Computer Science Journal of Moldova, vol.**17**, no.1(49), 2009

Structured knowledge management techniques for the development of interactive and adaptive decision support system

Iulian Secrieru

Abstract

The phase of knowledge acquisition and formalization is being considered as the key one for the development of decision support system (DSS). The main problem at this stage is to find a knowledge representation (KR) and a supporting reasoning system that can make the inferences your application needs. The main criterion of choice is what kind of inference the developers prefer and is more appropriate to the problem domain. However, already at the stage of the user interface creation there arise tension between KR and end-user "needs" (preferences and habits). So, multiple representations of the acquired and structured knowledge and management techniques could provide a solution for decision support system inference and interface requirements satisfaction.

Keywords: Knowledge representation, knowledge management techniques, decision support system, interactive and adaptive interface.

1 Introduction

The problem, associated with the physicians (doctors) diagnostic activities, acquire a special relevance in modern circumstances. First of all it is connected with the fact that the doctors have to work with weakly structured and formalized information. Besides, the volume of information is in a continuous growth thanks to the appearance of new methods of examination of patients.

The ultrasound investigation domain is not an exception. The appearance of new ultrasound devices or the improvement of the old scanners doesn't simplify but complicates the physician's diagnostic thinking, because he has to analyze a much larger number of diagnostic data, which typically reduces the accuracy and increases the time of determining the diagnosis.

The phase of knowledge acquisition and formalization is being considered as the key one for the development of decision support system (DSS). In order to obtain a well-structured description of the problem domain, the developers are forced to choose a "rigid" scheme of its representation. The problem domain is quite often represented as a decision tree or semantic net.

The development of the user's interface for the DSS based on the decision tree can lead to various problems and inconveniences.

The main lack is the fact that the interface does not correspond to the daily work and habits of the end-user.

The discrepancy of the user's interface of the decision support medical system with the form of the doctor's diagnostic thinking may become the reason of different mistakes or it may lead to the rejection of the user to utilize it in medical practice.

The principles and techniques of structured knowledge manipulation and management aiming to create an interactive and adaptive user interface for a decision support system in the ultrasound investigation domain, is being described in this article.

2 Knowledge Representation Schemes

The fundamental goal of KR is to represent knowledge in a manner as to facilitate drawing conclusions (inferencing) by decision support or another computer-aided systems.

We distinguish two approaches - single and hybrid KR schemes.

First we focus on the most popular single KR schemes.

Semantic nets, decision trees and their descendants (*frames* or *schemes*) [1] represent knowledge in the form of graph (or hierarchy). Nodes in graph represent concepts and the edges represent relations between the concepts. Nodes in a frame hierarchy also represent concepts, but they have internal structure that describes the corresponding concept via a set of attributes. All of these KR schemes are very natural and well suited for representing structural and relational knowledge. They can also make efficient inferences for small to medium graphs (hierarchies). However, it is difficult to represent heuristic knowledge, uncertain knowledge, and make inferences from partial inputs. Also, explanations are not provided and knowledge updates are difficult. *Conceptual graphs* are similar to semantic nets, whereas ontologies [2] refer to a representation scheme similar to frames, but more restrictive.

Symbolic rules are one of the most popular KR methods [1]. They represent general domain knowledge in the form of IF-THEN rules: if <conditions> then <conclusion>, where the term <conditions> represents the conditions of a rule, whereas the term <conclusion> represents its conclusion. The conditions are connected with one or more logical operators such as "and", "or", and "not". The inference engine uses the knowledge in the rule base as well as facts about the problem at hand to draw conclusions. Typically, facts are provided by the user during inference. There are two main inference approaches: backward chaining (guided by the conclusions) and forward chaining (guided by the input data). The explanation module provides explanations regarding the drawn conclusions. Rules are natural (easy to comprehend) and rule-base updates (removing/inserting rules) can be easily made. In addition, heuristic knowledge is naturally represented by rules. Efficiency of the inference process depends on the length of the inference chains. Additionally, conclusions cannot be derived if some of the inputs are unknown. Finally, pure rules cannot represent uncertain or vague knowledge and are not suitable for representing structural and relational knowledge.

Fuzzy rules (fuzzy logic) are good at representing imprecise and fuzzy terms, like "low" and "high". Fuzzy logic extends traditional logic and sets membership by defining membership functions over the range [0.0,1.0], where 0.0 denotes absolute falseness and 1.0 - absolute truth [3]. Given the above, fuzzy rules are good for representing vagueness. However, fuzzy rules are not as natural as symbolic rules, that complicates the knowledge acquisition and the updates processes. Inference is more complicated and less natural than in simple rule-based reasoning. Provision of explanations is feasible, but not all reasoning steps can be explained.

Case-based representations [4] store a large set of past cases with their solutions in the case base and use them whenever a similar new case has to be dealt with. A case-based system performs inference in four stages: (1) retrieve, (2) reuse, (3) revise, and (4) retain. In the retrieval stage, the stored case(s) most relevant to the new case is (are) retrieved. Similarity measures and indexing schemes are used in this context. In the reuse stage, the retrieved case is combined with the new case to create a solution. The revise stage validates the correctness of the proposed solution. Finally, the retain stage decides on retention (or not) of the new case. Cases are usually easy to obtain. Cases are natural. Explanations cannot be provided in a straightforward way as in rule-based systems. Even if some of the inputs are not known, conclusions can be reached through similarity to stored cases. Updates can be easily made. However, the efficiency of the inference process depends on the size of the case base. Finally, cases are not suitable for representing structural, uncertain, and heuristic knowledge.

Neural networks represent a totally different approach to artificial intelligence, known as connectionist [5]. A neural network consists of many simple interconnected processing units called neurons. Neural networks are very efficient in producing conclusions, since inference is based on numerical calculations, and can reach conclusions based on partially known inputs due to their generalization ability. On the other hand, neural networks lack naturalness of representation, that is, the encompassed knowledge is incomprehensible, and explanations for the reached conclusions cannot be provided. It is also difficult to make structural updates to specific parts of the network. Neural networks do not possess inherent mechanisms for representing structural, relational, and uncertain knowledge. Heuristic knowledge can be represented to some degree via supervised training.

Belief networks (or *probabilistic nets*) [6] are graphs, where nodes represent statistical concepts and links represent mainly causal relations between them. Each link is assigned a probability which represents how certain is it that the concept, where the link departs from, causes the concept, where the link ultimately arrives. Belief nets are good for representing causal relations between concepts. Also, they can represent heuristic knowledge. Furthermore, they can represent uncertain knowledge through the probabilities and make relatively efficient inferences (via computations of probabilities propagation). However, estimation of probabilities is difficult, making the knowledge acquisition process a problem. For the same reason, it is difficult to make updates. Also, explanations are difficult to produce, since the inference steps cannot be easily followed by humans. Furthermore, their naturalness is reduced.

Hybrid schemes are integrations of two or more single KR schemes. We mention the most popular ones.

Connectionist rule-based representations [5] combine neural networks with rule-based representation. The knowledge base is a network whose nodes correspond to domain concepts. Dependency information regarding the concepts is used to create links among nodes. The network's weights are calculated through a training process using a set of training patterns. In addition to the knowledge base, connectionist rule-based systems also consist of an inference engine and an explanation mechanism. Compared to neural networks, they offer more natural representation and can provide some type of explanation. Naturalness is enhanced due to the fact that most of the nodes correspond to domain concepts.

Another approach in hybrid knowledge representation is the *integrations of rule-based reasoning with case-based reasoning* [7]. Compared to "pure" case-based reasoning, their key advantage is the improvement in the performance of the inference engine and the ability to represent heuristic and relational knowledge. Furthermore, the synergism of rules and cases can cover up deficiencies of rules (improved knowledge acquisition) and also enable partial input inferences. The existence of rules in such hybrid schemes makes updates more difficult than "pure" case-based representations. Also, explanations can be provided but not as easily as in pure rule-based representations, given that similarity functions are still present.

There are various ways to integrate neural networks and fuzzy logic [8]. Such integrations are the fuzzy neural networks and the hybrid neuro-fuzzy representations. Fuzzy neural networks retain the basic properties and architectures of neural networks and "fuzzify" some of their elements. In a hybrid neuro-fuzzy system, both fuzzy techniques and neural networks play key role. Each does its own job in serving different functions in the system. Hybrid neuro-fuzzy systems seem to satisfy KR requirements to a greater degree than fuzzy neural networks. This hybrid approach enables the representation of incomplete, imprecise, and vague information and also exploits the generalization capability of neural networks.

Neurules are type of hybrid rules integrating symbolic rules with neurocomputing [9, 10]. In contrast to other hybrid approaches, the constructed knowledge base retains the modularity of rules, since it consists of autonomous units (neurules), and also retains their naturalness in a great degree, since neurules look much like symbolic rules. Neurules can be constructed either from symbolic rules [9], thus exploiting existing symbolic rule bases, or empirical data [10].

A conclusion that can be drawn is that there is no single or hybrid schemes that satisfy end-users preferences and/or all the requirements of decision support systems developers. So, taking into account only the system requirements on the knowledge acquisition and modeling stages, one can say that semantic nets, decision trees, frames, description logics are more suitable for representing knowledge in the domain model.

3 Knowledge base of the decision support system SONARES.

SONARES is a knowledge based system in the ultrasound investigation domain. Experts are its main source of knowledge. Expert knowledge was obtained in result of "knowledge engineer – expert group" communication and stored as a pyramid of meta-concepts. The common work of knowledge engineer with the experts revealed that in the ultrasound investigation domain the reasoning based on meta-concepts (facts) and knowledge representation in the form of a hierarchy (pyramid) totally corresponds to the expert's thinking and reasoning.

The semantic rules scheme was chosen as a model of acquired knowledge representation. Based on the principles of semantic rules scheme the knowledge base of the decision support system SONARES has been established. It consists of a pyramid of meta-concepts, and of a set of rules created on its basis.

As a result of 23 common working sessions of knowledge engineer with experts there was received a pyramid of knowledge (decision tree) which consists of 335 facts (9 root nodes with a maximum deep level equal to 9) and 54 rules [11, 12 , 13]. This knowledge represents formalized description of the ultrasound investigation process of gallbladder.

4 The user interface based on the decision tree scheme

Let's analyze the following example.

Suppose that during the stage of knowledge acquisition the developers have identified a group of 3 mutually exclusive facts: F1 = <gallbladder volume, normal>, F2 = <gallbladder volume, enlarged>, F3 = <gallbladder volume, reduced>. That is, there was identified the attribute A1=<GALLBLADDER VOLUME> with three possible values: V1="normal", V2="enlarged", V3="reduced". Due to the sources of knowledge, these three facts can be represented as a decision tree in three different ways (see Figure 1).

If we take into consideration all possible options for the interchangeability of the facts for each of these ways of representation, we'll get 18 different decision trees to describe these 3 facts.

Thus, the interface creation on the base of one of the 18 decision trees can be unusual for certain users, whose process of reasoning is described by one of the remaining 17 options. In this case, the created interface won't correspond to their daily work.

Besides, the user interface based on the decision tree scheme is

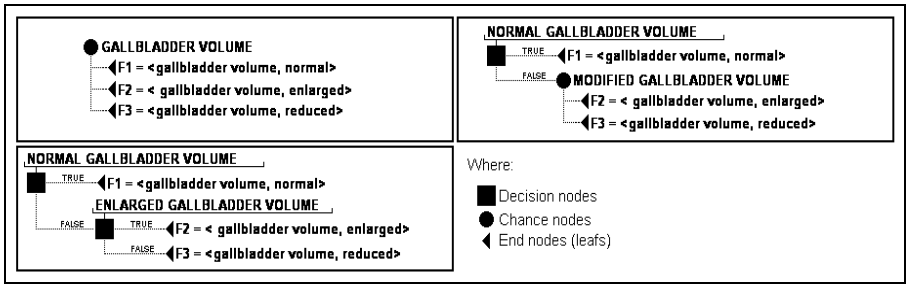

Figure 1. Decision trees for description of 3 exclusive facts.

inefficient because:

- unjustifiedly limits the end-user actions
- requires frequent appeals to the knowledge base
- requires a lot of screen space

It is obvious, in order to organize an effective dialog with the DSS users, it is necessary to develop an alternative representation scheme for the knowledge base represented in form of decision tree.

5 Alternative representation of the knowledge base as a means of effective user interface for the DSS.

Note that we are not talking about replacing the decision tree as a means of knowledge representation scheme at the stage of knowledge acquisition. Since, at this stage, the use of such representations is reasonable, and in some cases it is the necessary and the only right decision. Especially in cases when developers want to receive a well-structured knowledge base and have to deal with poorly formalized domain. In these cases it is meaningful and effective to use the decision tree scheme even as a means of visualization the knowledge base. [14, 15].

The aim of creation of the alternative representation scheme of the knowledge base is the organization on it's bases the effective dialog with the DSS users and the elimination of the deficiencies, inherent in the user interface based on decision tree scheme.

The source of information for alternative representation scheme is the DSS knowledge base, described as a decision tree.

In our case, the DSS SONARES knowledge base is represented in the form of a decision tree which consists of 335 facts that can describe any situation in the domain of gallbladder ultrasound investigation. There were formulated 54 rules on its base, which correspond to the anomalies and pathologies of this problem domain. It is necessary to propose such representation of this acquired knowledge in order to have the opportunities to realize the user interface of DSS SONARES with the following features:

<u>The interface must be simple and understandable.</u> The dialog with the end-users should take place in its usual rhythm and form, and should not be necessary required the change of his reasoning.

<u>The interface should correspond to the user's daily work and preferences.</u> The users must have the possibility to influence the dialog form.

<u>The interface must be "transparent".</u> The solution proposed by DSS should be easily verified.

<u>The user dialog should not have a linear structure.</u> The user should always have the opportunity to return to the appointed step back.

<u>The interface should be interactive.</u> It must change in dependence on the user available time to make a decision. In addition, the interface must conform to the basic forms of user's diagnostic thinking.

<u>The interface should not unnecessarily restrict the end-user's actions.</u>

<u>The interface should be oriented on a restricted screen space and a limited decision making time.</u> The medical DSS users often have to use them in an emergency or a network mode.

The most common form of communication and information transfer is the dialog. Therefore, the organization of the DSS user interface as an ordered set of questions is justified.

The facts of which the decision tree consist, in fact constitute a meta-knowledge, through which we can describe some situation in the selected problem domain and are involved in the inference to determine solutions. That is, these are the answers to those questions that we should address to the user in order to help him with the decision making. It is obvious that in the interface based on a decision tree, the questions themselves are missing (determined by the structure of the decision tree). It doesn't correspond to the usual way of the users reasoning, because for each of the fact from the decision tree representation the user is forced to formulate a question. A more common variant for him could be the option to answer to the specific questions by selecting from a list of all possible answers.

The essence of the proposed new representation approach is the separation of knowledges to those ones that can be used in the inference and those which are used only in the interface.

At the first stage of the creation of the alternative representation of the knowledge base there were determined those facts of the decision tree, which are involved in the inference. For each of them there has been formulated a question concerning the existence or not of this fact. For example, for the fact F1 = <gallbladder volume, normal> there was formulated the question Q1 = "Is the volume of gallbladder a normal one?", for the fact F2 = <gallbladder volume, enlarged> - the question Q2 = "Is the volume of gall bladder enlarged?", and for F3 = <gallbladder volume, reduced> - Q3 = "Is the volume of gallbladder reduced?".

As a result, 203 questions were formulated.

Answering to some of these questions, the user can describe the case of gallbladder ultrasound investigation domain, in which he needs assistance of DSS SONARES.

In terms of these questions all of 54 pathologies and anomalies of this domain are described. That is, each pathology or anomaly from the gallbladder ultrasound investigation domain can by described by the vector (Q1.value,Q2.value,...Qn.value), where Qi.value — is the answer to the question Qi, n - total number of questions, in our case n=203. Under these conditions, the whole diagnostic knowledge base (the information about all pathologies and anomalies, that is necessary for the inference to make a decision) can be represented in the form of decision making matrix [Pi,Qj.value], in our case i=1..54, j=1..203.

Concerning the matrix representation of diagnostic knowledge base, the proposed approach was named *alternative matrix representation* of KB.

On the second stage, we saved all existing relationships between facts. That is, we elaborated an interconnection system between all formulated questions.

There are two types of relationships between facts in the decision tree.

The first one indicates the position of given fact in the knowledge base hierarchy.

Let's analyze the subtree F4-F5-F6, where F4=<gallbladder form, abnormal (abnormality of conformation)>, F5=<gallbladder twist, present>, F6=<gallbladder twist form, circular>. There is a hierarchical relationship between the facts F4-F5, which indicates that it makes sense to show the fact F5 only in the case when the fact F4 is determined. The same relationship exists between the facts F4-F6 and F5-F6. They are not taken into account during the inference process, however, they are of great importance for the determination of the opportunity of a fact visualization. In our case, this information helps us to determine the opportunity of a question visualization and organize the dialog with the DSS user.

The second type of relationships indicates the existence of interdependence between facts.

For example, the above mentioned facts F1, F2, F3 are mutually exclusive (it means If F1 = TRUE then (F2 = FALSE) & (F3 = FALSE), If F2 = TRUE then (F1 = FALSE) & (F3 = FALSE), If F3 = TRUE then (F1 = FALSE) & (F2 = FALSE)). These relations does not depend on the form of visualization of the facts or the whole user interface, but form the basis of the system knowledge base and inference.

The separation of the existing relationships between questions in two groups, those that can be used in inference and those which are used only in the interface, allows us to create a high-quality interactive interface based on individual characteristics and habits of the end-user. This is achieved because the user can define himself the subject and the form of dialog (by changing the visualization relationships between questions), without any fear to influence the inference.

Some of the questions can be grouped. For example, the questions Q1 = "Is the volume of gall bladder a normal one?", Q2 = "Is the volume of gall bladder enlarged?", and Q3 = "Is the volume of gall bladder reduced?" could be grouped into the group, which describes the volume of the gallbladder. Now, if the user wants to visualize all the questions related to the volume of the gallbladder, he may do so through the visualization of the group.

Additionally, the questions association into the group will allow to diversify the form of dialog.

The resulting relational database is the alternative representation of DSS SONARES knowledge base.

6 Conclusions

Realization of the described approach has shown that the creation of alternative matrix representation of DSS knowledge base requires additional time for its creation, but it is justified, if we want to be able to organize an effective interactive dialog with the user. In addition, the user interface based on a matrix representation is simple, understandable and transparent, fully corresponds to the daily activities and habits of the user, not unreasonably restrict the user actions.

This approach allows realization of different versions of the interface with the restricted screen space and limited time of the decision making (for systems used in emergency cases).

Table 1 compares the KR schemes discussed in the previous sections with the proposed approach. Symbol "-" means "unsatisfactory"; "±" - "average"; "+" - "good"; and "V" - "very good".

Table 1. Comparison of knowledge representation schemes and approaches (data source [16])

	USER'S REQUIREMENTS									SYSTEM REQUIREMENTS					
	Naturalness	Ease knowledge update	Efficient inference	Explanations	Ease knowledge acquisition	Partial input inferences	Adaptive inference	Interactive interface	End-user's preferences update	Structural knowledge	Relational knowledge	Uncertain knowledge	Vague knowledge	Heuristic knowledge	New knowledge discovery
Semantic nets / decision trees / frames	V	±	V	-	+	-	±	±	±	V	V	-	-	-	±
Symbolic rules	V	V	+	V	±	-	+	+	±	-	±	-	-	V	±
Case-based representations	V	V	+	+	V	+	±	±	-	-	+	-	-	-	±
Belief networks	±	-	V	-	±	-	-	±	-	+	V	V	±	±	±
Neural networks	-	-	V	-	V	V	+	+	±	-	±	-	-	±	+
Fuzzy rules	+	-	+	-	±	-	±	±	-	-	±	±	V	V	±
Connectionist expert system	±	±	V	±	V	V	±	±	±	-	±	-	-	±	±
Neuro-fuzzy representations	±	-	+	-	+	±	±	±	±	-	±	±	V	+	±
Cases and rules	V	+	+	+	+	+	+	+	±	-	+	-	-	+	±
Neurules	+	+	V	V	V	V	+	+	+	-	±	-	-	V	±
Semantic nets/decision trees & symbolic rules in form of "matrix" representation	V	+	V	V	+	±	+	+	+	V	V	-	-	V	+

Also there have been identified some additional advantages of using a matrix representation of the DSS knowledge base.

1. Matrix representation of knowledge base allows to organize an interactive interface according to the type of user's diagnostic thinking. We realized a version of user interface with adaptive support for inductive reasoning ability.

2. Matrix representation of knowledge base has a cognitive value. It can be used as a means of visualization and detection of weakly described sub-domain in the problem domain in general and in the knowledge base in particular.

3. In matrix representation of the knowledge base every decision is described by the vector (Q1.value, Q2.value, ... Qn.value). By calculating the correlation coefficient between the vectors, the solutions can be grouped by various criteria. This will allow better knowledge formalization of problem domain.

Acknowledgments
This research is supported partially by Moldova State Program and the Science & Technology Center in Ukraine (STCU) grant #4035.

References

[1] Negnevitsky, M. *Artificial intelligence: A guide to intelligent systems.* Reading, MA: Addison Wesley, 2002.

[2] Staab, S., Studer, R. (Eds.). *Handbook on Ontologies.* Berlin: Springer-Verlag, 2004.

[3] Dubois, D., Pride, H., Yager, R. (Eds.). *Fuzzy rules in knowledge-based systems.* In Reading in fuzzy sets for intelligent systems. San Francisco: Morgan Kaufmann, 1993.

[4] Leake, D.B. (Ed.) *Case-based reasoning: Experiences, lessons & future directions.* Menlo Park, CA: AAAI Press/MIT Press, 1996.

[5] Gallant, S.I. (1993). *Neural network learning and expert systems.* Cambridge, MA: MIT Press, 1993.

[6] Russel, S., Norvig, P., *Artificial intelligence. A modern approach.* Englewood Cliffs, NJ: Prentice Hall, 2003.

[7] Golding, A.R,. Rosenbloom, P .S., *Improving accuracy by combining rule-based and case-based reasoning.* Artificial intelligence, 87, pp. 215–254, 1996.

[8] Nauck, D., Klawonn, F., Kruse, R. *Foundations of neuro-fuzzy systems.* New York: John Wiley & Sons, 1997.

[9] Hatzilygeroudis, I., Prentzas, J., *Neurules: Improving the performance of symbolic rules.* International Journal on Artificial Intelligence Tools, 9, pp. 113–130, 2000.

[10] Hatzilygeroudis, I., Prentzas, J., *Constructing modular hybrid rule bases for expert systems.* International Journal on Artificial Intelligence Tools, 10, pp. 87–105, 2001.

[11] Popcova, O., Secrieru, Iu., Sologub, D., Jantuan, E., Papanaga, V., *Decision Support System for Ultrasound Diagnostics,* Proceedings of the 1st International Conference of Young Scientists Computer Science & Engineering 2006, Lviv, Ukraine, pp. 30–31, 2006.

[12] Burtseva, L., Cojocaru, S., Gaindric, C., Jantuan, E., Popcova, O., Secrieru, I., Sologub, D., *SONARES – A decision support system in ultrasound investigations,* Computer Science Journal of Moldova, nr. 2 (44), vol. 15, Kishinev, pp. 153–177, 2007.

[13] Secrieru, Iu., Popcova, O., Puiu, S., Sologub, D.,. *Knowledge Structure Modeling in Ultrasound Investigation Domain.* International Conference BIT+ "Informational Technologies-2007". VIIth edition, Chişinău, Moldova, pp. 59–60, 2007.

[14] Secrieru, Iu., Popcova, O., Puiu, S., Sologub D., *Knowledge Structure Modeling in Ultrasound Investigation Domain.* Proceedings

of the 5th European Conference on Intelligent Systems and Technologies (ECIT 2008), Iasi, Romania, CD, Performantica. pp. 1–7, 2008.

[15] Popcova, O., Puiu, S., Secrieru, Iu., *Knowledge structurization and formalization for gallbladder ultrasound investigation domain.* Proceedings of the II International Conference on Advanced Information and Telemedicine Technologies for Health (AITTH 2008), Minsk, Belarus, pp. 98–102, ISBN 978-985-6744-45-0, 2008.

[16] Hatzilygeroudis, I., Prentzas, J., *Knowledge representation in intelligent educational systems.*, Web-based intelligent e-learning systems, Information Science Publishing, pp. 175–192, 2006

Iulian Secrieru

Received May 8, 2009

Institute of Mathematics and Computer Science
of Academy of Sciences of Moldova,
MD2028, Moldova, Chisinau,
str.Academiei 5, off. 337
E–mail: *secrieru@math.md*