

P-Graph-based Workflow Modelling

József Tick

Institute for Software Engineering, John von Neumann Faculty of Informatics,
Budapest Tech
Bécsi út 96/B, H-1034 Budapest, Hungary
tick@bmf.hu

Abstract: Workflow modelling has been successfully introduced and implemented in several application fields. Therefore, its significance has increased dramatically. Several work flow modelling techniques have been published so far, out of which quite a number are widespread applications. For instance the Petri-Net-based modelling has become popular partly due to its graphical design and partly due to its correct mathematical background. The workflow modelling based on Unified Modelling Language is important because of its practical usage. This paper introduces and examines the workflow modelling technique based on the Process-graph as a possible new solution next to the already existing modelling techniques.

Keywords: workflow management, workflow modelling, P-graphs

1 Introduction

Workflow management was originally a tool for the organisation, analysis and rearrangement of business processes. Information technology has gone through one of the most significant changes in the last decades while having proliferated and penetrated into business processes in each and every business field and has become a determining component. In parallel the application of workflow has improved as well. By nowadays the number of application fields has increased a lot and after successes in the production and administrative processes, workflow management gained ground in almost each professional field. Further application fields include a wide range of sectors ranging from technical design [4] via solving of various computing problems [3] to several application of the processing of information data [11].

The drastic development in workflow applications has meant the computerisation of the existing processes and later, with the appearance of the Business Process Management (BPM), the total restructuring of these business processes. BPM examines and models the processes from various aspects, which meant the information based restructuring and reorganisation of this field.

1.1 Key Definitions of Workflow and the Elements

The concept of workflow is widely interpreted, and it might mean simple activity steps but can be interpreted as a synonym of the total Business Process. The Workflow Management Coalition (WfMC) as the acknowledged professional association strives to reach standardisation and also conducts widespread marketing to spread the concepts and methodologies linked to Workflow technology. Consequently, this paper takes the WfMC [1] definitions and interpretations for granted when reviewing the definitions.

Workflow: ‘The automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules.’ [1]

Activity: ‘A description of a piece of work that forms one logical step within a process. An activity may be a manual activity, which does not support computer automation, or a workflow (automated) activity. A workflow activity requires human and/or machine resources(s) to support process execution; where human resource is required an activity is allocated to a workflow participant.’ [1]

Process: ‘The representation of a business process in a form which supports automated manipulation, such as modelling, or enactment by a workflow management system. The process definition consists of a network of activities and their relationships, criteria to indicate the start and termination of the process, and information about the individual activities, such as participants, associated IT applications and data, etc.’ [1]

Instance: ‘The representation of a single enactment of a process, or activity within a process, including its associated data. Each instance represents a separate thread of execution of the process or activity, which may be controlled independently and will have its own internal state and externally visible identity, which may be used as a handle, for example, to record or retrieve audit data relating to the individual enactment.’ [1] Instances are process instances, which are representations of a single enactment of a process, or activity instances, which are representations of an activity within a single enactment of a process (within a process instance). Process instances are often called cases.

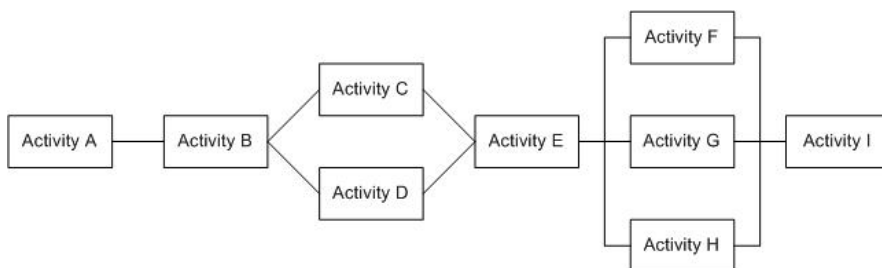


Figure 1

Example of a Workflow using WfMC's Notation

Figure 1 presents a simple example using the above definition and given the notation system provided by WfMC.

2 The Existing Workflow Modelling Methodologies

Taking workflow modelling into account five relevant workflow model perspectives can be differentiated based on the above [2], [5] and [8]. The **Control flow or process perspective** presents the workflow elements, their relationships, that is the static structure and organisation of the workflow. The control flow includes the time dependency between the elements, and the entire routing description valid for the workflow model. The **Resource or organization perspective** determines the types, the quantity and/or the availability and accessibility of the resources necessary for the execution of the tasks. It describes the roles from the perspective of functionality, and the groups from the aspect of the organisation as well as their labelled responsibility, authorisation and availability. The **Data or information perspective** includes the description of the control data needed for the operation of the workflow and the execution of the routine, and in the framework of production information the data, tables and documents containing the significant characteristics of production. The **Task or function perspective** defines the elementary operations carried out by the resources while performing a task. The **Operation or application perspective** specifies the elementary actions using specified applications. These applications could be general applications such as text editor, spreadsheet editor, or special applications developed for the given task. Traditional modelling focuses mainly on the first perspective.

This paper highlights two modelling techniques out of the several existing ones: the modelling technique based on the UML activity diagram and the modelling based on the Petri-Net.

Unified Modelling Language (UML) is a widespread modelling language in software engineering that is highly standardised and is rich in tool system. The control perspective of the Workflow model consists of a description process activities and its routing in cases [7]. This is the reason why UML's sequence diagram and activity diagram is useful to model the discussed aspects of workflow models. The research work done by W. van der Aalst et al. resulted in 21 several patterns which describe the behaviour of business processes. [7] S. A. White mapped this patterns from business process model to UML activity diagram. The usage of UML reflects a rather practice-oriented view, which is closely linked to the everyday work of software developers.

Petri-Net is one of the most important mathematical and graphical representation possibilities of distributed systems [12], networks and workflows. It has been a

popular workflow modelling tool for a long time [13]. Introducing Petri-Nets to the field of workflow modelling Aalst and Hee [5] specified a process using Petri-Net as the basic element of workflow. Van der Aalst with his more than three hundred publications worked out the whole theory and methodology of the Petri-Net based workflow management. He describes [14] how to map workflow management concepts onto Petri-Nets, defines the processes, the control flow possibilities, routing constructs, triggering, tasks, work items and activities. He focuses also on the analysis of workflow using Petri-nets.

3 Process Modelling Using P-Graphs

The P-graph based modelling as well as the process network synthesis generated by combinatorial methods are presented by using the works published by Friedler et al. The chapter discusses the mathematical description of the P-graphs mainly using the literature [15], [16], [17], [18], [19], [20].

3.1 Introduction of P-Graphs

The so called P-graph (Process graph), which is a directed bigraph, has been used for modelling network structures for some time. The vertices of the graph denote the operating units (O – operating units) and the materials (M – materials). The edges of the graph represent the material-flow between the materials and the operating units.

The P-graph is a bigraph, meaning that its vertices are in disjunctive sets and there are no edges between vertices in the same set. In case of P-graphs the assignment of operating units and materials are strictly determined by the tasks given, i.e. an edge can point to an M material type vertex from an O operating unit type vertex, only if M is element of the output set of O , that is O produces M material namely, $M \in \text{output}_O$. An edge can point from an M material type vertex to an O operating unit type vertex, only if M is element of the input set of O , that is O processes M material, namely $M \in \text{input}_O$. Thus, the P-graph can be presented by the pairs of operating unit and the assigned material vertices set like the (M, O) P-graph.

The material type vertices can be put into several subsets. There are various subsets like the raw-material type one, which contains the input elements of the whole process, the product-material type subset, which gathers the results of the entire process, the intermediate-material type one, the elements of which emerge or are used between the processing phases, and finally the by-product-material type set, which contains the non desired results of the process.

The applied operating unit and material element notations in the P-graph notation are presented in Figure 2. As an example let us consider a process network with 7

operating units, in which the operating units are $1,2,\dots,7$ and the materials are A,B,\dots,L . A,B,C and D are the materials available for the production of L . The possible structure is given in Figure 3.

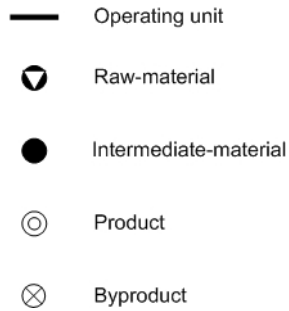


Figure 2
Notations used at P-graphs

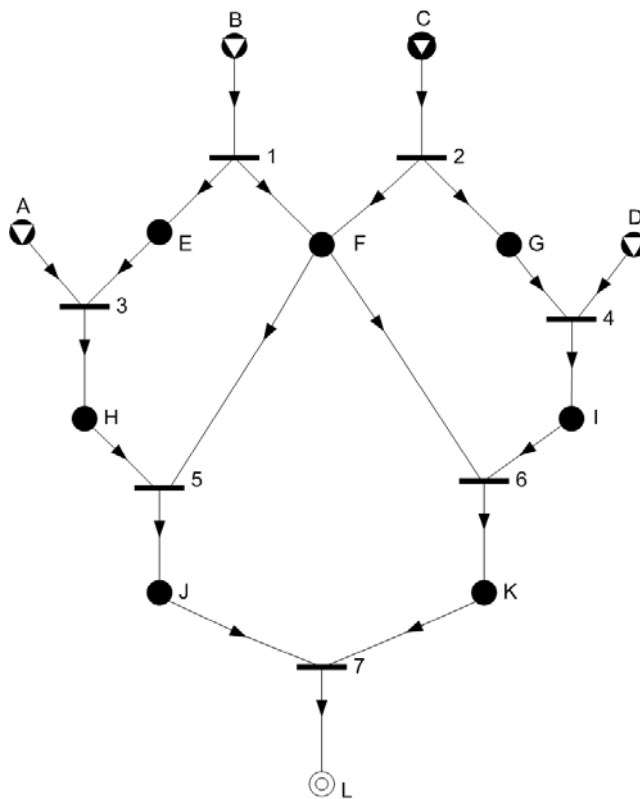


Figure 3
The notation system of a P-graph model

3.2 The Mathematical Definition of P-Graphs

There is a finite set of material M (which contains the sets of P products and R raw-materials) and the finite set of O operating units. Consequently, the set of P end-products and the set of R raw-materials must be subsets of M and the set of M materials and the set of O operating units are disjunctive. The basic relations between M, P, R and O are as follows:

$$P \subseteq M, R \subseteq M, \text{ and } M \cap O = \emptyset \quad (3.1)$$

As physical processes are defined, each operation unit produces output materials from input materials. Therefore two disjunctive sets can be assigned to each operating unit, i.e. the set of input and the set of output materials. Let an arbitrary operating unit (α, β) , then α is the set of input materials which are processed by the (α, β) unit and β is the set of output materials, which are produced by the given unit. Considering the process-network the output materials of each operating unit are the inputs of different operating units. In general, it can be proved that

$$O \subseteq \wp(M) \times \wp(M) \quad (3.2)$$

where O is the set of operating units, M is the set of materials and $\wp(M)$ is the power set, that is the set of subsets of M , and $\wp(M) \times \wp(M)$ represents the set of $\wp(M)$ and $\wp(M)$ pairs.

Supposing that there is a finite set m , which is a subset of M , i.e. it is true that $m \subseteq M$ and there is an o finite set, which is a subset of O , i.e. it is true that $o \subseteq O$ and supposing that there is such a material which is an input for one or more operating units, and there is such material which is the output of one or more operating units, then

$$o \subseteq \wp(m) \times \wp(m) \quad (3.3)$$

The P-graph is defined as a bigraph, where the set of V vertices is made of the elements of the union of m and o that is

$$V = m \cup o \quad (3.4)$$

Obviously, the vertices that are elements of m are M (material) type vertices, while the vertices that are elements of o are O (operating units) type ones.

It is true for the set of A edges of P-graph that the elements of A are the elements of the union of A_1 and A_2 , i.e.

$$A = A_1 \cup A_2 \quad (3.5)$$

where

$$A_1 = \{(x, y) \mid y = (\alpha, \beta) \in o \text{ és } x \in \alpha\} \quad (3.6)$$

and

$$A_2 = \{(y, x) \mid y = (\alpha, \beta) \in o \text{ és } x \in \beta\} \quad (3.7)$$

On the basis of (3.5)...(3.7) the edges of the graph can be A_1 and A_2 . The A_1 type edge, which is determined by (x,y) points from an input material vertex belonging to the operating unit to the operating unit vertex, considering that x is element of the input set α , while y is the element of o operating unit set, which is determined by the α and β material set pairs. Opposing the A_2 type edge, which is determined by (y,x) , points from an operating unit vertex to an output material vertex belonging to the operating unit, considering that y is the element of the o operating unit set determined by α and β material set pairs, while x is the element of the β output material set.

3.3 The Combinatorial Solution of the Process-Network Synthesis (PSN)

The primary aim of the process-network is to produce P products from R raw-materials. As the first step of PNS, all the plausible operating units O and intermediate-materials must be determined. By determining P, R and O , the set of all the material in the network, M is also defined.

The optimal solution structure generated by process-network synthesis must have several basic features that are taken for granted as axioms, and the introduction of which improves the efficiency of the combinatorial search during the process. These axioms are the following:

- (S1) Each final product is represented in the graph.
- (S2) An M -type vertex does not have an input if and only if it represents a raw-material.
- (S3) Each O -type vertex representing an operating unit is defined in the network synthesis problem.
- (S4) There must be at least one route from each O -type (operating unit) vertex represented in the structure leading to an M -type vertex representing a product.
- (S5) If an M -type vertex belongs to the graph, then there must be at least one route leading to an O -type vertex or a route from an O -type vertex to the given M -type vertex.

The number of all the combinatorially possible networks increases exponentially by the increase in the number of operating units. For example, in case of a process-network synthesis made up of 35 operating units [20] the number of all the potentially possible structures is $2^{35}-1$ that is 34,359,738,367. Supposing an average PC, continuous calculation, and assigning 10^{-2} sec processing time for each combination, the calculation and management of all the 34 billion combination would take more than 11 years.

The structures of the solutions must carry the features defined in the axioms. This is, however, only necessary but not satisfactory condition of the selection of the optimal structure. A practical reduction has been carried out in the number of the structures to be handled with the help of the axioms, thus leaving out and eliminating the redundant and not valid structures. The remaining structures, which fulfil the axioms, are called the combinatorially possible structures. With the help of this practical restriction the search field has been drastically reduced. (see Figure 4) Taking the previous example the number of the structures drops from 34 billion to 3465. Therefore the actual processing time drops from 11 years to 11 and half minutes.

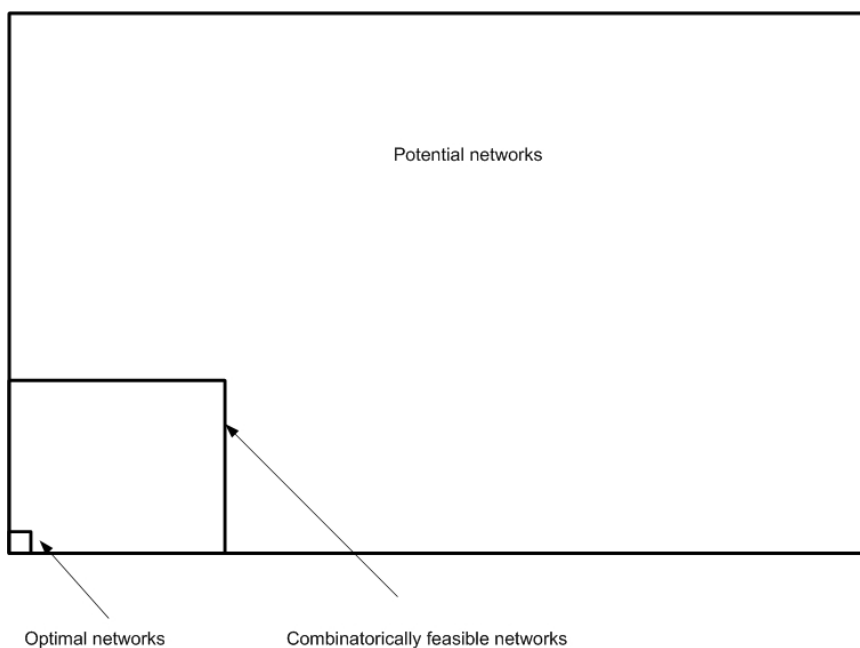


Figure 4

The decrease in search field with the help of the axioms

Such a reduction in the size of the search field, i.e. the exclusion of those structures which are at first sight cannot be taken as possible optimal structures that are the determination of the combinatorially possible structures cannot be

carried out in a mathematically precise way with the conventional methods of the process synthesis. The identification of all the possible structures and the selection of the optimal structure based on the conventional super-structure methods like MILP or MINLP cannot be used due to the features of the applied exponential algorithms. The Maximal Structure Generation (MSG) polynomial algorithm elaborated by Fiedler et al., which uses the 5 axioms, generates that maximal structure, all the combinatorially possible structures of which are its subsets.

3.4 The Generation of the Maximal Structure, the MSG Algorithm

The maximal structure of the synthesis problem (P, R, O) contains all the combinatorially possible structures, which make the production of defined products possible from given raw-materials. Therefore, it certainly contains the optimal structure as well.

The algorithm can be divided into four main phases:

The first phase is the input phase, in which the synthesis problem is defined (P, R, O) such a way, that the set of M all the plausible materials, the set of P end-products, the set of R raw-materials and the set of O operating units are given. M contains not only the intermediate-materials assigned to operating units defined in the set of O , but the raw-materials given in R and the end-products given in P as well.

The second phase is the elaboration of the input structure of the network, which is carried out by the linking of all the similar (same type) material type vertices.

The third phase is the elimination phase, where those materials and operating units are eliminated, which, taking the 5 axioms into account, are not and cannot be linked to the maximal structure for sure. The elimination is carried out step by step, starting from the deepest level of the input structure, from the raw-materials, and going from level to level by examining the vertices at the material level and the operating unit level in turns to see whether that fulfil all the 5 axioms. Certainly, the elimination of a vertex often leads to the elimination of further vertices linked to it.

During the fourth phase the vertices are linked again from level to level, starting from the highest, the end-product level.

The maximal structure generated this way contains all the combinatorially possible structures and all of its elements fulfil the 5 axioms.

3.5 The Generation of the Solution Structure, the SSG Algorithm

The maximal structure generated by the MSG algorithm contains all such combinatorically possible network structures that are able to produce the end-product from the given raw-materials. Consequently, it contains the optimal network as well. In most cases the optimisation means to find the most cost effective solution.

The application of the SSG (Solution Structure Generation) algorithm enables the production of all the solution structures. The SSG is a new mathematical tool based on the application of the Decision-mapping (DM) which has been developed by Friedler et al. [20].

On (M, O) P-graph let Δ denote a mapping between the M set of materials various groups of the (α, β) operating units in $\wp(O)$ power set including the empty set as well. The Δ mapping can be defined, which determines for $X \in M$ material the set of operation units producing X . The above defined mapping is $\Delta(X)$ which is a set itself:

$$\Delta(X) = \{(\alpha, \beta) \mid (\alpha, \beta) \in O \text{ és } X \in \beta\} \quad (3.8)$$

Let us suppose that m is a subset of M and X is an element of m , as well as $\delta(X)$ is one of the subsets of $\Delta(X)$ then a $\delta[m]$ decision mapping can be defined for m with the help of m and $\delta(X)$:

$$\delta[m] = \{(X, \delta(X)) \mid X \in m\} \quad (3.9)$$

The $\delta[m]$ decision mapping given for m determines the relation between m set and the subsets of O set, since the element of the set m i.e. the materials are generated as the various combinations of the operating units. Each element of the subsets of $\delta[m]$ that is the $\delta[m]$ s pair the given X materials of the set m with the groups of operating units producing the given material, that is with $\delta(X)$.

The complement of the decision-mapping $\bar{\delta}[m]$ is defined that in case of any $X \in m$, $\bar{\delta}(X) = \Delta(X) \setminus \delta(X)$, i.e. if $\delta(X)$ is the set of operating units producing X material then $\bar{\delta}(X)$ is the set of operating units, which produce X material but are not elements of $\delta(X)$, that is $\delta(X) \cup \bar{\delta}(X) = \Delta(X)$. The definition of the complement of the decision-mapping in mathematical terms is the following:

$$\bar{\delta}[m] = \{(X, Y) \mid X \in m \text{ és } Y = \Delta(X) \setminus \delta(X)\} \quad (3.10)$$

As the first step in the process representation made with decision-mapping an active set must be defined. Let us suppose that m' is an active set, and a subset of the

m material set of the (m, o) P-graph. This m' subset of the (m, o) P-graph is an active set if and only if at least one element of β output materials of each (α, β) operating unit is element of m' , i.e. $m' \cap \beta \neq \emptyset$. On the basis of this, the m' active set will definitely contain with certain restrictions one of the optimal solutions including all the operating units.

With the help and application of the algorithm generating the maximal structure and the decision-mapping given for P-graph, the steps of the realisation of the SSG algorithm can be defined. This procedure generates all the solution structures, i.e. produces the combinatorially possible solution structures.

In the input phase of the SSG algorithm the sets needed for the generation, the products in the P set, the raw-materials in the R set, and all the other materials (intermediate-materials and by-products) not defined in M , P and R are defined. The restriction given earlier must be true for the input sets, that is $R \subset M$, $P \subset M$ and $P \cap R \neq \emptyset$. A $\Delta[M]$ set must also be defined in the input phase, which contains the X materials and the $\Delta[X]$ mappings.

After the input phase the computer algorithm selects systematically and combinatorially the m active sets in a recursive way and executes the $[m]$ decision mapping on them. The algorithm works until the selection and execution of all the m active sets is done.

4 P-Graph Extension for Workflow Modelling

A workflow model can be considered as a network structure. As it was presented in the previous chapter, the so far described solutions have been based on graph modelling. A simple graph, however, does not contain enough information so as to give the model precisely. The Petri-Net is also a directed bigraph which, with a token supplement, enables modelling of demand of resources crucial for workflow management and/or of eventual information or of raw material.

The Process-graph or P-Graph has been implemented in the fields dealing with the network-oriented synthesis of process networks (PNS) [15, 16, 17, 18, 19]. The P-graph is also a directed bigraph, the vertices of which are of two types. One vertex type is operation unit type vertex, while the other one is the material type vertex.

The O operating units in the P-graph can be interpreted in the same way, thus they can be assigned regarding their functions and handling to the activities in the workflow. In the P-graphs the M materials are basically modelled as documents or/and document processes in the workflow. During the process a document behaves similarly to the materials, however, there are slight differences which can be assigned to materials but not to documents. This means smaller restrictions which, however, do not reduce the representation significance of the modelling.

Therefore, similarly to the previously written about P-graph based workflow model the following documents can be introduced deduced from the materials:

Raw-materials	—————▶	input documents
Intermediate-materials	—————▶	intermediate documents
By-product	—————▶	----- (cannot be interpreted)
End-product	—————▶	output documents

By-products cannot be interpreted in the workflow, since the processing of the documents is targeted, and it is not influenced technologically, namely no redundant documents are prepared on purpose.

The separation of the intermediate documents from the output documents is not unambiguous, since documents which seem to be intermediate documents in one phase will be an output documents in others. Therefore, such databases can be regarded as real intermediate documents, which are prepared during the workflow process and will become information sources in later phases.

In addition to the condition above, the axioms, (S1 ... S5) used during the traditional application of the P-graphs and presented in the third chapter, can also apply and be interpreted in the case of workflow modelling. The relations (3.1 ... 3.10) presented in the algorithms (MSG, SSG) used for generating PNS can also be interpreted. Consequently, the biggest advantage of the P-graph-based workflow modelling is that the PNS method introduced above can be applied for the determination of the optimal network.

A lot of problems are already solved with Petri-Net modelling. Model controlling problems have not been mentioned at all, and routing has not been dealt with at all. The advantage of the P-graph is that there is a sound mathematical background, the P-graph can be presented by set theory tools; a well elaborated methodology for the optimal synthesis of the network structure can be found in the [15], [16], [17] literature, which apparently have similar advantages in case of workflow modelling as in several other applications [18], [19].

Conclusions

Next to the already existing workflow modelling solutions, it is justified to search new solution methods. In parallel to the Petri-Net-based workflow modelling that has an extensive literature and is widely used in the profession, P-graph-based workflow modelling can take ground and prove to be a useful way of workflow modelling. As a future step, the task is to investigate further application fields of the P-graph based workflow modelling.

References

- [1] The Workflow Management Coalition Specification, Workflow Management Coalition Terminology & Glossary; Document Number WPMC-TC-1011; Document Status - Issue 3.0; February 1999

-
- [2] S. Jablonski, C. Bussler: Workflow Management: Modelling Concepts, Architecture, and Implementation. International Thomson Computer Press, 1996
- [3] Gy. Hermann: Distributed Computer System for Gauge Calibration, in Proceedings of 4th Serbian-Hungarian Joint Symposium on Intelligent Systems, SISY 2006, Subotica, Serbia, September 29-30, 2006, pp. 349-358
- [4] Gy. Hermann: The Design of a Submicron Precision Coordinate Measuring Machine, in Proceedings of 3rd Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence, SAMI 2005, Herlany, Slovakia, January 21-22, 2005, pp. 397-408
- [5] W. M. P. van der Aalst, K. M van Hee: Workflow Management – Models, Methods, and Systems, The MIT Press, Cambridge, Massachusetts, London, England, 2002
- [6] S. A. P. White: Process Modelling Notations and Workflow Patterns, [Online] http://www.omg.org/bp-corner/bp-files/Process_Modelling_Notations.pdf
- [7] R. Eshuis, R. Wieringa: Verification Support for Workflow Design with UML Activity Graphs, in the Proceedings of the 24th International Conference on Software Engineering Orlando, Florida, 19-25, May, 2002, pp. 166-176
- [8] J. Tick: Workflow Model Representation Concepts, in Proceedings of 7th International Symposium of Hungarian Researchers on Computational Intelligence, HUCI 2006, Budapest, Hungary, November 24-25, 2006, pp. 329-337, ISBN 963 7154 54X
- [9] P. Hruby: Specification of Workflow Management Systems with UML. In the Proceedings of the 1998 OOPSLA Workshop on Implementation and Application of Object-oriented Workflow Management Systems, Vancouver, BC 1998
- [10] T. Kövér, D. Vígh, Z. Vámosy: Improved Face Recognition in the MYRA System, in Proceedings 4th Serbian-Hungarian Joint Symposium on Intelligent Systems, Subotica, Serbia, September 29-30, 2006, pp. 187-195, ISBN 963 7154 50 7
- [11] Sz. Sergyán, L. Csink: Consistency Check of Image Databases, in Proceedings 2nd Romanian-Hungarian Joint Symposium on Applied Computational Intelligence, Timisoara, Romania, May 12-14, 2005, pp. 201-206
- [12] L. Horváth, I. J. Rudas: Evaluation of Petri Net Process Model Representation as a Tool of Virtual Manufacturing, in Proceedings SMC'98 Conference Proceedings, 1998 IEEE International Conference on Systems,

- Man, and Cybernetics, October 11-14, 1998, ISBN: 0-7803-4778-1, pp. 178-183
- [13] L. Horváth, I. J. Rudas: Modelling of Manufacturing Processes Using Object-oriented Extended Petri Nets and Advanced Knowledge Representations, in Proceedings IEEE International Conference on Systems, Man, and Cybernetics, Vancouver, BC, Canada, October 22-25, 1995, ISBN 0-7803-4778-1, Vol. 3, pp. 2576-2581
- [14] W. M. P. van der Aalst: The Application of Petri Nets to Workflow Management. The Journal of Circuits, Systems and Computers, Vol. 8(1), pp. 21-66, 1998, ISSN: 0218-1266
- [15] F. Friedler, K. Tarjan, Y. W. Huang, L. T. Fan: Combinatorial Algorithms for Process Synthesis, Computers Chem. Engng, Vol. 16, pp. 313-320 (1992)
- [16] F. Friedler, K. Tarjan, Y. W. Huang, L. T. Fan: Graph-Theoretic Approach to Process Synthesis: Axioms and Theorems, Chem. Engng Sci., Vol. 47, pp. 1973-1988 (1992)
- [17] F. Friedler, L. T. Fan, B. Imreh, Process Network Synthesis: Problem Definition, Networks, Vol. 28, pp. 119-124 (1998)
- [18] J. Varga: A folyamat-hálózatszintézis feladat kiterjesztései, PhD értekezés, Veszprémi Egyetem, Mérnöki Kar, Veszprém, 2000.
- [19] B. Bertók: Folyamathálózatok struktúráinak algoritmikus szintézise, PhD értekezés, Veszprémi Egyetem, Műszaki Informatikai Kar, Veszprém, 2003
- [20] F. Friedler, J. B. Varga, L. T. Fan: Decision Mapping: A Tool for Consistent and Complete Decisions in Process Synthesis, Chemical Eng. Sci. Vol. 50, pp. 1755-1768, 1995