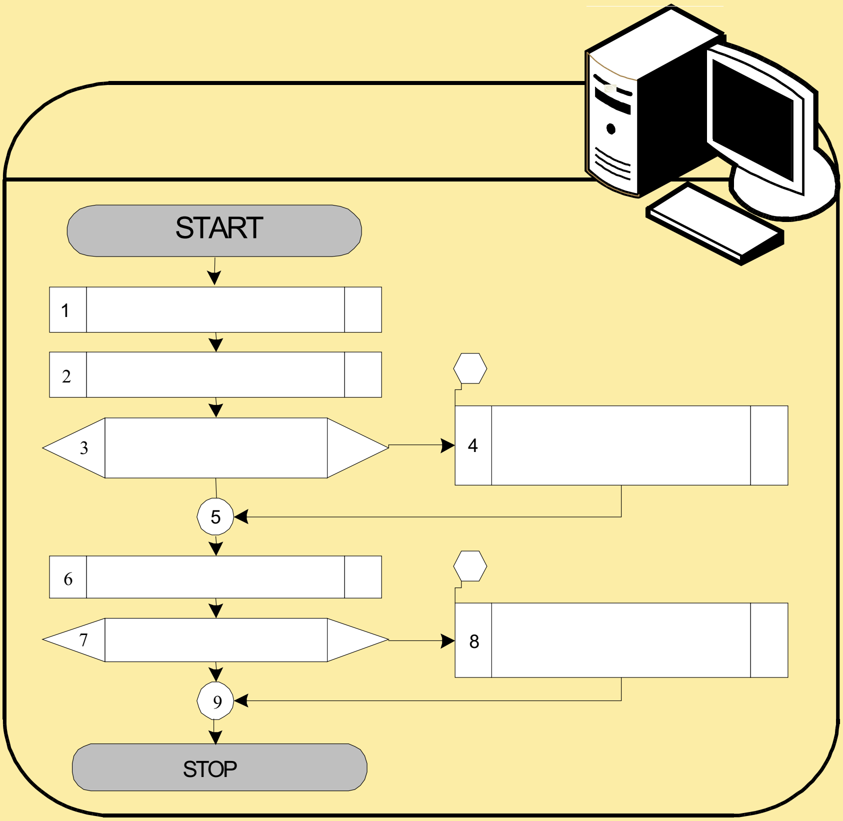
APPLIED COMPUTER SCIENCE

Vol. 9, No. 1, 2013



LUBLIN UNIVERSITY OF TECHNOLOGY

INSTITUTE OF TECHNOLOGICAL SYSTEMS OF INFORMATION

www.acs.pollub.pl

ISSN 1895-3735

Editorial staff:	**Arkadiusz Gola** (Editor-in-Chief, Secretary of the Scientific Board) **Jakub Szabelski** (Editorial secretary) **Dariusz Wołos** (Editorial secretary) **Łukasz Sobaszek** (Typesetting and text makeup) **Daniel Gąska** (Webmaster)
Redactorial staff:	**Jerzy Lipski** – Computer Science in Production Engineering (Subject Redactor) **José Luis Gascó Gascó** – Computer Science in Management and Economy (Subject Redactor) **Marek Janczarek** – Computer Science in Technology (Subject Redactor) **Alex Bodnar** – English Instructor at Kyushu University, Japan (Language Redactor) **Paweł Zaprawa** (Statistical Redactor)
Publisher:	Institute of Technological Systems of Information, Lublin University of Technology
Contact:	„Applied Computer Science" – Editorial Office 20-618 Lublin, ul. Nadbystrzycka 36, Poland Tel.: (+48 81) 538 42 76 Fax.: (+48 81) 538 44 96 e-mail: acs@pollub.pl www.acs.pollub.pl
Circulation:	550 copies

CONTENTS

project cash flow, constraint satisfaction problem, project portfolio

Marcin RELICH[*]

FUZZY PROJECT SCHEDULING USING CONSTRAINT PROGRAMMING

Abstract

The paper aims to present an application of constraint programming techniques for project portfolio scheduling taking into account the imprecision in activity duration and cost. Data specification in the form of discrete α-cuts allows combining distinct and imprecise data, and implementing a constraint satisfaction problem with the use of constraint programming. Moreover using α-cuts, optimistic, pessimistic, and several intermediate scenarios concerning the project scheduling and cash flows can be obtained and considered in terms of different risk levels.

1. INTRODUCTION

Project management is a complex decision-making process involving the time and cost estimations. The traditional approach to project management is to consider company projects as being independent of each other [1]. However, one of the characteristics of many industrial companies concerns the management of several simultaneously developed new products (projects) using the same resources. In order to maintain agility while avoiding wasteful investments, a strong discipline of project portfolio management is needed. This requires continuous attention and balancing company resources. In a multiple-project situation the vast majority of projects share resources with other projects and thus the major issue is to find a way of handling resource scarcity according to the overall strategic direction of the company and to rescue the project portfolio that tends to be at risk of failure [1, 2].

Project management problems typically consist of resource planning and scheduling decisions. In the context of resource management, it is often required to know how much a particular project will cost, what resources are

[*] University of Zielona Góra, Faculty of Economics and Management, Licealna 9,
65-216 Zielona Góra, Poland, e-mail: m.relich@wez.uz.zgora.pl

needed, what resource allocation procedure can ensure the completion of a project in target time, etc. Those requirements can be formulated as following standard, routine questions: Does the project can be completed before a given deadline? Is it possible to undertake a new project under given (constrained in time) resources availability? What is risk level for the project portfolio?

The scheduling problems that need to be solved in order to provide answers to the above questions belong to the class of NP-complete problems. The impact of real-life constraints on the decision-making is therefore of great importance, especially for designing interactive and task oriented decision support systems. Several methods and techniques have been proposed in this field, for instance, method for scheduling projects with resource constraints [3], resource levelling tools for resource constraint project scheduling problem [4], soft computing for optimization of an investment portfolio [5], the time-cost trade-off analysis in project scheduling [6], schedule-driven project management in multi-project environments [7].

In practice, managers frequently create programs and schedules based on the expected values of activity durations. However, many real-world planning and scheduling problems are subject to change, to resources becoming unexpectedly unavailable or tasks taking longer than expected [1]. The hypothesis made in critical path method that activity durations are deterministic and known is rarely satisfied in real life where tasks are often uncertain and variable [8]. The inherent uncertainty and imprecision in project scheduling have motivated the proposal of several fuzzy set theory based extensions of activity network scheduling techniques [9]. Among these extensions can be found, for instance, resource-constrained fuzzy project-scheduling problem [10], fuzzy critical chain method [11], criticality analysis of activity networks with uncertainty in task duration [12], fuzzy repetitive scheduling method [13]. Also, considerable research effort has been recently focused on the application of constraint programming frameworks [14-16].

Constraint Programming (CP) environment seem to be especially well suited for modelling real-life and day-to-day decision-making processes at an enterprise [14, 17]. CP is qualitatively different from the other programming paradigms, in terms of declarative, object-oriented and concurrent programming. Compared to these paradigms, constraint programming is much closer to the ideal of declarative programming: to say what we want without saying how to achieve it [18]. CP is an emergent software technology for a declarative Constraints Satisfaction Problem (CSP) description and can be considered as a pertinent framework for the development of decision support system software aims.

In the field of constraint-based scheduling two strengths emerge: natural and flexible modelling of scheduling problems as CSP and powerful propagation of temporal and resource constraints. Thus, the scheduling problem is modelled as CSP at hand in the required real-life detail and it enables to avoid the classical drawbacks of being forced to discard degrees of freedom and side constraints [19]. The model formulated in terms of CSP determines a single knowledge base and it enables effective implementation in constraint programming languages, as well as the development of a task-oriented decision support system for project portfolio scheduling. As a result, the problem specification is closer to the original problem, obtaining solutions that are unavailable with imperative programming. This provides motivation to consider project management in connection with the nature of a company and to develop a reference model that combines both these fields.

Although, several researchers have recognized the importance and necessity of applying fuzzy set theory or probability theory in project scheduling and project cash flow generation and analysis, there is still a lack of a use of a declarative approach in the field. The proposed approach aims at specifying project portfolio scheduling in terms of fuzzy CSP, using constraint programming to seek a solution to the problem, and enabling analysis of cash flow uncertainty at different α-levels.

The traditional approach for project scheduling is the well-known CPM (Critical Path Method) and PERT (Program Evaluation and Review Technique) [20]. The perception or estimation of uncertainty is encoded in the initial assignment of fuzzy activity duration and cost. Thereafter, in terms of project management, different α-cuts can be considered as separate risk levels [21]. Thus, a framework is provided for conducting risk analysis on the project cash flow with the appropriate α-cuts which limit the degree of fuzziness and essentially provide a measure of the prediction robustness. The risk levels can vary from "none", "low", "moderate" to "very high" as the α-cut moves from 1 towards 0. Moreover, at any given α-cut besides a delay or cost escalation there is also an opportunity to go ahead of schedule and reduce costs [22]. Difference between the proposed approach and PERT network diagrams concerns a number of scenarios and the use of integer numbers. PERT assumes only the absolute worst and best scenarios (everything goes worse or better than expected, respectively), whereas the proposed approach includes some possibility levels from 0 to 1.

The proposed approach for project portfolio planning allows a decision-maker perform analysis of cash flow uncertainty at different α-levels. During project implementation, the cash flow is crucial for the assessment of working capital requirements since the difference between project expenditures and payments determines the necessary capital reserves. Furthermore, an accurate cash flow is required in conducting project cost-benefit analysis, the determination of project financing requirements and in performing earned value analysis [22].

The remaining sections of this paper are organised as follows: Section 2 presents a problem formulation in terms of fuzzy CSP for project portfolio scheduling. A method for cash flow generation is shown in Section 3. An illustrative example of the approach is presented in Section 4. Finally, some concluding remarks are contained in Section 5.

2. PROBLEM FORMULATION OF FUZZY CSP

A considered reference model consists of a company and project portfolio. Its specification encompasses technical parameters, expert's experiences and user expectations in the form of knowledge base, i.e. as a set of variables, their domains, and a set of relations (constraints) that restrict and link variables. Such interpretation of model allows using the logic-algebraic method as a reference engine [14]. In this context, it seems natural to classify some decision problems as CSP. The problem formulation in terms of CSP enables a simplified description of actuality, i.e. a description encompasses the assumptions of object, implementing therein tasks, and a set of routine queries (the instances of decision problems).

In a classical form, the structure of constraints satisfaction problem may be described as follows [19]: $CSP = ((V, D), C)$, where: V – a set of variables, D – a set of discrete domains of variables, C – a set of constraints. Taking into consideration the imprecise characteristics of project management, it is assumed Fuzzy Constraints Satisfaction Problem (FCSP) as follows [23]:

$$FCSP = ((\tilde{V}, D), C) \tag{1}$$

where: $\tilde{V} = \{\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_n\}$ – a finite set of n fuzzy variables that in form of fuzzy number (a finite set of discrete α-cut) are described,
$D = \{d_1, d_2, ..., d_n\}$ – a set of domains for n fuzzy variables,
$C = \{c_1, c_2, ..., c_m\}$ – a finite set of m constraints limiting and linking decision variables.

FCSP is implemented according to the structure of the reference model, and can be also considered as a knowledge base. The knowledge base is a platform for query formulation as well as for obtaining answers, and it comprises of facts and rules that are characteristic of the system's properties and the relations between its different parts. As a consequence, a single knowledge base facilitates the implementation of a decision support system.

A knowledge base can be considered in terms of a system. At the input of the system are the variables concerning basic characteristics of an object that are known and given by the user. For instance, the variables concerning available resources in the enterprise and a sequence of project activities occur in the knowledge base describing the enterprise-project model. The output of the system is described by the characteristics of the object that are unknown or are only partially known. In the considered case, the variables include the cost and time of an activity as well as the resources usage.

A distinction of decision variables that are embedded in the knowledge base as an input-output variable permits the formulation of standard routine query containing a problem of cash flow planning in multi-project environment, such as: what is cash flow uncertainty for the given constraints (e.g. the deadline and budget, activity networks)? The method concerning the determination of admissible solutions for the above-described problem is presented in the next section.

3. CASH FLOW GENERATION

Given a set of projects $P = \{P_1, P_2, \ldots, P_I\}$, where the project P_i consists of J activities: $P_i = \{A_{i,1}, \ldots, A_{i,j}, \ldots, A_{i,J}\}$. The j-th activity of i-th project that is specified by the starting time of the activity $s_{i,j,1}$ (i.e. the time counted from the beginning of the time horizon H), the completion time of the activity $s_{i,j,2}$, and the duration of the activity $t_{i,j}$. The project P_i is described as an activity-on-node network, where nodes represent the activities and the arcs determine the precedence constraints between activities. According to this the precedence constraints are as follows: $s_{i,j} + t_{i,j} \leq s_{i,n}$ (for the n-th activity follows the i-th one), $s_{i,j} + t_{i,j} \leq s_{i,n}$; $s_{i,j+1} + t_{i,j+1} \leq s_{i,n}$; ...; $s_{i,j+n} + t_{i,j+m} \leq s_{i,n}$ (for the n-th activity follows other activities), and $s_{i,n} + t_{i,n} \leq s_{i,j}$; $s_{i,n} + t_{i,n} \leq s_{i,j+1}$; ...; $s_{i,n} + t_{i,n} \leq s_{i,j+m}$ (for the n-th activity is followed by other activities).

Imprecise variables determined by convex membership function $\mu(s)$ (e.g. a triangular fuzzy number $\tilde{s} = \langle a, b, c \rangle$) can be specified as α-cuts. An α-cut is a crisp set consisting of elements of A which belong to the fuzzy set at least to a degree of α ($0 < \alpha \leq 1$). An α-cut is a method of defuzzifying a fuzzy set to a crisp set at desired α-levels that correspond to the perceived risk (α=1 meaning no risk, α=0+ meaning the highest risk). Additionally, the low (α=0–) and high (α=0+) values of every α-cut represent the optimistic and pessimistic outcomes of that risk level. The main objective of fuzzy project scheduling is to apply fuzzy set theory concepts to the scheduling of real world projects where task duration can be specified as fuzzy numbers instead of crisp numbers [22].

The application of the FPS algorithm yields the fuzzy start and completion dates of activities. In the project network, the fuzzy start indicates the accumulation of uncertainty from preceding activities, whereas, the fuzzy completion date is the sum of the activity start with the activity duration. The fuzzy addition operator (+) for two triangular fuzzy numbers is defined as:

$$\tilde{s} + \tilde{t} = \langle a, b, c \rangle + \langle d, e, f \rangle = \langle a + d, b + e, c + f \rangle \quad (2)$$

The example of addition the fuzzy starting time of the activity to fuzzy the duration of the activity for three α-levels is illustrated in Fig. 1.

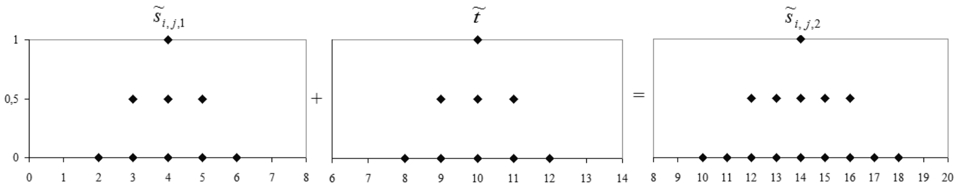


Fig. 1. Addition fuzzy numbers in terms of discretized α-cuts [source: own study]

In order to calculate the required cost per unit of time, the cost of every activity needs to be divided by its duration. However, the duration varies for different possibility measures and for optimistic and pessimistic scenarios. In the absolute best case (minD_α), the activity will start as early as possible and will last the minimum duration. In the absolute worst case (maxD_α), the activity will start as late as possible and will last the maximum duration. Considering an activity with an early start and completion dates, these intervals are defined as follows [22]:

$$minD\alpha = [\alpha(b-a) + a,\ \alpha(e-d) + d] \quad (3)$$

$$maxD\alpha = [\alpha(b-c) + c,\ \alpha(e-f) + f] \quad (4)$$

where: minD_α/maxD_α: is the interval of minimum/maximum duration of the activity at the respective α-cut,

D_α: is the α-cut of the activity duration.

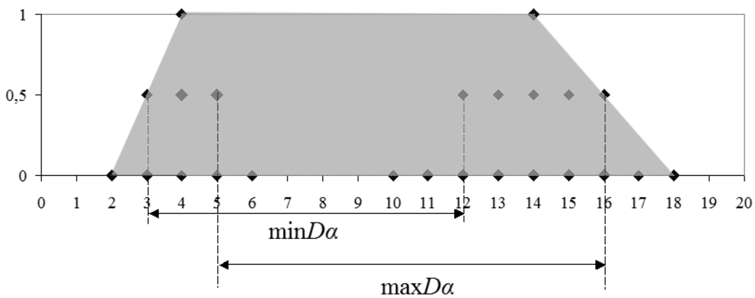


Fig. 2. Fuzzy activity start and completion date [source: own study]

Figure 2 shows an activity with a starting time of <2, 4, 6>, a duration of <8, 10, 12>, and a completion time of <10, 14, 18>. In this example, the duration intervals at α=0.5 are min$D_{0.5}$=[3, 12] and max$D_{0.5}$=[5, 16] and hence the activity cost is distributed in these intervals. In the best case, the activity begins as early as possible (3rd month) and lasts the minimum duration (9 months), whereas in the worst case, it starts as late as possible (5th month) and lasts the maximum duration (11 months). Similarly, minimum and maximum duration intervals representing optimistic and pessimistic scenarios for different possibility measures can be created for all α-levels (between 0 and 1). Consequently, the fuzzy start date and completion date mark the temporal start and completion boundaries of the activity within which the minimum and maximum duration intervals (minD_α/maxD_α) are defined for each α-cut.

The financial means are allocated to the activity $A_{i,j}$. taking into account all α-levels of a fuzzy number ($dp_{i,j,\alpha}$). If the cost for time unit is not an integer, then the cost is assigned as follows: for minD_α – the first value is less than the rest, for maxD_α – the last value is greater than the rest. Table 1 shows an example of the cost distribution for the activity equals 80 monetary units (m.u.) at five separate possibility levels. At time unit h = 3 for min$D_{0.5}$, there are 9 time units (see Fig. 2), the cost for a time unit equals 8.89 m.u., so for 8 time units (h = 4, …, 11) the cost distribution equals 9 m.u. and for first time unit (h = 3) equals 8 m.u. (80m.u. – 8*9 m.u.).

Tab. 1. An example of cost distribution for fuzzy duration of activity

h / α	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
minD_0			10	10	10	10	10	10	10	10											
min$D_{0.5}$				8	9	9	9	9	9	9	9	9									
D_1					8	8	8	8	8	8	8	8	8	8							
max$D_{0.5}$						7	7	7	7	7	7	7	7	7	7	10					
maxD_0							6	6	6	6	6	6	6	6	6	6	6	14			

The uncertainties of the duration and cost of an activity are positively correlated, so the minimum and maximum cost distribution per unit of time h of the j-th activity at the level α depict the best and the worst scenario, respectively. An example concerning the considered problem described in the constraint programming environment is presented in the next section.

4. ILLUSTRATIVE EXAMPLE

The example consists of three subsections: the description of project portfolio, the analysis of the first admissible solution of the fuzzy scheduling problem, and the analysis of cash flow distribution in project portfolio. The analyses contain the examination of project fuzzy Gantt charts and fuzzy project cash flows.

4.1. Project portfolio description

The example concerns new products development that can be considered as project portfolio $P = \{P_1, P_2, P_3\}$. It is assumed that the time horizon for the project portfolio equals 32 months ($H = \{0, 1, \ldots, 32\}$) and the budget of project portfolio is fixed at 1,150 m.u. The network diagrams of the activities in the project portfolio are shown in Fig. 3-5.

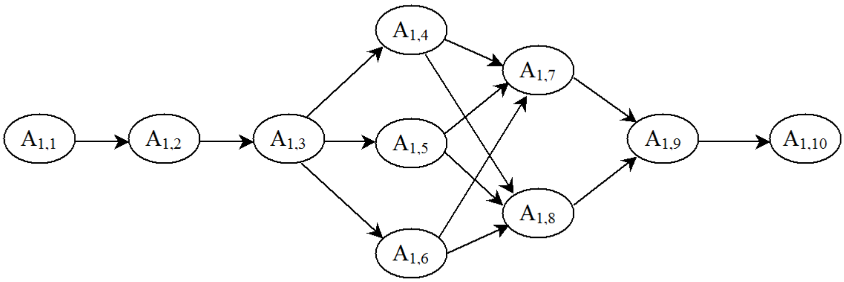


Fig. 3. Network diagram for project P_1 [source: own study]

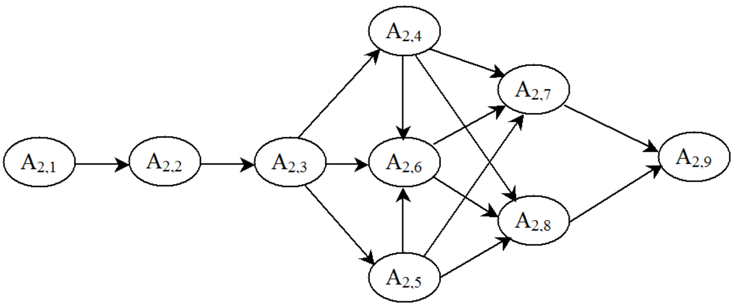


Fig. 4. Network diagram for project P_2 [source: own study]

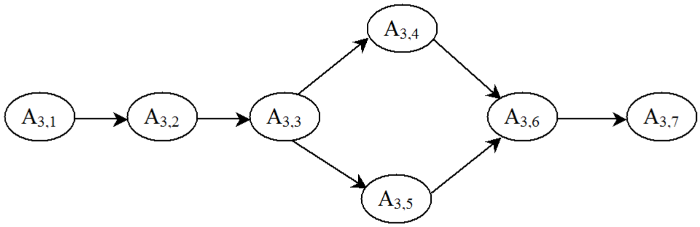


Fig. 5. Network diagram for project P_3 [source: own study]

Duration of project activities is estimated by using past experiences and/or an expert's knowledge. The different types of projects require the appropriate methods of forecasting that have been presented e.g. in [24-30]. The duration of some activities ($A_{1,7}$, $A_{1,10}$, $A_{2,4}$, $A_{2,7}$, $A_{2,9}$, $A_{3,4}$, $A_{3,5}$, $A_{3,6}$, $A_{3,7}$) is specified in the imprecise form. The sequences of activity duration for the considered projects can be described as follows: T_1 = (2, 1, 1, 6, 2, 2, "about 6", 1, 4, "about 6"), T_2 = (2, 2, 1, "about 9", 6, 4, "about 6", 4, "about 5"), T_3 = (1, 1, 1, "about 6", "about 6", "about 5", "about 4"). For instance, the duration of the activity $A_{1,7}$ is "about 6", i.e. the activity can be executed within the time period of 5 till 7 units of time.

4.2. Fuzzy scheduling

The standard routine queries formulated can be as follows: is there portfolio schedule (and if yes, what are its parameters) that follows from the given project constraints specified by the activity duration times, the deadline and budget of project portfolio? What are the fuzzy project cash flows for different risk levels? The answer to the questions is connected with the determination of the starting time of project portfolio activities s_{ij} and the allocation of financial means to the activities by different α-level $dp_{i,j,\alpha}$. For the considered project portfolio and α-level equals 1, the following sequences are sought: $S_1 = (s_{1,1}, \ldots, s_{1,10})$, $S_2 = (s_{2,1}, \ldots, s_{2,9})$, $S_3 = (s_{3,1}, \ldots, s_{3,7})$, $Dp_1 = (dp_{1,1,1}, \ldots, dp_{1,10,1})$, $Dp_2 = (dp_{2,1,1}, \ldots, dp_{2,9,1})$, $Dp_3 = (dp_{3,1,1}, \ldots, dp_{3,7,1})$.

Figure 6 presents the project portfolio schedule, in which the sequences of activity starting time are as follows: S_1 = (0, 2, 3, 4, 4, 4, 10, 10, "about 16", "about 20"), S_2 = (0, 2, 4, 5, 5, "about 14", "about 18", "about 18", "about 24"), S_3 = (0, 1, 2, 3, 3, "about 9", "about 14"). The completion time of project P_1, P_2, P_3 equals "about 26", "about 29", and "about 18" months, respectively.

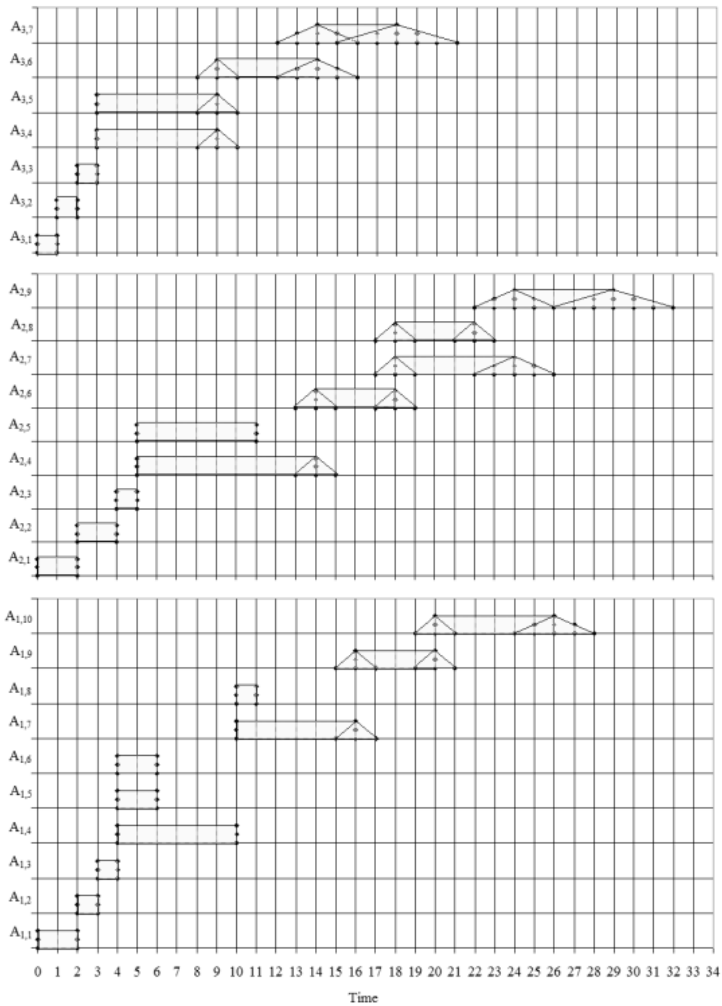


Fig. 6. Project portfolio schedule [source: own study]

It is noteworthy that using presented methodology, the level of uncertainty increases for subsequent activities according to the number of activities with the fuzzy duration. As a consequence, this can lead to the difficulties with the interpretation, because the fuzzy starting time of the activity can be greater than the fuzzy completion time. This case is presented in Fig. 6 for activity $A_{3,7}$ between 15 and 16 time unit. The proper interpretation and exploitation of the results attained by the presented method are subject to further research.

4.3. Fuzzy cash flow

Figure 7 presents five different cash flows for project portfolio (cumulative cost for project P_1, P_2, and P_3). At μ=1, the cash flow (dotted line) is equivalent to that generated from deterministic analysis. At μ=0.5, there is an optimistic scenario below and a pessimistic one above (dashed line). In turn at μ=0, the optimistic and pessimistic cash flows (solid line) have a wider spread indicating a higher degree of uncertainty. In the best case, the project portfolio will be completed in 26 months with the total cost of 921 m.u., whereas in the worst in 32 months with the total cost of 1,119 m.u.

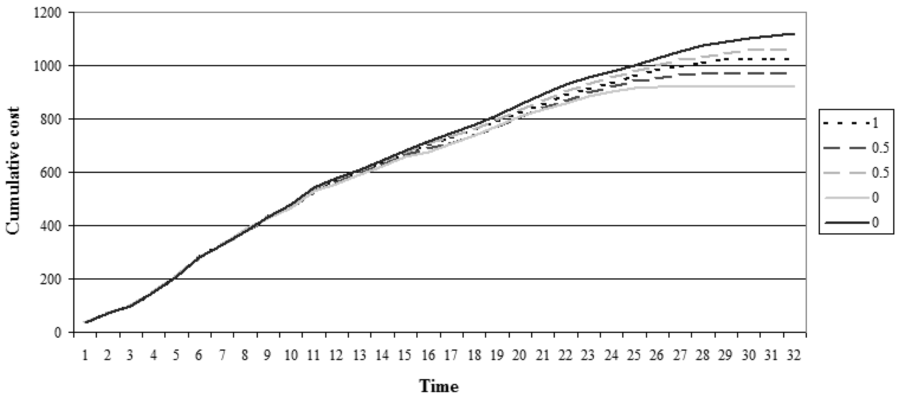


Fig. 7. Fuzzy cash flows for project portfolio [source: own study]

The presented approach allows the decision-maker to consider a wide range of analyses, including the requirement of the cost allocation in the horizon of project portfolio (Fig. 8), as well as to detail the analyses in the aspect of a single project.

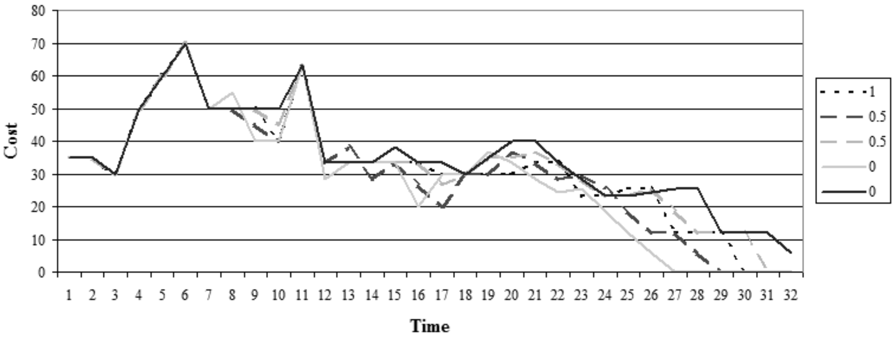


Fig. 8. Cost distribution for project portfolio [source: own study]

The feasible variants can be assessed according to a criterion such as the minimal total time, cost, relation time-cost, etc. Thus, the obtained variants provide a plan for project portfolio execution, and provide a base for further adjustment aimed at fitting to real-live execution.

5. CONCLUSIONS

Most of the projects are executed in the presence of uncertainty and are difficult to be managed, for they include many activities linked in a complex way. Hence, there is an increase in demand for new knowledge that enables the solution of problems encountered during complex project portfolio execution. In this case, knowledge concerning project management, especially project portfolio scheduling, is particularly significant. The proposed approach takes into account different form of variables (distinct, imprecise) and constraints as well as permits to formulate project scheduling problems. The model supports descriptive statement of the problem followed by its implementation in one of the constraint logic languages.

Constraint programming is an emergent software technology for declarative CSP description and can be considered as a pertinent framework for the development of decision support system software aims. CSP can always be solved with a brute force search, i.e. all possible values of all variables are enumerated and each is checked to see whether it is a solution. However, for many intractable problems, the number of candidates is usually too large to enumerate them all. CP has developed some manners (constraint propagation and variable distribution) to solve CSP that greatly reduce the amount of search needed. This is sufficient for solving many practical problems, including project portfolio planning and scheduling.

The proposed methodology can be easily incorporated into available fuzzy project scheduling software to provide a better perception of risk that is usually obscured in the conventional approach. A number of α-levels can be modified according to a decision-maker's requirements. As a result, it can assist project managers to gain deeper insight into the sources and extents of uncertainty, which consequently may lead to the avoidance of a project failure. Also, the presented approach is useful in the assessment of financial requirements during feasibility stage and project realization, as well as it may provide an evaluation of alternative proposals of a project completion. Moreover, the proposed approach tends to achieve a balance between a complexity of methodology and an intuitive, effective decision support system that is realistic in modelling uncertainty. Finally, its application in performing earned value analysis during project monitoring may also provide the useful results.

Further research focuses on the increase a number of the α-levels in order to eliminate the intersection of fuzzy starting and completion time of an activity. Moreover, future research can be aimed at comparing searching strategies in the aspect of a different number of the α-levels.

REFERENCES

[1] LASLO Z.: *Project portfolio management: An integrated method for resource planning and scheduling to minimize planning/scheduling-dependent expenses.* International Journal of Project Management, 28, 2010, 609-618.

[2] KORMANCOVÁ G.: *Project success and failure.* In: Zborník vedeckých prác: Theory of Management 6: The Selected Problems for the Development Support of Management Knowledge Base, Žilinská univerzita v Žiline, EDIS – vydavateľstvo ŽU, Žilina 2012, 117-119.

[3] CHELAKA M., ABEYASINGHE L., GREENWOOD D.J., JOHANSEN D.E.: *An efficient method for scheduling construction projects with resource constraints.* International Journal of Project Management, 19, 2001, 29-45.

[4] KASTOR A., SIRAKOULIS K.: *The effectiveness of resource levelling tools for Resource Constraint Project Scheduling Problem.* International Journal of Project Management, 27, 2009, 493-500.

[5] BUDÍK J., DOSKOČIL R.: *Soft computing as a tool to optimize an investment portfolio.* Intellectual Economics, 5(3), 2011, 359-370.

[6] KIM J.Y., KANG C.W., HWANG I.K.: *A practical approach to project scheduling: considering the potential quality loss cost in the time-cost tradeoff problem.* International Journal of Project Management, 30, 2012, 264-272.

[7] YAGHOOTKAR K., GIL N.: *The effects of schedule-driven project management in multi-project environments.* International Journal of Project Management, 30, 2012, 127-140.

[8] ZAMMORI F.A., BRAGLIA M., FROSOLINI M.: *A fuzzy multi-criteria approach for critical path definition.* International Journal of Project Management 27, 2009, 278-291.

[9] ATKINSON R., CRAWFORD L., WARD S.: *Fundamental uncertainties in projects and the scope of project management.* International Journal of Project Management, 24, 2006, 687-698.

[10] BONNAL P., GAURC K., LACOSTE G.: *Where do we stand with fuzzy project scheduling?* Journal of Construction Engineering and Management, 130, 2004, 114-123.

[11] LONG L.D., OHSATO A.: *Fuzzy critical chain method for project scheduling under resource constraints and uncertainty.* International Journal of Project Management, 26, 2008, 688-698.
[12] FORTIN J., ZIELINSKI P., DUBOIS D., FARGIER F.: *Criticality analysis of activity networks under interval uncertainty*. Journal of Scheduling, 13, 2010, 609-627.
[13] MARAVAS A., PANTOUVAKIS J.P.: *A fuzzy repetitive scheduling method for projects with repeating activities*. Journal of Construction Engineering and Management, 137, 2011, 561-564.
[14] BOCEWICZ G., BANASZAK Z.: *Abductive inference based approach to DSS designing for project portfolio planning*. Intelligent System for Knowledge Management, Series: Studies in Computational Intelligence, 252, 2009, 107-129.
[15] RELICH M.: *CP-based decision support for scheduling*. Applied Computer Science, 7(1), 2011, 7-17.
[16] RELICH M.: *Project prototyping with application of CP-based approach*. Management, 15(2), 2011, 364-377.
[17] RELICH M., WITKOWSKI K., WAŻNA L.: *The application of constraint programming to building the tool of liquidity planning in enterprise*. Management, 11(1), 2007, 129-137.
[18] VAN ROY P., HARIDI S.: *Concepts, techniques and models of computer programming*. Massachusetts Institute of Technology 2004.
[19] ROSSI F., VAN BEEK P., WALSH T.: *Handbook of Constraint Programming*. Elsevier Science 2006.
[20] DOSKOČIL R., DOUBRAVSKÝ K.: *Analysis of empirical characteristics in the PERT method*. Business Systems & Economics, 2(1), 2012, 7-19.
[21] MON D.L., CHENG C.H, LU H.C.: *Application of fuzzy distribution on project management*. Fuzzy Set and Systems, 73(2), 1995, 227-234.
[22] MARAVAS A., PANTOUVAKIS J.P.: *Project cash flow analysis in the presence of uncertainty in activity duration and cost.* International Journal of Project Management, 30, 2012, 374-384.
[23] RELICH M.: *An evaluation of project completion with application of fuzzy set theory*. Management, 16(1), 2012, 216-229.
[24] CHEBA K.: *The application of the index seasons method in forecasting decade data*. Prace Naukowe Akademii Ekonomicznej imienia Oskara Langego we Wrocławiu, 1112, 2006, 49-58.
[25] GONDEK A.: *Solutions for estimating the range of delay in time - area analogy method*. Management, 14(2), 2010, 161-170.
[26] RELICH M.: *A decision support system for alternative project choice based on fuzzy neural networks*. Management and Production Engineering Review, 1(4) 2010, 46-54.
[27] RELICH M.: *Assessment of task duration in investment projects*. Management, 14(2), 2010, 136-147.
[28] SOBIECHOWSKA-ZIEGERT A.: *Equilibrium price – modeling, forecast and application*. Econometrics: Forecasting, 28, 2010, 125-135.
[29] DITTMANN I.: *Forecasting in constructing scenarios for assessing profitability and risk of investing in real property*. *Econometrics*, 32, 2011, 83-91.
[30] KUŻDOWICZ P.: *Application of the concept of further approximations in production cost accounting*. In: Global Crises - Opportunities and Threats, Alumnipress, 2012, 109-116.

information technology, technological equipment, design, synthesis, structure

*Bogdan PALCHEVSKYI**

INFORMATION ASPECTS OF OPTIMIZATION SYNTHESIS OF FUNCTIONAL-MODULAR STRUCTURE OF TECHNOLOGICAL EQUIPMENT

Abstract

The paper presents the information approach to the design of equipment with functional-modular structure. The advantages and drawbacks of various technologies of synthesis are shown.

1. INTRODUCTION

The production cycle in machine building companies, which is the time that parts, components and finished goods spend in the shops, is only 1% of the total time from design to finished goods production. The other 99% fall on the research and development, design and technological preparation of production. Automation and computerization of these phases significantly accelerates the design process.

The most important trend in the development of methods for designing technological equipment is the creation and development of technology design with use of modern information technologies. The value of information technology in designing great – it creates the information foundation for the development of corresponding sphere of science and its applications.

* Lutsk National Technical University, 75 Lvivska str., 43018 Lutsk,
tel.: +380 (332) 767395, e-mail: pac@tf.dtu.lutsk.ua

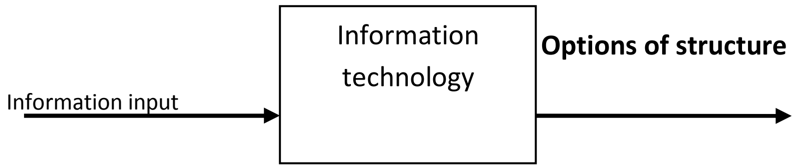


Fig. 1. The scheme of the information technology
[source: own study]

Information technology in the design – a combination of methods, processes, programming and technical means combined into a chain that provides collection, processing, storage and display of information in order to reduce the complexity of the processes of designing.

2. INFORMATION-AIDED DESIGN OF TECHNOLOGICAL EQUIPMENT

The peculiarity of using the information technology in the design of technological systems is to use them for modelling of technological equipment and workflow in it. Informational technologies in design are perfectly suited to meet the information needs in creating projects of technological equipment. It is necessary to solve the basic problem of designing – to determine dependences between the structure of the designed object and characteristics of its effectiveness.

Having some idea of the object of design (conceptual model), a person acquires background information and comparing thereafter the object with its representation forms the design action I_y.

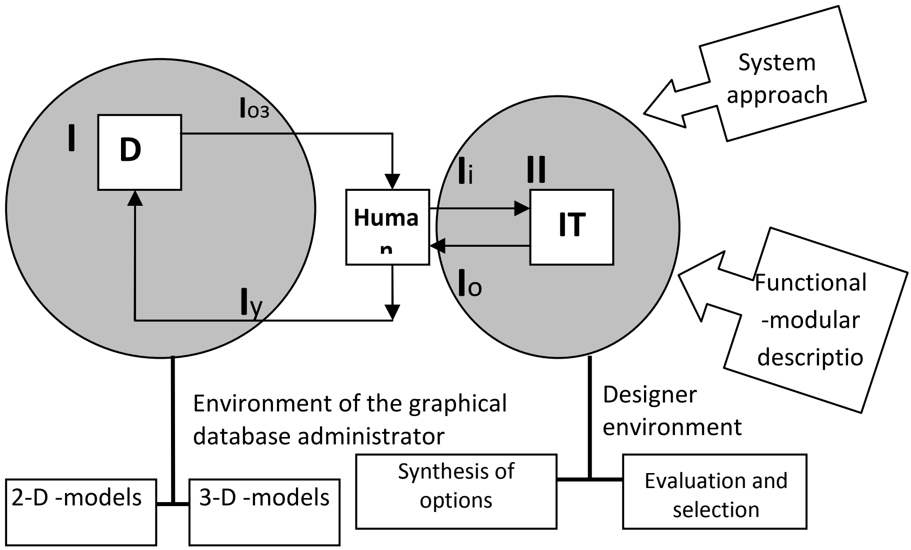


Fig. 2. Information system of computer-aided design [source: own study]

It is seen that circuit of information processing stands out, it is separated from the circuit of processing graphical information and becomes independent. The information technologies that convert input information I_{in} the product in the form of background information I_{out} are highlighted in the system. Depending on the human role in the chain I_{ip} – I_{in} – I_{out} – I_y the manual computer-aided, automated and automatic design is distinguished.

The environment of graphical database administrator allows creating and editing drawings and geometric models as well as it is designed to run on hierarchical image databases. Most CAD is reduced to drafting of the already finished technical solution with use of computer programs.

The designer environment provides a solution of all complex tasks of technological equipment designing needed to build the structural, functional and schematic diagram, carry out of necessary calculations, modelling and design. Thus, the design process includes the following stages:

- The components of technological process of packaging are distinguished – technological operations that can be conveniently called functions. These functions arrange the structure of technological process, while several types of functional modules are put in line for each of them.
- The optimal set of functional modules to perform a given task is searched for.

- Next comes the allocation of the functional modules in a certain sequence and establishing connections between them, namely to form spatial layout and functionality for material, information and energy flows in the machine.

The interaction of the defined environments is carried out by using a set of models of object that are divided into two distinctive groups.

1. The first group of models describe the functioning (working) of the future equipment and its component functional modules and processes occurring in it. These models allow us to identify the characteristics of workflows and assess their impact on the technological characteristics of the machine (its performance, reliability and other characteristics).
2. Models of the second group describe the spatial structure of the machine. They allow to determine its size, area, volume, mass and strength under loading.

Models of the first group are usually called functional. The second group of models is based on geometric proportions. They are usually referred as geometric or spatial models.

3. KEY INFORMATION TECHNOLOGIES OF TECHNOLOGICAL EQUIPMENT DESIGN

Arguably, the level of CAD automation is defined by the designer environment. Applying of the information technology in this case should allow obtaining the technical solutions for further comprehensible graphic representation. Therefore, systematic approach that involves presenting of technological equipment as a system that performs a given utility function and functional modules acting as the elements of its structure should be put as the basis for the designer environment. It allows to synthesize the equipment versatile in composition and functions from a limited set of standardized functional modules (FM).

The task of informatization of the part of the technological equipment design process is stated and solved as optimization synthesis of the equipment structure from a unified set of functional modules, i.e. on the basis of functional modularity of design.

Functional-modular principle of technological equipment design has the following features:

1. Technological machines of the same purpose perform about the same amount of technological functions.

2. The general utility function of any technological machine is performed by the finite set of functional modules, each executing a complete part of the overall service function of the machine.
3. The structure of technological machine is created by ensuring the functional and spatial connections between the functional elements.

A combination of information technology of synthesis and information technology of optimization (analysis) must be put in the basis of the information process of designing the technological equipment with functional-modular structure.

The design process of manufacturing equipment is customary split into stages (functional and technical, systematic and structural, design). Proceeding from such partition, it is a natural requirement the CAD system to support all phases and levels of the design in full. Unfortunately, in practice this approach is not fully realized.

Information process of design has a set of elementary operations – procedures, which quantity is limited. The main ones are the synthesis procedure and the procedure for selecting options. Of these various combinations of design stages are composed, while the information process of design is composed of the stages combinations.

In developing the methodology of technological systems designing is important to differentiate the process of solutions synthesis (generating options) and the procedure of limiting the diversity of solutions for selecting the best (analysis, evaluation, selection, optimization).

Direct generating of options of the machine structure can lead to their redundancy. Let the facade of 10-storey building have 24 windows on each floor. In the evening, each window can be lit or shaded, that is remain in one of two states. By defining a set of possible states of the building façade, where windows different by their location on the facade whose total number is 10 • 24 = 240 are in one of two states, we obtain:

$$N = 2^{240} \approx 10^{73}$$

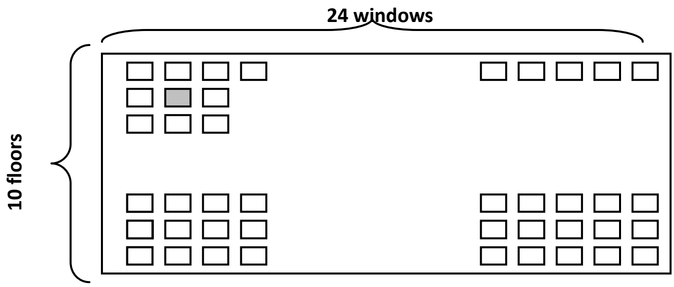


Fig. 3. Determining the number of options [source: own study]

The resulting number is extremely large; it corresponds to the number of atoms in the visible universe, which is also equal to approximately 10^{73}.

This example shows that with significant number of structure elements the set of options becomes very large, and the review and evaluation of all possible structures is impossible. For this purpose the special methods of synthesis, a feature of which is to impose additional conditions on the obtained structure, are used.

In practice, during the design process the space of solutions is restricted by studying only certain combinations of elements. As a result of selection only those options remain that have “better” value of the optimization function compared to the previous iteration. In order not to miss the best solution in the design process one need to include the information synthesis procedure reducing capacities of sets obtained at different stages, namely:

1. Technology of the directed synthesis. The design process is an analysis of the product and determining the sequence of the formation of components. The result is a sequence of technological operations, which specifies the sequence of functional modules during the synthesis of structure.
2. Synthesis technology with limited set of options based on typical mathematical models which allows to consider a number of restrictions for the options of structure.
3. Synthesis technology with limited set of structure options by using the "and-or" tree. This tree contains technical solutions in a compact form of information about a specific, pre-formed set of options for the structure of a technical system.
4. Technology of stepwise synthesis and optimization. The design process is an alternation of procedures for forming subsets on different hierarchical levels and their evaluation to select the best one.

5. Technology of hierarchical synthesis. The design process is an aggregated representation of the workflow so that every bigger technological operation can be qualitatively assessed using quality measures for its implementation. Each larger function is associated with a block of functional modules that implements the function. Then when combining options modular structure the set of options considered is much smaller.

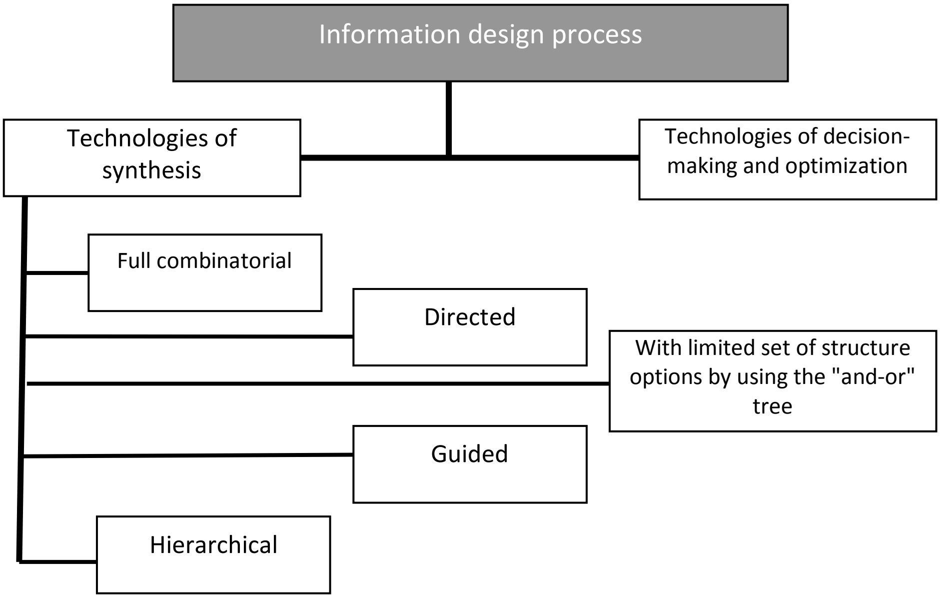


Fig. 4. Specialized methods of the synthesis procedure [source: own study]

4. TECHNOLOGY OF THE DIRECTED SYNTHESIS

It is advisable to improve design technology in a direction that every step taken from concept to finished project was an information procedure on multiple models of the object design. It creates conditions that, starting with the first and ending with the last stage of the project, all information on the results of one phase is used in the following stages during the design. Every previous model actually specifies the additional restrictions that reduce the number of options obtained with combinatorial synthesis on the next step (Figure 5). In this case, the process of structural synthesis enables the "top-down" design and gets the optimal structure that fits in the above specification and boundary conditions.

The sequence of steps of the workflow defines the product design, analysis of which allows to set the sequence of the forming its components. As a result, a sequence of technological conversions that form a workflow is determined. The structure of a workflow defines a sequence of functional modules in the structure of the machine.

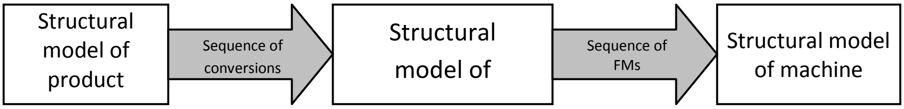


Fig. 5. Sequence of the system models use in the directed synthesis [source: own study]

The design begins with the analysis of the product. Analysis of the structure of traditional machining processes, for example, shows that the sequence of creating surfaces of a detail is chosen in most cases purely arbitrary. This sequence is set by a designer's intuition and is far from optimal. If an elementary technological conversion is put in line with the creation of each elementary surface, then it can be assumed for any a single conversion that the conditions for its realization will depend on some previous conversions. On the other hand, this technological conversion, in turn, can affect the implementation of the following ones. That is the surface created on the step is necessary to implement some of the next conversions. Thus, each conversion is characterized by technological input and output cause-ties of in the scheme of machining and therefore should occupy an appropriate place in the structure of processing. In the most general case, the functional, design and technological limitations affect the sequence of surface processing, that can distinguish three groups of contradictions namely:

- functional connection of precedence imposed by the conditions of functioning parts (auxiliary surfaces at first, next the functional surfaces);
- design connection of precedence imposed by the terms of spatial arrangement of surfaces in parts (datum references are processed initially);
- technological connection of precedence imposed by the conditions of part machining.

To determine the logical sequence of processing the functional, geometric and technological precedence ratio that are imposed on the surface details can be written in the form of a precedence matrix.

The surface to be processed previously						
П –the processed surface	П1	П2	П3	П4	П5	**Bo** Level of dependence of the surface
П1						0
П2	1					1
П3	1					1
П4		1	1			2
П5				1		1
Level of following of the surface	2	1	1	1	0	

In this matrix, each connection of precedence between two surfaces is marked as 1. Through formalized procedures developed by A. Kaufmann, a sequence of technological conversions of processing the surfaces of a part is derived divided into 4 stages of processing (Fig. 6).

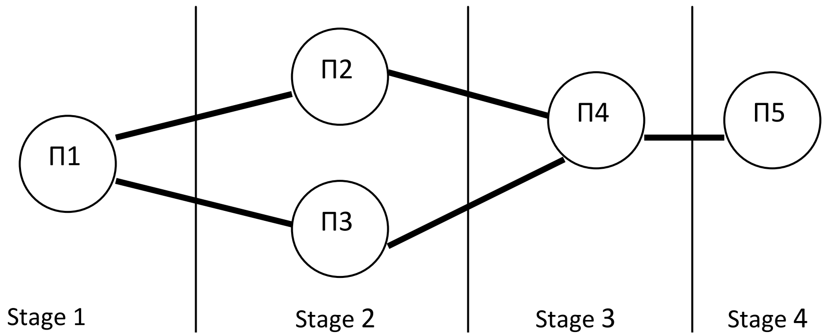


Fig. 6. Graph of precedence connections for processing the П1-П5 surfaces [source: own study]

The function of the equipment that performs technological operation implements complex technological transformation that includes a number of simple conversions performed by individual functions and subfunctions φ_i:

$$F \to \{\varphi_1 \cdot \varphi_2 \cdot \varphi_3 \cdot ......\}$$

A subfunction generally consists of two or more elementary functions. The advantage of this method is that for the required performance function several alternative models of workflow can be developed and selected for their optimal suitability at a relatively small cost of time. The structure of the workflow, shown in Figure 6, defines partially ordered sequence of formation of the product, for example, by processing blanks with the state S (0) to the final state S (P5):

$$F = C_0(0)\varphi_1 C_1(\Pi 1) \begin{matrix} \nearrow \varphi_2 C_2(\Pi 2) \\ \searrow \varphi_3 C_3(\Pi 3) \end{matrix} \Bigg\} \varphi_4 C_4\,(\Pi 4)\varphi_5 C_5(\Pi 5)$$

This record of a workflow structure means the following. Originally the workpiece transferred by means of transition φ1 from the state C0 to state C1 (P1) by processing the surface P1. Next a conversion φ2 or φ3 is possible. A part is shifted into a state C4 (A4) with the φ4 conversion and so on down to a final state C5 (P5). One can see that the order of conversions φ2, φ3 is not regulated. Because even this simplified scheme can build several options of processing that differ by type of machining equipment, degrees of processing concentration, parallel or serial execution of certain operations, the next step is to search for the optimal structure of the part machining process, but the estimated number of options will be significantly reduced.

Under the procedure of constructing the system model of the process is to be understood such activities that are necessary to ensure that workflow in equipment is provided by a structure derived from subfunctions or elementary functions that meet the technological conversions at various levels of complexity.

Functional modeling is a key moment of forming the structure of technological machine because the function model serves like a roadmap that helps designers to reach the desired layout of the machine.

The design process is based on the presence of a design scheme, with alternating design procedures of two types – synthesis and optimization of technical solutions (Fig. 7). Structural design is a combination of possible combinations of technology machine's FMs, getting set of layouts and their evaluation for choosing the best. The applied here method of directed search (as the machine workflow defines the sequence of FMs) provides consistent consideration of each of the options structure during sorting.

Consider, as an example, the structural design of the technological machines for the packaging of loose substances.

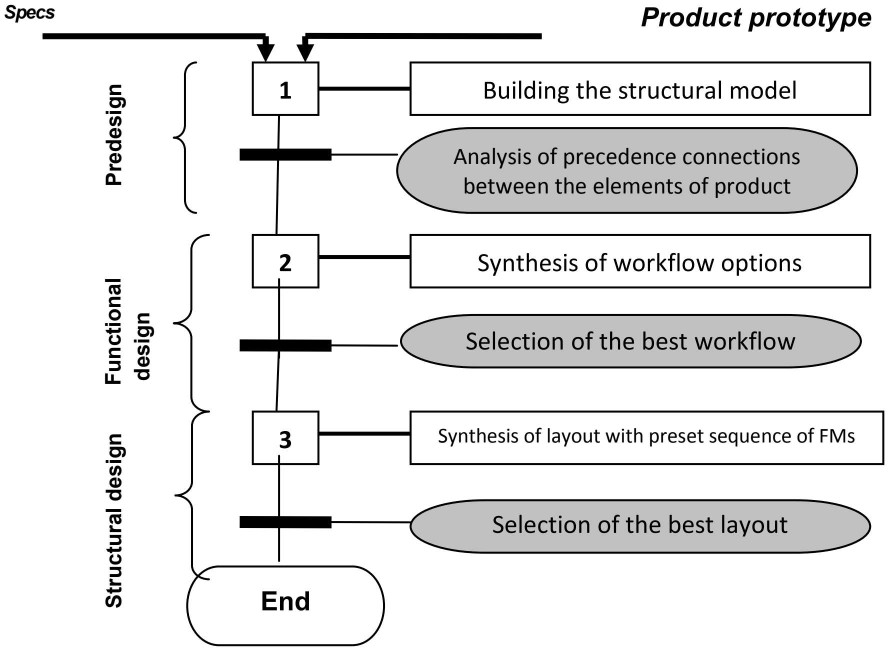


Fig. 7. Algorithm for the directed synthesis of technological equipment [source: own study]

Technological packaging operation includes technological transitions (x1 – x5) sequence of which can be considered as given:

- forming a tube blank of the package from the polymer film (x1);
- dosing the mass of substance (x2);
- longitudinal welding of package (x3);
- transverse welding of package with a dose of substance in it (x4);
- pulling a tube blank of the package (x5).

Since the implementation of each conversion makes possible the use of FM differing by design and characteristics (reliability, speed, energy consumption, cost, etc.), the problem of packing machine structure synthesis becomes multivariate. Methods of directed search are based on the possibility of gradual rejection of unpromising directions finding the optimal solution (Fig. 8). The process of designing is reduced to finding a way (1 – n) on the model that provides the best value of the machine quality. Problems that hidden behind such a simple algorithm become evident while increasing the number of FM in the structure of the machine.

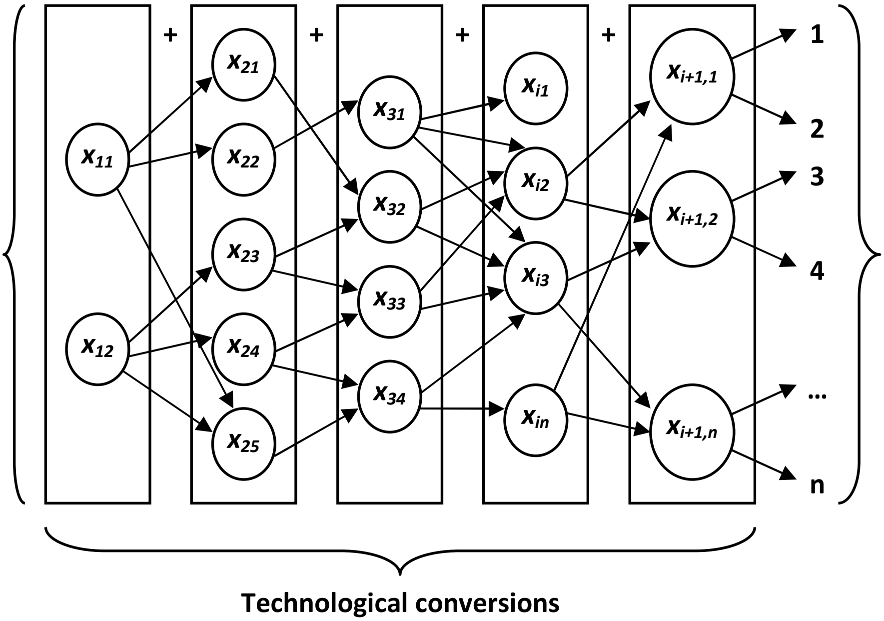


Fig. 8. Model of optimization synthesis of a packaging machine of functional-modular structure with preset sequence of FMs [source: own study]

5. SYNTHESIS TECHNOLOGY WITH LIMITED SET OF OPTIONS BASED ON TYPICAL MATHEMATICAL MODELS

It allows to take into account a number of constraints on the existence of options structure. The following characteristics of differences between versions of the structure are taken into account:

1. qualitative composition of elements, ie, possible variations of elements that create the object. This feature of the mathematical model is denoted as FE;
2. number of elements that create the object of synthesis. This attribute of the model is denoted as FN;
3. the order of elements in the structure of the object of synthesis. This attribute model is denoted as Fp.

If the content of elements of all the options while generating the structure is the same, then FE = 1 else – that FE = 0. If all the options of structure have the same number of elements, then FN = 1 else FN = 0. If a sequence of combining elements in the structure of all the options is the same, then fp = 1 else – Fp = 0.

Depending on the generated object with certain features of differences between the structure options, the following classes and subclasses of standard mathematical models (Table 1) are used.

Tab. 1. Classification of mathematical models used for generating the options of structure

Model			Characteristic of generating the option of structure		
Class	Subclass	Name	F_E	F_N	F_{Π}
Conjunctive	0	MS01	1	1	-
		MS02	0	1	-
		MS03	0	0	-
Organizing	Table – 1	MS11	1	1	1
	Network – 2	MS21	0	1	1
		MS22	0	0	1
	Permutation 3	MS31	1	1	0
		MS32	0	1	0
		MS33	0	0	0

Let us consider example of a typical model MS31 to generate variants of the technological process of assembling the product. Technological operation of assembly is associated with installation of one of the components (Figure 9).

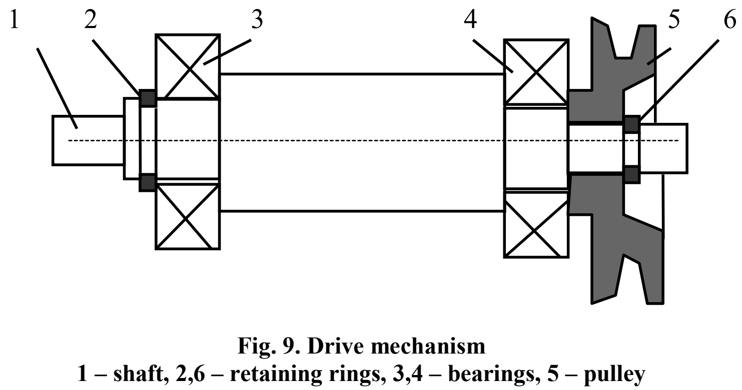


Fig. 9. Drive mechanism
1 – shaft, 2,6 – retaining rings, 3,4 – bearings, 5 – pulley
[source: own study]

All possible sequences of combinations of parts are described completely by graph on six vertices. Each of the options of assembly process is set by a covering tree of this graph. Total number of covering trees and thus options of assembly sequence is determined by Cayley's theorem and is:

$$n_{TП} = p^{p-2} = 6^4 = 1296,$$

where: p – is a number of vertices.

Obviously, the set of assembly processes obtained by formal enumeration of graphs include many unrealistic and absurd choices, the analysis of which is unnecessary. Therefore, the use permutation model which significantly limits the number of options eligible for consideration, greatly simplifies the process of optimization synthesis. First of all, only the variations of the process, starting from the operation of setting the part 1 (basic part of the mechanism) are taken into account. When considering the possibilities to access other parts the number of options process will be reduced to ten (Table 2). In Table 2 the following notation is used: 1 – install the shaft 1 in the assembly device; 2 – install retaining ring 2; 3 – install the bearing 3; 4 – install the bearings 4; 5 – press the pulley 5; 6 – install retaining ring 6.

Tab. 2. Options of the drive mechanism assembly process

№	Assembly sequence	№	Assembly sequence
1	1-3-2-4-5-6	6	1-4-3-5-6-2
2	1-3-4-5-6-2	7	1-4-3-5-2-6
3	1-3-4-2-5-6	8	1-4-5-3-6-2
4	1-3-4-5-2-6	9	1-4-5-3-2-6
5	1-4-3-2-5-6	10	1-4-5-6-3-2

6. TECHNOLOGY OF THE STEPWISE OPTIMIZATION SYNTHESIS

The methods controlled search are based on the possibility of gradual rejection of unpromising directions while finding the optimal solution. The design process is the successive addition of the specific functional elements to the machine at each stage and the evaluation of the result. The assessment procedure is shown as a logical transition to the next stage of design (Fig. 10).

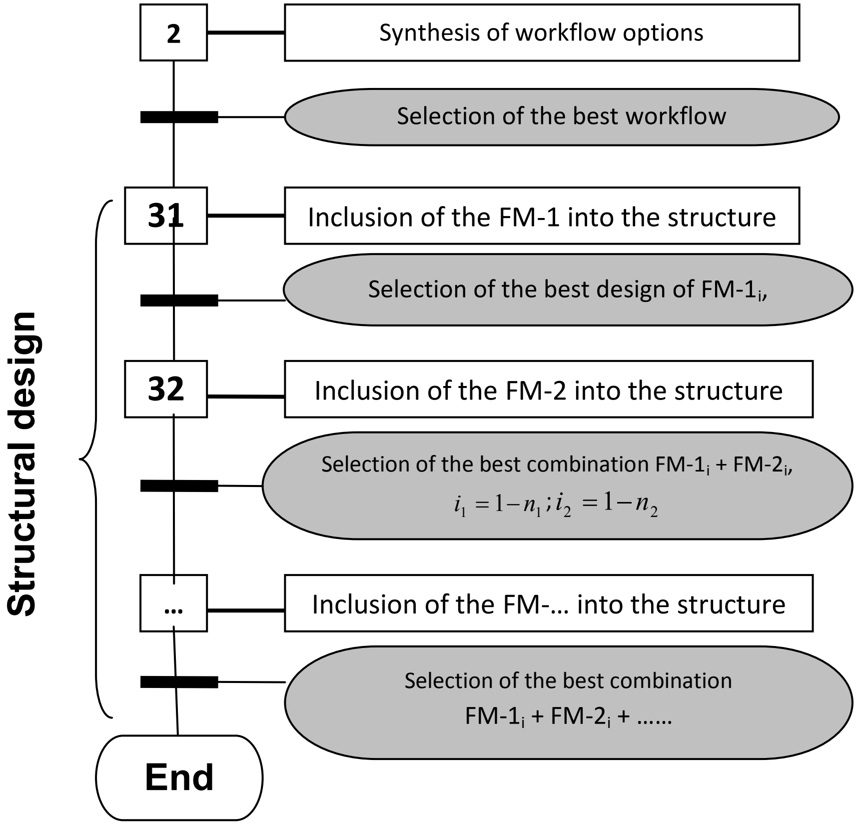


Fig. 10. Algorithm for the stepwise directed synthesis of technological [source: own study]

The peculiarity of this synthesis is additivity of the criterion for evaluation and selection of layout design at each step. These criteria include the cost of FM, its weight, energy consumption, etc. It should also be noted that the sequence of joining FMs to a machine can be both direct, i.e. starting from first technological conversion FM and reverse when step by step synthesis starts from the last (terminal) FM.

7. TECHNOLOGY OF HIERARCHICAL SYNTHESIS

To facilitate the search for proper layout and accelerate virtual search extensive libraries on the principle of building blocks performing the most complicated part of the machine workflow, are currently being formed.

The design process is an aggregated representation of the workflow of equipment so that every bigger technological operation or conversion can be assessed using quality measures for its implementation. Each enlarged function is associated with a correspondent block of functional modules. Thus, when combining the modular structure the set of options considered becomes significantly smaller.

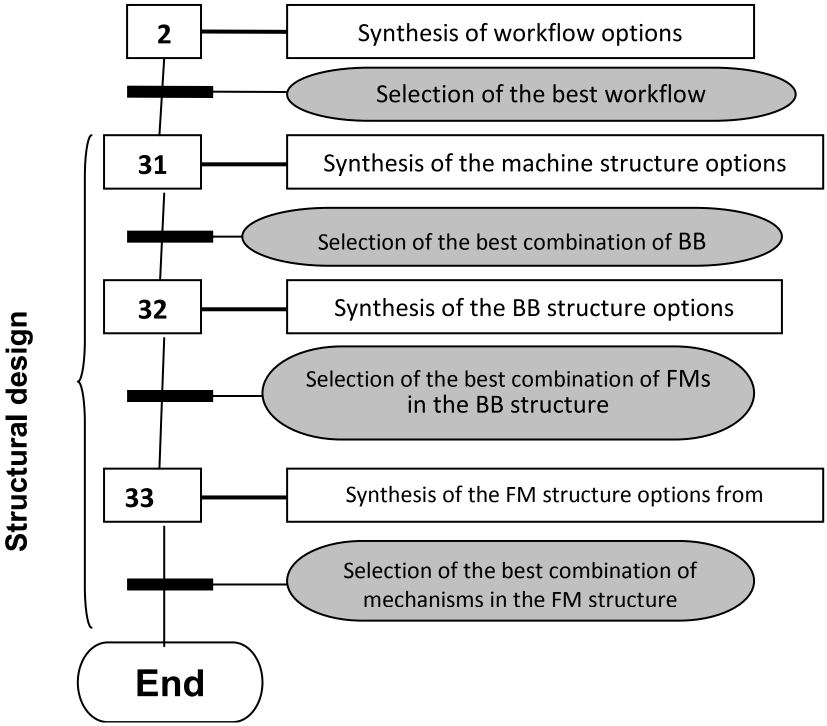


Fig. 11. Algorithm of hierarchical synthesis [source: own study]

Applying the hierarchical synthesis of functional-modular structure allows presenting a packing machine as a combination of three units: BB-1 – a dosing unit, BB-2 – package forming unit and BB-3 – auxiliary functions unit (Fig. 12).

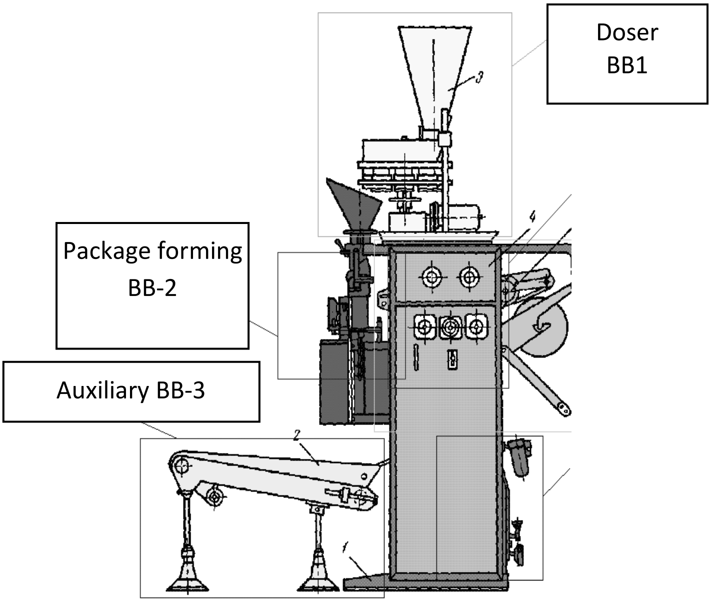


Fig. 12 Packaging machine design of three BBs [source: own study]

A distinctive feature of the use BB is as follows:

1. Few parts in BB can be developed more thoroughly using the methods of analysis and modeling that are inaccessible to a large array of connections. This allows to obtain more reliable results.
2. Second, a smaller number of structural units facilitates combinatorial processes.

REFERENCES

[1] ПАЛЬЧЕВСЬКИЙ Б.О. *Інформаційні технології проектування технологічного устаткування.* Монографія.-Луцьк: Луцький НТУ, 2012.-572 с.

[2] PALCZEWSKI B., *Zasady modularyzacji w projektowaniu technologicznych systemów pakowania*. Zeszyty naukowe Politechniki Rzeszowskiej, 2006, № 230, MECHANIKA, z.67.- s. 51-60.

[3] ПАЛЬЧЕВСЬКИЙ Б.О. *Оптимізаційний синтез модульної технологічної системи.* Машинознавство, 2001, № 6.- с.26-30.

genetic algorithm, simulation modeling, manufacturing system

*Tomasz WIŚNIEWSKI, Szymon RYMASZEWSKI**

THE USE OF SIMULATION AND GENETIC ALGORITHM WITH DIFFERENT GENETIC OPERATORS TO OPTIMIZE MANUFACTURING SYSTEM

Abstract

The article depicts an evolutionary approach to simulation based optimization of a typical manufacturing system. Genetic algorithm with four different variants of genetic operators (crossover operator and type of selection) is compared to find the best optimization method. A comprehensive discussion of the genetic algorithm results obtained from the simulation model was also presented.

1. INTRODUCTION AND A LITERATURE REVIEW

Profit from operating activities is an obvious aim of any manufacturing system. Therefore, optimization of production systems increases their productivity, flexibility and generally financial benefits. The aim of this paper is to compare the genetic algorithm of different types of genetic operators (selection method and crossover operator) for the optimization task in a typical manufacturing system. Exact algorithms within a reasonable time frame can only solve small problems. Thus, heuristic and metaheuristic algorithms (especially evolutionary algorithms) have been widely applied to solve problems in complex manufacturing systems [1]. The computer simulation environment Arena from Rockwell has been used for this purpose. Typical model of the manufacturing system will be optimize using Visual Basic (VBA) implemented with genetic algorithm.

* West Pomeranian University of Technology, Szczecin,
Department Of Information Systems Engineering, 71-210 Szczecin, Żołnierska 49,
e-mail: twisniewski@wi.zu.edu.pl, srymaszewski@wi.zu.edu.pl

A simulation modelling as an evaluative tool for stochastic systems has facilitated the ability to obtain performance measure estimates under any given system configuration [2]. Simulation is a very powerful tool often used in the design phase of manufacturing systems. Performance of various layout alternatives can be studied using simulation. Important features of the simulation indicated in [3] are: a possibility to study production systems without interfering with the real production system, and the ability to compress time, depending on your needs.

Genetic algorithm (GA) is the universal tool for combinatorial optimization problems. It belongs to evolutionary algorithms and have been applied to a variety of function optimization problems. GAs were shown to be highly effective in searching a large, complex response surface even in the presence of difficulties such as high dimensionality, multimodality and discontinuity. Many evolutionary algorithms have been developed in literature and implemented to solve manufacturing problems, due to the qualitative character of the variable and scale of the problem. This methodology is used in many fields such as manufacturing, engineering, business, science, etc [4].

The GA is good enough tool that is used not only to the optimization of the production systems, but also in forecasting such as energy consumption [5] and job shop scheduling problems [6]. The publication of Paul and Chanev [4] shows how, in practice, one can use computer simulations with GA to optimize the real production system. They created a simulation model of foundry, depending on control parameters such as the number of cranes, furnaces, and the number and size of cars. The use of optimization using a simple GA caused that the system has become less expensive and it did not generate waste. Entriken, Vössner [7] used the simulation with GA to optimize the production line of printed circuit boards. Zhang and Gen [6] presented a multistage operation-based GA, simplifying the chromosome representation to apply crossover and mutation operators in an optimal strategy. Cakar and Yildirim [8] propose a neuro-genetic decision support system coupled with simulation to design a job shop manufacturing system to obtain the optimal amount of resources in each workstation in conjunction with the right dispatching rule to schedule.

A popular control parameter analyzed in manufacturing systems is the size of the buffer. Its performance affects the allocation of the production line and the cost of production in progress. Storing too much Work-In-Progress will also be a cost associated with the freezing of capital. So it is important to optimally determine the value of the buffer in order to ensure maximum system performance with the lowest possible cost of Work-In-Progress. Can, Beham and Havey [9] devoted a whole article to the problem of buffer allocation and optimization of this problem by using GA and simulation. Many authors examine different variants of GA [1]. Eskandari et al. [10] presented an improved GA applied to the multicriteria optimization problems of simulation modeling. Konstam et al. [11] indicated the effectiveness of distributed GAs with multivariate crossover in optimization of function with a large number of independent variables. Their results showed that this algorithm has the unique potential to optimize such a function.

In this paper the multi-product manufacturing system is analyzed, i.e. parts for different kinds of final products are processed at the same time within a single manufacturing system. Each workstation in the system is preceded by a buffer. The criterion adopted in optimization task is profits derived from production. Control parameters during the optimization process are the size of buffers and amounts of resources (people, machines) in a given workstation. The study will use four variants of the genetic algorithm which differ from each other in types of crossover operator and the type of selection. Two types of crossover operator will be applied, and two types of selection methods. Each operator will be examined with any type of selection which will allow to compare four different types of genetic algorithm, and select the most suitable to the task of this kind.

The rest of the paper is organized as follows. Built simulation model with the control parameters was presented in Section 2. Section 3 contains a general scheme of optimization with a descriptions of criterion function and genetic algorithm. Section 4 is an analysis of the results obtained in simulations. Section 5 contains a summary of the discussion.

2. PRESENTATION OF THE MODEL

A typical manufacturing system was modeled to show the opportunities of using computer simulation. The model was constructed in Rockwell Arena simulation environment. There are five workstations and three types of product. Each product in the model passes through the system according to different sequences, which results in that the workstations have not equally utilization. Each workstation is preceded by a buffer of different capacity. An important parameter for the workstation is the amount of resources (capacity). Best value for this parameter for each of the workstation will be searched during the optimization process. Capacity values for individual workstation may range from 1 to 4. Customers orders' arrival patterns are exponential (100 pieces of each type of product a day). Components wait before first buffer, which do not generate the cost of Work-In-Progress. If only space in first buffer is available component goes through it technological itinerary and since then starts to generate the cost of Work-In-Progress (0.5 units per hour). Processing times for each workstation, which is different for each type of product, has been shown in

Table 1. Utilization costs is set at 60 units and cost of idle time is set at 50 units. Utilization costs is higher because of the assumption that while work the material is consumed.

Tab. 1. Processing times for each workstation (min.)

	Station 1	Station 2	Station 3	Station 4	Station 5
Product 1	10	8	10		10
Product 2	10	12	10		14
Product 3	10			12	12

2.1. Control parameters

Buffer sizes and resource capacity at each workstation have been selected as control parameters that could be used to optimize manufacturing system performance. Choice was dictated due to the fact that these are the parameters that you can easily change in practice compared to the others, and they allow significant improvement of the system through small changes. The parameters are discrete – different size of the buffer before each workstation takes integer values in the range from 1 to 32. While resource capacity at each workstation takes integer values in the range from 1 to 4. Properly selected system control parameters allow to affect the amount of Work-In Progress, eliminate the bottlenecks and provide the possibility of adjusting the productivity of the system. Number of different combinations of all these parameters is: 32^5 (buffer size) x 4^5 (resource capacity) = 34 359 738 368 combinations. It means that there are so many different settings of the system using only these indicated parameters, that will cause a very time-consuming calculations. Therefore, to find the optimal settings the GA will be used. Simulation runs parameters were selected in order to assure reliable results, hence the length of each replication was set at 30 days and 16 working hours a day. It can be assumed that this corresponds to a month of work in two shifts, 7 days a week. Each simulation run consists 3 replication and warm-up time is set to 1 day.

3. GENETIC ALGORITHM

Genetic algorithm (GA) belong to heuristic method that is frequently used with simulation-based optimization. GAs are based on the mechanisms of natural evolution. In this paper we compare four types of GA with different kind of selection and crossover. Algorithm was implemented in VBA. It changes control parameters of the model and the fitness function value is returned for each created solution. A general structure of the developed algorithm is described below:

STEP 1 Initialization (Create a population with random chromosome)
STEP 2 Fitness function evaluation (Evaluate a function value for all chromosome in population)
STEP 3 Stop condition (If true then end)
STEP 4 Selection (Select parents for offspring in next population)
STEP 5 Crossover (Crossover select parent and create offspring for next population)
STEP 6 Mutation (Go to STEP 2) (Mutation gene in offspring and create new population)

In GA solution is called chromosome which is a set of genes (one gene is one variable). A set of chromosome is a population. The main principle of operation in GA is created new population by using previous population of chromosomes and genetic operators. In this case the chromosome is encoded binary. Each parameter is encoded by the corresponding number of genes, what is shown in table 2. First five genes corresponding with buffer size before workstations and next are corresponding to the resource capacity at workstations.

Tab. 2. Chromosome encoding

Buffer 1	Buffer 2	Buffer 3	Buffer 4	Buffer 5	Cap1	Cap2	Cap3	Cap4	Cap5

Step 1 Initialization

The first step in the action of GA is generating, in a random way, a new population of chromosomes. Generate of new population occurs only once when the simulation is starting. The population select randomly *k* individuals with random genes. The value of each gene – 0 or 1 occurs with probability ½.

Step 2 Fitness function evaluation

A discrete-event simulator is used to evaluate the fitness function of each chromosome, which is calculated as a mean value obtained from set of replications. In our case the value of fitness function is maximized. The fitness function is given by equations 1 and 2.

$$CF = \sum_{i=1}^{n}[(A_i - B_i)\cdot C_i] - D - E \quad (1)$$

$$B_i = \sum_{i=1}^{n}\sum_{j=1}^{k}(W_{ij}\cdot K_j) \quad (2)$$

where: i – type of product (1,…n),
k – number of workstation (1,…k),
CF – fitness (criterion) function,
Ai – vector of selling prices of i-th product,
Bi – vector of unit costs of i-th product,
Ci – vector of quantities of i-th product,
D – idle cost,
E – cost of Work-In-Progress,
Kj – vector of costs for j-th workstation,
Wij – matrix of production time posts for i-th product at j-th workstation (table 1).

Step 3 Stop condition

The GA stops after constant number of generations. In our case – 100 generations. Until stopping condition is not fulfilled, the algorithm moves to the next step.

Step 4 Selection

GAs are inspired by nature, and therefore are more likely to draw individuals in order to have a better fitness function value. The selection process involves the selection of chromosomes to be parents of the offspring in the next population. In this paper the algorithm used a two different selection strategy, roulette wheel and tournament selection. The roulette wheel selection (proportional selection) operator is developed by Holland [12] and is used in many GA studies. The principle of this type of selection is to assign the probability of selecting to the each individual in each population. Chromosomes with better value of fitness function have better genetic material so they should go into the next population. The probability of drawing is formed by the following scheme:

$$\Pr_j = \frac{CF_j}{\sum_{j=1}^{k} CF_j} \quad (3)$$

where: Pr_j – probability of drawing j-th chromosome,
j – index of a chromosome in the population (1,..,k),
CF_j – fitness function value for j-th chromosomes.

Second type of selection is tournament. First step is draw a sets of chromosomes from current population. In each set the tournament between individuals is executed. Individual with the best value of fitness function win the tournament. This chromosome with best genetic material goes to the next population. Additionally the algorithm applies an elitist strategy, which means that the chromosomes with the best fitness function value go to the next population. This avoids losing the best genetic material.

Step 5 Crossover

After selection, a crossover operation is carried out. This procedure generates offsprings from the selected parents. The offspring are a combination of both parents. In this paper we are used two types of this genetic operator. First is one point crossover. It creates an offspring by copying the genes from 1 to *cp* from first parent, where *cp* is a random crossover point, and genes from *cp* to last gene form second parent. The second offspring is formed from the remaining genes. One point crossover is shown by pseudocode:

1. Select random point of crossover *cp{1,…,n-1}*
2. for *i = 1* to *cp* do
3.   *offspring1(i) = parent1(i)*
4.   *offspiring2(i) = parent2(i)*
5. end do
6. for *i = cp+1* to *n* do
7.   *offspring1(i) = parent2(i)*
8.   *offspring2(i) = parent1(i)*
9. end do

where *n* is the length of chromosome

Second type of crossover is multivariate crossover. In this type the chromosome is divided into vectors of parameters and crossing these parameters depending on the random factor. This is a pseudo code for multivariate crossover type:

1. Each parent chromosome is divided into a p_{ij} chains and offspring chains are of_{ij}, *q* is a number of parameters contained in the parent chromosome, *i* is the number of individuals and *j* is the number of parameter.
2. for *j=1* to *q* do
3.   if $Rnd \leq cp$ then
4.   for *i = 1* to *cp* do
5. $of_{1j}(i) = p_{1j}(i)$
6. $of_{2j}(i) = p_{2j}(i)$
7.   end do
8.   for *i = cp+1* to *n* do

$of_{1j}(i) = p2_j(i)$
$of_{2j}(i) = p1_j(i)$
end do
Else
for $i=1$ to p_{ij} length
$of_{1j}(i) = p_{1j}(i)$
$of_{2j}(i) = p_{2j}(i)$
end do
end if
end do

Step 6 Mutation

After selection and crossover operations all obtained offsprings may be subjected to mutation operation. For set probability one gene (from chromosome) is selected for mutation, then its value is reversed to the contrary. After mutation offsprings and champion chromosome created a new population and the algorithm returns to step 2.

4. RESULTS ANALYSIS

Simulation study was carried out with a presented earlier typical model of manufacturing system and different combinations of genetic operations: selection parameter and crossover in GA. Number of individuals in the population is 30, and the amount of replication of the population is 100 (amount is result of large simulation time). The probability of mutation is 0,035. Here are the four analyzed combinations of the GA:

1 – proportional selection, one-point crossover
2 – proportional selection, multivariate crossover
3 – tournament selection, one-point crossover
4 – tournament selection, multivariate crossover

Progress of GA will be presented for each combination. Figures from 1 to 4 showed the average criterion function (CF) for each generation and obtained the best value of CF. In each case repetiton with the best average value of CF gave the best result for the maximum value of CF. For each combination of selection and crossover was carried out three independent repetitions. In one repetition was performed 3000 simulation runs (30 individuals x 100 generation) with parameter values dependent on the chromosome. In figures 1 to 4 bottom line on the graphs represents average values of CF for entire population in 3 independent runs of GA. Top line indicates maximum value of CF from these runs.

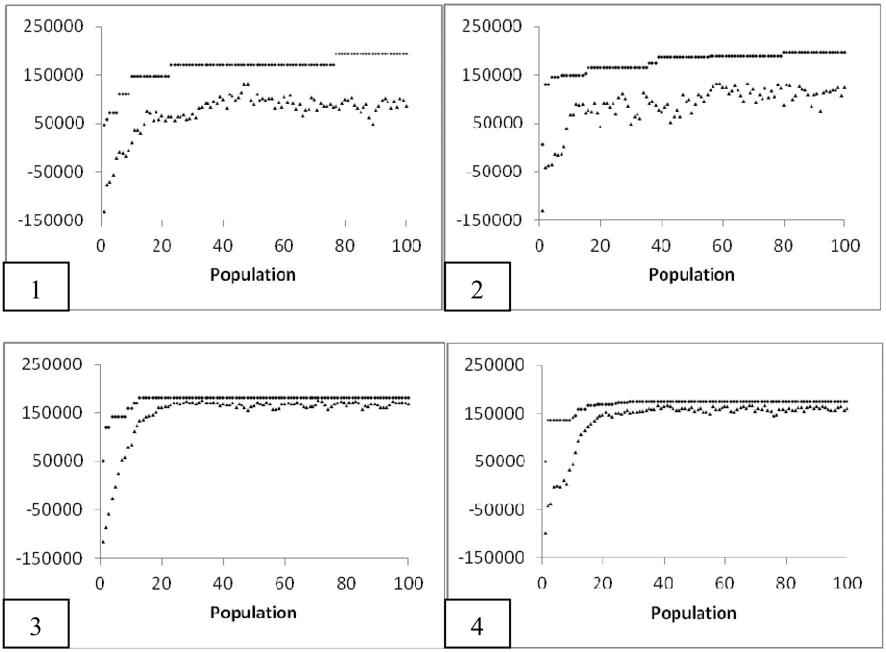


Fig. 1 – 4. The average and maximum value of CF for each algorithm combination [source: own study]

The results obtained for combinations 1 and 2 (Figures 1 and 2) showed the diversity of individuals in the population, as the average value for the entire population is significantly different from the maximum result obtained in all repetitions performed for these settings. The average value of CF, does not stabilize with each successive generation, which again shows the wide variety of individuals in successive generations. Wide variety of average values in successive generations gives a chance to not get stuck in local extremes. In the initial 20 generations it can be seen the considerable increase in the average value of CF. The large difference between the average and maximum value of CF leads to the conclusion that the best individuals was not often preferred as the parent. All three replications gave the average results which oscillating in the same bounds, which means similar adaptation of each population. Average values of CF obtained for the combinations 3 and 4 (figures 3 and 4) are close to the maximum values of CF obtained for these combinations. This proves a good adaptation of the population. Average values of CF significantly increased in the first 20 generations, which indicates a continuous improvement in the adaptation of the population. The average values close to the maximum result means frequent choice the best individual as a parent.

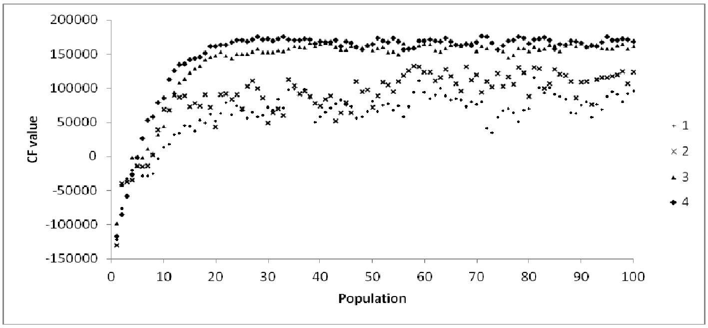


Fig. 5. The best average value of CF for each combination (symbols from 1 to 4 indicate best average value of CF from each combination) [source: own study]

Figure 5 is a comparison of the average value of the best average value of CF for each combination. Combination 1 gave the worst average result. It is very unstable and indicates the wide variety of individuals. Combination 2 gave a slightly better result, the average values are higher. Combination 3 gave a much improvement in terms of the average value of CF, is significantly higher than combinations 1 and 2, and slightly worse than the combination 4. The results showed that the average population adaptation depends on the method of selection and type of crossover. If we want to obtain a population where most individuals are well adopted and the average value of CF is close to the maximum value it should be used tournament selection method, which explicitly promotes individuals with better adaptation than the proportional selection, which gives worse average value of CF, but provides a greater variety of individuals with much different adaptation. When it comes to crossover, the average values indicate the minimum advantage multivariate crossover. If we assume that much greater impact on the average value of CF has the selection method, the results for combination 2 and 4 are better than the results for combination 1 and 3.

Figures form 6 to 9 show the value of CF for all individuals in successive generations of the GA populations for combination 1 to 4. For combination 1 and 2 populations contain a very diverse individuals as can be seen by a large range of values which CF reaches across generations. By way of elitism in any population there is an individual with the maximum CF value achieved in the previous populations, so that there is a chance of crossing the best individuals.

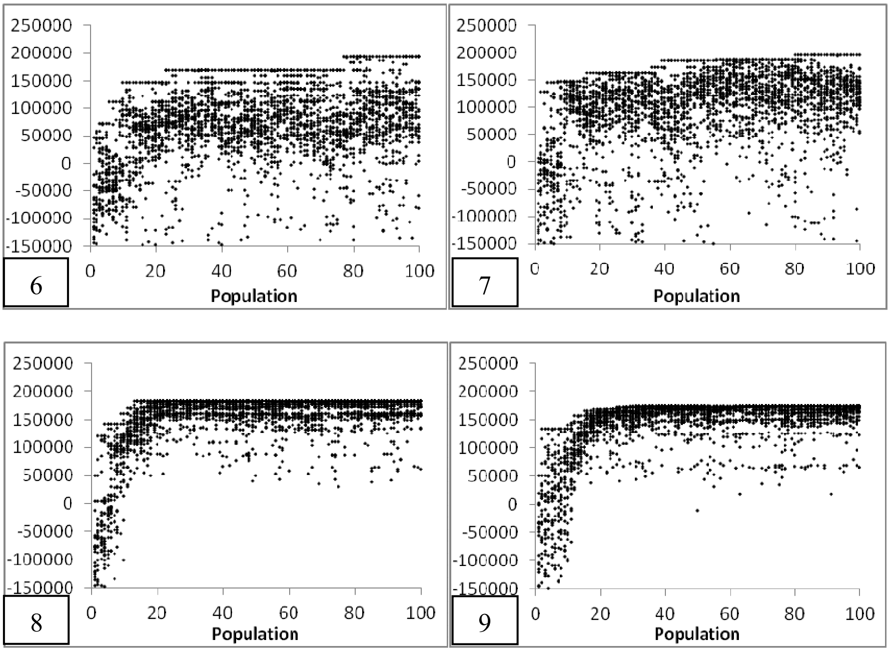


Fig. 6 – 9. CF for individuals in the one population for each combination [source: own study]

Although in the combination 2 with the multivariate crossing it can be seen that more individuals approach the maximum result of that particular repetition. In figures 8 and 9 a large concentration of CF values for each population can be seen. Most of the CF values are close to the maximum value. The concentration of results in a fairly narrow range already takes place in the first 20 generations, which may mean that chances of getting a better result for the next population is decreasing.

From the observation of all individuals in the population it can be concluded that proportional selection gives more diverse populations, and the tournament selection results in faster convergence of CF. Crossover operator does not have much influence on the result, one can assume that in this example a random factor has a strong influence on the maximum result obtained depending on the type of crossover.

5. SUMMARY

In the article, the optimization of typical manufacturing system using computer simulation was executed. For optimization genetic algorithm with different variants of genetic operators was used. Executed analysis shown that suitable buffer allocation and selection of resource capacity increase manufacturing system performance which gnerates higher profits. Various combinations of GA give different results. Proportional selection provides greater diversity in the given population. With tournament selection, the better value of whole population adaptation can be achieved. The crossover does not cause apparent effect, but it can be seen that the multivariate crossover operator will reduce the range in which there is the majority of CF values.

The analysis of the problem showed that it is worthwhile to conduct research towards the optimization of manufacturing systems by computer simulation and GAs. From the diversity of the obtained results one can observed how significantly a single parameter affects the profits received from the manufacturing system.

REFERENCES

[1] GOLDBERG D.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Boston, 1989.

[2] SCOTT L.R., HARMONOSKY C.M.: *An improved simulated annealing simulation optimization method for discrete parameter stochastic systems.* Computers & Operations Research, 32, 2005, p. 343–358.

[3] BANKS J., CARSON J.S., NELSON B.L., NICOL D.M.: *Discrete-Event System Simulation*. Prentice Hall Inc., New Jersey, 2001.

[4] PAUL R.J., CHANEV T.S.: *Simulation optimisation using a genetic algorithm.* Simulation Practice and Theory 6, 1998, p. 601–611.

[5] AZADEH A., TARVERDIAN S.: *Integration of genetic algorithm, computer simulation and design of experiments for forecasting electrical energy consumption*. Energy Policy, 35, 2007, p. 5229–5241.

[6] ZHANG H., GEN M.: *Multistage-based genetic algorithm for flexible job shop scheduling problem.* Complex Int., 11, 2005, p. 223–232.

[7] ENTRIKEN R., VÖSSNER S.: *Genetic Algorithms with cluster Analysis for Production Simulation.* Proceedings of the 1997 Winter Simulation Conference, p. 1307 – 1314.

[8] CAKAR T., YILDIRIM M.B.: *A neuro-genetic approach to designand planning of a manufacturing cell.* Journal of Intelligent Manufacturing, 16, 2005, p. 453–462.

[9] CAN B., BEHAM A., HEAVEY C.: *A Comparative study of Genetic Algorithm Components in Simulation-Based Optimisation*. Proceedings of the 2008 Winter Simulation Conference, p. 1829-1837.

[10] ESKANDARI H., RABELO L., MOLLAGHASEMI M.: *Multiobjective Simulation Optimization using an Enhanced Genetic Algorithm.* Proceedings of the 2005 Winter Simulation Conference, p. 833-841.

[11] KONSTAM, A.H., HARTLEY, S.J., CARR, W.L.: *Optimiaztion in a Distributed Proccesing Environment using Genetic Algorithm with Mutivariate Crossover*. Proceeding of the 1992 ACM annual conference on Communications, p. 109-116.

[12] HOLLAND J.K.: *Adaptation in Neural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI., 1975.

automated and robotic system, robot, simulation, teaching, virtual reality

Vladimír BALÁŽ, Marek VAGAŠ,
Mikuláš HAJDUK, Ján SEMJON*

MODULAR TRAINING SYSTEM IN VIRTUAL REALITY ENVIRONMENT – MSEVR

Abstract

MSEVR is a system created for teaching of automated and robotic systems by means of new advanced teaching aids, including virtual reality. One of its constituents is also virtual laboratory that has been created as a specialized website. The presented paper acquaints with structure and performance of the virtual laboratory created by a work team at KVTaR SjF TU at Košice. It is a laboratory that is segmented to a set of independent robots for learning basics of industrial robots and Modular system for experimentation in virtual reality. The Modular system has been designed for work with bigger, easily configurable workplaces.

1. INTRODUCTION

MSEVR is an education system designed for training at automated and robotized systems for employees of small and medium-sized enterprises as well as for professional trainers. One of its parts is a virtual laboratory, which is intended to teach in virtual reality environment. Created laboratory simulates real environment of workplace, where you can insert mechanical equipment and then makes experiments with it.

Basic building component for virtual laboratory is VRML language. By this language are created models that form 3D view of laboratory. Laboratory has a certain level of interaction. This feature is secured by using of VRML language sensors in cooperation with ECMA script (standard form language of JAVA script). This solution allows to retain ongoing events in virtual reality and respond to their processing in any way.

* Technical University of Kosice, Faculty of Mechanical Engineering, Department of Production Systems and Robotics, 042-00 Kosice, Němcovej 32, tel./fax.: (++421) 55 602 21 93,
e-mail: mikulas.hajduk@tuke.sk

Virtual laboratory can be run directly in VRML browser, or distributed through web site. In this case is VRML browser integrated as plug-in web browser. Created virtual laboratory has been developed and tested at configuration: Cortona3D Viewer v6 and MS Internet Explorer 8.

Virtual laboratory does not support a Mozilla Firefox of Google Chrome yet. At the beginning of the starting page is checked a configuration at client site to prevent an unresponsiveness of system. In case of inappropriate configuration of system follows next steps:

- If laboratory does not start at MS Internet Explorer browser follows a starting block.
- If client does not have a plug-in of VRML viewer Cortona3D from Parallel Graphics firm follows an instruction to download and half-automatic installation.

2. VIRTUAL LABORATORY

From the design point of view to virtual environment built through VRML language, which visually introduce a space of built workplace with empty feature of ground (see Fig. 1. basic view to environment system MSEVR). From the user design point of view is system before starting of experiment similar to status, when are finished building treatments of workplace and begins installation processes of mechanical equipments in real situations. In real situations is necessary at this moment proceed to buy, to import and to build all machines. In virtual environment we will follows similar steps, but without any costs and mainly much easier. Instead suppliers we have a gallery of machines, robots and equipments (see right part of Fig. 1). From it we can choose and move objects by drag & drop method to ground of workplace. Similarly we can edit properties of feature structure of ground workplace. For better orientation is possible to move textures, which have dimensions as an indicative lines from 1m x 1m to 0,1m x 0,1m.

Basic equipment MSEVR contains, (Fig. 2):

- models of robots,
- models of machines,
- models of transport equipments,
- models for trajectory optimization of robot,
- auxiliary models,
- textures for better orientation.

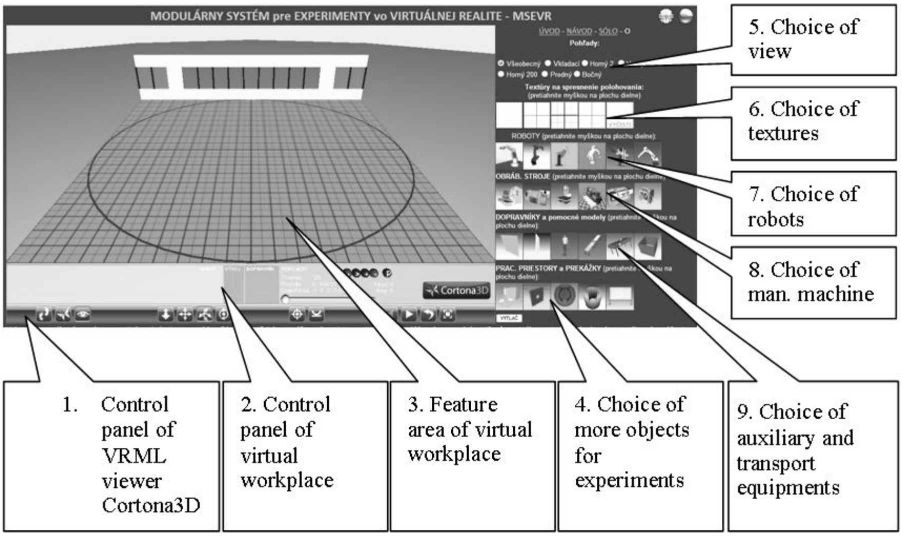


Fig. 1. Basic view to environment system MSEVR [source: own study]

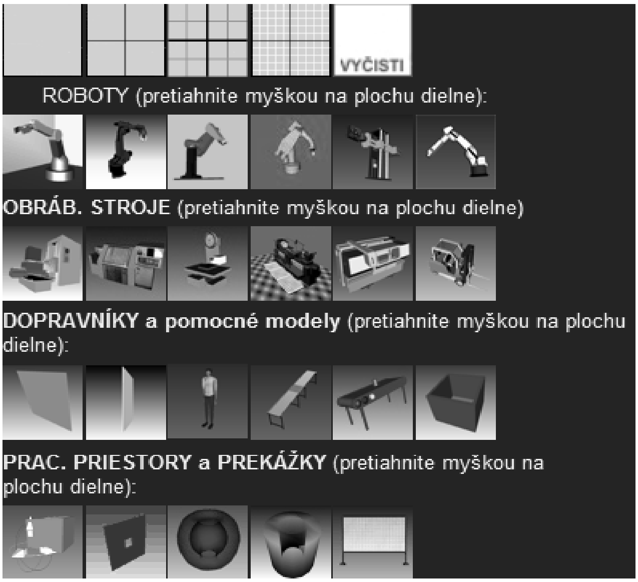


Fig. 2. Detail of MSEVR menu for select object at workplace [source: own study]

MSEVR allows to experiments in virtual reality quite simply through simply controlling of similar activities to real build of automated and robotized systems, see Fig. 3. So teacher can directly on lecture placed machines and equipments into feature ground of workplace with relative position and created automated and robotized systems from them. Students can saw not only results of its work, but also watch its progress, that is mainly increasing of effectively education process. For creating workplace is available pre-prepared models, but system allows using also external models. Due to creating of workplaces from individual models MSEVR will laboratory provides follow functions:

1. Visual
- representation at 3D virtual reality,
- freedom in movement at 3D space,
- change the type of motion at 3D space,
- set of view (cameras),
- to record trajectory of movement and than to start movement in automatic mode.

2. Preparatory
- change textures of feature ground area – guides lines for better orientation,
- insert (also remove) individual objects (machines, fixtures and other equipment),
- multiple insert,
- insert the foreign objects with minimal requirements to their adaptation.

3. Implementing
- movement with models,
- rotation with models,
- through scaling to change dimensions of inserting objects (only at necessary models),
- measurement of feed movement (axis X and Y) also rotation of objects, (around axis Z) for accurate determining their position,
- to retrieve all functions of own machines.

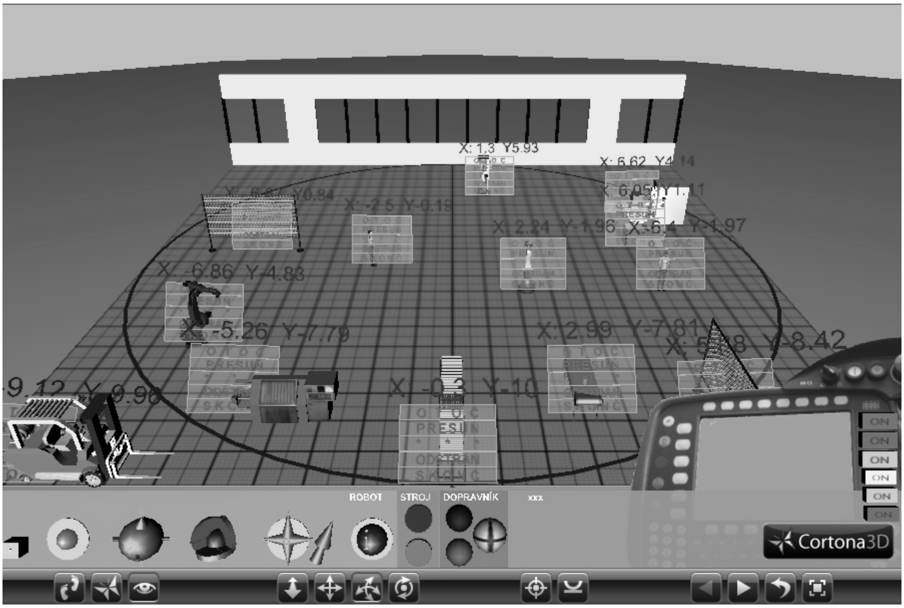


Fig. 3. Building phase of workplace in MSEVR system
(Mutual position setting of equipment) [source: own study]

Except for those functions is possible for created virtual workplace also more functionality because each model by inserting to the workplace brings not only its own shapes, but also all properties, which were assigned during its creating including interactive properties. Some standard models of workplace brings to us possibility for control, for example, also performance technology movements, controlling of movement speed, programming by teach method, sound and visual effects, possibility for extended edit and manipulation with model and so on. Complete list of functions from created workplace is depended on using models and their functions. Algorithm for building of virtual workplace in MSEVR environment can be seen at Fig. 4.

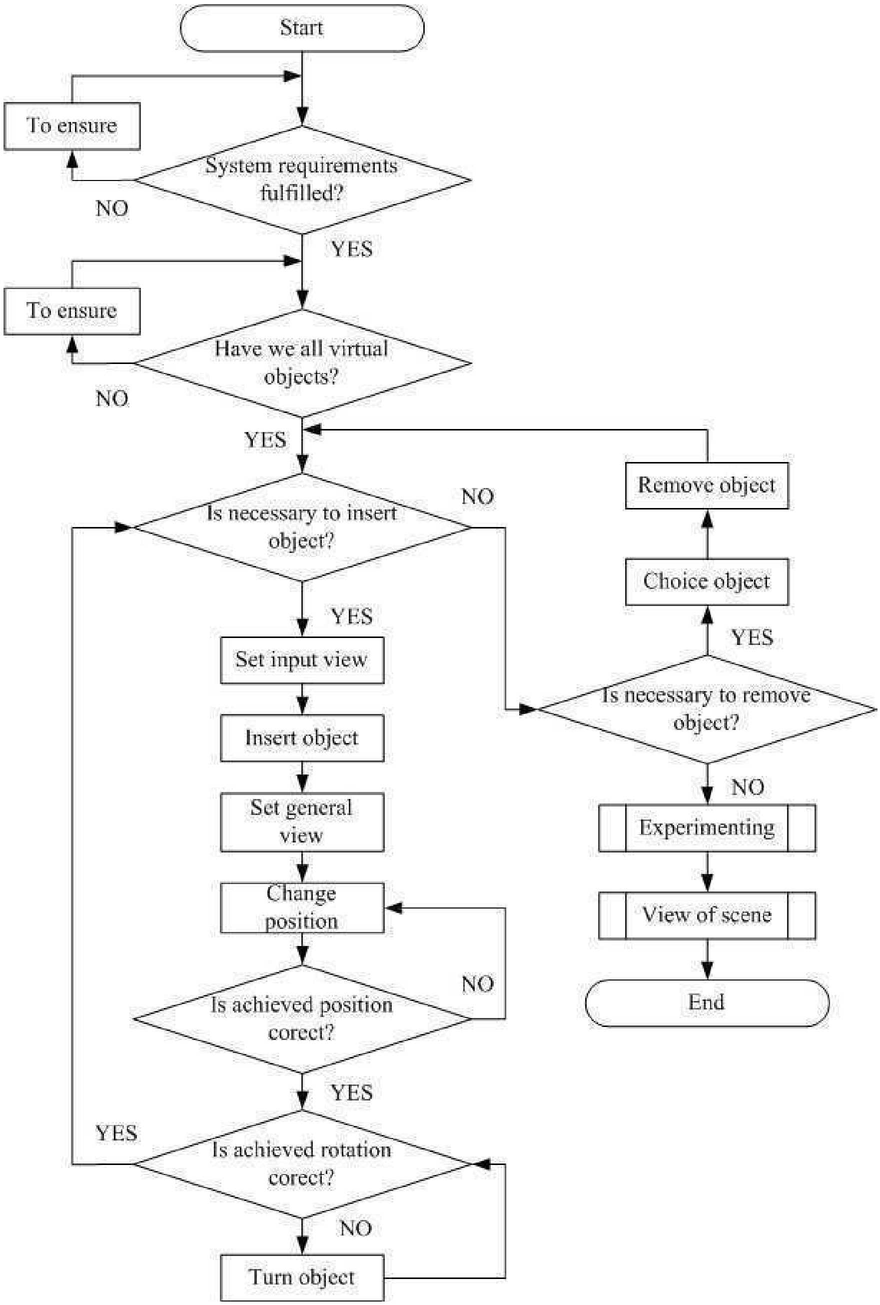


Fig. 4. Algorithm for building of virtual workplace in MSEVR environment [source: own study]

If we want to optimize the layout of machines is convenient to will be visualized some of them, because in real situation is not visible some spaces. Most of them are envelopes of machines, what is the reason how we can without any problems solve situation if are input / output places in its range. Visualized envelope is technically realized similarly as standard model, for example models of machines. Due to this we are able to their edit similarly as for standard models. Only different is their transparency set. It depends on the user, which invisible envelopes (in real situation) to display (in virtual reality) and for what purpose we use them. For us are preferred to display not only working envelopes, but also security zones and creating some fixtures to built virtual reality. MSEVR models have also dynamic characters. Working movement is animated, manually scheduled and controlled by program. All this cases were experimented. The use of dynamic properties allow to accurate optimalization at layout of workplace. Fig. 5 shows virtual workplace built during development of MSEVR system. MSEVR system allows addition of modules, which brings some new properties, so it is possible to their adapt at however implementation, if is virtual reality deployed. In its early it was tested to implementation in education process and optimized to automated and robotized systems. By most of verification experiments were confirmed our assumptions. In next we show our researched possibilities for MSEVR system:

- dispositional solution of workplace,
- use of animation on technological movements,
- use of visualization of security and technological envelopes,
- optimizing of robot trajectory for circumvent obstacles.

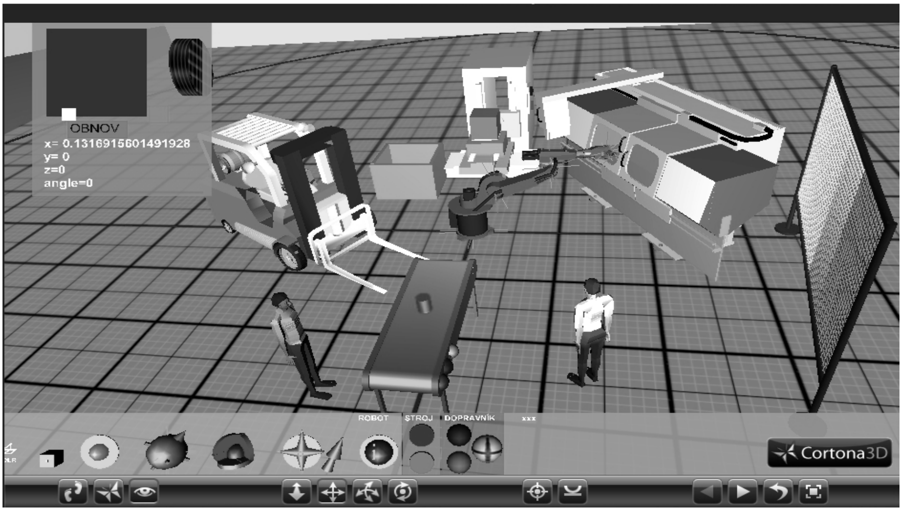


Fig. 5. Example of virtual workplace in MSEVR system [source: own study]

3. CONCLUSIONS

Virtual laboratory brings virtual reality into education process, what changed to better quality. Its interactive components allow to teacher and students are active part same as persons in real practice. This form of education process is in maximum volume similarly to real situation, and all costs for its realisation are negligible in compare of real reality. Due to the fulfilment of targets, which were set before its design, still exist the scope for its next improvement. Here is also possibility to insert free available models from different internet sources. Their disadvantage is necessary adaptation for inserting to MSEVR system for more effectively way as its own creating, because some firms that produce industrial robots (for example, ABB, KUKA, MOTOMAN) provides 3D models their products as support for designers without any conditions.

This contribution is the result of project implementation: KEGA 047TUKE-4/2011 E-learning robotiky s implementáciou virtuálneho laboratória s diaľkovým riadením reálnych zariadení na báze internetu.

REFERENCES

[1] BALÁŽ V., VAGAŠ M., SEMJON J.: *Využitie štruktúr informačných systémov pre pružné výrobné bunky*. Atp journal plus 1/2013, 2013, p.42-44.

[2] BALÁŽ V., BELOVEŽČÍK A., PÁCHNIKOVÁ L.:*Virtuálne laboratórium – Vitralab*. International Scientific Herald. Vol. 4, no. 23 (2012), 2012, p. 43-47.

[3] BELOVEŽČÍK A.: *Využitie virtualizačných prostriedkov pri projektovaní robotických pracovísk.* Písomná práca k dizertačnej skúške, Košice, SjF, 2010.

[4] CRAIG A. B.: *Understanding virtual reality, Morgan Kaufmann*. 2003, p. 448.

[5] HAJDUK M.*: Dva spôsoby spriahnutia vizuálny system – robot*. Zborník príspevkov z 11. medzinárodnej konferencie ROBTEP 2011, Automatizácia / Robotika v teórii a praxi, 2011, p.87-90.

[6] NIESZPOREK T., PIOTROWSKI A.: *Parametric Programming of CNC Machine Tools*. Applied Mechanics and Materials Vol.282, Robotics in Theory and Practice, 2013, p. 203-210.

[7] ŚWIĆ A., ZUBRZYCKI J., TARANENKO V.: *Modelling and Systematic Analysis of Models of Dynamic Systems of Shaft Machining*. Applied Mechanics and Materials Vol.282, Robotics in Theory and Practice, 2013, p.211-220.

[8] http://www.parallelgraphics.com/products

[9] http://wwwvrl.umich.edu

[10] http://www.web3d.org/x3d/specifications

packaging, machine, design, genetic algorithm, structure

Taras VARANITSKYI, Bogdan PALCHEVSKYI[*]

GENETIC-BASED APPROACH TO THE FUNCTIONAL-MODULAR STRUCTURE DESIGN OF PACKAGING MACHINES

Abstract

The paper presents a method for decomposition of packaging machine structure in order to provide the designer with choices for feasible assemblies. The aim is at providing a systematic approach to explore a large number of decompositions prior to the detailed component design phase. The structure is transformed to a graph with equivalent topology by a genetic algorithm.

1. INTRODUCTION

Most of the packaging devices, especially packaging machines with modular structure are made by collecting various components of the simpler structure than the final product (Fig. 1). The synthesis of such a structure is the choice of what components should be added together to get a result in the construction with the necessary parameters. The method presented in this paper aims to systematize the process of structural synthesis of packaging machine with a main focus on its versatility, taking also into account productivity, cost, reliability, weight and overall performance. This approach aims to provide the designer feedback on possible designs of packaging machine before detailed design stage.

[*] Lutsk National Technical University, 75 Lvivska str., 43018 Lutsk, tel.: +380 (332) 767395, e-mail: varanitskyi@i.ua, pac@tf.dtu.lutsk.ua

In this method, the design of packaging machine is obtained by optimization structure synthesis of a set of heterogeneous functional modules (FMs). This process includes three stages:

1. Creating a topological graph of the packing machine structure based on a set of elementary technological operations (ETO) and their corresponding functional modules;
2. Building the two-dimensional matrices for genetic coding of packing machine structure on the basis of topological graph;
3. Optimizing the formed matrix using a genetic algorithm with regard to the above criteria and consistency of the packing process.

The issue of synthesis of topological graph of the packing machine structure is considered as a discrete optimization problem that is solved by means of genetic algorithm.

2. CREATING A TOPOLOGICAL GRAPH STRUCTURE PACKING MACHINE

Fig. 2 shows the course of forming the packing machine structure matrix. At the beginning the distinct parts of the packaging process are distinguished – the elementary technological operations (Fig. 2a). Each ETO is an action, which results in a new element of package to be formed or changing the shape, position or properties of the existing one.

Next, each ETO should be put in correspondence with a specific functional module (Fig. 2b). In view of technology or possible design of FM a separate ETO is not always performed apart so it is necessary to combine ETOs into sets. Each set describes a certain type of FM. For example, a set of ETO "to weld a longitudinal seam" and "move film sleeve" match to design of the roller mechanisms for longitudinal welding. Thus two types of functional modules that implement ETO sets are different, if these sets have at least one different ETO.

Obviously, the set may contain only one ETO. Sets of ETO significantly expand the number of FM designs that are used during the synthesis of structure packing machine.

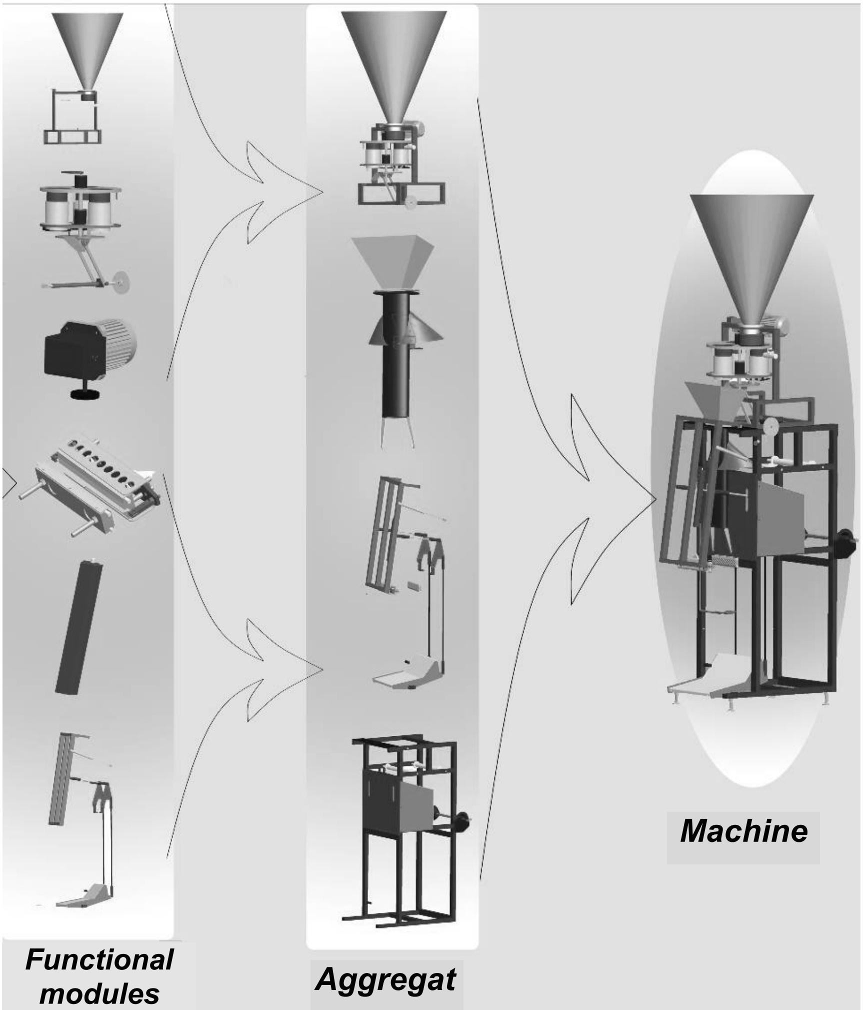


Fig. 1. Scheme of the packing machine decomposition [source: own study]

Each vertex in the graph is given a number. The numbering is carried out in accordance to a set of ETO, which the appropriate FM implements: first in the sequence of operations of a generalized process and then two digits – a combination of numbers of single ETO. Thus the obtained set of vertices of the structure topological graph includes the FM, presented by two types of sets: single and paired ETO.

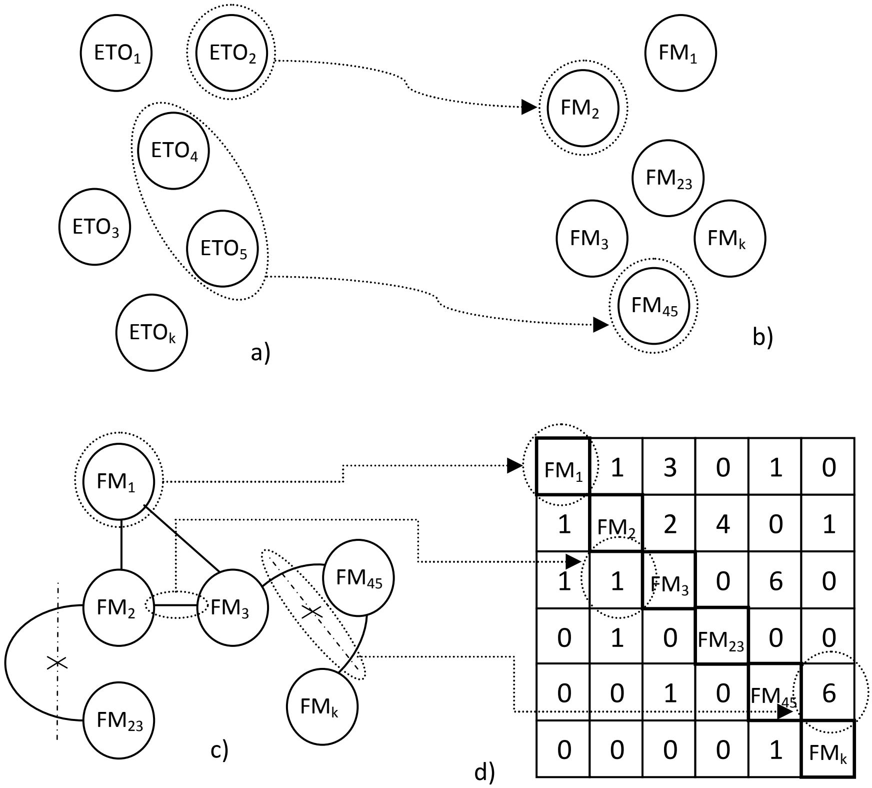


Fig. 2. Course of forming the packing machine structure matrix [source: own study]

The next step is ordering the vertices: placing them in certain sequence and establishing connections between them (Fig. 2c). This process tends to generate a graph that describes the structure of the packaging machine and is performed in the following way:

1. A vertex that has the smallest number, usually 1, is selected from the set.
2. The value of degree is generated for each vertex. It is a natural number, limited to a certain value.
3. Steps 1 and 2 are repeated until the k vertices are collected. The maximum number of vertices in the graph k is limited by the number of structural elements in a family of products.

The generated graph should be checked on several grounds. The main purpose of this test is to set those options that represent actual allowable structures of packing machine. Their separation will greatly accelerate the optimization process, reducing costs to process options that go beyond the space of feasible solutions.

Graph must be connected. Obviously, the machine cannot consist of a set of functional modules that are not linked, even if each of them separately can provide the relevant operations of package forming. Therefore a disconnected graph does not depict acceptable structure at all. In validation process all graphs the degree of at least one vertex is zero are discarded immediately, while others are checked by any known method (e.g. Kruskal's algorithm) for connectivity.

Graph must be simple, i.e. it should not contain loops or multiple bonds. The loop in this case means that the manufacture of the product may be stuck in one position in the machine. Multiple bond between the vertices of the graph leads to the loss of uniqueness in determining relative positions of FM in the machine. This problem is solved by creating a structure matrix: its completion is organized in a way that eliminates the appearance of loops and multiple bonds.

The FM describing the graph should implement all the ETO necessary for the formation of all structural elements of a product family in a given sequence. The longest path in the graph must include all ETO necessary for making at least one type of product. Otherwise the basic aim of the machine – producing the necessary range of products – is not provided.

As a result the topological graph of packing machine structure shows all its working positions and options the product routes in the machine.

3. BUILDING A MATRIX OF PACKING MACHINE STRUCTURE

A matrix of packing machine structure is obtained by sequential cellpadding of the two-dimensional array. Each cell of the matrix is associated with the corresponding vertex of topological structure graph. The matrix size is $k \times k$ elements. The adjacency matrix by which undirected graphs are usually describe is taken as a basis. However, a way its completion is modified the gain an opportunity to describe not only the interconnections between the vertices of the graph, but also the characteristics of arcs and the vertices themselves. A matrix of packing machine structure (Fig. 2d) can be divided into three parts, each comprising an encoded description of certain aspects of the structure of the packaging machine.

The first part is represented by cells that are below and to the left of the main diagonal of the matrix. The content of cells is based on the topological graph of structure and describes the links between the FMs in the machine. Filling of this part is performed likewise cellpadding of classical adjacency matrix and fully describes the process of generating the graph. Since the topological graph of packing machine structure is a simple graph, the cells can only take values 0 or 1.

The second part of the matrix structure – is the cells that are above and to the right of the main diagonal. As usual adjacency matrix for an undirected graph is symmetric, it was decided to use these cells for expanding the characteristics of the graph. They show the actual links between FMs in a machine. Each cell in this part describes mutual position of a pair of FMs to which it relates. The value of the cell is a number that describes a one-way communication between the modules in the sequence of their connection. There may be six types of such links:

- Placement along the vertical axis
- Placement along the horizontal axis
- Placement along the inclined axis
- Placement in a circle with a vertical axis of rotation
- Placement in a circle with a horizontal axis of rotation
- Placement in a circle with oblique rotation axis

Therefore, each cell gains value from 1 to 6. So 6^x connection methods between the neighbouring FMs can be specified.

The third part of the matrix – its main diagonal – describes the FM types included in the structure of the machine. Cells contain information about the design of FM and thus set of ETO it performs. The value of a cell consists of two digits. They point to the ETO implemented by a particular module. In case a set consists of one ETO, a zero is written in place of the second one. Filling of cells of the main diagonal is conducted according to the order of the process of packing. That is, cell a_{11} describes the FM, a set of ETO for which necessarily contains the first operation. Further cells respectively contain the entire set of ETO required for the formation of all structural elements from a given family of packages.

Thus, each cell of the main diagonal of the matrix of structure meets specific FM and its corresponding diagonal cells – additionally describe the location and characteristics of FM.

4. MATHEMATICAL MODEL

Optimization synthesis of structure can be considered as a problem of creation of the graph. In such case the elements of the machine structure correspond to vertices, and the links between them – to edges that connect two or more vertices. Therefore, the whole structure can be represented as a graph $G = (V, E)$, where V – the set of vertices, E – set of edges. The problem of synthesis optimization then can be described as finding the optimal set P of nodes V, such that the objective function $f(P)$ is maximized.

It should be noted that the number of vertices k, which form the best option is not defined in general. However, a set of characteristics F for each FM should be specified in order to get the opportunity to fully evaluate the various options of structure.

Thus, the optimization problem can be put as:

Given G a topological graph of structure, F a set of characteristics for each FM, and k the number of vertices. Find a vector x, which determines the optimal set of FM, the vector v – way to link FMs and vector y – description of each FM, subject to $p(x,v,y)$ – the objective function that defines the quality of packing machine structure.

The objective function assesses the structure of the following criteria:

- Versatility – increased by introducing a FM, carrying out new functions into the structure of packing machine and depends on the combination method thereof;
- Productivity – rises by increasing the number of FMs;
- Reliability – reduced when FM, carrying out new features introduced into the structure of packing machine and depends on the combination method thereof.

To assess a packaging machine according to the criterion of universality, the total number of types and sizes of products, which the machine can produce is calculated primarily, and on this basis the coefficient of universality is determine. This value is determined by the structure of the machine and the set structural elements of products that specific FM produces. In general, the ratio of universality is the product of the coefficients of universality for serial and parallel connection of FMs. Thus arbitrarily complex TM is reduced to a chain of series-connected parts. Several FM, which operate in parallel or sequentially on the same level are considered as one group.

To evaluate the productivity of packaging machine the theoretical quantity of product it produces during a specified time is calculated. Raising the value of this parameter can be achieved by increasing the number of FM as the slowest ETO, as well as throughout the whole packing machine.

Reliability of a packaging machine is estimated by the availability function. This value is calculated based on the known factors of availability for each FM and the method of connection thereof. Thus in case of serial connection, the reliability of a machine declines. This shortcoming can be partially removed by reserving FM on operations where there the failures are most often.

The following objective function is a result of combining the criteria:

$$f(x,v,y) = w_1 \cdot (1-(\prod_{j=1}^{\alpha}(1-K_{Yi}) \cdot \prod_{j=1}^{\beta}\frac{1}{\sum_{i=1}^{q}(\frac{1}{1-K_{Yi}})})) + w_2 \cdot (\frac{1}{T_{\max}}) + w_3 \cdot (\frac{1}{1+\sum_{i=1}^{k}(\frac{1}{K_{\Gamma i}}-1)}) \quad (1)$$

where: $x=(x_i)$ – x_i is a variable that refers to the presence of the edge e_i in the x,
$v=(v_i)$ – v_i is a variable that represents the weight coefficient of the edge e_i in the set x,
$y=(y_i)$ – y_i is a variable that refers to the characteristics of the FM for the correspondent vertex,
w_i – weights of partial criteria,
α and β – number of series-connected FM and parallel connected sets of FM respectively,
K_{Yi} – actor universality of the i-th FM,
q – number of parallel-connected FM,
T_{max} – the duration of the longest ETO,
$K_{\Gamma i}$ – availability factor of i-th FM,
k – number of FMs in the packaging machine.

5. OPTIMIZATION METHOD WITH USE OF GENETIC ALGORITHM

Synthesis of a graph is NP-complete problem even using simple linear relationships for calculating the value of criteria. As a result, all known algorithms that exactly solve the problem of graph synthesis, do it for the time that depends exponentially on the size of the graph. In our case objective function is nonlinear and since we cannot afford exponential computation, heuristic algorithms are found to be suitable. More specifically, a steady-state genetic algorithm (SSGA) has been used to solve the problem approximately, which goes through the following stages:

1. Randomly create a population P of n chromosomes (encoded representation of the parameters x, v and y), estimate the value of the objective function and keep the best chromosome. Create an empty subpopulation Q;
2. Select two chromosomes c_i and c_j in P with a probability:

$$\Pr(c_i) = \frac{f_i}{\sum_{k=0}^{n} f_k} \tag{2}$$

where: f_i – value of the objective function for chromosome c_i;

3. Crossover the chromosomes c_i and c_j and generate new chromosomes c'_i and c'_j;
4. Mutate c'_i and c'_j with a certain low probability;
5. Evaluate the objective function for c'_i and c'_j and add them to Q. If Q contains less than m chromosomes move to step 2;
6. Replace m chromosomes in P with the ones in Q and empty Q. Update the best chromosome and increment the generation counter. If the generation counter has reached a pre-specified number, terminate the process and return the best chromosome. Otherwise go to 2.

The main advantage of SSGA is empirically set prevention of premature convergence and optimal results obtained in fewer computations.

Each version of the packaging machine is encoded in a matrix structure, which is also its chromosome. Values of genes are identical and describe the same characteristics as the corresponding matrix cell.

It should be noted that the information is the first part of the chromosome refers to the vector *x*, the second – to the vector *v*, and – the third to vector *y* in a mathematical model. Also is necessary to stress that not every chromosome describes the "true" structure of packing machine. Eliminating some edges may not lead to destruction of the structure, so each option should be checked whether it meets the unconnected graph. Obviously, the ideal case is all the k-connected vertices.

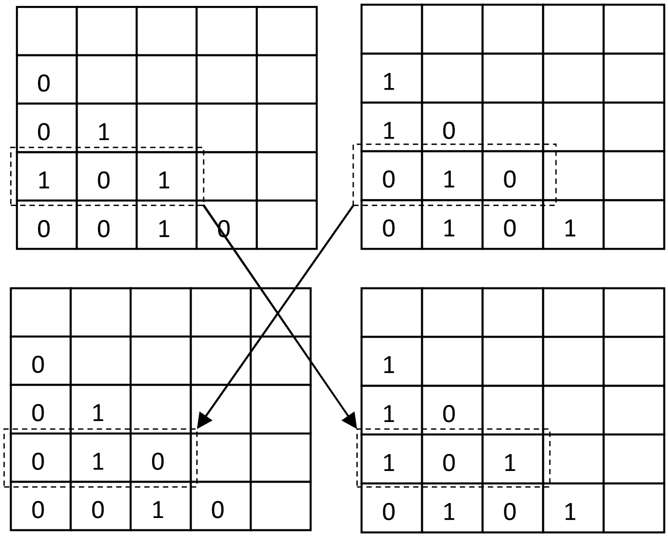


Fig. 3. Cross-over of two chromosomes [source: own study]

Since chromosomes, describing the structure of packaging machine contain three types of information there is a need to configure crossing-over and mutation operators. Cross-over operator influences on each part of chromosome separately. It is performed in two steps where the first third mates with the first third, and third – with the third part of the chromosome. Such crossing takes place in lines (Fig. 3), the line number being also generated randomly. Crossover is not performed for the main diagonal. Thus when deploying a matrix structure in line, we have actually multipoint crossover.

The mutation operator differs from the typical one by the higher probability of the second part of the chromosome.

6. EXAMPLE OF PRACTICAL IMPLEMENTATION

For example, consider the process synthesis and optimization of the structure of a packaging machine for loose products. By the condition the machine must produce a family of three types of packages: three-seam "cushion"-type package, package with a flat bottom and a bent top package. ETO set for this family of packages will be as follows:

1. Bend the film into sleeve
2. Weld the lower cross seam
3. Weld the top cross seam
4. Weld the longitudinal seam
5. Bend the bottom of the package
6. Fold the top of package
7. Cut the finished package
8. Pull a film sleeve on a step

15 sets of ETO and the same number of FM structures can be put in correspondence with the above set, namely:

- single FM1, ... FM8
- double FM23, FM25, FM27, FM36, FM37, FM48, FM56

The first generation of the genetic algorithm generates the following matrix of packing machine structure size 8x8:

01	0	2	0	0	0	0	0
0	25	0	1	0	3	0	0
1	0	06	0	0	5	0	0
0	1	0	36	2	0	0	0
0	0	0	1	56	0	2	4
0	1	1	0	0	04	1	0
0	0	0	0	1	1	08	0
0	0	0	0	1	0	0	37

As can be seen from the genetic code, this chromosome does not represent an acceptable construction of packaging machine because the necessary sequence of package producing process is not provided. The structure graph of the FM set has the form shown in Fig. 4. This also will be the rough location of FMs in the machine.

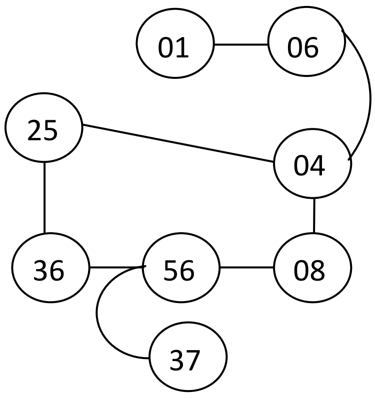


Fig. 4. The structure topological graph, formed at the first generation [source: own study]

During the genetic algorithm process each part of the chromosome undergoes the following processes:

- as a result of mutations in the main diagonal the required set of ETO is determined and composition of FMs in the machine is optimized. This may also change the universality – the presence of vertices with identical ETO in sets can create the alternative ways of passing the product through the machine without its readjustment and form routes that lead to the creation of a new product type;

- genetic changes in the first part of the chromosome lead to the establishment of the required order of ETO. Basically, this process occurs due to crossing-over, mutation only accelerates finding the final result;
- the third part of the chromosome is responsible for setting such geometric relationships between FM, in which the machine design is not too cumbersome and provides a logical and efficient product passing through all positions.

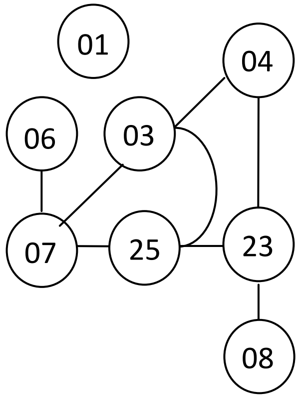
Fig. 5 shows two couples of "matrix structure – topological graph", illustrating the possible intermediate stages of optimization synthesis of packing machine structure.

The first pair (Fig. 5a) represents already established set of ETO. Links between modules do not describe the acceptable design of packaging machine yet. The lack of connections between some FM is evident so this structure still needs refinement so that permissible design of packaging machine could be put in correspondence to it.

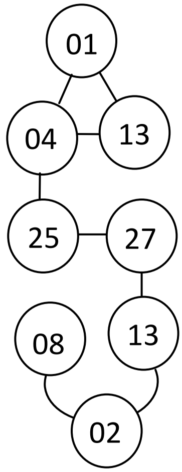
The second pair (Fig. 5b) shows the case where the manner and types of connections between vertices of the graph can represent realistically possible machine design. However, the main diagonal of the matrix is filled chaotically: an unreasonable duplication of some ETO and the absence of others occurs. As a result of mutations, the set of encoded ETO should be reduced to the "rules" to achieve the acceptable solution of the problem.

a)

01	0	0	0	0	0	0	0
0	04	0	1	0	0	0	3
0	0	08	1	0	0	0	0
0	1	1	23	2	0	0	0
0	0	0	1	25	0	2	4
0	0	0	0	0	06	1	0
0	0	0	0	1	1	07	3
0	1	0	0	1	0	1	03



01	3	3	0	0	0	0	0
1	04	2	1	0	0	0	0
1	1	13	0	0	0	0	0
0	1	0	25	2	0	0	0
0	0	0	1	27	1	2	0
0	0	0	0	1	13	4	0
0	0	0	0	0	1	02	4
0	0	0	0	0	0	1	08



b)

Fig. 5. Interim results of packing machine structure optimization synthesis using GA [source: own study]

Comparison of the interim results for multiple chromosomes on intermediate generation leads to identify trends that demonstrate how a matrix of packing machine structure mutates when approaching the optimal solution. The first trend is the location of FM numbers in ascending order along the main diagonal of the matrix. The second trend is reflected in the "clustering" of nonzero values of cells on both sides of the main diagonal.

Of course, in both cases, there may be some deviation, but the preservation of classical logic and sequence of ETO for packing loose products finds confirmation in the application of genetic algorithm.

Finally, Fig. 6 shows a chromosome of a possible structure of packing machine, and its topological structure graph.

01	3	0	0	0	0	0	0
1	04	1	0	0	0	0	0
0	1	08	3	3	0	0	0
0	0	1	25	2	3	0	0
0	0	1	1	23	3	2	0
0	0	0	1	1	07	2	3
0	0	0	0	0	1	06	1
0	0	0	0	0	1	1	03

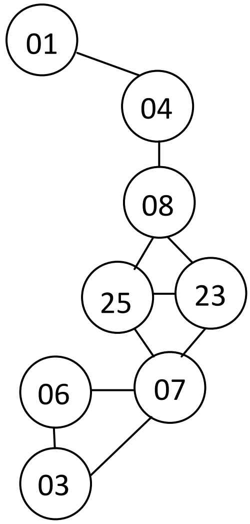


Fig. 6. Feasible final structure of packing machine obtained with GA [source: own study]

The process of packing in this case will take place in the following sequence:

- film rolled up in the sleeve moves at an angle to the longitudinal seaming mechanism,
- seamed sleeve moves to pulling mechanism that is placed vertically under the longitudinal seaming mechanism,
- before filling the blank with the product one of two mechanisms of cross seaming triggers: mesh sponge, forming both the upper seam of next package and the bottom of the previous one or the mechanism which bends the bottom of the package and welds only the bottom seam,
- next the package moves vertically downward for cutting,
- when secondary mechanism of cross seaming triggered, the filled and cut package passed horizontally to the mechanism of folding the top of package, and then – straight down – the upper cross seam is formed,
- if the folding the top of package is not necessary – the blank is immediately transmitted along an inclined trajectory to the transverse seaming mechanism.

REFERENCES

[1] F. ARKIN YETIS, KAZUHIRO SAITOU: *Decomposition-Based Assembly Synthesis Based On Structural Considerations.* Proceedings of DETC’00 ASME 2000 Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2000.

[2] PALCHEVSKYI B., VARANITSKYI T.: *Analisys Of Variability Of The Family Of Packages And The Versatile Packaging Machine Design.* Applied Computer Science Vol. 8, №2, 2012, p.22-33.

[3] PALCHEVSKYI B.: *Avtomatyzacja technologichnych protsesiv (vygotovlenia i pakuvannia vyrobiv)* (in Ukrainian). Svit, Lviv, 2007.

ear-clipping, figure recognition, calculation of missing dimensions, undirected consistent and cyclical graph, 3D visualization

Aneta BERA[*]

AUTOMATION OF ROOF AREA CALCULATION PROCESS

Abstract

The paper presents approach that allows to automate roof area calculation process as well as other accompanying actions. These actions include: figure recognition, calculation of necessary dimensions, recognition of connections between roof surfaces, visualization of the roof in three-dimensional space and calculation of the roof area based on processed data. The aim of proposed approach is to automate as much actions as it is possible to reduce calculation errors and moreover increase the efficiency of whole process. Presented idea has been implemented and then tested for accuracy of calculations and capabilities. Performed tests showed that presented method is flexible and gives relatively small error in case of roof area calculation.

1. INTRODUCTION AND MOTIVATION

Nowadays many roofing companies employ person who is responsible for pricing roof plans. This pricing is based on architectural project, that is delivered by customer, which shows projection of the roof. To calculate the number of needed tiles it is necessary to calculate roof area. This process is preceded by a series of other steps – all of them done without using the computer. That kind of approach takes much time and is not efficient enough. Using the computer to automate this process would result in considerable economic and time benefits. It would also eliminate human errors and increase accuracy of calculations. Author has analyzed whole process in one of the companies in Szczecin and based on gathered observations has created new solution which is presented in following part of this paper.

[*] West Pomeranian University of Technology, Szczecin, Faculty of Computer Science and Information Technology, 71-210 Szczecin, Żołnierska 49, tel./fax.: 607339458, e-mail: abera@wi.zut.edu.pl

2. COMPUTATIONAL ALGORITHMS

Proposed approach to solve the problem of calculating roof area consists of ten steps. Those steps are shown in the figure 1 and discussed in further part of this paper. Solution assumes that basic data about the architectural design of the roof, like the shape of the roof surface or slope angle, and at least one length of the line segments, are entered by the human. All further steps could be automated and executed by the computer without user supervision.

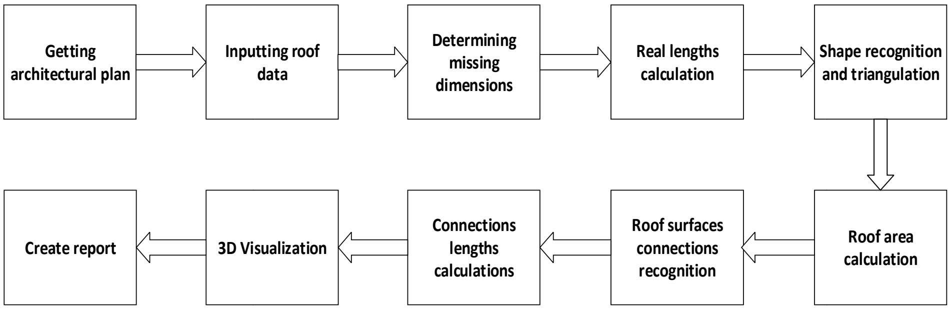


Fig. 1. Steps in proposed solution [source: own study]

2.1. Data collecting

The main part of this step is to get an architectural plan and create the model of the roof. It is created by user by drawing roof surfaces. This model should be useful for further analyzing by the computer. Necessary information that is required in further parts of presented idea are: slope angle as well as base line for each roof surface and at least one length of line segment.

Base line in figure is a line segment which is situated the lowest and is real length – not the slope line.

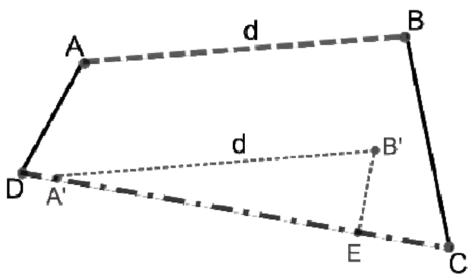


Fig. 2. Real line segment's lengths calculation [source: own study]

2.2. Getting line segments data

To calculate the roof area it is requisite to know real measurement of each line segment – not its projection. To accomplish this task first it is needed to find all missing measurements on the architectural plan and then calculate their lengths in three-dimensional space.

To calculate missing lengths it is necessary to know at least one measurement and then designate the scale for this one specific project. With this information it is possible to move measurements from local coordinate system to real one. When all dimensions are known, the real lengths of line segments in 3D can be calculated. To do that slope angle (α) and base line for each figure needs to be known. Figure 2 shows sample figure in which *d* is the line that length is needed to be known. First of all points A and B are projected to base line (|CD|), and then length d can be calculated from Pythagorean theorem [1] as:

$$d = \sqrt{|A'E|^2 + \left(\frac{|B'E|}{\cos\alpha}\right)^2} \tag{1}$$

2.3. Figures recognition

Next step, that needs to be done, is recognition of the shapes of each roof surface. It can be done with information that has been already collected – vertices positions. With that knowledge it is possible to recognize shape of the surface based on simple rules. Sample rules can be written as:

IF (number_of_vertices == 3) THEN (shape='triangle'),
IF (number_of_vertices == 4) AND (sides_are_parallel_in_pairs) AND (angle_between_two_adjacent_walls != pi/2) THEN (shape='parallelogram').

Similarly two other rules for rectangle and trapeze could be written.

Figures which shapes could not be recognized by presented rules (for example pentagon or deltoid) need to be triangulated – divided into n-2 triangles, where n is number of polygon vertices. Method that can be easily applied, in case when coordinates of vertices in polygon are already given, is triangulation by ear-clipping. This algorithm is also useful when studied figure is concave. The main idea in this method is to identify ears of polygon and their subsequent removal until remaining part of polygon is not a triangle. Ear of the polygon is a triangle formed from three consecutive vertices where two outer vertices form a diagonal. Necessary condition for created triangle to be called ear is that all diagonal must be inside the polygon [2,3]. To find it out odd parity rule can be applied, which says when the point is placed inside or outside the polygon. When ray drawn from processed point in any direction intersects the polygon edges odd number of times then the point is placed inside the polygon, when even outside [4].

2.4. Roof area calculation

After figure shapes recognition or possible triangulation and calculating real lengths of lines segments, roof area can be calculated. Those areas are calculated using basic formulas used in math. One exception is the case of triangle – it is better to use Heron's formula rather than basic formula [1]. Heron's formula does not require additional calculations such as height of the triangle – those calculations would be the reason of greater error.

2.5. Recognition of connections between roof surfaces based on graphs

To be able to calculate the total lengths of different types of connections between roof surfaces it is necessary to recognize their types. It is possible with undirected, consistent and cyclical graph. In this case, nodes in the graph are figure's vertices and the edges are line segments. Each node has an information about its neighbors. Created graph will be traversed using modified version of breadth first search algorithm [5].

Changes in the algorithm can be presented as:

- start with queue filled with all vertices belonging to the edge of the roof, instead of one vertex,
- due to relatively small number of vertices, it is suggested to use an array of size n, where n is a number of vertices, instead of the queue,
- during graph traversal, instead of marking vertex as visited, number describing how far from the roof edge is placed is assigned to him.

Described algorithm presented as flowchart is shown on figure 3. Sample graph that was created using presented algorithm is shown on figure 4.

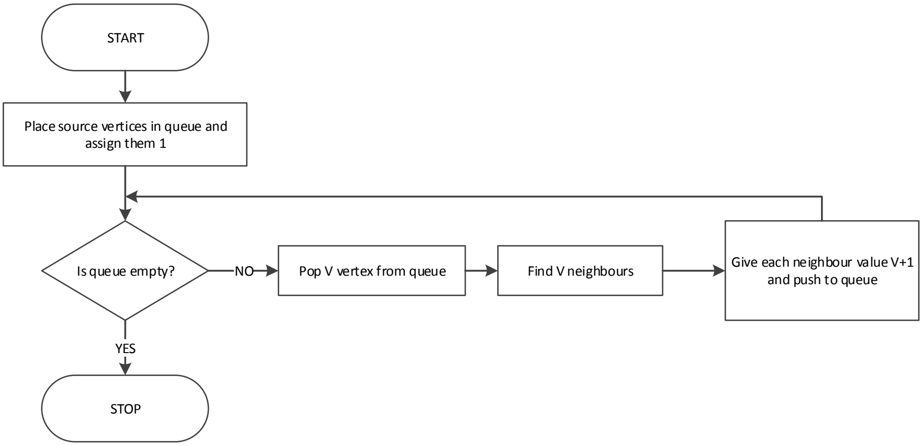


Fig. 3. Modified BFS algorithm [source: own study]

After the types are assigned to all vertices, it is possible to begin recognition of the connections between surfaces. Types that could be recognized are: valley, ridge/hip or roof edge. Beside the information about type of the vertices it is also needed to know slope angle.

There are two possible scenarios:

- line segment consists of two vertices of the same type,
- line segment consists of two vertices of different types.

In first scenario, type of connections could be determined using information about slope angles of the figures that contain examined line segment. In second, it is necessary to create three vectors and then calculate angles between them. Vectors that need to be formed are:

- vector that describes processed line segment,
- vector that is between the lowest point of processed line segment and its neighbor situated on the same level – the same type,
- vector that is between the lowest point of processed line segment and its other neighbor, that is the same type.

Example vectors that need to be created are presented in figure 5.

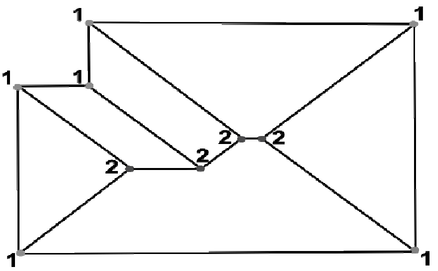


Fig. 4. Sample graph with vertices types [source: own study]

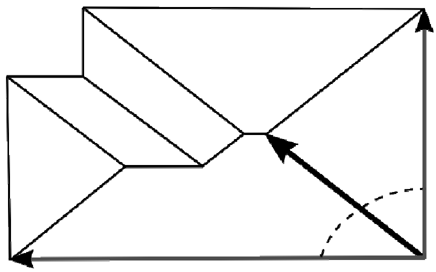


Fig. 5. Vectors creation [source: own study]

After creating said vectors, angles between examined line segment and other two created vectors are summed. The result of that operation determines type of connection. The algorithm that is used to determine type of connection between roof surfaces and edges is described using flowchart in figure 6.

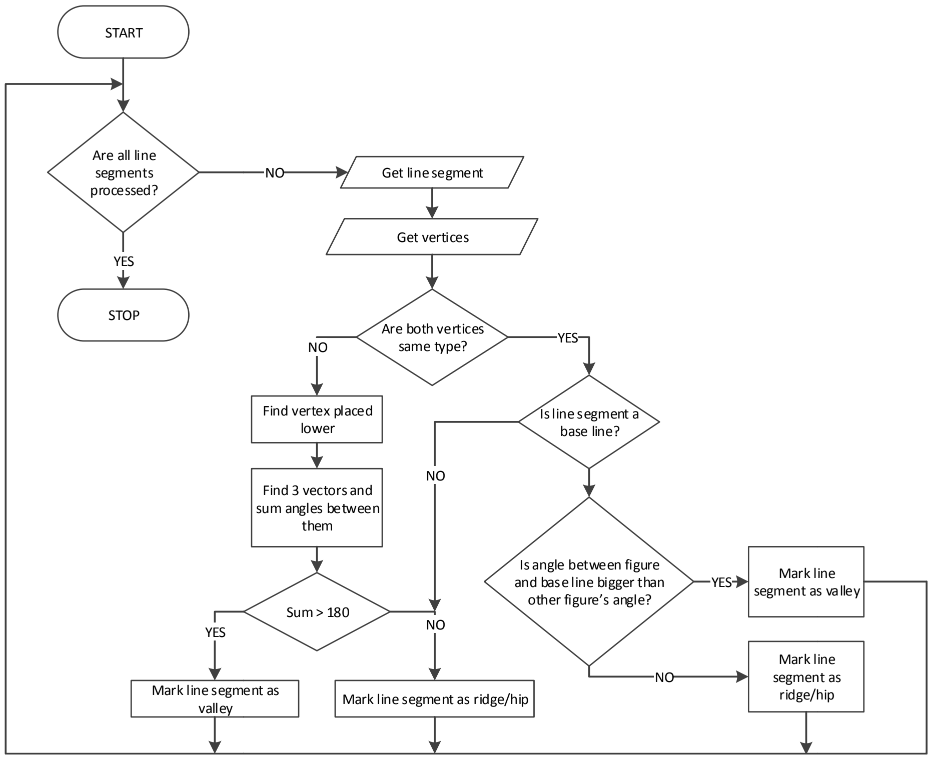


Fig. 6. Connection type recognition solution [source: own study]

2.6. Visualization of roof shape in three-dimensional space

Before roof surfaces shapes can be visualized in three-dimensional space, it is necessary to use graph, that was created already, again. When the graph was being created, each vertex had value 0 assigned to its height. Traversing the graph again, we can determine if height of each vertex is 0, or if it needs to be calculated. Algorithm that describes process of creating 3D model of the roof is shown in the figure 7.

The method of calculating height of the nodes that do not belong to base line is based on projecting processed vertex on the base line, calculate distance *d* between processed vertex and its projection. Then using trigonometric function [1] height can be calculated as:

$$h = \tan\alpha \cdot d \tag{2}$$

where: α – angle slope to which processed point belongs.

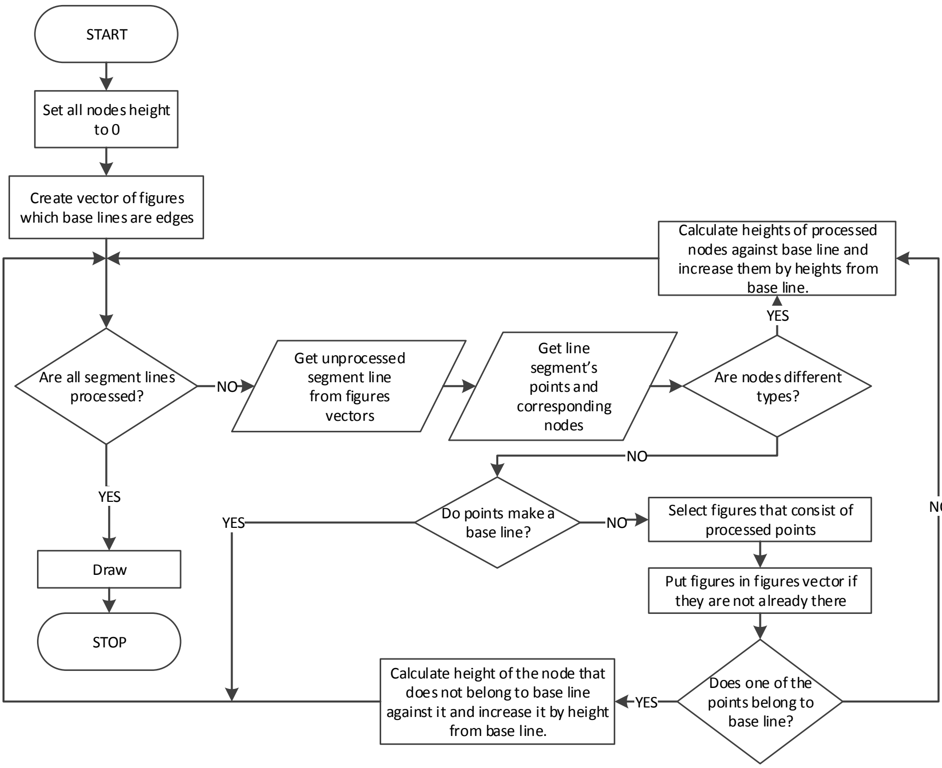


Fig. 7. 3D model creation [source: own study]

3. TESTING PROCEDURE AND RESULTS

Proposed solution, that solves the problem, has been implemented in Matlab [6] and tested for:

- errors in roof area calculation,
- correctness of connections between roof surfaces recognition,
- correctness of 3D visualization.

Tests were conducted for three sample architectural projects: hip roof, gable roof with outhouse and roof with six surfaces. Projections of these roofs are shown on figure 8. To determine the accuracy while calculating roof area relative error concept have been used:

$$\delta x = \frac{|x - x_0|}{|x_0|} \cdot 100\% \quad [7] \qquad (3)$$

Main purpose was to determine how number of lengths, that are known at the beginning (does not have to be determined by computer), affects the accuracy of calculated area. Graph in figure 9 shows relative mean error for all tested projects against number of specified lengths. As can be seen, with increasing the number of lengths, relative mean error is decreasing. It is related to the fact that the less known data means that more needs to be approximated what could cause errors.

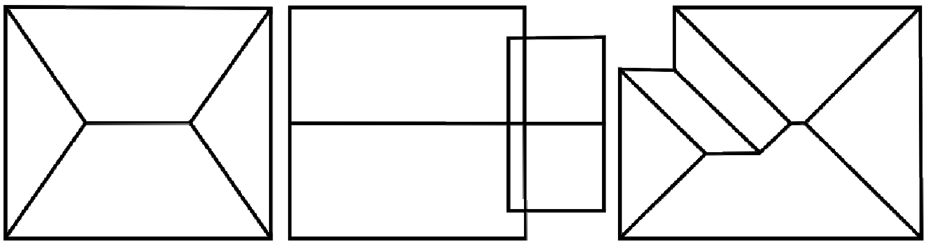


Fig. 8. Architectural plans: hip roof, gable roof with outhouse, roof with six surfaces [source: own study]

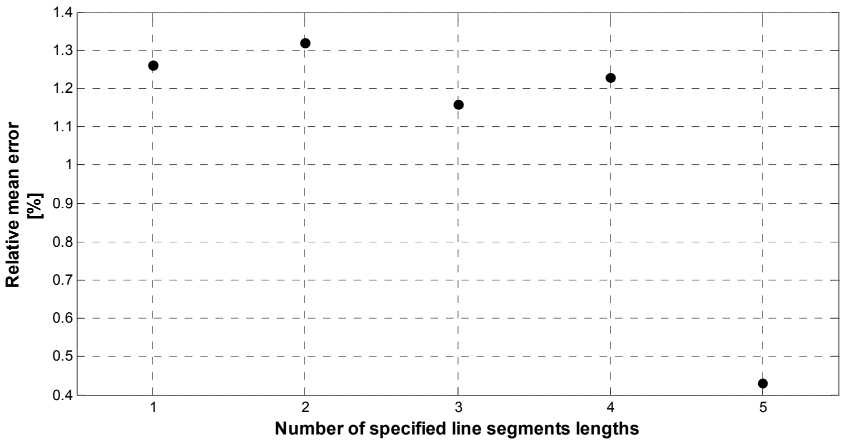


Fig. 9. Dependencies between relative mean error and number of specified data [source: own study]

Figure 10 shows recognized connections between roof surfaces. Solid lines indicate roof edges, the brightest one ridge\hip and the other color means that the line segment is a valley. In all three tested roofs, with the help of proposed algorithm it was possible to correctly identify types of connections.

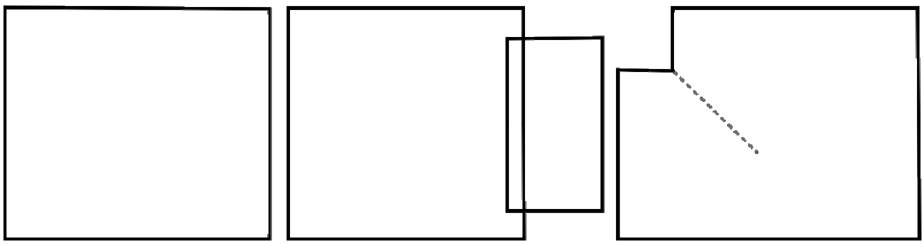


Fig. 10. Recognized connections [source: own study]

Figure 11 shows three-dimensional visualization of two tested roofs: roof with six surfaces and hip roof. Those models were created based on collected data and knowledge about architectural design such as base lines and slope angles. As it can be seen, proposed solution works well allowing to draw roof surfaces shapes in 3d.

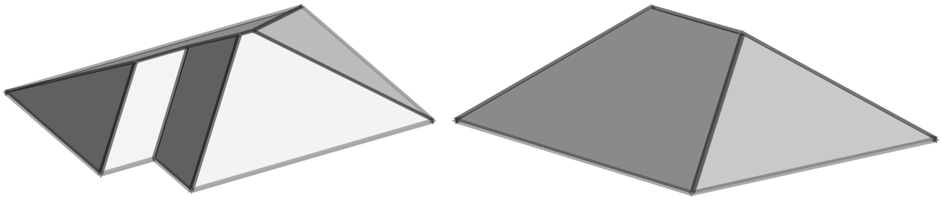


Fig. 11. 3D roof shape visualization [source: own study]

Further tests were designed to compare the effectiveness of proposed solution against two chosen free applications available online: calculator from company CB S.A. [8] and application from TONDACH [9]. Two criteria were taken into account. First of them – calculation – determines whether the solution allows to calculate roof area in specific cases. Both of the free applications allow only to calculate roof area of predefined types – they do not allow user to specify his own, custom project. Second criterion was the whether additional calculations are required before two applications or proposed solution may proceed to calculate the area. Both free applications require the user to provide real lengths of line segments – not its projection – which are given on the architectural plan. Proposed solution is more versatile, it is not needed to specify any lengths that are not given on plan. Table 1 shows comparison of proposed solution and two free applications.

Tab. 1. Comparison of two free applications CB S.A.,TONDACH and proposed solution capabilities. Roof types: A – shed roof, B – gable roof, C – hip roof, D – gable roof with outhouse, E – gambrel roof, F – roof with six surfaces

	CB S.A.		TONDACH		Proposed solution	
	calculation	requires additional calculation	calculation	requires additional calculation	calculation	requires additional calculatio n
A	+	+	+	+	+	-
B	+	+	+	+	+	-
C	+	+	+	+	+	-
D	+	+	+	+	+	-
E	+	+	-	X	+	-
F	-	X	-	X	+	-

4. SUMMARY

Paper presents the solution of presented problem that allows to automate roof area calculation process in substantial degree. All the steps, beside entering information about roof digital model, have been automated and those are for example: recognition of connections between roof surfaces, shape of surfaces recognition and 3D visualization. Main advantages of proposed solution are that it deals with incomplete information about roof surface and that it is possible for user to provide own architectural plans instead of using only predefined ones. Based on performed tests it can be said without hesitation that proposed solution can be used to solve described problem. Considering that roofing companies sell tiles with some excess, mean relative error on 1.02% level is not a problem.

To make presented idea even more universal, new functions, like calculating area of roof surfaces that are not flat, have to be added.

REFERENCES

[1] CEWE A.: *Tablice matematyczne: definicje, twierdzenia, własności, wzory*. Podkowa, Gdańsk, 2000.

[2] MEISTERS G.H.: *Polygons Have Ears*. The American Mathematical Monthly, vol. 82, no. 6, 1975.

[3] EBERLY D.: *Trangulation by Ear Clipping*. Geometric Tools, 2008.

[4] KLAWONN F.: *Introduction to Computer Graphics: Using Java 2D and 3D (Undergraduate Topics in Computer Science)*. Springer, 2012, pages: 87-91.

[5] SEDGEWICK R.: *Algorytmy*. Helion, Gliwice, 2012, pages: 147-149, 160-163.

[6] THE MATHWORKS, *Inc.: Matlab Documentation*. Version R2008A.

[7] DAHLQUIST G.: *Numerical methods in scientific computing*. Society for Industrial and Applied Mathematics, Philadelphia, 2008, page: 90.

[8] CB S.A.: *Corotop*. http://corotop.com.pl/pl/kalkulator, access: 2012-11-23.

[9] TONDACH: *Tondach*. http://www.tondach.pl/dane-techniczne/program-do-obliczen, access: 2012-11-23.