CEAI, Vol. 4, No. 1, pp. 65-75, 2002 Printed in Romania

ABOUT THE USE OF INFERENCE STRATEGIES WITH APPLICATIONS IN MEDICAL EXPERT SYSTEMS

Dorina PETRICĂ

"Politehnica" University of Timişoara, Department of Automation and Industrial Informatics

Bd. V. Pârvan no.2 1900 Timişoara

*Tel.*40-56-204333 ext.701, 703, fax *40-56-192049*

e-mail: dpetrica@aut.utt.ro

Abstract: *The inference strategies represent the logical nuclei of expert systems. Those pure mathematical approach makes difficult the access of practitioners involved in the design and use of expert systems, at the interpretation and deciphering of the inference mechanisms. The present paper, dedicated first of all to the practitioners, has as objective the presentation of some easy and engineering techniques of approach, based upon means that are easy to interpret such as the qualitative inference table and the analysis in the state space. The introduced notions are sampled on a segment of an existing expert system, designed for the diagnosis of the isolated paralysis of extraocular muscles. The segment is reported on the gone over at the design of the mentioned expert system by the identification of facts and rules and at the build of the logical model associated to the examination based on the inference strategy "forward chaining".*

Keywords: *expert system, inference strategies, state space, medical diagnosis*

1. INTRODUCTION

The artificial intelligence proposes a specific informatic technology that, by reproducing the reasoning on a artificial way allows the setting up of systems with intelligent character. A product of this technology are the expert systems that, based on the storage of a great number of knowledge and on the solving of difficult, nonspecific problems, are part of the class of systems with pragmatic artificial intelligence. Thus, the expert systems represents one of the top concern of the artificial intelligence, the actual research being directed to the setting up of some models of structuring knowledge associable of some solving methods directed toward the problem[1].

In the case of expert systems based on rules, this orientation leads to taking evidence of a "cognitive-resolutive" ensemble resulted from the association of some structuring methods, applied on rules and facts with adequate inference strategies. The choice of structuring methods, as well as that of inference strategies influences considerable the performances of expert systems, that implies an analysis of inference strategies under different aspects.

A category of frequently used problems by applying expert systems are the diagnosis problems. In a simplified mood the diagnosis might be joined by the plotting in Fig. 1. The application consists in establishing some correspondences between all possible combinations of the values joined with the symptoms and the lots of diagnosis. The inference, the deduction of some solutions based on the observation and knowledge is realized in most cases by the examination of intermediate steps, joined by an intermediate diagnosis structure, wherefrom the final diagnosis is established.

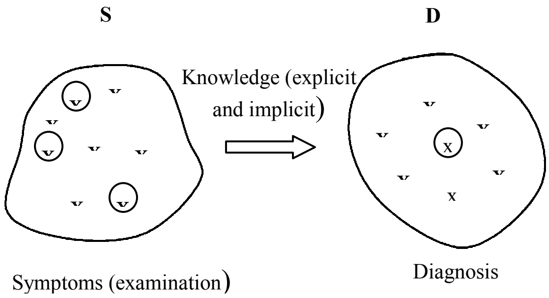

Fig.1. Structural elements associated with the diagnosis of a process

In the medical field, this procedure might be applied while establishing the proper diagnosis (ex.:MED1, INTERNIST/QMR), at the choice of a therapy (ex.:MYCIN, DIACONS), at the interpretation of the laboratory data (ex.:PUFF), at monitoring/supervising (VM) [2],[3], [4], [5]. The performances of medical expert systems, such as the duration of the diagnosis process, the decreasing of the number of symptoms that have to be supervised, the quality of the obtained diagnosis might be improved by the choice of an adequate inference strategy.

The present paper presents some aspects regarding the inference strategies associated to the expert systems based on rules, obtained as a result of a comparative analysis of these strategies, in order to realize some design mechanisms applied on a study case [6]. At the same time there are presented the work tools for the evaluation of the inference strategies and there is made a sampling of their use in the case of a simple medical expert system [6].

2. INFERENCE STRATEGIES

The inference strategies, named also control strategies, assure a correct succession of the inferences, in fact realizing the process of constitution for the inferential chaining, by the successive crossing of some states. The inference mechanism uses rules and facts in order to build dynamic reasonings, the transition from one state to another in the inference process being realized by the starting (execution) of a single rule. This requires moreover the application of a solving strategy of the conflict according to a choice criterion. The usual criteria are: “first rule” (criterion 1), “the rule with the highest priority” (criterion 2), “the rule with the most premises” (criterion 3), “the rule that gives reference to the most recent element added to the fact base”.

2.1. *The FRE-cycle*

The inference process is in fact realized by the repetition of a base cycle formed of three steps, named FRE (filtering-(re)solving-execution). The filtering might be realized on more levels. The first filtering level (I) assures the constitution of fact (F) and rules (R) sets considered as being of interest in the phase of activating the inferential network, that leads to a substantial decreasing of the needed time for the rules processing. The second filtering level (II) assures the choice of rules that might be activated in the current step and their placing into the set of conflicts C. In Fig. 2 there is represented the structure of a FRE cycle with multiple filtering.

The operations performed by the inference mechanism in the filtering steps and in the implementation differ according to the strategy type (Table 1).

Table 1.

Strategy Type	Filtering	Execution
“forward chaining”	Checking the premises match with the facts in the base of facts	Adding the conclusion of the activated rule as fact
“backward chaining”	Checking the conclusion match with the fact mentioned as scope	Adding the premises of the activated rule as scope

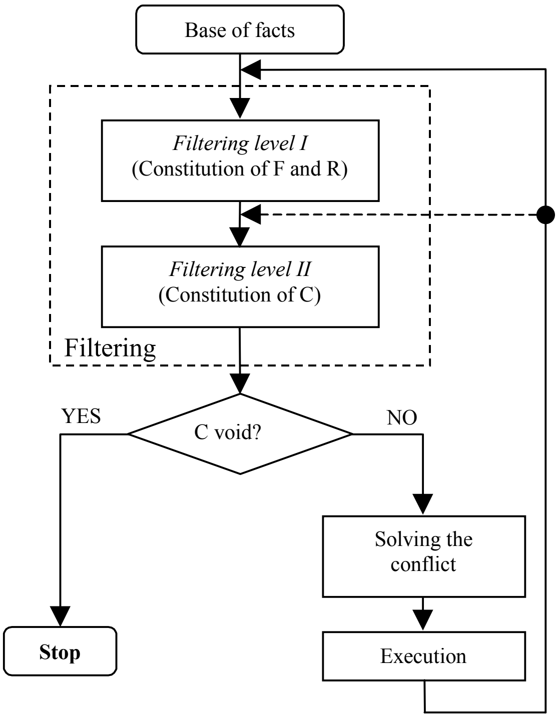

Fig.2. Structure of the FRE cycle with multiple filtering

2.2. *The qualitative inference table*

In order to illustrate the way of interaction between facts and rules out of the problems enunciation in the inference problem of [7] there has been proposed the use of a scheme named qualitative inference table (QIT). Under such circumstances there is considered the case of study represented by the knowledge base **BC1**:

BC1 **R1.** If $\boldsymbol{A}$ then $\boldsymbol{U}$.

R2. If $\boldsymbol{B}$ then $\boldsymbol{H}$.

R3. If $\boldsymbol{D} \wedge \boldsymbol{E}$ then $\boldsymbol{I}$.

R4. If $\boldsymbol{E} \wedge \boldsymbol{F}$ then $\boldsymbol{V}$.

R5. If $\boldsymbol{H}$ then $\boldsymbol{J}$.

R6. If $\boldsymbol{C}$ then $\boldsymbol{K}$.

R7. If $\boldsymbol{I} \wedge \boldsymbol{G}$ then $\boldsymbol{W}$.

R8. If $\boldsymbol{J} \wedge \boldsymbol{K}$ then $\boldsymbol{X}$.

R9. If $\boldsymbol{K} \wedge \boldsymbol{I}$ then $\boldsymbol{Y}$.

R10. If $\boldsymbol{X}$ then $\boldsymbol{Z}$.

R11. If $\boldsymbol{Y}$ then $\boldsymbol{Z}$.

R12. If $\boldsymbol{W}$ then $\boldsymbol{Z}$.

BC1 includes only rules that are brought under the Horn clausal form. The joined QIT has the aspect represented in Fig. 3. A, B… Z are facts, their detail is of no importance.

QIT realises a hierarchy of the rule, from the point of view of the order in which they might be activated and respectively, a hierarchy of the facts that might be used in order to highlight the evolution of the reasoning within the state space. By crossing the table “from the bottom to the top” there result the fact that might be defined as scope of the inference process. These facts are marked with interrupted line. By crossing the table “from the top to the bottom” there result the fact needed in order to demonstrate the truth of the facts defined as scopes. Thus, this table is usable for strategies forward chaining, named hereafter PC strategies (Premises → Conclusions), as well as for strategies backward chaining, named hereafter CP strategies (Conclusions → Premises).

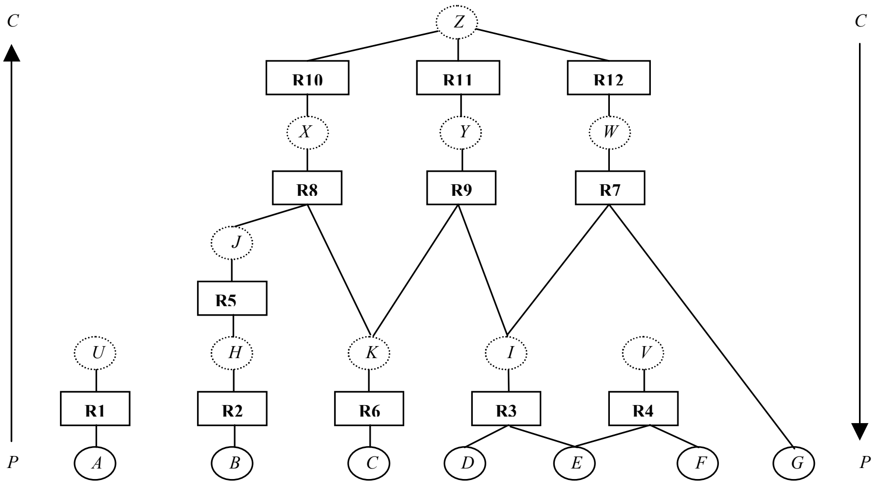

Fig.3. QIT associated to **BC1**

QIT allows a complete visualization of all possible elementary inferences during the inference processes joined to an enunciation. Any inference process represents a succession of elementary inferences extracted from QIT according to an inference strategy

2.3. *PC inference strategies*

The implementation of this strategy in case of a reasoning with implicit scope follows exactly the structure of the FRE cycle in Fig. 2. In order to emphasize this thing, in Fig. 4 the plotting has been considered in the state space for the evolution of the inference process realized with the PC strategy and the criterion 1 of solving the conflict, associated to the knowledge base **BC1** and of the enunciation (**E1**):

(**E1**): "B,C,D,E initial facts"

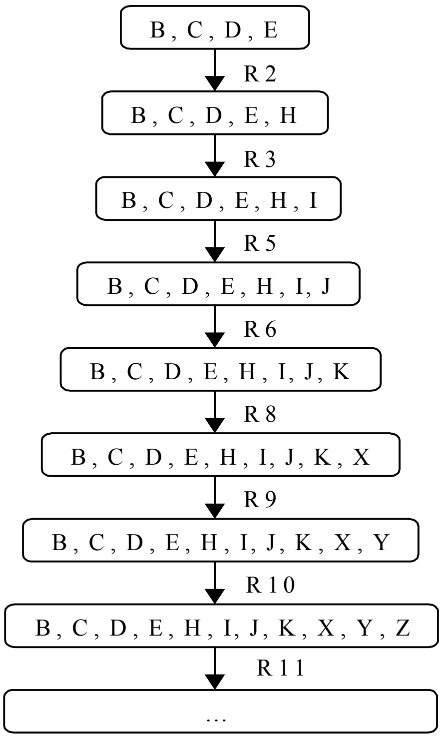

Fig. 4. PC inference for (**E1**) and criterion 1.

The resulted inference process consists of eight transitions, each of them corresponding to a FRE cycle. The seventh transition, realized by applying rule R10, brings the base of facts at the state of *saturated base of facts*. Starting from this reason the eighth FRE cycle might be eliminated. The PC strategy can demonstrate as being true in a non-interactive way, any scope belonging to the saturated base of facts, by the implementation of a number of transitions and the gathering of an informational lumber.

The implementation of the PC strategy in the case of a reasoning with explicit scope might be developed non-interactive, as well as interactive. Thus, for **BC1** and for the enunciation (**E2**):

(**E2**): "B,C,D,E initial facts, Y scope"

the inference process associated to the criterion 4 has the aspect in Fig.5.

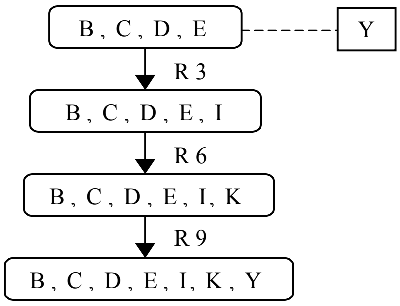

Fig.5. PC inference for (E2) and criterion 4

Comparing the processes in fig. 4 and 5 it results that the facts H, J and X are irrelevant for the deduction of Y, representing an informational lumber. The number of transitions and implicit the informational lumber might be controlled by the correlation of the informations in QIT and the order of facts, respectively the order and prioritization of rules.

The most situations of reasoning correspond to the case in which the scope can not be demonstrated non-interactive. The inference for **BC1** and the enunciation (**E3**)

(**E3**): "B,D,G initial facts, Z scope"

leads to such a case. QIT in Fig. 3 allows the fast identification of deductive facts (H,I), as well as the identification of facts needed to be interrogated for the demonstration of scope Z (C or C,E or E,G).

In order to generate all possible evolutions for the inference process adequate to a complete reasoning, in case that the criterion "the order of the appearance of facts at the discrimination of states" has been abandoned leads to the structure in Fig. 6, named the general table of states GTS. The group of facts B,C,D,E adequate to the initial state has been epitomized in the next states by a point. GTS represents in a compact form, structured on

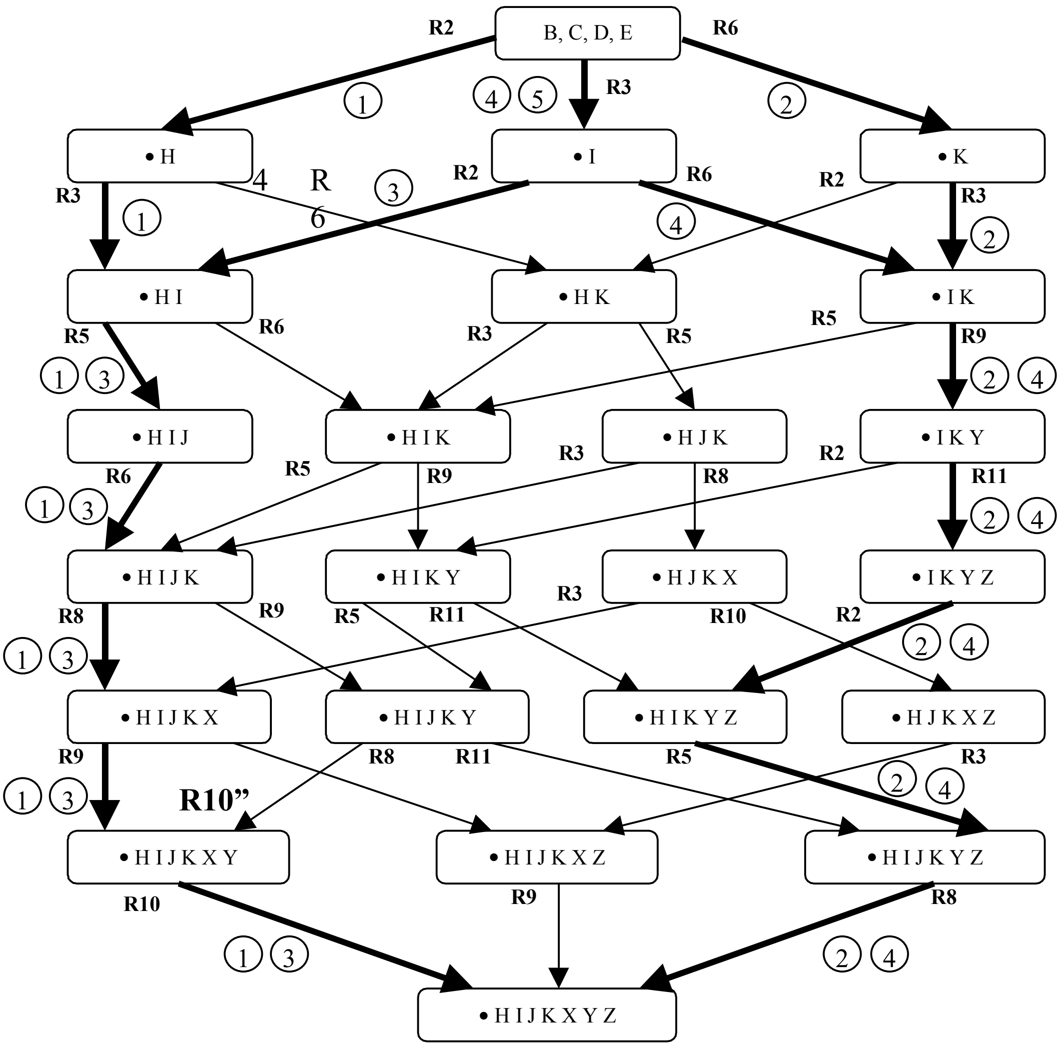

Fig.6. The general table of states associated to the enunciation (**E1**).

levels, all possible transitions from a initial state to a final state, by applying the PC inference strategy. The evolution of inference processes adequate to the four choice criteria has been marked in Fig. 6 by a thick line and the numbers 1, 2, 3, 4 encircled. The GTS structure emphasizes the fact, that the implementation of any inference process is in fact, a process of search, usually directioned, in the state space.

2.4. *The CP inference strategy*

This kind of strategy is usually applied in the situation where the enunciation of the problem comprises a scope that has to be demonstrated. Essentially, the demonstration of such a scope means the aggregation of a group of valid premises, named *primary premises*, that might hypothetically lead, by the application of rules, to the adequate scope. The principle of solving this kind of problem consists in the iterative decomposition of a problem in sub-problems, whose scopes are closer in the inferential chain to the primary problems. The process of decomposition is repeated until there are obtained *primitive sub-problems*, that means sub-problems whose premises are only primary premises. From the point of view of the problems enunciation, the primary premises are of two types:

- nonexaminable facts mentioned by the enunciation of the problem;
- examinable facts, demanded by the expert system in the solving process.

The value of truth for the examinable facts is “true” (valid premise) or “false” (invalid premise), according to the response of the user of expert systems. The primary premises might appear during a solving process and in the sub-problems that are not primitive problems.

The implementation of the *execution* step of the FRE cycle associated to the CP strategy takes different forms according to the logical connective presented in the part "If" of the released rule. Considering for **BC1** and the enunciation (**E4**):

(**E4**): "Z scope"

the plotting by trees AND/OR shown in Fig.7, it results following interpretation: Z might be demonstrated if, by aggregation it might be reached one of the groups of primary premises {B,C},{C,D,E},{D,E,G}. These information correspond exactly to those extracted by the QIT interpretation associated to the base BC1 (fig. 3), QIT being more suggestive, from the point of view of the identification of rules applied at the performing of a FRE cycle and of the place of each elementary inference relatively to a inferential chain.

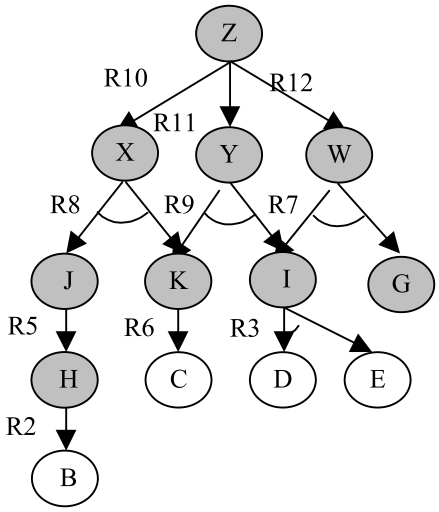

Fig.7. The AND/OR tree for **BC1** and the enunciation (**E4**).

The plotting in the state space of the inference process obtained as a result of a choice strategy in the depth, with return joined to the criterion 1 of choice, for **BC1** and the enunciation (**E5**):

(**E5**): " B,C initial facts, Z scope"

is shown in Fig. 8.

The transitions might be established either by applying a rule, either by operations of "aggregation of primary premises" (APP), either by intermediate operations of "state decreasing" (SD). The presented alternative in Fig. 8 combines these means in the frame of the implementation step of the FRE cycle.

The joining of a heuristic search strategy instead of the search strategy in depth might improve the performances of the CP strategy. The estimation of the tree searching structure is made relatively simple if the QIT afferent to the ensemble of facts and rules is known. The building of the QIT is, at its turn, very simplified if the primary premises afferent to the demonstration of a scope are known. In order to generate primary premises associated to a scope a result of an exhaustive crossing of the search tree afferent to the scope, following the CP strategy by searching in depth without considering the base of initial data has to be obtained. Besides, each time in the FRE cycle there is reached a scope for that the set of releasable rules is empty, that is introduced into a PP set that will store the primary premises.

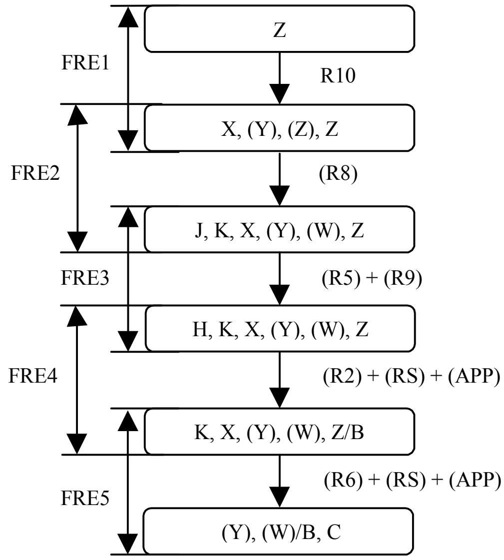

Fig.8. CP inference for (**E5**) and criterion 1.

2.5. *The inference strategy of combined type*

In lots of practical situations, the application of only one inference strategy type does not bring the wished performances for the expert system. Relatively to the QIT structure, that gives all possible elementary inferences, there might be observed that the performances are influenced a lot by the position of the demonstrated scope in the hierarchy. If it is situated on lower levels, i.e. much nearer to the primary premises, it results that the use of a PC strategy is more efficient. If the scope is situated on the higher levels, i.e. it is far from the primary premises and in the same time it possesses more aggregations of the primary premises for demonstration, the CP strategy is recommended. Relatively to the type of facts circulated by the expert system, their performance is influenced by their structuring in facts:

- Known facts (presented in the initial base of facts);
- Unknown, but deducting facts;
- Unknown, but examinable facts;
- Unknown and non-examinable facts.

A solution of compromise between the advantages and disadvantages brought by the two strategies is the inference strategy of combined types. This strategy combines the two mentioned strategies, by the repetition of some inference processes of the type “backward chaining followed by forward chaining” (CP_PC) or "forward chaining followed by backward chaining" (PC_CP). Both alternatives are based on the deduction of the base of rules in two separate and complementary sets, a set of rules where by filtering there is made a checking of the match for the conclusion (R_CP) and a set of rules where, by filtering there is checked the match of premises (R_PC).

3. EXPERT SYSTEM FOR THE DIAGNOSIS OF ISOLATED PARALYSIS OF EXTRAOCULAR MUSCLES

Two of the performance criteria imposed to the medical expert systems of diagnosis do follow the minimization of the duration of the diagnosis process and the number of symptoms that have to be observed. From the point of view of the inference strategies the shortening of the duration and the decreasing of the number of symptoms is obtained by the choice of some strategies that might lead to the finding of the solution after a short number of transitions and by the aggregation of a number of very few primary premises.

In order to sample the elements in chapter 2, there is considered an expert system designed to the diagnosis of isolated paralysis of extraocular muscles, detailed in [6]. The expert system establishes the paralysis of a extraocular muscle according to the analysis of elements of clinical interactive semiology doctor-patient supplied to the system by the implementation of four different phases of examinating the patient:

- Defective position of head and neck (Phase I);
- Repose position of the eyeballs (Phase II);
- The direction of the limitation of the eyeball movement (Phase III);
- Diplopia in its dynamic (Phase IV).

The expert system has been realized in lots of implementation alternative, based upon the CP strategy, as well as on the PC strategy, the process of diagnosis following different functional models associated to the examination. For the mentioned sampling, below, there has been extracted out of this system only aspects regarding the identification of facts and of rules and at the building of a logical model associated to the examination based on the PC strategy.

The process of diagnosis implemented by the expert system has as its fundamentals anatomic-physiological knowledge, as well as procedure and strategic knowledge overtaken from the human expert [8]. The eyes are moved and they are maintained in a physiological repose position by the action of six pairs of extraocular muscles: four straight muscles (external M1, internal M2, superior M3, inferior M4) and two oblique muscles (inferior M5, superior M6). The pairing actions of the six pairs of muscles are schematic given in Fig. 9.

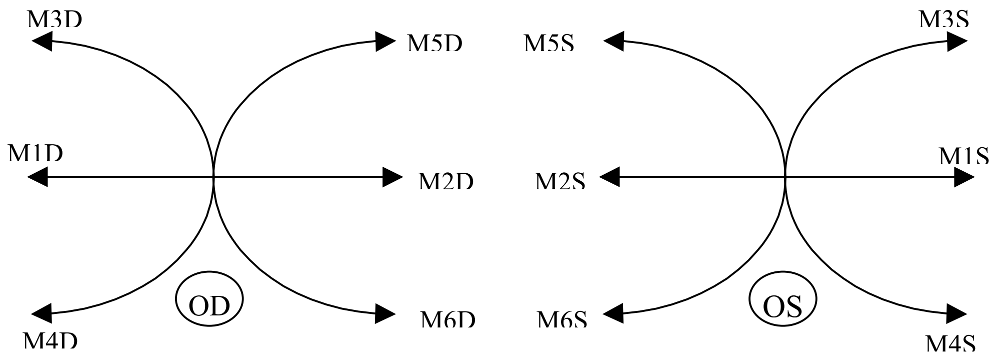

Fig.9. Scheme “pairing actions” of extraocular muscles

The typical concepts for the examination phases have been structured under the form of facts and rules. The inference process realized at the phase level has been analysed with the QIT structure, that made also possible the construction of a logical model of examination.

For sampling, the analysis over Phase I lead to the identification of six facts **F1…F6**:

- Position front right - **F1**
- Position front left - **F2**
- Position raised chin - **F3**
- Position lowered chin - **F4**
- Inclination head to the right - **F5**
- Inclination head to the left - **F6**

And eight rules **R1… R8**:

R1. If **F1** and **F3** and **F6** then **pM3D**.

R2. If **F1** and **F4** and **F5** then **pM4D**.

R3. If **F2** and **F4** and **F6** then **pM6D.**

R4. If **F2** and **F3** and **F5** then **pM5D**.

R5. If **F2** and **F3** and **F5** then **pM3S**.

R6. If **F2** and **F4** and **F6** then **pM4S**.

R7. If **F1** and **F4** and **F5** then **pM6S**.

R8. If **F1** and **F3** and **F6** then **pM5S**

The prefix p in the notation mentions "paralysed muscle".

QIT joined to the rules **R1... R8** has the aspect in fig.10. From the QIT structure it results that **pM3D** and **pM5S** have the same primary premises, such as the group of facts {**F1, F3, F6**}. In a similar way, there might be observed that {**F1, F4, F5**} are the primary premises for the conclusions **pM4D** and **pM6S**, {**F2, F4, F6**} are primary premises for **pM6D** and **pM4S**, while {**F2, F3, F5**} are primary premises for **pM5D** and **pM3S**. Thus, Phase I furnishes no sure result in the process of diagnosis, the complete inference joined to phase I for different groups of primary premises leading simultaneously at the presence of two possible conclusions in the saturated base of facts.

In the same manner there have been finally identified for the four phases either 60 facts and 110 fixed rules, either 60 facts and 44 variable rules [2]. The set of facts builds the set of premises, all of them being examinable facts whose truth values are obtained in the same time with the implementation of phases. The set of conclusions is build of the set of six pairs of paralysed muscles. In case that the truth values joined to the facts as a result of the implementation of each phase assures the aggregation of some groups of primary premises, there might be obtained a part result specific to the phase that competes at the establishing of the final diagnosis. The structure of part results differs according to the phase.

The analysis of joining "primary premises – conclusions" resulted from QIT made for the Phases II … IV, together with the above mentioned, lead to following observations [6]:

- The paralysis of the right extern and right intern muscle might not be established with Phase I.
- The isolated implementation of phases I and IV can not lead to the establishing of a diagnosis. The implementation of those phases must be joined each time by the implementation of some phases that allow the discrimination of the affected eye, that will implicit lead to the choice of one of the two furnished ones by the Phases I and IV.
- If the implementation of phases II and III leads to the aggregation of a single group of primary premises, afferent either to the right eye, either to the left eye, the part result consists in the statement of a single conclusion.
- If the implementation of phases II and II leads to the aggregation of two groups of primary premises, one afferent to the right eye and the other afferent to the left eye, the part result consists in the statement of two conclusions.

According to these observations there have been emphasized in the decisional process joined to the consultation (examination) two types of decisions: *decision at the phase level* with the role of establishing a part diagnosis and *decision at the level of examination,* with the role of establishing the succession of phases during the consultation, as well as the final diagnosis, respectively the cease of consultation. Both decisions might be approached either by PC strategies, either by CP strategies, either by combinations of these strategies.

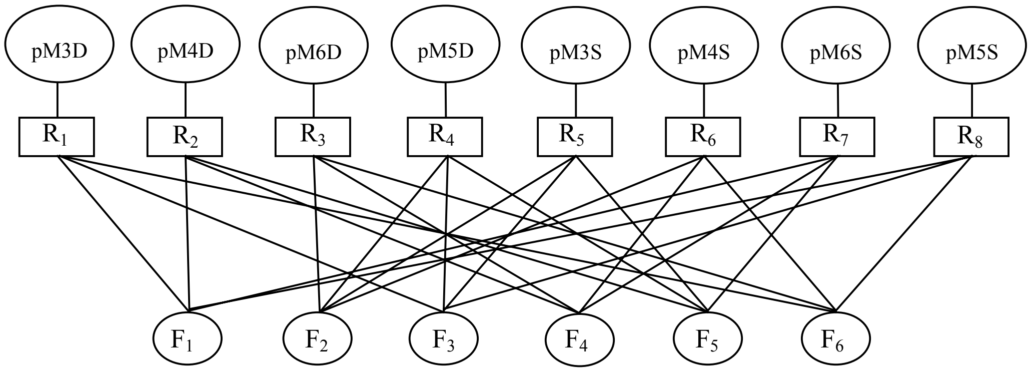

Fig. 10. QIT associated to the rule **R1…R8** belonging to Phase I

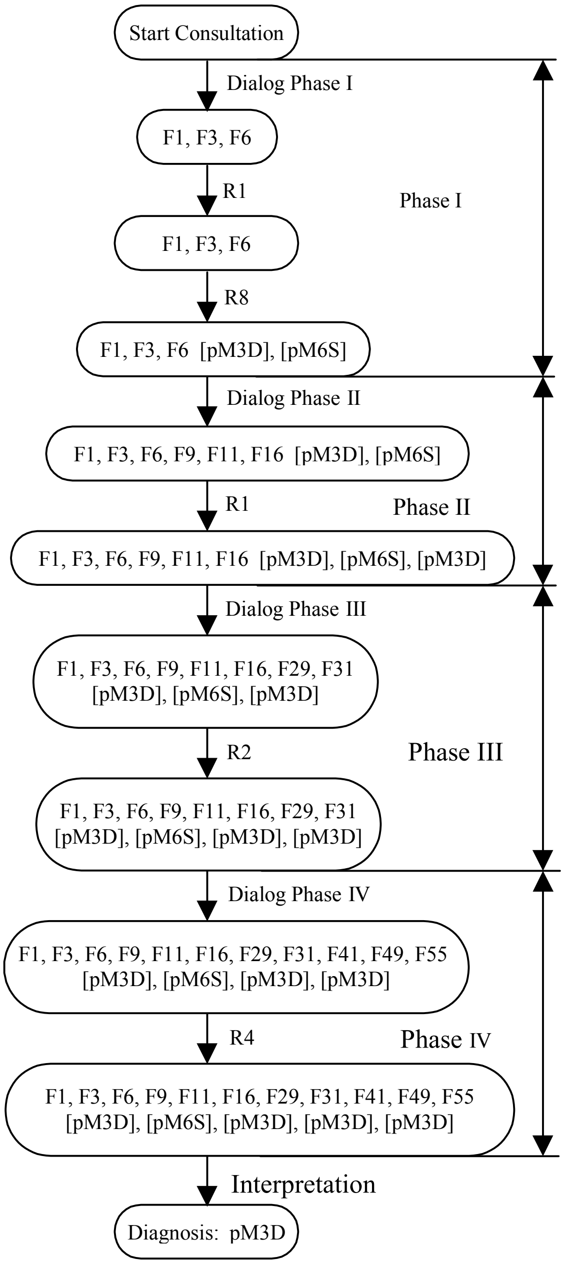

Fig.11. The diagnosis of paralysis M3D.

The evaluation of the number of needed transitions for the establishing of a hypothetical diagnosis might be made with the help of the plotting in the state space. The implementation of a phase through the expert system includes a dialog with the user in order to establish the truth value of examinable facts and the proper inference. This fact is rendered in the state space by two separate states. As example therefore, in the particular case of diagnosis through the PC strategy of the hypothetical diagnosis "paralysis of the right superior muscle of the right eye", by crossing all four phases, in the order I, II, III and IV, all symptoms adequate to the four present phases, the evolution in the space of states has the aspect shown in fig. 11. The considered inference for Phase I in this case, correspond to a complete inference realized according to the PC strategy, criterion 1.

The medical practice considers that the implementation of all phases is not necessary, observing, from this point of view, two diagnosis levels [8]. The level 1 supposes the unfurling of Phases I and II and if the result is conclusive (Phase II confirms the result of Phase I), the doctor can stop the consultation or he might check the diagnosis by implementing Phase IV. The result obtained after Phase II is named "result of level 1". If Phase I and II do not allow the establishing of a conclusive result, then it is implemented Phase III compulsory followed by Phase IV. The result furnished by Phase IV, at its turn conclusive or not, is named "result of level 2". Besides, the implementation of Phase I is optional. This strategic knowledge has been formed into an ensemble of rules **RE-PC1...RE-PC13** integrated in QIT in fig. 12. QIT emphasizes the fact that Phase II is a decisive procedure for the determination of a disease of the extraocular muscles. Thus, if the symptoms specific to Phase II are present and if the decision of not checking any more Phase

II has been taken, the disease might be established as a result of the inference process realized by the activation of rules in the sequence <**RE-PC1, RE-PC3, RE-PC5, RE-PC6, RE-PC8** >.

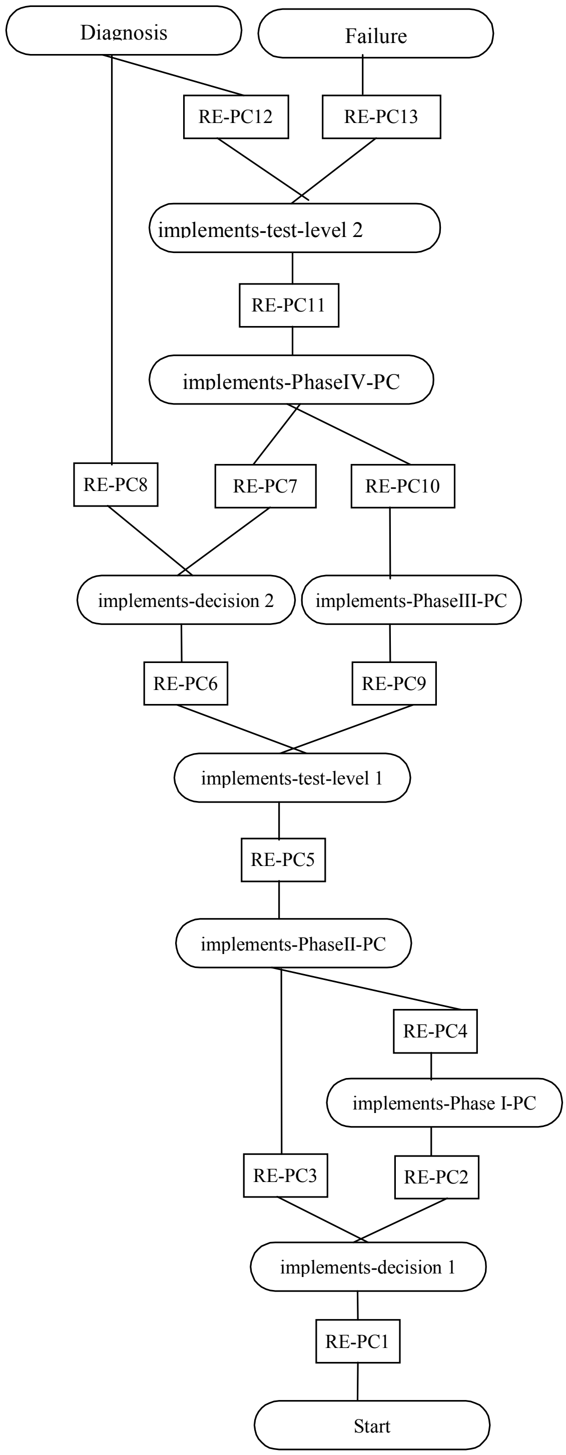

Fig.12. QIT associated to the PC strategy.

The longest process of inference at the level of examination, represented by the sequence <**RE-PC1, RE-PC2, RE-PC4, RE-PC5, RE-PC9, RE-PC10, RE-PC11, RE-PC13**>, is obtained in the case that phases I, II are implemented, the result after Phase II being non-conclusive, fact that determines the implementation of Phases III and IV.

4. CONCLUSIONS

The medical field represents a very good ground for the development of expert systems, because it makes available a very good structured knowledge and it needs decisional processes at each step.

The inferential mechanisms of expert systems yield to a intuitive analysis and systematization in order to facilitate the access at the conceiving of a mixed team formed of experts in the knowledge engineering and experts in the field of applications. That's why the paper presents some tools and elements of suggestive and accessible analysis, as the qualitative inference table, the plotting in the space of states, their use in the application of forward and backward chaining strategies.

The sampling of their use is made by considering a example in the field of extraocular muscles diseases. The chosen example, deliberate simplified, allows an overview of the types of problems and methods of solving, by the involvement of inferential mechanisms.

5. REFERENCES

[1] Puppe, F. - "Systematic Introduction to Expert Systems", Springer-Verlag, Berlin Heidelberg, 1993.

[2] Fieschi, M. - "Artficial Intelligence in Medicine", Ghapman and Hall, 1990.

[3] Pesonen, E. - "Studies on the Experimental Construction of Neural Network/Based Decision Suppot System for Acute Abdominal Pain", *Doctoral dissertation*, Department of Computer Science and Applied mathematics, University of Kuopio, 1998.

[4] Zahlman, G. - "Decision methods for medical expert systems", vde/verlag, Berlin, 1994.

[5] Schill, K. - "Medizinische Expertensysteme", Oldenbourg Verlag

Muenchen Wien, 1990.

[6] Petrică, D. - “Contribuții la aprofundarea mecanismelor inferențiale alesistemelor expert cu aplicații în domeniul medical”, *Teză de doctorat*, 2000.

[7] Petrică, D. and Dragomir, T.L. - “Qualitative Table of Inference (QTI) – a tool for expert system design”, *CONTI 2000*, Buletinul Stiințific, Vol. 45 (59), ISSN 1224-600X, 2000.

[8] Von Norde, G.N. - “Physiology of the ocular movements, binocular vision and ocular motility”, *Mosby Yearbook*, St LouisMo, 1996.