RIGA TECHNICAL UNIVERSITY

Ivars KARPIČS

DEVELOPMENT OF INTELLIGENT COMPUTER METHODOLOGY FOR SYSTEM OPERATION RESTARTING TASKS

Summary of Doctoral thesis

Riga 2012

RIGA TECHNICAL UNIVERSITY
Faculty of Computer Science and Information Technology
Institute of Computer Control, Automation and Computer Engineering

Ivars KARPIČS
PhD student of doctoral study program „Automation and computer technique. Computer control systems, Decision support systems”

DEVELOPMENT OF INTELLIGENT COMPUTER METHODOLOGY FOR SYSTEM OPERATION RESTARTING TASKS

Summary of Doctoral thesis

Scientific supervisor
Dr. habil. sc. ing., professor
Z. MARKOVIČS

Riga 2012

UDK 004.89(043.2)
Ka 643 i

Karpičs I. Development of intelligent computer methodology for system operation restarting tasks. Summary of Doctoral thesis. - R.:RTU, 2012. 44 p.

Printed in accordance with the decision of the board of Institute of Computer Control, Automation and Computer Engineering, Faculty of Computer Science and Information Technology, Riga Technical University No 105 of 26 June, 2012.

This work has been supported by the European Social Fund within the project “Support for the implementation of doctoral studies at Riga Technical University”.

ISBN 978-9934-10-354-4

DOCTORAL THESIS SUBMITTED FOR THE DOCTORAL DEGREE AT RIGA TECHNICAL UNIVERSITY

The defence of the thesis submitted for doctoral degree of computer control systems will take place at an open session on 5th of November, 2012 in 1/3 Meza street, auditorium 202, Riga Technical University Faculty of Computer Science and Information Technology.

OFFICIAL OPPONENTS:

Dr. habil. sc. ing., professor Jānis Osis
Riga Technical University, Latvia

Dr. habil. sc. ing. professor Pēteris Rivža
Latvia University of Agriculture, Latvia

Dr. sc. ing. asoc. professor František Zbořil
Brno University of Technology, Czech Republic

APPROVAL

I confirm that I have developed this thesis submitted for the doctoral degree at Riga Technical University. This thesis has not been submitted for the doctoral degree in any other university.

Ivars Karpičs ……(signature)

Date: ……

The doctoral thesis is written in Latvian and includes introduction, 4 sections, conclusions, bibliography, 9 appendices, 66 figures and 26 tables in the main text, 157 pages. The bibliography contains 211 references.

CONTENTS

INTRODUCTION

Functioning systems and processes are observed in various fields of engineering. For example in technical sciences, engineers design and construct machines in various level of complexity. Engines, turbines, conveyors and other mechanisms are based on the well defined structure. While these mechanisms and systems functionate, structural parameters of them change. In some cases they become even worse and system goes in malfunction state. To re-establish conditions of a correct work it is necessary to reconfigure or even restart the system. In the critical case a full diagnostic procedure should be performed with a goal to find the cause of these nonworking conditions. After the procedure of diagnostics a recovery strategy is defined, which mostly is based on external influences. Correct recovery strategy can restore normal work conditions of the system. Functional processes are observed also in nature and definite processes of physiology could be defined as processes consisting of initial, main performance and ending states. Within the research a methodology for analysis of functioning systems has been developed. The doctoral thesis is devoted to the development of new methods for correct work resumption based on topological modelling, optimization and computer control. The main goal of these methods is to assign an optimal recovery strategy which could re-establish a fully working state of a system. In the thesis an example from physiology and medicine is carried out. A progress of a disease (pathogenesis) is defined as a physiological process containing initial phase, main execution and end stage. Disease is initiated by definite causes of a pathogenesis. Further it evolves by influencing organs and subsystems of human body. At the end clinical symptoms are observed. To re-establish a normal state of a patient (rehabilitate a patient) a drug therapy could be used. A correctly selected drug therapy can eliminate negative effects of a disease and enhance overall health conditions of a patient.

Motivation of the research

The main motivation of the research is described by the necessity to develop a methodology for analysing functioning systems and processes in the case when the available information is heterogeneous and insufficient. The main goal of this methodology is to model the functioning process and re-establish normal conditions of the system. For technical systems mostly various mathematical approaches exist. While, for example, for medical tasks these methods are not always applicable. A inaccessibility of these “classical” methods is explained by the nature of processes from physiology, biology and chemistry. These types of

processes mostly do not have a wholesome analytic description. The thesis is dedicated to solve these types of tasks. As an example, diseases from cardiovascular group are selected. An overall mathematical analyse of multiple cardiovascular diseases is inconvenient because a united methodology to analyze them is not available. Each cardiovascular disease, like arterial hypertension, atherosklerozis, diabetes mellitus and others, is modelled and analyzed by diverse mathematical methods. From this follows that in practice it is hard to assign a therapy for patient having several cardiovascular diseases. The practical necessity of the research is based on the need to develop a medical computer system for the selecting an appropriate healing strategy.

Goal of the thesis

Develop an intelligent computer control methodology for system operation restarting tasks. Examine an application from medicine where optimal drug therapy for patient having cardiovascular diseases should be selected. To achieve this goal several tasks have been proposed and solved.

Tasks:

1. Examine methods for the analysis of functional systems and processes. Define basic characteristics and structural elements of a mathematical model which is used to analyse a system. Examine topological modelling approach as a possible method for constructing a mathematical model in medicine task domain;
2. By using expert enquiry methods develop a topological model for multiple cardiovascular diseases;
3. Process mathematical model with a goal to obtain quantitative and qualitative characteristics which are used to analyse the system and predict further performance;
4. Develop a formal method for the correct performance recovery of a functional system:
   - Define a concept and mathematical description of a recovery strategy;
   - Develop evaluation criteria of a recovery strategy;
   - Develop an approach for selecting an optimal recovery strategy, based on the methods of brute-force search multi-objective optimization;
   - Develop an approach for the synthesis of an optimal recovery strategy which is based on algorithms from metaheuristic class;
5. Implement and verify formal methods in medicine computer system:
   - Formalize and adapt the mathematical model for creating a knowledge base of the computer system;

- Develop decision making module based on the developed formal method;
- Verify the medicine computer system by cases from real patient history. Compare the results with recommendations provided by a doctor and guidelines of the management of hypertension;

Research objects

Research objects are mathematical models of functioning systems, topological models of diseases and processing methods of them. Also intellectual medicine computer systems for the recovery selection are included under research objects.

Research methods

Research methods used in the thesis are based on the analysis of available literature with a goal to develop a united approach for the task to be solved. Mathematical model is created by using expert enquiry and analysis methods. Graph theory for quantitative and qualitative analysis of the model is used. All the required evaluations are based on mathematical analysis and statistical methods. In the case when the solutions space is limited and foreseeable for selecting an optimal recovery strategy, methods of multi-objective optimization are used. If the set of solutions is large or even unlimited then metaheuristic methods (multi-objective genetic algorithms) are used.

Scientific and theoretical novelty of the thesis

1. By using expert enquiry methods a united topological model of arterial hypertension, atherosclerosis and diabetes mellitus has been created. An analysis of the model and cycle detection has been done. Also graph homomorphism and decomposition was applied to expand usage possibilities of the model;
2. Nodes description of the topological model was carried out:
   - Functional level of a node;
   - Importance level of a node. Two methods for computing the importance level of a node by using graph analysis and expert evaluation methods has been developed;
   - Influence level is used to describe how a node is reachable- if it is located in the “middle” of the graph or in the periphery.
3. Efficiency criteria of an external influence (recovery strategy) has been defined:
   - Criterion describing a functional change of the model;
   - Influence rate on the important parts of the model;
   - Influence distribution describes how a recovery reaches periphery of the model;

- Side effects of a recovery strategy;
- Cost ratio of a recovery strategy. The parameter has been evaluated by using pair comparison method (an expert evaluation technique);

4. A method for selecting an optimal recovery strategy has been developed:
   - Seven methods of brute-force search multi-objective optimization has been adapted for selecting an optimal recovery strategy in the case of limited solution space;
   - Three multi-objective genetic algorithms have been adapted for the synthesis of an optimal recovery strategy in the case of a large or unlimited solutions' space.

Practical value of the thesis

The main practical value of the thesis is the developed medicine computer system, which could be used as a decision support system for doctors. The computer system can be used for a treatment selection when one or several cardiovascular diseases (arterial hypertension, atherosclerosis or diabetes mellitus) have been diagnosed.

Approbation of the obtained results

The main results of the thesis are presented in 16 international conferences:

1. 16th IEEE International Conference on Intelligent Engineering Systems 2012 (INES 2012), (13.-15.06.2012, Lisbon, Portugal);
2. Asemundus contact seminar for higher education representatives (15.-16.05.2012, Seoul, South Korea);
3. The 3rd United World Congress of Latvian Scientists and the 4th Letonika (Latvian Studies) Congress (25.10.2011, Riga Technical University, Riga, Latvia);
4. RTU 52. International scientific conference, subsection „Technologies of Computer Control" (13.10.2011, Riga Technical University, Riga, Latvia);
5. 9th IEEE International Symposium on Intelligent Systems and Informatics (8.-11.09.2011 Subotica, Serbia);
6. 9th IEEE International Symposium on Applied Machine Intelligence and Informatics (27.-29.01.2011, Smolenice, Slovakia);
7. FAIRS'10 - Forum for AI Research Students (13.12.2010, Cambridge University, Cambridge, England);
8. 14th International Biomedical Engineering Conference (28.-29.10.2010, Kaunas Technical University, Kaunas, Lithuania)- diploma;
9. FAIRS'09 - Forum for AI Research Students (14.12.2009, Cambridge University, Cambridge, England);

10. 13th International Biomedical Engineering Conference (29.-30.10.2009, Kaunas Technical University, Kaunas, Lithuania);
11. RTU 51. International scientific conference, subsection „ Technologies of Computer Control" (13.10.2010, Riga Technical University, Riga, Latvia);
12. 12th International Biomedical Engineering Conference (23.-24.10.2008, Kaunas Technical University, Kaunas, Lithuania);
13. RTU 50. International scientific conference, subsection „ Technologies of Computer Control" (15.10.2009, Riga Technical University, Riga, Latvia);
14. 2th International Conference on advanced information and telemedicine technologies for health (1.-3.10.2008, National Academy of Sciences of Belarus, Minsk, Belarus);
15. RTU 49. International scientific conference, subsection „ Technologies of Computer Control" (13.10.2008, Riga Technical University, Riga, Latvia);
16. RTU 48. International scientific conference, subsection „ Technologies of Computer Control" (12.10.2007, Riga Technical University, Riga, Latvia);

All results of the thesis are published in 13 articles in scientific proceedings, international journals and other publications:

1. Karpics I., Markovics Z., Markovica I. Topological Modelling as a tool for analysis of functioning systems. Intelligent Systems: Models and Applications (Topics in intelligent engineering and informatics (edited by E. Pap, Editors-in-Chief: J. Fodor, I. J. Rudas), Springer Verlag – in print;
2. Karpics I. Personalized therapy selection by using multi-objective optimization. Proceedings of 16th IEEE International Conference on Intelligent Engineering Systems 2012, ISBN 978-1-4673-2694-0, Lisbon, Portugal, 2012, pp. 537- 542 (IEEExplore, SCOPUS)
3. Karpics I. A comparison of Hypertensive Therapies' Estimated Costs by Using Expert Evaluation Methods. Scientific Journal of Riga Technical University, Computer science, series. 5, vol. 48., ISSN 1407-7493, Riga, Latvia: RTU, 2011, pp 30-35 (EBSCO, ProQuest, Versita, VINITI)
4. Karpics I., Markovics Z., Markovica I. Composition of United Multiple Diseases Evolution Topological Model. Proceedings of Intelligent Systems and Informatics (SISY 2011): IEEE 9th International Symposium, ISBN 978-1-4244-7429-5, Subotica, Serbia, 2011, pp. 65-69 (IEEExplore, SCOPUS)
5. Karpics I., Markovics Z. Development and evaluation of normal performance recovery method of a functional system. Scientific proceedings of 9th IEEE International

Symposium on Applied Machine Intelligence and Informatics (SAMI 2011), ISBN 978-1-4244-7429-5, Smolenice, Slovakia, 2011, pp. 171-175 (IEEExplore, SCOPUS)

6. Karpics I., Markovics Z. Development of pathogenesis topological model node evaluation complex. Proceedings of 14th International Conference "Biomedical Engineering", ISSN 2029-3380, Technologia, Kaunas, Lithuania, 2010, pp. 163- 166
7. Karpics I., Markovics Z. Development of improvement complex influence assessment**.** Scientific Journal of Riga Technical University, Computer science, series. 5, vol. 42., ISSN 1407-7493, Riga, Latvia: RTU, 2010, pp. 31-37. (EBSCO, ProQuest, Versita, VINITI)
8. Karpics I., Markovics Z. Improvement and Assessment of the Effective Therapy Selection Method. Proceedings of 13th International Conference "Biomedical Engineering", ISSN 2029-3380, Technologia, Kaunas, Lithuania, 2009, pp. 219-223
9. Karpičs I., Markovičs Z. Extension of Pathogenesis Topological Model and Processing Methods. Scientific Journal of Riga Technical University, Computer science, series. 5, vol. 39., ISSN 1407-7493, Riga, Latvia: RTU, 2009, pp. 43-49 (EBSCO)
10. Karpics I., Markovica I., Markovics Z. Most effective two therapies combination detection approach. Proceedings of 12th International Conference "Biomedical Engineering", ISSN 2029-3380, Kaunas, Lithuania: Technologia, 2008, pp. 234-236
11. Markovics Z., Karpics I., Markovica I. Arterial Hypertension Therapy Selection by Using Topological Modeling and Production Law Logic, Proceedings of AITH'08-Advanced information and telemedicine technologies for health. ISBN 978-985-6744-45-0, Minsk, Belarus: National Academy of Sciences of Belarus, 2008, pp. 54-58
12. Karpičs I., Markoviča I., Markovičs Z. Method for detection and estimating the combination of system correction. Scientific Journal of Riga Technical University, Computer science, series. 5, vol. 35., ISSN 1407-7493, Riga, Latvia: RTU, 2008, pp. 55-62
13. Karpičs I., Markovičs Z. Calculation method of final evaluation in the higher education system. Scientific Journal of Riga Technical University, Computer science, series. 5, vol. 32., ISSN 1407-7493, Riga, Latvia: RTU, 2007, pp. 34-43

The structure of the thesis

The doctoral thesis consists of Introduction, 4 sections and Conclusion. In the *Introduction* of the thesis the problem sphere and actuality of the topic is introduced. Based on this actuality a main goal and tasks of the thesis are defined. Also in the introduction research objects, methods and the achieved novelty of science is described.

1. Section is devoted for the development of a mathematical model for functioning systems. At the beginning of the section a survey and formal aspects of topological modelling are described. Then a topological model for multiple diseases has been created;

2. Section is based on the analysis and processing of the developed topological model. At first decomposition is performed to provide detailed views of the model. After the decomposition a cycle detection procedure is performed. Within the section three parameters of topological models' elements are introduced. For the evaluation of an external influence (in diseases case: a therapy) five performance criteria are used. Further these criteria are used as an input information in decision making process;

3. Section is devoted to the method of an optimal solution's selection. Within the section two subtasks are solved. In the first one a full search multi-objective optimization methods are used and in the second case multi-objective genetic algorithms are applied;

4. Section covers a description of a medical computer system which is created as an implementation of the proposed method. A literature review of available medicine systems for similar tasks is carried out. The structure of the system is based on the typical structure of an expert system. The performance and precision of the system is evaluated by several tests. Tests also include the analysis and comparison of the results provided by the system, a doctor and by the overall treatment guidelines.

In the *Conclusion* a summary, results and the goals of the further research are provided.

The outline of the thesis: 157 pages, 66 figures, 26 tables. Bibliography consists of 211 records and thesis has 9 appendices: 1. Functions of organism subsystems, 2. Side effects of therapies, 3. Topological model in Powersim software, 4. Tables of initial values of nodes; 5. Parameters of organism subsystem's nodes; 6. Parameters of side effect nodes, 7. Link parameters, 8. Therapy cost comparison, 9. Test example of the computer system.

1. TOPOLOGICAL MODEL FOR MULTIPLE DISEASES

Functioning processes are observed in nature, society, technology and other real- world domains. Each process involves characteristics, features and other structural elements, which all are mutually connected via cause-effect relations. Various approaches and methods exist to describe and analyse the functioning of a system. One part of these methods are based on statistical methods and approaches of mathematical modelling. While statistical computation methods are inflexible, use voluminous processing resources and do not always provide a solution, mathematical modelling and heuristic approaches are flexible with regard to changes in the input of information and can be adjusted to an acceptable solution [Zhou2002]. The main advantage of the heuristic approaches, however, is that they guarantee an acceptable solution regarding the stated preciseness, constraints and evaluation of the final solution.

During the functioning process, all systems lose their resources and working capacity, which leads to overall malfunction. The state of a normal functioning needs to be re-established by adding external influences. These external influences mostly are formed like recovery strategies. A correctly selected strategy resumes the state of normal functioning. When using a mathematical modelling and employing a heuristic approach, a recovery strategy is selected in three steps [Zbign2004] (Figure 1.1.):

1. Full investigation of a functioning system by using expert inquiry methods: summary, formalisation and arrangement of all available knowledge in a common knowledge base;
2. Development of the mathematical model;
3. Recovery strategy selection by using the created mathematical model and decision making techniques.

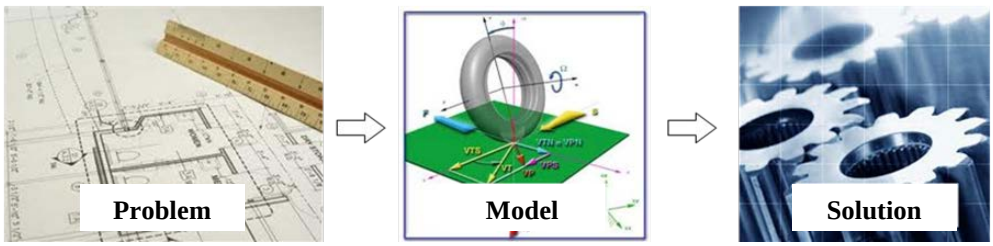


Figure 1.1. Selection of recovery strategy by using mathematical modelling

A mathematical model represents an abstract, simplified mathematical construction that reflects reality and is created for the defined research purposes. Such a model is described by a group of characteristics and logical regularities and defines an idealised performance of the analysed system. A model connects input/ output information, external influences and

disturbances in one united mathematical abstraction [Bender1978]. The main requirements for a model are:

- The model includes all available knowledge about the analysed functioning process;
- It is developed by knowledge engineers and expert system developers;
- The model needs to be capable to include various types of knowledge (quantitative, linguistic, rule based and other types);
- The construction of the model has to be unsophisticated and foreseeable/predictable, but at the same time it needs to be detailed enough to manage projective research tasks;
- Structural changes of the system can be easily added and implemented in the model;
- The mathematical calculus of the model is unsophisticated and can easily be implemented in the expert system.

One of the mathematical methods for analysing a functioning system is the topological modelling. This method is widely used for the analysis of various systems. Different applications for technical and electrical systems, information systems, biochemical and biological processes, robot control and other systems have been created. Lately, a fast progress of TM for software development and Model Driven Architecture (MDA) applications can be observed.

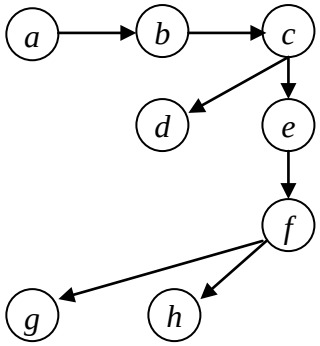
The fundamentals of the topological modelling are based on the assumption that a complex functioning system can be described by abstract concepts like the topological space. The principle of the system's functioning, i.e. of the functional mathematical model, is represented by a topological space in the form $T= (X, \Theta)$. X is a functional characteristic space $X = \{x_1, x_2, ..., x_l\}$ and Θ is a topology that describes the relationships between these functioning characteristics in a binary form. The topology [Alex1998] in the set X is each system Θ that consists of open subsets A, and which satisfy two Kolmogorov axioms: (1.1. and 1.2.), [Kolm1957, Kolm1961]:

- Set X belongs to the topology $\Theta(X \in \Theta)$ and also empty set belongs to the topology $(\varnothing \in \Theta)$; (1.1)
- Each split or union of the subsets belongs to the topology:

$$\forall_\eta \left(\bigcup_\eta A_\eta \in \Theta\right); \ \forall_\varphi \left(\bigcap_{\varphi=1}^{k} A_\varphi \in \Theta\right). \quad (1.2)$$

The topological model is created in the space of all systems' characteristics, but it also includes information about the structure of the system (defining the topology). One of the solutions how to represent the topology is the description in X open subsets. After the

collection of all available information, all characteristics are defined and the structure of the topology is known. As the number of characteristics is a finite, it is possible to represent the topology in an oriented graph (i.e. arcs have directions) of the form *G(X, U)*, where *X* is the set of the functioning characteristics and *U* is the link set (Figure 1.2.a). Mostly, in calculations and computer applications an incidence matrix is used (Figure 1.2.b).



a

	a	**b**	**c**	**d**	**e**	**f**	**g**	**h**
a	0	1	0	0	0	0	0	0
b	0	0	1	0	0	0	0	0
c	0	0	0	1	1	0	0	0
d	0	0	0	0	0	0	0	0
e	0	0	0	0	0	1	0	0
f	0	0	0	0	0	0	1	1
g	0	0	0	0	0	0	0	0
h	0	0	0	0	0	0	0	0

b

Figure 1.2. Representation of a topological model in graph (a) and incidence matrix (b) form

In the process of the model construction it is assumed that the system is closed and independent from the environment. Accordingly the graph is bounded, which defines logical chains within the model. The model allows analysing the performance of the functioning system in the normal and in the incorrect state, considering a cyclic or non-cyclic performance. For a large scale and hierarchical system, a structural distributed form can be used. In the construction of a large scale model it is possible to use various methods for an extraction and combination of separated models.

1.2. Topological modelling for medicine tasks

One of the main advantages of the topological modelling is the possibility to use various types of knowledge that can be expressed in numerical, functional or non-analytic form as well as to include them into the model. As a result, the model can be used to mathematically describe a heterogeneous system in a situation when the available information is insufficient. Physiological, biological and chemical processes often have complex structures. They are complex because their features are not structurally even and it is not possible to use differential or other strict type of the mathematical equations. In the field of the medicine, pathogenesis has been defined as a functioning process, which consists of complex physiological processes that are interrelated with each other and arise when the organism is affected by pathological exogenous or endogenous factors. These processes emerge in the form of clinical syndromes and symptoms. The topological model of a pathogenesis is created by summarizing the knowledge of medics and the overall used treatment keynotes of a

disease. To analyze a process of the pathogenesis physical type, models of the graph structure can be used. Miscellaneous topological models for differential diagnostic, diagnostic parameter selection, prediction of a diseases' state and therapy selection have been created. To perform all these tasks one common type of a topological model is used and slightly adjusted to the each subtask. It is possible to adjust the topological model for modelling and analysing a state of disease by defining all structural elements of the model.

1.3. Elements of topological model

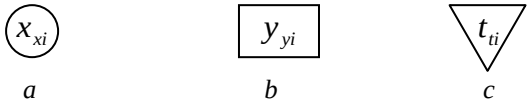
Topological model is described as a graph, containing nodes. Graph nodes describe occurrences, processes, parameters and other features of a pathogenesis:

- The circle nodes *x* are used to describe the organism subsystems, also called as pathogenesis basic mechanisms (Figure 1.3.a, an example in Table 1.1.). A node describes each organism subsystem and its functioning level, which is involved in the process of pathogenesis. All *x* nodes form a set $X = \{x_1, x_2 ... x_{xi} ... x_{xn}\}$;

Table 1.1.

Node	Title
x_1	Hyperactivity of hypotholamus zone
x_2	Hyperactivity of sympathetic nerve system
x_3	Arteriol hyper alfa energy
x_4	Heart hyper beta energy
x_5	Rise of heart frequency
...	

- The triangle nodes *t* defines external treatments, which are used as recovery strategies (Figure 1.3.c). In the pathogenesis example, recovery strategies are medicament groups. All medical treatments *t* form a set $T = \{t_1, t_2 ... t_{ti} ... t_{tn}\}$;
- The square nodes *y* describes side effects of a therapy (Figure 1.3.b) and they are collected in a set $Y = \{y_1, y_2 ... y_{yi} ... y_{yn}\}$. Frequently therapy side effects overlap and create a united set of them.



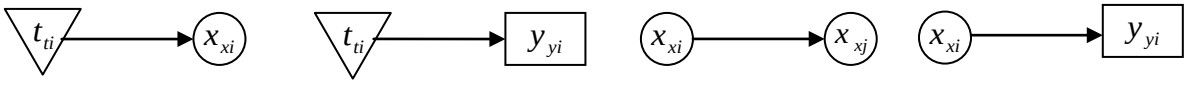
a *b* *c*

Figure 1.3. Node types of topological model

Links show cause-effect relations between graph nodes, where one node by changing its functioning level initiates a performance change of the other node. The link between two nodes can express a linear or non-linear correlation between these functioning levels. The

analytic expressions and the production law (If..., Then...) rules can also be used to describe the essence of the link. In the process of the topological model development, the link definition is a time consuming procedure and it has the highest importance because links define the topology of the model, i.e. the structural behaviour of the system. Within the process of a link definition a large amount of physiological, chemical and biological regularities are taken into account. Usually, interrogatory methods of an expert group are used [Mark2009]. The model can include four different link types:

- A link between therapy t_{ti} and node x_{xi} defines the therapy influence on the basic mechanisms of a pathogenesis (Figure 1.4.a). The link is used to connect the overall process of the pathogenesis with the external influence (therapy);
- A link from therapy t_{ti} and side effect is used to describe an association between the dose of a therapy and the corresponding side effect (Figure 1.4.b);
- A link between two organism subsystem nodes x_{xi} and x_{xj} (Figure 1.4.c). This link is used to describe a basic mechanism of a pathogenesis by using a functional expression. Each link is unique and has an individual description of the relationship between these two nodes;
- A link from x_{xi} to y_{yi} (Figure 1.4.d) is used to describe the connection between the organism subsystem and the therapy side effect. The probability of the existence of this link is low, but in the real-life practice this occurrence is observed.



a *b* *c* *d*

Figure 1.4. Link types of the topological model

Model development steps:

1. Knowledge acquisition by using expert evaluation methods;
2. Definition of *x* nodes starting from cause aspects of a disease till clinical symptoms. Cause- affect link evaluation and definition;
3. Enclosure of external influences: therapy *t* nodes. All links from therapies to subsystem nodes are defined;
4. Definition of quantitative and qualitative measurements of all nodes and links:
   - Each node has its initial functioning level;
   - Each link has a corresponding weight or influence ration. It defines how first node of a link influences the second node.

Characteristics, functions and solved tasks of the topological model:

1. To create a model it is not necessary to structure available information about the system. All knowledge should not be homogenous (i.e. same format);
2. The pathogenesis model includes a symptomatic level and it is created till the level of important mechanisms of a pathogenesis;
3. A model in medicine allows solving a diagnosis task of a pathogenesis;
4. A model can be used to perform a modelling of an external influence (in pathogenesis case: therapy) and to evaluate efficiency of this influence;
5. After the modelling process, ending state of the process, functional and structure parameters are known;
6. By using the pathogenesis model it is possible to model a complex[1] of multiple therapies with a goal to estimate an optimal therapy for the patient.

1.4. Detailed definition of structural elements

Each node of the model has an initiate functional level, which for every individual human is different. Altogether, these functional levels define the clinical conditions of the patient. Model nodes use the following functioning indicators:

- α_{xi}: Functioning level of the node x_{xi}. The value is given in the range [0, 1], where 1 indicates a fully functioning and 0 describes malfunction. The initial value is evaluated by a doctor by using value definition table. Table 1.2. describes an example of a functional level of the organism subsystem "Rise of systolic blood pressure". In this case, the functioning value needs to be entered by analyzing systolic blood pressure (SBP). For example, if SBP is around 160 mm/Hg, then the value of the node x_8 is 0.6;

Table 1.2.

Functional level definition

ID	x_8
Name	Rise of systolic blood pressure
Evaluation	**SAS mm/Hg**
	$\geq$ 200 - 0.2
	180 - 0.4
	160 - 0.6
	140 - 0.8
	$\leq$ 120 - 1.0

[1] Therapies complex is a multiple drug combination which is used as a treatment for the patient. In the following when referring to 'therapies complex' it will be used the term 'therapy'.

- β_{yi}: Functioning level of the side effect y_{yi}. Also this value is in the scale [0, 1] and given by a doctor;
- τ_{ti}: Therapy t_{ti} dose. Quarter ¼, half ½ and full doses (i.e. values) are used in equations. The level of a dose does not change during the modelling an analysing a pathogenesis.

The existence of a link describes a cause effect relation between two nodes and each type of links defines different influence mechanism. In the topological model four links with different influence mechanisms are used:

1. A link describing a therapy influence on the organism subsystem (Figure 1.3.a). The change of functional level of the node x_{xi} is depended from t_{ti} dose and it is evaluated by (1.3.)

   $\alpha_{xi}^{b} = k \cdot (\tau_{ti}^{b} - \tau_{ti}^{(0)}) + \alpha_{xi}^{(0)}$, where (1.3)

   τ_{ti}^{b} - Therapy dose [¼, ½ or 1];

   $\tau_{ti}^{(0)}$ - Used therapy in the previous period [0, ¼, ½, and 1]. If the therapy is used first time then this parameter is 0;

   k - Influence ratio evaluated by (1.4.).

   $$\begin{cases} k = I_0, \text{ if } 0 \le \alpha_{xi}^{(0)} \le \alpha_{xi}^{kr}, \\ k = \dfrac{1-\alpha_{xi}^{(0)}}{1-\tau_{ti}^{(0)}}, \text{ if } \alpha_{xi}^{kr} < \alpha_{xi}^{(0)} < 1. \end{cases} \quad (1.4)$$

   If the initial functioning level of the node x_{xi} is larger than α_{xi}^{kr}, then by applying a full therapy t_{ti} it is possible to normalize the organism subsystem x_{xi}. Contrary, if the functioning level is under the critical value, then resumption is not possible;

2. A link from the therapy t_{ti} to the side effect y_{yi} (Figure 1.3.b). The functional relationship between the nodes is defined as an analytic expression $\beta_{yi}^{b} = f(\tau_{ti})$, which is specified in (1.5.) and (1.6.)

   $\beta_{yi}^{b} = \beta_{yi}^{(0)} - k \cdot (\tau_{ti}^{b} - \tau_{ti}')$, where (1.5)

   $$\begin{cases} \text{if } 0 \le \tau_{ti}^{b} \le \tau_{ti}', \text{ then } \beta_{yi}^{b} = \beta_{yi}^{(0)}, \\ \text{if } \tau_{ti}' < \tau_{ti}^{b} \le 1, \text{ then } k = \dfrac{I_0}{1-\tau_{ti}'}. \end{cases} \quad (1.6)$$

3. An in-between relation between two organism subsystems (Figure 1.3.d). Topological model consists of large amount of organism subsystem nodes which are linked by cause-affect links. A logical chain between nodes defines a mechanism of a disease,

i.e. the character of a pathogenesis. The functional relation is described by equations (1.7) and (1.8).

$$\alpha_{xk}^{b} = k \cdot (\alpha_{xi}^{b} - \alpha_{xi}^{(0)}) + \alpha_{xk}^{(0)}, \tag{1.7}$$

$$\begin{cases} \text{if } 0 \le \alpha_{xk}^{(0)} \le \alpha_{xk}^{k}, \text{ then } k = I_0, \\ \text{if } \alpha_{xk}^{kr} < \alpha_{xk}^{(0)} \le 1, \text{ then } k = \dfrac{1-\alpha_{xk}^{(0)}}{1-\alpha_{xi}^{(0)}}. \end{cases} \tag{1.8}$$

4. Relation between the organism subsystem and side effect (Figure 1.3.d). This link describes the situation when the functioning level change of an organism subsystem node causes a side effect. By the influence of the node x_{xi} the functioning level of y_{yi} is evaluated by (1.9) and (1.10.).

$$\beta_{l}^{b} = k \cdot (\alpha_{i}^{b} - 1) + \beta_{l}^{(0)}, \tag{1.9}$$

$$\begin{cases} \text{if } 0 \le \alpha_{i}^{b} \le 1, \text{ then } \beta_{l}^{b} = \beta_{l}^{(0)}, \\ \text{if } 1 \le \alpha_{i}^{b} \le 2, \text{ then } k = I_0 < 0. \end{cases} \tag{1.10}$$

1.5. Development of pathogenesis topological model

Within the framework of the research, a model for united arterial hypertension (AH), atherosclerosis (AS) and diabetes mellitus (DM) has been developed. These three diseases have been selected consciously, because they show common manifestations, has a similar progress mechanism and they partly involve equal organism subsystems. They are mutually connected and frequently one disease provokes the other one and they cannot be divided. The AH model is based on the results of the previous research and on developed medicine computer systems [Mark2000b, Mark2002a, Karp2008a, Karp2008b]. The model is edited, simplified and used for the construction of the united pathogenesis model. In contradiction to the previous AH model, in which diagnose of the pathogenesis is the main task to be solved, the contemporary model is used to provide information for the selection of optimal therapies in the situation, when multiple diseases appear. By adding new AS and DM models, the united pathogenesis model has been created (Figure 1.5.) to solve new tasks:

- The united pathogenesis model is created to investigate mutual relations between three diseases;
- The model can be used to predict the influence of a therapy;
- The efficiency of each therapy is evaluated by performance criteria. These criteria are the basic information for the optimal selection of a recovery strategy.

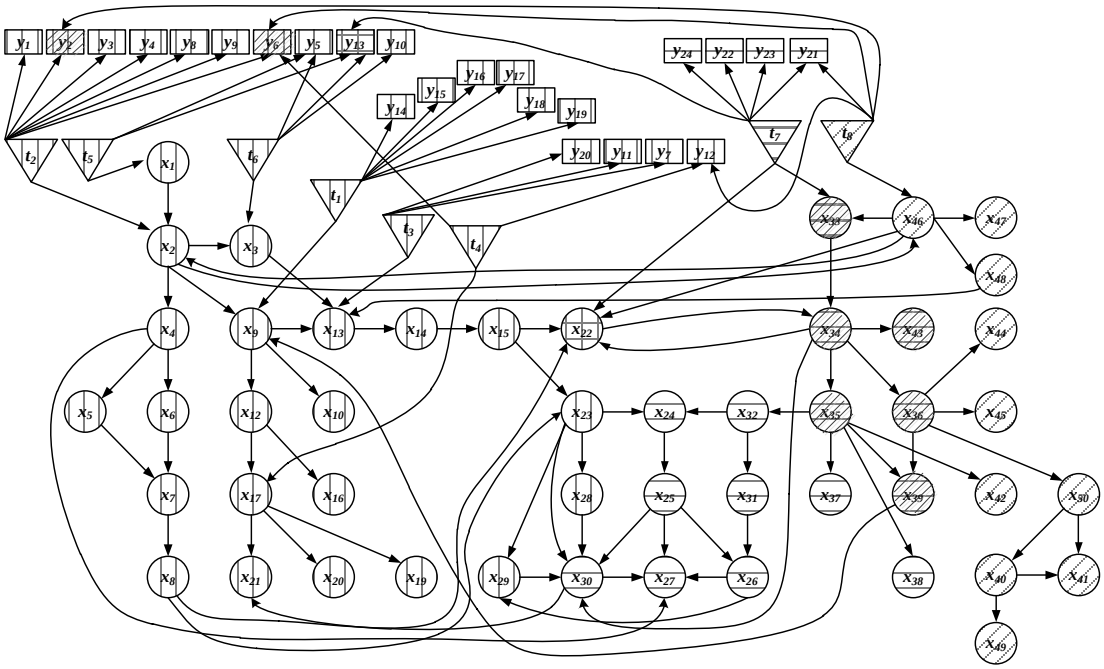


Figure 1.5 Topological model of arterial hypertension, atherosclerosis and diabetes mellitus

2. PROCESSING AND ANALYSIS OF MATHEMATICAL MODEL

To analyse a pathogenesis and predict ending conditions of a disease it is necessary to process a topological model, describing this pathogenesis. The main goal of this mathematical processing is to estimate structural characteristics of the model. This processing formally is a graph structural analysis observing all elements of the model. New characteristics for elements of the model are introduced. These characteristics allow to evaluate an importance of each element and to predict the ending functional state of the system.

2.1. Evaluation of model elements

If the topological model is created for a heterogeneous system, then the structure of the model is uneven and elements have various character and nature. To reduce and formalize this inequality, characteristics of model elements are introduced. These characteristics allow performing overall mathematical operations with the model. Characteristics are like lowest level parameters which are used in the process of modelling and analysis of the system. Also they are used as qualitative and quantitative measures to select an optimal recovery strategy. The main characteristic of a node is ***functional level***, described in the section 1.4. Other characteristics of a node:

1. ***Importance level of a node.*** This parameter describes the importance of a node in the context of the topological model. For the node x_{xi} the importance level $a_{xi}^{(1)}$ also indicates a role of the organism subsystem in the pathogenesis process. For the side effect node y_{yi} an importance level $b_{yi}^{(1)}$ is assigned and indicates how much the side effect is dangerous and can cause new symptoms and diseases. Both parameters are evaluated by an expert in the scale [0, 1], where 1 describes highest importance;
2. ***Influence sensitivity of a node***. The parameter indicates an influence rate of a node. Each node in the graph is reachable by defined chains. Each chain consists of the set of nodes and links. The influence sensitivity parameter $a_{xi}^{(2)}$ of the node x_{xi} is inversely proportional to the therapy count which can reach x_{xi}. Therapies that reach nodes of organism subsystems are evaluated by performing an algorithm of graph traverse, for example Dijkstra's or Floyd-Warshall algorithms [Cormen2009, Dasgupta2006]. Parameter is in the range [0, 1], where value 1 indicates that the nodes is reachable just from one therapy.

2.2. Evaluation criterion of external influence

The main task to be solved by the pathogenesis topological model is the modelling and analyse of a recovery strategy (drug therapies). A strategy is evaluated by its performance after the modelling process. By using separated node parameters, five efficiency criteria of a recovery strategy are proposed. These criteria allow comparing recovery strategies in-between and selecting the optimal solution for further improvements of the system. Projective optimality criteria of a recovery strategy:

1. ***The efficiency rate*** of a therapy is used to describe how a therapy changes the health state of a patient. The rate is evaluated by modelling the performing therapy, which mathematically is a result of recalculating all functional levels of nodes and summarising the changes in one value (2.1.). The recalculation of all nodes is based on the best-first search method (a graph traverse algorithm) [Dasgupta2006], performed from each therapy node. Coefficients V_1 and V_2 defines the evaluation importance of subsystem nodes and side effect nodes.

$$Eff(t_{ti}) = \sum_{xi=1}^{xn} V_1(\alpha_{xi}^{(0)} - \Delta_{xi}) + \sum_{yi=1}^{yn} V_2(\beta_{yi}^{(0)} - \Delta_{yi}). \quad (2.1)$$

2. ***A complex influence rate on essential parts of the model***. As the pathogenesis process is heterogeneous, diverse parts of the model have a diverse influence (importance) rate. By using expert evaluation methods an importance level is assigned to each node ($a_{xi}^{(2)}$ of x_{xi}). The final importance level of a recovery complex is calculated by summing up the importance levels of all nodes being covered by the therapy t_{ti} (2.2.).

$$A_1(t_{ti}) = \sum_{xi=1}^{xn} a_{xi}^{(1)}(1-\alpha_{xi}^{(0)}). \quad (2.2)$$

3. ***The recovery coverage level*** is used to describe an influence spectrum of a therapy. The criterion indicates how a therapy covers a periphery and hardly reachable nodes of the model. The parameter is evaluated by (2.3.).

$$A_2(t_{ti}) = \sum_{xi=1}^{xn} a_{xi}^{(2)}(1-\alpha_{xi}^{(0)}). \quad (2.3)$$

4. ***Provoked side effects*** is the first negative parameter to be minimised. The parameter is calculated by summing up the functioning levels of the provoked side effects of the therapy (2.4.). The criteria describe the negative side effect of a recovery.

$$B_1(t_{ti}) = \sum_{yi=1}^{xn} b_{yi}^{(1)}(1-\beta_{yi}^{b}), \quad (2.4)$$

5. ***A cost of a recovery complex*** is the second criteria with the negative manner. The criterion is calculated by using expert group inquiry methods. 12 medicine experts were provided with a test sheet, which contained all possible medicament combinations. The task of each expert was to evaluate which medicament is more expensive. Also equality sign was allowed. Then concordance of experts were evaluated and final cost coefficients of therapies were evaluated (Table 2.1.) [Karp2011].

Table 2.1.

Cost coefficients of therapies

No.	Therapy	Coeff.
1.	ACEI	0,8
2.	Beta blockers	0,4
3.	Calcium antagonists	0,6
4.	Diuretics	0
5.	Central simpatolitics	0,2
6.	Selective alfablockers	1

3. SELECTION OF OPTIMAL RECOVERY STRATEGY

The pathogenesis topological model allows modelling of external influences. By defined efficiency criteria a comparison between these external influences can be performed. The main task to be solved is the development of an appropriate prototype of a decision making, which could allow selecting the most optimal solution for a patient. Before the definition of the decision making technique a review of problem sphere should be done. It is necessary to theoretically examine methods of possible decision making of a medic in the task of therapy selection.

In the real praxis a patient comes to the doctor with a list of diagnosis (results from laboratory tests) and conditions (linguistic evaluations). The doctor respecting this list and by estimating the clinical history of the patient, tries to define the ending diagnosis and prescribe a therapy (Figure 3.1.). Frequently a therapy complex containing multiple drugs is assigned because the ending diagnosis is uncertain.

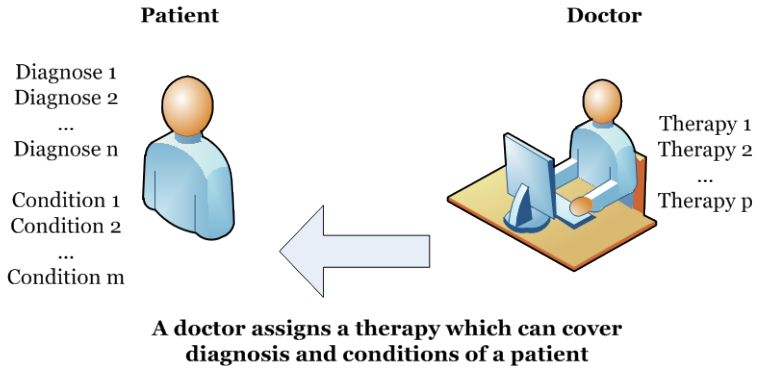


Figure 3.1. Procedure of therapy selection

Based on the combinatory laws (combinations and permutations), Table 3.1. describes the amount of solution set to be taken into account. The table is divided in two groups:

- Combinations where therapies do not repeat. In the creating process of a therapy just a unique drugs are included. This is an overall practice in medicine, where one therapy complex includes several unequal drugs;
- Combinations with repeating therapies. In this case repeating drugs are possible.

The amount of drugs included in final therapy complex is indefinite concept and a guideline of the therapy dimension is not defined. By studying widely used guidelines of management of arterial hypertension [Erglis2007] it is possible to conclude that recommended size of a multi therapy is 3 or 4 drugs.

Table 3.1.

The amount of possible therapies

Therapy count	Therapies do not repeat			Therapies repeat		
	1	2	3	1	2	3
1	8	16	24	8	16	24
2	28	112	252	64	256	576
3	56	448	1512	512	4 096	13 824
4	70	1120	5670	4 096	65 536	331 776
5	56	1792	13608	32 768	1 048 576	7 962 624
6	28	1792	20412	262 144	16 777 216	181 102 976
7	8	1024	17496	2 097 152	268 435 456	4 586 471 424
8	1	256	6561	16 777 216	4 294 967 296	1 100 754 314 176

If therapies in the possible combination should not repeat then the maximal count of combinations is in the case when a therapy contains 6 drugs. The set of solutions is restricted and foreseeable (approximately 20 thousand solutions). To select a solution from such amount set, methods of multi-objective optimization could be used. The basis of the usage of these methods is an overall examination of all solutions and a mutual assessment of them (*Brute-force search*). In the case of therapies containing repeated drugs an exponential rise of an amount of possible solutions is observed. For example, if a therapy can include 8 repeating drugs then the amount of possible solutions is over trillion. The size of this set is a restriction to perform a modelling and mutual comparison of all therapies. A possible calculation time and used computation resources are limited and also the result should be provided in a reasonable time. The full search of a large space of solutions is inadequate and ineffective. These constraints are the basics why the solution should not be selected from the set of all solutions, but synthesised by a step procedure. Metaheuristic methods are used to search the space of solutions to select the optimal recovery strategy.

3.1. Selection of optimal recovery by using multi-objective optimization

In mathematics and computer science optimization is defined as a methodology which allows to select the best (optimal) solution from a defined set of solutions. Principles of optimization are widely used to estimate a suitable solution of a goal function. In multi-objective optimization (also called multi-goal, vectorial, multi-criterion optimization) there are multiple goal functions which conflict in-between. In this conflicting condition one

solution which satisfies all goals mostly does not exist. The task of multi-objective optimization is defined as equation (3.1.).

$$\text{Min/ Max} \quad \left.\begin{array}{ll} f_m(x), & m = 1,2,...,M; \\ g_j(x) \geq 0, & j = 1,2,...,J; \\ h_k(x) = 0, & k = 1,2,...,K; \\ x_i^{(L)} \leq x_i \leq x_i^{(U)}, & i = 1,2,...,n. \end{array}\right\} \tag{3.1.}$$

A possible solution *x* is a decision vector $x = (x_1, x_2, ..., x_n)^T$. Last part of equation defines constrains, which mark a set *D* of possible solutions. Second and third part of the equation system describes *J* inequality and *K* equality constraints. A solution *x* which satisfy all constraints and is within the set of solutions is regarded as feasible solution. If one of the conditions is not satisfied then the solution is infeasible. The task of optimization consists of *m* goal functions $f_m(x)$, which should be minimized or maximized. [Deb2011]

Not always a goal function *f(x)* is defined as mathematical function, but could be as an evaluation of a solution. An evaluation by certain criteria could be used to mutually compare solutions. In the analysis of a functioning system by using topological modelling, a possible solution of a recovery is evaluated by previously stated performance criteria (Section 2.2.).

In contradistinction to single-objective optimization, a possible solution in multi-objective is more a concept than definition. Mostly one optimal solution in multi-objective space does not exist, but instead a set of possible solutions exists. Within this set all solutions satisfy previously defined criteria of optimality. One of the concepts, defining a set of optimal solutions is the concept of Pareto optimality [Pareto1906]. The main task of multi-objective optimization is to determinate the Pareto set, which satisfy conditions [Zitler2000]:

1. A Pareto set generated by multi-objective optimization should be as close as possible "true" Pareto set. In the best case the generated set is a subset of Pareto set;
2. Elements of generated Pareto set should be distributed equally over "true" Pareto set;
3. Solutions of boarders of "true" Pareto set should be discovered and examined;

If the Pareto set is estimated then the next task is to select a final solution from this set. In literature and practice one universal method does not exist, which always could provide an optimal solution. Instead various methods and classes for definite mathematical models, goal functions and decision making preferences exist [Coll2003, Marl2004, Parlos2000]. Methods to be selected should satisfy projective tasks:

1. A full browse of the set of all possible solutions is performed and solutions are evaluated by five optimality criteria. The final solution should be selected from the set regarding these criteria. To solve this task combinatory multi-objective optimization methods should be used;

2. The pathogenesis topological model is created to analyse events in discrete time. No dynamics of the system are taken into account, but discreet states are analysed. It means that selected optimization methods should be from discreet optimization class;
3. Methods of gradient type cannot be used, because goal functions are not defined as analytic functions;
4. It is necessary to use a prior optimization, where decision maker defines weights for each criterion. A prior method best of all reflects a strategy of a decision maker. Weights can characterize the preference of a decision maker.

To obtain criteria weights an expert survey was executed. Together 8 experts evaluated each criterion and gave a weight in the range from 0 till 10. All the results and evaluated weights are provided in Table 3.2.

Table 3.2.

Survey of experts

Criteria\Experts	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Sum**	**Coeff.**
1. Efficiency	10	10	9	10	9	9	10	10	77	**0.24**
2. Importance	10	8	10	8.5	9	9	10	10	74.5	**0.23**
3. Spectrum	5	5	8	8	9	7	10	9	61	**0.19**
4. Side effects	7	9	8	8	9	7	8	8	64	**0.2**
5. Costs	5	4	8	4	7	6	5	4	44	**0.14**
									320.5	1

Concordance W= 0.79

Decision making includes 7 methods of multi-objective optimization:

1. Weighted sum method;
2. Weighted goal method;
3. Exponential Weighted method;
4. Absolute two criteria method;
5. Relative two criteria method;
6. E-constraint method;
7. Lexicographic method;

The strategy of decision making is based on methods of expert enquiry and evaluation [Mark2009]. At the beginning of the decision making all methods are executed and each solution gets an evaluation by these methods. Then the concordance of methods is evaluated by Kendall concordation [Kend1990]. If the unity is low (concordation coefficient is under 0.6), then a rejecting procedure is performed. The rejecting procedure disposes separate solutions or entire multi-objective method to obtain a higher level of unity. If the concordance level is high then a solution with a lower average rank is selected as a final one.

3.2. Synthesis of optimal recovery by using multi-objective genetic algorithms

Metaheuristic methods are based on the tactic of a decision maker. These methods are more like concepts and represent the class of approximation algorithms. Methareuristics are like high level methodologies which are used as an overview guideline for the usage of heuristic methods in the case of complex optimisation. The main advantage of these methods is the guarantee of the result in the case when "classical" optimization methods cannot provide an acceptable solution. [Talbi2009]

Evolutionary algorithms (EA) belong to the class of metaheuristic algorithms and are based on the Darvins' evolution theory. First papers where published in the middle of 20^{th} century and a fast grow and popularity was shortly observed [Holland 1975]. In the end of 20^{th} century genetic algorithms, which represent evolutionary algorithms, became popular among the tasks of hard optimization [Goldb1988, Goldb1989]. Nowadays evolutionary and genetic algorithms are widely used in various optimization tasks. The main advantage of evolutionary algorithms comparing strong "classical" optimization is the population based search which simultaneously searches solutions in different directions. Other optimization methods like hill climbing, tabu search, simulated annealing operates with one solution which by the algorithm is improved. [Jaime2008]

Genetic algorithms

Genetic algorithms (GA) represent the class of evolutionary algorithms. Within the research Multi-objective genetic algorithms will be used to synthesize a recovery strategy. To use these methods basic concepts should be defined.

Individual- a possible solution. In the usage of genetic algorithms an individual ind_i is assumed as one complex of therapies. For example, an individual could be a complex containing one therapy t_1, two therapies t_1/ t_2 and so on. If 8 therapies with 3 doses are analysed then together there are 1 100 754 314 176 individuals (Table 3.1.).

Chromosome- an abstract representation of an individual. The term further will be used in the context of genetic algorithms and defines an abstract representation of an individual. Mostly chromosome is coded in binary form containing ***genes***. In practical usage also real number and linguistic coding is possible. In the recovery selection a integer number coding will be used. The length of chromosome is twice longer than the amount of therapies, being used, because one gene defines the index of therapy, but second index is used to describe a dose of therapy. The maximal possible length of chromosome is 16 genes (8 therapies and 8

doses). In figure 3.2. an example of chromosome is represented. The therapy consists from t_1 with half dose, t_3 with full dose, t_4 with full does and t_6 with quarter dose. In this example the coding is of the chromosome is 01 22 32 50.

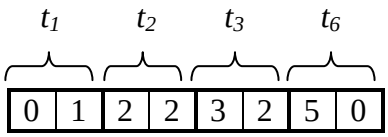


Figure 3.2. Example of chromosome

Generation- a group of individuals used in genetic operations. The performance of GA is based on the evolution of a generation by using genetic operators. A generation is defined as $Pa_i = \{ind_1, ind_2, ..., ind_u\}$, where *u*- number of individuals in the generation. The first generation is made by random operator and next generations are made by genetic operators.

Fitness function- an evaluation function of individuals. A fitness function is used to evaluate the suitability of an individual in the generation. Not always a fitness function is in the form of mathematical equation, but could be also given as an evaluation by previously defined criteria. In the recovery selection the fitness function is an evaluation by defined five performance criteria.

Genetic operators- operations used to generate new individuals based on individuals from the previous generation. Used genetic operators:

1. ***Crossover*** is the main operator used to generate new offspring from parents. In the method one point two parent crossover method will be used, which generates two new offspring by exchanging chromosome parts of two parents;
2. ***Mutation*** is the easiest genetic operator and is used to change a state of one gene in chromosome. In this case a random number generator is used to change a state of a gene. If for the mutation a even index of gene is selected (therapy) then a number from range [0, 7] is selected and if odd number is chosen then new number will be in range [0, 2];
3. ***Elitism*** is an operator allowing to pass best individuals to next generation. In direct way elitism is not allowed in multi-objective genetic algorithms, because the rule of nondominace should be satisfied. In this case a modification of elitism is used.

For the synthesis of an optimal therapy a basic genetic algorithm will be used (Figure 3.3.), [Goldb1989].

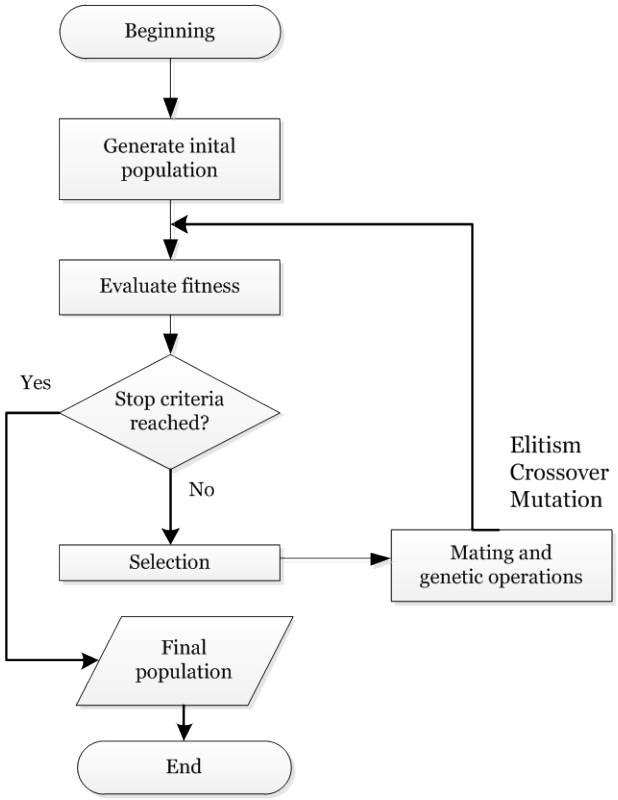


Figure 3.3. Basic genetic algorithm

For the further usage three multi-objective genetic algorithms are adjusted and used:

1. Weighted sum genetic algorithm (WSGA);
2. Vector evaluated genetic algorithm (VEGA);
3. Multi-objective genetic algorithm (MOGA).

If the genetic algorithm is designed correctly then the overall fitness of generation should improve and move to global optimum. This tendency is called convergence. A genetic algorithm is reached the level of convergence if from the population at least 95% of individuals have the same genetic information (have the same genes) [Dejong1975].

4. IMPLEMENTATION OF DEVELOPED METHOD- MEDICINE DECISION SUPPORT SYSTEM

Intellectual computer systems are widely used in medicine and other task domains. The main task of these systems is to provide a decision or advice similar to one provided by an expert. A computer system is an intelligent system if it has decision making capabilities and corresponds to the concept of artificial intelligence. An overall definition of artificial intelligence does not exist, but by analysing them, it is possible assume, that the computer system to be developed will be from the class of intelligent systems. Computer systems containing knowledge base and decision module to solve tasks in specific domain are called production systems or expert systems. These systems are defined as a model, capable to make a decision like a real expert. The model of knowledge based systems is based on the paradigm that a man solves a problem by using his knowledge [Newell1958, Newell1972]. The overall structure of medicine computer system is based on the typical structure of an expert system (Figure 4.1.), [Durkin1994, Durkin1998, Giarr1994, Negnev2002]:

- **Knowledge base** is used to store all available knowledge, rules and relations about the problem, which is observed. In this case, developed topological model and additional expert evaluations form knowledge base. The development of the model and testing tools are included into the system;
- **Data base module** is used to store a historical data from previous cases. It provides additional information for decision making;
- **Decision making module** is the central element of the medical system and is based on graph computation, multi-objective optimisation and expert evaluation methods;
- **Explanation model** and user interface provide easy readable and understandable conclusions and explanations, which are generated by the expert system, which includes multi language interface and all patients data input. As a result, the computer software provides all suggested therapies in a ranked manner, obtained by each of the optimisation techniques;
- **User interface** provides an intuitive and correct data input/output. The computer system is based on the windows structure and includes multi language (Latvia, Russian, and English) support.

Participants in the development and usage of the medical computer system:

- An expert provides his knowledge about the problem to help creating the knowledge base of the expert system;

- Knowledge engineer collects information from the experts and formalizes them into the knowledge base;
- End user of the system is not involved in the development process but is the main person for whom this system is created and who will use it in his regular praxis.

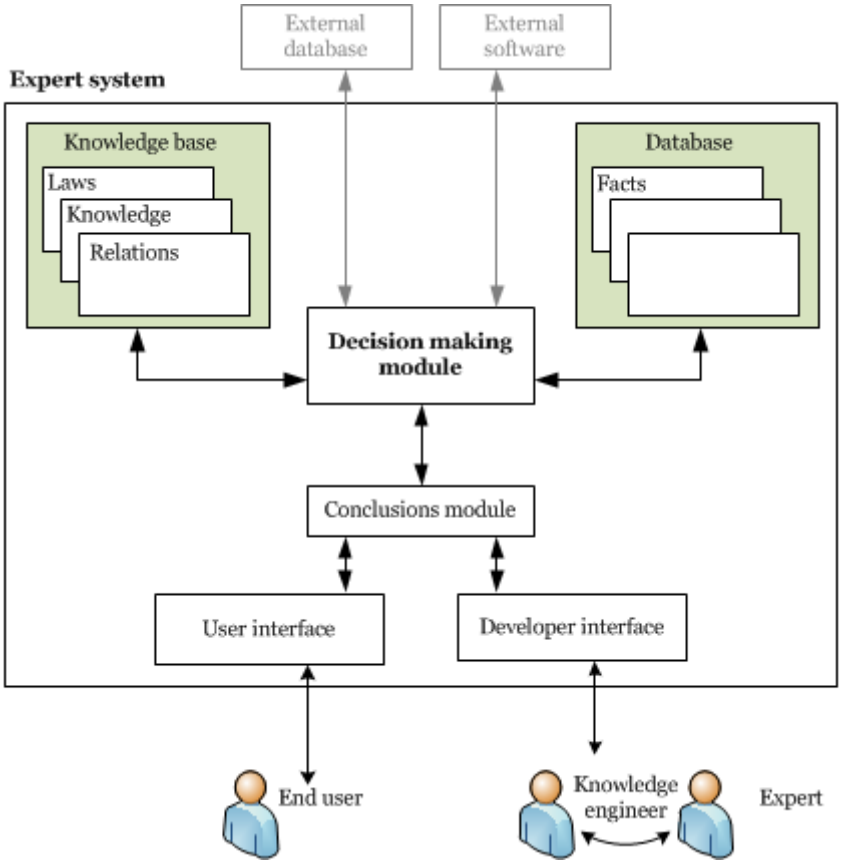


Figure 4.1. Structure of expert system

Main characteristics and tasks to be solved by the medicine computer system:

1. Medicine computer systems for multiple cardiovascular diseases are not common, because the diagnosis methods of them vary. For example, computer system for arterial hypertension mostly are based on knowledge base and decision making module, systems for atherosclerosis detection use pattern recognition and diabetes mellitus systems mostly are long period screening systems. A therapy selection for these systems is not even. Within the thesis developed computer system uses topological modelling to analyse all three diseases;
2. The computer system is innovative because include a therapy selection for all three diseases by using a joint analyse methodology. The methodology is based on mathematical modelling and multi-objective optimization;
3. The computer system is based on the structure of expert system and includes knowledge base (topological model and all evaluations), decision and explanation module. Decision making is based on the methods of expert evaluation;

4. The main goal why such a medicine system is developed is not case to take a place of the expert or exclude him at all. As the knowledge base includes high level knowledge of experts the main task is to hold these knowledge and if it is necessary provide a decision for the task to be solved;
5. A decision provided by this medicine system should be as an advice and final decision about the patient should be made by the doctor.

4.1. Performance evaluation of the system

To evaluate the performance of the computer system a patient from historical data of a hospital was examined. In the test a therapy should be selected by including not repeating drugs. For the test an average home computer was used (Core 2 Duo E8400 3 Ghz, 4Gb RAM, Windows experience index 5.8.) with Windows 7 32- bit operating system. Results of the test are described in Figure 4.2. As it is seen from the figure, a possible number of therapies when a treatment should be found by genetic algorithms is 6. If the amount of therapies, included in the treatment is less than 6, then the average modelling and selection time is low and the user of the system could receive results immediately.

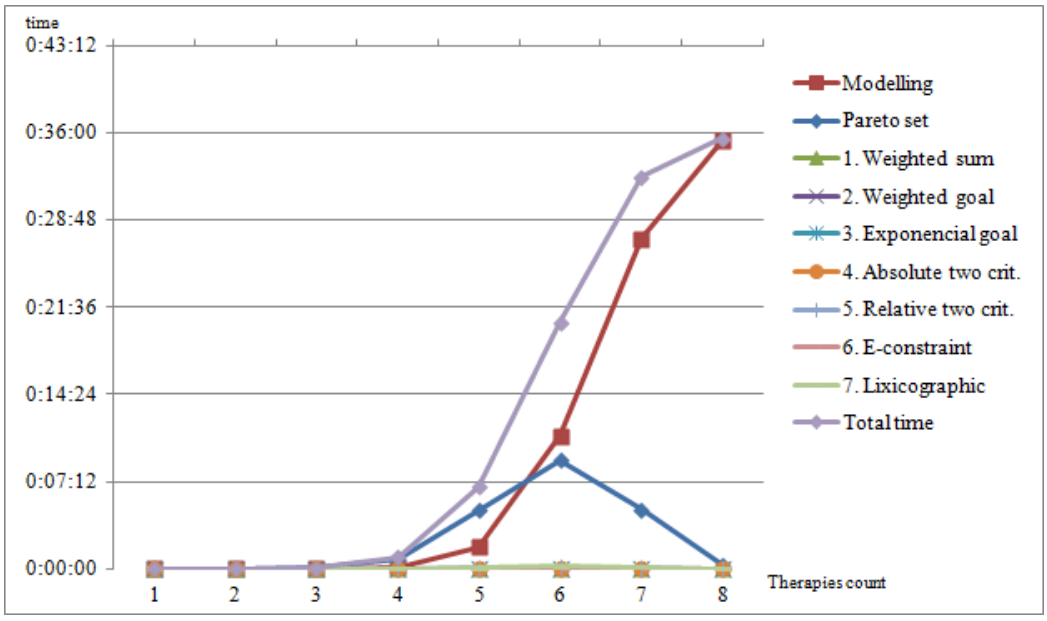


Figure 4.2. Performance time of multi-objective optimization

To evaluate a correctness of all methods a test patient from real clinical data has been examined. Together three different test scenarios where performed.

Test no. 1 – Methods of Brute- force search multi-objective optimization

- Therapies to be examined: 6;
- Doses used: 3;
- Therapy count in the treatment complex: 3 (drugs do not repeat);
- Possible amount of treatments: 1512;
- Pareto set is evaluated by weighted sum and weighted goal methods;

Results of all methods are provided in Table 4.1. In the table just solutions from Pareto set are included. By columns all seven methods are listed. A number in the table describe a rank of a solution acquired by a certain method. In Figure 4.3. the set of all solutions is provided. On the abscise all positive criteria are aggregated, but on the ordinate is used to described the component of negative criteria.

Table 4.1.

Results obtained by multi-objective optimization

No.	Therapy	Met1	Met2	Met3	Met4	Met5	Met6	Met7	Sum
1	T2-0.5/T4-0.5/T5-1	1	1	1	1	10	4	11	**29**
2	T2-0.5/T4-0.5/T5-0.5	2	2	2	2	4	17	41	**70**
3	T4-0.5/T5-1/T6-0.5	8	10	6	21	42	11	16	**114**
4	T4-0.5/T5-1/T6-1	11	13	5	33	72	3	22	**159**
5	T3-0.25/T4-0.5/T5-0.5	32	35	53	17	13	191	35	**376**
6	T2-0.5/T5-1/T6-1	55	76	15	119	175	1	146	**587**

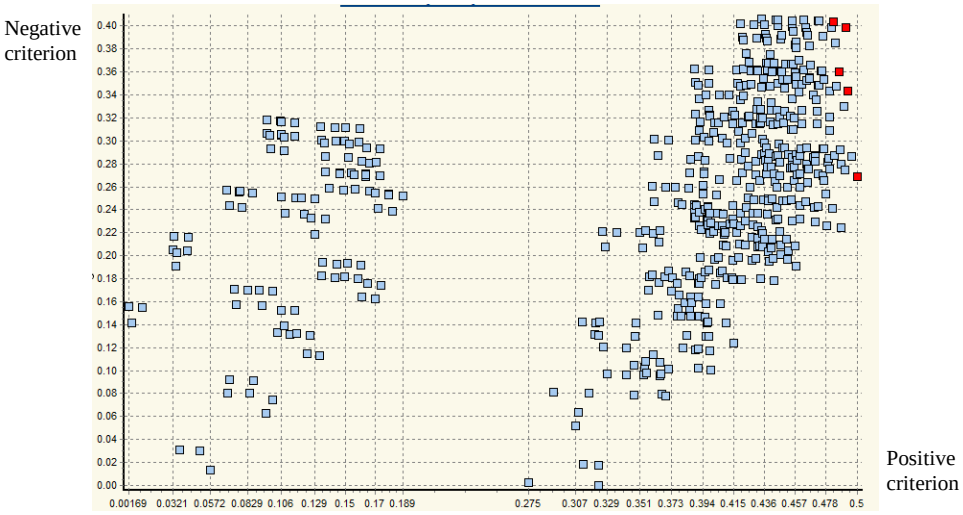


Figure 4.3. Set of solutions (Elements from Pareto set are with dark marks)

Results:

1. Methods of multi-objective optimization within the class provide comparable results, which justify the correctness of the mathematical model created for particular task;
2. E- constraint and lexicographic methods provide results which differ from the results calculated by other methods;

3. Pareto set satisfy the definition and geometrical interpretation. Solutions within the Pareto set are distributed equally and also solutions from the boarders are included;
4. Most of the solutions within the Pareto set have a large sum of ranks. This situation is the cause of criteria's weights provided by experts.

Test no. 2 – Multi-objective genetic algorithms

- Therapies to be examined: 8;
- Doses used: 3;
- Therapy count in the treatment complex: 3 (drugs repeat);
- Possible amount of treatments: 13 824;
- Population size: 300, Elitism index: 10, sequential population selection;
- Stop criteria: maximal amount of populations: 400; normalization of individual fitness: 50 individuals.

After execution of all genetic algorithms, best individuals in the last population repeat. Main differences are in the sequence of drugs and their doses, which are sufficient condition. All genetic algorithms provide the convergence. In Figure 4.4. results of weighted sum multi-objective optimization are provided. As it is seen from figure, maximal and average fitness is increasing till the convergence.

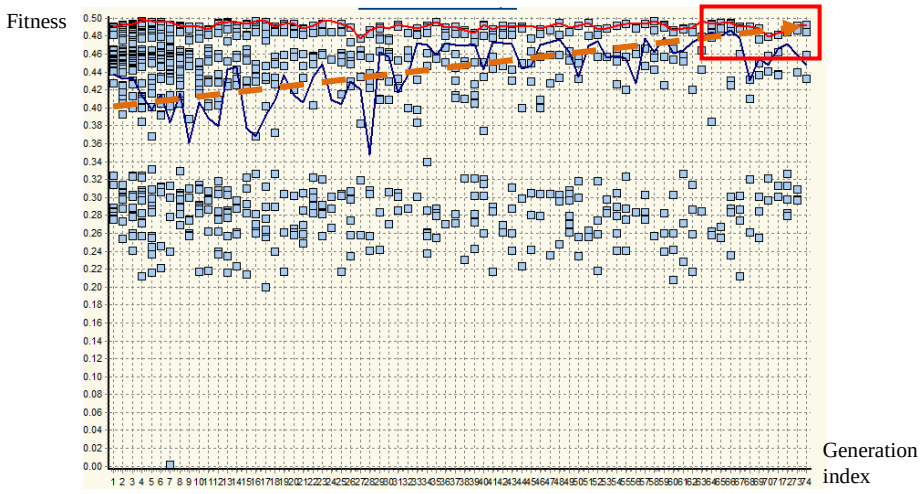


Figure 4.4. Results of weighted sum multi-objective genetic algorithm 1

Results:

- Results obtained by all genetic algorithms repeat;
- All algorithms ensure convergence. For example in figure 4.4. the are of convergence of weighted sum genetic algorithm is marked with a square. In the region of convergence an overall fitness does not improve and after fixed number of iteration the algorithm stops;

- Results calculated by vector evaluated genetic algorithm differ from other methods. Vector evaluated method splits a population in several subpopulation and then optimize solutions within the group by each criterion. Other two genetic algorithms in iterative procedure generate Pareto front;
- Genetic algorithms irregularly have a low average fitness (unprovoked peaks of the average fitness). These peaks are generated because a higher amount of useless solutions (mutants) with a low fitness are generated.

Test no. 3 – Common test for all methods of full browse and synthesis

- Therapies to be examined: 8;
- Doses used: 3;
- Therapy count in the treatment complex: 4 (drugs repeat);
- Possible amount of treatments: 331 776;
- Population size: 300, Elitism index: 10, sequential population selection;
- Stop criteria: maximal amount of populations: 400; normalization of individual fitness: 50 individuals.

Two best solutions from Pareto set calculated by brute- force search multi-objective optimization (execution time 10:12):

1. T4-0.5/T7-0.5/T7-0.5/T8-0.25;
2. T4-0.5/T7-0.5/T7-0.5/T8-0.5;

Two best solutions from Pareto set calculated by multi-objective genetic algorithms (execution time 00:43):

1. T7-0.5/T4-0.25/T7-0.5/T8-0.25;
2. T2-0.5/T7-0.5/T7-0.5/T8-0.5;

A set of solutions calculated by brute-force methods are provided in Figure 4.5. Solutions marked with red are from Pareto set. Results of weighted sum genetic algorithm are displayed in Figure 4.6.

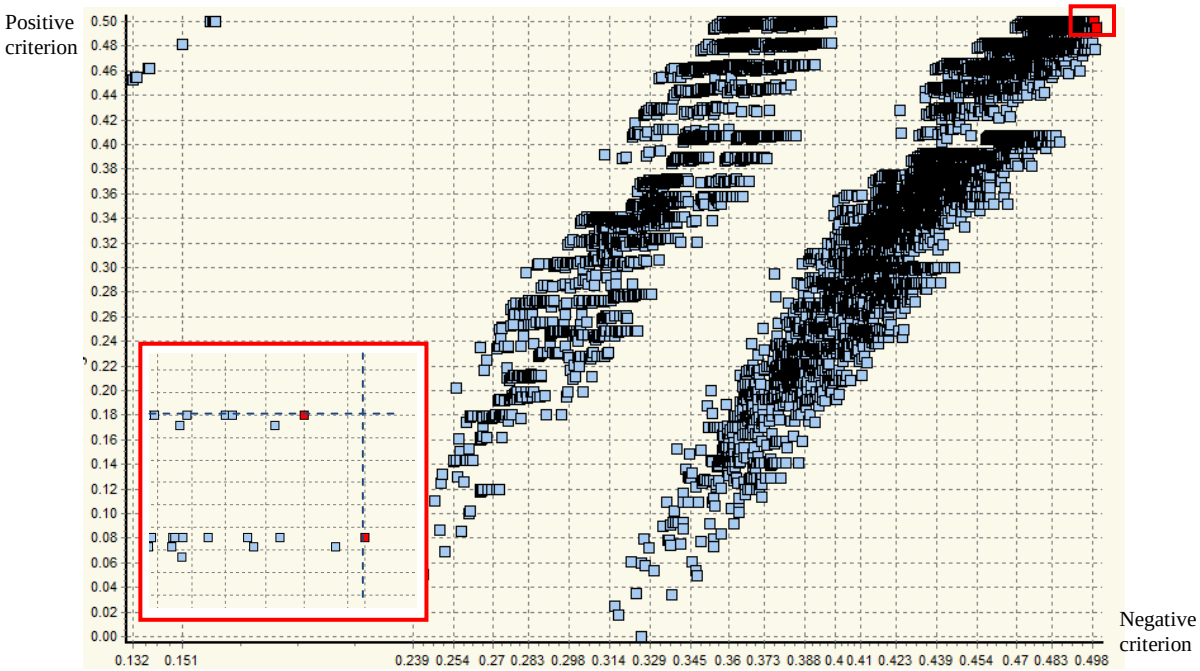


Figure 4.5. Solutions calculated by brute-force multi-objective optimization

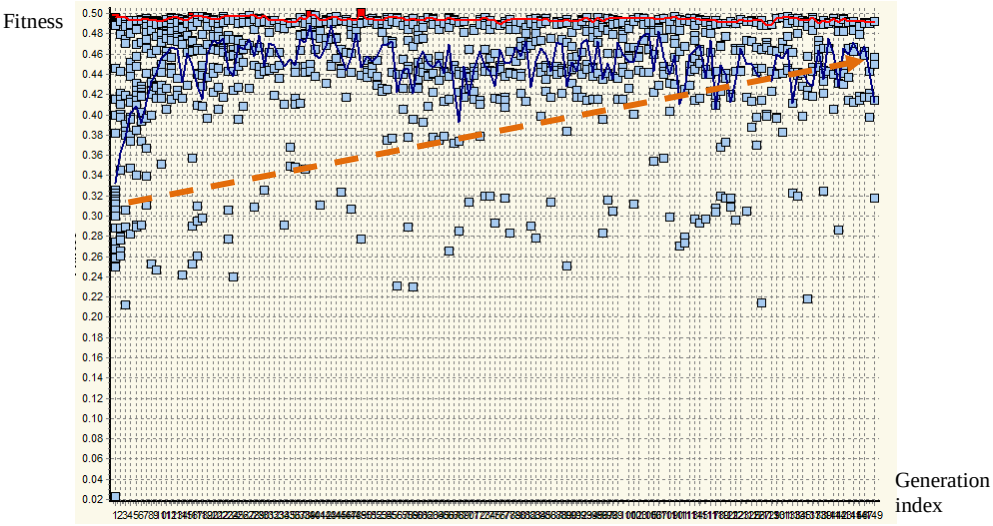


Figure 4.6. Results of weighted sum multi-objective optimization 2

Results:

- Genetic algorithms guarantee the same solutions which were obtained by methods of brute- force search.
- This test is used to adjust genetic operands and parameters (population size, elitism index, stop criteria and others) for more precise performance;
- If the set of solutions is large then both types of methods work correctly and estimate Pareto set. In figure 4.5. solutions of Pareto set are provided. If the set of solutions is large then it is reasonable to use genetic algorithms because they can provide result much faster. For example, in the test brute force methods where executed in around 10 minutes, but genetic algorithms in 43 seconds;

- In the case of an infinite set of solutions genetic algorithms should be used, because brute force methods will not be able to provide results.

4.2. Comparison and analyse of results acquired by the medicine computer system, doctor and guidelines for the management of arterial hypertension

Previously performed tests were made to evaluate the accuracy of the topological model. Other tests should be made to evaluate how a decision provided by the system corresponds to the decision made by a doctor. A set of clinical data from real patient history was collected and formalized into the input information of the computer system. Results were compared also with widely used guidelines for the management of arterial hypertension [Guid2003, Erglis2007]. Comparision results were analyzed in collaboration with dr. med. I Markovica (Table 4.2.: full compatibility, partly compatible, discrepancy).

Table 4.2.

Therapies assigned by medicine system, doctor and AH treatment guidlines

Patient	Doctor	Medicine system	AH guidelines
1.P.S.	T1/T2/T3	T1-0.5/T2-0.5/T3-0.5	T1/T10
2.K.M.	T1/T4	T4-0.5/T5-0.5	T1/T3/T4/T10
3.S.V	T1/T2	T2-0.5/T5-0.5	T10
4.L.D.	T1/T2/T5(T6)	T2-0.5/T4-0.5/T5-0.25	T1/ T10
5.E.A.	T2	T2-0.5	T10
6.Z.1.	T2/T4/T7	T2-0.5/T4-0.5/T7-0.5	T1/T2/T3/T4/T10
7.Z.2.	T1 /T2/T4/T6	T2-0.5/T5-0.5/T6-0.5/T7-0.5	T1/T3/T4/T10
8.Z.3.	T2/T5	T2-0.5/T5-0.5	T2
9.Z.4.	T2/T8	T2-0.5/T8-0.25	T2
10. Pac1	T1/T2/T3/T8	T2-0.5/T3-0.25/T8-0.25	T1/T2
11.Pac2	T2/T3/T4	T2-05/T3-0.25/T4-0.25	T1/T2 /T3/T10
12.Pac3	T2/T5	T2-0.5/T5-0.5	T2
13.Pac4	T2/T6	T2-0.5/T6-0.5	T2
14.Pac5	T1/T3/T4 /T7	T3-0.25/T4-0.5/T6-0.5/T7-0.5	T10
15.Pac6	T1/T3/T4/T7/T8	T3-0.25/T4-0.5/T6-0.5/T7-0.5/T8-0.25	T1/T3/T4/T10
16.Pac7	T1/T2/T8	T2-0.5/T6-0.5/T7-0.5	T2/T3
17.Pac8	T1/T2/T4/T7	T1-0.5/T2-0.5/T3-0.25/T7-0.5	T1/T4/T10
18.Pac9	T3	T2-0.5	T3
19.Pac10	T1/T2/T4/T7	T2-0.5/T3-0.25/T6-0.5/T7-0.5	T1/T4/T9

T1- AKEI, T2- Beta-Blockers, T3- Calcium antagonists, T4- Diuretics, T5- Central simpatolitics,
T6- Selective alfa-blockers, T7- Statins, T8- Anti diabetes therapy, T9- AT1-blockers, T10- Alfa-Blockers.

Conclusions made after the comparision:

1. The computer system provides similar decisions provided by a doctor (the computer system can repeat decisions made by an expert), accordingly computer system belongs to the intelligent systems (branch of artificial intelligence). In overall from 19 cases in 10 was a complete relapse, in 8 examples decisions made by the computer system

were better than decisions made by a doctor. Just in one case computer system provided a wrong decision. In the cases where decisions were better by the computer system it is visible that the therapy selected by the doctor will not improve the health state of a patient. The computer system more detailed evaluates the health state of a patient. In the case were the therapy selected by the computer system was wrong a cause could be in the formalization of the clinical data of a patient. Another reason for this difference could be weights for all criteria. Differently selected weights change the final decision;

2. The guidelines for the management of arterial hypertension in overall perform a lower adequacy (in 6 cases: full compatibility, 11: partial compatibility and in 2 case were discrepancy). The lower adequacy could be explained with a research method used in the guidelines. The guidelines for the management of arterial hypertension cover a wide geography and are used to guide doctors. They describe results from statistics and mostly recommend a direction of AH treatment. The developed computer system observes a patient more detailed and the final treatment is patient- specific. The computer system uses 70 observations, but guidelines use 30;
3. The computer system mostly recommends treatments which cover symptomatic of a patient. Non adequate drugs are not selected;
4. The cause of discrepancy between the computer system and guidelines could be explained by the weights of all criteria. In the testing an average weights obtained from the group of experts were used. These average weights cannot fully match with the priorities of a doctor;
5. The computer system recommends treatments with lower doses which is widely used praxis in the treatment of cardiovascular diseases;
6. The guidelines for the management of arterial hypertension are wide and frequently include too many treatments. Recommendations are more like directions and not always are effective. A large amount of drugs can cause non willing side effects and new diseases.

CONCLUSIONS AND FURTHER OBJECTIVES

Structural analyze of functioning systems and processes is an actual task domain in engineering and science. To solve this task mathematical models, describing the processes, has been developed. Topological modelling is an appropriate tool for constructing a mathematical model for heterogeneous systems when the available information is insufficient. Typical cases of heterogeneous processes are observed in biology, physiology and nature. In the doctoral thesis a new methodology for analyzing functional systems and processes has been developed. As a practical implementation of the method, an example from medicine is described. The methodology is adapted to analyze diseases, predict further condition of a patient and to select an optimal treatment for the recovery. At the beginning by using expert enquiry, a topological model of cardiovascular diseases (arterial hypertension, atherosclerosis and diabetes mellitus) has been developed. The mathematical model also includes external influences- drug treatments. In the modelling it is possible to predict the effect of a treatment and evaluate the efficiency of it by defined criteria. Efficiency criteria are used as a metric for the selection of an optimal solution. The recovery treatment is selected by using multi-objective optimization or genetic algorithms.

As an implementation of the methodology an intellectual medicine system has been created. The computer system is based on the typical structure of expert system and includes knowledge base, decision making module, conclusion module and user interface. The computer system is developed as a decision support system and ending conclusion about the treatment of a patient should be made by a doctor.

Further Objectives:

1. Testing and improvement of the medicine system;
2. Development of a module for automatic data input based on data mining. In the current system all the data about the patient should be entered manually;
3. Functional analyze of different processes. Formation of a topological model by using genetic programming;
4. Development of a dynamical structural analyze of a system.

LITERATURE

[Belev1962] Belevitch V. Summary of the history of circuit theory. Proceedings of the IRE, vol. 50, iss. 5, 1962, pp. 848-855

[Bender1978] Bender E. A. An Introduction to Mathematical Modelling, Dover, New York, USA, 1978, p. 256

[Cauer2000] Cauer E., Mathis W., Pauli R. Life and Work of Wilhelm Cauer (1900 – 1945), Proceedings of the Fourteenth International Symposium of Mathematical Theory of Networks and Systems (MTNS2000), p4, Perpignan, 2000

[Cauer1941] Cauer W. Theorie der linearen Wechselstromschaltungen, vol. I, Akad. Verlags-Gesellschaft Becker und Erler, Leipzig, 1941

[Coll2003] Collette Y., Siarr P. Multiobjective optimization: principles and case studies. 1st ed. Springer 2003. Corr 2nd printing, 2003, p. 293

[Cormen2009] Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. Introduction to Algorithms (3rd ed.), MIT Press and McGraw-Hill, 2009, p. 1312

[Dasgupta2006] Dasgupta S., Papadimitriou C. H., Vazirani U. V. Algorithms 1 edition, McGraw-Hill Science/Engineering/Math, 2006, p. 336

[Deb2001] Deb K. Multi-Objective Optimization Using Evolutionary Algorithms, John Wiley & Sons inc., New York, USA, 2001, p. 518

[Dejong1975] DeJong K. The Analysis and behaviour of a Class of Genetic Adaptive Systems. PhD thesis, University of Michigan, 1975

[Durkin1994] Durkin J. Expert System: Design And Development, Macmillan Publishing Company inc., New York, USA, 1994, p. 800

[Durkin1998] Durkin J. Certainty Therory, 1st eidition: Expert System: Design and Development, Macmillan Publishing Company inc., 1998, pp. 350-353

[Erglis2007] Ērglis A, Kalvelis A, Lejnieks A, Dzērve V., Latkovskis G., Mintāle I, Zakke I., Rasa i.. Kardiovaskulāro slimību (KVS) profilakses vadlīnijas, Rīga, 2007

[Giarr1994] Giarratano J. C., Riley G. Expert Systems: Principles and Programming, PWS Publishing Company/ International Thomson, Boston, USA, 1994, p. 644

[Goldb1988] Goldberg D. Genetic Algorithms, Addison Wesley, 1988

[Goldb1989] Goldberg D. Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Publishing Company, Reading, Massachusetts, USA, 1989, p. 432

[Holland1975] Holland J. H. Adaptation in natural and artificial system, Ann Arbor, The University of Michigan Press, 1975, p. 183

[Jaimes2008] Jaimes A. L., Coello Coello C. A. An Introduction to Multi-Objective Evolutionary Algorithms and Some of Their Potential Uses in Biology. Applications of Computational Intelligence in Biology, 2008, pp. 79-102

[Karp2011a] Karpics I., Markovics Z., Markovica I. Composition of United Multiple Diseases Evolution Topological Model. Proceedings of Intelligent Systems and Informatics (SISY 2011): IEEE 9th International Symposium, Subotica, Serbia, 2011, pp. 65-69

[Karp2011b] Karpics I., Markovics Z. Development and evaluation of normal performance recovery method of a functional system. Scientific proceedings of 9th IEEE International Symposium on Applied Machine Intelligence and Informatics, 2011, pp. 171-175

[Karp2010a] Karpics I., Markovics Z. Development of pathogenesis topological model node evaluation complex. Proceedings of 14th International Conference Biomedical Engineering, Technologia, Kaunas, Lithuania, 2010, pp. 163-166

[Karp2009a] Karpics I., Markovics Z. Improvement and Assessment of the Effective Therapy Selection Method. Proceedings of 13th International Conference Biomedical Engineering, Technologia, Kaunas, Lithuania, 2009, pp 219-223

[Karp2008a] Karpics I., Markovica I., Markovics Z. Most effective two therapies combination detection approach. Proceedings of 12th International Conference Biomedical Engineering, Technologia, Kaunas, Lithuania, 2008, pp 234-236

[Karp2008b] Karpics I., Markovics Z., Markovica I. Arterial Hypertension Therapy Selection by Using Topological Modeling and Production Law Logic. Proceedings of AITH'08- Advanced information and telemedicine technologies for health, National Academy of Sciences of Belarus, Minsk, Belarus, 2008, pp 54-58

[Kend1990] Kendall M. G., Gobbons J. D. Correlation methods, 5th edition, Arnod, London, UK, 1990, pp. 114-121

[Kolm1957] Kolmogorov A. N., Fomin S. V. Elements of the theory of functions and functional analysis Volume 1. Metric and Normed Spaces, Dover publications, New York, USA, 1957, p. 129

[Kolm1961] Kolmogorov A. N., Fomin S. V. Elements of the theory of functions and functional analysis Volume 2. Measure. The Lebesgue integral. Hilbert space, Dover publications, New York, USA, 1961, p. 128

[Mark2000a] Markovitch Z., Markovitcha I. Modeling and diagnostics. In E. Carson & E. Salzseider (ed.), Modeling and control in biomedical systems 2000 IFAC Symposium, Germany, 2000, pp. 157-160

[Mark2000b] Markovitch Z., Markovitcha I. Modelling as tool for therapy selection. In simulation and modelling 14- th European simulation multiconference, ESM Belgium, 2000, pp. 621-623

[Mark2002a] Markovičs, Z., Markoviča, I., Makarovs, J. Alternatīva koncepcija terapijas izvēlei, The Alternative Conception on Therapy Selection, Rīgas Tehniskās universitātes zinātniskie raksti. 5.sēr., Datorzinātne. - 11.sēj.: Datorvadības tehnoloģijas (2002), pp. 19-27

[Mark2009] Markovičs Z. Ekspertu novērtējumu metodes. Rīga : RTU Izdevniecība, 2009, 110. lpp.

[Marl2004] Marler R. T., Arora J. S. Survey of multi-objective optimization methods for engineeringStruct Multidisc Optim 26, 2004, pp. 369-395

[Negnev2002] Negnevitsky M. Artificial Intelligence. A Guide to Intelligent Systems, Pearson Education Limited, 2002, p. 394

[Newell1958] Newell A., Shaw J. C, Simon H. A. Elements of a theory of human problem solving, Psychological Review, vol. 65, 1958, pp. 151-166

[Newell1972] Newell A., Simon H. A. Human problem solving, Prentice hall, Englewood Cliffs, Nj, USA

[Osis1969] Osis J. Topological Model of System Functioning. Automatics and Computer Science, -J. of Acad. Of Sc, nr. 6, Riga, Latvia, 1969, pp. 44-50

[Osis1991a] Osis J. Topoloģiskie modeļi tehniskajā un medicīniskajā diagnostikā, attēlu pazīšanā un ekspertu sistēmās Latvijā. Vispasaules Latviešu Zinātņu Kongress, Rakstu krājums, 5.sēj., Rīga, 1991.

[Pareto1906] Pareto V. Manuale di Economica Politica, Societa Editrice Libraria. Milan; translated into English by A.S. Schwier as Manual of Political Economy, edited by A.S. Schwier and A.N. Page, 1971, New York, USA

[Parlos2000] Parlos P. M. Multi-Criteria Decision Making Methods: A Comparative Study, Kluwer Academic Publishers, 2000, p. 288

[Talbi2009] Talbi E. G. Metaheuristics - From Design to Implementation, Wiley, 2009, p. 593

[Zbign2004] Zbigniew M**,** Fogel D. B. How to Solve It: Modern Heuristics. 2nd ed. Revised and Extended, 2004, XVIII, p. 554

[Zhou2002] Zhou X. H., McClish D. K., Nancy A. O. Statistical Methods in Diagnostic Medicine (Wiley Series in Probability and Statistics), Wiley- interscience. John Wiley & Sons., New York. USA, 2002, p. 464

[Zitzler1999] Zitzler E., Thiele L. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. IEEE Transactions on Evolutionary Computation, vol. 3(4), 1999, pp. 257-271

[Zitzler2000] Zitzler E., Deb K., Thiele L. Comparison of multiobjective evolutionary algorithms: empirical results. Evol Comput, vol 8(2), 2000 pp. 173-195