Computer Science Journal of Moldova, vol.**1**, no.3(3), 1993

Automatized systems for knowledge acquisition

G.Andrienko N.Andrienko G.Ginkul S.Soloview

The knowledge acquisition is the transmission of the experience for decision of some class of problems from some source to AI system. In a case of Expert Systems (ES) the experienced human expert is the main source of such information. It is necessary to elicite his professional knowledge and to represent it adequately in the knowledge base (KB) of ES. The person which does all these things and in general is responsible for creating of KB is called knowledge engineer.

The creation of Automatized Systems of knowledge acquisition (ASKA) based on the idea that the expert can build KB in the dialogue with computer. By another words, computer must replace knowledge engineer absolutely. Our experience in the field of knowledge acquisition enable us to confirm that today the most of ASKA demonstrate another and non-correct approach to knowledge acquisition, when knowledge representation and inference mechanism are fixed beforehand. In this case, knowledge engineer is forced to articulate his (or her) knowledge by the formal structure language, which he has in its disposal. It is coming out, that ASKA predetermine the decision of some principal questions of AII System technology.

Such situation in the field of ASKA is volunteered by peculiarities of current stage, which could be characterized as the stage of accumulation of information. Current ASKA are automatizated tools for supporting knowledge engineer's work. Potentially they are "usable" only in some typical cases and so knowledge engineer's skills are to profit from all known means. In general, knowledge acquisition is the art and not the formal procedure.

We'll tell about our practice in the ASKA building. In our work we base on two well known ideas: the repertory grid technique (RGT) and technology based on consultation by correspondence. As auxiliary

program tools we use an instrumental Expert System Shells FIACR[1] and ESKIZ[2], which were created in our Laboratory and deal with attribute models of problem domains. We had built program systems for such tasks, as: the diagnostics of gormonal balance of tomatoes, irido-diagnostics (medical diagnostics based on examination of iridescent membrane of eye), monitoring of functioning of Thermo-Electric Station and for some other tasks.

According to our experience we can conclude, that the future of project is often determined on the very first stage, when knowledge engineer elicites and coordinates with expert the set of the right decisions for which the ES is meaned. After that it becomes clear: either knowledge engineer faces with known "good" case or he will forces to go along he's own unexplored way. The common features of such "good" cases are simplicity of its names and possibility to enumerate all decisions directly.

For "good" cases there are some methods, which proved their advantages. The repertory grid technique is one of the most known among them. Its main idea was addopted from experimental psychology [3] and assumes that knowledge engineer shows the expert triples of different decisions and asks him to fulfil a task. Expert must name a charasteristic feature (finaly we used it as attribute with the set of values), which separates three decisions on two similar ones and one different. With the help of that characteristic feature it can be elicited knowledge about properties of decisions. If to speak more formally, the expert's answer is the fragment of semantic network with binary connections.

At the begining, we presumed, that repertory grid was the ideally balanced method. Factually, on the one hand expert has ONLY three variants of separation, and it is not a problem to look over all of them. On the other hand, expert has EVEN three variants of separation, and he has enough freedom to choose acceptable separating features. In its turn, the requirement of similarity for two decisions, in the first place, restricts the field of searching feature, and, in the second place, directs the process of expert's reasoning.

Another advantage of repertory grids is connected with the fea-

tures, which express the qualitative characteristics of decisions. Every such feature defines evident or implied consents about what manifestations of quality are similar and what ones are different. And only expert, basing on problem pecularities, has rights to attribute different manifestations of the feature to the same value of the attribute. For example, the contrast of colors of flowers (yellow, green, purple, etc.) is significant from the point of view of global recognition of sorts of tomatoes. But from the point of view of plant growth (which is caused by the availability of microelements and manganese, in particular), then hues of colors (dark-green, light-green, etc.) play the main role in diagnostics. Thus the repertory grids direction of the searching of similarity and difference literally prompts the expert the approach for the building of scale of values of qualitative characteristics.

The repertory grids allows to elicite only two values of attribute, so we complete it with procedure for eliciting the full spectrum of values. For that we asked the expert about the values of attribute for another decisions.

Gradually we had ascertained the formal lacks of repertory grids. Some of them are very special. For example, in some cases the displayed triples of decisions irritiated the expert with its nonsense. Also sometimes the directive for searching the difference created the state of discomfort for expert. To get over these difficulties we carried out the preliminary genus-species classification of decisions. Such addition allow us to separate evidently discompared decision on different branches of hierarhy and to make easy the next selection of triples.

Another problem was volunteered by non-verbalized character of expert's knowledge. Really, often it is very difficult for expert to articulate his knowledge with natural language. At the same time, RGT demands the expert to formulate the future property exactly and nobody else can make this for expert.

We attempted to make easy for expert the difficult process of searching of acceptable words by dividing it into sub-tasks. At the begining expert must name the property values. Such "rough" names are only for "inward using", they cannot have the real name of the attribute, which is usual, for example, for qualitative attributes. The main circle

of improvement of wording consists three stages [4].

On the first stage expert must formulate interrogative proposition (question) for data answers, which are the "rough" names of values. In the case when finding attribute is qualitative, the name of this quality will appear in the wording of interrogative proposition without fail. On the second stage expert must formulate for each value of the attribute answer to the question in full form. The full proposition has a subject and a predicate. On the third stage expert must elicite the constant (prefix) part of full answers, which can consider as a search definition of the attribute.

Described procedure was welcomed by experts. Such circumstance was a very pleasant surprise for us, because it is known that many peoples prefer to hide his routine work for searching of definitions. Perhaps this is the rare case, when the help of inanimate computer is more preferable than the service of knowledge engineer.

Formal lack of universality of repertory grids was unpleasant surprise for us. For example, if table, chair and bed is the triple of decisions, then RGT cannot elicit their functional purpose as classificatory property. In rare cases such decisions are renaming the scale of values, so there are not any decisions which are similar by that property. Therefore we are forced to refuse a traditional analysis of triples for eliciting of classificatory properties.

It is possible to avoid the described lacks of RGT, if to "dilute" it with another techniques. For example, classificatory properties could be elicited, if to show the expert two decision and to ask him about its similarity and difference [5]. Also often we had additional questions to expert concerning his actions during the concrete comparison of triples. However, when we attempted to explain the expert new task, we often failed, because so-called (by psychologists) "the replacing of goal direction" situation took place.

In such situation the picture could be ideal means for dialoque, because it is possible to express very complex thought. We created some variants of pattern languages [5], extending in some sense the expressive ability of repertory grids. As we thought, if expert and knowledge engineer could reach agreement about using pictures (and

using computer graphics), then they have in their disposal very strong means for pattern dialogue, which allow to formulate strickly complex questions and answers. But in practice we faced so much objections, mostly psychological ones, that to say about something successful is too early.

At the same time, when we deal with problem domains where tasks of choice had appeared, we had successfully used so-named "role-dialogue technique" [2]. It allows us "to hide" the replacing of goal direction by using of some scenario.

For problems of choice (we dealt with the problem of supplying of physical laboratory) we created scenario named "Advertisement agent" allowed to elicite expert's criterions of preference and its decision appraisals. In "Advertisement agent" decisions are named as goods, expert as advertisement agent and knowledge engineer or ASKA as customer. The role without words belongs to competitor: although he keeps the silence, expert must consider him. Before the begining all goods are divided into own (expert's) goods and strange one. The first ones are in range of interests of advertisement agent, the second ones – are produced by competitor. So the advertisement agent attempts to outstrip the competitor and to sell the customer his goods. But the customer is very fastidious: he has doubts and demands explanations.

The dialogue are in the form of question-answer, where customer is asking the agent. In his answers agent tells about another criterion of preference and places appraisals for goods. The "Advertesement agent" includes about 40 different emotional questions. For each of them there is a number of the answers-remarks of agent. One of these remarks allow to avoid the real answer, but others mean the introducing of criterion.

As in RTG the most of questions of customer correspond the comparison of goods-decisions. But buying-sale logic causes principal new situations, which extends the limits of comparison. In particular, questions which "hide" the case of considering only one decision can be constructed. Of course, the advantages of role technique in many respects depend on suitability (naturality) of advertisement agent's role for concrete human-expert. For examination of this circumstance we

had created special control program procedures.

Only the idea of consultation by correspondence could compete with the popularity of repertory grids. Not by chance, there is a number of well known interview techniques based in that idea: modelling of scenario [6], decomposition of goal [7], protocol analysis [8], diagnostic games [9,10] and etc. As known, during the consultation by correspondence knowledge engineer suggests the expert to manisest his skills in the laboratory conditions, by another words, without touching the reality.

The main lack of all interview methods is their monotonous and boring from the expert's point of view. As we were convinced in practice, the expert glad to work with knowledge engineer only during about first 30 minutes. Then interview become in truth infernal employment for expert and effectiveness of his work fall down. We had presented and investigated non-traditional approach to knowledge acquisition, which was called the Expert Games (EG) [11]. The essence of EG idea is to elicite expert's knowledge during the fascination computer game is organized in the investigated problem domain. Such approach allows to automatize the process of the eliciting the expert's knowledge and to make it (process) more attractive and convenient for problem specialist.

First of EG made by us was "Blackbox" [12]. We used it in iridodiagnostic problem domain. But, in general, the game works by crib prepared by knowledge engineer and it isn't conected with concrete domain of use. The set of some known examples of decisions made by expert is used as the capacity of crib. Also knowledge engineer must place into the KB the set of all potential decisions. The stage of preparing examples was the obligate characteristic feature of all games create by us.

The main idea of "Blackbox" based on the desire "to torment" the expert, to insist him to advance suppositions based on incomplete data. Such way makes clear the "inward kitchen" of expert's reasonings. In such case the knowledge engineer's task is to verbalize expert's actions accordind to some of methods of knowledge representation. For analysis of expert's game actions we made special algorithms of game protocol analysis, which finally create formal knowledge structures. Below we

describe "Blackbox" expert game.

At the begining of the game the expert's opponent (his role play computer) "places into the Blackbox" the description of some problem situation and correspondent true decision. Expert must solve hidden decision – that is his game task. Computer can tell the expert about some fact from the situation by the stipulated fee. If he "pays" once again, expert could learn the next fact and e.g. – the game goes on tour by tour. The points are used for payment in the game. At the begining expert has some start points. Rules of the game force the expert to make suppositions (as stakes made on concrete decisions) and what's more – true suggests for successful gaming. Besides that, there is direct profit for expert to recall wrong stakes as soon as possible and so to save at least its part.

The game is finished either when opponent tells about all facts from the description or when expert guesses the decision. At the end of the game computer displays hidden object, then correctly made stake is twiced and its points are added to points, which aren't used up. Incorrect stakes (its points) are lost. Subtraction between start and final points indicates the success of expert's gaming.

If we look through the game's protocol from the point of view of knowledge engineer, then we'll see the next situation. For example, in the third tour the expert knew facts $F1$, $F2$ and $F3$ and put 150 points (from start 1000) on decision $D1$. And futher that stake was decreased. We can suppose, that expert tried to use the next rule:

IF	facts $F1$, $F2$ and $F3$ take place and also condition C,
THEN	$D1$ is the correct decision.

But condition C doesn't appear in further and expert rejects this rule. Of course, if we don't know condition C, it is not possible to put that rule into the KB, but we can keep it in mind for the next. It is interesting, that the value of stake can be considered as information for determining the confidence coefficient of the rule.

During the reasoning expert also eliminates decisions, and it is indicates in the annulation of stakes. For example, in the 4-th tour expert

recalls the stake for decision $D2$. It is means, that enumerated facts and $D2$ form the combination of facts which is not available in the problem domain, as "Flying fish". Such restriction can be represented as the set (in our case as five) of rules.

Further more, we created the number of EG [12], based on different game principles and for different types of knowledge. During its practical realization we faced many interesting problems. So during game protocol analysis we must take into account the game motivation of expert's behaviour. For that we estimated the efficiency of game protocol by checking the function of deviation of it from optimal strategy.

The discussion about EG may be continued and we will devote to it the individual paper. We should remember, that in some situations (for example, when the expert hates games) the EG approach could be useless. Neverthless, in some cases it helps us to fascinate the expert and to mask real goals of dialogue.

References

[1] Soloview S.Y., Soloviewa G.M. The comercial system FIACR. The Handbook "Artificial Intelligence". Vol.1, Moscow, Radio and communication, 1990 (Russian).

[2] Andrienko G.L. Knowledge base forming in the ESKIZ system. Applied systems of AI (Matematical researches, vol.123) Kishinev, Stiinca, 1991 (Russian).

[3] Kelly G.A. The Psychology of Personal Constructs, Norton, New-York, 1955.

[4] Ginkul G.P., Soloview S.Y. The techniques for attribute models of problem domains forming. III AI Conference, vol.1, Tver, 1992 (Russian).

[5] Andrienko G.L., Andrienko N.V. Game comparison procedures in knowledge engineering. III AI Conference, vol.1, Tver, 1992. (Russian)

[6] Slatter Ph. Building Expert Systems: Cognitive Emulation, Ellis Horwood Limited, 1987.

[7] Parsaye K., Chignell M. Expert Systems for Experts. John Wiley & Sons, Inc., 1988.

[8] A.Hart Knowledge Acquisition for Expert Systems. London, Kogan Page, 1986.

[9] Gelfand I.M., Rosenfeld B.I., Shifrin M.A. Structure organization of data in medical diagnostics and forecasting. Questions of cybernetics and medical diagnostics tasks from the mathematician's point of view. Moscow, Nauka, 1985 (Russian).

[10] Gelfand I.M., Rosenfeld B.I., Shifrin M.A. Articles about collaborate working of mathematicians and physicians. Moscow, Nauka, 1989 (Russian).

[11] Ginkul G.P., Soloview S.Y. Approach to knowledge base forming. Automatization of manufacture with using of microprocessors. Kishinev, 1986 (Russian).

[12] Ginkul G.P. Expert Games: a means to acquire the expert knowledge. Journal of computer science of Moldova, Vol.2, 1993.

G.Andrienko,
N.Andrienko,
G.Ginkul,
S.Soloview
Institute of Mathematics,
Academy of Sciences of Moldova,
5 Academiei str.,
Kishinev, 277028, Moldova
e-mail: 23*lsii*@*math.moldova.su*

Received August 23, 1993