

# Virtual Experts for Imagery Registration and Conflation

Boris Kovalerchuk, Artemus Harper, Michael Kovalerchuk  
Dept. of Computer Science  
Central Washington University, Ellensburg, WA, U.S.A

**Abstract.** *The unique human expertise in imagery analysis should be preserved and shared with other imagery analysts to improve image analysis and decision-making. Such knowledge can serve as a corporate memory and be a base for an imagery virtual expert. The core problem in reaching this goal is constructing a methodology and tools that can assist in building the knowledge base of imagery analysis.*

*This paper provides a framework for an imagery virtual expert system that supports imagery registration and conflation tasks. The approach involves two strategies: (1) recording expertise on-the-fly and (2) extracting information from the expert system in an optimized way. The second strategy is based on the theory of monotone Boolean functions.*

**Keywords:** imagery virtual expert, knowledge base, rule generation optimization, monotone Boolean function, registration, conflation.

## 1. Introduction

The goal of *imagery registration* is providing geospatial coordinates to the image. The goal of the *imagery conflation* is correlation and fusion of two or more images or geospatial databases. “The process of transferring information (including more accurate coordinates) from one geospatial database to another is known as ‘conflation’” [5]. Typically, the result of the conflation is a combined image produced from two or more images with: (1) matched features from different images and (2) transformations that are needed to produce a single consistent image. Note, registration of a new image can be done by conflating it with a registered image.

Such a way of registration can be useful if there is a lack of reliable metadata that provide registration directly. This consideration motivates us to concentrate on the conflation task in this paper. Recently the conflation has

been viewed as a *matching technique* that fuses imagery data and preserves inconsistencies (e.g., inconsistencies between high and low resolution maps, “best map” concept, [4]). This approach tries to preserve the pluralism of multisource data. The traditional approach [3] uses an “artistic” match of elevation edges. If the road has a break on the borderline of two maps then a “corrected” road section starts at some distance from the border on both sides and connects two disparate lines. This new line is artistically perfect, but no real road may exist on the ground in that location (see Figure 1).

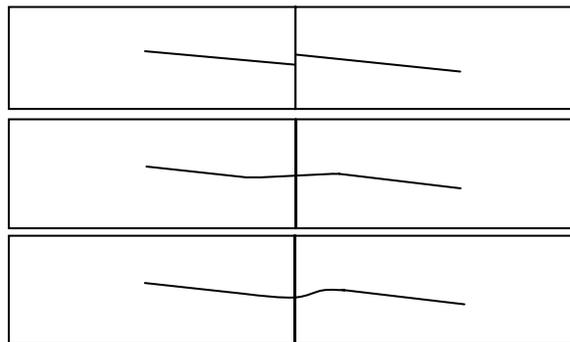


Figure 1. Initial mismatch and “artistic” conflations.

*Why design virtual experts for conflation?*

Can the conflation problem be solved by designing a sophisticated mathematical procedure without relying on an expert’s knowledge? In essence, the conflation problem is a *conflict resolution* problem between disparate data. Inconsistencies in multisource data can be due to scale, resolution, compilation standards, operator license, source accuracy, registration, sensor characteristics, currency, temporality, or errors [4]. The conflict resolution strategies are highly *context and task dependent*. Dependency of conflation from a specific task is discussed in [7].

In solving a conflation problem, experts are unique in extracting and using non-formalized context and in linking it with the task at hand (e.g., finding the best route). Unfortunately, few if any contexts are explicitly formalized and generalized for use in conflating other images. It is common that the context of each image is unique and not recorded. For example, an expert conflating two specific images may match feature F1 with feature F3, although the distance between features F1 and F2 is smaller than the distance between features F1 and F3. The reasoning (that is typically not recorded) behind this decision could be as follows. The expert analyzed the whole image as a context for the decision. The expert noticed that both features F1 and F3 are small road segments and are parts of much larger road systems A and B that are structurally similar, but features F1 and F2 have no such link. This conclusion is very specific for a given pair of images and roads on these images. The expert did not have any formal definition of structural similarity in this reasoning. Thus, this expert's reasoning may not be sufficient for implementing in an automatic conflation system for conflating other images. Moreover, informal similarity the expert used for one pair of images can differ from similarity the same expert will use for two other images.

There are two known approaches to incorporate context: (1) *formalize context* for each individual image and task directly and (2) *generalize context* in the form of expert rules. In the first approach, the challenge is that there are too many images and tasks and there is no unified technique to for context formalization. The second approach is more general and more feasible, but in some cases may not match a particular context and task, thus a human expert needs to take a look.

## 2. Shortcomings of previous attempts to deal with the subject

Currently, even large knowledge bases cannot answer many questions, which are in their scope. The real world is too dynamic, uncertain, and complex for even modern knowledge bases. The conflation/registration

problem is an example of such a real world problem. DARPA's program "High-Performance Knowledge Bases" [1] that started in 1997 has set up the critical size barrier for large knowledge bases around the *10,000 axiom/rule* limit. At that time, DARPA's goal was to build technology, which will scale up to *100,000 axiom/rule* knowledge base systems [1]. The DARPA program "Rapid Knowledge Formation" [2] formulated new requirements that include parallel entry of knowledge by teams of 25-50 individuals (end users) for test tasks such as crisis management and battlespace understanding.

According to DARPA using High Performance Knowledge Base technology, a 5-person team can create knowledge at a rate of *40 axioms* per hour and *100K* of axioms per year. After that, DARPA stated a new goal: the creation of new knowledge at a rate of *400 axioms* per hour. Next, DARPA identified the criterion of *comprehensiveness* of the knowledge base at the level of *a million axioms*. The PARKA project research team at the University of Maryland extracted *125913 assertions* from CIA World Fact book pages on the World-Wide Web using a *web robot* [10]. Thus, a million axioms can be comparable with encoding knowledge from eight books like the CIA World Fact book. Next, notice that these assertions are not rules, but facts such as "economy\_imports#tadjikistan#\$690 million {1995}", that is Tjikistan's import is \$690 million (1995). In addition, those assertions have been extracted from the written text using a web robot. For the conflation task, there is *no text available* to use as a source for a web robot. This means that we need to build such written sources and extract rules from experts directly.

To understand what the million-axiom size of the knowledge base means in more detail we need to clarify the concept of axiom used by DARPA. The same DARPA source provides an example of the axiom:

$$\forall x, p1, p2 \text{ vehicle}(x) \Leftrightarrow \text{physical\_object}(x) \\ \text{and self-propelled}(x) \\ \text{and can}(\text{move}(x), p1, p2).$$

This is a relatively simple statement with three basic statements combined using AND operator. For complex tasks with interdependent

attributes, axioms can involve more than ten statements connected by the AND operator. Respectively the time for extracting these rules can be much longer. They also can be much less trivial and certain as in conflation problems.

Let us consider the rate and quality of knowledge base development reached in 1999 in the High Performance Knowledge Base program [2]. The maximum number of axioms was 90,000 per 10 months and the smallest number of axioms was 2300 per seven months. Depending on the domain and the problem, 90K may be not enough or 2300 may be enough. Also note that  $90,000 < 2^{17}$ , which means that for designing a complete knowledge base with 17 binary attributes we need even more axioms.

*Building comprehensive virtual experts.*

Let us illustrate the important question of completeness and comprehensiveness of the knowledge base. A knowledge base will be called *complete* if for a given set of attributes the knowledge base can generate an answer for every combination of values of these attributes. We will say that a knowledge base has *comprehensive coverage* if a set of attributes of rules in the knowledge base covers most of attributes used in the domain. For instance, we may include in the knowledge base all the attributes used in NIMA's Vector Product Format (VPF) to get a comprehensive coverage. Nevertheless, this knowledge base may not be complete because rules cannot produce answers for many questions formulated as AND combinations of VPF attributes.

However, there are some *positive examples* of a complete knowledge base. For instance, in medical imaging [8, 9] having 11 binary attributes of X-ray images (mammograms) entered into the knowledge base for a particular patient the knowledge base should output one of the target values: "highly suspicious for malignancy" or "probably benign". Another target with the same 11 binary features for the same patient will be biopsy (should biopsy be done or not). Both of these questions are life critical questions (more than 100,000 die each year of breast cancer in the US). An axiom "mined" from the experienced radiologist may look like the following:

*If variation in shape of calcifications is marked AND the number of calcifications is between 10 and 20 AND the irregularity in shape of calcifications is moderate THEN the case is highly suspicious for malignancy.*

The common way to extract such rules is to ask an expert to write down rules that the expert uses and to provide software for converting the rules to computer-readable knowledge base (KB) form. However, it is unlikely that the expert will enter complex rules involving, say 12 attributes. Often it is above a human's capabilities to keep in mind more than 5 to 9 attributes for analysis. Later testing may show that the rule with only 3 attributes is wrong for some cases and the knowledge base should be refined. The refinement can take years and for life critical applications the systems should not be used before the process of cleaning rules will successfully finish. The problem is that this process can be exponential in time. For instance, having 14 binary attributes we may search among  $2^{14} = 16384$  potential rules like the rule shown above.

To avoid refining and testing the knowledge base for years, we need to be sure that the set of rules is complete enough from the beginning. Asking the expert does he believe that 10 rules he entered are complete may not be the right choice. We need to be sure that the rest of potential  $16384 - 10 = 16374$  rules are not rules at all. Thus, DARPA's design time should also measure both rules included in the knowledge base and the rules *rejected*. If we know that something is not a rule, this is also useful knowledge. There is a big difference between a rejected rule and a rule unconfirmed by the expert or not tested against independent data yet. DARPA's current goal of 1,000,000 axioms, would correspond to a design of a complete knowledge base with less than 23 binary attributes. More exactly, in the worst case scenario for 22 attributes we may need to record 1,144,066 axioms (using formula from Hansel Lemma [8, 6]), which exceeds the 1,000,000 axioms. To find all these axioms (rules) we may need to search in a much larger set of possible axioms, which is  $2^{23}$ .

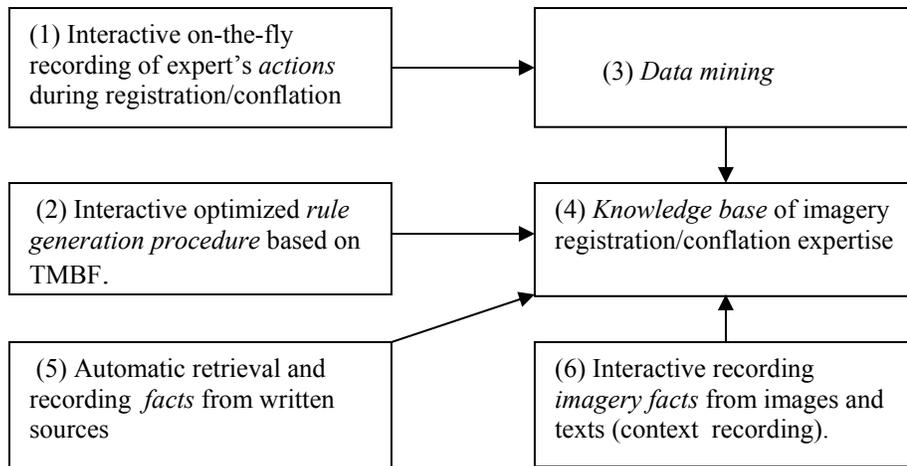


Figure 2. General system design

Let us note that the problem of designing a complete knowledge base above the limit of 20 attributes ( $10^9$  potential rules) is especially critical when knowledge is not presented in any printed form (book, articles) and should be extracted from an *expert as a sole knowledge body*. This is the case of the virtual expert for imagery conflation/registration problem.

### 3. System Architecture

The general design of our system is presented in Figure 2. The complete system design contains six components. The first component- Interactive on-the-fly recording of expert's *actions* during registration/conflation serves as major source of the raw information for rule generation. It is also useful as a quality control tool of the analyst's conflation results (a kind of airplane "black box"). An interactive optimized *rule generation procedure* based on the theory of monotone Boolean functions (TMBF) is intended to speed up direct generation rules by the imagery analyst. We implemented these two components in Java as a web portal. These components are described in more detail in the following sections. The data mining component (block 3) is designed to generalize the record of an expert's actions. The results of such generalizations are conflation rules. The major problem for successful mining of rules from recordings is that the system actually records lower level expert's actions, such

as rotation, scaling, translation. Mining these lower level rules may not be very beneficial, thus the system provides a tool for the expert to record an identification of upper level categories such as "selecting main feature", "conflating main features", etc. The system design includes recording expert's actions and mined rules in XML format, for rules it is RML (rule markup language).

Automatic retrieval and recording of facts from written sources (Block (3)) is designed to fulfill functions similar to PARKA project [10]. At this moment, we do not consider this source as a main source to fill the KB with facts related to registration and conflation of images, but potentially it can provide useful facts for the KB. In contrast, interactive recording of imagery facts from images and texts (block (6)) can be one of the major sources of conflation related facts right now. The reason is that such recording provides context for conflation and registration tasks.

### 4. Interactive optimized rule generation

The rule generation contains five steps:

- (2.1) Interactive recording *characteristics* to be used as arguments of rules;
- (2.2) Interactive optimized *rule generation* based on the theory of monotone Boolean functions;
- (2.3) Recording *test cases* for testing rules;
- (2.4) *Testing* rules against test cases, and

(2.5) *Recording rules* to the knowledge base of imagery registration/conflation expertise.

The logic of *rule generation* (step (2.2)) based on the theory of monotone Boolean functions is as follows. Assume that the analyst identified  $n$  rule arguments (an example of such arguments is shown in Figure 3 for the task of creating expert rules to judge that two images can be conflated by using a single affine linear transformation). Below we list the first three arguments: (1) a simple dominant geometric feature exists, (2) a simple unique geometric feature exists, and (3) an asymmetric unique features exist.

The full list, in this example, includes 14 arguments. Now the imagery expert can be asked if all these 14 arguments are true, would he/she conclude that two images can be conflated with a single affine linear transformation. If the answer is “yes”, then it will be encoded as 1. Taking into account that all 14 arguments are Boolean too, we matched a Boolean vector (11111111111111) to (1). We may ask the expert the same question, but do not require some arguments to be true. For instance, we can ask about situation represented by the vector ((11011101110111)). The total number of questions could be  $2^{14}$ . Assume that we did not ask the expert all these  $2^{14}$  questions. Then if in the real conflation case we have a situation (11011101110111) that was not asked about then the knowledge base does not provide the answer and the KB is not complete. The theory of monotone Boolean functions allows us to avoid asking  $2^{14}$  and still generate a complete set of rules.

Two fundamental ideas permit to accomplish this optimization of the expert interview process: (1) dynamic sequence of questions – each further question depends on the expert’s answers to previous questions, (2) rule arguments are designed in a such way that the property of monotonicity is fulfilled. The software system for questioning the expert using the property of monotonicity is shown in Figure 3.

Monotonicity means that if the expert believes that in the situation presented by the vector  $\mathbf{v}=(11011101110111)$  there is no way to

have a single linear affine transformation then for every situation that has some “1” in  $\mathbf{v}$  substituted by “0” the answer should be negative too. This means that we can eliminate a large number of questions to the rational expert.

As was described above in the process of interactive rule generation, the expert knowledge is modeled by the system as a set of binary vectors, which represent combinations of predefined image characteristics. This approach has been successful in the medical application. The medical imagery analyst (radiologist) was asked only 40 questions and a complete set of rules (out of potential 2048 questions) has been extracted [8, 9] just in 30 minutes. A prototype web-based medical expert consultation system has been created and a demonstration version is available at <http://www.cwu.edu/~borisk/rulegen/rulegen.html>

The expert can define the rules (as shown in Figure 4) as binary vectors, which are then tested against a database of test cases for which the answer is “known” (in the medical application verified though biopsy). In the image conflation, the known cases can be the cases that are actually georeferenced and the correct transform can be computed from reference points. This would be testing against georeferenced data.

## 5. Interactive on-the-fly recording of expert’s actions

The system component that supports interactive on-the-fly recording of expert’s actions is implemented as a web portal that allows an expert to conflate images while having the knowledge presented by conflating these images recorded on-the-fly. The tool called Case Recorder currently allows the user to load a set of images, then conflate these images using basic scaling, translation and rotation tools.

The user can view these images overlapped and change the opacity of the images for conflating. The system provides facilities for marking up the sections of the images using various shape tools. Each of these markups can be named by the user using a basic name of

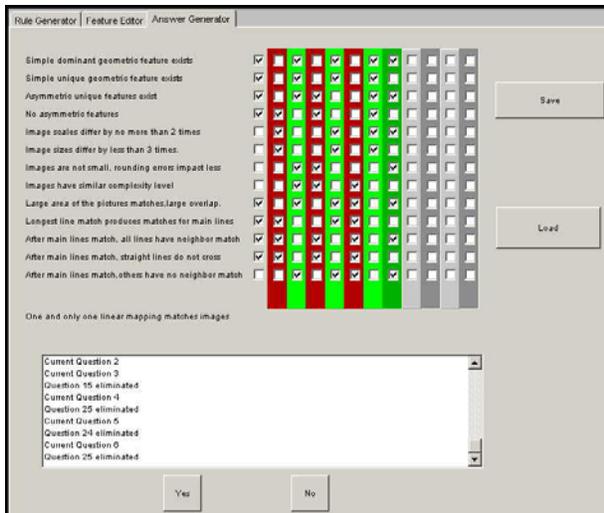


Figure 3. Expert questioning system using monotonicity principle.

his/her choice or by choosing from a list of predefined terms. Currently, the system permits to load DAML Geofile ontology [11] and DAML CIA World Fact book ontology [12] and markup selected parts of images with these

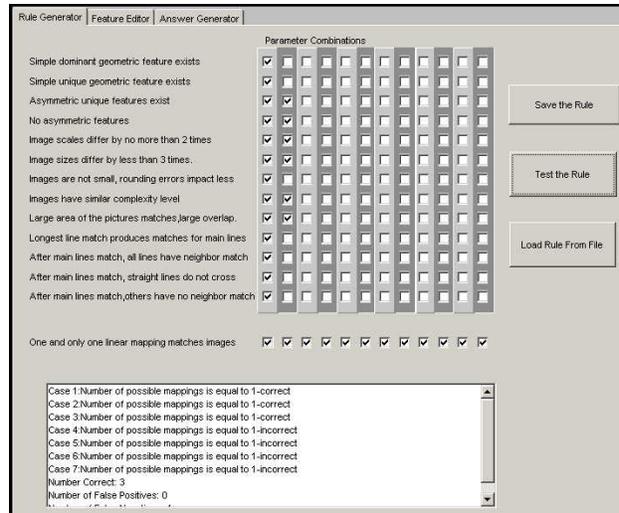


Figure 4. Defining Expert Rules

terms (see figure 5). There is also a detailed list of actions the user has performed that can be undone and redone to any point. A basic magnifier is available for taking a detailed look at the image. Any of these markups

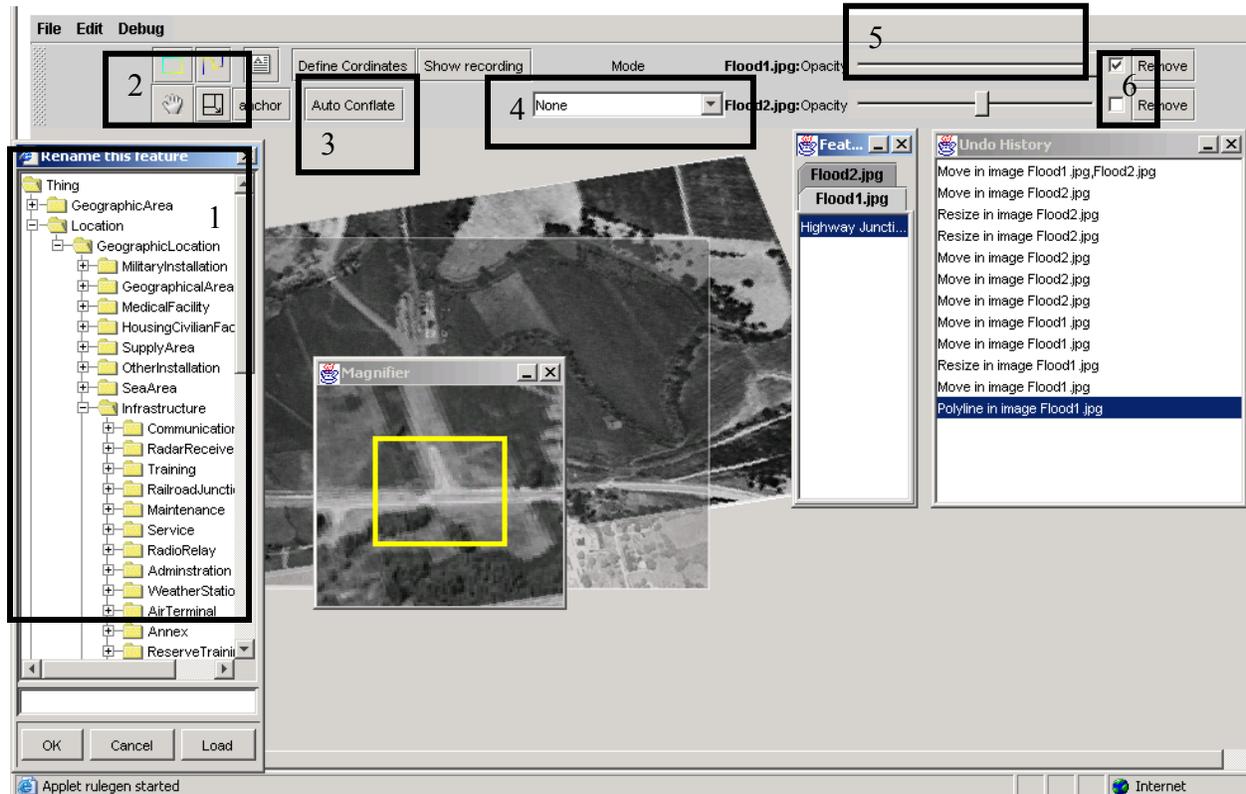


Figure 5. Image of sample conflation using the case recorder tool

and conflatons can be applied to multiple images to allow two (or more) images already conflated to be conflated to a third image or simply zoom in all images. All of these actions are recorded, can be presented in a human readable form and are available for playback.

The list below presents comments to numbered items shown in Figure 5.

1. The Markup tools can be used to mark major features on the image and name them, these names can be defined in a DAML ontology [11].
2. The move, resize, and rotate tools are basic operations used to conflate images
3. The Auto Conflate tool allows the user to choose 3 points on two images and using transformations, match those points together (and hopefully the images as well).
4. The Show recording button allows the user to see how the conflation was broken up via the current mode that was used in various segments of the conflating.
5. The Opacity slider is used to set the transparency of the images, allowing one image to be seen through another.
6. The checkboxes are used to select the image(s) that will receive the operations such as move or draw polyline. This allows a user to resize or rotate both images together, or once two images have been conflated together, a third image can be brought in and conflated against the other two together.

## 6. Conclusion

In moving toward the goal of preserving human expertise in imagery analysis and capturing non-formalized context, two tools have been developed to assist knowledge engineers and image analysts in populating the knowledge base of the virtual expert. The first tool records an imagery analyst's actions on the fly; and the second tool generates expert rules by questioning the imagery analyst and minimizing questioning time using the theory of Monotone Boolean functions. These tools are implemented as a web portal using Java.

## 7. References

- [1] High-Performance Knowledge Bases, DARPA, 1996, <http://www.darpa.mil/iso/documents/hpkb/baa96-43pip.html>
- [2] Rapid Knowledge Formation, DARPA, 1999, <http://www.darpa.mil/iso/RKF/RKFOverview.ppt>
- [3] Digital Elevation Model Standards, USGS <http://rockyweb.cr.usgs.gov/nmpstds/demstds.html>
- [4] Edwards D., Simpson J. Integration and access of multi-source vector data, In: *Geospatial Theory, Processing and Applications*, ISPRS Commission IV, Symposium 2002, Ottawa, Canada, July 9-12, 2002, <http://www.isprs.org/commission4/proceedings/pdfpapers/269.pdf>
- [5] Federal Geographic Data Committee FGDC-STD-999.1-2000, NSDI Framework Transportation Identification Standard -- *Public Review Draft*. [www.bts.gov/gis/fgdc/introduction.pdf](http://www.bts.gov/gis/fgdc/introduction.pdf)
- [6] Hansel G., Sur le nombre des fonctions Boolenes monotones den variables. *C.R. Acad. Sci.*, Paris (in French), 262(20):1088-1090, 1966.
- [7] Kovalerchuk, B., Schwing J., Mathematical support for combining geospatial data, In: *Proceedings of International Conference on Imaging Science, Systems, and Technology (CISST'2001, June 25-28, 2001)*, Las Vegas, pp. 485-491.
- [8] Kovalerchuk, B., Triantaphyllou, E., Despande, A.S., and Vityaev, E. Interactive Learning of Monotone Boolean Functions, *Information Sciences*, Vol. 94, issue 1-4, 1996, pp. 87-118.
- [9] Kovalerchuk, B., Vityaev E., Ruiz J.F., Consistent and Complete Data and "Expert" Mining in Medicine, In: *Medical Data Mining and Knowledge Discovery*, Springer, 2001, pp. 238-280.
- [10] Very Large Knowledge Bases, 1998, <http://www.cs.umd.edu/projects/plus/Parka/parka-kbs.html>
- [11] DAML GeoFile ontology, <http://www.daml.org/2001/02/geofile/geofile-ont>, 2001
- [12] DAML WFB ontology <http://ontolingua.stanford.edu/doc/chimaera/ontologies/world-fact-book.daml>, 2002