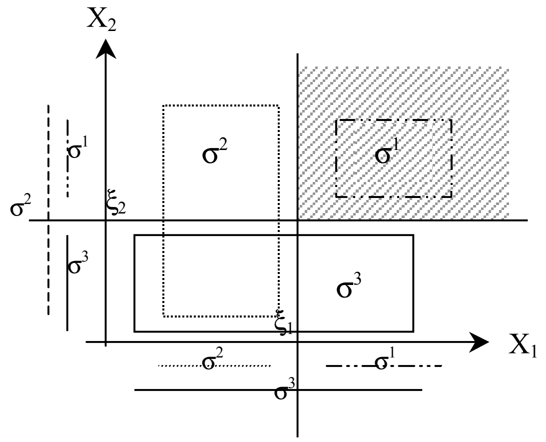

Fig. 1. Separation of classes σ^1 and σ^2 from class σ^3 by the conjunctive SDF $Q_1 \wedge Q_2$

The coordinates X_1 and X_2 in Fig.1 represent two parameters describing classes σ^1, σ^2, σ^3. The bi-dimensional regions of the parameter space occupied by the examples of each class are approximated by rectangles in the coordinate plane. Lines also indicate the one-dimensional projections of these regions below and left to the coordinate axes. Let ξ_1 and ξ_2 be threshold values for parameters X_1 and X_2. Then the feature Q_1 describes the relation $\Pr_{X_l}(s_v^k) \ge \xi_1$, and Q_2 describes the relation $\Pr_{X_l}(s_v^k) \ge \xi_2$. The conjunction $Q_1 \wedge Q_2$ will be true for any training example in the shaded area, and false in the rest of the plane.

The figure above illustrates the fact that two sets of classes $\{\sigma^1\}$ and $\{\sigma^2, \sigma^3\}$ overlapping in both one-dimensional parameter spaces X_1 and X_2 can be separated in the two-dimensional space by the conjunction $Q_1 \wedge Q_2$ if some particular relation is hold. This relation can be expressed by the following statement: "class σ^3 is unstable with respect to threshold ξ_1 but stable with respect to threshold ξ_2 and all of its examples have their X_2 values less than ξ_2; and class σ^2 is unstable with respect to threshold ξ_2, but stable with respect to threshold ξ_1 and all of its examples have their X_1 values less than ξ_1". The generalized proposition for any class k and for any pair of features Q_i, Q_j, can be described using symbols of a meta-features of the first order and define *necessary* condition of existence of the *stable* conjunction $Q_i \wedge Q_j$:

$$\forall k\{\sim P_i(k) \rightarrow [P_j(k) \wedge \sim \Pi_j(k)]\} \wedge \forall k\{\sim P_j(k) \rightarrow [P_i(k) \wedge \sim \Pi_i(k)]\}, \quad (3)$$

$(k = 1, \ldots, K)$,

which can be rearranged into the normal conjunctive form:

$$\forall k[(P_i \vee P_j) \wedge (P_i \vee \sim \Pi_j) \wedge (P_j \vee \sim \Pi_i)], \quad (4)$$

where $k=1, \ldots, K$, and expressions $P_i(k)$ and $\Pi_i(k)$ are shortly denoted as P_i ir Π_i. Existence conditions for other conjunctions are derived using the same methodology.

Expert information can be conveniently described using meta-feature formalism. For example, the proposition: "for all patients, suffering from high blood pressure disease, diastolic blood pressure exceeds 110 mm Hg" will be described as "P_r*(high blood pressure)*$\wedge\Pi_r$*(high blood pressure)*", where P_r(high blood pressure) and Π_r(high blood pressure) are meta-features describing the feature Q_r = "patient *s* diastolic pressure ≥ 110 mm Hg".

3 The second-order meta-features

The second-order meta-features are predicates specifying relations between features Q_r and subclasses of class σ^k.

We will say that class σ^k has a *subclass* t_k, if it has a subset of examples denoted by $\{s_w^k\}$: $t_k=\{s_w^k\}\subset\sigma^k$, where w=1, …, N_t^k, and N_t^k is number of examples in the subclass t_k. We will call the *complement* of the subclass t_k the set of remaining examples of the class σ^k and will denote it $\tau_k = \sigma^k - \{s_w^k\}$. Each class can be partitioned into subclasses in more than one way.

Let us define the intermediate predicates:

$p_s(k,t_k)$ = "Example *s* belongs to the subclass t_k of the class ", or "$s\in t_k$".

$p_s(k,\tau_k)$ = "Example *s* belongs to the subclass τ_k of the class *k*", or "$s\in \tau_k$".

$q_s(i) = q_s(Q_i)$ = "The value of feature Q_i for example *s* is 1"

$q'_s(i) = q_s(\sim Q_i)$ = "The value of feature $\sim Q_i$ for example *s* is 1"

The second-order meta-features will be defined as follows:

1. $L_i(k,t_k) = \forall s\in \sigma^k(p_s(k,t_k)\rightarrow q_s(i))$, or "$Q_i$ is the necessary feature of the subclass t_k of class σ^k ".
2. $Ł_i(k,t_k) = \forall s\in \sigma^k(p_s(k,\tau_k)\rightarrow q_s(i))$, or "$Q_i$ is the necessary feature of the subclass τ_k of class σ^k ".
3. $L'_i(k,t_k) = \forall s\in \sigma^k(p_s(k,t_k)\rightarrow q'_s(i))$, or "$\sim Q_i$ is the necessary feature of the subclass t_k of class σ^k ".
4. $Ł'_i(k,t_k) = \forall s\in \sigma^k(p_s(k,\tau_k)\rightarrow q_s(i))$, or "$\sim Q_i$ is the necessary feature of the subclass τ_k of class σ^k ".

Let us define:

5. $P_i(k,t_k) = L_i(k,t_k) \vee L'_i(k,t_k)$ – "Q_i is stable in respect to subclass t_k",

6. $P_i(k,\tau_k) = Ł_i(k,t_k) \vee Ł'_i(k,t_k)$ – "Q_i is stable in respect to subclass τ_k".

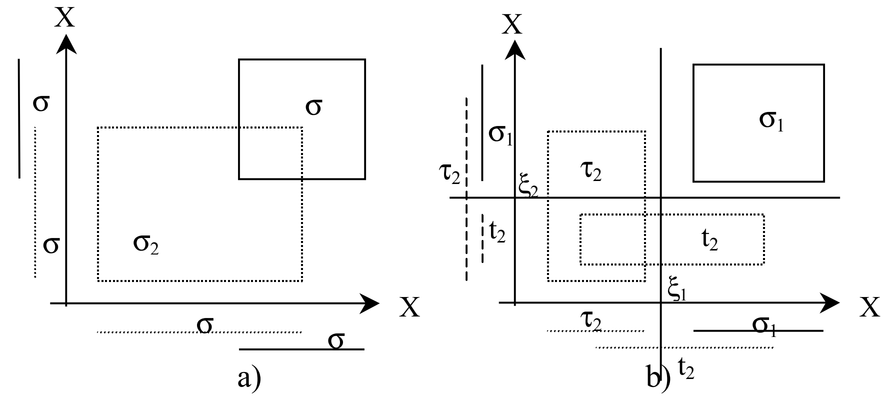

Fig. 2. Separating classes σ^1 and σ^2 by the discriminant conjunction $Q_1 \wedge Q_2$ in case where σ^2 has subclass structure.

We will illustrate the use of the second order meta-features for definition of SDF existence. Let us analyze two classes σ^1 and σ^2 in the parameter space X_1, X_2. Fig. 2 (a) depicts the parameter space regions occupied by these classes approximated by two rectangles. The first order meta-features are helpless as rectangular regions intersect, consequently SDF cannot be found.

Let the class σ^2 have a subclass t_2 and its complement τ_2, as shown in Fig. 2 (b) Then SDF can be obtained by selecting thresholds ξ_1 and ξ_2 for parameters X_1 and X_2 respectively, i.e. by constructing predicates $Q_1 = \Pr_{X_1}(s_v^1) \geq \xi_1$ and $Q_2 = \Pr_{X_2}(s_v^2) \geq \xi_2$. Thus, the SDF can be given by conjunction $Q_1 \wedge Q_2$.

The layout of subclasses must respect some restrictions similar to the restrictions imposed on classes in case of the first-order meta-features. Otherwise SDF will be non-existent. Nevertheless, more logical expressions are needed when dealing with subclasses. The multiplication of logical expressions is due to the fact that both subclasses t_k and its complement τ_k must be taken into account as well as the possibility of t_k and τ_k to be stable or unstable. Every restriction is composed of two parts each being dedicated to one variable and taking into account class meta-description or meta-description of subclasses if they exist.

Restrictions are function-dependent. Sufficient condition for the existence of the stable discriminating conjunction $Q_1 \wedge Q_2$ (Fig.2) using meta-features is given in the following expression (5).

$$\begin{aligned}
&\forall k\ [(\sim P_i(k) \rightarrow P_j(k) \wedge \sim \Pi_j(k)) \vee \\
&\exists t\ \{[\sim P_i(k, t) \rightarrow L_j'(k, t)] \wedge [\sim P_i(k, \tau) \rightarrow Ł_j'(k, t)] \wedge \\
&\quad [\sim L_i(k, t) \rightarrow L_j'(k, t)] \wedge [\sim Ł_i(k, t) \rightarrow Ł_j'(k, t)]\}] \wedge \\
&\forall k\ [(\sim P_j(k) \rightarrow P_i(k) \wedge \sim \Pi_i(k)) \vee \\
&\exists t\ \{[\sim P_j(k, t) \rightarrow L_i'(k, t)] \wedge [\sim P_j(k, \tau) \rightarrow Ł_i'(k, t)] \wedge \\
&\quad [\sim L_j(k, t) \rightarrow L_i'(k, t)] \wedge [\sim Ł_j(k, t) \rightarrow Ł_i'(k, t)]\}], \\
&k = 1, \ldots, K
\end{aligned} \tag{5}$$

Using shorter notation:

$$\begin{array}{llll}
P_i(k) - P_i\ , & P_i(k, t) - P_{it}\ , & P_j(k) - P_j\ , & P_i(k, \tau) - P_{i\tau} \\
\Pi_i(k) - \Pi_i\ , & P_j(k, t) - P_{jt}\ , & \Pi_j(k) - \Pi_j\ , & P_j(k, \tau) - P_{j\tau} \\
L_i(k, t) - L_i\ , & L_j(k, t) - L_j\ , & L_i'(k, t) - L_i', & L_j'(k, t) - L_j' \\
Ł_i(k, t) - Ł_i\ , & Ł_j(k, t) - Ł_j\ , & Ł_i'(k, t) - Ł_i', & Ł_j'(k, t) - Ł_j'
\end{array} \tag{6}$$

Expression (5) in normal conjunctive form will be:

$$\begin{aligned}
&\forall k\ [(P_{it} \vee L_j') \wedge (P_{i\tau} \vee Ł_j') \wedge (P_{jt} \vee L_i') \wedge (P_{j\tau} \vee Ł_i') \wedge \\
&(L_i \vee Ł_i' \vee Ł_j') \wedge (L_i' \vee L_j' \vee Ł_i) \wedge (Ł_i' \vee Ł_j' \vee L_j) \wedge (L_i' \vee L_j' \vee Ł_j)], \\
&k = 1, \ldots, K.
\end{aligned} \tag{7}$$

Similar expressions were derived for the remaining three conjunctions, i.e. for remaining useful Boolean functions of two variables.

4 Analysis of three variable SDF

This paragraph is dedicated to the analysis of existence conditions of SDF of three variables in case where subclass partition of classes is known. We will describe these conditions using the second-order meta-features.

Out of all 256 possible logical functions of three variables only 128 are interesting, because of each function having its symmetrical with respect to negation, which provides the same division of the feature space. Logical constants and functions that can be reduced to functions of two or one variable will not be considered. Functions that express equivalence relations will be considered neither. The remaining functions will be grouped together according to the number and configuration of their disjuncts and each group will be analyzed separately.

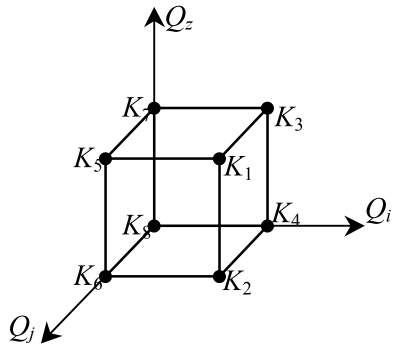

Fig. 3 Eight basic conjunctions in the Boolean feature space Q_i,Q_j,Q_z.

The first group of functions includes basic conjunctions (8) Every such function can be represented by single vertex of the eight-angular Boolean feature space Q_i,Q_j,Q_z (Fig. 3).

$$K_1 = Q_i \wedge Q_j \wedge Q_z \quad , \quad K_5 = \sim Q_i \wedge Q_j \wedge Q_z$$
$$K_2 = Q_i \wedge Q_j \wedge \sim Q_z \quad , \quad K_6 = \sim Q_i \wedge Q_j \wedge \sim Q_z \qquad (8)$$
$$K_3 = Q_i \wedge \sim Q_j \wedge Q_z \quad , \quad K_7 = \sim Q_i \wedge \sim Q_j \wedge Q_z$$
$$K_4 = Q_i \wedge \sim Q_j \wedge \sim Q_z \, , \quad K_8 = \sim Q_i \wedge \sim Q_j \wedge \sim Q_z$$

Let us analyze the first basic conjunction $K_1 = Q_i \wedge Q_j \wedge Q_z$ in greater details. Analogously to the case of two variables, let features Q_i,Q_j,Q_z be predicates relating some parameters to three thresholds ξ_i,ξ_j,ξ_z. and lets use the abbreviated meta-feature notation (6). Then the sufficient condition for existence of stable discriminant conjunction K_1 can be described by (9):

$$
\begin{aligned}
&\forall k\, [(\sim P_i \rightarrow (P_j \wedge \sim\Pi_j \vee P_z \wedge \sim\Pi_z)) \vee \\
&\exists t\, \{[\sim P_{it} \rightarrow (L_j' \vee L_z')] \wedge [\sim P_{i\tau} \rightarrow (Ł_j' \vee Ł_z')] \wedge \\
&\quad [L_i \rightarrow (L_j' \vee L_z')] \wedge [Ł_i \rightarrow (Ł_j' \vee Ł_z')]\}] \wedge \\
&\forall k\, [(\sim P_j \rightarrow (P_i \wedge \sim\Pi_i \vee P_z \wedge \sim\Pi_z)) \vee \\
&\exists t\, \{[\sim P_{jt} \rightarrow (L_i' \vee L_z')] \wedge [\sim P_{j\tau} \rightarrow (Ł_i' \vee Ł_z')] \wedge \\
&\quad [L_j \rightarrow (L_i' \vee L_z')] \wedge [Ł_j \rightarrow (Ł_i' \vee Ł_z')]\}] \wedge \\
&\forall k\, [(\sim P_z \rightarrow (P_i \wedge \sim\Pi_i \vee P_j \wedge \sim\Pi_j)) \vee \\
&\exists t\, \{[\sim P_{zt} \rightarrow (L_i' \vee L_j')] \wedge [\sim P_{z\tau} \rightarrow (Ł_i' \vee Ł_j')] \wedge \\
&\quad [L_z \rightarrow (L_i' \vee L_j')] \wedge [Ł_z \rightarrow (Ł_i' \vee Ł_j')]\}], \\
&k = 1, \ldots, K
\end{aligned}
\tag{9}
$$

Expression (9) includes three parts: one for each feature. The first fragment of each part (the row beginning with quantifier $\forall$) takes into account relations between classes. Other two rows accounts for the possible configurations of subclasses within classes. Rewriting (9) in normal conjunctive form we obtain:

$$
\begin{aligned}
&\forall k\, [(P_{it} \vee L_j' \vee L_z') \wedge (P_{i\tau} \vee Ł_j' \vee Ł_z') \wedge \\
&(P_{jt} \vee L_i' \vee L_z') \wedge (P_{j\tau} \vee Ł_i' \vee Ł_z') \wedge \\
&(P_{zt} \vee L_i' \vee L_j') \wedge (P_{z\tau} \vee Ł_i' \vee Ł_j') \wedge \\
&(L_i' \vee L_j' \vee L_z' \vee Ł_i) \wedge (L_i' \vee L_j' \vee L_z' \vee Ł_j) \wedge (L_i' \vee L_j' \vee L_z' \vee Ł_z) \wedge \\
&(Ł_i' \vee Ł_j' \vee Ł_z' \vee L_i) \wedge (Ł_i' \vee Ł_j' \vee Ł_z' \vee L_j) \wedge (Ł_i' \vee Ł_j' \vee Ł_z' \vee L_z), \\
&k = 1, \ldots, K
\end{aligned}
\tag{10}
$$

Similar expressions were derived for the remaining 7 basic conjunctions.

The second group of functions is composed of 24 functions, which are disjunctions of three adjacent basic conjunctions. Spatial configuration of the function $K_4 \vee K_6 \vee K_8$ belonging to this group is represented in Fig. 4.

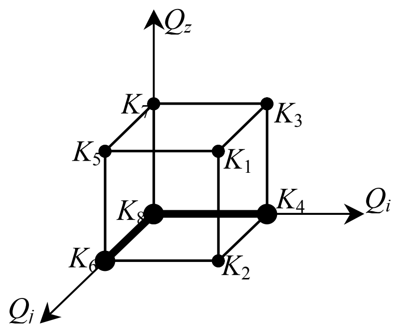

Fig. 4 Logical function of three variables $K_4 \vee K_6 \vee K_8$ in the boolean feature space Q_i, Q_j, Q_z.

Similarly, the second-order meta-features were used to state the necessary conditions of existence for all 24 instances of SDF of this group. These necessary conditions were rewritten to be in the conjunctive normal form. For example, sufficient condition of existence of the SDF $K_4 \vee K_6 \vee K_8$ is given by formula (11):

$$\begin{aligned}&\forall k\,[(P_i \vee P_z) \wedge (P_j \vee P_z) \wedge (P_z \vee \Pi_i) \wedge (P_z \vee \Pi_j) \wedge \\ &(P_i \vee \Pi_z \vee L_i' \vee L_j') \wedge (P_j \vee \Pi_z \vee L_i' \vee L_j') \wedge \\ &(P_i \vee \Pi_z \vee Ł_i' \vee Ł_j') \wedge (P_j \vee \Pi_z \vee Ł_i' \vee Ł_j')], \\ &k = 1, \ldots, K.\end{aligned} \qquad (11)$$

The third group of logical functions has eight instances. These functions represent configuration obtained by disjuncting one central basic conjunction and all its neighbors in the Boolean feature space. Spatial configuration of the function $K_4 \vee K_6 \vee K_7 \vee K_8$ of this type is shown in Fig. 5.

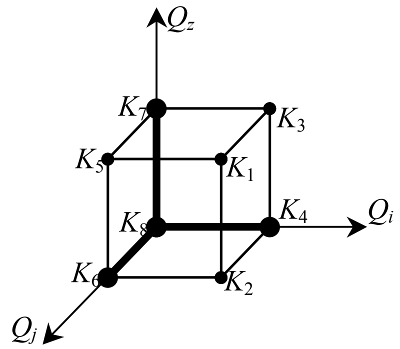

Fig. 5. Boolean function of three variables $K_4 \vee K_6 \vee K_7 \vee K_8$ in the boolean feature space Q_i, Q_j, Q_z.

Half of the eight functions composing this group are symmetrical with respect to negation, consequently sufficient conditions of existence were constructed only for four SDF. These SDF have more complex discriminating surface and permit discrimination of more complex configurations of classes and subclasses.

$$
\begin{aligned}
&\forall k\, [(P_{it} \vee P_{jt}) \wedge (P_{it} \vee P_{zt}) \wedge (P_{jt} \vee P_{zt}) \wedge \\
&(P_{i\tau} \vee P_{j\tau}) \wedge (P_{i\tau} \vee P_{z\tau}) \wedge (P_{j\tau} \vee P_{z\tau}) \wedge \\
&(P_{it} \vee L_j' \vee L_z) \wedge (P_{it} \vee L_j \vee L_z') \wedge \\
&(P_{i\tau} \vee \text{Ł}_j' \vee \text{Ł}_z) \wedge (P_{i\tau} \vee \text{Ł}_j \vee \text{Ł}_z') \wedge \\
&(P_{jt} \vee L_i' \vee L_z) \wedge (P_{jt} \vee L_i \vee L_z') \wedge \\
&(P_{j\tau} \vee \text{Ł}_i' \vee \text{Ł}_z) \wedge (P_{j\tau} \vee \text{Ł}_i \vee \text{Ł}_z') \wedge \\
&(P_{zt} \vee L_i' \vee L_j) \wedge (P_{zt} \vee L_i \vee L_j') \wedge \\
&(P_{z\tau} \vee \text{Ł}_i' \vee \text{Ł}_j) \wedge (P_{z\tau} \vee \text{Ł}_i \vee \text{Ł}_j') \wedge \qquad (12) \\
&(L_i' \vee L_j' \vee \text{Ł}_i \vee \text{Ł}_j) \wedge (L_i' \vee L_j' \vee \text{Ł}_i \vee \text{Ł}_z) \wedge (L_i' \vee L_j' \vee \text{Ł}_j \vee \text{Ł}_z) \wedge \\
&(L_i' \vee L_z' \vee \text{Ł}_i \vee \text{Ł}_j) \wedge (L_i' \vee L_z' \vee \text{Ł}_i \vee \text{Ł}_z) \wedge (L_i' \vee L_z' \vee \text{Ł}_j \vee \text{Ł}_z) \wedge \\
&(L_j' \vee L_z' \vee \text{Ł}_i \vee \text{Ł}_j) \wedge (L_j' \vee L_z' \vee \text{Ł}_i \vee \text{Ł}_z) \wedge (L_j' \vee L_z' \vee \text{Ł}_j \vee \text{Ł}_z) \wedge \\
&(L_i \vee L_j \vee \text{Ł}_i' \vee \text{Ł}_j') \wedge (L_i \vee L_j \vee \text{Ł}_i' \vee \text{Ł}_z') \wedge (L_i \vee L_j \vee \text{Ł}_j' \vee \text{Ł}_z') \wedge \\
&(L_i \vee L_z \vee \text{Ł}_i' \vee \text{Ł}_j') \wedge (L_i \vee L_z \vee \text{Ł}_i' \vee \text{Ł}_z') \wedge (L_i \vee L_z \vee \text{Ł}_j' \vee \text{Ł}_z') \wedge \\
&(L_j \vee L_z \vee \text{Ł}_i' \vee \text{Ł}_j') \wedge (L_j \vee L_z \vee \text{Ł}_i' \vee \text{Ł}_z') \wedge (L_j \vee L_z \vee \text{Ł}_j' \vee \text{Ł}_z')], \\
&k = 1, \ldots, K
\end{aligned}
$$

Consequently the sufficient conditions of existence of SDF of this type are described by rather complex expressions. These expressions were rewritten in conjunctive normal form. One example of such sufficient condition of existence is given by (12) expression for the SDF $K_4 \vee K_6 \vee K_7 \vee K_8$:

5 Conclusions

In this paper mathematical formalism of the second-order meta-features was introduced. The formalism provides basis for constructing necessary and sufficient discriminant functions called stable discriminant functions (SDF) in symbolic recognition tasks. The obtained discriminant functions take into account possible different configurations of classes and subclasses in the parameter space.

Sufficient conditions of existence for 4 SDF of two variables and 36 SDF of three variables were derived using meta-feature formalism.

Sufficient conditions of existence stated in conjunctive normal form appear to have many conjuncts in common. Consecutively, as in case of the first-order meta-features [9], an effective algorithm for simultaneous verification of these conditions could be possibly constructed for the second-order meta-features.

The second-order meta-feature formalism described in this paper is based on 6 meta-features: 4 basic and 2 composite ones. There are more ways to define the set of meta-features that allow the same description of possible configurations of subclasses within classes. Addition of complementary meta-features may simplify logical expressions that describe sufficient conditions of SDF existence. Nevertheless the set of conjuncts taken from all these conditions will grow, the conditions will have less conjuncts in common so simultaneous verification of multiple conditions will become less effective. Thus, the investigation of the basis of meta-feature formalism is an important task.

Human intelligence has the characteristics of using the same mechanisms for the investigation of external world as well as for retrospective analysis of reasoning. Introduction of meta-features can be viewed as an attempt to introduce retrospective properties into artificial intelligence systems.

The sufficient conditions of existence for the remaining SDF of three variables are going to be constructed in future. Extending the set of SDF's to include the equivalence relation is important as this will allow the discrimination of classes that have disjoint, distant and non-overlapping configurations of their subclasses. Investigation of efficient algorithms for SDF construction using meta-feature is foreseen in future.

6 References

1. Ganascia J.-G. "TDIS: an Algebraic Formalization", In: *International Joint Conference on Artificial Intelligence*, p. 1008-1013, 1993.
2. Clark P., Niblett T. "The CN2 Induction Algorithm", *Machine Learning* 3, p. 261-283, 1989.
3. Cohen W.W. "Fast Effective Rule Induction", In: *12th International Conference on Machine Learning*, 1995.
4. Quinlan J.R. "Improved Use of Continuous Attributes in C4.5", *Journal of Artificial Intelligence Research* 4, p. 77-90, 1996
5. Ragavan H., Rendell L. "Lookahead Feature Construction for Learning Hard Concepts", *10th International Conference on Machine Learning*, p. 252-259, 1993.
6. Raškinis A., Stanionis K. "Meta-features as Mean for Constructing Stable Boolean Discriminating Functions in Symbolic Learning", *Lietuvos TSR Mokslų Akademijos darbai. B serija, t. 3(166)*, p. 141-147, 1988. (in Russian)
7. Paulauskas M., Raškinis A., Stanionis K. "Application of Pattern Recognition Techniques to the Problem of Optimal Resource Allocation of Large Energetic System. (1.Formulating Pattern

Recognition Problem)," *Lietuvos TSR Mokslų Akademijos darbai*. B serija, t. 5(150), p. 67-71, 1985. (in Russian)

8. Paulauskas M., Raškinis A., Stanionis K. "Application of Pattern Recognition Techniques to the Problem of Optimal Resource Allocation of Large Energetic System. (Description of the Recognition System)," Lietuvos TSR Mokslų Akademijos darbai. B serija, t. 2(153), p. 98-108, 1986. (in Russian)
9. Stanionis K. "The Meta-feature based Algorithm for Finding Stable Boolean Discriminating Functions", *Lietuvos TSR Mokslų Akademijos darbai*. B serija, t. 6(175), p. 125-131, 1989. (in Russian)