APPLICATIONS OF KNOWLEDGE DISCOVERY METHODS IN EDUCATION

Vladimir Jotsov, Stefan Kojnov

Abstract: *The topic of the presented investigation is contemporary education threats that will lead to big problems in the nearest future. It is shown that the prevention of such threats is impossible without applications of intelligent technologies .Even more, the contemporary education system should possess some features of inteligent technologies, smart systems and data mining elements. Original results are presented on how to build more effective applications of classical and/or presented novel methods: kaleidoscope, puzzle, contradiction, etc. It is shown that all the demonstrated advantages may be successfully combined with other known methods from e-learning, m-learning or other advanced education fields. It is shown how the stated applications enhance the quality of teaching, improve the emotional connection between the teacher and the students and ultimately serve the sustainable development of the students. It has been proved that qualitative forms of education do not require the usage of expensive hardware.*

Key words: *education, crisis, information security, national security, intelligent system, agent, knowledge discovery, data mining, web mining, ontology, human-centered systems, knowledge management, automation of creative processes.*

1. Introduction

Contemporary Information Security Systems (ISS) and especially their web-based versions represent a wide field for applications of modern methods and technologies. Contemporary training of intelligent agents reaches such a high level that some elements of machine learning became practicable also in contemporary education for such rapidly changing subjects as Information Security (IS). IS is an example for a single fast changing trend the qualitative training of which is neither easy nor cheap. The following sections propose a number of original methods not only for pupils but also for informatics students, particularly in IS. The research is based on a series of methods for intelligent knowledge/data processing elaborated within ULSIT-Sofia. The versatility of the proposed approaches is demonstrated also by their successful application for training soft agents or holohy-type robots in the area of IS. The presented series of experimental results showed accumulated problems that require the application of sophisticated modern intelligent methods for educational purposes: tendency to memorize rather than thinking, we name it shallow learning, which starts from small classes of school education, the lack of attention due to the use of computer games and the too active use of Web forums and resources for communication, the desire to receive high marks without a deep understanding of the presented subject, etc. These accumulated problems can not be solved using traditional teaching methods.

Also these problems cannot be solved using a single method but rather a series of methods under the control of a synthetic metamethod. In the ongoing crisis in the country we cannot allow the use of expensive technological solutions such as using smartphones to connect directly with the course data and video images of lecture material, G4 methods for m-learning, etc. [1-10]. Even if a project allows the installation of corresponding tools, policies and practices, it is not practical to hire teams to support these latest systems because lecturers alone can not make all the necessary flash animations, implement and maintain mobile connectivity and other related things. Under these conditions we developed a relatively inexpensive technology for modern teaching based on a minimum of technological solutions that do not require high costs of purchasing and support [11].

2. Qualitative Education during Stagnation Times

This research develops the thesis that all the best modern e-learning systems have in common with research systems and vice versa. The considered methods include a number of elements for applications in research systems. It is shown that almost the same method applied in other way gives new results for applications in IS training. On the other hand, teaching IS stands out from other undergraduate and graduate programs that it operates in rapidly changing models, requirements, concepts, tools, instruments and policies where the problem for qualitative education is particularly difficult. On the other hand, teaching IS highlights in the brightest way the advantages of the proposed new methods and applications in ISS. The following applications are presented for e-learning in IS. For the sake of better visualization of the presented material, the kaleidoscope method is described in the next section.

One of the most instructive facts about the audience for training is that current neural-based machines in the form of software agents/holons repeat the actions of the teacher on a much higher level, even related to an excellent student. Therefore, rote memorization of lectures makes no sense: sooner or later machines will displace us from such activities; instead applied science of matter should be deeply rationalized through comparisons with other works accessible through the Internet, by experimenting with numerous courses and with open source products which become more pervasive in our practice together with contemporary Web-applications. The example shown here illustrates how to achieve simultaneously two goals: explain artificial neural networks and modern principles of quality training through effective and deep understanding. Similarly, using neural networks explains best practices for control knowledge, for more drills for lagging students, etc. The application of the described methods makes the lecture process more easy, understandable and transparent. Hence sustainable education results are obtained.

On the other hand, monitoring the quality of education most often uses different statistical methods. We used the same in a general data mining module together with logic applications described in this paper and the results of logical and statistical parts come together on an evolutionary basis. 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 resolve other logic problems 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 metamethod 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 [12]. 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 uses SMM (Synthetic MetaMethod) metamethod to control the process of consecutive replacement of applications by other ones. 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. In general, the use of SMM is one of the most important and demanding elements of the efficiency of intelligent applications. But in the educational field combining the proposed methods depends on the lecturer's experience, which makes the performance much simpler than just in the automatic case. At this level, the operation of the data mining module is limited to implementation-voluntary advices to the lecturer. Suggested innovations serve the more effective application of education practice altogether with data mining, Web mining, collective evolutionary components in multiagent systems [13, 14].

3. Kaleidoscope: Graphical Interpretations Highlighting Cognitive Elements of the Presentation

Essentially the method of kaleidoscope of [11] consists in a form of visualization, which leads to a creative work in various forms of training or for research purposes. We would like to remind that learning inevitably leads to independent research. In the present study it is shown that in some cases the graph itself of the obtained partial solution from Fig. 1 is sufficient for decision-making by the user or the expert.

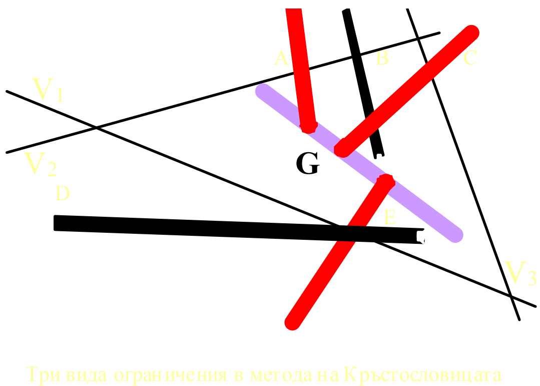

Figure 1. Graphical representation of linguistic material

Different versions of the modeled knowledge contained in Fig. 1 provide a base to describe things that are difficult to express in words. Not surprisingly, a Chinese aphorism states ‘A picture replaces thousand words.’ It should be added: ‘... and is understood and memorized better.’ Below are given different linguistic equivalents of combinations of interpretations, similar to that of Fig. 1 and obtained by puzzle method.

- Things are moving way that …
- ... neither ... nor ... but it is too close to ...
- Looked at from another angle ... Other findings ...
- An object can be represented as intersection of two dynamic areas ...
- In the future, these two concepts should be used jointly ...

etc.

These natural-language fragments can be presented in plain text for students, but further it is shown that by the puzzle method this is done much better. For example, Fig. 2 and 3 show another example of puzzle method clarifying certain processes and relationships between objects. Concentrating the attention in the area of Fig. 2 supports by narrowing the set of analyzed elements, the opening of implicative relation referred to in Fig. 3. Also in the case of effective learning using informal causal relations that are associated with that implication ‘from M it follows that N’.

In Fig. 3 the system of different limitations helps detect a causal relationship between M and N, with no correlation between them in Fig. 2. The connection occurs in imposing additional constraints in Fig. 3. The process is dynamic in nature and it is almost impossible to be properly explained with words.

This image (or media) is so much more convenient and efficient to visualize, transform and use different dynamic processes through ontologies. Examples and figures can lead us to the conclusion that ontologies are introduced, used and dynamically changed applying the methods of the puzzle or of the kaleidoscope.

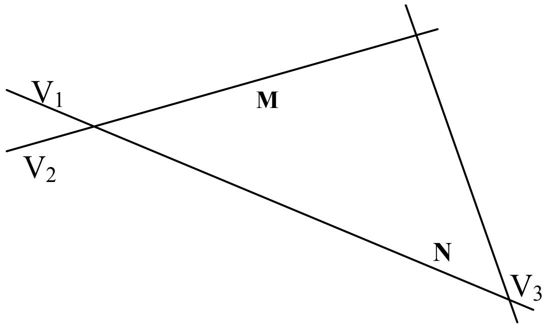

Fig. 2. M and N are unrelated

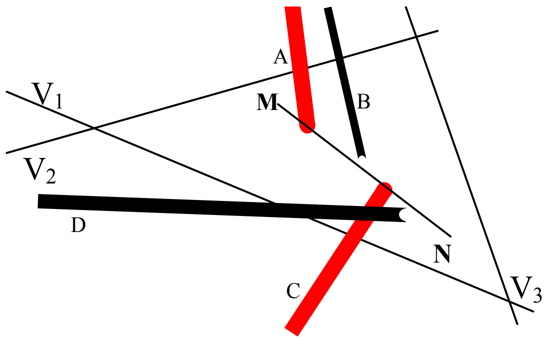

Fig. 3. Detecting a connection M–N

Unlike oral presentation, the material visualized by the puzzle method gives more directions for own research and for detecting and correcting gaps in knowledge. For example, 'it is neither ... nor ...' and many other natural-language interpretations incorporate a lot of fuzziness that can sometimes be misunderstood. Careful study of graphs similar to Fig. 3 excludes misunderstanding and furthermore – a misunderstanding of the material. On the other hand, the whole learning process is concentrated in one place, it is not necessary to look for other Internet sources to clarify the question e.g. **'why', 'what', 'how'** to get results, etc. The quoted relations also have been used for pointing the lecture main points. Intelligent training tools not only accelerate but also intensify the training.

Unlike IS, a wide range of tools for modeling and presentation of material is used in this section. The modeling must be used not for ISS software agents but for people. Respectively it is easier to introduce audio or multimedia ontologies, transmitting the meaning of things. For example, in Fig. 4 the answer to the question 'How the passive defense works' is given at a schematic level. To present the meaning of things it is enough to show one or several key points on the subject altogether with corresponding explanations. Sometimes it is better than an entire movie. The relevant ontology is not limited by the photo which has just a specifying role. More important are the descriptions to a picture to describe what the picture is essentially, in detail. The picture explains many details clear to people, but intelligent agents will learn nothing from this picture, especially if the agents are designed for IS.

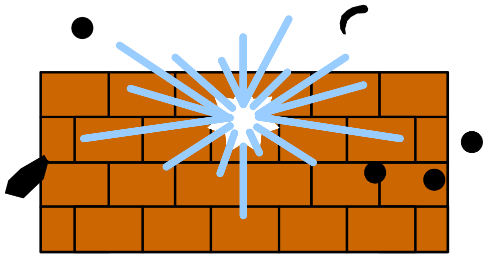

Figure 4. Passive defense

Presenting information by meaning is a key element of modern higher and university education. The proposed methods of the puzzle, of the kaleidoscope and the other tested methods greatly expand opportunities for qualitative teaching IS, one of dynamic interdisciplinary specialties with increased complexity and also of various types of requirements for trained students.

4. Applications of Puzzle Methods

Another original method discovers new or hidden knowledge by connecting the unknown, the sought solutions with previous experience in the form of knowledge bases. Let the constraints of the defined problem form a curve in the space as depicted in Fig. 1 – 3. The main goal of the puzzle method is to reduce the multidimensional search space for the solution. For this purpose we used several limitations, in [11] the study is focused at the case of using ontologies instead of individual limitations. Furthermore, the research of process dynamics aimed at reducing the field in certain cases allows us to derive new knowledge in the form of rules. This inference process and using constraints allows us to significantly simplify formal and evidence material in lectures, to attract the attention to details and simultaneously to increase the activity of learners. Showing the process of connecting known with unknown improves understanding and retention of the presented material.

For example, if a bachelor who has graduated ULSIT 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 instead of curves with complicated forms obtained as a result of crossing surfaces. Therefore, below we investigate the usage in systems of constraints by lines of first or higher orders.

The following example below shows how the search process is reduced using ontologies. Let's admit that the search space is presented on fig. 5 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. A right, blue-coloured 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. 5 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. There are two red dots on the same fig. 5 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.

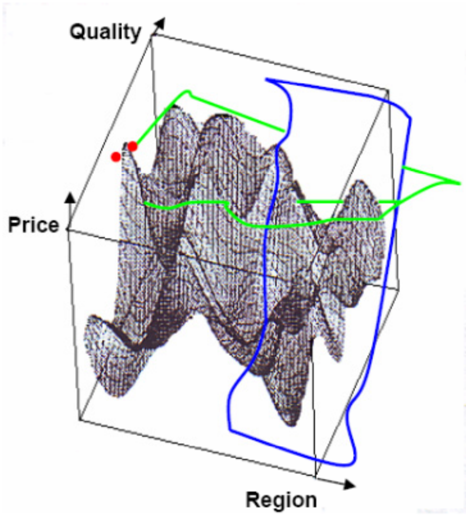

Figure 5. Binding and other constraints

By introducing new constraints, our goal is to show that it is possible to use causal links that are different from implications.

5. Control of Emotions

There are many analogies between intelligent technology and IS training [15]. With no specific game elements, if the information is presented in a monotonous manner, then it is possible to provoke negative consequences and so affect the quality of education. This ‘game element’ means the number of applications, starting from conflict resolution and contradictions using the apparatus of game theory and/or mathematical logic to applications used in computer games. The application of this aims at establishing an emotional contact between teacher and student, and this must abandon outdated notions that emotional teaching scientific disciplines is an obstacle, not a contribution. The term ‘emotional teaching’ no longer means that it is necessary to ‘read’ enthusiastically, with emotion. Surely this will not appeal to everyone in the audience. The creation of comfortable conditions for the system or the students will produce the necessary emotions.

There are two main approaches to winning contemporary education in IS. First, it is necessary to draw the attention of the student to the teaching material, for example by introducing a competitive element (which will first solve the problem), through a system of bonuses and at the same time – the introduction of additional work for lagging students (topics for further study, additional tasks, etc.). Those elements of the training system should unite in a clear and explained in advance system in which excellent performing students are being tested for a much shorter time than laggards. One of the goals is to avoid monotony in teaching modern dynamic disciplines, where the lack of attention can lead to misunderstanding or misconception of the material. Second, the gained attention of the students should be held. Without the introduction of certain emotional elements in the lecture it is difficult to attract attention and it is practically impossible to keep long-term.

The inclusion of emotional components in the presentation is not enough for modern teaching information and communication technology (ICT). In the direction emotion-aware systems it is widely discussed the application of measuring the emotions of the personalization data for making individual training scenarios, modeling every person working with the system. These funds shall not apply in the study because they require greater financial resources (and therefore apply mainly expensive developments in ambient intelligence) and are not effective:

response and appearance of people often lie about their emotional state; models allow for individualized learning, but their development is a slow and costly process. It is therefore proposed not to measure positive emotions but to produce them under the right conditions. To create conditions in the first place it is necessary to exclude factors that prevent them:

a) Complex and voluminous material for which there are not enough hours. If the material is cut, then should be taught not different methods, but large groups of methods or strategies and there should be required additional study of the material of substance. This prevents a proper understanding of the subject by students which copy texts from web in the best case.

b) To understand complex material with lots of formal descriptions or with many included fragments of programs. Formal or program descriptions provide a detailed understanding of IS and other disciplines in ICT. On the other hand, understanding formal descriptions takes time, which can 'eat' the majority of the lectures.

c) It is not easy to teach dynamically changing material with an interdisciplinary focus on people who are accustomed to traditional teaching methods, algorithms, data sources and references for briefly studying with the requirement that everything must be written in Bulgarian. In this we encounter with the collective mindset to do it the fastest way (hence – badly), for example the decisions for homework problems can be found in 'manuals' or in other Internet sources. Independent copying from the Web space is mistakenly considered as an ownership, especially if a translation from a foreign language is performed. When the methods are based on data then they are not largely algorithmic, they cannot be briefly described and fully understood without additional programming skills; the resources are in English and most of them are from the Internet, so teaching of IS using traditional tools leads to misunderstanding and low effectiveness.

d) Emotional contact cannot be built without a clear description of things, without interest and bilateral trust. Here the following thesis is used:

What is more convincing, better structured, clearer – it is emotional (i.e. it produces positive emotions in students).

The following discussion is an example of a situation where you need to clarify the term 'confusion of thought.' To illustrate it sufficiently complete including the evolution of its meaning in different situations requires a certain creative act on the part of the teacher. Fig. 6 presents the necessary illustration of the 'confusion of thought'. Whether intentionally mess banknotes in Bernanke style or they were badly packed, or the packaging is opened by accident inside the cabin, there will be different explanations depending on the situation and the explanations will even relate to different subject areas – from economy to security. One of the open questions is whether the helicopter is military accidentally or it is specifically used for security purposes. But good visualization combines all these situations in a concept – confusion of thought with its emotional load. It is a big problem with intelligent machines to get the least basic meaning and context of the shown situation. Thus, the example shows how the study of unresolved ISS problems supports school/university education.

On the other hand, if the lecturer wants to convey not so much the meaning as the emotional coloration of the situation, so far in training intelligent agents, he/she is even less likely to succeed compared to described in Fig. 6. For example, Fig. 7 depicts the situation of an accident with a car. The emotions of the people inside the car are evident and they easily affect trained people, indicating that what is very useful in the lecture hall is now useless in advanced IT.

Figure 6. Typical example for 'badly thought'

Figure 7. Emergency situation

But **how** to describe clearly intelligent data processing, serious statement which even experts in the field perceived as complicated enough? To solve the problems presented without an expensive technical devices starting from 'breaking the ice' by explaining that more complicated matter leads to higher valued knowledge, by surprise, 'awards', a change of the environment, etc. For example, it is better to use the idea of the leadership of ULSIT, where there are lectures on oral explanation, to use comfortable armchairs for outdoor classes. The main objective of these measures is some of the best and most active students to participate and help other students. Counter questions are promoted and also comments from students if they have a direct connection with the material. For example, you can use a system of bonuses from +0.25 up to +1 as score to evaluate the following self-tests for a given topic, task solved on the spot, etc. Experience shows that even the smallest bonus then stimulates less organized students to devote more time to the subject, because they are convinced that they are better than others by something. Organisational arrangements for quality teaching must be numerous: checking attendance, interim control knowledge, grouping students according to the results, varied approach to different groups, etc. Sometimes are important also least important details such as teaching easier to understand material in weeks with most other control workload from other disciplines.

In general teaching requirements of ISS are contradictory. In manual applications to the problem of teaching proposed methods solving the controversy, the obtained results are described in [11].

The problem of the time shortage for lessons both inside and outside the audience halls is solved by introducing emotional elements to reduce tension during the more intensive work and also by issuing more textbooks and school supplies thanks to the support from management. Along with this more complex material should be presented in a more understandable way, preferably through informal means and explanations. For example, the concept of a prime number, which is widely used for protection of information, is inevitably linked with the concept of infinity. In the above illustrated method of kaleidoscope the two concepts are linked through graphical interpretation. Graphics and the issues they raise, are relatively easy to connect to the evidence, and without them, this material would become 'tedious grind material.' Graphical interpretations are more easily stored, easier to explain, they give more possibilities to continue separately the study and thus to compensate for the lack of time to explain. It is good formal results only illustrate, to accompany graphs showing the path of their receipt and not their memory is the main objective of the course. Thus, teaching is clear and unambiguous. To avoid the opposite, the presented results are accompanied by full details of **who**, **when**, **where**, under **what** conditions has developed the method or the application. It is not necessarily all information to be presented in one place, it is important to have it somewhere, e.g. in a textbook. For the same reasons for recording knowledge in KB it must be recorded this kind of meta-information, it helps to resolve the conflict between the confluence of knowledge information created by many experts in other situations. The visual presentation of the material is intended to provoke lateral thinking in a large group of students interested in the subject. Those are also emotional elements of teaching and the use of ontologies leads to more in-depth study of the subject.

Available students (people or machines) are offered to use another new relationship – '**important**' – for structuring the material in varying degrees of importance. For example, when describing applications of radar to protect airports, relations '**why**' and '**important**' should be directed to the following: Doppler radar is significantly more expensive than the usual one, but it not only allows to show the location of monitoring sites but also their speed and one of its main functions is now detecting micro hurricanes and the movement of winds in the storm. Naturally, the detailed technical descriptions of the radar are also important, but they do not give new information about the meaning of the studied objects and about their use, therefore they are not associated with the relation '**important**'. If an agent, without understanding the meaning of things, depending on the situation transfers to a human texts on relations with '**why**', '**background**', '**how**' or, for example, '**mportant**', it will complete the task of proper information and there will be no need for further actions. In a scenario with two software agents, this technology will not work without additional actions.

When it comes to emotions, everyone thinks of emoticons. Their use of the screen or on the board is considered frivolous, maybe a bad taste, perhaps tacky to play around with the students. Meanwhile, we think their entry into ISS pending with other analogues of icons of the past and thus their roles and status will change.

The proposed method is from the category of blended teaching where ICT are widely used in the teaching process, but they are no substitute of the lecturer. For ease of learning the disciplineir is necessary the most important things concerning the strategies and fundamental new concepts, paradigms, changes in the subject area to be studied on the spot and educational tools to help study the technical details by checking the knowledge of exercise. Thus giving place to more easily adopting the discipline and to giving more examples and presentations.

As a main conclusion it can be stated as follows. The emotional contact, driving the student to creative events, is not necessarily based on means of measuring emotions. Emotional elements in teaching together with other measures of organizational complexity reduce complexity of the taught subject and there should be left 'doors for independent research'. Using innovative teaching elements in ICs is necessary. This largely relates teaching ICT as a whole. A number of approaches to this are presented. Last but not least, the presented teaching brainstorming process cannot be successfully implemented without the approaches listed here. Therefore, there is a direct link between proper teaching today and create technologies for future IA. Without emotion-based elements it is very difficult to provoke students' interest – a base for deep understanding and high-quality learning.

6. Conflict Resolution Method and its Machine Learning Applications

This section will discuss issues related to improving the quality of schooling, respectively university education through self-study or correction of knowledge on the subject. It is shown that the learner should not regard the material verbatim as a description of classic and always solved problems. Instead, students can learn to high quality fabric and connect it with practice, if we do not seek semantic conflicts of the new material accumulated to date knowledge and do not allow them, for example by studying additional sources. In this situation, it is a good idea to consult with the lecturer, especially if learners are sure they have conscientiously learned topic. Often conflict in knowledge arises from gaps in the description of the material and the study alone or with a specialist the misunderstanding is easily solved. On the other hand, if the learner does not ask any questions to the lecturer, it most often means that the surface is studied and it is about hard work for the quality control of the subject.

The creation and using conflict situations is sometimes useful to be artificially induced by experienced lecturer to improve the understanding. Below are methods and tools for detecting and resolving semantic conflicts in an informal form used in our lectures. On the other hand, if these intelligent agents are sufficiently developed through the implementation of projects in education, the assessment of each student can be obtained dynamically at any point of time and the total scenario in hall training will be completed of information, personalized for the individual learner. For example, most learners memorize the term 'semantic conflict', and the excellent learners will examine this difference with other types of conflicts. But it is possible always there to be people who stop learning before the clarifying any unknown terms. Here a multiagent system with their search and resolution of conflicts will provide the necessary descriptions at any moment of time.

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 '$\neg$' and the weak (conditional, paraconsistent) negation 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 '$\neg$' or '~'. For example: A and $\neg$A; B and not B (using $\neg$ 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 (1).

$$\{U\}[\eta: A, B] \tag{1}$$

where η 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 *an 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 (1) 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: $\chi_1, \chi_2, \ldots \chi_z$.

$$\{U\}[\eta: A_1,A_2,\ldots A_p] <\tilde{\chi}_1 * \tilde{\chi}_2 * \ldots * \tilde{\chi}_z>. \quad (2)$$

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

The present research allows a transition from models of contradictions to ontologies 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 (3):

{ user : faults }[~: accidental, tendentious] <T,¬M,O;¬S> (3)

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. 8 to Fig. 11.

In situations from Fig. 8, 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. 9, 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.

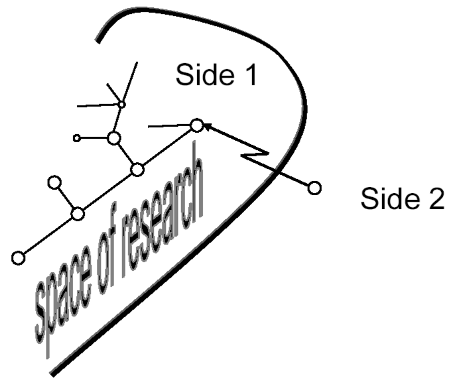

Figure 8. Avoidable (postponed) conflicts when Side 2 is outside of the research space

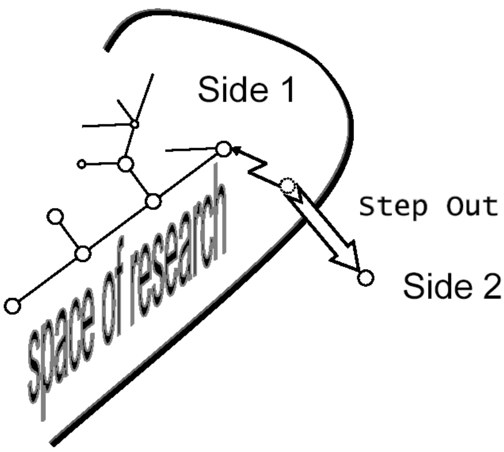

Figure 9. Conflict resolution by stepping out of the research space (postponed or resolved conflicts)

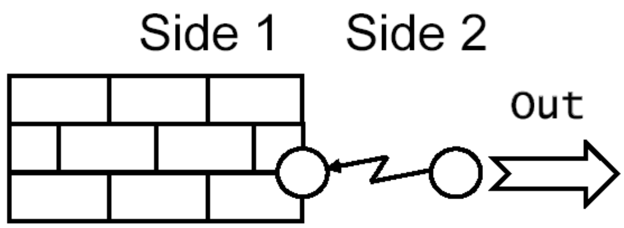

Figure 10. Automatically resolvable conflicts

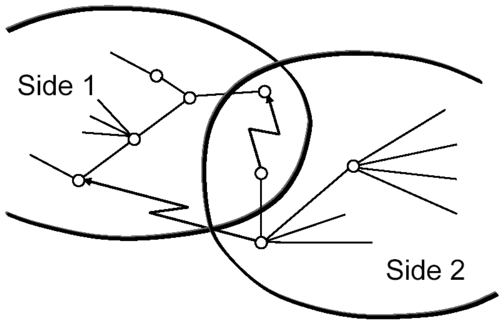

Figure 11. Conflicts resolvable using human-machine interaction

The situation from Fig. 10 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. 11 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.

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. 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.

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.

We rely on nonsymmetrical reply 'surprise and win'.

7. Conclusions

The main conclusion is that to overcome the contemporary education shortcomings, methods and applications are considered concerning the logical parts of knowledge discovery and data mining. Special attention is paid to methods for identification and resolution of conflicts, and to (self-) learning based on them. The role of the above methods for the education purposes is discussed.

Analysis is represented for technologies used for sending information by sense, and for understanding the semantics of the information. Common advantages and disadvantages for different existing groups of contemporary applications are revealed.

Same methods in different combinations are effectively used to enhance security stuff possibilities or in contemporary e-learning systems in the field of Information/National Security [16]. Applications outside the field of information security have been made since a long time, but their explanation goes beyond the field of the considered research concerning the most difficult education cases.

It is shown how, by applying the principles of data mining technology for lecturing to pupils or students, there is a substantial rise in the quality of lecturing even for the most difficult disciplines, and there is sustainability in the implementation of the studied material, which leads to sustainable development of the regions.

Acknowledgement

The paper is published with financial support by the European Union under Bulgaria-Serbia IPA cross-border programme, project IPA MORESYN EDU, CCI number 2007CB16IPO006-2011-2-14.

Bibliography

[1] I. Armuelles et al. On ad hoc networks in the 4G integration process. Proc. of MedHocNet, 3rd Int. Mediterranean Ad Hoc Networking Work shop, ISBN 975-98840-1-1. Bodrum, Turkey, 27-30 June, pp. 45-56, 2004 (online version available).

[2] A. Nayak et al. Enterprise Architecture for Semantic Web Mining in Education. In: Proc. of Second Internatioinal Conference on Computer and Electrical Engineering ICCEE'09, vol.2, pp. 23-26, 2009.

[3] Y.-T. Qu et al. Research about the Application of Web Mining in Distance Education Platform. In: Proc. of International Conference on Embedded Computing SCALCOM-EMBEDDEDCOM'09, pp. 508-513, 2009.

[4] J. Huang, A. Zhu, Q. Luo. Personality mining method in web based education system using data mining. In: Proc. of IEEE International Conference on Grey Systems and Intelligent Services GSIS 2007, pp. 155-158, 2007.

[5] M. Ida. XBRL extension for knowledge discovery in higher education. In: Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) 2011, pp. 2120-2123, 2011.

[6] Vo Thi Ngoc Chau; Nguyen Hua Phung A Knowledge-Driven Educational Decision Support System. In: Proc. of International Conference of Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 2012 IEEE RIVF, pp. 1-6, 2012.

[7] C. J. Carmona et al. Subgroup discovery in an e-learning usage study based on Moodle. In: 7th International Conference on Next Generation Web Services Practices (NWeSP) 2011, pp. 446-451, 2011.

[8] Yang, Tie-li; Ping-Bai; Gong, Yu-Sheng. Spatial Data Mining Features between General Data Mining Education Technology and Training, 2008, pp. 541-544. 2008.

[9] B. K. Baradwaj, S. Pal, Mining Educational Data to Analyze Students' Performance, (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, Vol. 2, No. 6, 2011, pp. 63-69, 2011.

[10] C. Vialardi et al. Recommendation in Higher Education Using Data Mining Techniques, *Educational Data Mining 2009*, pp. 190-199, 2009.

[11] В. Йоцов. Сигурност и защита на информацията. - София: За буквите - О писменехь, 2006 (V. Jotsov. Information Security Systems. Sofia: Za bukvite-o pismeneh, 2006, 156 p).

[12]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.

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

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

[15] 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).

[16]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.

Author information

Vladimir S. Jotsov (В.С. Йоцов) *- University of Library Studies and Information Technologies (ULSIT); P. O. Box 161, Sofia 1113, BULGARIA, EU;*
e-mail: bgimcssmc@gmail.com

Stefan L. Kojnov (С. Койнов) *- Intsitute of Information and Communication Technologies of the Bulgarian Academy of Sciences (IICT-BAS); Block 2 – BAS, acad. G. bonchev Str., Sofia 1113, BULGARIA, EU*
e-mail: slkoynov@gmail.com