A Concept of Self-Supervised Logical Rule Inference in Symbolic Classifications

Xenia Naidenova[1[0000-0003-2377-7093]] and Vladimir Parkhomenko[2[0000-0001-7757-377X]]

[1] Military Medical Academy, Saint Petersburg, Russian Federation
E-mail: ksennaidd@gmail.com
[2] Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russian Federation
E-mail: parhomenko.v@gmail.com

Abstract. An approach to modelling self-supervised learning for automated inferring good classification tests is proposed. The concepts of internal and external learning contexts are formulated. A model of intelligent agent, capable of improving own learning process of inferring good classification tests in the external context is advanced. Internal evaluation is used in an internal process of learning with the aim of tuning the external learning process. The same learning algorithm is used for supervised learning both in the external context and in the internal context. The structure of good test inferring is described and a procedure to recognize the end of inferring process is proposed.

Keywords: Self-supervised learning, Good classification tests, Internal context, External context, Intelligent agent, Deep learning.

1 Introduction

Self-learning embodies one of the essential properties of human intelligence related to an internal evaluation of the mental process quality. A deeper level of learning – self-learning – allows to manage the learning process in an external context in terms of its effectiveness through the internal evaluation and developing rules to select the best learning strategies and parameters without a teacher.

We shall understand self-learning as a process of improvement of an agent's (or system's) actions on the basis of self-evaluation of his (its) actions in a variable context. When the agent selects sub-contexts and some actions in learning process, he (it) uses some criteria. The self-learning is related to the ability to change these criteria, to form new criteria, which is essentially to improve the learning algorithms, making them more consistent with the external context and more effective.

The purpose of this paper is to model a self-learning process in the logical or symbolic supervised algorithms of machine learning. This mode of learning covers mining logical rules and dependencies from data: "if-then" rules, decision trees, functional, implicative and associative dependencies. We shall consider a special kind of symbolic machine learning, namely, inferring good tests from data [1] in multi-valued dynamic contexts (external contexts) for recognizing classes of objects represented by their symbolic descriptions. The self-learning at the internal (deep) level implements the analysis and internal evaluation of classification rule inferring in the external context and allows one to reveal the relationships between the external contexts (subcontexts) and the parameters of learning. The implementation of self-learning in the internal context can be based on the same algorithm of symbolic machine learning that works in the external context.

The paper is organized as follows. The related works are discussed in Section 2. Sections 3 and 4 deal with defining a software agent capable of self-learning and the structure of the internal context. Sections 5, 6, and 7 cover the description of self-learning in inferring good maximally redundant classification tests from data. To complete the paper, we give a short conclusion.

2 Related works

The analysis of modern researches has been implemented in the following directions: modeling of self-learning (self-supervised learning), deep learning [2] and models of learning in robots and robotic systems.

In the first direction, it is particularly interesting [3] the principles and technologies of creating a robot that can move in the environment, manipulate objects and avoid obstacles. The robot is designed as an autonomous system. It requires from the robot a good spatial and semantic understanding of the environment. The self-learning robot should be aware of its own localization and realize an internal reflection of spatial situation taking into account different scenes (semantic understanding) in order to recognize new objects. It is declared by the author that the robot should be self-esteemed and self-managed on the basis of previous experience. It must constantly adapt its spatial and semantic models in order to improve the performance of its tasks. Some concepts and algorithms are proposed to evaluate the robot's own movement (Self-Supervised Visual Ego Motion Learning) [4]. Note that the concept of self-learning proposed in [3] coincides with the concept of self-learning offered by us.

In [5], the role of curiosity in self-learning is analyzed and the concepts of self-learning with the phenomenon of curiosity are developed.

It is an ordinary practice to associate self-learning with deep learning. Impressive successes in deep learning achieved in simulation games [6] and image analysis [7-16]. However, deep learning does not mean self-learning. Using neural networks for segmenting images traditionally requires a large quantity of training data marked manually. In [14] an algorithm is proposed on the basis of which 130000 images were generated with automatic marking for 39 objects. In [15], a robot’s internal evaluation of its future path cost is based on the probabilistic Bayesian method.

Neural networks recognize classes of objects and form a feature hierarchy of classes, but do not form their symbolic logical descriptions or rules to recognize them. There are a number of works in which attempts are made to find the interconnection between artificial neural networks and symbolic machine learning within the framework of the analysis of formal concepts (FCA) [17-20]. The main purpose of these works is to use the algorithms of constructing the concept lattice to configure the artificial neural networks in order to make it interpretable in terms of concepts. However, an improvement of the artificial neural network learnability has not yet obtained.

In some works, the authors propose the use of robot's manipulation reflection in learning algorithms for improving and accelerating robot's training. For example, industrial Robot of Japanese Company Fanuc uses a method known as "training with reinforcement" to grab objects by a manipulator. In this process, the robot fixes its work on video and uses this video for correcting own activity. Domestic development of robots is also based on the use of artificial neural networks [21-24].

3 Software agent capable of self-learning

Intelligence acts always in a changing context. Several examples of changing contexts can be: the descriptions of patient's conditions supplemented by doctor's decisions and patient's responses, images of the Earth's surface, and student personal characteristics. The task of self-learnable individual or an automatic device in a such changing context is to support any purposeful action or function (search of food, search for exit from a labyrinth, etc.). Intelligence must have some abilities to act in the context by choosing sub-contexts and/or actions in them, as well as by assessing the extent to which its actions bring it closer to the goal. We shall refer to the context in which an intellectual being or device is acting as the external context.

The objects in the external context (training samples) are described in terms of their properties (features, attributes) and they are specified by splitting into classes. The task of learning is to find rules in a given space of object descriptions in order to repeat the classification of objects represented by splitting objects into disjoint classes. Good tests approximate the specified object classification in the best way and give the minimum sets of attributes (values) that carry out the greatest possible generalization within object classes and distinguish in pairs all objects from different classes [1]. As a task in the external context, we have chosen the task of constructing good maximally redundant classification (diagnostic) texts, because the algorithms developed for this task have a number of convenient properties for self-monitoring the process of inferring tests [25]:

- external context is partitioned into sub-contexts in which good tests are inferred independently;

- sub-contexts are chosen and formed by the logical rules based on analyzing sub-contexts' characteristics; the choice of sub-context determines the speed and efficiency of classification task.

The strategies for selecting sub-contexts of the external context and the algorithms to find good tests in them are easy to describe (to represent) with the use of special multi-valued attributes. In what follows, we shall call the intellectual being an agent, although it does not mean that we identify it with the agent in multiagent systems. Summing up the foregoing, we conclude that for self-learning the agent should have:

1. A display of the external context in terms of the internal context);
2. A set of rules (possible actions) for selecting context (sub-context);
3. A display of the desired target (state);

4. An operation (a function) for comparing the desired target with the achieved result.

During the training process, the agent must develop a sequence of actions that will lead to the goal. We shall consider the permanent external context and its changes only in connection with the activity of the agent, for example, a sub-context can be deleted when the agent has completely solved the problem for this sub-context. Decomposition of contexts into sub-contexts in the tasks of inferring good classification tests have been considered in [25-26].

When the agent selects sub-contexts and its (his) actions in learning process, it (he) uses some criteria. These criteria can be: the number of sub-contexts to be considered, the number of tests already extracted in sub-context, the number of objects and values of attributes in sub-context, the number of essential objects and values of attributes (attributes) in sub-context [1], temporal characteristics and some others. The agent needs to memorize the situations of learning and the activity associated with them.

Let us assume that the internal context necessarily contains:

1. Description of selected sub-context in terms of its properties;
2. Description of selected action and the rule for its selection;
3. Internal estimation of learning process with the use of some criteria of its efficiency.

4 The structure of the internal context and realizing self-learning

Let K be the descriptions of external sub-context via its properties, $A = \{A_1, A_2, \ldots. A_n\}$ be the descriptions of algorithms of good tests inferring via their properties in this sub-context, $R = \{R_1, R_2, \ldots,.R_m\}$ be the set of rules for selecting sub-contexts, and $V = \{V_1, V_2, \ldots, V_q)$ be the set of rules for evaluating the process of good test inferring. Then the internal context is described by the direct product of sets K, A, R and its mapping on V: $K \times A \times R \rightarrow V$. There are more simple variants of the internal context: $K \times A \rightarrow V$ and $K \times R \rightarrow V$.

The same algorithm can be used in both the external and the internal context in order to infer the logical rules for distinguishing the variants of learning in the external context evaluated as good ones from the variants evaluated as not good ones. A few algorithms for good test inferring have been elaborated: ASTRA [27], DIAGARA, NIAGARA, and, INGOMAR [28].

We come to the realization of deep learning for the symbolic machine learning tasks. The internal context is a memory of the agent, the rules extracted from the internal context represent the agent's knowledge about the effectiveness of its actions in the external context. Actions in the internal and external contexts can be represented as actions of two agents functioning in parallel and exchange data (Fig. 1).

Agent A1 transmits the data (the descriptions of contexts, algorithms, rules for selecting sub-contexts) to Agent A2. Agent A2 acts in the internal context (obtained from agent A1) and passes to agent A1 the rules, which the latter applies to select the best variant of learning with each new external sub-context.

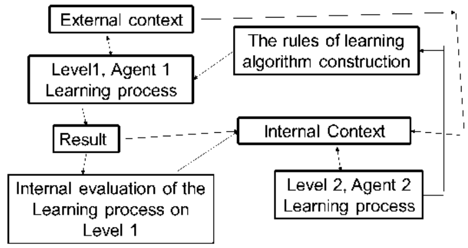

Figure 1. Scheme of self-learning with the interaction of two agents

For Agent A2, the internal context (memory) should not be empty, but this agent (as well as Agent A1) can use an incremental mode of learning [28]. A few incremental algorithms for good test inferring in symbolic contexts are described in [28].

5 The structure of good maximally redundant test inferring

Good test analysis (GTA) deals with the formation of best descriptions of a given object class (class of positive objects) against the objects do not belonging to this class (class of negative objects) on the basis of lattice theory. We assume that objects (or patterns) are described in terms of values of a given set U of attributes. The key notion of GTA is the notion of classification. To give a target classification of objects, we use an additional attribute $k \notin U$. This attribute partitions a given set of objects into disjoint classes the number of which is equal to the number of values of this attribute. We need in the following series of definitions.

Denote by M the set of attribute values such that $M = \cup a \in U\ rng(a)$, where rng(a) is the set of all values of a. Let $G = G+ \cup G-$ be the set of objects, where G+ and G− are the sets of positive and negative objects, respectively.

Let T be a table with many-valued data, where lines correspond to objects and columns correspond to attributes. For representing data, we do not use any scaling.

Denote a description of $g \in G$ by $\delta(g)$, and descriptions of positive and negative objects by $D+ = \{\delta(g)|\ g \in G+\}$ and $D- = \{\delta(g)|\ g \in G-\}$, respectively. The Galois connections [29] between the ordered sets $(2^G, \subseteq)$ and $(2^M, \subseteq)$, i.e. $2^G \rightarrow 2^M$ and $2^M \rightarrow 2^G$, are defined by the following mappings called derivation operators [30]:

for $A \subseteq G$ and $B \subseteq M$, $val(A) = \cap g \in A\ \delta(g)$ and

$obj(B) = \{g|\ B \subseteq \delta(g), g \in G\}$.

There are two closure operators [30, 31]: generalization_of(B) = val(obj(B)) and generalization_of(A) = obj(val(A)). A is closed if A = obj(val(A)) and B is closed if B = val(obj(B)). If (val(A) = B) & (obj(B) = A), then a pair (A,B) is called a formal concept [30, 32], subsets A and B of which are called concept extent and intent, respectively. A triplet (G,M,I), where I is a binary relation between G and M, is a formal context K. According to the values of a goal attribute, we get some possible forms of the formal contexts: $K\varepsilon := (G\varepsilon,M,I\varepsilon)$ and $I\varepsilon := I \cap (G\varepsilon \times M)$, where $\varepsilon \in rng(k)$, $rng(k) = \{+,-\}$ (if necessary the value τ can be added to provide undefined objects) [32]. A classification context $K\pm$ is formed by the sub-position of contexts $K+$ and $K-$, and the apposition of the resulted context with $(G\pm, k, G\pm\times k)$, i.e. after adding the classification attribute k. Let us rewrite the definitions of tests by using notation of classification contexts and semi-concepts [33]: pairs like (obj(B),B), $B \subseteq M$, the left side of which is called an extent, and pairs like (A,val(A)), $A \subseteq G$, the right side of which is called an intent. Here and later words "diagnostic test" (and GMRT) will be used for semi-concepts (or concepts), the right part of which is a test.

Definition 1. A diagnostic test (DT) for K+ is a pair (A,B) such that $B \subseteq M$, $A = obj(B) \neq \emptyset$, $A \subseteq G+$, and $obj(B) \cap G \neq \emptyset$.

Definition 2. A diagnostic test (A,B) for K+ is to be said maximally redundant if $obj(B \cup m) \subset A$ for all $m \in M \setminus B$.

Definition 3. A diagnostic test (A,B) for K+ is to be said good iff any extension $A1 = A \cup i$, $i \in G+ \setminus A$, implies that (A1,val(A1)) is not a DT for K+ .

A maximally redundant test which is simultaneously good is called a good maximally redundant test (GMRT).

Definitions of tests (as well as other definitions), associated with K+, are applicable to K−.

If a good DT (A,B) for K+ is maximally redundant, then any extension $B1 = B \cup m$, $m \notin B$, $m \in M$ implies that (obj(B1),B1) is not a good DT for K+.

In the general case a set B is not closed for DT (A,B), consequently, DT is not obligatorily a formal concept. A GMRT can be regarded as a special type of formal concept [1]. Note that the definition of GMRTs is equivalent to the definition of inclusion-minimal concept-based hypothesis in the FCA [30].

To transform inferring GMRTs into an incremental process, we introduce two kinds of subtasks for K+ (K−), called subtasks of the first and second kind, respectively [34]:

1. Given a positive object g, find all GMRTs (obj(B),B) for K+ such that B is contained in $\delta(g)$. In the general case, instead of $\delta(g)$ we can consider any subset of values B1, such that $B1 \subseteq M$, $obj(B1) \neq \emptyset$, $B1 \not\subseteq \delta(g)$, $\forall g \in G-$.

2. Given a non-empty set of values $B \subseteq M$ such that (obj(B),B) is not a DT for positive objects, find all GMRTs (obj(B1),B1) such that $B \subset B1$.

Accordingly, we define two kinds of sub-contexts of a given classification context called object and attribute value projections, respectively. If (G,M,I) is a context and if $H \subseteq G$, and $N \subseteq M$, then $(H,N,I \cap H \times N)$ is called a sub-context of (G,M,I) [35].

Definition 4. The object projection $\psi(K+,g)$ returns sub-context $(N,\delta(g),J)$, where $N = \{n \in G+ \mid n$ satisfies $(\delta(n) \cap \delta(g)$ is a test for $K+)\}$, $J = I+ \cap (N \times \delta(g))$.

Definition 5. The attribute value projection $\psi(K+,B)$ returns sub-context (N,B,J), where $N = \{n \in G+ \mid n$ satisfies $(B \subseteq \delta(n))\}$, $J = I+ \cap (N \times B)$. In the case of negative objects, symbol + is replaced by symbol − and vice versa.

The decomposition of inferring GMRTs into the subtasks requires the following actions:

1. Select an object or value to form a subtask.
2. Form the subtask.

3. Reduce the subtask.
4. Delete the object or value when the subtask is over.

The following theorem gives the foundation for reducing sub-contexts formed by object and attribute value projections [27, 28].

Theorem 1. Let B ⊆ M, (obj(B),B) be a maximally redundant DT for positive objects and obj(m) ⊆ obj(B), m ∈ M. Then m cannot belong to any GMRT for positive objects different from (obj(B),B).

6 A procedure for mining the all GMRTs in the projections of both kinds

Let Sgood+ (Sgood−) be the partially ordered set of obj+(m), m ∈ M satisfying the condition that (obj+(m), val(obj+(m))) is a current good DT for K+ (K−). The basic recursive procedure (BRP) for K+ is defined in Fig. 2, where

• the first step of recursion is omitted for simplicity;

• the output Sgood+ is implicitly given via a globally defined set, which is modified during the procedure; algorithm formSgood is given in Fig. 3;

• variable ψtype has two possible values: object or attribute value projection;

• algorithm choiceOfprojection returns ψtype, and X, which can be either g or B w.r.t. value of ψtype;

• algorithm formSubcontext implements a definition of object or attribute value projection and returns new subcontext K∗; conditions for the end of recursion are described in steps 7, 25;

• after the end of the current recursion iteration the control goes to the previous recursion iteration from steps 13, 31;

• checking whether (obj+(m),val(obj+(m))) is a DT for K+ is performed as follows: val(obj+(m)) is a test for K+ iff obj(val(obj+(m))) = obj+ (m).

Procedure BRP

Input: K+,K−,Sgood+

Output: Sgood+

1. f := 0;
2. forall m ∈ M do
3. if val(obj+(m)) is a test for K+ then
4. formSgood(obj+(m),Sgood+);
5. M := M \m, f := 1;
6. end
7. if |M|≤ 1 then
8. return;
9. if f = 0 then
10. ψtype, X choiceOfprojection (K+,K−);
11. K∗ + formSubcontext(ψtype,X,K+);
12. BRP (K∗ +,K−,Sgood+);
13. if ψtype = object projection then

```
14. G+ := G+ \X; 15. else
16. M := M \X;
17. else
18. f := 0;
19. end
20. forall g ∈ G+ do
21. if val(g) is not a test for K+ then
22. G+ := G+ \g;
23. f := 1;
24. end
25. if |G+|≤ 1 then
26. return;
27. if f = 0 then
28. ψtype, X choiceOfprojection (K+,K−);
29. K∗ + formSubcontext(ψtype,X,K+);
30. BRP (K∗ +,K−,Sgood+);
31. if ψtype = object projection then
32. G+ := G+ \X;
33. else
34. M := M \X;
35. else
36. go to 1;
37. end
```

Figure 2. Pseudo code of basic recursive procedure

7 Forming SGOOD as the main problem of good test inferring

Essentially, the process of forming Sgood is an incremental procedure of finding all maximal elements of a partially ordered (by inclusion relation) set. It is based on topological sorting of partially ordered sets. Thus, when the algorithm is over, Sgood contains the extents of all the GMRTs for K+ (for K−) and only them. The operation of inserting an element A∗ into Sgood (in algorithm formSgood) under lexicographical ordering of these sets is reduced to lexicographically sorting a sequence of k-element collections of integers.

A sequence of n-collections whose components are represented by integers from 1 to |M|, is sorted in time of O(|M| + L), where L is the sum of lengths of all the collections of this sequence [36]. Consequently, if Lgood is the sum of lengths of all the collections A of Sgood, then the time complexity of inserting an element A∗ into Sgood is of order O(|M| + Lgood). The set Tgood of all the GMRTs is obtained as follows: Tgood = {t|t = (A,val(A)), A ∈ Sgood}.

Algorithm formSgood

Input: A∗ ⊆ G+,Sgood+

Output: Sgood+

```
forall A ∈ Sgood do
if A ⊂ A∗ then
Sgood+ := Sgood+ \A;
else
if A∗ ⊆ A then
return;
end
Sgood+ := Sgood+ ∪A∗;
return;
```

Figure 3. Pseudo code of algorithm formSgood

8 Some problem to be solved

In self-learning, it is very important determining the nearness of the current result to the goal of learning process. The goal in mining GMRTs is to find the all GMRTs for a given external context. Generally, a situation can be when there exist sub-contexts of the external context to be solved, but the saturation of S_{GOOD} is already achieved (i. e., all GMRTs are obtained). A procedure of determining the saturation of S_{GOOD} can be based on the properties of the set of all GMRTs of a formal context to be the Sperner System [37].

It is important to formulate some unsolved and nontrivial problems related to the decomposition considered in this paper. These problems are:

- How to recognize a situation that current formal context contains only the GMRTs already obtained?
- How to evaluate the number of recurrences necessary to resolve a subtask in inferring GMRTs? (in case we use a recursive algorithm like DIAGARA)?
- How to evaluate the perspective of a selected sub-context with respect to finding any new GMRT?

These problems are interconnected and the subject of our further research. The effectiveness of the decomposition depends on the properties of the initial classification context (initial data). Now we can propose some characteristics of data (contexts and sub-contexts) useful for choosing a projection:

- The number of objects;
- The number of attribute values;
- The number of the GMRTs already obtained and covered by this projection.

Some unsolved problems cited above are difficult for analytical solution. It is possible that realizing the proposed approach to self-improving learning algorithms permits one to investigate these problems and enables us to overcame the above difficulties.

One of the advantages of our approach is related to the possibilities to reduce the process of choosing sub-contexts and to obtain the best variant of learning to the plausible deductive reasoning, one of the models of which is described in [28]. Modeling of on-line human reasoning is a key problem in creating intelligent computer systems.

However, any attention is hardly paid to this topic in computer science. Knowledge engineering has arisen from a paradigm in which knowledge is considered as something to be separated from its bearer and to function autonomously with a problem-solving application. This paradigm ignores the very essential feature of intelligence, namely, its continuous cognitive activity. Knowledge is corrected constantly. This means that the mechanism of using knowledge cannot be separated from the mechanism of discovering knowledge. The future realization of our approach to self-improving good test inferring will support using logical rules extracted from the internal context for deductive process of choosing variants of learning.

9 Conclusions

The concept of self-learning in the processes of inferring good classification tests is proposed in the paper. The inferring of good classification tests is a task of symbolic machine learning, for which the questions of self-learning has been not considered earlier. The results of this article are the following:

A model of self-learning was proposed allowing to manage the process of inferring good tests in terms of its effectiveness through an internal evaluation of the learning process and the development of rules for choosing the best strategies, algorithms, and learning characteristics.

The concepts of internal and external learning contexts were formulated.

The structure of the internal context was proposed.

A model of intelligent agent, capable of improving own learning process of inferring good classification tests in the external context was advanced;

It was shown that the same learning algorithm can be used for supervised learning both in the external context and in the internal context. The proposed approach is a model of deep learning implemented by inferring logical rules from examples.

Acknowledgments. The research is partially supported by RFR grant № 18-07-00098A.

References

1. Naidenova, X.: Good diagnostic tests as formal concepts. In: F. Domenach, D.I. Ignatov, J. Poelmans (eds.) ICFCA-2012, LNCS vol. 7278, Springer, pp. 211-226 (2012).
2. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. The MIT Press (2016).
3. Pillai, S.: Towards richer and self-supervised perception in robots. PhD Thesis Proposal (2017) http://people.csail.mit.edu/spillai/research/ and http://people.csail.mit.edu/spillai/data/papers/2017-phdthesis-proposal-nocover.pdf
4. Pillai, S., Leonard, J.: Towards Visual Ego-motion Learning in Robots. Submitted to IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS) (2017). http://people.csail.mit.edu/spillai/learning-egomotion/learning-egomotion.pdf
5. Pathak, D., Agraval, P., Efros, A., Darrell, T.: Curiosity-driven exploration by self-supervised prediction. In: Proc. of the 34th Int. Conf. on Machine Learning, JMR: W&CP, pp. 12 (2017) https://pathak22.github.io/noreward-rl/

6. Silver, D., Huang, A., Maddison, Ch. J. et al.: Mastering the game of Go with deep neural networks and tree search. Nature **529**, 484-489 (2016). doi: 1038/nature 16961
7. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556v6 [cs.CV] (2015).
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. Computer Vision and Pattern Recognition, arXiv:1512.03385v1[cs.CV] (2015).
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: F. Pereira, C.J.C. Burges et al. (eds.) Advances in neural information processing systems, pp. 1097-1105 (2012).
10. Potapov, A., Batishcheva, V., Pan Shu: Improving the quality of recognition in deep learning networks using the method of simulation of annealing. Scientific and Technical Bulletin of Information Technologies, Mechanics, and Optics, 17(4), (2017). (in Russian)
11. Hossain, D., Capi, G. Jinday, M.: Evolution of Deep Belief Neural Network Parameters for Robot Object Recognition and Grasping. Procedia Computer Science 105, 153-157 (2017).
12. Schmidt, T., Newcombe, R., Fox, D.: Self-supervised Visual Descriptor Learning for Dense Correspondence. IEEE Robotics and Automation Letters, 2(2), 420-427 (2016). doi: 10. 1109/LRA.2016.2634089
13. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. arXiv:1408.5093 [cs.CV] (2014).
14. Zeng, A., Yu, Kuan-Ting, Song, S., Suo, D., Walker, Ed., Rodrigues, A., Xiao, J.: Multi-View Self -Supervised Learning for 6D Pose Estimation in the Amazon Picking Challenge. In: Proceedings of IEEE International conference on Robotics and Automation (ICRA), pp. 1986-1993 (2017). doi: 10.1109/ICRA.2017.7989165
15. Sofman, B., Line, E. et al.: Improving Robot Navigation Through Self-Supervised Online Learning. Journal of Field Robotics, 23(11-12), 1059-1075 (2016). doi: 10.1002/rob.20169
16. Long, J., Shelhamer, E., Darrell, T.: Fully Convolutional Networks for Semantic Segmentation. arXiv:1605.06211v1 [cs.CV] (2016).
17. Endres, D., Foldiak, P.: Interpreting the neural code with formal concept analysis. In: Koller, D., Schuurmans, D., Bengio, Y., Bottou, L. (eds.) Advances in Neural Information Processing Systems, vol. 21, pp. 425–432. MIT Press, Cambridge (2008).
18. Kuznetsov, S.O., Makhazhanov, N., Ushakov, M.: On Neural Network Architecture Based on Concept Lattices, LNAI, vol. 10352, pp. 653–663 (2017). Available https//doi.org/10.1007/978-3-319-60438-1_64
19. Rudolph, S.: Using FCA for Encoding Closure Operators into Neural Networks. In: Priss, U., Polovina, S., Hill, R. (eds.) ICCS 2007, LNAI, vol. 4604, pp. 321–332. Springer, Berlin- Heidelberg (2007).
20. Tsopzé, N., Nguifo, E.M., Tindo, G.: CLANN: Concept lattice-based artificial neural network for supervised classification. In: Proceedings of the Fifth International Conference on Concept Lattices and Their Applications, vol. 331, pp. 153–164 (2007).
21. Pavlovsky, V. Savitsky, A: A Quadrocopter neural network control algorithm on typical trajectories. Nonlinear World, 13(6), 47-54 (2016). (in Russian)
22. Aliseychik, A., Orlov, I., Pavlovsky, V., Smolin, V., Podoprosvetov, A., Shishova, M.: Pneumatic manipulation with neural network control. LNCS, vol. 9719, pp. 292-301 (2016).
23. Savitsky, A. V., Pavlovsky, V. E.: Model of quadrotor and algorithm of vehicle control based on neural network. Keldysh Institute preprint 077 (2017). http://library.keldysh.ru/preprint.asp?id=2017-77 (in Russian)

24. Pavlovsky, V.E., Pavlovsky V.V.: Technologies SLAM for moving robots: State and Prospects. Mechatronics, Automation, Management, 17(6), 384-394 (2016). (in Russian)
25. Naidenova, X., Parkhomenko, V., Shvetsov, K.: Context-Dependent Incremental Learning Good Maximally Redundant Tests. SAI Intelligent Systems Conference 2015, pp. 1-6. London, UK: IEEE (2015). DOI:10.11.1109/IntelliSys.2015.7361258 https://www.researchgate.net/publication/292608731_Context-Dependent_Incremental_Learning_of_Good_Maximally_Redundant_Tests
26. Naidenova, X., Parkhomenko, V.: Context-Dependent Classification Reasoning Based on Good Diagnostic Tests. Proc. of FCA&A' 2015 (co-located with ICFCA'2015), J. Baixeries, Ch. Sacarea, and M. Ojeda-Aciego (eds), pp. 65-80. University de Malaga (2015). ISSN-84-606-7410-8. http://ceur-ws.org/Vol-1434/proceedings-fcaa.pdf
27. Naidenova, X., Plaksin, M. Shagalov, V.: Inductive inferring all good classification tests. Proceedings of International Conference "Knowledge-Dialog-Solution", vol. 1, pp.79-84. Jalta, Ukraine (1995).
28. Naidenova, X.: An incremental learning algorithm for inferring logical rules from examples in the framework of the common reasoning process. In: Triantaphyllou, E., & Felici, G. (eds.), Data mining and knowledge discovery approaches based on rule induction techniques, pp. 89–146. New York, NY: Springer. (2006).
29. Ore, O.: Galois connections. Trans. Amer. Math. Soc 55 (1944) 494–513.
30. Ganter, G., Kuznetsov, S. O.: Pattern Structures and Their Projections. In: H. S. Delugach, G. Stumme (eds.), Conceptual Structures: Broadening the Base, Proceedings of the 9th International Conference on Conceptual Structures, 129–142 (2001).
31. Naidenova, X.: The Data-Knowledge Transformation. In: V. Solovyev (ed.), Text Processing and Cognitive Technologies, vol. 3, Pushchino, 130–151 (1999).
32. Ganter, B., Kuznetsov, S. O.: Formalizing Hypotheses with Concepts. In: Conceptual Structures: Logical, Linguistic, and Computational Issues, Proceedings of the 8th International Conference on Conceptual Structures, 342–356 (2000).
33. Luksch, P., Wille, R.: A Mathematical Model for Conceptual Knowledge Systems. In: H.-H. Bock, P. Ihm (eds.), Proceedings of the 14th Annual Conference of the Gesellschaft fur Klassifikation (GfKl 1990), 156–162 (1991).
34. Naidenova, X., Parkhomenko, V.: Attributive and Object Sub-contexts in Inferring Good Maximally Redundant Tests. In: K. Bertet, S. Rudolph (eds.), Proceedings of the Eleventh International Conference on Concept Lattices and their Applications, Košice, Slovakia, October 7-10, 2014., vol. 1252 of CEUR Workshop Proceedings, 181–193 (2014).
35. Ganter, B., Wille, R.: Formal concept analysis: mathematical foundations. Springer, Berlin (1999).
36. Aho, A. V., Hopcroft, J.E., Ullman, J.D.: The design and analysis of computer algorithms. Addison-Wesley (1975).
37. Sperner, E.: Ein Satz über Untermenge einer Endlichen Menge. Math. Z. 27, 544-548 (1928).