Preface

This book is an attempt to build the theory of classification both as a cognitive process and as its result. Until now, the theory of logical inference did not include classification reasoning as its inalienable component, although precisely the classification reasoning constitutes an integral part of any mode of reasoning. Furthermore, the current models of commonsense reasoning also do not include classification. However, the role of classification in inferences is enormous. Classification as a process of thinking performs the following operations: 1) forming knowledge and data contexts adequate to a current situation of reasoning; 2) reducing the domain of the search for a solution of some problem; 3) generalizing or specifying object descriptions; 4) interpreting logical expressions on a set of all thinkable objects; 5) revealing essential elements of reasoning (objects, attributes, values of attributes, etc); 6) revealing the links of object sets and their descriptions with external contexts interrelated with them. This list can be continued.

In this book, commonsense reasoning is understood as a process of thinking, on the basis of which the causal connections between objects, their properties and classes of objects are revealed. In fact, commonsense reasoning is critical for the formation of conceptual knowledge or ontology in the contemporary terminology.

Studying the processes of classification within the framework of machine learning and knowledge discovery led to the necessity of reformulating the entire class of symbolic machine learning problems as the problems of finding the approximations for the given classifications of objects.

The approach to machine learning problems we propose is based on the concept of a good diagnostic test for the given classification of objects. The problem of inferring good diagnostic tests is formulated as searching for the best approximations of the given classification (partition) on the given set of object examples. The good classification test has a dual nature: on the one hand, it is a logical expression in the form of implication or functional dependency (strict or approximate one), on the other hand, it generates the partition of the training set of objects equivalent to the given classification of this set or the partition that is nearest to the given classification with respect to the inclusion relation.

The concept of good diagnostic test (GDT) has been initially formed within the framework of searching for functional and implicative dependencies in relational databases. But later, the fact has been revealed that the task of inferring all good diagnostic tests for a given object classification is this task that some well-known machine-learning problems can be reduced to: finding keys and functional dependencies in database relations, finding association rules, finding implicative dependencies, inferring logical rules (if-then rules, rough sets, “ripple down” rules), decision tree construction, learning by discovering concept hierarchies, eliminating irrelevant features from the set of exhaustively generated features.

Since the tasks of classification can be reduced to deductive-inductive commonsense reasoning, the same reduction proves to be possible and for the tasks of machine learning.

We consider the theory of algebraic lattices as a mathematical language for constructing algorithms of inferring good classification tests. The advantage of the algebraic lattice approach is based on the
fact that an algebraic lattice can be described both as an algebraic structure that is declarative, and as a system of dual operations with the use of which the elements of this lattice and the links between them can be generated.

The problem of constructing good diagnostic tests is formulated in algebraic terms as constructing the chains of elements of dual lattice or Galois lattice ordered by the inclusion relation.

The analysis of the algorithms of searching for all GDTs for a given object classification in terms of constructing Galois lattice allowed us not only to determine the structure of inferences but also to decompose the algorithms into sub-problems and operations that represent known deductive and inductive modes (modus operandi) of commonsense reasoning. Each step of constructing a classification lattice is interpreted as a mental act. These are mental acts that can be found in any reasoning: stating new propositions, choosing the relevant part of knowledge and/or data for further steps of reasoning, involving a new rule of reasoning (deductive, abductive, inductive, traductive, etc.).

The concept acts as the principal “atom” in commonsense reasoning. This reasoning is based on the knowledge system, with objects, properties (values of attributes) and classifications (attributes) being its elements. If we take into account that implications express relations between concepts (the object ↔ the class, the object ↔ the property, the property ↔ the class), we can assume that schemes of inferring and applying implications (rules of the “if–then” type) form the core of classification processes, which, in turn, form the basis of commonsense reasoning. Deductive steps of commonsense reasoning imply using known facts and statements of the “if–then” type to infer consequences from them. To do it, we apply deductive rules of reasoning, the main forms of which are modus ponens, modus tollens, modus ponendo tollens and modus tollendo ponens. Inductive steps imply applying known facts and existing knowledge to infer new implicative assertions and correct those that turned out to be in contradiction with the existing knowledge. These steps rely on inductive rules of reasoning represented by inductive canons stated by British logician John Stuart Mill. These canons are known as five inductive methods, viz. the Methods of Agreement, the Method of Difference, the Joint Method of Agreement and Difference, the Method of Concomitant Variations and the Method of Residues.

The analysis of inferences for lattice construction allows demonstrating that these inferences engage both inductive and deductive reasoning rules. The implicative dependencies (implications, interdictions, rules of compatibility) generated in a process of good tests construction are used immediately in this process with the aid of deduction for pruning the space of search for new tests.

The theory of classification, stated in the book, proves to be somewhat more general with respect to the formal conceptual analysis (FCA) since this theory allows constructing concepts that are not closed on the lattice operations. In particular, such concepts are good irredundant classification tests.

Thus we come to a new view on modeling commonsense reasoning and machine learning algorithms.

In the book, some methodological approaches to the organization of data and knowledge in intelligent computer systems are discussed. We give some examples of expert system construction based on the integration of data and knowledge via machine learning mechanism. One of the chapters is dedicated to the development of a CASE-technology of automated programming of psycho-diagnostic expert systems. This technology combines two interconnected processes: knowledge specification and knowledge interpretation. A generator of knowledge is capable of using inductive inference for constructing some incompletely specified constituent elements of expert systems and an interpreter uses deductive inference for interpreting knowledge. The ideas placed in the realization of this technology can be developed and personified in a new programming language with the built-in mechanism of commonsense reasoning.

In the book, the basic difficulties of realizing commonsense reasoning in computers are formulated. The solution of these difficulties is the matter of future studies and setting of new problems in the
computer sciences. To similar problems the author relates the integrating of conceptual clustering and supervised machine learning in such a way that an intelligent system would be capable of finding the best interpretations of the results of clustering by using its internal knowledge rather than some external instructions of the supervisor.

To the problems requiring future studies, we also relate the representation of data and knowledge as algebraic lattice structures. Some steps in this direction have been already done within the framework of OLAP and OLAM technologies.

The crucial problem of commonsense reasoning implementation in computers is also connected with the need for the pre-processes of analyzing and synthesizing objects of different nature—image, words and proposals of natural languages, speech, music and other objects of the external world.

It is the author’s hope that this book will be interesting for the specialists of different fields, first of all for the specialists in artificial intelligence. The problems touched upon in the book will draw the attention of the developers of machine learning algorithms, knowledge engineers, programmers who create intelligent computer systems, and also psychologists and philosophers who are interested in questions of the psychology of thinking.

The book can draw the attention of logicians and mathematicians who will develop the theory of classification at the higher professional level and advance new models of logical inference, which will include, finally, the theory of classification expanding logical inference by commonsense reasoning. Moreover, the book can contribute to stimulate new ideas, new collaborations and new research activity in this research area.

The arrangement of the chapters follows a natural exposition of the main subjects in classification theory and practice.

**Chapter 1** gives a view of historical development of the concepts of knowledge and human reasoning both in mathematics and psychology. Mathematicians create the formal theory of correct inferences; psychologists study cognitive mechanisms that underpin knowledge construction and thinking as the most important functions of human existence. They study how the human mind works. The progress in understanding human knowledge and thinking will be undoubtedly related to combining the efforts of scientists in these different disciplines. Considering the problems of knowledge and human reasoning unable to be treated separately, we strive to cover in this chapter the key ideas of knowledge and logical inference that have been manifested in the works of outstanding thinkers and scientists of past time. These ideas reveal all the difficulties and obstacles on the way to comprehending the human mental processes.

In **Chapter 2**, we focus on the tasks of knowledge engineering related mainly to knowledge acquisition and modeling integrated logic-based inference. We have reviewed the principal and more important directions of research that pave the ways to understanding and modeling human plausible (commonsense) reasoning in computers.

**Chapter 3** develops a conception of commonsense reasoning based on mutually coordinated classification operations on objects, classes of objects, and properties of objects. This conception goes back to the model of classification processes given by J. Piaget, the outstanding psychologist of the 20th century.

The operations of classification are an integral part of any reasoning about time, space, things, events, motions and so on. They consolidate all the forms of reasoning and make it possible to present knowledge as the system of interconnected relations. We analyze the logical semantics of classification connections between objects, classes of objects, properties of objects and the role of commonsense reasoning operations in creating classification structures.

In **Chapter 4**, we describe a model of commonsense reasoning that has been acquired from our numerous investigations on the human reasoning modes used by experts for solving diagnostic problems
in diverse areas such as pattern recognition of natural objects (rocks, ore deposits, types of trees, types of clouds etc.), analysis of multi-spectral information, image processing, interpretation of psychological testing data, medicine diagnosis and so on. The principal aspects of this model coincide with the rule-based inference mechanism that has been embodied in many expert systems.

We can assume that schemes of inferring and applying implications (rules of the “if–then” type) form the core of classification processes, which, in turn, form the basis of commonsense (plausible) reasoning. Deductive steps of commonsense reasoning imply using known facts and statements of the “if–then” type to infer consequences from them. To do it, we apply deductive rules of inference, the main forms of which are modus ponens, modus tollens, modus ponendo tollens and modus tollendo ponens. Inductive steps imply applying known facts and existing knowledge to infer new implicational statements and correct those that turned out to be in contradiction with the existing knowledge. These steps rely on inductive rules of reasoning represented by inductive canons known as five inductive methods formulated by English logician Mill, viz. the Methods of Agreement, the Method of Difference, the Joint Method of Agreement and Difference, the Method of Concomitant Variations and the Method of Residues.

Chapter 5 contains some examples of natural human commonsense reasoning related to both scientific pattern recognition problems and logical games. An analysis of inference structure shows that inductive and deductive rules communicate in reasoning. An automated model for detecting the types of woodland from the incomplete descriptions of some evidences is also given in this chapter. An interesting part of this model is a small knowledge base containing the representation of experts’ knowledge about natural woodlands as biological formations.

Chapters 6 and 7 give the main ideas of the classification theory and its connections with machine learning problems.

Chapter 6 discusses a revised definition of a classification (diagnostic) test advanced earlier within the framework of information theory. This revised definition allows considering the problem of inferring classification tests as a problem of searching for the best approximations of a given classification on a given set of training data.

In this quality, the concept of a classification (diagnostic) test became a key concept of machine learning problems dealing with inferring conceptual knowledge of the following types: conceptual descriptions of object classifications and logical links (expressed in the form of logical rules) between these classifications.

We demonstrate that a class of well-known machine learning problems related to inferring logical dependencies (implicative, functional, and associative), decision trees from examples, discovering concept hierarchies (ontology), eliminating irrelevant features from the set of exhaustively generated features and some others can be reduced to the task of inferring diagnostic (classification) tests. In fact, the diagnostic task covers all the tasks of symbolic supervised machine learning.

An algebraic model of diagnostic task (algebra of classifications) is brought forward on the foundation of the partition lattice in which object, class, attribute, and value of attributes take their interpretations. This model possesses both declarative and procedural properties.

In Chapter 7, the definition of good diagnostic test and the characterization of good tests are introduced and the concepts of good maximally redundant test (GMRT) and good irredundant test (GIRT) are given.

The definition of good test is based on the partition model of classifications that has been given in the previous chapter. Some characteristics of good tests allow choosing a strategy for inferring all kinds of good diagnostic tests. We describe an algorithm called Background Algorithm based on the method of mathematical induction. This algorithm is applicable to inferring all kinds of good classification tests and, consequently, for inferring functional, implicative dependencies and association rules from a given
data set. We discuss also, in this chapter, the possible ways of constructing an efficient algorithm for inferring good tests of any kind.

The concept of good classification test is redefined in Chapter 8 as an element of the dual lattice or Galois lattice. Inferring the chains of Galois lattice elements ordered by the inclusion relation lies in the foundation of generating all types of diagnostic tests. The concept of an inductive transition from one element of a chain to its nearest element in the lattice is determined. The following special rules are introduced for realizing the inductive transitions: generalization rule, specification rule, dual generalization rule, and dual specification rule.

Note that reasoning begins with using a mechanism for restricting the space of searching for tests: for each collection of attributes’ values and objects, to avoid constructing all its subsets. For this goal, the concepts of admissible and essential values (objects) are introduced. Searching for admissible or essential values (objects) is based on inductive diagnostic rules. During the lattice construction, the implicative assertions are generated and used immediately. It means that the knowledge acquired during the process of generalization (specification) is used for pruning the search in the search space. All the operations of lattice generation take their interpretations in human mental acts.

In this chapter, we propose a non-incremental learning algorithm NIAGaRa based on a reasoning process realizing one of the ways of Galois lattice generation. Next we discuss the relations between the good test construction and the Formal Concept Analysis (FCA).

The methodology presented in this chapter provides a framework for solving diverse and very important problems of constructing machine learning algorithms based on a unified logical model in which it is possible to interpret any elementary step of logical inferring as a human mental operation. This methodology is more general than the FCA because it deals with some objects that are not formal concepts in terms of the FCA (for example, good irredundant tests, contained in a good maximally redundant test).

The important steps in the direction to an integrative model of deductive-inductive commonsense reasoning are made in Chapter 9. The incremental approach to developing machine learning algorithms is one of the most promising directions in creating intelligent computer systems. The decomposition of inferring good classification tests is advanced into two kinds of subtasks that are in accordance with human mental acts.

This decomposition involves searching for essential values and objects, eliminating values, cutting off objects, choosing values or objects for subtasks, extending or narrowing collections of values, extending or narrowing collections of objects, using forbidden rules, forming subtasks and some others actions. This decomposition allows, in principle, to transform the process of inferring good tests into a “step by step” reasoning process.

We give two basic recursive procedures based on two kinds of subtasks for inferring all good maximally redundant classification tests (GMRTs): ASTRA and DIAGaRa. An incremental algorithm INGOMAR for inferring all GMRTs is presented too. The problems of creating an integrative inductive-deductive model of commonsense reasoning are discussed in the last section of this chapter.

Chapter 10 presents some fast heuristics for inferring approximately minimal diagnostic tests based on which a model of commonsense reasoning by analogy is constructed. This model has been implemented in the system called DEFINE. The results of this system’s application for recognizing the type of tree species with the use of aerial photographs is described.

Next, we discussed the different approach to mining approximate rules. Mining approximate functional, implicative dependencies and approximate association rules is based on the same criteria and on the application of one and the same algorithm of machine learning realized in the Diagnostic Test Machine described shortly in this chapter. The Diagnostic Test Machine is destined for inferring the
broad class of logical dependencies from raw data: functional dependencies (strict and fuzzy), implicative dependencies (strict and fuzzy), decision trees based on obtained dependencies, logical rules based on obtained dependencies (implicative ones), association rules.

A technology for rapid prototyping and developing expert systems or intelligent systems as a whole is proposed in Chapter 11. The main parts of the technology are the object-oriented model for data and knowledge representation and the mechanism for data-knowledge transformation on the basis of an effective algorithm of inferring all good classification tests. An approach to expert system development by means of the technology proposed is analyzed. The toolkits for expert system generation are described and the application of these tools is demonstrated for the development of a small geological expert system.

Chapter 12 proposes an automated technology for creating the applied psycho-diagnostic expert systems (APDS) the main peculiarity of which consists in using machine learning to choose, validate, define and redefine the main constituent elements of testing and decision making procedures utilized in the developed psycho-diagnostic systems.

Machine learning and knowledge acquisition from experts have distinct capabilities that appear to complement one another. The integration of these approaches to knowledge discovering can both improve the accuracy of system knowledge bases and reduce the time of APDS development. The expert systems created by means of the integrated approach possess higher accuracy than those created only by knowledge elicitation from experts without using machine learning methods. We describe a software, called GENINTES (GENERATOR + INTERPRETER of EXPERT SYSTEMS) realizing the integrated CASE – technology for expert system rapidly prototyping, creating, and evolution.

We consider both statistical and logical (symbolic) methods of machine learning, so our approach encompasses the automated knowledge acquisition methods for a wide range of knowledge types.

The CASE-technology is based on two models of knowledge: the model of knowledge embodied in APDS (M1) and the model of knowledge of the designing of APDS (M2). Both models are object-oriented. M1 is a meta-structure describing the concepts, connections between them and processes of knowledge and data transformation used in any APDS. M1 includes the following main concepts: psychological characteristic (Ch) (measured and inferred), computational scheme (CS), computational expression (CE), operation (Op), scale (Sc), logical rule (LR), inference model (IM), and conclusion or diagnosis (C-D).

The process of constructing APDS is a sequence of transforming incomplete specification of a projected PDS from one state to another until its final state will be obtained under which the union of the specification with the program of an INTERPRETER gives the ready APDS.

Models M1 and M2 are object-oriented and operational: the relationships between concepts can be viewed as both the schemes of inference and the procedures of concepts transformation when the operations are added to the schemes of relationships.

The process of APDS specification is analogous to the process of specifying computations with the use of any programming language. The peculiarity of our programming language consists in the fact that the forward deductive and inductive engines are incorporated in it.

Chapter 13 deals with the description of possible mechanisms for data-knowledge organization and management in intelligent computer systems. Challenges and future trends are discussed in the last section of this chapter, followed by the concluding remarks.

The term “intelligent system” means that the system is capable of performing commonsense reasoning operations at any stage of its functioning (i.e., under performing query answering and updating data and knowledge). More exactly, the functioning of intelligent system as a whole is a commonsense reasoning process. The incorporation of commonsense reasoning mechanisms into data-knowledge
processing is becoming an urgent task in intelligent computer system and conceptual data-knowledge
design. There is not a methodology supporting the solution of this task. We tried briefly to describe the
problems awaiting a solution. We began with the consideration that data-knowledge communication at
all stages of functioning an intelligent computer system is central for realizing commonsense reasoning.
We believe that a new methodology for data-knowledge communication has to be created by the use of
a systemic approach to modeling the cognitive activity.

All the above chapters contain a number of examples explaining the ideas of the author and the
proposed algorithms.