# TOWARDS A NOVEL DESIGN OF CREATIVE SYSTEMS

**Vladimir Jotsov**

***Abstract***: *The topic of the presented investigation is the automation of creative processes via one Synthetic Mata Method (SMM) and few analytic methods considered in the paper and used under SMM control. The prevention of contemporary web threats is discussed at an agent/application level. Advantages and disadvantages of synthetic data mining methods are investigated, and obstacles are revealed to their application in contemporary systems. Novel results for juxtaposing statistical vs. logical data mining methods aiming at possible evolutionary fusions are described. Recommendations are made on how to build more effective applications of classical and/or presented novel (meta) methods: SMM, KALEIDOSCOPE, FUNNEL, PUZZLE, and CONTRADICTION. The usage of ontologies is investigated with the purpose of information transfer by sense. Practical aspects of agent applications, intrusion detection, intrusion prevention, cryptography applications, multiple software and other research results are mentioned aiming to show that intelligent and classical technologies should be carefully combined in one software/hardware complex to achieve the creative goals. It is shown that all the demonstrated advantages may be combined with other known methods and technologies.*

***Keywords:*** *automation of creative processes, human-machine brainstorming methods, knowledge discovery, data mining, web mining, ontology, information security systems, intrusion detection, intrusion prevention, human-centered systems, knowledge management, agent, collective evolution.*

## 1. Introduction

There exist a plenty of philosophy - or psychology-based research on creative systems [1] but the technically-oriented papers like [2] are rather scarce. On the other hand, creative systems are welcome to the field where contemporary statistical, data mining, web mining systems are still not efficiently used, especially in the dynamic environments of information security systems. Methods which in my opinion are essential to every creative system or to next generation knowledge discovery system are discussed. One of them, a conflict resolution method, allows the agent or other application element to be really autonomous or reinforces the possibilities of the contemporary systems.

Contemporary Information Security Systems (ISS) and especially the web-based systems represent a wide field for applications of modern creative methods and technologies. The need to create sufficiently effective and universal tools to protect computer resources grows every year in systems for detection and prevention from intrusions (Intrusion Detection Systems IDS, Intrusion Prevention Systems IPS). For this reason different applications of intelligent data processing are initiated based on a combination of methods from statistical and logical information processing [3, 4]. Other elaborations with growing influence in this domain are artificial immune systems and multiagent systems [5, 6]. The unifying factor is the longer life cycle, elaborations require bigger teams and time for introduction. Due to the complex structure the prevention from direct attacks against these systems is very challenging.

Modern applications of statistical methods are effective and convenient to use at the expense of information encapsulation. In other words, it is impossible to construct tools to acquire new knowledge or to solve other problems of logical nature in this area. If we split methods in two groups (quantitative and qualitative) then statistical methods belong to the first group and logical ones belong to the second group. For this reason, their mechanical union is of no perspective. We do not attempt to propose any isolated solutions; instead we offer a combination of novel methods that is well adjustable to the existing ones. Our research includes a new evolutionary meta-method for joint control of statistical and logical methods where the statistical approach is widely applied on the initial stage of the research when the information about the problem is scanty and it is possible to choose the solution arbitrarily [7]. The accumulation of knowledge makes logical applications more and more effective and more universal than the probabilistic ones, as well as fuzzy estimates and similar applications. The paper introduces the SMM (Synthetic Meta-Method) meta-method to control the process of consecutive replacement of applications by other ones and is synthetic by nature. The difference from the classical analytic methods is in the fact that the design of systems controlled by synthetic (meta) methods is not just science, sometimes it is an art. If we make an analogy with the set of traditional methods and fashion clothes then the synthetic method will apply the design of the display window with the fashion clothes. In the common case during intelligent data processing, there is no convergence of the results but this does not hamper practical applications of these systems. In other words, bad and good designers will arrange the display window in quite different ways and there is no guarantee that every user understands the technology and that his access to the system will have positive results.

The cited innovations are made and demonstrated here for the following reason. The problems presented above show that there is a need to introduce elements of machine creative work in ISS. It is demonstrated in the paper that this goal is accessible, if the usage of possibilities for human-machine contact and a set of comparatively simple intelligent technologies are done in the right way. On the other hand, the innovations can hardly be described in a single paper using the traditional academic style. For this reason, in the paper we avoided when possible the technical details and formalizations, and for the sake of the contents reduction descriptions by analogy are used, illustration visual aids and other nonstandard approaches.

How can security agents operate autonomously? In the first place this is because of the usage of neural networks where the agent most conscientiously copies the acts of the teacher. At that the agent itself does not understand the sense of teacher's transfer of knowledge; it operates as universal approximator instead. It is shown schematically at Fig. 1 in the following way.

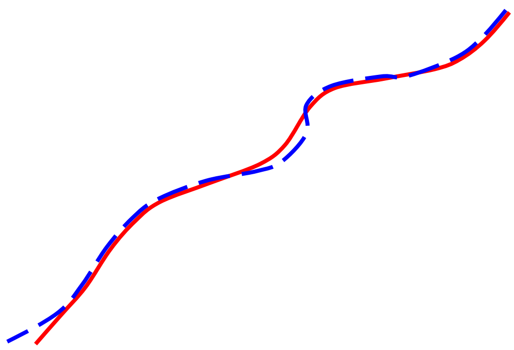

**Figure 1**. Approximation of teacher's activities

Teacher's acts are principally presented as a dotted curve. The neural network approximates this curve via a continuous line on the same figure; the deviation between the two lines must not exceed 3%. One of the main disadvantages of this approach is related to the necessity the agents in ISS to apply the learned knowledge in a rather creative way because frequently they operate in unexpected by the teacher situations. And as they are poorly trained or they are not at all trained how to act in unexpected situations, this approach as a whole is not very effective in its classical appearance. The presented approaches in anomaly ISS are combined with statistical applications which count for the normal traffic and other mean statistical values related to the operation of the guarded place. In Fig. 2 this process is schematically shown in the following way.

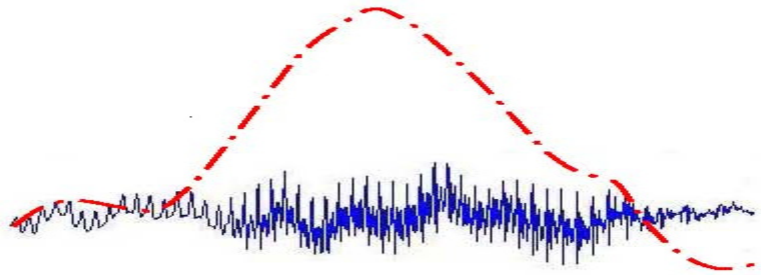

**Figure 2**. Anomaly detection

The normal traffic is presented via a continuous line and the dotted line shows a case with drastic deviation from the normal traffic which is considered as an anomaly and it is analyzed whether it is a consequence from an intrusion. One of the disadvantages of this approach lies in the fact that intruder(s) know the principles of operation of such systems and they can 'fit' in the constraints of the normal traffic.

There is a variety of other applied approaches when the system is overloaded by heuristic information but they are not discussed here due to their evident weakness for ISS.

Here we introduce a new way for agents' operation. Best of all is the usage of ontologies to model the domain but this is not obligatory. On the other hand, as it is shown in the next section, the multiagent system functions more effectively if a system of ontologies is included.

But ontologies also do not contribute to a great extent in order to understand the sense of matters by the agents or with respect to transmit the sense of matters during communications between the agents. On the way to produce an analogy to how agents think we offer evolutionary methods to process data or knowledge because thinking (and understanding) is an evolutionary process. The problem is how to direct evolutionary methods with no exaggerations of heuristics, statistical information and other relative methods. Our approach is rather untraditional. We have elaborated for more than twenty years methods to detect and to resolve contradictions. A method for machine learning is built based on them. Searching and solving conflicts and contradictions the agent improves its knowledge and at the same time it may solve other problems. Detection of contradictions is based on using models of contradictions that can also be improved gradually. The agent may request an external help to solve the conflicts but this takes place only in extraordinary situations. At the same time it is shown how to change the reasoning component of security agents. Different logical methods are used that are rather analogous to means-ends analysis, constraint satisfaction, variable fitness function, brainstorming, and cognitive graphics. The combination of new methods to a great extent mechanizes creative efforts and also it serves agents' operation to improve the abilities of security experts working with similar types of systems.

Suggested innovations serve the more effective application of data mining, Web mining, collective evolutionary components in multiagent systems. They are very well combined especially for applications of evolutionary approaches with classical neuro-fuzzy, statistical applications, genetic algorithms, and etc. methods for the domain. For example such systems critically accept teacher's acts in cases of supervised learning (Fig. 3): they may precise or argue teacher's acts and in this way they can learn more effectively and deep.

A wide application of intelligent agents is forecasted in the field of information security systems. This will lead to the situations when the agent has no possibility to learn from the expert (teacher) but should swiftly learn from other agents or should self-learn without teacher. Then the role of the above considered critical learning will significantly increase.

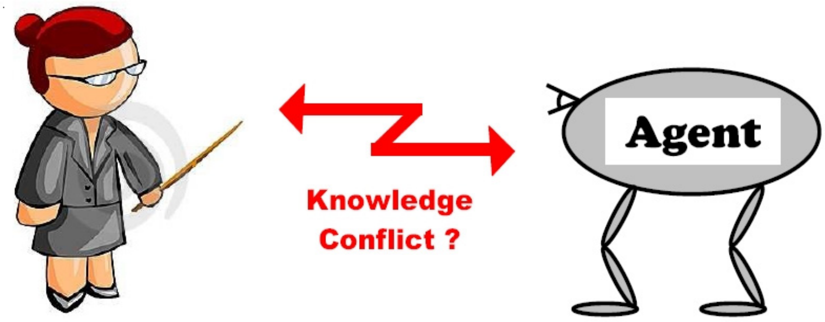


**Figure 3**. Example of supervised learning by using critical analysis

## 2. Synthetic Methods: a Necessary Background to Future Creative Systems

A synthetic meta-method (SMM) is elaborated and applied aiming at application of a set of ‘creative’ elements in agent environments. They work most effectively as a system, but even particular elements of them, let’s say, applied in semantic reasoners, are proved to be very useful. Their principles are easy-to-be explained: bind the unknown knowledge from the goal with the knowledge from the knowledge base, apply a flexible contraint system to manage a system of variable fitness/goal functions, make the goals automatically via self-improvement of the existing knowledge, etc. The consize goals are better applied in the intelligent agents. The set of methods necessary to every creative system is considered in this Section. The coordination activities under the control of SMM is discussed in Section 3. Ontologies and other applications are represented in Sections 4 and 5.

### 2.1 PUZZLE Method

The basic methods of the suggested meta-method SMM are presented below. Let the constraints connected to the defined problem form a line in the space described by the equation (1).

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1} \qquad (1)$$

For example if a bachelor who has graduated our university SULSIT lives in Sofia and he/she does not want to work anywhere else, then the line restricts the search space and in this way a lot of unnecessary work is avoided. It is also possible to inspect a case when the constraint is defined as a type of surface but as a result a more general solution is obtained where a special interest is provoked by the boundary case of the crossing of two or more surfaces. When the common case is inspected in details, then in the majority of cases the problem is reduced to exploring the lines of type (1) or to curves with complicated forms obtained as a result of crossing surfaces by constraints. Therefore, below we investigate the usage in systems of constraints by lines of first or higher orders.

If the mentioned curves have common points of intersection and if they lie in a common plain so that a closed figure (triangle, tetragon, etc.) is formed then the search space is significantly reduced and it is searched much easier. The practical usage of the classical technology, as well as the constraint satisfaction, is complicated by the following. The viewed plains are not only nonlinear in the common case but they also include fuzzy estimates. The usage of fuzzy logic significantly raises the algorithmic complexity of the problem and it can make the application ineffective. Even when the usage of constraints significantly reduces the number of the inspected solutions, for example up to 10, this does not mean that the problem is solved and that all that must be done is to explore the possibilities one by one.

Let's admit that $k_i \in K$ are the elements of the solution, or, same, of the objective goal G, and $w(k_i)$ are the weights of the elements on the condition that the weight of the goal equals to one. In different situations a concrete measure is chosen and that is why the order of weights' processing is different. A popular example is measuring the length to the goal and then $w(k_i)$ equals to the length of the piece $k_i$ from the segment G. Then S from the next formula shows the reliability of the inference 'G is true'. If $k_i \in K$ are continuous values then there integral calculus must be applied:

$$S = \frac{\sum_i k_i}{1 - \sum_i k_i}$$

It follows from the above formula that the bigger part of the goal G is confirmed by the knowledge, the easier is the goal confirmation. The detection of point of intersection between the available knowledge and the unknown goal uses an original method of the PUZZLE that binds the unknown knowledge (the goal) to the known knowledge (the accumulated up to the moment information).

The following example below shows how the search process is reduced via using ontologies. Let's admit that the search space is presented on fig. 4 where statistical data about ISS are generalized about the regions depending on their price and quality. It is necessary to select an acceptable ISS to our project.

In fig. 5 a subset of feasible solutions is chosen without ISS designed outside Europe. The space of feasible solutions is to the left of the separating surface that is depicted on the figure in blue color.

In fig. 6 another surface in green is shown delimiting the search space of the solutions. In our case it is 'systems with unknown principles of operation'. It is accepted that in the data bases there is no clear distinction related to the presented criteria so the search of the feasible solutions is nonlinear of high dimensionality and practically it cannot be solved using traditional methods. Nevertheless, by applying ontologies analogous to the ones from the previous section the problem is solvable via the PUZZLE method.

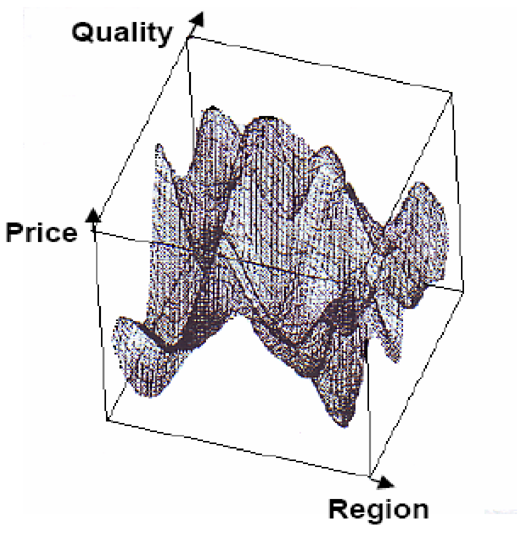


**Figure 4**. Example of a space of solutions

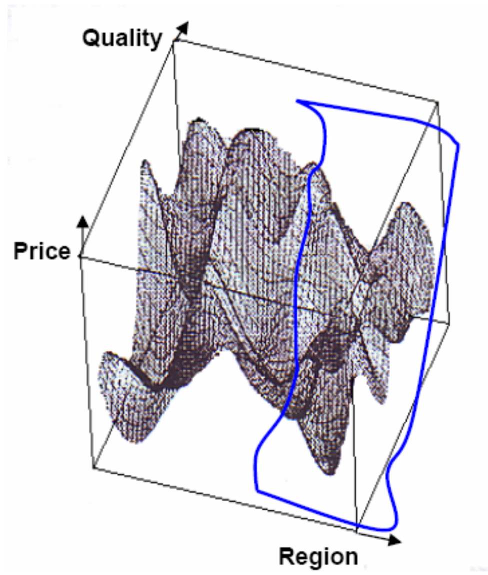


**Figure 5.** Nonlinear space division(s) of the region

There are two red dots on the same fig. 6 in its left corner. Each of them is a kind of constraint but of another type which we name a binding constraint and it is introduced by us. Its semantics is the following: it is not a solution but it resides close to the searched solution. For example we have the information that Fensel's elaborations are a good solution to the problem and that they define the left dot; the right dot has semantics of some other type. By introducing new constraints, our goal is to show that it is possible to use causal links that are different from implications.

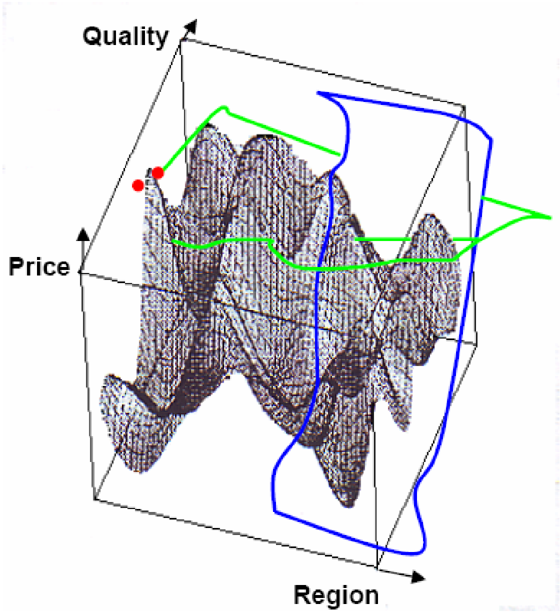


**Figure 6**. Binding and other constraints

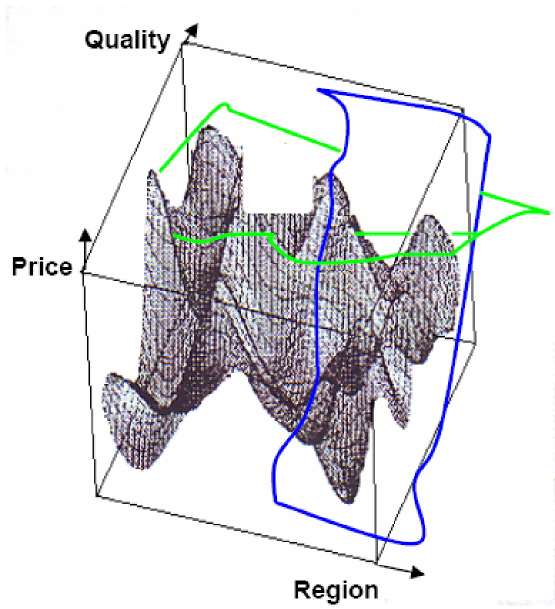


**Figure 7.** Two nonlinear intersections best executed by using ontologies

The same situation is presented on fig. 7 but some of the solutions are absent and this is evident in comparison to the images from the previous figures. It will be demonstrated below that the pointed incompletenesses are often met often and, even in this situation which is an obstacle for other existing methods, we offer an effective solution. In Section 3 it is discussed that the nonlinear surfaces could be efficiently represent by ontologies.

## 2.2. FUNNEL Method

Below, we discuss in brief the next proposed FUNNEL method. Fig. 8 presents the main elements of the method: a system of constraints in the form of a funnel around a center which is a goal (fitness) function which points to the desired direction for information output or to search for new knowledge.

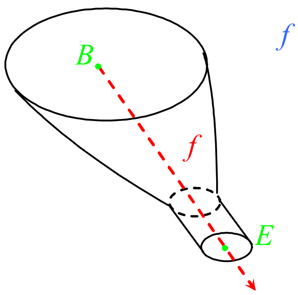


**Figure 8**. A funnel-type system of constraints

As it is evident, the goal of this method is the gradual narrowing of the space of the feasible solutions, together with the progress of the dynamic information processes. Usually the FUNNEL method operates properly when combined with the other methods introduced here and that is why its peculiarities are viewed in detail in the next section where the interactions between the methods are examined. For example, it is convenient to concentrate on fig. 7 shown above over one of the peek values of the diagram by fixing a funnel above it.

## 2.3. Conflict Resolution Method and its Machine Learning Applications

Any lack of collaboration in a group of agents or intrusion could be found as an information conflict with existing models. Many methods exist where a model is given and every non-matching it knowledge is assumed as contradictory. Let's say, in an anomaly intrusion detection system, if the traffic has been increased, it is a contradiction to the existing statistical data and an intrusion alert has been issued. The considered approach is to discover and trace different logical connections to reveal and resolve conflict information. The constant inconsistency resolution process gradually improves the system DB and KB, and leads to better intrusion detection and prevention. Models for conflicts are introduced and used, and they represent different forms of ontologies.

Let the strong (classical) negation be denoted by '¬' and the weak (conditional, paraconsistent) negation [8, 9, 10] be '~'. In the case of an evident conflict (inconsistency) between the knowledge and its ultimate form–the contradiction–the conflict situation is determined by the direct comparison of the two statements (the conflicting sides) that differ one from another by just a definite number of symbols '¬' or '~'. For example: A and ¬A; B and not B (using ¬ equivalent to 'not'), etc.

In the case of implicit (or hidden) negation between two statements, A and B can be recognized only by an analysis of preset models of the type of (2)

$$\{U\}[\eta\text{: } A, B], \tag{2}$$

where $\eta$ is a type of negation, U is a statement with a validity including the validities of the concepts A and B, and it is possible that more than two conflicting sides may be present. It is accepted below that the contents in the figure in brackets U is called unifying feature. In this way, it is possible to formalize not only the features that separate the conflicting sides but also the unifying concepts joining the sides. For example, the intelligent detection may be either automated or of a human-machine type but the conflict cannot be recognized without the investigation of the following model

{detection procedures}[¬: automatic, interactive].

The formula (2) formalizes a model of the conflict the sides of which unconditially negate each another. In the majority of the situations, the sides participate in the conflict only under definite conditions: $\chi_1, \chi_2, \ldots\chi_z$.

$$\{U\}[\eta\text{: } A_1, A_2, \ldots A_p] <\tilde{\chi}_1 * \tilde{\chi}_2 * \ldots * \tilde{\chi}_z>, \tag{3}$$

where $\chi$~ is a literal of $\chi$, i.e. $\chi$~ $\equiv\chi$ or $\chi$~ $\equiv\eta\chi$, * is the logical operation of conjunction, disjunction or implication.

The present research allows a transition from models of contradictions to ontologies [11] in order to develop new methods for revealing and resolving contradictions, and also to expand the basis for cooperation with the Semantic Web community and with other research groups. This is the way to consider the suggested models from (2) or (3) as one of the forms of static ontologies.

The following factors have been investigated:

T – time factor: non-simultaneous events do not bear a contradiction;

M – place factor: events that have taken place not at the same place, do not bear a contradiction. In this case, the concept of place may be expanded up to a coincidence or to differences in possible worlds;

N – a disproportion of concepts emits a contradiction. For example, if one of the parts of the contradiction is a small object and the investigated object is very large, then and only then it is the case of a contradiction;

O – identical object. If the parts of the contradiction are referred to different objects, then there is no contradiction;

P – the feature should be the same. If the parts of the contradiction are referred to different features, then there is no contradiction;

S – Simplification factor. If the logic of user actions is executed in a sophisticated manner, then there is a contradiction;

W – Mode factor. For example, if the algorithms are applied in different modes, then there is no contradiction;

MO – contradiction to the model. The contradiction exists if and only if (*iff*) at least one of the measured parameters does not correspond to the meaning from the model. For example, the traffic is bigger than the maximal value from the model.

**Example:** We must isolate errors that are done due to lack of attention from tendentious faults. In this case we introduce the following model (4):

$$\{ \text{user : faults} \}[\sim\text{: accidental, tendentious}] <T,\neg M,O;\neg S>. \quad (4)$$

It is possible that the same person does sometimes accidental errors and in other cases tendentious faults; these failures must not be simultaneous on different places and must not be done by same person. On the other hand, if there are multiple errors (e.g. more than three) in short intervals of time (e.g. 10 minutes), for example, during authentications or in various subprograms of the security software, then we have a case of a violation, nor a series of accidental errors. In this way, it is possible to apply comparisons, juxtapositions and other logical operations to form security policies thereof.

Recently we shifted conflict or contradiction models with ontologies that give us the possibility to apply new resolution methods. For pity, the common game theoretic form of conflict detection and resolution is usually heuristic-driven and too complex. We concentrate on the ultimate conflict resolution forms using contradictions. For the sake of brevity, the resolution groups of methods are described schematically.

The conflict recognition is followed by its resolution. The schemes of different groups of resolution methods have been presented in Fig. 9 to Fig. 12.

In situations from Fig. 9, one of the conflicting sides does not belong to the considered research space. Hence, the conflict may be not be immediately resolved; only a conflict warning is to be issued in the future. Let's say, if we are looking for an intrusion attack, and side 2 matches printing problems, then the system could avoid the resolution of this problem. This conflict is not necessary to be resolved automatically; experts may resolve it later using the saved information. In Fig. 10, a situation is depicted where the conflict is resolvable by stepping out from the conflict area. This type of resolution is frequently used in multi-agent systems where conflicting sides step back to the pre-conflict positions and one or both try to avoid the conflict area. In this case a warning on the conflict situation has been issued.

The situation from Fig. 11 is automatically resolvable without issuing a warning message. Both sides have different priorities, say side 1 is introduced by a security expert, and side 2 is introduced by a non-specialist. In this case, side 2 has been removed immediately. A situation is depicted on Fig. 12 where both sides have been derived by an inference machine, say by using deduction. In this case, the origin for the conflict could be traced, and the process is using different human-machine interaction methods.

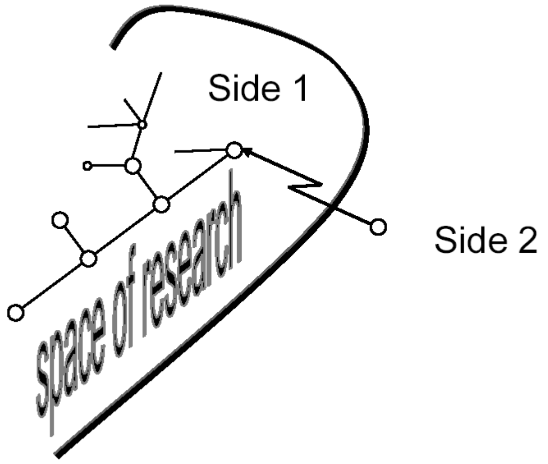


**Figure 9.**Avoidable (postponed) conflicts when Side 2 is outside of the research space.

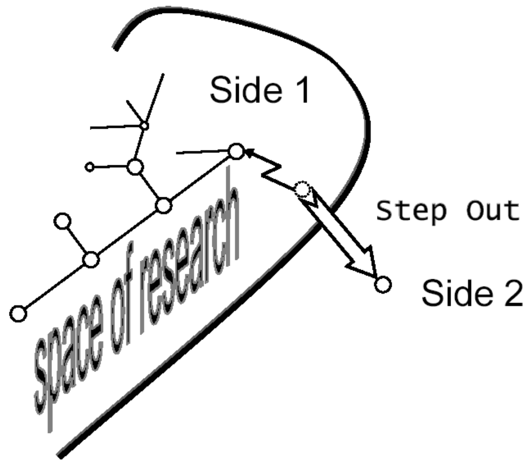


**Figure 10.** Conflict resolution by stepping out of the research space (postponed or resolved conflicts).

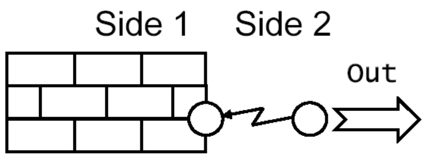


**Figure 11**. Automatically resolvable conflicts

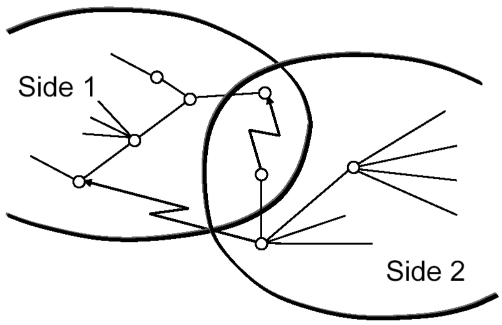


**Figure 12.**Conflicts resolvable using human-machine interaction

Knowledge bases (KBs) are improved after isolating and resolving contradictions in the following way. One set is replaced by another while other knowledge is supplemented or specified. The indicated processes are not directed by the elaborator or by the user. The system functions autonomously and it requires only a preliminary input of models and the periodical updates of strategies for resolving contradictions. Competitions to the stated method may be methods for machine supervised – or unsupervised – learning. During supervised learning, for example by using artificial neural networks, training is a long, complicated, and expensive process, and the results from the applications outside the investigated matter are unreliable. The 'blind' reproduction of teacher's actions is not effective and it has no good prospects except in cases when it is combined with other unsupervised methods. In cases of unsupervised training via artificial neural networks the system is overloaded by heuristic information and algorithms for processing heuristics, and it cannot be treated as autonomous. The presented method contains autonomous unsupervised learning based on the doubt-about-everything principle or on the doubt-about-a-subset-of-knowledge principle. The contradiction-detecting procedure can be resident; it is convenient to use computer resources except for peak hours of operation.

The unsupervised procedure consists of three basic steps. During the first step, the contradiction is detected using models from (2) to (4). During the second step, the contradiction is resolved using one of the resolution schemes presented above, depending on the type of conflict situation. As a result from the undertaken actions, after the second stage the set *K* is transformed into *K'* where it is possible to eliminate from *K* the subset of incorrect knowledge $W \subseteq K$, to correct the subset of knowledge with an incomplete description of the object domain $I \subseteq K$, to add a subset of new knowledge for specification $U \subseteq K$. The latter of cited subsets includes postponed problems, knowledge with a possible discrepancy of the expert estimates (problematic knowledge), and other knowledge for future research which is detected based on the heuristic information.

In cases of ontologies, metaknowledge or other sophisticated forms of management strategies, the elimination of knowledge and the completion of KBs becomes a non-trivial problem. For this reason the concepts of orchestration and choreography of ontologies are introduced in the Semantic Web and especially for WSMO [12, 13]. The elimination of at least one of the relations inside the knowledge can lead to discrepancies in one or in several subsets of knowledge in *K*. That is why after the presented second stage, and on the third stage, a check-up of relations is performed including elimination of modified knowledge and the new knowledge from subsets *W*, *N*, *I*, *U* are tested for non-discrepancies via an above described procedure. After the successful finish of the process a new set of knowledge *K'* is formed that is more qualitative than that in *K*; according to this criterion it is a result from a machine unsupervised learning managed by models of contradictions defined a priori and by the managing strategies with or without the use of metaknowledge.

## 2.4. KALEIDOSCOPE Method

The next SMM method is called KALEIDOSCOPE [6]. It is used to visualize the results. The difference from other analogical methods is the usage of cognitive approaches to human-machine communication. Similarly to the classical caleidoscope, the machine makes things that it performs better than humans: it arranges knowledge, it searches repeated elements and other regularities in the data, it presents the results to the humans. It is the user that makes the results creative and interesting for other researchers in the domain; the machine just assists him/her analogously to the caleidoscope which cannot estimate the beauty and the interest in the drawings produced during rotations and the arbitrary movements of the glass pieces inside of it. In both cases the human estimates the results and the developers apply methods to attract the user's attention and to make the work with the system equipped with elements from games, less boring and consequently more effective. The below application is from Number Theory, it is used to reveal the way of automatic detection of regularities, their research, automatic or manual, leads to very deep theoretic knowledge.

KALEIDOSCOPE consists of a visualization approach, a method for information transfer-by-sense and applied approaches to maintain a natural-style dialog. It is shown in [6] how different machine-done visualized patterns lead to student's perceptions that cannot be represented by the machine and when this is helpful during the educational process. Aiming to show domain independent approaches, the examples are introduced from Number Theory to language expressions and nonclassical logic applications. For example, on fig. 7 and 10 are shown multiplication cycles for numbers 9 and 11 in a set of 8 arithmetic series leaving after filtering out the first three prime numbers and their co-multipliers as well. The figures provoke questions why the images are repeated periodically across 11, or, respectively 19 rows, why their tour sequence is in a reverse order related to an imaginary center – the series $15+30k$ and many other questions the answers to which led us to interesting theorems from Number Theory together with cryptology applications. In the case of CALEIDOSCOPE the machine performs mechanical arrangement searches of repeated elements, etc. by presenting the results in a dialog with the human; it is expected that the human will see the unformal elements of the picture and will draw his/hers conclusions. The correct organization of the dialog and the presentation of the results is the essence of the CALEIDOSCOPE method.

The last one from the methods discussed here is the inference from contradictory information. This method is presented in details in [16]. On the first stage, based on the check for preliminary input knowledge as *models for contradictions,* the system determines whether there is a contradiction. If the sides of the contradiction are more than two, then the problem is reduced to the resolution of a set of 2-sided contradictions. Depending on the situation, the methods for the solution are automatic or interactive.

It is the inference method from contradictory information that allows to eliminate the incompletenesses from situations similar to the one depicted on fig. 13, and also to correct insignificant elements and/or to present the situation more clearly.

In this chapter basic methods of the synthetic meta-method SMM are presented. The next chapter is dedicated to the coordination of their operation under the control of SMM.

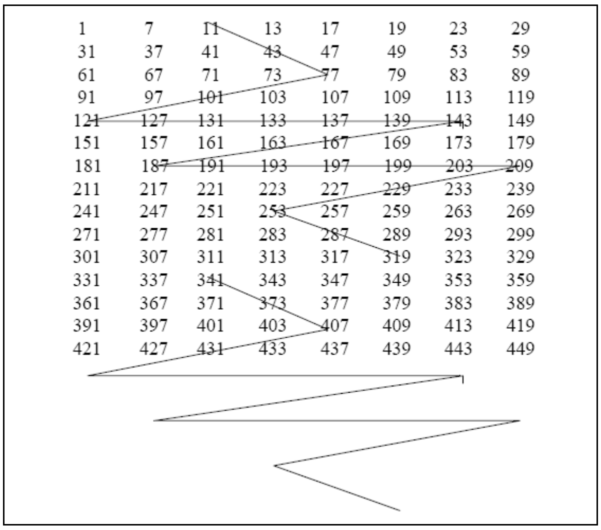


**Figure 13**. Cycle 11 in a set of 8 arithmetic progressions containing all primes ≥7

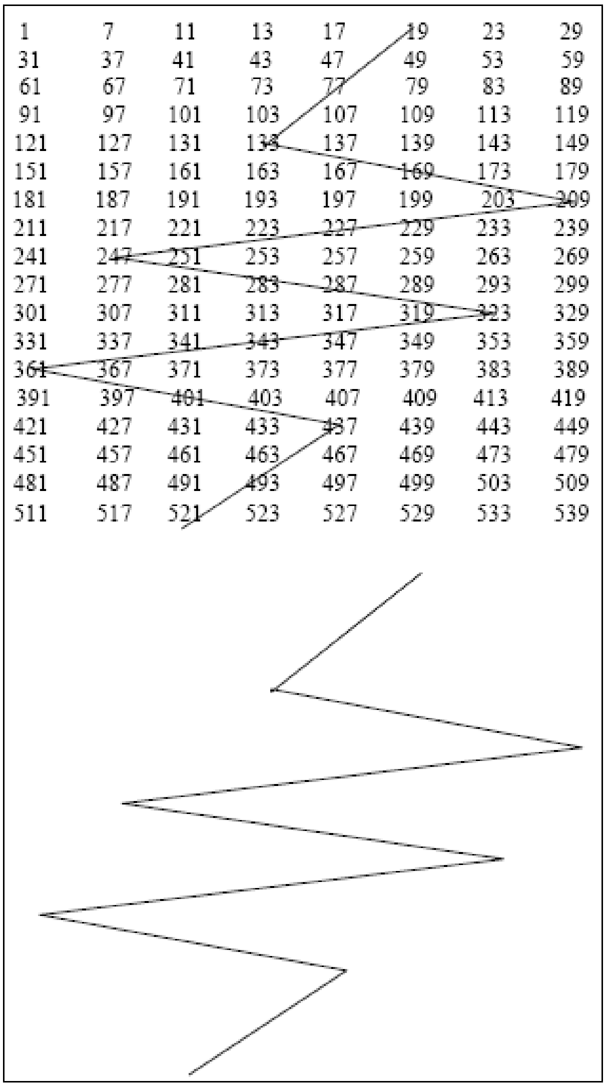


**Figure 14.** Repeating regularities similar to represented on Figure 13

## 3. Method coordination

Comparisons to well-known analogs may be given. Superiority of PUZZLE over classical case of constraint satisfaction, of CALEIDOSCOPE over cognitive graphics and other human-centered methods, of SMM/CONTRADICTION/FUNNEL applications over evolutionary fitness functions (variable or not, or a system of fitness functions), and of CONTRADICTION over conflict resolution methods may be discussed via comparisons of the above presented information to the well known cases. Altogether the advantages will allow construction of creative systems in future.

SMM is a synthetic method. Results from synthetic methods are more influenced by the coordination and the control among the components than by the components themselves. Usually the components of the synthetic methods are various algorithms from traditional analytic methods but we also investigated communications between several synthetic methods under the control of one synthetic meta-method. It is possible to make an analogy with the design of a display-window: if clothes are really good but the designer is not skillful enough, then the effect from the display-window dress make fail. The effect from an unqualified control in the domain of information security is even more significant and it leads to a complete crash.

It is recommended that at any time during the operations with SMM a direct control of the user's activities should be possible. For example, to detect the causes that led to a breakthrough in the system, the user inputs a {PUZZLE, FUNNEL, KALEIDOSCOPE, CONTRADICTION} and in this way he/she engages all resources to detect the cause. The cited sequence of commands is applied by default. In this case, every method operates independently, computer resources are allocated equally and it is the user intervention that may change the course of the investigation. It is provided that the results can be collected in two ways. The first way is the *blackboard* option. Here results are transferred between different methods and also to the user via a central node that is called the *blackboard*. In the second option methods from multiagent systems are used and results are transferred via the so called pheromones. This is information left by the agent which is kept intact for some time and which may be read only by agents that are close enough in space. The second option is especially good for transfer of operative information. Multiagent applications are discussed at the end of this section.

Another version for coordination is based on the use of statistical data and the personalization of information. For example, let's assume that statistics show that user X works most effectively with the PUZZLE method and also that he/she is interested least of all of the data visualization. In such case for him/her it is just the PUZZLE method that must be executed with the top priority, and the KALEIDOSCOPE must be activated only on special demands. On the other hand, as based on experts' estimates there must be a common system of priorities tied to the presented methods where the priorities are linked to the number of successive runs of the respective method. In this case the methods with higher priorities will occupy more computational resources.

The role of tests for contradiction(s) and the KALEIDOSCOPE method play a special role in the system. KALEIDOSCOPE is a visualization tool for results by constructing special modes for the human-machine contact. Therefore it must be activated after terminating the execution of any other method under the control of SMM. If the results do not require special technologies for their representation, then the method may be disabled manually. In the example presented below the set for coordination {FUNNEL-KALEIDOSCOPE} is used in a way shown on fig. 15 and 16. Let's assume that the goal function or the fitness function which are equivalent in the terms of evolutionary computation be marked by a central (red) line on fig. 15. Let's assume that the solution process of the defined problem is presented by a dotted trajectory on the figure. The intelligent data processing from the example is a type of approaches that are data driven. One of the properties of these approaches is the one that the solution may deviate from its way to the goal when lateral unexpected solutions are sufficiently interesting. A case with a deviation from the predicted trajectory is presented in the following way. Let the function G be denoted by the central (red) line on fig. 15.

The set of data accumulated in the system and knowledge M is filled via applying various logical inference rules, machine learning and/or other ways for knowledge acquisition. Each inference and any other change in the set M is traced via the gradient of information where the coordinate system is tied to the goal direction G.

The sum of gradients is shown on the figure via the dotted line. In the case when the change of M is inconsistent, for example the gradient calculus often change their direction because there are not enough significant intermediate results, then the dotted trajectory will change its direction when it reaches the funnel edge or, in other words, new knowledge will be searched only in the direction delimited by the position of the solution in the funnel which is presented for the point – an intermediate solution Y in (5):

$$A \le grad(Y) \le B. \quad (5)$$

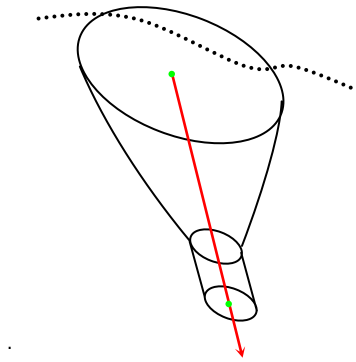

**Figure 15.** Funnel is not working

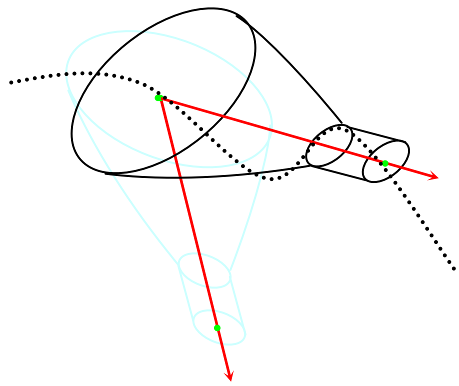

**Figure 16**. Funnel adjustments to work

The influence of G over the process of changes in M may be compared to the one of the gravity. Gravity is by no means the strongest force on Earth and we overcome it by walking, etc. On the other hand elevating (a motion along the antigradient) requires a lot of energy. If energy is not saved then the first cosmic speed is reached i.e. the earth attraction is completely overcome. Therefore limitations A and B from (5) are useless if in the flow – the dotted trajectory in fig. 15 there is much accumulated moment of inertia. There is an equivalent of the inertia in the case linked to the dotted trajectory in fig. 15. The following relations are introduced in relation to the problem defined. Let $s_i$ be an intermediate solution that is interpreted as a point from the trajectory and then by analogy with (5) the following constraints (6) are introduced:

$$A \le \sum_{j=1}^{i-1} grad\left(s_j\right) \le B. \quad (6)$$

This is based on the following explanations. Gradients were the necessary interval between A and B for steps from the first up to $i-1$-th. Then, the bigger $i-1$ is, the bigger the inertial property ad hoc the solution $s_i$ will be. When *i* is greater than an a priori threshold value then the constraints of the FUNNEL change their effect.

If the sum of gradients is bigger than an a priori set threshold value $T_1$, then the cited limitations may be overcome (7):

$$\sum_{j=1}^{i} grad\left(s_j\right) \ge T_1. \quad (7)$$

The next version includes a new constant $T_2$ and a solution $s^*$ which is different from the defined goal G but which is marked by experts as sufficiently interesting. It can also be marked by the machine because for example, it is the missing link to solve other earlier defined problems. $s^*$ does not participate in the series of $s_i$ but if the distance between them is calculated via an additional p-adic metrics, then it is possible to avoid the limitations given by the FUNNEL:

$$\left\| s^* - s_i \right\| \le T_2. \quad (8)$$

In the three cases, cited above the 'inertia stream' pierces the FUNNEL under conditions given in the formulas as it is shown in fig. 15. Otherwise, all consecutive solutions will be inside the FUNNEL as it is presented in the description of the method above. In the case of a break, the FUNNEL stops running. Therefore, if the FUNNEL method does not adapt to the new conditions of the information flow as it is shown on fig. 16, then the method ceases to perform its main function and starts to gradually narrow the limitations to direct the set of dynamically changing information to the required place. During the tuning from fig. 16, this problem does not exist. On the other hand, by solving the problem from fig. 15 and 16 we had to construct more than one system of fitness functions where the first (classical) function is directed downwards and the second one operated temporarily only between the beginning and the end of the depicted FUNNEL. In this way, conflicts are avoided between classical evolutionary methods and the requirements of contemporary intelligent data processing.

## 4. Ontologies

Intelligent data processing by itself is a complex process of transforming data into knowledge. If data and the knowledge accumulated in the system are not structured in advance, then in the majority of cases their processing will be ineffective or it will produce poor results. For this reason, ontologies are used [11]. The tool to create and manage ontologies is Protégé OWL [14]. In the majority of cases ontologies are presented as graphs. For example, the fragment of ontology from fig. 17 illustrates the types of ISS: intrusion detection systems, intrusion prevention systems and others. The presented ontologies give an idea about the links between the objects and the relations between them.

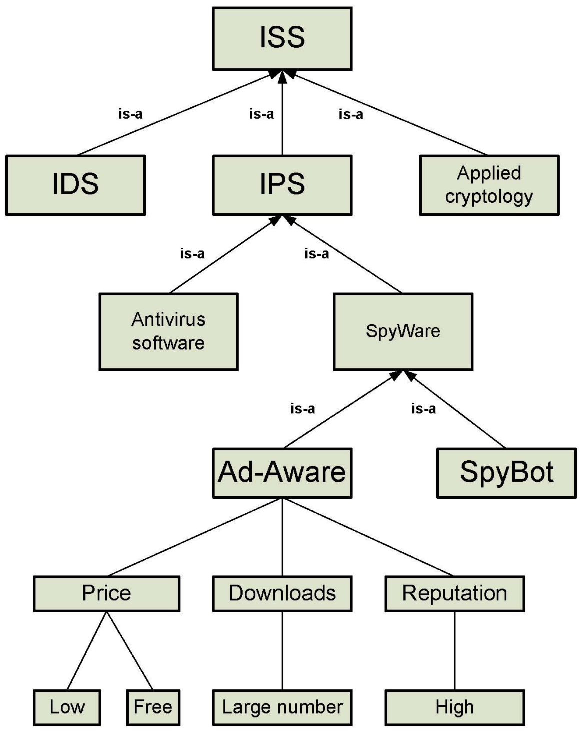


**Figure 17.** An example of ISS ontology

The relations on fig. 17 use the standard type 'is-a', and the taxonomy of the 'ancestor-successor' type is built based on them. The presented graphs are widely used to search information on the Internet, e.g. if the keywords are ISS then there is no need to write 'anomaly detection' because the machine will make the necessary links between the concepts. Another widely used relation is 'has-a' meaning 'the object possesses the properties …'.

Examples of these relations are the meanings of the attributes in the lower part of the figure, e.g. 'low cost'. Based on these two relations there are many applications of descriptive logics for the needs of the Semantic Web [13] and others. Recently various other relations are introduced in the KAON projects and other initiatives [15] that include operations with quantifiers and other logical operations.

The proposal to structure information is as follows. We propose the usage of tools close to the so called means-ends analysis. For this reason we introduce relations like WHY, HOW, WHAT FOR. In this way, new possibilities are formed to operate with knowledge that can be applied by traditional tools. For example, let's propose an input in the system with a large number of differential equations. Let's have a relation 'WHAT FOR' applicable to this block: 'descriptions of trajectories during signal processing'. Using the relation WHY, we shall determine that 'this way is more effective than traditional methods because …'; via the relation ORIGIN we shall be able to compare the offered fragment of knowledge with other popular methods, tools and approaches and so to discover the reasons for the problem. Applying the relation 'WHAT IS THE WAY', we shall discover texts with explanations how to use the differential method and other technical information. Using the enumerated tools, there will be a better structuring of information and prerequisites will be created for its more qualitative processing: in this way, we shall be to able more effectively operate with large blocks of formalized knowledge without going in technical details.

The usage of ontologies is not only building large sets of them. In the Web services, it goes through application of tools of the W3C consortium like WSMO and others [12]. The technology of these applications is more complicated. It includes specific concepts like orchestration and choreography. However, this is the way to apply a contemporary ontology processing - addition, deletion or update without breaking the links with other knowledge that is input in the system. In this way, the problems connected to the semantics of the problem are taken under consideration. Any other knowledge may change its meaning depending on the context, the point of view, etc. The version of Protégé-OWL is used in the creation of ontologies.

The necessity to introduce the following new relations in ontologies for intelligent applications of security is demonstrated. For example there is a large number of differential equations introduced in the system. A new relation '**what for**' is introduced for this block: 'to describe trajectories for signal processing'. The relation '**why**' determines that 'this way is more effective than traditional methods because …'; the relation '**origin**' will be able to compare the objective fragment of knowledge with other known methods, tools and approaches to determine the cause for the problem; the relation '**in what way**' will assist during localizations of explanatory texts about the usage of the concrete differential method and other information. It will be possible to structuralize information in a better way in the presented tools, and prerequisites will be created for its more qualitative processing. So we may operate with large blocks of formalized knowledge more effectively without deepening in technical details; at the same time we shall avoid much too detailed formalizations, i.e. here ontologies contribute to use good practice in educational processes. The more relations are examined, the more precise categorization of objects will be. Substantially all other used for the moment relations just illustrate this: in what way the relation arguments are connected , otherwise '$\mathbf{R}_i$ links the arguments (the objects) from A to N in a way determined in the definition':

$$\mathbf{R}_i\ (A, B, C, D \ldots N). \qquad (9)$$

The set of $\mathbf{R}_i$ is used to determine that 'A is connected with C', etc. and the way of connecting is set in the definition of $\mathbf{R}_i$. It is suggested to introduce such $R_i^*(A, B, C, D \ldots Z)$, the usage of which will form automatically answers not only of the type 'B is connected with T' but **why** 'B is connected with T', **in what way**, **when** the pointed link is actual, etc. Innovations do not deny previous ontological elaborations, they just supplement them; so the sense of matters is shown much more clearly, unambiguously that can be changed by the context. Also another ontology is introduced: '**important**'. At the first sight this type of information is subjective: the importance of knowledge is determined by the situation, the object, etc. But if information is not graded by importance then the agents will ‘drown’ in this sea of information provided the knowledge is not well structured; in the case structuring is performed best of all via the introduced new relation ‘important’.

Same as proposed ontologies may be used in the opposite direction of human-computer interaction: **what** is the reason to the question, in **what way** it is connected to the investigated goal G, why the user thinks the answer is **important**, etc. The new relations improve the human-machine interations and lead to human-machine brainstorming methods [6], one of the best ways to apply creative systems.

Part of ontologies may be succesfully shifted by PUZZLE and/or other applications of methods from Section 2. In the represented research is introduced that in some cases the graphic itself of the partly received resolve from figure 18 is enough to take decision from the side of the consumer or the expert.

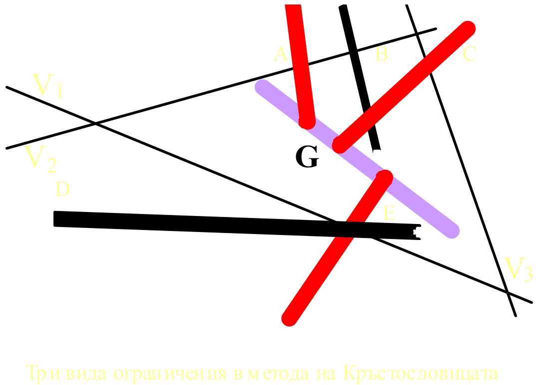


**Figure 18.** *Graphic presentation of the linguistic material*

The different variants of the modeling knowledge, consisted in the figure 18, represent resources for presentation of the things which are difficult to discover verbally. It is not unusual for the Chinese to use the aphorism “a picture replaces thousands words”. Here should to be added: “…and it is understood and remembered better”. Below are given different linguistic equivalents of the combinations of interpretations similar to that one from the figure 18 and received through the proposed and graphically described method of the puzzle.

- The things are moving to the way that…
- It is neither…nor…but it is too close to… (Without dilutions!)
- Look from the other angle … Other conclusions…
- The object can be represented as crossroad of the two dynamically developed areas…
- In the future these two concepts have to be used together…

The pointed naturally linguistic fragments can be represented and in direct text to the students but below is shown that trough the method of the puzzle this can be done significantly better. Modeling through the use of the method of the puzzle helps for the emotional contact with the audience. For example on the fig. 19 and fig. 20 is represented other example for the use of the method of the puzzle when are clarified particular processes and connections between the objects.

On the fig. 20 the system from different types of limitations, described in the second chapter, helps for discovery of the casual connection between M and N at the same time between them not existing any connection on the fig. 19. The connection is displayed in the impose of additional limitations on the fig. 20. The process is dynamic as nature and it is almost impossible to be correctly transferred by words only.

## 5. Other applications

The presented system source codes are written in different languages: C++, VB, and Prolog. It is convenient to use the applications in freeware like RDF, OWL, Ontoclean or Protégé. Many of the described procedures rely on the usage of different models/ ontologies in addition to the domain knowledge thus the latter are metaknowledge forms. In knowledge-poor environment the human-machine interactions have a great role, and the metaknowledge helps make the dialog more effective and less boring to the human. The dialog forms are divided in 5 categories from 1='informative' to 5='silent' system. Knowledge and metaknowledge fusions are always documented: where the knowledge comes from, etc. This is the main presented principle: every part of knowledge is useful and if the system is well organized, it will help us resolve some difficult situations.

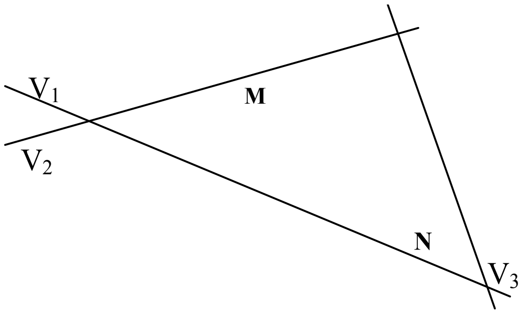


**Figure 19.** *M and N are not connected but located using linear contraints*

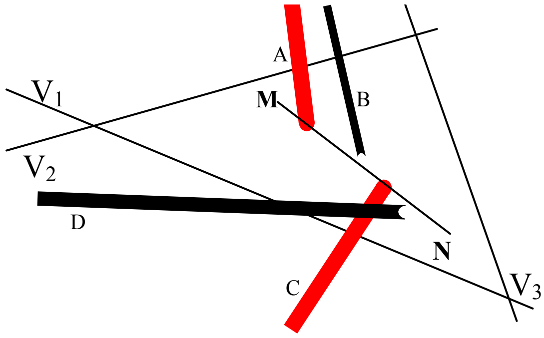


**Figure 20.** *Discovery of the connection between two objects M and N*

We rely on nonsymmetrical reply 'surprise and win', on the usage of unknown codes in combination with well known methods, and on the high speed of automatic reply in some simple cases e.g. to halt the network connection when the attack is detected. If any part of ISS is infected or changed aiming at reverse engineering or other goals, then the system will automatically erase itself and in some evident cracking cases a harmful reply will follow. The above presented models of users and environment are used in the case. Therefore different SMM realizations are not named IDS but ISS because they include some limited automatic reply to illegal activities.

Conflict situations occur when the agent can't execute its goal in an acceptable time or in other situations. Say, the agent's way to the goal is interrupted by a route for many other agents. The standard way frequently quoted in multiagent systems is 'step out' for all conflicting agents. Then another routes should be found, otherwise same conflicts will repetitively occur. Meanwhile some goals may be lost. In contemporary dynamic environments consisting of hundreds and thousands of agents the old-fashioned 'step out' strategy leads to creation of new conflicts rather than to their resolution. On the other hand, one or more of conflict resolution methods considered above resolve the problem using well known class Agent methods halt(); suspend(); activate(); getState(); move(Location where); and sometimes clone(Location where, java.lang.String newName). Several libraries for class Agent had been used. All of them comply to FIPA standards.

Upon the formulation of goal *G*, few variants of control strategies could be applied. Most of the presented methods are data driven, and the user may trace the intermediate results and interrupt or change the resolution process if necessary. Also he may apply his personal plans because the goal is seldom formulated without any plan or rough idea how to resolve it. It is possible to write the plan like 'use the FUNNEL method in direction *X*, then use PUZZLE and in the end KALEIDOSCOPE to represent the result'. This example of methods interaction is interpreted in Fig. 4 where the solution is located in spot E, the dashed lines are linear constraints, and the spotted lines give us the parts of the solution as described in the PUZZLE section above. If we compare the results represented on figures from Section 2, we may conclude that PUZZLE method works better in combination with FUNNEL because we need less constraints using the fitness function f, and narrow search in FUNNEL helps to locate the goals. The same conclusion is to be applied to combinations with other existing methods from the scope: combining any of the above methods with other statistical or data mining methods is a way to better applications. The only obstacle here is the overall complexity of the obtained combinations. Almost the same result could be reached under SMM automatic control. The FUNNEL usage narrows the resolution process direction but in the beginning it doesn't exclude unpredictable results, hence the best possibilities for data-driven methods are obtained. If no interesting results have been obtained, then PUZZLE is activated to bind the goal to knowledge related to it.

## 6. Conclusions

The main conclusion is that to overcome the shortcomings of contemporary creative-related methods and systems, methods and applications are considered concerning the logical parts of knowledge discovery and/or human-machine creation of a brainstorming type. Special attention is paid to methods for identification and resolution of conflicts, and to machine (self-) learning based on them. The role of the above methods to the creation of more autonomous agents is discussed.

Analysis is represented for methods used for machine learning in intelligent agents, for sending information by sense, and for understanding the semantics of the information, all the parts essential to every creative system. Common disadvantages for different existing groups of contemporary applications are revealed. Ways for method coordination under a synthetic method SMM are discussed.

## References

[1] M. Boden. The Creative Mind: Myths and Mechanisms. London: Routledge, Taylor and Francis Group, 2004.

[2] G. Wiggins. Creative Systems. A preliminary framework for description, analysis and comparison of creative systems. Knowledge-Based Systems, Vol. 19, No. 7, November 2006, p. 449-458.

[3] V. Jotsov. Intelligent Information Security Systems. Sofia: Za bukvite-o pismeneh, 2010, 211 p.

[4] B. Thuraisingham. Data Mining Technologies, Techniques, Instruments and Trends. NY etc.: CRC Press, 1999.

[5] D. Dasgupta. Artificial Immune Systems. Moscow: Fizmatlit, 2006.

[6] V. Jotsov. Novel Intrusion Prevention and Detection Systems. *In: Proc. 4th International IEEE Conference on Intelligent Systems*, Yager R., Sgurev V and Jotsov V. (Eds.), Vol. II, Varna, Bulgaria, September 6-8, 2008, pp. 14.20-14.27.

[7] V. Jotsov. "Evolutionary parallels," *In: Proc. First Int. IEEE Symp. 'Intelligent Systems'*, T. Samad and V. Sgurev (Eds.), Varna, Bulgaria, vol. **1**, pp. 194-201, 2002.

[8] A. Arruda, "A survey on paraconsistent logic," *in: Math. Logic in Latin America*, A. Arruda, C. Chiaqui, N. Da Costa, Eds. North-Holland, Berlin NY, pp. 1-41, 1982.

[9] V. Jotsov. Semantic Conflict Resolution Using Ontologies, Proc. 2nd Intl. Conference on System Analysis and Information Technologies, SAIT 2007, RAS, Obninsk, September 11-14, 2007, vol. 1, pp. 83-88.

[10] V. Jotsov. "Knowledge acquisition during the integer models investigation," *Proc. XXXV Int.Conf. "Communication, Electronic and Computer Systems"*, Tecnical University of Sofia, pp. 125-130, 2000.

[11] T. Gruber. A translation approach to portable ontologies. *Knowledge Acquisition*, Vol. 5, No. 2, pp.199-220, 1993.

[12] Web Service Modeling Ontology *(WSMO).* http://www.wsmo.org/TR to date.

[13] D. Fensel. Ontologies: A Silver Bullet for Knowledge Management and Electronic Systems. Berlin Heiderberg New York: Springer-Varlag, 2004.

[14] The Protégé Ontology Editor and Knowledge Acquisition System. http://protege.stanford.edu to date

[15] Welcome to KAON — KAON - The KArlsruhe ONtology and Semantic Web Open Source Management Infrastructure. http://kaon.semanticweb.org/ to date

[16] V. Jotsov. Emotion-Aware Education and Research Systems. J. Issues in Informing Science and Information Technologies, USA, vol. 6, pp. 779-794, 2009 (online version available).

## Author information

**Vladimir S. Jotsov (В.С. Йоцов):** e-mail: bgimcssmc@gmail.com

State University of Library Studies and Information Technologies (SULSIT);

Intsitute of Information Communication Technologies of the Bulgarian Academy of Sciences (IICT-BAS);

P. O. Box 161, Sofia 1113, BULGARIA, EU;