On the Non-Uniqueness of the Solution to the Least-Squares Optimization of Pairwise Comparison Matrices

András Farkas
Institute for Entrepreneurship Management
Budapest Polytechnic
1034. Doberdó út 6., Budapest, Hungary
e-mail: farkasa@gsb.ceu.hu

Pál Rózsa
Department of Computer Science and Information Theory
Budapest University of Technology and Economics
H-1521 Budapest, Hungary
e-mail: rozsa@cs.bme.hu

Abstract: This paper develops the "best" rank one approximation matrix to a general pairwise comparison matrix (PCM) in a least-squares sense. Such quadratic matrices occur in the multi-criteria decision making method called the analytic hierarchy process (AHP). These matrices may have positive entries only. The lack of uniqueness of the stationary values of the associated nonlinear least-squares optimization problem is investigated. Sufficient conditions for the non-uniqueness of the solution to the derived system of nonlinear equations are given. Results are illustrated through a great number of numerical examples.

Keywords: nonlinear optimization, least-squares method, multiple-criteria decision making

1 Introduction

Let $\mathbf{A}=[a_{ij}]$ denote an $n\times n$ matrix with all entries positive numbers. $\mathbf{A}$ is called a *symmetrically reciprocal* (SR) matrix if the entries satisfy $a_{ij}a_{ji}=1$ for $i\neq j$, $i,j=1,2,...,n,$ and $a_{ii}=1$, $i=1,2,...,n$. These matrices were introduced by Saaty [1] in his multi-criteria decision making method called the analytic hierarchy process (AHP). Here an entry a_{ij} represents a *ratio*, i.e. the element a_{ij} indicates the strength with which decision alternative A_i dominates decision alternative A_j with respect to a given criterion. Such a *pairwise comparison matrix* (PCM) is usually constructed by eliciting experts' judgements. Then the fundamental objective is to derive implicit *weights*, $w_1,w_2,...,w_n$, for the given set of decision alternatives according to relative importance (priorities) measured on a ratio-scale.

Let **B**=[b_{ij}] denote an n×n matrix with all entries positive numbers. **B** is called a *transitive* matrix if $b_{ij}b_{jk}=b_{ik}$, for i,j,k=1,2, ...,n. In [2] it is proven that any transitive matrix is necessarily SR and has rank one. Two strongly related notations will be used for the weights: **W**=diag[w_i], i=1,2, ...,n, is the diagonal matrix with the diagonal entries $w_1,w_2,...,w_n$, and the (column) vector from $\mathbb{R}^n$ with elements $w_1,w_2, ...,w_n$ is denoted by **w**. Thus, **W** is a positive definite diagonal matrix if and only if **w** is an elementwise positive column vector. With these notations, and defining the n vector $\mathbf{e}^{\mathrm{T}}$=[1,1, ...,1] to be the row vector of $\mathbb{R}^n$, any transitive matrix **B** can now be written in the form

$$\mathbf{B} = \mathbf{W}^{-1}\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{W} = \left[\frac{w_j}{w_i}\right], \quad i,j = 1,2,...,n. \tag{1}$$

Using (1) it is easy to show that

$$\mathbf{B}\mathbf{W}^{-1}\mathbf{e} = n\,\mathbf{W}^{-1}\mathbf{e}. \tag{2}$$

From (2) it is seen that the only nonzero (dominant) eigenvalue of **B** is n and its associated Perron-eigenvector is $\mathbf{W}^{-1}\mathbf{e}$, i.e. a vector whose elements are the reciprocals of the weights.

In decision theory, a transitive matrix **B** is termed *consistent* matrix. Otherwise a PCM is termed *inconsistent*. Saaty [1] showed that the weights for a consistent PCM are determined by the elements, u_i, i=1,2, ...,n, of the principal right eigenvector **u** of matrix **B**, if **B** is a *consistent* matrix, i.e. if it is transitive. This solution for the weights is *unique* up to a multiplicative constant. Hence, this Perron-eigenvector becomes $\mathbf{u}=\mathbf{W}^{-1}\mathbf{e}$. (In the applications of the AHP these components are usually normalized so that their sum is unity.)

During the last decades several authors have advocated particular best ways for approximating a *general* (*not* transitive) SR matrix **A**. There are various possible ways to generate approximations for **A** in some sense. Saaty [3] proposed the eigenvector approach for finding the weights, even if **A** is an inconsistent PCM. Extremal methods have also been considered, like the direct least-squares method [4], the weighted least-squares method [5],[6], the logarithmic least-squares method [7],[8], furthermore, the logarithmic least-absolute-values method [9]. A graphical technique that is based on the construction of the Gower-plots was also proposed which produces the “best” rank two matrix approximation to **A** [10]. The most comprehensive comparative study that has appeared thus far both in terms of the number of these scaling methods and the number of the evaluation criteria used was presented by Golany and Kress [11]. They concluded that these methods have different weaknesses and advantages, hence, none of them is dominated by the other.

2 The Least-Squares Optimization Method

The authors have developed a method that generates a "best" transitive (rank one) matrix **B** to approximate a general SR matrix **A**, where the "best" is assessed in a least-squares (LS) sense [12]. There it is shown that a common procedure to find a positive vector of the weights can be done by minimizing the expression

$$S^2(\mathbf{w}) := \|\mathbf{A} - \mathbf{B}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - \frac{w_j}{w_i}\right)^2 . \tag{3}$$

(Here, the subscript F denotes the Frobenius norm; the square root of the sum of squares of the elements, i.e. the error.)

Given the subjective estimates a_{ij} for a particular PCM, it is always desired that $a_{ij} \approx w_j / w_i$. In other words, the *weights* w_i, and thus the *consistency adjustments*, $a_{ij} - b_{ij}$, $i,j = 1,2, \ldots, n$, should be determined such that the sum of the *consistency adjustment error, S*, is minimized. In [12], with appropriately chosen initial values, the Newton-Kantorovich (NK) method was applied for this optimization procedure due to its computational advantages. There the authors asserted that a stationary value **w**, of the error functional $S^2(\mathbf{w})$, called a *stationary vector*, satisfies the following *homogeneous* nonlinear equation

$$\mathbf{R}(\mathbf{w})\mathbf{w} = \mathbf{0}, \tag{4}$$

where

$$\mathbf{R}(\mathbf{w}) = \mathbf{W}^{-2}(\mathbf{A} - \mathbf{W}^{-1}\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{W}) - (\mathbf{A} - \mathbf{W}^{-1}\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{W})^{\mathrm{T}}\mathbf{W}^{-2} \tag{5}$$

is a variable dependent, *skew-symmetric* matrix. In [13] it is shown that expression (5) can be more generally used in the approximation of merely *positive* matrices (which are not necessarily SR).

In the present paper, with regard to the basic properties of a PCM, investigations are made for general *symmetrically reciprocal* (SR) matrices **A** with positive entries. If matrix **A** is in SR, the homogeneous system of n nonlinear equations (4) can be written in the form

$$f = (w_1, w_2, ..., w_n) = \sum_{k=1}^{n} \left[\frac{a_{ik}}{w_i^2} + \frac{1}{a_{ik} w_k^2} - 2\left(\frac{w_i}{w_k^3} + \frac{w_k}{w_i^3} \right) \right] w_k = 0, \quad i = 1,2,...,n. \qquad (6)$$

Note that each of the n equations in (6) represents a type of homogeneous function in the variables as

$$f_i(cw_1, cw_2, ..., cw_n) = c^v f_i(w_1, w_2, ..., w_n), \qquad i = 1,2,...,n, \qquad (7)$$

where c is an arbitrary constant and $v = -1$ is the degree of the homogeneous function in (6). Substituting (7) for the set of equations (6) we get

$$\frac{1}{c} \sum_{k=1}^{n} \left[\frac{a_{ik}}{w_i^2} + \frac{1}{a_{ik} w_k^2} - 2\left(\frac{w_i}{w_k^3} + \frac{w_k}{w_i^3} \right) \right] w_k = 0, \qquad i = 1,2,...,n. \qquad (8)$$

It is apparent from (8) that any constant multiple of the solution to the homogeneous nonlinear system (4) would produce an other solution. To circumvent this difficulty, equation (4) can be reformulated, as any one of its n scalar equations can be dropped without affecting the solution set [12]. Denoting the jth row of any matrix $\mathbf{M}$ by $\mathbf{M}_{j*}$ and introducing the nonzero vector $\mathbf{c} \in \mathbb{R}^n$, let (4) have a positive solution $\mathbf{w}$ normalized so that $\mathbf{c}^{\mathrm{T}}\mathbf{w}=1$. Then, for any j, $1 \leq j \leq n$, apparently, the stationary vector $\mathbf{w}$ is a solution to the following *inhomogeneous* system of n equations

$$\mathbf{c}^{\mathrm{T}}\mathbf{w} = 1, \qquad \mathbf{R}_{k*}(\mathbf{w})\mathbf{w} = \mathbf{0}, \quad k \neq j, \quad 1 \leq k \leq n. \qquad (9)$$

Here it is convenient to use $j=1$. Thus $\mathbf{c}^{\mathrm{T}}=[1,0, ...,0]$, i.e. the normalization condition in (9) is then $w_1=1$.

Although a great number of numerical experiments showed that authors strategy always determined a convergent process for the NK iteration, however, a possible *non-uniqueness* of this solution (local minima) have also experienced. The occurrence of such alternate stationary vectors for a PCM was first reported by Jensen [4, p.328]. He argued that it is possible to specify PCMs "that have certain symmetries and high levels of response inconsistency that result in multiple solutions yielding minimum least-squares error." Obviously, an eventual occurrence of a possible multiple solution may not seem surprising since it is well-known in the theory of nonlinear optimization that $S^2(\mathbf{w})$ does not necessarily have a unique minimum. In what follows now, some respective results of the authors is discussed for this problem.

3 On the Non-Uniqueness of the Solution to the Least-Squares Optimization Problem

In this Section *sufficient* conditions for a *multiple* solution to the inhomogeneous system of n equations (9) are given. The following matrices will play an important role in this subject matter:

Definition 1 An n×n matrix $\mathbf{Z}$=[z_{ij}] is said to be *persymmetric* if its entries satisfy

$$z_{ij} = z_{n+1-j,n+1-i}, \qquad i, j = 1,2, \ldots, n, \tag{10}$$

i.e., if its elements are symmetric about the counterdiagonal (secondary diagonal).

Definition 2 An n×n matrix $\mathbf{P}_n$ is called a *permutation* matrix and is described by $\mathbf{P}_n = [\mathbf{e}_{j_1}\ \mathbf{e}_{j_2} \ldots \mathbf{e}_{j_n}]$, where the n numbers in the indices, p=($j_1\ j_2\ \ldots\ j_n$), indicate a particular permutation from the standard order of the numbers 1,2, ...,n.

It is easy to see that any permutation matrix $\mathbf{P}_n$ is an orthogonal matrix, since

$$\mathbf{P}_n^{\mathrm{T}}\mathbf{P}_n = \mathbf{I}_n, \tag{11}$$

where $\mathbf{I}_n$ denotes the n×n identity matrix.

Definition 3 An n×n matrix $\mathbf{M}$=[m_{ij}], i,j=1,2, ...,n, is called a *symmetric permutation invariant* (SPI) matrix if there exists an n×n permutation matrix $\mathbf{P}_n$ such that

$$\mathbf{P}_n^{\mathrm{T}}\mathbf{M}\mathbf{P}_n = \mathbf{M}. \tag{12}$$

is satisfied [14].

Definition 4 By a *circulant matrix*, or *circulant* for short, is meant an n×n matrix $\mathbf{C}$=[c_{jk}], j,k=1,2, ...,n, where

$$c_{jk} = \begin{cases} c_{1,k+1-j}, & \text{if} \quad j \le k, \\ c_{1,n+k+1-j}, & \text{if} \quad j > k. \end{cases} \tag{13}$$

The elements of each row of **C** in (13) are identical to those of the previous row, but are moved one position to the right and wrapped around. Thus, the whole circulant is evidently determined by the first row as

$$\mathbf{C} = \mathrm{circ}[c_{11}, c_{12}, \ldots, c_{1n}]. \tag{14}$$

It is meaningful to use a different notation for a special class of the permutation matrices. Among the permutation matrices the following matrix plays a fundamental role in the theory of circulants. This refers to the forward shift permutation, that is to the cycle p=(1,2, ...,n) generating the cyclic group of order n, since its factorization consists of one cycle of full length n (see in [15]).

Definition 5 The special n×n permutation matrix $\mathbf{\Omega}_1$ of the form

$$\mathbf{\Omega}_1 = [\mathbf{e}_n \quad \mathbf{e}_1 \quad \mathbf{e}_2 \ldots \mathbf{e}_{n-1}], \tag{15}$$

is said to be the *elementary (primitive) circulant* matrix, i.e. $\mathbf{\Omega}_1$ = circ[0, 1, 0, ..., 0]. The other n×n *circulant* permutation matrices $\mathbf{\Omega}_k$ of the form

$$\mathbf{\Omega}_k = [\mathbf{e}_{n-k+1} \quad \mathbf{e}_{n-k+2} \ldots \mathbf{e}_n \quad \mathbf{e}_1 \ldots \mathbf{e}_{n-k}], \qquad k = 1,2,\ldots,n, \tag{16}$$

are the powers of matrix $\mathbf{\Omega}_1$ defined by (15).

Notice in (16) that the relation $\mathbf{\Omega}_k = \mathbf{\Omega}_1{}^k$, holds for all k=1,...,n−1, and, obviously, $\mathbf{\Omega}_1{}^n = \mathbf{I}_n$. It follows from (14) that a circulant **C** is invariant to a cyclic (simultaneous) permutation of the rows and the columns, hence

$$\mathbf{\Omega}_k^{\mathrm{T}} \mathbf{C} \mathbf{\Omega}_k = \mathbf{C}, \qquad k = 1,2,\ldots,n-1, \tag{17}$$

where $\mathbf{\Omega}_k$ is a particular circulant permutation matrix. Thus, by Definition 3, any circulant matrix is an SPI matrix. Also, it can be readily shown that a circulant **C** may be expressed as a polynomial of the elementary circulant matrix in the form of

$$\mathbf{C} = \mathrm{circ}[c_{11}, c_{12}, \ldots, c_{1n}] = c_{11}\mathbf{I}_n + c_{12}\mathbf{\Omega}_1 + c_{13}\mathbf{\Omega}_1^2 + \ldots + c_{1n}\mathbf{\Omega}_1^{n-1}. \tag{18}$$

Definition 6 The special *n*×*n* permutation matrix $\mathbf{K}_n$, which has 1's on the main counterdiagonal and 0's elsewhere, i.e.,

$$\mathbf{K}_n = \begin{bmatrix} 0 & 0 & . & . & . & 0 & 1 \\ 0 & 0 & . & . & . & 1 & 0 \\ . & . & & & . & . & . \\ . & . & & . & & . & . \\ . & . & . & & & . & . \\ 0 & 1 & . & . & . & 0 & 0 \\ 1 & 0 & . & . & . & 0 & 0 \end{bmatrix}, \tag{19}$$

is called a *counteridentity* matrix.

Using (19), it may be easily shown that the following expression,

$$\mathbf{K}_n \mathbf{A} \mathbf{K}_n = \mathbf{A}^{\mathrm{T}}, \tag{20}$$

holds for a persymmetric SR matrix **A**.

Remark 1 The special *n*×*n* permutation matrices, $\mathbf{\Omega}_k$ and $\mathbf{K}_n$, defined by (16) and (19), respectively, are persymmetric matrices.

In the sequel we will provide *sufficient* conditions for the occurrence of multiple solutions to the inhomogeneous system of *n* equations (9).

Proposition 1 Let $\mathbf{A}=[a_{ij}]$ be an *n*×*n* SR matrix with positive entries. Let a (positive) stationary vector of the error functional (3) be derived and be denoted by $\mathbf{w}^*$. If **A** is a symmetric permutation invariant (SPI) matrix to a certain permutation matrix $\mathbf{P}_n$, then $\mathbf{P}_n^{\mathrm{T}}\mathbf{w}^*$ produces an alternate stationary vector, provided that $\mathbf{P}_n^{\mathrm{T}}\mathbf{w}^*$ and $\mathbf{w}^*$ are linearly independent. If this permutation is consecutively repeated (not more than *n* times over) then the vectors, $\mathbf{P}_n^{\mathrm{T}}\mathbf{w}^*, \mathbf{P}_n^{\mathrm{T}^2}\mathbf{w}^*, \mathbf{P}_n^{\mathrm{T}^3}\mathbf{w}^*, \ldots$ represent alternate stationary vectors, provided that they are linearly independent.

Proof. Write the Frobenius norm of the nonlinear LS optimization problem (3) in the form

$$S^2(\mathbf{w}) = \left\| \mathbf{A} - \mathbf{W}^{-1}\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{W} \right\|_F^2. \tag{21}$$

Let $\mathbf{P}_n$ be an arbitrary $n\times n$ permutation matrix. Considering the fact that the sum of squares of the elements of a matrix is not affected by any permutation of the rows and the columns of this matrix the Frobenius norm does not vary by postmultiplying the matrix ($\mathbf{A}-\mathbf{W}^{-1}\mathbf{ee}^{\mathrm{T}}\mathbf{W}$) by an arbitrarily chosen permutation matrix $\mathbf{P}_n$, and then by premultiplying it by its transpose $\mathbf{P}_n^{\mathrm{T}}$. Therefore,

$$S^2(\mathbf{w}) = \left\|\mathbf{P}_n^{\mathrm{T}}(\mathbf{A} - \mathbf{W}^{-1}\mathbf{ee}^{\mathrm{T}}\mathbf{W})\mathbf{P}_n\right\|_F^2 = \left\|\mathbf{P}_n^{\mathrm{T}}\mathbf{AP}_n - \mathbf{P}_n^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{P}_n\ \mathbf{P}_n^{\mathrm{T}}\mathbf{e}\ \mathbf{e}^{\mathrm{T}}\mathbf{P}_n\ \mathbf{P}_n^{\mathrm{T}}\mathbf{WP}_n\right\|_F^2. \qquad (22)$$

Observe that in (22) $\mathbf{P}_n^{\mathrm{T}}\mathbf{e} = \mathbf{e}$ and $\mathbf{e}^{\mathrm{T}}\mathbf{P}_n = \mathbf{e}^{\mathrm{T}}$. For an SPI matrix $\mathbf{A}$, by (12), $\mathbf{P}_n^{\mathrm{T}}\mathbf{A}\ \mathbf{P}_n = \mathbf{A}$ holds. Thus,

$$S^2(\mathbf{w}) = \left\|\mathbf{A} - \mathbf{P}_n^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{P}_n\ \mathbf{ee}^{\mathrm{T}}\ \mathbf{P}_n^{\mathrm{T}}\mathbf{WP}_n\right\|_F^2. \qquad (23)$$

In (23), the terms $\mathbf{P}_n^{\mathrm{T}}\mathbf{WP}_n$ and $\mathbf{P}_n^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{P}_n$ represent the permutations of the elements of $\mathbf{W}$ and $\mathbf{W}^{-1}$, respectively. After they have been permuted by the permutation matrix $\mathbf{P}_n = [\mathbf{e}_{j_1}\ \mathbf{e}_{j_2} \ldots \mathbf{e}_{j_n}]$, the elements of $\mathbf{P}_n^{\mathrm{T}}\mathbf{WP}_n$ (and the elements of $\mathbf{P}_n^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{P}_n$) are: $w_{j_1}, w_{j_2}, \ldots, w_{j_n}$, (and their inverses). If the derived stationary vector, $\mathbf{w}^*$ is linearly independent of the vector $\mathbf{P}_n^{\mathrm{T}}\ \mathbf{w}^*$, i.e., if $\mathbf{P}_n^{\mathrm{T}}\ \mathbf{w}^* \neq c\ \mathbf{w}^*$, where c is an arbitrary constant, then

$$\mathbf{P}_n^{\mathrm{T}}\mathbf{WP}_n\mathbf{e} = \mathbf{P}_n^{\mathrm{T}}\mathbf{We}$$

becomes an alternate stationary vector. By repeating this procedure we may get

$$\mathbf{P}_n^{\mathrm{T}^2}\mathbf{We},$$

which constitutes an other stationary vector, provided that this solution is linearly independent of both of the previous solutions. This way, the process can be continued as long as new linearly independent solutions are obtained. This completes the proof.◊

Corollary 1 If an $n\times n$ SR matrix $\mathbf{A}$ is a circulant matrix then its factorization consists of one cycle of full length by the circulant permutations, $\mathbf{\Omega}_k\mathbf{w}^*$**,** k=1,2, ...,n, (i.e. if $\mathbf{P}_n^{\mathrm{T}}$ is an elementary circulant matrix) and the total number of alternate stationary vectors of the error functional (9) is n.

It is well-known that any permutation, p=($j_1\ j_2 \ldots j_n$), may be expressed as the product of the circulant permutations, p=(p_1)(p_2)(p_3) ... (p_r), where p_i is a circulant permutation of s_i elements called a *cyclic group*, where $\sum_{i=1}^{r} s_i = n$. Thus, after an appropriate rearrangement of the rows and the columns, any permutation matrix $\mathbf{P}_n$ may be written in the form of a *block diagonal* matrix with the circulants $\mathbf{\Omega}^{(s_i)}$ of order s_i, i=1,2, ...,r, being placed on its main diagonal as follows

$$\mathbf{P}_n = \begin{bmatrix} \mathbf{\Omega}^{(s_1)} & & & & & \\ & \mathbf{\Omega}^{(s_2)} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \mathbf{\Omega}^{(s_r)} \end{bmatrix}. \tag{24}$$

Using this notation, after performing an appropriate rearrangement of the rows and the columns, it implies that any SPI matrix **M**, defined by (12), can be partitioned in the form

$$\mathbf{M} = [\mathbf{M}_{ij}],$$

where every s_i×s_j block, $\mathbf{M}_{ij}$, i,j=1,2, ...,r, satisfies the relation

$$\mathbf{\Omega}^{(s_i)^{\mathrm{T}}}\mathbf{M}_{ij}\mathbf{\Omega}^{(s_j)} = \mathbf{M}_{ij}, \qquad i, j = 1,2, ..., r. \tag{25}$$

Apply now the above considerations to an SPI matrix **A,** which is in SR. Perform a (simultaneous) rearrangement of the rows and the columns of **A**. Let the resulting matrix be denoted by $\tilde{\mathbf{A}}$. Then, obviously, for $\tilde{\mathbf{A}}$ the following relation holds

$$\mathbf{\Omega}^{(s_i)^{\mathrm{T}}}\tilde{\mathbf{A}}_{ij}\mathbf{\Omega}^{(s_j)} = \tilde{\mathbf{A}}_{ij}, \qquad i, j = 1,2, ..., r. \tag{26}$$

Hence, in case of s_i=s_j, the matrix $\tilde{\mathbf{A}}$ has circulant SR matrices of order s_i for the blocks, $\tilde{\mathbf{A}}_{ii}$, on the main diagonal, where the order s_i is odd (otherwise an SR matrix cannot be a circulant), or all elements of $\tilde{\mathbf{A}}_{ii}$ are equal to 1, if the order s_i is even. It follows from the definition of an SR matrix that any other block, $\tilde{\mathbf{A}}_{ij}$, (i≠j, s_i=s_j), might be a circulant of order s_i satisfying

$$\tilde{\mathbf{A}}_{ji} = \tilde{\mathbf{A}}_{ij}^{-\mathrm{T}}, \qquad (i \neq j), \tag{27}$$

where $\Sigma^r_{i=1}s_i=n$ and $\tilde{\mathbf{A}}_{ij}^{-\mathrm{T}}$ denotes the transpose of the block containing the reciprocals of the elements of $\tilde{\mathbf{A}}_{ij}$. If for any pair (i,j), $s_i \neq s_j$, $i,j=1,2, ...,r$, then the off-diagonal block, $\tilde{\mathbf{A}}_{ij}$, is an $s_i \times s_j$ rectangular block. Since for $s_i \neq s_j$,

$$\mathbf{\Omega}^{(s_i)^{\mathrm{T}}} \tilde{\mathbf{A}}_{ij} \mathbf{\Omega}^{(s_j)} = \tilde{\mathbf{A}}_{ij} \tag{28}$$

holds, a sufficient condition for $\tilde{\mathbf{A}}_{ij}$ is that its elements are equal.

Corollary 2 Let $\tilde{\mathbf{A}}$ be an $n \times n$ positive SR matrix whose rows and columns have been appropriately rearranged to be an SPI matrix. Let a (positive) stationary vector of the error functional (3) be determined. Let this solution be denoted by $\mathbf{w}^*$. Then the permutations $\mathbf{P}_n^{\mathrm{T}}\mathbf{w}^*, \mathbf{P}_n^{\mathrm{T}^2}\mathbf{w}^*, \mathbf{P}_n^{\mathrm{T}^3}\mathbf{w}^*, ...$ are also solutions, where $\mathbf{P}_n{}^{\mathrm{T}}$ is defined by (24). The total number of the alternate stationary vectors as solutions to equation (9) cannot exceed the least common multiple of $s_1,s_2, ...,s_r$. (see the proof in [13].)

Proposition 2 Let $\mathbf{A}=[a_{ij}]$ be an $n \times n$ SR matrix with positive entries. Let a (positive) stationary vector of the error functional (3) be determined and let this solution of eq. (9) be denoted by $\mathbf{w}^{*\mathrm{T}(1)} = [1, w_2^{*(1)}, ..., w_n^{*(1)}]$. If $\mathbf{A}$ is a persymmetric matrix, then

$$\mathbf{w}^{*\mathrm{T}(2)} = \left[\frac{1}{w_{n-1}^{*(1)}}, \frac{1}{w_{n-2}^{*(1)}}, \ ..., 1 \right], \tag{29}$$

is an alternate stationary vector as an other solution of equation (9), provided that the latter solution $\mathbf{w}^{*(2)}$ is linearly independent of $\mathbf{w}^{*(1)}$, i.e., if

$$w_n^{*(1)} \neq (w_i^{*(1)})(w_{n+1-i}^{*(1)}), \qquad i = 1,2, ...,n. \tag{30}$$

Proof. Write the Frobenius norm of the nonlinear LS optimization problem (3) in the form

$$S^2(\mathbf{w}) = \left\| \mathbf{A} - \mathbf{W}^{-1}\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{W} \right\|_F^2. \tag{31}$$

Consider the $n \times n$ counteridentity (permutation) matrix, $\mathbf{K}_n$ defined by (19). Since $\mathbf{K}_n$

is an involutory matrix, therefore, $\mathbf{K}_n^2 = \mathbf{I}_n$. Let $\mathbf{P}_n$ be an arbitrary n×n permutation matrix. Recognize that $\mathbf{I}_n = \mathbf{P}_n\,\mathbf{P}_n^{\mathrm{T}} = \mathbf{P}_n\mathbf{K}_n\mathbf{K}_n\mathbf{P}_n^{\mathrm{T}}$. Now apply the same technique that was used for the proof of Proposition 1. Thus, one may write that

$$S^2(\mathbf{w}) = \left\|\mathbf{K}_n\mathbf{A}\mathbf{K}_n - \mathbf{K}_n\mathbf{W}^{-1}\mathbf{K}_n\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{K}_n\mathbf{W}\mathbf{K}_n\right\|_F^2 = \left\|\mathbf{A}^{\mathrm{T}} - \mathbf{K}_n\mathbf{W}^{-1}\mathbf{K}_n\mathbf{e}\mathbf{e}^{\mathrm{T}}\mathbf{K}_n\mathbf{W}\mathbf{K}_n\right\|_F^2. \qquad (32)$$

Making use of (20), the transpose of the matrix in the right hand side of (32) is, apparently,

$$S^2(\mathbf{w}) = \left\|\mathbf{A} - \mathbf{K}_n\mathbf{W}\mathbf{K}_n\,\mathbf{e}\mathbf{e}^{\mathrm{T}}\,\mathbf{K}_n\mathbf{W}^{-1}\mathbf{K}_n\right\|_F^2. \qquad (33)$$

It is obvious from (33) that the elements of the matrix $\mathbf{K}_n\mathbf{W}^{-1}\mathbf{K}_n$ are composed of the elements of a vector $\mathbf{w}^{*(2)}$, which also constitutes a stationary vector. If this solution is linearly independent of $\mathbf{w}^{*(1)}$, then it must represent an alternate stationary vector as the entries of $\mathbf{K}_n\mathbf{W}^{-1}\mathbf{K}_n$ are: $\dfrac{1}{w_{n-1}^{*(1)}}, \dfrac{1}{w_{n-2}^{*(1)}}, \ldots, 1.$ If (30) is satisfied, then they are linearly independent. This completes the proof. ◊

Corollary 3 Suppose that for the stationary vector $\mathbf{w}^{*(1)}$ the equality

$$\left[\,1, \frac{w_n^{*(1)}}{w_{n-1}^{*(1)}}, \frac{w_n^{*(1)}}{w_{n-2}^{*(1)}}, \ldots, w_n^{*(1)}\right] = [\,1, w_2^{*(1)}, w_3^{*(1)}, \ldots, w_n^{*(1)}] \qquad (34)$$

is satisfied, i.e., the relation

$$w_n^{*(1)} = (w_i^{*(1)})(w_{n+1-i}^{*(1)}), \qquad i = 1,2, \ldots, n, \qquad (35)$$

holds. Then, (34) provides *one* solution to the nonlinear optimization problem (3). In this case no trivial alternate stationary vector can be found. It should be noted, however, that one might not call this solution a *unique* solution until the *necessary* conditions for the non-uniqueness problem of the solution of equation (9) have not been found, because, at this point, the existence of an other stationary value cannot be excluded.

To summarize the results of the developments made in this Section the following theorem gives *sufficient* conditions for the occurrence of a *non-unique* stationary vector of the error functional (3) as a solution to equation (9).

Theorem 1 Let **A** be a general *n*×*n* SR matrix with positive entries.

(i) If **A** is a *circulant* matrix, or

(ii) if **A**=[$\mathbf{A}_{ij}$] is a block SR matrix with $s_i \times s_j$ blocks, where $\mathbf{A}_{ii}$ are circulant SR matrices, $\mathbf{A}_{ij}$ are circulant matrices for $i \neq j$, $s_i = s_j$ and $\mathbf{A}_{ij}$ has equal entries for $i \neq j$, $s_i \neq s_j$ and all blocks satisfy (27), or

(iii) if **A** is a *persymmetric* matrix, and for a given solution the relation

$$w_i^* \neq \frac{w_n^*}{w_{n+1-i}^*}, \qquad i = 1,2,...,n, \tag{36}$$

is satisfied, and

(iv) if a (positive) solution to equation (9), under the condition (i), or (ii), or (iii), represents a *stationary vector* $\mathbf{w}^*=[1,w_2^*, ...,w_n^*]$ (a local minimum),

then, this solution, $\mathbf{w}^*$, of the nonlinear least-squares optimization problem (3) is a *non-unique* stationary point.

4 Numerical Illustrations

The illustrations presented in this Section were selected to demonstrate the results of our paper. The numerical computations are made by "Mathematica". Seven examples for given positive SR matrices, **A**, (PCMs) are discussed in some detail below. For these examples Saaty's reciprocal nine-point scale: [1/9, ...,1/2, 1, 2, ..., 9] is used for the numerical values of the entries of **A**. The numerical experiments reported below include computation of the Hessian matrices. In every case they were found to be positive definite, thus ensuring that each stationary value computed was a local minimum.

EXAMPLE 1. The first example concerns data of a 5 ×5 PCM and demonstrates the occurrence of alternate optima, if **A** is a circulant matrix. Since matrix **A** is in SR by definition, for sake of simplicity the entries below the main diagonal are not depicted. This response matrix $\mathbf{A}_1$ is specified as

$$\mathbf{A}_1 = \begin{bmatrix} 1 & a & 1/a & a & 1/a \\ & 1 & a & 1/a & a \\ & & 1 & a & 1/a \\ & & & 1 & a \\ & & & & 1 \end{bmatrix}.$$

Let a=9. Applying the (elementary) circulant permutation matrix, $\mathbf{P}_n = \mathbf{\Omega}_k$, k=1,2, ..,5, the following linearly independent solutions (five alternate minima) are obtained (see Corollary 1):

$$\mathbf{w}^{*\mathrm{T}(1)} = [\, 1 \,, 0.3354 \,, 1.9998 \,, 0.5000 \,, 2.9814 \,],$$

$$\mathbf{w}^{*\mathrm{T}(2)} = [\, 1 \,, 5.9623 \,, 1.4909 \,, 8.8890 \,, 2.9814 \,],$$

$$\mathbf{w}^{*\mathrm{T}(3)} = [\, 1 \,, 5.9623 \,, 1.9998 \,, 0.6708 \,, 3.9993 \,],$$

$$\mathbf{w}^{*\mathrm{T}(4)} = [\, 1 \,, 0.3354 \,, 0.1125 \,, 0.6708 \,, 0.1677 \,],$$

$$\mathbf{w}^{*\mathrm{T}(5)} = [\, 1 \,, 0.2500 \,, 1.4909 \,, 0.5000 \,, 0.1677 \,],$$

with the same error, $S(\mathbf{w}^{*})^{(i)} = 23.8951, \quad i = 1,2, \ldots,5.$

Since each of the above stationary vectors, $\mathbf{w}^{*(i)}$, i=1,2, ...,5, directly gives the first rows of their corresponding "best" approximating (rank one) transitive matrices, thus these approximation matrices, $\mathbf{B}_1^{(i)}$, i=1,2, ...,5, to $\mathbf{A}_1$ could now be easily constructed.

EXAMPLE 2. The second example refers to a 6×6 PCM whose rows and columns have been rearranged appropriately. This response matrix $\tilde{\mathbf{A}}_2$ is specified as

$$\tilde{\mathbf{A}}_2 = \begin{bmatrix} 1 & 9 & 1/9 & \vdots & 5 & 4 & 1/3 \\ 1/9 & 1 & 9 & \vdots & 1/3 & 5 & 4 \\ 9 & 1/9 & 1 & \vdots & 4 & 1/3 & 5 \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots \\ 1/5 & 3 & 1/4 & \vdots & 1 & 7 & 1/7 \\ 1/4 & 1/5 & 3 & \vdots & 1/7 & 1 & 7 \\ 3 & 1/4 & 1/5 & \vdots & 7 & 1/7 & 1 \end{bmatrix}.$$

Observe that $\tilde{\mathbf{A}}_2$ contains two 3×3 circulant SR block matrices along its main diagonal. Note that $\tilde{\mathbf{A}}_2$ consists of a circulant off-diagonal block of size 3×3 as well. Using an appropriate permutation matrix, $\mathbf{P}_n$, consisting of two circulant permutation matrices along its main diagonal the following linearly independent solutions (three alternate minima) are obtained:

$$\mathbf{w}^{*\mathrm{T}(1)} = [\,1\,,5.5572\,,1.9022\,,5.8397\,,3.9612\,,2.5166\,],$$

$$\mathbf{w}^{*\mathrm{T}(2)} = [\,1\,,0.5257\,,2.9215\,,1.3230\,,3.0700\,,2.0825\,],$$

$$\mathbf{w}^{*\mathrm{T}(3)} = [\,1\,,0.3423\,,0.1799\,,0.7128\,,0.4529\,,1.0508\,],$$

with the same error, $S(\mathbf{w}^{*})^{(i)} = 18.7968, \quad i = 1,2,3.$

As for EXAMPLE 1, the "best" approximating transitive matrices, $\mathbf{B}_2^{(i)}$, i=1,2,3, to $\tilde{\mathbf{A}}_2$ could readily be constructed. EXAMPLE 2 demonstrates the occurrence of a multiple solution for an SR matrix $\tilde{\mathbf{A}}_2$ which is neither circulant nor persymmetric. The reason that even in such a case alternate stationary vectors may occur is attributed to the SPI property (see Definition 3) of the SR matrix $\tilde{\mathbf{A}}_2$, which could be permuted by elementary circulants since it consists of cyclic groups.

EXAMPLE 3. The third example contains data of a 7×7 PCM whose rows and columns have been rearranged appropriately. This response matrix $\tilde{\mathbf{A}}_3$ is specified as

$$\tilde{\mathbf{A}}_3 = \left[\begin{array}{ccccccccc}
1 & 9 & 1/9 & \vdots & 9 & \vdots & 1 & 9 & 1/9 \\
1/9 & 1 & 9 & \vdots & 9 & \vdots & 1/9 & 1 & 9 \\
9 & 1/9 & 1 & \vdots & 9 & \vdots & 9 & 1/9 & 1 \\
\cdots & \cdots & \cdots & \vdots & \cdots & \vdots & \cdots & \cdots & \cdots \\
1/9 & 1/9 & 1/9 & \vdots & 1 & \vdots & 9 & 9 & 9 \\
\cdots & \cdots & \cdots & \vdots & \cdots & \vdots & \cdots & \cdots & \cdots \\
1 & 9 & 1/9 & \vdots & 1/9 & \vdots & 1 & 9 & 1/9 \\
1/9 & 1 & 9 & \vdots & 1/9 & \vdots & 1/9 & 1 & 9 \\
9 & 1/9 & 1 & \vdots & 1/9 & \vdots & 9 & 1/9 & 1
\end{array}\right].$$

Observe that $\tilde{\mathbf{A}}_3$ consists of two 3×3 circulant block matrices along its main diagonal and a single element on its midpoint. Note that $\tilde{\mathbf{A}}_3$ has a 3×3 circulant off-diagonal block with the same entries as those of the block matrices along the main diagonal. Using an appropriate permutation matrix, $\mathbf{P}_n$, consisting of two circulant permutation matrices on its main diagonal and the unity at the midpoint (in other words, there are three cyclic groups here: two of size three and one fix-point), the following linearly independent solutions (six alternate minima) are obtained:

$\mathbf{w}^{*T(1)} = [\,1\,, 2.6729\,, 2.1324\,, 0.8259\,, 1.7821\,, 7.6524\,, 4.9210\,],$

$\mathbf{w}^{*T(2)} = [\,1\,, 0.4689\,, 1.2535\,, 0.3873\,, 2.3077\,, 0.8357\,, 3.5886\,],$

$\mathbf{w}^{*T(3)} = [\,1\,, 0.7978\,, 0.3741\,, 0.3090\,, 2.8629\,, 1.8410\,, 0.6667\,],$

$\mathbf{w}^{*T(4)} = [\,1\,, 0.6431\,, 2.7613\,, 5.9584\,, 2.3077\,, 1.8410\,, 4.9210\,],$

$\mathbf{w}^{*T(5)} = [\,1\,, 4.2940\,, 1.5551\,, 9.2657\,, 2.8629\,, 7.6524\,, 3.5886\,],$

$\mathbf{w}^{*T(6)} = [\,1\,, 0.3621\,, 0.2329\,, 2.1578\,, 1.7821\,, 0.8357\,, 0.6667\,],$

with the same error, $S(\mathbf{w}^*)^{(i)} = 32.0030, \quad i = 1, \ldots, 6.$

Similarly to the previous examples the "best" approximating transitive matrices, $\mathbf{B}_3^{(i)}$, i=1, ...,6, to $\tilde{\mathbf{A}}_3$ may be constructed easily.

EXAMPLE 4. The fourth example shows data of a 8×8 PCM whose rows and columns have been rearranged appropriately. This response matrix $\tilde{\mathbf{A}}_4$ is specified as

$$\tilde{\mathbf{A}}_4 = \begin{bmatrix} 1 & 9 & 1/9 & \vdots & 9 & 9 & 9 & 9 & 9 \\ 1/9 & 1 & 9 & \vdots & 9 & 9 & 9 & 9 & 9 \\ 9 & 1/9 & 1 & \vdots & 9 & 9 & 9 & 9 & 9 \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1/9 & 1/9 & 1/9 & \vdots & 1 & 9 & 1/9 & 9 & 1/9 \\ 1/9 & 1/9 & 1/9 & \vdots & 1/9 & 1 & 9 & 1/9 & 9 \\ 1/9 & 1/9 & 1/9 & \vdots & 9 & 1/9 & 1 & 9 & 1/9 \\ 1/9 & 1/9 & 1/9 & \vdots & 1/9 & 9 & 1/9 & 1 & 9 \\ 1/9 & 1/9 & 1/9 & \vdots & 9 & 1/9 & 9 & 1/9 & 1 \end{bmatrix}.$$

Observe here that $\tilde{\mathbf{A}}_4$ contains two circulant block matrices along its main diagonal of size 3×3 and 5×5, respectively. Note that $\tilde{\mathbf{A}}_4$ consists of a rectangular off-diagonal block of size 3×5 with identical entries. It represents a "trivial"circulant matrix. By applying an appropriate permutation matrix, $\mathbf{P}_n$, which consists of two circulant permutation matrices along its main diagonal, then the following linearly independent solutions (fifteen alternate minima) are obtained:

$$\mathbf{w}^{*\mathrm{T}(1)} = [\,1\,,1.1943\,,1.0028\,,3.3427,2.0231\,,5.9230\,,5.3371\,,5.9300],$$

$$\mathbf{w}^{*\mathrm{T}(2)} = [\,1\,,0.9972,1.1910,2.0175,5.9066,5.3224,5.9136,3.3334],$$

$$\mathbf{w}^{*\mathrm{T}(3)} = [\,1\,,0.8396,0.8373,4.9593,4.4687,4.9651,2.7988,1.6939],$$

$$\vdots$$

$$\mathbf{w}^{*\mathrm{T}(15)} = [\,1,0.8396,0.8373,4.9651,2.7988,1.6939,4.9593,4.4687],$$

with the same error, $S(\mathbf{w}^*)^{(i)}$ = 29.3584, i = 1, ...,15.

Now, the "best" approximating transitive matrices, $\mathbf{B}_4^{(i)}$, i=1, ...,15, to $\tilde{\mathbf{A}}_4$ may be constructed easily.

EXAMPLE 5. The fifth example exhibits data of a 5×5 PCM. Since **A** is in SR by definition, for sake of simplicity the entries below the main diagonal are not depicted. This response matrix $\mathbf{A}_5$ is specified as

$$\mathbf{A}_5 = \begin{bmatrix} 1 & 9 & 8 & 5 & 1/9 \\ & 1 & 3 & 2 & 5 \\ & & 1 & 3 & 8 \\ & & & 1 & 9 \\ & & & & 1 \end{bmatrix}.$$

Observe that $\mathbf{A}_5$ is a persymmetric SR matrix. Applying Proposition 2, the following linearly independent solutions (two alternate minima) are obtained:

$$\mathbf{w}^{*\mathrm{T}(1)} = [\,1\,,\,0.9988\,,\,0.6409\,,\,0.5834\,,\,4.4176\,],$$

$$\mathbf{w}^{*\mathrm{T}(2)} = [\,1\,,\,7.5719\,,\,6.8926\,,\,4.4229\,,\,4.4176],$$

with the same error, $S(\mathbf{w}^*)^{(i)}$ = 16.0449, i = 1,2.

The "best" approximating transitive matrices, $\mathbf{B}_5^{(i)}$, i=1,2, to $\mathbf{A}_5$ are:

$$\mathbf{B}_5^{(1)} = \begin{bmatrix} 1 & 0.9988 & 0.6409 & 0.5834 & 4.4176 \\ & 1 & 0.6417 & 0.5841 & 4.4229 \\ & & 1 & 0.9103 & 6.8926 \\ & & & 1 & 7.5719 \\ & & & & 1 \end{bmatrix},$$

and

$$\mathbf{B}_5^{(2)} = \begin{bmatrix} 1 & 7.5719 & 6.8926 & 4.4229 & 4.4176 \\ & 1 & 0.9103 & 0.5841 & 0.5834 \\ & & 1 & 0.6417 & 0.6409 \\ & & & 1 & 0.9988 \\ & & & & 1 \end{bmatrix}.$$

Observe here that neither $\mathbf{B}_5^{(1)}$ nor $\mathbf{B}_5^{(2)}$ is a persymmetric matrix. The two independent solutions, $\mathbf{w}^{*(1)}$ and $\mathbf{w}^{*(2)}$, are in the first rows and they also appear in the last columns of $\mathbf{B}_5^{(1)}$ and $\mathbf{B}_5^{(2)}$ in opposite order.

EXAMPLE 6. The sixth example shows data of a 6×6 PCM. Since **A** is in SR by definition, for sake of simplicity the entries below the main diagonal are not depicted. This response matrix $\mathbf{A}_6$ is specified as

$$\mathbf{A}_6 = \begin{bmatrix} 1 & 9 & 8 & 7 & 6 & 1/9 \\ & 1 & 1 & 3 & 9 & 6 \\ & & 1 & 1 & 3 & 7 \\ & & & 1 & 1 & 8 \\ & & & & 1 & 9 \\ & & & & & 1 \end{bmatrix}.$$

Observe that $\mathbf{A}_6$ is a persymmetric SR matrix. By applying Corollary 3, the following solution (a local minimum) is obtained:

$$\mathbf{w}^{*\mathrm{T}} = [\,1\,, 0.8753\,, 1.7818\,, 3.0458\,, 6.2006\,, 5.4271\,],$$

with the error, $S(\mathbf{w}^*) = 18.8874$.

The “best” approximating transitive matrix, $\mathbf{B}_6$, to $\mathbf{A}_6$ is:

$$\mathbf{B}_6 = \begin{bmatrix} 1 & 0.8753 & 1.7818 & 3.0458 & 6.2006 & 5.4271 \\ & 1 & 2.0358 & 3.4799 & 7.0843 & 6.2006 \\ & & 1 & 1.7094 & 3.4799 & 3.0458 \\ & & & 1 & 2.0358 & 1.7818 \\ & & & & 1 & 0.8753 \\ & & & & & 1 \end{bmatrix}.$$

Observe here that now $\mathbf{B}_6$ is a persymmetric matrix, i.e. its entries are symmetric about its counterdiagonal (secondary diagonal). Therefore, in such a case, no trivial alternate stationary vector can be found. It is easy to check that each of the conditions (35) for the elements of the stationary vector (which are in the first row *and* in opposite order in the last column of matrix $\mathbf{B}_6$) holds.

EXAMPLE 7. The authors carried out a comprehensive analysis for a large set of different 3×3 SR matrices $\mathbf{A}$. Although in the applications of the AHP these matrices represent the simplest cases only, yet they seem to be adequate to show us a certain tendency of the occurrences of non-unique solutions to the nonlinear LS optimization problem (3). For this purpose we utilize some of our results presented in Section 3. Let these response matrices $\mathbf{A}_7$ be given in the form

$$\mathbf{A}_7 = \begin{bmatrix} 1 & a & b \\ 1/a & 1 & a \\ 1/b & 1/a & 1 \end{bmatrix}.$$

Observe that $\mathbf{A}_7$ is persymmetric. Here the entries a correspond to a_{12}=a_{23} and entry b corresponds to a_{13}. Using Saaty’s nine-point scale for a great number of appropriately chosen 3×3 PCMs the multiple stationary vectors (global minima) have been determined and are displayed in Figure 1 over the entire range of the possible values of these entries (recall that the respective entry in $\mathbf{A}$ expresses the relative strength with which decision alternative A_i dominates alternative A_j). Here a particular solution represents, in fact, a *global* minimum. Namely, by the NK method, applying a heuristics approach *all* solutions were generated and examined in a numerical way for each interval by using “Mathematica”. Consider now Figure 1. Here the selected scale-values of the entries, a_{12}=[1, 2, ... ,9] (or for a_{12}=[1/9,1/8, ... ,1]) are plotted as function of the NE corner entry, a_{13}. The locations of the black dots indicate the numerical values with which a 3×3 PCM, $\mathbf{A}_7$ becomes a circulant matrix (here at these

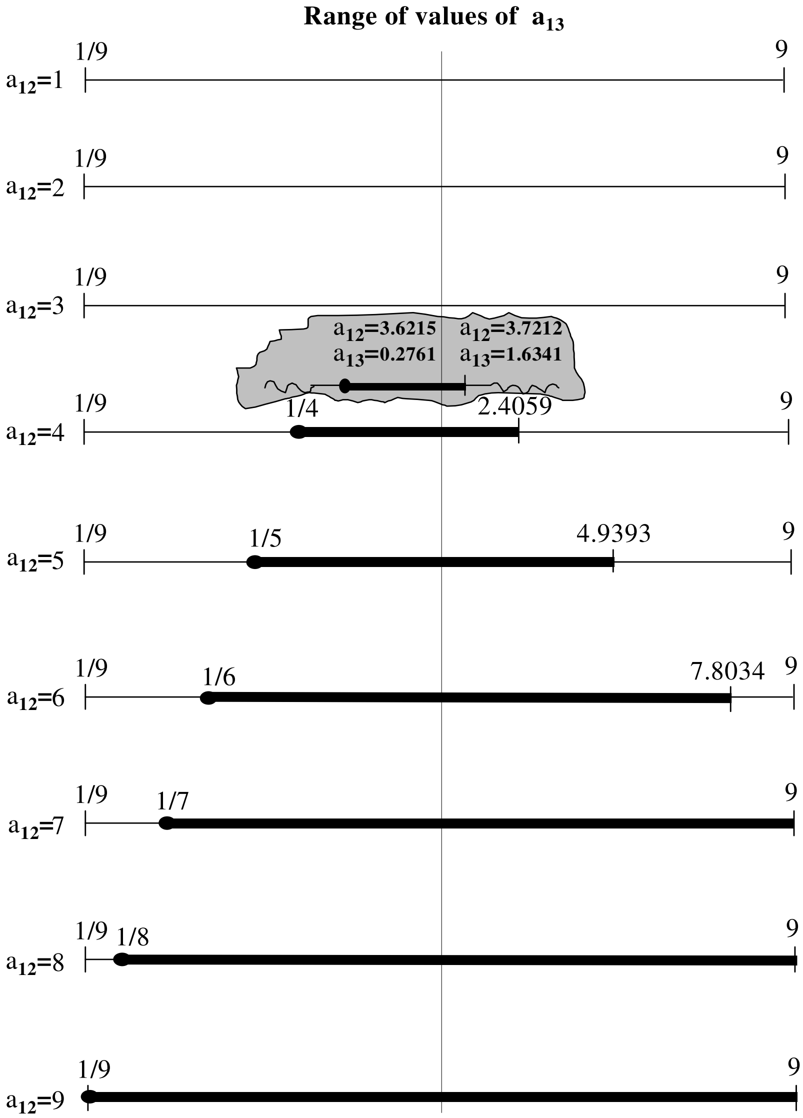


Figure 1. Domains of multiple optima for 3x3 SR matrices **A** (PCMs)

points there are three alternate optima). The intervals drawn by heavy lines indicate the regions over which linearly independent solutions occur (here, there are two alternate optima). The regions drawn by solid lines indicate the intervals over which there is one solution (a global minimum). Bozóki [16] used the resultant method for analyzing the non-uniqueness problem of 3×3 SR matrices. It is interesting to note that on applying our method to EXAMPLE 7 exactly the same results are obtained; giving confidence in the appropriateness of both approaches.

Figure 1 exhibits a remarkable tendency concerning the likelihood of a multiple solution. Note that with a growing level of inconsistency (as the entries a_{12}=a_{23} are increased relative to the entry a_{13}) the range of values over which a multiple solution occurs will be greater and greater. Note that for a_{12}=9, only multiple solution occurs within the whole possible range of the values of entry a_{13}. One may recognize from this chart that initially (i.e. at low levels of inconsistency of the matrix $\mathbf{A}_7$) the optimal solution (global minimum) is unique. It is interesting to note that up to a turning point the solution yields $\mathbf{w}^{*T}$=[1,1,1] evaluated at the entries of a and b=1/a. If, however, the entries of $\mathbf{A}_7$ are increased to: a_{12}=a_{23}=3.6215 and a_{13}=1/a_{12}=1/a_{23}= 0.2761, then three other independent stationary vectors achieve the minimum in (3) also (thus at this intersection point there are four alternate optima). For $\mathbf{A}_7$, the numerical values of the entries a_{12}=a_{23} and a_{13} when a unique solution switches to a multiple solution, or reversed, when it switches back to a unique one, can be determined explicitly (see in the Appendix the formulation of the system of nonlinear equations which considers a more general case than is discussed by EXAMPLE 7).

Conclusions

A system of nonlinear equations has been used to determine the entries of a transitive matrix which is the best approximation to a general pairwise comparison matrix in a least-squares sense. The nonlinear minimization problem as the solution of a set of inhomogeneous equations has been examined for its uniqueness properties. Sufficient conditions for possible non-uniqueness of the solution to this optimization problem have been developed and the related proofs have also been presented. For a great number of different sized positive SR matrices having certain properties, the results have been demonstrated by the numerical experiments. Further research will include the investigation for finding the necessary conditions for the non-uniqueness of the solution to the nonlinear optimization problem.

Acknowledgement

This research was partially supported by the Hungarian National Foundation for Scientific Research, Grant No.: OTKA T 043034.

Appendix

Suppose that the positive $n \times n$ SR matrix **A** is specified as

$$\mathbf{A} = \begin{bmatrix} 1 & a & a & a & . & . & . & b \\ & 1 & 1 & 1 & . & . & . & a \\ & & 1 & 1 & . & . & . & a \\ & & & 1 & . & . & . & a \\ & & & & . & & & . \\ & & & & & . & & . \\ & & & & & & . & . \\ & & & & & & & 1 \end{bmatrix}.$$

Let the entries of **A** be denoted by $a=a_{1j}=a_{jn}$, $j=2,...,n-1$, and $b=a_{1n}$, the linearly independent solutions to equation (9) given in Proposition 2, by $\mathbf{w}^{*(1)}$ and $\mathbf{w}^{*(2)}$, respectively, and the solution (34) by $\mathbf{v}^*$. At a stationary point of $S^2(\mathbf{w})$ where a multiple optima occurs the elements of the unknown vectors $\mathbf{w}^{*(1)}$, $\mathbf{w}^{*(2)}$, $\mathbf{v}^*$ and the entries of **A** can be determined by solving the following constrained nonlinear optimization problem [for a particular problem, the size of matrix **A** has to be properly adjusted, the weights v_i and v_k should be inserted in the equations according to relation (35) and recall that $\mathbf{w}^{*(2)}$ can be obtained from (29)]:

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\left(a_{ik}-\frac{w_k}{w_i}\right)^2 - \sum_{i=1}^{n}\sum_{k=1}^{n}\left(a_{ik}-\frac{v_k}{v_i}\right)^2 = 0,$$

$$\sum_{k=1}^{n}\left[\frac{a_{ik}}{w_i^2}+\frac{1}{a_{ik}w_k^2}-2\left(\frac{w_i}{w_k^3}+\frac{w_k}{w_i^3}\right)\right]w_k = 0, \qquad i=2,\ ...,n,$$

$$\sum_{k=1}^{n}\left[\frac{a_{ik}}{v_i^2}+\frac{1}{a_{ik}v_k^2}-2\left(\frac{v_i}{v_k^3}+\frac{v_k}{v_i^3}\right)\right]v_k = 0, \qquad i=2,\ ...,n,$$

$$\mathbf{c}^{\mathrm{T}}\mathbf{w} = 1, \qquad \text{and} \qquad \mathbf{c}^{\mathrm{T}}\mathbf{v} = 1,$$

where $\mathbf{w}^{\mathrm{T}}:=\{w_1,w_2, ...,w_n\}$, $\mathbf{v}^{\mathrm{T}}:=\{v_1,v_2, ...,v_n\}$, and $\mathbf{c}^{\mathrm{T}}:=\{1,0,0, ...,0\}$.

References

[1] Saaty,T.L., “A scaling method for priorities in hierarchical structures”, Journal of Mathematical Psychology”, 15 (1977) 234-281.

[2] Farkas,A. Rózsa,P. and Stubnya,E. ”Transitive matrices and their applications” Linear Algebra and its Applications, 302-303 (1999) 423-433.

[3] Saaty,T.L., “Axiomatic foundation of the analytic hierarchy process”, Management Science, 32 (1986) 841-855.

[4] Jensen,R., “An alternative scaling method for priorities in hierarchical structures”, Journal of Mathematical Psychology, 28 (1984) 317-332.

[5] Chu,A.T.W., Kalaba,R.E. and Springarn,K.,“A comparison of two methods for determining the weights of belonging to fuzzy sets”, Journal of Optimization Theory and Applications, 27 (1979) 531-538.

[6] Blankmeyer,E,“Approaches to consistency adjustment” Journal of Optimization Theory and Applications”, 54 (1987) 479-488.

[7] Crawford,G., “The geometric mean procedure for estimating the scale of a judgment matrix”, Mathematical Modeling, 9 (1987) 327-334.

[8] Genest,C. and Rivest,L.P., “A statistical look at Saaty’s method of estimating pairwise preferences expressed on a ratio scale”, Journal of Mathematical Psychology, 38 (1994) 477-496.

[9] Cook,W.D. and Kress,M., “Deriving weights from pairwise comparison ratio matrices: An axiomatic approach”, European Journal of Operations Research, 37 (1988) 355-362.

[10] Genest,C. and Zhang,S.,”A graphical analysis of ratio-scaled paired comparison data”, Management Science, 42 (1996) 335-349.

[11] Golany,B. and Kress,M., “A multicriteria evaluation of methods for obtaining weights from ratio-scale matrices”, European Journal of Operations Research, 69 (1993) 210-220.

[12] Farkas,A., Lancaster,P. and Rózsa,P.,”Consistency adjustments for pairwise comparison matrices”, Numerical Linear Algebra with Applications, 10 (2003) 689-700.

[13] Farkas,A., Lancaster,P. and Rózsa,P.,”On approximation of positive matrices by transitive matrices”, Computers & Mathematics, (to appear).

[14] Lee,A., “Über permutationsinvariante Matrizen”, Publicationes Mathematicae, 11. Institutom Matematicum Universitatis Debreceniensis, Hungaria, (1964) 44-58.

[15] Davis,P.J., “Circulant Matrices”, John Wiley, New York, 1979.

[16] Bozóki,S., “A method for solving LSM problems of small size in the AHP”, Central European Journal of Operations Research, 11 (2003) 17-33.

Simulation Based Verification of the Applicability of a Novel Branch of Computational Cybernetics in the Adaptive Control of Imperfectly Modeled Physical Systems of Asymmetric Delay Time and Strong Non-linearities

József K. Tar, Imre J. Rudas, János F. Bitó

Budapest Polytechnic, John von Neumann Faculty of Informatics, Centre of Robotics and Automation, 1081 Budapest, Népszínház utca 8.
E-mail: tar@nik.bmf.hu, rudas@bmf.hu, bito@nik.bmf.hu

Abstract:
In this paper the applicability of an adaptive control based on a novel branch of Computational Cybernetics is illustrated for two different, imperfectly and inaccurately modeled particular physical sytems. One of them is a water tank stirring cold and hot water as input and releasing the mixture through a long pipe. The mass flow rate and the temperature are prescribed at the free end of the exit pipe while the taps at the input side can diretly be controlled. Due to the incompressibility of the fluid the variation of the mass flow rate of the output is immediately observableat the pipe's end and is related to the control action at the input taps, while its effect on the temperature becomes measurable at the free end of the pipe only after a delay time needed for the fluid to flow through the pipe. This results in asymmetric and non-constant delay time. The other paradigm is the thermal decay of the molecular nitrogen during a throttling down process. As is well known chemical reactions hav very drastic non-linearities and it is not easy to construct their "exact" or satisfacorily avccurate model. The fundamental principles of this new branch of Computational Cybernetics are briefly presented in the paper. To some extent it is similar to the traditional Soft Computing, but by using a priori known, uniform, lucid structure of reduced size, it can evade the enormous structures so characteristic to the usual approach. Clumsy deterministic, semi-stochastic or stochastic machine learning is replaced by simple, short, explicit algebraic procedures especially fit to real time applications. The costs of these advantages may manifest themselves in the expected limitation of the applicabilityof this new approach. However, the simulation results exemplify the applicability of the new method in the control of systems of strong non-linearities and asymmetric delay time.

1 Introduction

A new approach for the adaptive control of imprecisely known dynamic systems under unmodeled dynamic interaction with their environment was initiated in [1]. In the family of the adaptive control methods this new one lays between the linear PID/ST and the parameter identification approaches.

Instead of the supposed analytical model's parameters the control is tuned as in the PID/ST, but it offers the possibility of using several parameters of some abstract Lie groups fit to the needs of the „non-linear control". In the same time these parameters may be considered as that of the system model's, though they are not the part of its detailed analytical description. This „non-analytical modeling" is akin to the Soft Computing philosophy.

In this approach adaptivity means that instead of the simultaneous tuning of numerous parameters, a fast algorithm finding some linear transformation to map a very primitive initial model based expected system-behavior to the observed one is used. The so obtained „amended model" is step by step updated to trace changes by repeating this corrective mapping in each control cycle. Since no any effort is exerted to identify the possible reasons of the difference between the expected and the observed system response, it is referred to as the idea of "Situation-Dependent Partial System Identification". This anticipates the possibility for real-time applications.

Regarding the appropriate linear transformations several possibilities were investigated and successfully applied. E.g. the „Generalized Lorentz Group" [2], the „Stretched Orthogonal Group", the "Partially Stretched Orthogonal Transformations" [3], and a special family of the „Symplectic Transformations" [4] can be mentioned.

The key element of the new approach is the formal use of the „Modified Renormalization Transformation". The „original" transformation was widely used e.g. by Feigenbaum in the seventies to investigate the properties of chaos [5-7]. Its useful property from our special point of view is that this (originally scalar) transformation modifies the solution of an x=*f*(x) fixed-point problem, since the adaptive control was formulated as a fixed-point problem, too [8]. The modification of the original transformation was necessary due to phenomenological reasons. Satisfactory conditions of the complete stability of the so obtained control for Multiple Input-Multiple Output (MIMO) systems were also highlighted in [8] by the means of perturbation calculation. This means the most rigorous limitation regarding the circle of possible application of the new method. To release this restriction to some extent "ancillary" but simple interpolation techniques and application of "dummy parameters" were also introduced in [8].

The applicability of the method was investigated for electro-mechanical and hydrodynamic systems via simulation [9-10]. These systems were exempt of any kind of delay or lag. In this paper a quite simple but lucid typical non-linear paradigm, a water tank of open outlet is chosen to be the subject of the new type adaptive controller. It contains continuous non-linearities due to the velocity-dependent resistance of the pipelines, saturated (bounded) non-linearities set by the temperature of the „warm" and the „cold" input water to be mixed in the tank, and the open input of the tank making it impossible for the fluid to flow back in the input pipes. Further non-linear limitation is that the velocity of the flow leaving the tank is unique function of the density and full mass of the fluid exiting the tank, so it cannot be directly controlled: only the mass flow rate of the cold and warm input is controllable. Furthermore, since the mass flow rate and the temperature of the required output is defined and measured only at the end of the pipe serving as the outlet, while the input is directly controllable at the location of the tank, the temperature signal contains considerable lag. (Due to the incompressibility of the liquid the velocity signal of the flow doesn't suffer from considerable delay.)

In the sequel at first the basic principles of the adaptive control are described, then the models and the simulation results for the particular paradigms considered are given. Following the presentation of the typical simulation results the conclusions are drawn.

2 The basic principles of the adaptive control

From purely mathematical point of view the control task can be formulated as follows. There is given some imperfect model of the system on the basis of which some excitation is calculated to obtain a desired system response $\mathbf{i}^d$ as $\mathbf{e}=\boldsymbol{\varphi}(\mathbf{i}^d)$. The system has its inverse dynamics described by the unknown function $\mathbf{i}^r=\boldsymbol{\psi}(\boldsymbol{\varphi}(\mathbf{i}^d))=\boldsymbol{f}(\mathbf{i}^d)$ and resulting in a realized response $\mathbf{i}^r$ instead of the desired one, $\mathbf{i}^d$. Normally one can obtain information via observation only on the function $\boldsymbol{f}()$ considerably varying in time, and no any possibility exists for directly "manipulaing" the nature of this function: only $\mathbf{i}^d$ as the input of $\boldsymbol{f}()$ can be "deformed" to $\mathbf{i}^{d*}$ to achieve and maintain the $\mathbf{i}^d=\boldsymbol{f}(\mathbf{i}^{d*})$ state. [Only the *model function* $\boldsymbol{\varphi}()$ can directly be manipulated.] On the basis of the modification of the method of renormalization widely applied in Physics the following "scaling iteration" was suggested for finding the proper deformation:

$$\begin{gathered}\mathbf{i}_0;\mathbf{S}_1\mathbf{f}(\mathbf{i}_0)=\mathbf{i}_0;\mathbf{i}_1=\mathbf{S}_1\mathbf{i}_0;\ldots;\mathbf{S}_\mathrm{n}\mathbf{f}(\mathbf{i}_{n-1})=\mathbf{i}_0;\\ \mathbf{i}_{n+1}=\mathbf{S}_{n+1}\mathbf{i}_n;\mathbf{S}_\mathrm{n}\xrightarrow[n\to\infty]{}\mathbf{I}\end{gathered}\tag{1}$$

in which the $\mathbf{S}_n$ matrices denote some linear transformations to be specified later. As it can be seen these matrices maps the observed response to the desired one, and the construction of each matrix corresponds to a step in the adaptive control. It is evident that if this series converges to the identity operator just the proper deformation is approached, therefore the controller „learns" the behavior of the observed system by step-by-step amendment and maintenance of the initial model. (The response arrays may contain a „dummy", that is physically not interpreted dimension of constant value, in order to evade the occurrence of the mathematically dubious 0→0, 0→finite, finite→0 cases.)

Since (1) does not unambiguously determine the possibly applicable quadratic matrices, we have additional freedom in choosing appropriate ones. The most important points are fast and efficient computation, and the ability for remaining as close to the identity transformation as possible. In the present paper an orthogonal transformation is created which transforms the realized vector into a vector parallel with the desired one while leaves the orthogonal sub-space of these two vectors unchanged. Then proper stretching/shrinking factor is calculated which makes the absolute value of the realized vector equal to that of the desired one. On this basis two linear operators are created which apply the appropriate stretches/shrinks in the "realized" one-dimensional sub-spaces, rotate them to be parallel to the "desired" directions, and leave the orthogonal sub-spaces unchanged [3]. This operation evidently equals to the identity operator if the desired response just is equal to the desired one, and remains in the close vicinity of the unit matrix if the non-zero desired and realized responses are very close to each other. In the application of the above method it was implicitly supposed that practically the „desired" and the „observed" responses were simultaneously observable/available.

3 Description of the water tank

The water tank considered is an open vessel into which hot and cold water of fixed temperatures T_1=10 °C, and T_2=90 °C is purred from the top. The mass flow rates of the input components $\dot{M}_1, \dot{M}_2$ [kg/s] are directly controllable via electric valves. According to [11] the density of the water in the above temperature range is 999.7 kg/m^3 within 3.4 % precision, so it is approximated with the mean value over this interval as ρ=982.48 kg/m^3 as a constant. The cross-sectional area of the tank is A=1 m^2, and it is supposed to be high enough to contain all the amount of the liquid occurring in the calculations. At the bottom level of the tank a pipe of diameter D=1.8×10^{-1} m, length L=10 m, and relative internal surface roughness of k_{rel}=1.5×10^{-2} is attached. The pressure increase with respect to the environmental pressure, that is the actual pressure difference driving the water flow in the pipe is Δp=$M(t)g/A$ Pa if g=9.87 m/s^2 is the gravitational acceleration, and $M(t)$ in kg units denotes the actual mass of the fluid in the tank. By neglecting the minor pressure losses at the exit at the tank and the free end of the pipe, the velocity of the flow in the pipe, u is determined by the equation

$$\left(\frac{D}{4L}\right)\frac{\Delta p}{0.5\times\rho u^2}=f\left(\frac{\rho uD}{\mu},k_{rel}\right) \tag{2}$$

in which *f* is the non-dimensional *friction factor*, and μ denotes the dynamic viscosity of the fluid. The viscosity mainly depends on the fluid temperature, and in the given range it varies within the range of $[3.11\times10^{-4}, 1.3\times10^{-3}]$ *kg/*(*m*×*s*). The non-dimensional expression $Re:=\rho uD/\mu$ defines the *Reynolds Number*. The $f(Re,k_{rel})$ function is given in the well-known *Moody Diagram* [12]. At the given numerical value of k_{rel} *f* practically is constant (1.21×10^{-2}) if *Re* is greater than 10^{-5}. Allowing M_{min}=100 *kg* minimum mass of water in the tank and supposing that f=1.21×10^{-2} (1) yields the minimum seeped of water flow as u_{min}=0.86 *m/s* to which the $Re\cong1.16\times10^5$ values belongs if the maximum value of the viscosity in the given range is taken into account. Therefore, if the mass of the fluid in the tank remains over 100 *kg*, the flow in the pipe will be fully turbulent with a constant f=1.21×10^{-2} friction factor. For the given pipe length a delay time of about a few seconds can be expected for the temperature signal.

Regarding the mixing of the cold and warm water, the heat capacity of the fluid mainly depends on the temperature and varies in the interval [4.193, 4.208] *kJ*/(*kg*×°*K*), that is it can also be considered to be constant.

Under the above conditions the operation of the tank can be approximated by the following differential equations:

$$\dot{T}=\frac{T_1-T}{M}\dot{M}_1+\frac{T_2-T}{M}\dot{M}_2 \tag{3}$$

$$\dot{M}_3=\frac{\pi D^2\rho}{4}\sqrt{\frac{2gM}{AKf\rho}},\quad K=\frac{4L}{D} \tag{4}$$

in which *T* denotes the temperature of the mixed fluid in the tank, and $\dot{M}_3$ means the mass flow rate at the output. While $\dot{T}$ can directly be controlled by the valves at the input, the output mass flow rate cannot. This gives the system a kind of „inertia". Only the time-derivative of the output mass flow rate can be directly controlled due to the conservation of the mass of the fluid as

$$\ddot{M}_3=\frac{\pi D^2}{4}\sqrt{\frac{2g\rho}{4AKfM}}\dot{M} \tag{5}$$

$$\dot{M}=\dot{M}_1+\dot{M}_2-\dot{M}_3 \tag{6}$$

For the directly controllable quantities therefore the following pair of equations is obtained:

$$\begin{aligned}\dot{T} &= \frac{T_1 - T}{M}\dot{M}_1 + \frac{T_2 - T}{M}\dot{M}_2 \\ \ddot{M}_3 &= \frac{\pi D^2}{8}\sqrt{\frac{2g\rho}{AKfM}}\left(\dot{M}_1 + \dot{M}_2\right) - \frac{\pi^2 D^4}{32}\frac{2g}{AKf}\end{aligned} \tag{7}$$

in the integration of which (4) and (6) can also be used. Regarding the problem of the delay of observation, the quantities in (4-7) are to be taken in common time instant if they are measured/observed immediately at the tank. However, if the temperature is measured at the outlet of the pipe, one has to distinguish between the actual values in the tank and in the outlet. It can be stated, that if *t* is the time of the observation, and the input valves are controlled by fast electronic signals, than

$$T^{Obs}(t) = T^{Tank}(t - \delta(t)) \tag{8}$$

n which the lag $\delta(t)$ is determined by the equation

$$L = \int_{t-\delta(t)}^{t} u(\tau)d\tau \tag{9}$$

Due to the incompressibility of the liquid and the fast electric signals the mass flow rates are immediately observable and no such distinction has to be done.Principles of the adaptive control

However, in the case of the present paradigm the effect of the control action immediately can be observed on the output mass flow rate, but its observation suffers from a lag $\delta(t)$ as far as temperature is concerned. This „asymmetry" is tackled in the control in the following way. If a P-type controller is applied, an exponentially asymptotic trajectory reproduction is prescribed by defining certain „desired" time-derivatives in the following manner:

$$\begin{bmatrix} \ddot{M}_3^D(t) \\ \dot{T}^D(t) \end{bmatrix} = \begin{bmatrix} \ddot{M}_3^N(t) \\ \dot{T}^N(t) \end{bmatrix} + \alpha \begin{bmatrix} \dot{M}_3^N(t) - \dot{M}_3^R(t) \\ T^N(t) - T^R(t) \end{bmatrix} \tag{10}$$

where the indices *D, N,* and *R* refer to the „desired", „nominal", and the „realized" (actual) values, and α controls the speed of the desired error-relaxation. In the adaptive version, in the lack of any time lag, the matrices in (10) were constructed from the pair

$$\mathbf{S}\begin{bmatrix} \ddot{M}_3^R(t) \\ \dot{T}^R(t) \\ C \end{bmatrix} = \begin{bmatrix} \ddot{M}_3^D(t) \\ \dot{T}^D(t) \\ C \end{bmatrix} \tag{11}$$

where C denotes the „dummy” parameter introduced due to pure technical reasons only. In the „asymmetric” case, if t measures the time at the outlet of the pipe the error term fed back in (11) can be replaced by

$$\begin{bmatrix} \dot{M}_3^N(t)-\dot{M}_3^R(t) \\ T^N(t-\delta(t))-T^R(t-\delta(t)) \end{bmatrix} \tag{12}$$

expressing the fact that the actual response observable at the end of the pipe at time „t” can be related to a control action based on a desired derivative computed previously at t-$\delta(t)$, since the observed values at t correspond to the available „freshest” information on that control action. On the same basis, the **S** matrices of the adaptive law at time t are calculated from the pair of vectors

$$\begin{bmatrix} \ddot{M}_3^R(t) \\ \dot{T}^R(t) \\ C \end{bmatrix}, \begin{bmatrix} \ddot{M}_3^N(t)+\alpha\left[\dot{M}_3^N(t)-\dot{M}_3^R(t)\right] \\ \dot{T}^N(t-\delta)+\alpha\left[T^N(t-\delta)-T^R(t-\delta)\right] \\ C \end{bmatrix} \tag{13}$$

In similar way, if instead of a P-type a PI-type control becomes necessary to calculate the desired derivatives in the linear control approach, it is reasonable to compute the integrated error with the same delay as above, and the adaptive matrices have to be computed from the so obtained counterpart of (12).

Since amongst the conditions for which the convergence of the method was proved near-identity transformations were supposed in the perturbation theory, a parameter ξ measuring the „extent of the necessary transformation”, a „shape factor” s, and a „regulation factor” λ can be introduced in a linear interpolation with small positive ε_1, ε_2 values as

$$\xi := \frac{\left|\mathbf{f}-\mathbf{i}^d\right|}{\max\left(\left|\mathbf{f}\right|,\left|\mathbf{i}^d\right|\right)},\quad \lambda = 1+\varepsilon_1+\left(\varepsilon_2-1-\varepsilon_1\right)\frac{s\xi}{1+s\xi},\quad \hat{\mathbf{i}}^{\mathrm{d}} = \mathbf{f}+\lambda\left(\mathbf{i}^{\mathrm{d}}-\mathbf{f}\right) \tag{14}$$

This interpolation reduces the task of the adaptive control in the more critical session and helps to keep the necessary linear transformation in the vicinity of the identity operator.

4 Simulation results for the water tank

In the simulations the non-adaptive and the adaptive controls' results are compared to each other α=0.25 *1/s* proportional, and β=$10^{-3}\times\alpha\, s^{-2}$, that is with a very small integrating coefficient in Fig. 1. As a rough system model, as an analogy of (6), constant coefficients *a, b, c,* and *d* were used as $\dot{T} = a\dot{M}_1 + b\dot{M}_2, \ddot{M}_3 = c\left(\dot{M}_1+\dot{M}_2\right)-d$.

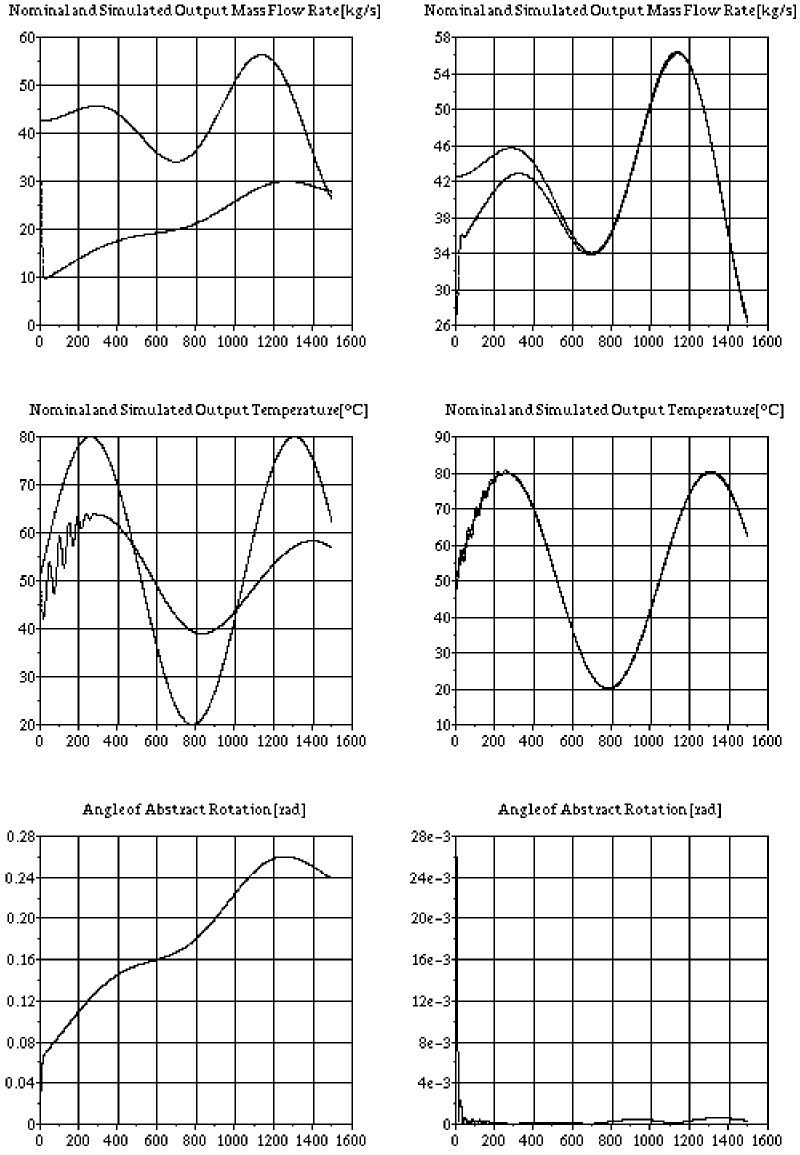


Figure 1. The operation of the simple PID (left column) and the adaptive (right column): prescribed and simulated mass flow rate [*kg*/*s*], prescribed and simulated temperature [°C], and the angle of the necessary step-by-step abtsract rotation (for the non-adaptive version it is calculated only without being used) vs. time [*s*]

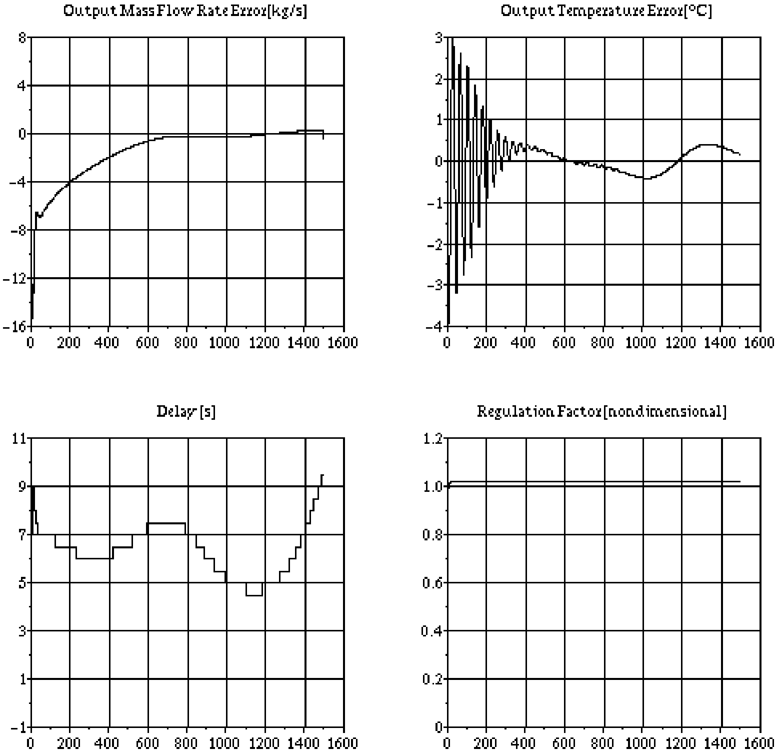


Figure 2. The tracking error for the mass flow rate [*kg/s*] and the temperature [°C], the time delay [*s*], and the regulating factor λ of the adaptive controller vs. time [*s*].

In Fig. 2 the mass flow rate and temperature tracking error, the delaya time, and the regulating factor λ are given for the adaptive controller to reveal some details. It can be seen that adaptivity causes considerable amendment in the accuracy of the control.

5 Thermal decay of the molecular nitrogen

The simplest examples of the chemical reactions are the reactions taking place in the mixtures of ideal gases. The thermodynamic model of these gases can be reconstructed by the use of certain "basic data" belonging to the temperture dependence of the equilibrium constant, the stoichiometric coefficients of the reaction and the chemical potential of the appropriate components in the mixture. If the chemical reaction is written in the form using positive or negative rational stoichiometric coefficients ν_i and symbolic notations for the components A_i as

$$\sum_i \nu_i A_i \Leftrightarrow 0\,, \tag{15}$$

and the model of the components is built up from the temperature-dependence of the molar heat of the gas $c_v(T)$ at constant volume by using the functions

$$\Psi_i(T;T_0) = 1 - \frac{\hat{s}_{i0}}{R} - \frac{1}{R}\int_{T_0}^{T} \frac{c_{vi}(t)}{t}\,dt \tag{16}$$

(R denotes the universal gas constant, T denotes the actual temperature in [°K] units, and T_0 is an arbitrary positive starting point of integration), the internal energy and the entropy of the mixture take the form as

$$E(T,V,\{N\}) = \sum_i N_i\left[\hat{\varepsilon}_{0i} - R\int_{T_0}^{T}\left(t\,\frac{d\Psi_i(t)}{dt}\right)dt\right] \tag{17}$$

$$S(T,V,\{N\}) = \sum_i N_i\left[R\text{-}R\Psi_i(T;T_0) + R\ln\frac{V}{N_i}\right]. \tag{18}$$

It can clearly be seen that $\hat{\varepsilon}_{0i}$, $\hat{s}_{0i}$ denote the molar internal energy and entropy of the components at temperature T_0. The pair of equations (16) and (17) makes it possible to deduce all the thermal data of the mixture. Via applying the 2nd Postulate of Thermodynamics for the thermal equilibrium of the mixture we obtain the socalled "Mass Action Law" stating that the exclusively temperature-dependent "equilibrium constant" $K(T)$ and the full pressure of the system p imposes a restriction to the possible chemical composition of the mixture.

$$K(T) = p^{\sum_i \nu_i} \prod_i \left(\frac{N_i}{N}\right)^{\nu_i}, \quad N = \sum_i N_i \tag{19}$$

On one hand, by measuring the temperature, the pressure, and the chemical composition of the mixture validity of (19) can be verifyed and $K(T)$ can be tabulated. On the other hand it is related to the model of the mixture in the following form:

$$\sum_i \nu_i \psi_i = -\frac{d}{dT}\left[T\ln K(T)\right] + \sum_i \nu_i(\ln RT + 1) \tag{20}$$

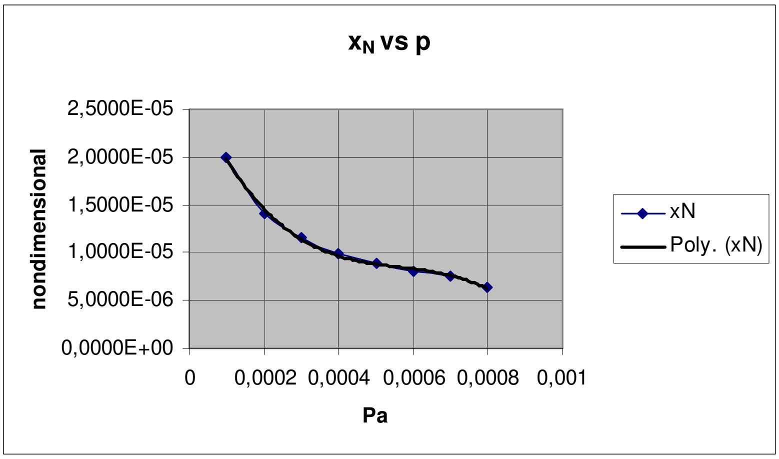


Figure 3. The mole fraction of the atomic nitrogen [non-diemnsional] during the process of throttlin down at low pressure [*Pa*] and high temperature

The chemical potential of the "pure" components just are equal to their molar Gibbs potential also related to the model. By the use of tabulated data describing the molar Gibbs potential of certain components the individual model functions of these components can be found as, e.g. for the molecular nitrogen as

$$\Psi_{N_2}(T;T_0)=\frac{1}{R}\frac{\partial\mu_{N_2}(T,p,1)}{\partial T}+\ln\frac{RT}{p}+1 \tag{21}$$

If our mixture consists of atomic and molecular nitrogen only (20) and (21) togeher determines ψ_N, too, for the atomic nitrogen. Since the entropy constants are built in in the ψ functions, and for since for the internal energy constants the relation

$$\sum_i \nu_i\hat{\varepsilon}_{0i}=-RT\ln K(T)-\sum_i \nu_i\left(RT\Psi_i+R\int_{T_0}^{T}t\frac{d\Psi_i}{dt}dt-RT\ln(RT)\right) \tag{22}$$

can be deduced, too, the energy constants of the atomic nitrogen can also be computed rom that of the molecular one.

Whenever a well defined amount of mixture of N and N_2 gases is in thermal equilibrium at a given temperature and pressure is throttled down to a prescribed pressure p, its full enthalpy $H(T,p,N_{N2},N_N)$, and its full mass $M(N_{N2},N_N)$ is conserved, furthermore the mixture has to satisfy the Mass Action Law at this lower pressure p. These three equations determine the new value of T, N_{N2}, *and* N_N at this lower pressure. Via applying various numerical fitting techniques no detailed here, by the use of the MICROSOFT EXCEL's SOLVER a third order polynomial was fitted to the $x_N(p)$ mole fraction of the atomic component of the mixture. The result is illustrated in Fig. 3.

6 Adaptive control of the thermal decay

In this case the controller's task is to giarantee an appropriate $p(t)$ function to produce a gas of nominal $x_N^N(t)$ composition. For this purpose the time derivative of the $p(t)$ function can directly be controlled. As a rough system model a constant value serving as the estimation of the $\partial x_N / \partial p$ derivative is used. Being a SISO system, for the control of this reaction scalar multiplication factors are used in the following form:

$$s(n) := \begin{cases} sign\left(\dot{x}^{Des}\right) \times sign\left(\dot{x}^{R}\right) \dfrac{\left|\dot{x}^{Des}\right| + \varepsilon}{\left|\dot{x}^{R}\right| + \varepsilon} & if\ \dot{x}^{Des} \neq 0\, and\ \dot{x}^{R} \neq 0 \\ 1 \quad otherwise & \end{cases} \tag{23}$$

in which ε is a very small number of about 10^{-25} order of magnitude to avoid both division and multiplication by zero in the control algorithm.

In Fig. 4 the nominal and simulated mole fraction values are described for the the non-adaptive and the adaptive approach, while Fig. 5 describes the tracking error in more details. It is evident that the adaptive completion of the control significantly increases the quality of the control.

To reveal details in Fig. 6 describes the desired and simulated speed of change in the pressure for the non-adaptive and the adaptive control. The differences are quite significant in the non-adaptive case.

Finally, in Fig. 7 the variation of the adaptive parameter (the scalar $s(n)$ multiplication factors) are described for the non-adaptive [not used but calculated only] and the adaptive cases. It is evident that the consecutive corrections of the adaptive control are very close to 1, while the similar graph pertaining to the non-adaptive case conveys information on the modeling errors and inaccuracies.

5 Conclusions

In this paper the behavior of the conventional PID and that of an adaptive control based on a novel branch of Computational Cybernetics were compared to each

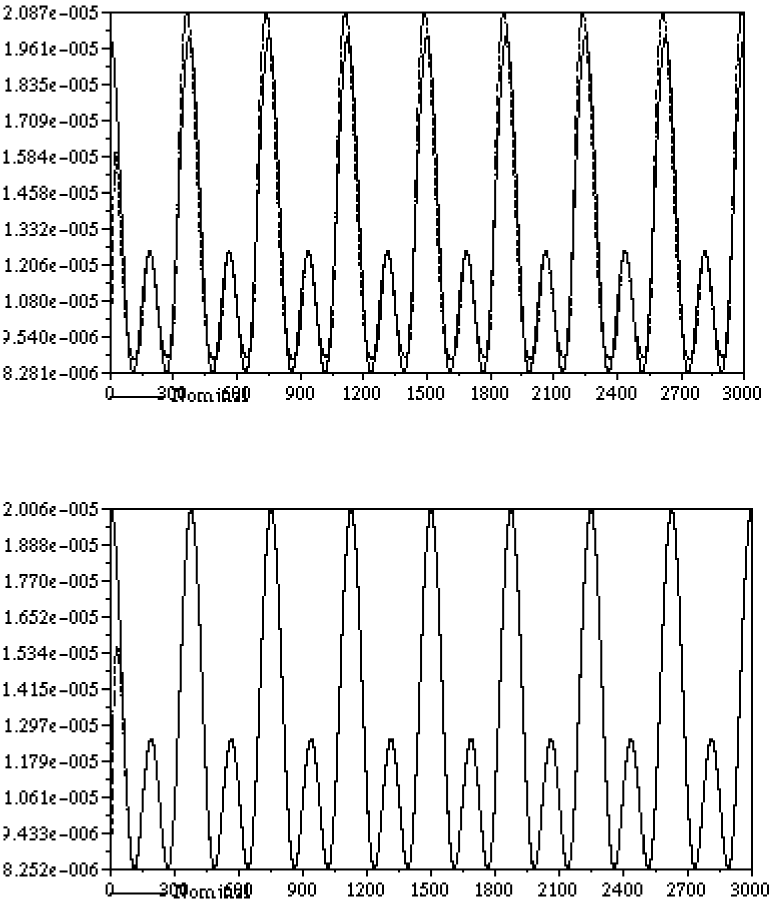


Figure 4. The nominal and the simulated mole fraction of atomic nitrogen in the non-adaptive and the adaptive case

other in the case of controlling an approximately modeled non-linear system having considerable and non-constant delay time.

The simulation results made it clear that a simple increase in the integrating coefficient can cause considerable improvement in the control but cannot approach the accuracy of the adaptive control when the delay time is important.

The here presented approach evades the sizing and learning problems having central significance in the rather traditional branch of soft computing [e.g. 14-20] by applying simple uniform operations in finite number of algebraic steps. The

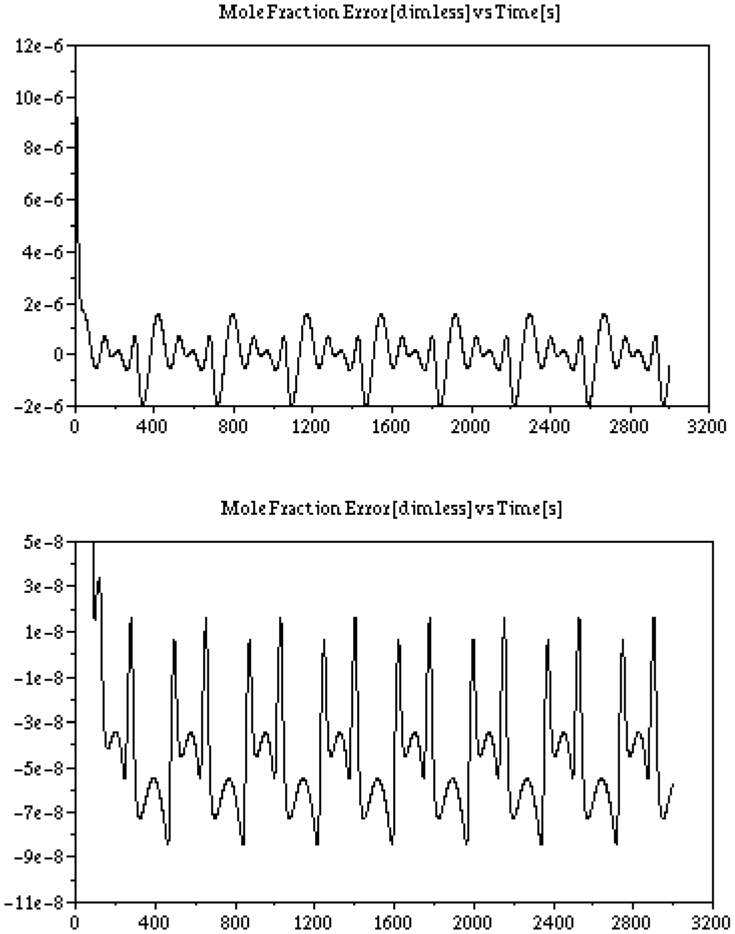
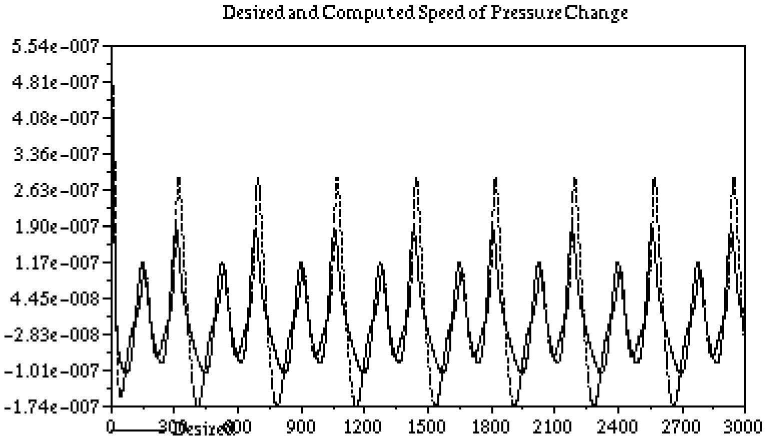


Figure 5. The tracking error of the non-adaptive and the adaptive control

size of the vectors and matrices used by it is simply determined by the modeled number of the degree of freedom of the system to be controlled. The “costs” of these advantages appear in the relatively limited class of problems for which the novel method can be applied.

The critical point is the proper convergence of the series of the linear transformations.

However the here-investigated paradigms suggests that from practical point of view the class of problems for which the new approach can be applied may be quite wide and may have drastic non-linearities, and time lag, too.



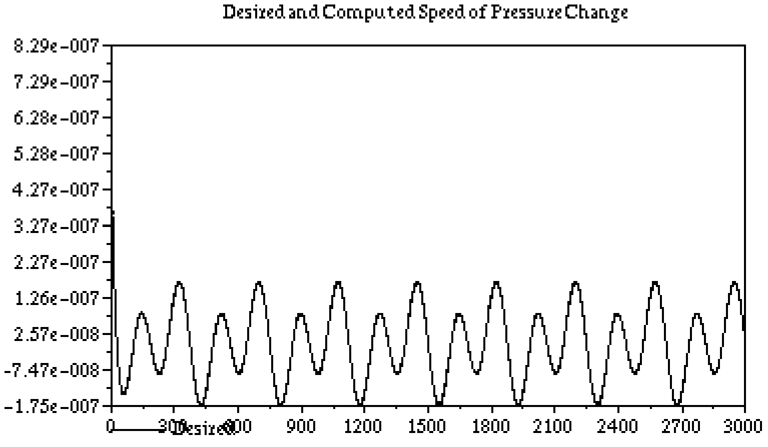


Figure 6. The desired and simulated speed of change in the pressure for the non-adaptive and the adaptive control

5 Acknowledgment

The authors gratefully acknowledge the support by the Hungarian National Research Fund (OTKA T034651, T034212 projects).

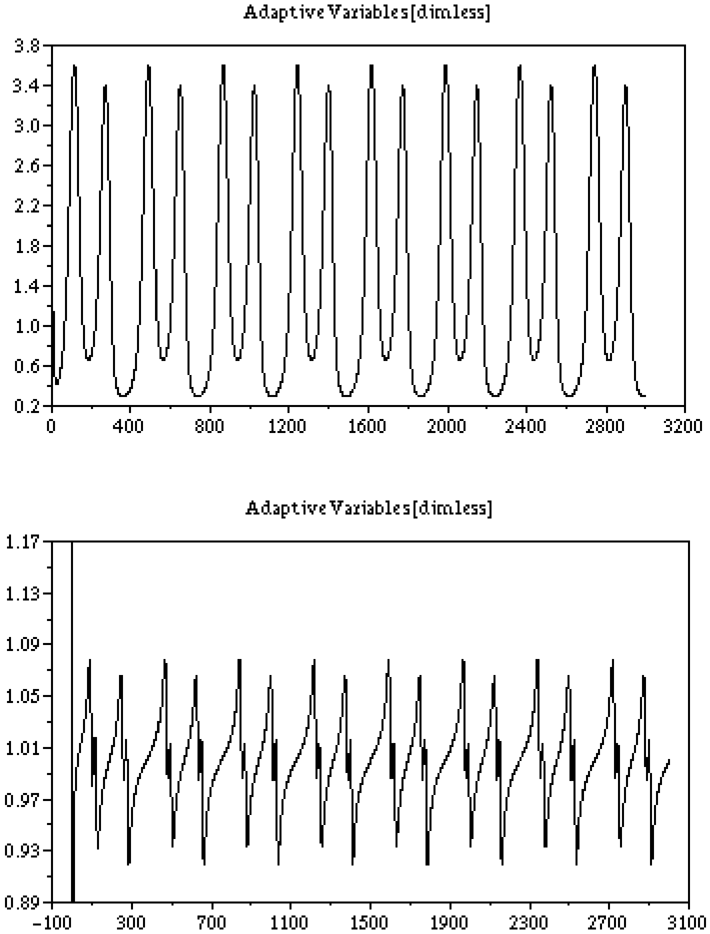


Figure 7. The adaptive variable in the case of the non-adaptive and the adaptive control (calculated only but not used in the non-adaptive case)

6 References

[1] M. Bröcker, M. Lemmen: "Nonlinear Control Methods for Disturbance Rejection on a Hydraulically Driven Flexible Robot", in the Proc. of the Second International Workshop On Robot Motion And Control, RoMoCo'01, October 18-20, 2001, Bukowy Dworek, Poland, pp. 213-217, ISBN: 83-7143-515-0, IEEE catalog Number: 01EX535.

[1] J. K. Tar, I. J. Rudas, J. F. Bitó: "Group Theoretical Approach in Using Canonical Transformations and Symplectic Geometry in the Control of Approximately Modeled Mechanical Systems Interacting with Unmodelled Environment", Robotica, Vol. 15, pp. 163-179, 1997.

[2] J. K. Tar, I. J. Rudas, J. F. Bitó, K. Jezernik: "A Generalized Lorentz Group-Based Adaptive Control for DC Drives Driving Mechanical Components", in the Proc. of The 10th International Conference on Advanced Robotics 2001 (ICAR 2001), August 22-25, 2001, Hotel Mercure Buda, Budapest, Hungary, pp. 299-305 (ISBN: 963 7154 05 1).

[3] Yahya El Hini: "Comparison of the Application of the Symplectic and the Partially Stretched Orthogonal Transformations in a New Branch of Adaptive Control for Mechanical Devices", Proc. of the 10th International Conference on Advanced Robotics", August 22-25, Budapest, Hungary, pp. 701-706, ISBN 963 7154 05 1.

[4] J. K. Tar, A. Bencsik, J. F. Bitó, K. Jezernik: "Application of a New Family of Symplectic Transformations in the Adaptive Control of Mechanical Systems", in the Proc. of the 2002 28th Annual Conference of the IEEE Industrial Electronics Society, Nov. 5-8 2002 Sevilla, Spain, Paper SF-001810, CD issue, ISBN 0-7803-7475-4, IEEE Catalog Number: 02CH37363C.

[5] M.J. Feigenbaum, J. Stat. Phys. 19, 25, 1978;

[6] M.J. Feigenbaum, J. Stat. Phys. 21, 669, 1979;

[7] M.J. Feigenbaum, Commun. Math. Phys. 77, 65, 1980;

[8] J. K. Tar, J. F. Bitó, K. Kozłowski, B. Pátkai, D. Tikk: "Convergence Properties of the Modified Renormalization Algorithm Based Adaptive Control Supported by Ancillary Methods", Proc. of the 3rd International Workshop on Robot Motion and Control (ROMOCO '02), Bukowy Dworek, Poland, 9-11 November, 2002, pp. 51-56, ISBN 83-7143-429-4, IEEE Catalog Number: 02EX616.

[9] J. K. Tar, I. J. Rudas, J. F. Bitó, L. Horváth, K. Kozłowski: "Analysis of the Effect of Backlash and Joint Acceleration Measurement Noise in the Adaptive Control of Electro-mechanical Systems", Accepted for publication on the 2003 IEEE International Symposium on Industrial Electronics (ISIE 2003), June 9-12, 2003, Rio de Janeiro, Brasil, CD issue, file BF-000965.pdf, ISBN 0-7803-7912-8, IEEE Catalog Number: 03th8692.

[10] J. F. Bitó, J. K. Tar, I. J. Rudas: “Novel Adaptive Control of Mechanical Systems Driven by Electromechanical Hydraulic Drives”, in the Proc. of the 5th IFIP International Conference on Information Technology for BALANCED AUTOMATION SYSTEMS In Manufacturing and Services, Sheraton Towers and Resort Hotel, Cancún, Mexico September 25-27, 2002, Cancún, Mexico.

[11] G. F. C. Rogers, Y. R. Mayhew: „Thermodynamic and Transport Properties of Fluids – SI Units”, 4th Edition, Blackwell Oxford UK & Cambridge USA, 1980, ISBN 0-631-90265-1.

[12] B. S. Massey: „Mechanics of Fluids”, Sixth Edition, Chapman & Hall, 1989, ISBN 0 412 34280 4.

[13] Dr. Harmatha András: “Termodinamika műszakiaknak”, (in Hungarian), Műszaki Könyvkiadó, Budapest, 1982, ISBN 963 10 4467 X, p. 197 and 276.

[14] R. Reed: "Pruning Algorithms - A Survey", IEEE Transactions on Neural Networks, 4., pp.- 740-747, 1993.

[15] S. Fahlmann, C. Lebiere: "The Cascade-Correlation Learning Architecture", Advances in Neural Information Processing Systems, 2, pp. 524-532, 1990.

[16] T. Nabhan, A. Zomaya: "Toward Generating Neural Network Structures for Function Approximation", Neural Networks, 7, pp. 89-9, 1994.

[17] G. Magoulas, N. Vrahatis, G. Androulakis: "Effective Backpropagation Training with Variable Stepsize" Neural Networks, 10, pp. 69-82, 1997.

[18] C. Chen, W. Chang: "A Feedforward Neural Network with Function Shape Autotuning", Neural Networks, 9, pp. 627-641, 1996.

[19] W. Kinnenbrock: "Accelerating the Standard Backpropagation Method Using a Genetic Approach", Neurocomputing, 6, pp. 583-588, 1994.

[20] A. Kanarachos, K. Geramanis: "Semi-Stochastic Complex Neural Networks", IFAC-CAEA '98 Control Applications and Ergonomics in Agriculture, pp. 47-52, 1998.

Modeling Course for Virtual University by Features

László Horváth*, Imre J. Rudas**

John von Neumann Faculty of Informatics, Budapest Polytechnic, Népszínház u. 8.,
Budapest H-1081 Hungary
*lhorvath@zeus.banki.hu
**rudas@zeus.banki.hu

Abstract:

Environments with large number of interrelated information uses several advanced concepts as computer description of different aspects of modeled objects in the form of feature based models. In this case a set of features is defined then used for the purpose of modification of an initial model to achieve a final model as a description of an instance of a well-defined complex object from a real world environment. Utilization this approach and some relevant methods have been investigated by the authors to establish course modeling in virtual university environments. The main objective is definition generic model entities for courses and instance model entities for student course profiles. Course model entities describe virtual university activities. The modeling can be applied generally but it is being developed for the domain of higher education in virtual technologies. The paper introduces some virtual university related concepts and the approach of the authors to virtual university. Following this feature driven associative model of virtual course developed by the authors is explained. Some issues about the conceptualized application oriented virtual course features are discussed as a contribution to implementation of a virtual classroom model proposed by the authors. Finally, possibilities of integration of the university model with engineering modeling systems are discussed taking into account present day virtual universities and possibilities to communicate with prospective students both in professional design and home computer environments.

1 Introduction

Spending days, weeks even months for attending campus courses is impossible for most of the people engaged in industrial employment as engineers. At the same time, substantial changes in knowledge in some domains, changes in demands by employers against employees, changes in field of activities of humans and other motivation of humans to improve their knowledge in some of the possible directions resulted a demand lifetime learning for more and more people. This is why distance learning has been expanded in recent years. However, conventional forms of distance learning in higher education have a lot of drawbacks in comparison to campus courses. If students can not attend the campus the campus should be brought to students. This has made possible by development of Internet technology and virtual classroom models as proposed by the authors. Virtual classrooms can be established as special purpose portals. Numerous virtual classrooms and universities offer excellent programs on the Internet. The related amount and complexity of teaching information and classroom activities make design and maintenance of these portals very difficult. At the same time the flexibility of classroom programs demanded by potential students are hard to provide by the existing portals [4]. The authors analyzed the related problems and decided investigations on application of advanced computer modeling together with well proved knowledge technology on the basis of Internet technology for the purpose of virtual classroom.

Internet portals for advanced distance learning are often called as virtual universities. Virtual universities offer services similar as of conventional universities but their purpose is not simply a solution to replace them [3]. Existing virtual universities have been established for different purposes and programs in higher education. The authors would like to contribute to methodology basics of virtual universities by following a model-based approach. Different aspects of a comprehensive virtual university concept and methodology by the authors are included in [1] and [2] as earlier results utilized by the reported research.

Existing virtual classroom methods do not offer direct tools for customization of existing course models. The author's approach involves description of effects of new components on modified course models. This needs description both of the consequences of modifications and the modified relationships. An obvious solution is feature driven associative modeling. The research reported in this paper is about the above-mentioned approach to virtual classroom in higher education especially in the field of education in engineering. The only solution is taking the advance of computer modeling. Authors decided to establish virtual classroom model by using of advanced concepts as knowledge intensive feature and associativity based modeling for description of virtual classroom objects. It is the main topic of this paper.

The extending field of virtual universities motivated the authors to adapt virtual university principles to teaching an other large group of virtual technologies. The authors propose in the paper a representation that describes virtual university and can be integrated with virtual engineering modeling systems. Internet technologies and proven methods of computer based training are used as a basis of this research.

The paper introduces some virtual university related concepts and the approach of the authors to virtual university. Following this feature driven associative model of virtual course developed by the authors is explained. Some issues about the conceptualized application oriented virtual course features are discussed as a contribution to implementation of a virtual classroom model proposed by the authors. Finally, possibilities of integration of the university model with engineering modeling systems are discussed taking into account present day virtual universities and possibilities to communicate with prospective students both in professional design and home computer environments.

2 Virtual University

Platform-independent Internet software enhances advanced forms of distance learning. This requires substantial computer resources both on university and student sides. Engineers are working in a similar system in their every day company practice so that university activities can be done in the same system as professional engineering activities. Students from the non-professional area can join to this system. Finally, companies engaged in development, production or consult of engineering modeling, in common sense words CAD/CAM systems are interested in participation at higher education systems and may offer substantial computing and knowledge resources. Virtual university offers services similar as of conventional university using this environment for this purpose of campus and distance type of higher education [3].

The outline of the scene of virtual university can be seen on Fig. 1. Teachers are operating virtual university services. Virtual university is installed on a computer system that can provide the necessary services to students through network. Students use local services, e. g. at a company, or services of some providers. The virtual system establishes both off line and on line communication amongst teachers and students.

Virtual universities are extended learning communities and constitute virtual campuses on the basis of advanced communication tools as World Wide Web and telephone systems [7]. Motivated, keen instructors, classroom helpers, etc. share their knowledge with students in a large computer system. Dramatic development of distance communication technologies and virtual technologies are anticipated. The authors think that this is the high time to make research in the above outlined topic.

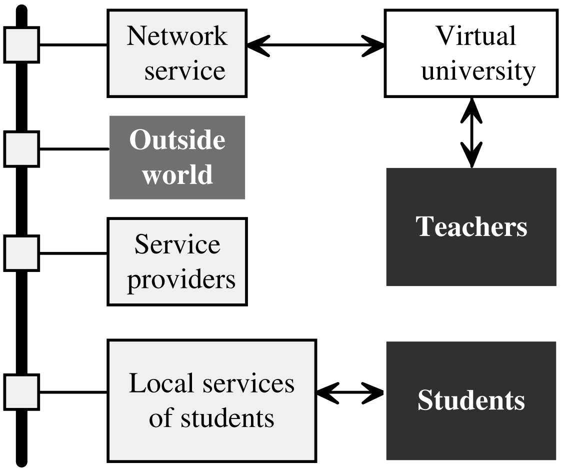


Fig. 1 Essential connections

The VU concept is growing from advanced distance learning [6], [8]. The main resources of a virtual university are published lectures, course materials, assignments for homework, on-line conferencing or consultation and live chats. Materials are browseable and a lot of links are inserted for background information. Virtual systems open new possibilities for the virtual university [5], [4]. Special education versions of modeling procedures can behave as instructor with navigation, correction and explanation features. Advanced commercial modeling systems can be tailored for this purpose. Lectures can be illustrated by live modeling etudes. Virtual laboratory makes it possible for students to login from remote computers within the virtual university. Individual and group work tasks, directed drills and case studies can be made available for students. The engineering model is created and annotated by the student then evaluated and annotated by the teacher. Exam questions can be assigned for solution by the using of modeling procedures. Video materials can be applied to carry records of modeling procedures and can be displayed step by step at learning. Where virtual laboratory can not be accessed on line, special education versions of modeling procedures can be downloaded by students of these special virtual courses.

Multimedia lectures can be applied by hyperstructure to give explanations on different levels of knowledge. More detailed lectures can be chosen by students who are interesting in a given topic.

3 An approach to Model Based Virtual University

In the author's approach for virtual university [1] teachers operate virtual university functions and provide the necessary services to students through computer network. Students use local services, e. g. at a company, or services of some outside providers. Both off line and on line communications are to be operated amongst teachers and students. Virtual university constitutes virtual campus [5]. Motivated, keen instructors, classroom helpers, etc. share their knowledge with students using advanced functions offered by large computer systems. Virtual classroom can be considered as an up-to-date solution for distance learning.

Strongly interrelated information structure about virtual classroom objects to be represented is to be created and handled in computer systems. It is obvious that the only way for handling this information in computer is establishment of a well structured, attributed and related description.

An approach to modeling of the related virtual university activities has been outlined in [1]. Model of a virtual university consists of a set of function entities grouped according to tasks and connected by relationships defined between them. Managers (Fig. 2.) handle function entities. A manager consists of a set of computer procedures for handling creation, modification and application of well-defined function entities. Course manager handles modules of the teaching program. Enrollment manager does credit and fee related affairs. Communication manager supervises communication tools available for teachers and students. Teaching material manager downloads materials, offers on line video service, sends materials as E-mail attachments automatically and establishes links to outside sources of materials. Process manager deals with processes in managing of courses. There are several other managers as it can be seen on the Fig. 1.

The another important problem area is modeling courses (Fig. 2). In the approach by the authors to virtual classroom model structure, a course is a sequence or network of modules. In other words the main structural elements of courses are modules. A module consists of blocks. A block involves topics. A topic consists of topic related procedures for handling principles, methods, relationships, examples, questions, materials and instructor activities. Links can be defined to other topics and outside world objects. Modules are arranged in courses or can be applied individually. Core studies contain basic and essential knowledge. They are modules or blocks. A course offers a choice of modules, blocks and topics.

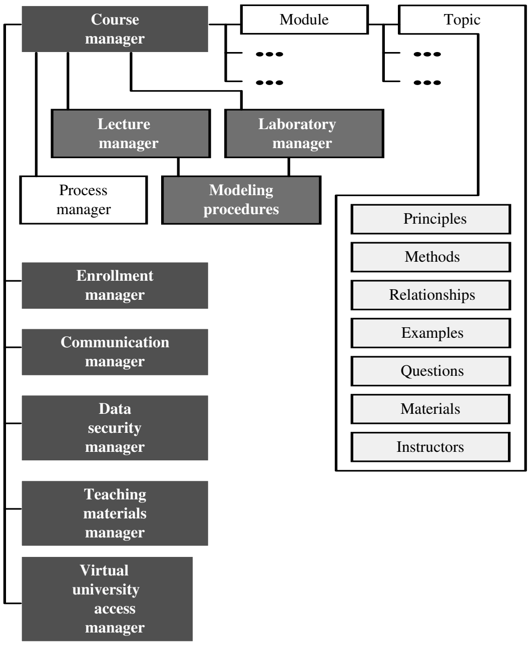


Fig. 2 Functional structure of a virtual university

Background analysis of virtual classroom revealed its components as curriculum, teaching processes, credits, students and virtual laboratories [2] (Fig. 3.). Curriculum as an organized learning experience involves content of a degree program, provides conceptual structure and time frame to get that degree. The course is an organized learning experience in an area of the education. A curriculum can be composed using courses or courses can be defined according to predefined curriculum. Virtual laboratories are composed using software modules, software arrangements for assignments as well as results of student work as assignments and degree works.

Virtual classroom is active in an environment where students, teachers and related humans and objects from the outside world are integrated (Fig. 3.). Classroom model, course instance model and outside world model communicate teachers, students and outside sites through the Internet.

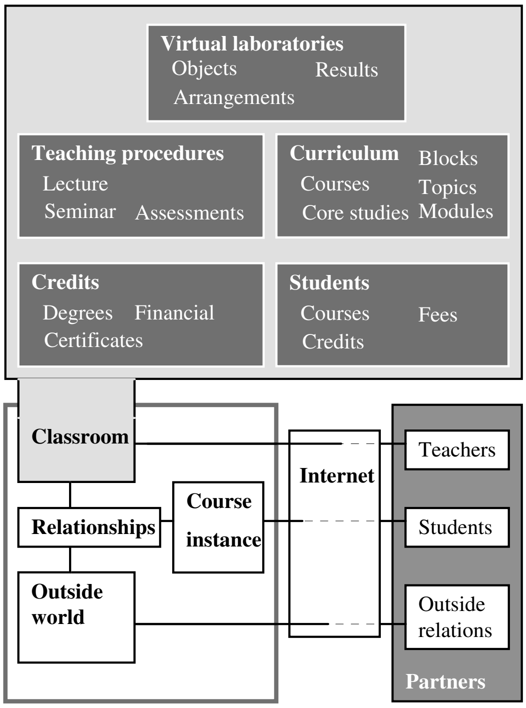


Fig. 3 Virtual classroom and its environment

The above outlined approach to virtual university and virtual classroom constitutes basic considerations for modeling of virtual classroom by the authors.

4 Model of Virtual Course

The course model as proposed by the authors uses structure of its elements, feature driven construction of modules and associativities between course elements. Track has been introduced as a course element comprising a set of modules for a well-defined purpose. Tracks and modules can be involved in different courses as instances. In this case model descriptions are not duplicated.

In the feature driven modeling approach a module is considered as a base feature modified by module modification features to create a customized module instance. Content of a module is defined by the teachers engaged in the related teaching program and customized on the basis of student demands. Consequently, generic models are applied and used at creation their instances. At the same time types of base and module modification features with basic model related characteristics are defined by course modeling experts. In this context base and module modification feature types are frames final content of which are defined in feature instances.

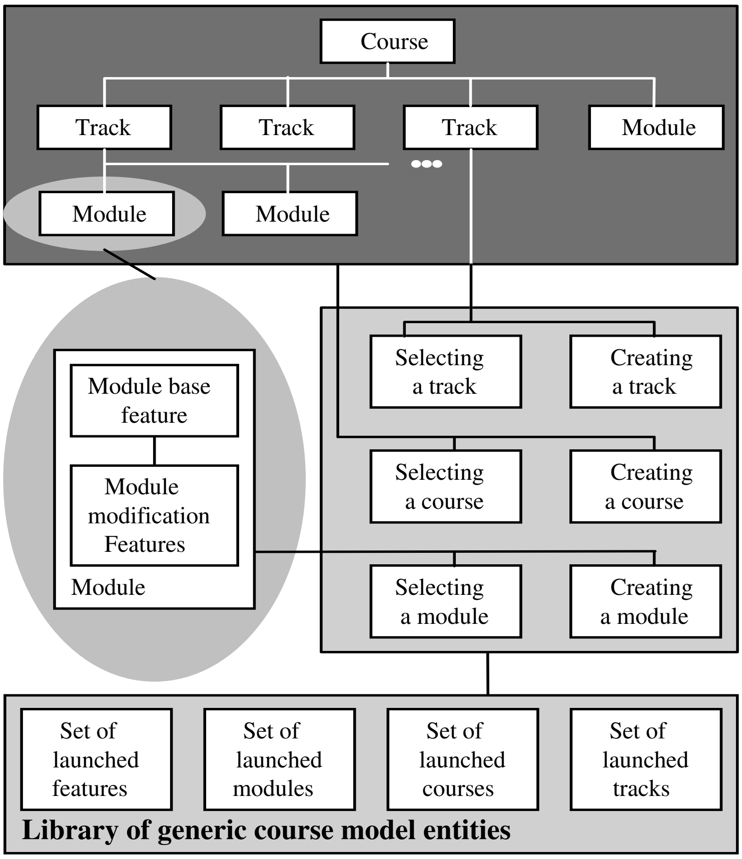


Fig. 4 Construction of feature based course models

Basic construction schema of feature based course models is summarized by Fig. 4. Generic or frame model entities are launched then stored in libraries, as it will be explained in chapter IV of this paper. A course entity is selected then related with track and module entities or customized by appropriate changes. Customization covers selection then adaptation purpose configuration. Track models also created or configured. Module entities are created by modification of a base feature by appropriate module modification feature entities selected from the choices stored in appropriate libraries.

Predefined classroom features are used for modification of modules to create module instance for customary higher education teaching programs. Fig. 5 summarizes a possible set of classroom features. A module is modified feature by feature if it has necessary places and surfaces to create the modification. In other words information carried by the feature should be accepted by the description structure of the actual module instance. Features have been grouped according to their requirements for place and surface of modification. Structural, contact, assessment, content and handout groups of features have been defined by the authors.

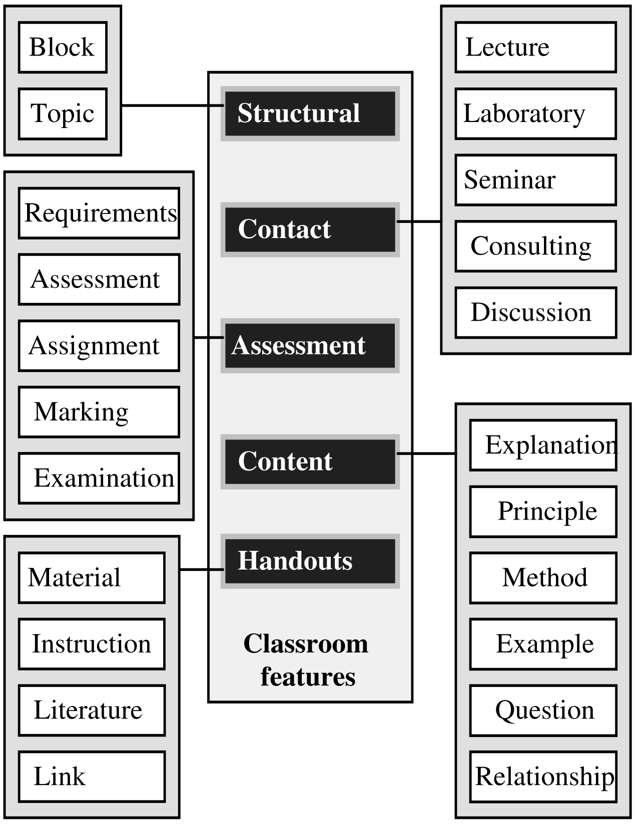


Fig. 5 A possible set of classroom features

Structural feature modifies structure of a module by introducing a new block or topic. Contact features place course elements on the module to establish contact activities between students and teachers. Consulting and discussion are inherently interactive. Lectures, laboratories and seminars can be also interactive. Semi interactive contact features substitutes teacher by using of sets of typical answers and explanations together with effective searching. Content feature contributes to teaching content of the module by purposeful explanations, description of principles and methods, representative examples, putting questions with or without answers and relating things by relationships. Assessment features complete module by description of requirements, composition of assessment, assignment, marking schemas and examinations. Finally handout features include materials, instructions, literatures and links to outside materials.

Surface for a feature can be placed on the module or on one of the existing module modification features according to the group of the feature and the decisions for modifications. Some of the features can modify only the base feature (module) whereas others can also modify previously placed module modification features.

Fig. 6 shows two examples for placing features on modules at initial stage of their creation. Base feature on Fig. 6/a has been defined to provide surfaces for placing of structural, handout and assessment features. *Topic A* modifies base feature as a structural feature. *Topic A* has been defined to provide surfaces for placing of contact and content features. Feature *Lecture B* is placed at the contact feature surface of feature *Topic A*. Fig. 6/b illustrates an alternative solution when feature *Lecture B* is placed at the contact feature surface defined directly on the base feature. Feature *Lecture B* has surface for placing contact features.

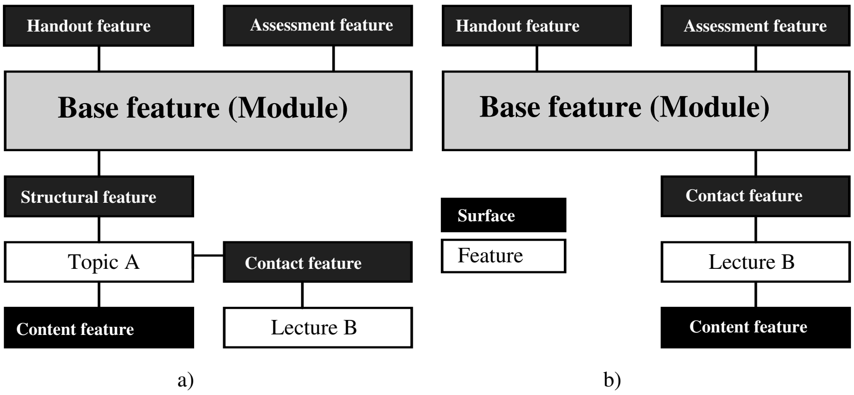


Fig. 6 Placing module modification features

Modules are built into a structure defined for a course or a track. Conventional course plans relate modules each other only by information for other modules that act as preliminary study requirements. However, generally not a complete module but only its some elements are necessary to understand given element of an other module. If the module instance configure omit that element, preliminary requirement is no more exist. To achieve a flexible description and avoid unnecessary preliminary study requirements, modules are integrated by relationship entities defined in a course or a track (Fig. 7). The course or tack model defined by using of this method is brief and consistent without redundancies. Consistency can be checked by an appropriate computer procedure.

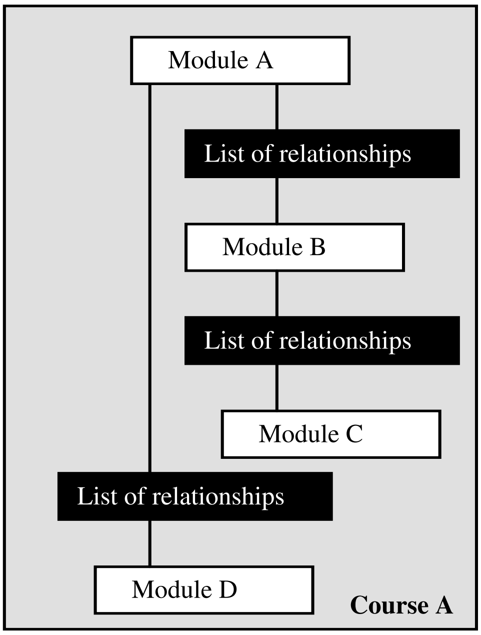


Fig. 7 Integration of modules by relationships

Associativities are also defined within modules. Basically, the structure of modification describes associativities. Other associativities can be defined between features or between any description elements in their content. Similarly, associativities can be defined between the modules and their lower level elements and the outside world in the form of links. Associativity describes dependency while a simple link only points to something. In case of change in an element, the associative elements change according to the existing associativity definitions. Associativities are often defined for the purpose of saving teacher intent.

5 Application Oriented Virtual Course Features

An absolute free definition of features would require unreal amount of analyses on features to reveal these characteristics. It is impossible to define a complete set of virtual course features that can be applied in all possible courses in all possible fields and purposes of higher education. On the other hand feature-processing procedures must be informed on some basic feature characteristics. The solution in the author's approach is application oriented feature definition in the course model relied upon general feature type definitions (Fig. 8). This method has been proved in engineering modeling at solutions for similar problems.

An other problem to be solved is the high amount of custom feature variants that can be anticipated in the higher education practice. It is impossible to define them as individual features. Instead, configurable generic features are applied. Instance features can be easily configured by adaptation of generic features. Generic entities are also applied on the levels of module, track and course. Modules and features can be suppressed or their parameter values can be changed in order to gain purposefully configured instances of module modification features, modules, tracks and courses.

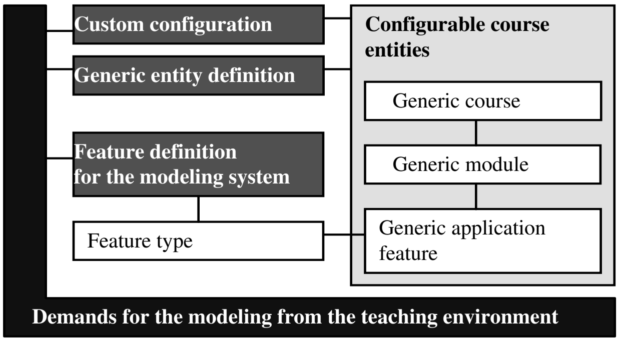

Custom configuration
Generic entity definition
Feature definition for the modeling system
Feature type
Configurable course entities
Generic course
Generic module
Generic application feature
Demands for the modeling from the teaching environment

Fig. 8 Concept of configurable course model

It is very important to emphasize that all features on the application level are defined by teachers who do the education program. Also teachers define generic modules and courses. Configuration just simplifies work of a teacher or group of teachers at the definition of flexible customer oriented courses and modules. It would be a bad idea to define entities on the application level by teachers or researchers other than doing the offered courses and modules. The knowledge and teaching skill are of personal nature. Teachers must describe them for computer models to make the use of computer system at education. One of the main advances of the above outlined application oriented modeling is that teachers can define entities without any advanced skill in course modeling.

6 Teaching Engineering Modeling in Virtual University

Quick development of modeling principles, methods and systems requires frequent training of engineers. Product related training courses at companies are not appropriate to deliver new higher education related topics. Consequently the proposed virtual university concept is most important for further education in the industry, however they can be utilized with equally success in undergraduate and graduate courses. A concept of configurable course model is outlined in Fig. 9.

Modeling tools are organized in comprehensive Computer Aided Design/Computer Aided Manufacturing (CAD/CAM) systems in the present day engineering design practice. Users of modeling procedures use company support through Internet and Internet communication with other users. On the application site on line help, tutorials and manuals are available as knowledge and practice sources (Fig. 10). This environment can be used in virtual university environment as a courtesy of CAD/CAM manufacturers.

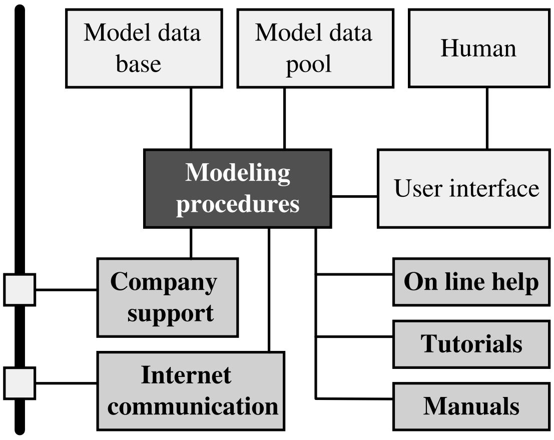


Fig. 9 Concept of configurable course model

Virtual university managers connect laboratory with industrial modeling support and application environments. Industrial application environments offer real world examples and case studies for the virtual university and use training services from the virtual university.

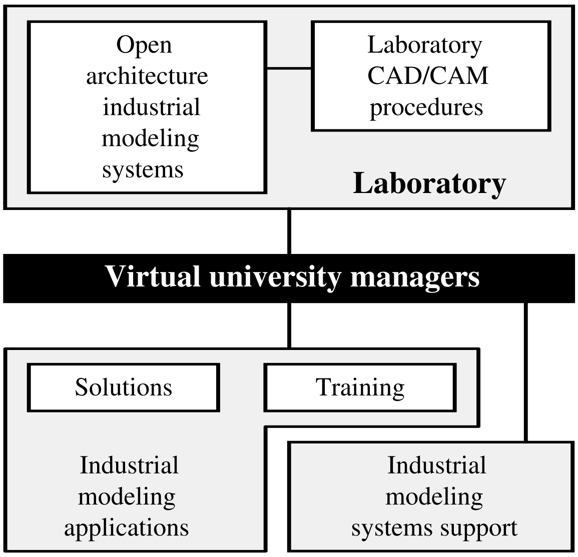

Open architecture industrial modeling systems
Laboratory CAD/CAM procedures
Laboratory
Virtual university managers
Solutions
Training
Industrial modeling applications
Industrial modeling systems support

Fig. 10 Concept of configurable course model

7 Conclusions

The authors have proposed a model aimed as a contribution to the dream of virtual university. Using this model an advanced application and utilization of Internet is realized in a manner that opens the system for the teachers to make custom configured course, track, module and module modification feature model entities by the method of configuration instance entities using generic entities.

A concept and an approach have been outlined for a virtual university system in this paper. The purpose is to model a virtual university that is appropriate for teaching virtual technologies for engineering design. The proposed virtual university model consists of functions that are handled by functional managers. Modeling methods are modeled as topics. Topics are organized in modules and modules are organized in courses. The involved teachers also define generic entities. The same teachers use these entities then for creating instance entities on the basis of student demands. Making and application of course entities are tied closely to teachers to save the personal nature of teaching and outstanding performances of teacher individuals. Wide application of associativities guarantees saving intent of teachers. The proposed course model consists of module entities. Modules are constructed as series of modifications by module modification features. Modules are organized into tracks and courses by relating them using relationships.

References

[1] Imre J. Rudas, László Horváth, "Teaching Virtual Technologies in Virtual University", *Proceedings of the International Conference on Information Technology based Higher Education and Training*, Istanbul, Turkey, pp: 127-131, 2000.

[2] László Horváth, Imre J. Rudas, Okyay Kaynak: "Modeling Virtual Classroom for Education in Engineering”, *Proceedings of International Conference of Information Terchnology based Higher Education and Training*, Kumamoto, Japan, pp.395-398, 2001.

[3] Richard Teare, David Davies, Eric Sandelands, “The Virtual University: An Action Paradigm and Process for Workplace Learning”, *Cassell Academic*, 1999.

[4] Starr Roxanne Hiltz, The Virtual Classroom: "Learning Without Limits Via Computer Networks", *Human/Computer Interaction*, 1997

[5] Gerald C. Van Dusen, The Virtual Campus, Technology and Reform in Higher Education, *George Washington University School of Education & Human Dev.*,1997

[6] Leadership in Continuing and Distance Education in Higher Education, Allyn & Bacon, 1998

[7] Rena M. Palloff, Keith Pratt,Building Learning Communities in Cyberspace : Effective Strategies for the Online Classroom, Jossey-Bass Publishers, 1999

[8] Thomas E. Cyrs, Teaching and Learning at a Distance : What It Takes to Effectively,Design, Deliver, and Evaluate Programs, Jossey-Bass Publishers, 1997

Topological Characterization of Cellular Structures

Tamás Réti

Budapest Polytechnic, Doberdó u.6, H-1031 Budapest, Hungary

E-mail: reti@zeus.banki.hu

Károly Böröczky, Jr.

Alfréd Rényi Institute of Mathematics, P.O. Box 127, H-1364 Budapest, Hungary

E-mail: carlos@renyi.hu

Abstract: In order to characterize quantitatively the local topological structure of cellular systems a new method has been developed. First, we analyzed the topological properties of infinite periodic cellular structures, and then the general theoretical results obtained have been adapted to the local topological characterization of 2-dimensional finite cellular surface systems. The concept of this new approach is based on the use of the so-called double toroidal embedding (DT embedding) by which a finite cellular system defined on a torus can be generated from an infinite periodic cellular system. The DT embedding is a special mapping, which enables to preserve all the local topological properties of the original infinite periodic cellular system. As a result of performing a DT embedding, so-called neighborhood coefficients can be generated. The neighborhood coefficients are scalar topological invariants, by which the local topological structure of cellular systems can be quantitatively evaluated and compared. Moreover, by investigating the relationship between the neighborhood coefficients and other local topological quantities, we verify that the validity of the Weaire-Fortes identity can be extended to a broad class of infinite periodic cellular systems and 2-d finite cellular surface systems (i.e. generalized fullerene-like surface structures). Finally, it has been shown that the traditional definition of fullerenes can be generalized by introducing the notion of the cellular fullerene, which is considered as a finite cellular system defined on a 2-d unbounded, closed and orientable surface.

Keywords: cell, embedding, toroidal graphs, Weaire-Fortes identity, corona, fullerene

1. Introduction

In various fields of material sciences, many interesting 2- and 3-dimensional structures (fullerenes, nanotubes, froths, metal foams, polycrystals) can be modeled by a special arrangement of space filling polygons and polyhedra (i.e. 2- or 3- dimensional polytopes) and thus can be considered as finite or infinite cellular systems. Over the past two decades, most studies have concentrated on 2-d cellular structures which may be represented by infinite, planar networks, usually with trivalent vertices (i.e. three edges at each vertex) [1-6]. This paper presents a general method, which is designated primarily to the topological evaluation of infinite periodic and finite cellular systems composed of d-dimensional polyhedra (polytopes) where d ≥ 2.

The proposed method is based on the application of a double toroidal embedding (DT embedding) by which a finite space-filling cellular system defined on a torus can be generated. The DT embedding is considered as a one-to-one mapping of the topological types, which enables to preserve all the local topological properties of the original infinite periodic cellular system. It will be shown that, after performing a DT embedding, so-called neighborhood coefficients can be computed, by which the local topological structure of periodic cellular systems can be simply analyzed and compared. Additionally it will be verified that the validity of the Weaire-Fortes identity [2-4] playing a key role in the topological description of 2-dimensional random cellular patterns, could be extended to finite dimensional periodic cellular systems. The fundamental results concerning the extension of the Weaire-Fortes identity are represented by Eqs. (33 and 34). Finally, it is shown that the traditional definition of fullerenes can be generalized by introducing the notion of the cellular fullerene, which is considered as a finite cellular system defined on a 2-d unbounded, closed and orientable surface.

2. Locally finite periodic cellular systems

The most important type of infinite d-dimensional cellular systems is the so-called countable cellular system [7]. A countable cellular system is considered as a face-to-face tiling (tessellation) of d-dimensional Euclidean space denoted by $\boldsymbol{E}^{(d)}$ by a countable set of d-dimensional compact combinatorial polyhedra (polytopes). Each d-dimensional polyhedron called a cell is topologically equivalent (homeomorfic) to a d-dimensional sphere. A countable cellular system denoted by $\boldsymbol{\Omega}_d$ is defined by taking into consideration the fulfillment of the following requirements:

i. $\boldsymbol{\Omega}_d$ can be represented as

$$\mathbf{\Omega_d} = \left\{ \mathrm{A_j} \,\middle|\, \mathrm{j} \in \mathrm{I_P}\text{...and...}\bigcup_{\mathrm{j}} \mathrm{A_j} = \mathbf{E^{(d)}} \right\} \qquad (1)$$

where I_P is the index set of positive integers, A_j is the jth cell (polyhedron) in $\boldsymbol{\Omega}_d$.

ii. The k-dimensional faces of polyhedra included in $\boldsymbol{\Omega}_{\mathbf{d}}$ (k=0,1,2,…,d-1) are also compact combinatorial polyhedra, and the maximum number of k-dimensional faces is less then γ_k, where γ_k are finite positive integers for k=0,1,2,…d-1. (The 0-dimensional and 1-dimensional faces of polyhedra are called vertices and edges, respectively.)

iii. Polyhedra can be included in a d-dimensional sphere with a finite radius, which guaranties that the “size” of cells is finite [7].

iv. All of the k-dimensional faces of a d-dimensional polyhedron have a positive k-dimensional volume (measure) for k=1,2,…d.

v. Each (d-1) -dimensional face between cells is the common face of two different cells (polyhedra) exactly.

vi. Additionally it is assumed that $\boldsymbol{\Omega}_{\mathbf{d}}$ is locally finite [7]. By definition, a countable cellular system is called locally finite if there exists a positive number ρ for any arbitrary point P_x in $\boldsymbol{E}^{(\mathbf{d})}$, such that every d-dimensional sphere $G(P_x,\rho)$ with radius ρ and center point P_x, contains finite number cells from $\boldsymbol{\Omega}_{\mathbf{d}}$ only. This definition implies that there are no singularity points of cells in the cellular system. For each vertex X (0-dimensional face) in $\boldsymbol{\Omega}_{\mathbf{d}}$ the number of edges (1-dimensional faces) incident to X is called the valency of X, denoted by r (or r(X)). If all of the vertices of have the same valency R, then $\boldsymbol{\Omega}_{\mathbf{d}}$ is said to be a regular, or R-valent cellular system.

For purposes of our investigations the most important groups of locally finite cellular systems are the periodic cellular systems. A locally finite cellular system $\boldsymbol{\Omega}_{\mathbf{d}}$ is called periodic, if there exists a d-dimensional parallelepiped $\mathbf{\Pi_d}$ represented by a linearly independent vector system ($\mathbf{v_1}$, $\mathbf{v_2}$,… $\mathbf{v_k}$, ...$\mathbf{v_d}$) for which relationships

$$\mathbf{\Pi_d} \subset \left\{ \bigcup A_j \middle| A_j \in \mathbf{\Omega_d} \right\} \tag{2a}$$

and

$$\mathbf{E^{(d)}} = \left\{ \bigcup \mathbf{B}_v \middle| \mathbf{B}_v = \mathbf{\Pi_d} + \sum_{k=1}^{d} \varepsilon_k \mathbf{v_k}, \ldots \varepsilon_k \in I_\varepsilon \right\} \tag{2b}$$

are fulfilled, where ε_k are integers for k = 1,2,….d, and I_ε is the set of integers [7].

In the following, it is supposed that $\boldsymbol{\Omega}_{\mathbf{d}}$ is a locally finite periodic cellular system (LFPC system). From the previous considerations it follows, that parallelepiped $\mathbf{\Pi_d}$ can be covered by the union of a finite set of cells belonging to $\boldsymbol{\Omega}_{\mathbf{d}}$. This implies that a LFPC system is generated from a finite set of polyhedra of combinatorially different types.

It will be shown that the topological description of a locally finite periodic cellular system (LFPC system) can be traced back to the topological characterization of an appropriately constructed finite cellular system. Parallelepiped $\mathbf{\Pi_d}$ has been chosen in such a way, that it has a minimum volume. It should be emphasized that this parallelepiped $\mathbf{\Pi_d}$ is not uniquely defined. They can be constructed in different manners; however, their common property is that their d-dimensional volumes are identical.

There is no loss in generality in assuming the following: By using an appropriately selected homogenous linear transformation, parallelepiped $\mathbf{\Pi_d}$ can be mapped into a d-dimensional unit cube. This unit cube $\mathbf{\Pi_{d,U}}$ which is called "a unit domain" in the classical crystallography is given by

$$\mathbf{\Pi_{d,U}} = \{\mathbf{x} = (x_1, x_2,\ldots x_k,\ldots x_d) \mid 0 \leq x_k \leq 1 \text{ and } k = 1,2,\ldots d\} \qquad (3)$$

This simple transformation makes it possible to replace the original LFPC system by a "standardized" periodic cellular system generated by translations of $\mathbf{\Pi_{d,U}}$. The only difference is that the standardized LFPC system is composed of unit cubes instead of parallelepipeds. Since a linear transformation represents a "topology preserving" onto-to-one mapping, this implies that the original and the transformed periodic cellular systems are topologically equivalent. In the further investigations, it will be supposed that the LFPC system $\boldsymbol{\Omega_d}$ is a standardized cellular system.

3. Finite cellular systems constructed by using a double toroidal embedding

From a LFPC system, finite cellular systems of a toroidal type can be constructed in several ways. In the following, it will be demonstrated that starting with a d-dimensional LFPC system and by using the so-called double toroidal embedding, it is always possible to construct a uniquely defined finite cellular system represented by a torus in the (d+1) dimensional Euclidean space, which is advantageously applicable to the local topological evaluation of infinite periodic cellular systems.

In order to generate a finite cellular system from a standardized LFPC system, consider a unit domain $\mathbf{\Pi_{d,U}}$ defined by Eq. (3). As a first step, let us construct a so-called identification region $\mathbf{S_d}$, which is composed of 2^d unit domains, as follows

$$\mathbf{S_d} = \{\mathbf{x} = (x_1, x_2,\ldots x_k,\ldots x_d) \mid 0 \leq x_k \leq 2 \text{ and } k = 1,2,\ldots d\} \qquad (4)$$

As can be stated, $\mathbf{S_d}$ is also a d-dimensional cube with edge length of 2. As a second step, let us construct a finite toroidal cellular system $\boldsymbol{R}_\mathbf{d}$ (FTC system) by gluing (identifying) the opposing k-dimensional face pairs (edges, vertices, etc.) of $\mathbf{S_d}$ (k=0,1,2,….d-1).

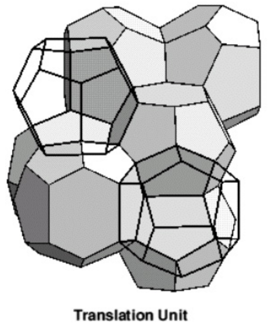


Fig.1 The 3-dimensional, periodic Weaire-Phelan cellular system

This mapping is called the double toroidal embedding (DT embedding) of the d-dimensional LFPC system, because $\boldsymbol{R_d}$ represents a torus in the (d+1)-dimensional Euclidean space. As an example, **Fig. 1** shows a two-component, space-filling periodic polyhedral system. In this 3-dimensional LFPC system that was discovered by Weaire and Phelan, the space-filling unit domain consists of six tetrakaidecahedra (14-sided Goldberg polyhedra) and two irregular pentagonal dodecahedra (12-sided polyhedra) [8].

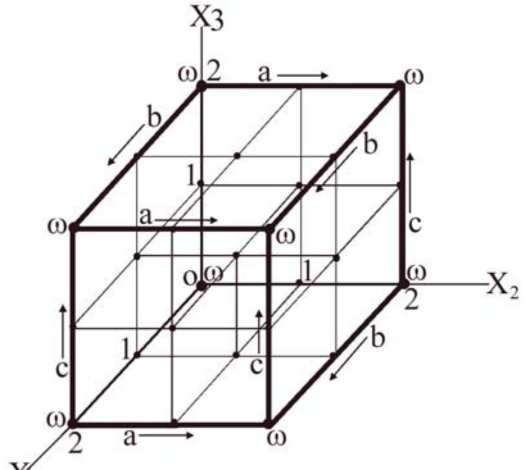


Fig.2 Identification region $\mathbf{S_3}$ generated by 8 unit domains to the DT embedding of a 3-d LFPC system

In **Fig. 2**, the construction of the identification region of a 3-dimensional LFPC system is illustrated. As can be seen, this is the union of 8 unit domains. The arrows a, b and c are used to specify a direction for the edges, and this direction must be respected when gluing is done. The eight vertex points of the identification region **S₃** are joined to form a single point ω of the resulting toroidal system.

Fig. 3 demonstrates the general concept of the DT embedding of a 2-dimensional LFPC system. As an example, in **Fig. 4**., the DT embedding is shown for a 2-d periodic cellular system, which includes 4- and 8-sided polygons. The resulting FTC system is also composed of four 4-sided and four 8-sided cells (See Fig 4.c). The number of cells is 8, the number of edges is 24, and the number of vertices is 16. As it is expected, the Euler-characteristic of this finite system defined on the torus surface is zero.

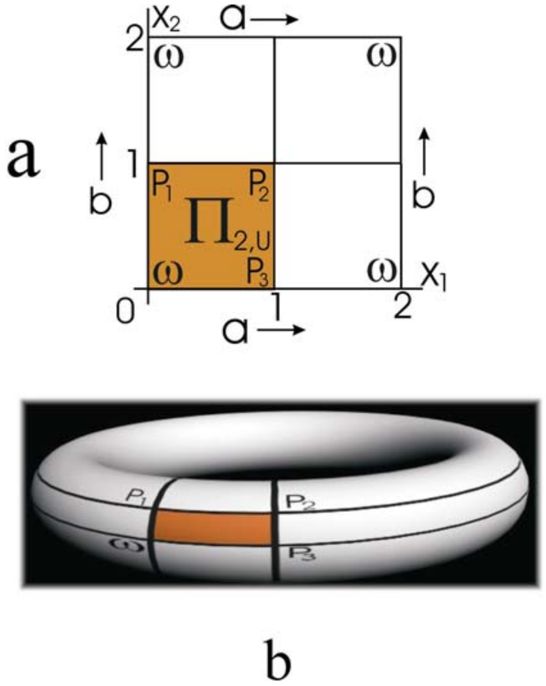


Fig.3 Principle of the DT embedding of a 2-d LFPC system (a) The 2-d identification region **S_2** composed of 4 unit domains, (b) The corresponding FTC system defined on a torus

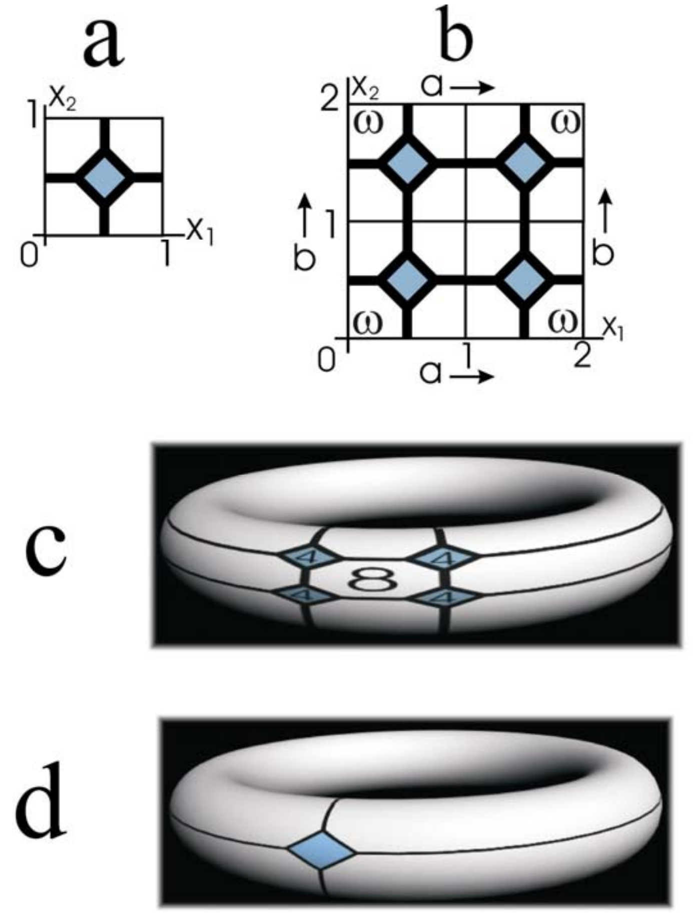


Fig.4 Example of the toroidal embedding of a 2-d LFPC system (a) The unit domain $\mathbf{\Pi_{2,U}}$, (b) The corresponding identification region $\mathbf{S_2}$, (c) The DT embedding performed on the torus, (d) The traditional toroidal embedding of the 2-d LFPC system by using a single unit domain only

The basic properties of the DT embedding and of FTC systems are as follows:

a. The total number of cells in $\mathbf{S_d}$ is equal to $N_{d,U}$ x 2^d, where $N_{d,U}$ denotes the number of cells in the unit domain.

b. The resulting FTC system preserves all the topological properties of the original LFPC system. This is due to the following fact: In cellular systems generated by a DT embedding, each (d-1)-dimensional face is shared by two different neighbor cells exactly. Since a DT embedding is a topology preserving one-to-one mapping between the LFPC and FTC system this implies that an LFPC system can be unambiguously reconstructed from the corresponding FTC system generated by a DT embedding. It should be noted, that using a single unit domain could also perform a toroidal embedding. Unfortunately, in certain cases, if we use only one unit domain to generate a finite toroidal cellular system, as a result of this procedure the local topological properties characterizing the neighborhood structure of cells can change radically. Consequently, the classical toroidal embedding cannot be applicable to every case. As an example, this is illustrated in Fig 4.d. Due to the toroidal embedding with a single unit domain, the original topological structure of the 2-dimensional LFPC system depicted in Fig. 4.a, has been degenerated. In this case, the number of cells is 2, the number of edges is 6, and the number of vertices is 4, the Euler-characteristic is zero. As it can be stated Fig 4.d shows the conventional 2-cell embedding of the complete graph K_4 in the torus, where one of the two cells is 4-sided, while the other is 6-sided. This is explained by the fact that in this toroidal cellular system there exist two edges, which belong to the same 6-sided cell. This implies that the original LFPC system shown in Fig. 4.a, cannot be reconstructed from the finite graph depicted in Fig 4.d.

c. Every FTC system generated by the DT embedding of a 2-dimensional LFPC system can be represented by a finite toroidal graph. (i.e. 2-dimensional FTC systems are considered as a subset of toroidal graphs). This implies that the topological analysis of 2-dimensional LFPC systems can be reduced to the characterization of traditional toroidal graphs embedded on a genus 1 surface.

4. Topological properties of FTC systems

In order to simplify the treatment of problems to be outlined, we introduce some definitions. Let us denote by $N_{d,k}$ the number of k-dimensional faces of FTC system $\boldsymbol{R_d}$ where k=0,1,2,…d. By definition, $N_{d,d}$ is the number of cells (polyhedra), $N_{d,0}$ is the total number of vertices, $N_{d,1}$ is the total number of edges, $N_{d,2}$ is the total number of traditional faces of cells.

The finite toroidal cellular system $\boldsymbol{R_d}$ generated by space filling polyhedra (polytopes) can be represented as

$$\boldsymbol{R_d} = \{A_{n,i} \,|\, i=1,2,..N_n,\ n_{min} \leq n \leq n_{max},\ n \in I_R\}. \qquad (5)$$

In Eq.(5), $A_{n,i}$ denotes the ith cell (d-dimensional polyhedron) with n-faces, where i= 1,2,...N_n and N_n is the total number of d-dimensional, n-faceted cells in $\boldsymbol{R_d}$. By definition, an n-faceted cell stands for a d-dimensional cell having (d-1) dimensional faces of number n. It is supposed that $n_{min} \geq 2$ and $n_{max} > n_{min}$ are positive integers, I_R is a finite index set for n.

The total number $N_{d,d}$ of cells is $N_{d,d} = \sum N_n$ where n=2,3,... n_{max}. The fraction (or frequency) p_n of n-faceted cells is $p_n = N_n/N_{d,d}$, where $p_n>0$. Consequently, $\sum p_n = 1$. For a FTC system, the mean number of (d-1) dimensional faces per cell denoted by $\langle n \rangle$ can be calculated as $\langle n \rangle = \sum np_n$. Generally, in the cell statistics, expression $\langle U(n) \rangle$ is the average value of the quantity U(n) with frequency p_n, i.e. $\langle U(n) \rangle = \sum p_n U(n)$ by definition.

Since any (d-1) dimensional face is a common face of two different neighbor cells, this implies that

$$\langle n \rangle = \sum_n np_n = \frac{2\sum_n nN_n}{N_{d,d}} = \frac{2N_{d,d-1}}{N_{d,d}} \tag{6}$$

and

$$\sum_n \sum_{k \leq n} e_{d-1}(n,k) = \sum_n \sum_{k \geq n} e_{d-1}(n,k) = N_{d,d-1} \tag{7}$$

where $N_{d,d-1}$ is the total number of (d-1) dimensional faces, and $e_{d-1}(n,k)$ is the number of the common (d-1) dimensional faces of n-faceted and k-faceted neighbor cells.

In a FTC system, vertices do not all have the same valency, consequently, we may define an average valency [r] as follows:

$$[r] = \frac{1}{N_{d,0}} \sum_r rV_r^{(d)} \tag{8}$$

where $V_r^{(d)}$ is the number of r-valent vertices in $\boldsymbol{R_d}$ and $N_{d,0} = \sum_r V_r^{(d)}$. For every FTC system we have

$$2N_{d,1} = [r]N_{d,0} = \sum_r rV_r^{(d)} \tag{9}$$

which is due to the fact, that each edge has two different ends (endvertices).

The component number of a FTC system is defined by $\Phi = \sum \mathrm{sgn}(p_n)$. It follows from the definition that $\Phi \geq 1$. On the other hand, $\Phi=1$, if and only if the FTC system is a so-called face-homogenous system which is composed only of polyhedra with identical face numbers. It should be noted that there exist face-homogenous LFPC systems including combinatorially different cells (i.e. topologically non-equivalent polyhedra) with identical face numbers. (The simplest 3-dimensional LFPC system of such type is composed of two combinatorially different 5-sided polyhedra.)

4.1 Euler-equation for FTC systems

It is has been shown, that the traditional Euler-formula can be extended to the topological description of FTC systems [7, 9-11]. This modified Euler-equation, which valid even for a d-dimensional FTC system can be formulated as follows:

For an arbitrary FTC system $\boldsymbol{R}_\mathbf{d}$ where all the k-dimensional faces are topologically equivalent to a k-dimensional sphere, the equality

$$\chi(\mathbf{R_d}) = \sum_{k=0}^{d} (-1)^k N_{d,k} = 0 \tag{10}$$

is valid. In Eq.(10), $\chi(\boldsymbol{R}_\mathbf{d})$ is the Euler-characteristic of the finite toroidal cellular system $\boldsymbol{R}_\mathbf{d}$. Particularly, for the case of d=2, we have

$$N_{2,2} - N_{2,1} + N_{2,0} = 0 \tag{11}$$

while for the case of d=3,

$$-N_{3,3} + N_{3,2} - N_{3,1} + N_{3,0} = 0 \tag{12}$$

yields.

Because the unit domain $\Pi_{\mathbf{d},\mathrm{U}}$ representing the corresponding LFPC system has a minimum volume, it follows that the total numbers of k-dimensional faces in $\boldsymbol{R}_\mathbf{d}$ (k=0,1,2,…d), i.e. quantities $N_{d,k}$ in Eqs. (10-12) are uniquely defined positive integers.

For the 2-dimensional case, identity (11) coincides with Euler's theorem for the torus [7,9]. For the 3-dimensional case, Eq.(12) has been proven by Kinsey [10], who verified that if $\boldsymbol{R}_3$ is a compact connected 3-manifold without a boundary then $\chi(\boldsymbol{R}_3) = 0$. The proof of the general case is based on the following concept: Considering the Euler-characteristic of a d-dimensional torus, we argue as follows: The d-dimensional torus can be represented as the direct product of d circles (meaning d circular arcs). Since the Euler-characteristic is multiplicative with respect to direct products and the Euler-characteristic of a circle is zero, this implies that the Euler-characteristic of a d-dimensional torus is zero, as well. (See Exercise B4 on page 205 in Ref. [11]).

4.2 Cell coronas

The analysis of local topological properties can be traced back to the evaluation of the correspondences between the individual cells and their first neighbor cells. Cells A and B are called adjacent (neighbors) if they have common (d-1) dimensional faces. The cell corona C(A) of a cell A in $\boldsymbol{R}_\mathbf{d}$ is the union of neighbor cells of A. According to this definition, cell A is not included in C(A).

FTC systems can also be characterized on the basis of the topological properties of their cell coronas. For this purpose, we define the corona frequency vector $\mathbf{f_A}$ (CF-vector) of cell A included in $\boldsymbol{R}_\mathbf{d}$ as follows:

$$\mathbf{f_A} = \left(f_2^{(A)}, f_3^{(A)}, ...f_k^{(A)}....f_{n_{max}}^{(A)}\right) \tag{13}$$

where component $f_k^{(A)}$ is the number of (d-1) dimensional, k-faceted cells in C(A), and index n_{max} denotes the maximum (d-1) dimensional face number of cells included in $\boldsymbol{R}_\mathbf{d}$. It is obvious that for any k and $f_k^{(A)}$ relationships $0 \leq f_k^{(A)} \leq n_{max}$ and $n_{min}=2 \leq k \leq n_{max}$ are valid, and $f_k^{(A)} = 0$ if and only if, there is no k-faceted neighbor cell in C(A). It is clear that, if A is an n-faceted cell, then the sum of components of $\mathbf{f_A}$ is equal to n.

Consider two n-faceted cells A_n and B_n characterized by their corresponding CF-vectors denoted by $\mathbf{f_{A,n}}$ and $\mathbf{f_{B,n}}$. Cells A_n and B_n are called topologically similar, if $\mathbf{f_{A,n}} \equiv \mathbf{f_{B,n}}$ is fulfilled. As can be stated, this topological similarity is an equivalence relation by which all the cells of a FTC system can be classified into disjoint subsets.

This implies that all the topologically similar n-faceted cells denoted by $A_{n,j}^{(1)}, A_{n,j}^{(2)}, ..., A_{n,j}^{(R_{n,j})}$ are the elements of the same configuration set $\boldsymbol{R}_\mathbf{n,j}$ defined as

$$\boldsymbol{R}_\mathbf{n,j} = \left\{A_{n,j}^{(1)}, A_{n,j}^{(2)}, ..., A_{n,j}^{(R_{n,j})}\right\} \tag{14}$$

where $R_{n,j}$ is the number of topologically similar n-faceted cells in $\boldsymbol{R}_\mathbf{n,j}$ (j=1,2,…J(n)). It follows that $\boldsymbol{R}_\mathbf{d}$ can be described as a union of disjoint subsets

$$\mathbf{R_d} = \bigcup_{n} \bigcup_{j=1}^{J(n)} \mathbf{R_{n,j}} \tag{15}$$

where J(n) stands for the number of configuration sets including topologically similar, n-faceted cells. It is obvious that cells belonging to $\boldsymbol{R}_\mathbf{n,j}$ is characterized by the same FC-vector $\mathbf{f}_{R_{n,j}}$. Let us denote by $p_{n,j}$ the fraction of topologically similar n-faceted cells included in $\boldsymbol{R}_{n,j}$. Because $p_{n,j}= R_{n,j}/N_{d,d}$ it follows that

$$\sum_{n}\sum_{j=1}^{J(n)} p_{n,j} = \sum_{n} p_{n} = 1 \qquad (16)$$

It is easy to see that $p_{n,j}$ and $\mathbf{f}_{R_{n,j}}$ are unambiguously defined quantities which are independent of the particular choice of the unit domain of the LFPC system. It follows that the total number J of possible configuration sets $\boldsymbol{R}_{n,j}$ can be calculated as

$$J = \sum_{n} J(n) = \sum_{n}\sum_{j} \operatorname{sgn}(p_{n,j}) \qquad (17)$$

It is obvious that for any FTC system inequality J≥Φ is fulfilled. This means that the total number of possible configuration sets is not less than the component number Φ of the FTC system. The quotient φ=Φ/J ≤1 which is called the complexity index of the cellular system, gives information on the fraction of topologically distinct cell coronas in the LFPC and the corresponding FTC system.

4.3 Face-coordination number

The face-coordination number m_A of an arbitrary n-faceted cell A belonging to the configuration set $\boldsymbol{R}_{n,j}$ is defined as

$$m_A = \frac{1}{n}\sum_{k} k f_k^{(R_{n,j})} \qquad (18)$$

where $f_k^{(R_{n,j})}$ is kth component of the CF-vector $\mathbf{f}_{R_{n,j}}$.

The face-coordination number m_A is the mean number of (d-1) dimensional faces of the neighbors of A. It should be emphasized that m_A is a local topological parameter, which gives some information on the arrangement of the cells included in the cell-corona. For the FTC system $\boldsymbol{R}_d$ which is composed of cells $A_{n,i}$ (i=1,2,..N_n), the mean face-coordination number m(n) of n-faceted cells is defined as

$$m(n) = \frac{1}{N_n}\sum_{i=1}^{N_n} m_{A_{n,i}} \qquad (19)$$

where $m_{A_{n,i}}$ is the face coordination number of cell $A_{n,i}$.

Knowing the set of FC-vectors $\mathbf{f}_{\mathbf{R}_{n,j}}$ and the corresponding fractions $p_{n,j}$ of topologically similar cells, the mean face-coordination number m(n) of n-faceted cells can be calculated as

$$m(n) = \frac{1}{np_n} \sum_{j=1}^{J(n)} p_{n,j} \left\{ \sum_k k f_k^{(R_{n,j})} \right\} \tag{20}$$

Starting with Eqs.(19 and 20) we define the total face-coordination number ⟨m(n)⟩ of $\boldsymbol{R}_d$ as follows

$$\langle m(n) \rangle = \sum_n p_n m(n) = \sum_n \frac{1}{n} \sum_{j=1}^{J(n)} p_{n,j} \left\{ \sum_k k f_k^{(R_{n,j})} \right\} \tag{21}$$

It is conjectured that for all FTC systems inequality ⟨m(n)⟩ ≥ ⟨n⟩ holds, and ⟨m(n)⟩ = ⟨n⟩ if and only if, the cellular system is a face-homogenous system including cells with identical face numbers only (i.e. Φ = 1 is fulfilled).

4.4 Neighborhood coefficients

Now, let us define quantities denoted by H(n,k) as

$$H(n,k) = \sum_{j=1}^{J(n)} p_{n,j} f_k^{(R_{n,j})} \tag{22}$$

where $n_{min} \leq n,k \leq n_{max}$. Quantities H(n,k) are called the neighborhood coefficients of the FTC system. The neighborhood coefficients are non-negative numbers, which have a special property of symmetry

$$H(n,k) = \sum_{j=1}^{J(n)} p_{n,j} f_k^{(R_{n,j})} = \sum_{j=1}^{J(k)} p_{k,j} f_n^{(R_{k,j})} = H(k,n) \tag{23}$$

The neighborhood coefficients can be interpreted geometrically as follows:

$$H(n,k) = \begin{cases} \dfrac{1}{N_{d,d}} e_{d-1}(n,k) & \text{if} \quad n \neq k \\ \\ \dfrac{2}{N_{d,d}} e_{d-1}(n,n) & \text{if} \quad n = k \end{cases} \tag{24}$$

where $N_{d,d}$ is the number of d-dimensional cells, and $e_{d-1}(n,k)$ is the number of the common (d-1) dimensional faces of n-faceted and k-faceted neighbor cells.

Especially, for the case of d=2, we have

$$H(n,k)=\begin{cases}\frac{1}{N_{2,2}}e_1(n,k) & \text{if} \quad n\neq k\\ \\ \frac{2}{N_{2,2}}e_1(n,n) & \text{if} \quad n=k\end{cases} \tag{25}$$

where $N_{2,2}$ is the number of 2-dimensional cells (polygons), and $e_1(n,k)$ is the number of the common edges (i.e. 1-dimensional faces) of n-sided and k-sided neighbor cells.

For the case of d=3,

$$H(n,k)=\begin{cases}\frac{1}{N_{3,3}}e_2(n,k) & \text{if} \quad n\neq k\\ \\ \frac{2}{N_{3,3}}e_2(n,n) & \text{if} \quad n=k\end{cases} \tag{26}$$

where $N_{3,3}$ is the number of 3-dimensional cells (polyhedra) and $e_2(n,k)$ is the number of the common 2-dimensional faces of n-faceted and k-faceted neighbor polyhedra.

It is easy to verify, that for quantities H(n,k) the following relationships are valid:

$$p_n=\frac{1}{n}\sum_k H(n,k)=\frac{1}{n}\sum_k H(k,n) \tag{27}$$

$$\sum_n\left\{\sum_k H(n,k)\right\}^z=\sum_n p_n^z n^z \tag{28}$$

$$\sum_n\sum_k k^z H(n,k)=\sum_n n^z\sum_k H(n,k)=\sum_n p_n n^{z+1}=\left\langle n^{z+1}\right\rangle \tag{29}$$

where z is an arbitrary integer. Additionally, from Eq.(6), identity

$$N_{d,d}\sum_n\sum_k H(n,k)=N_{d,d}\langle n\rangle=2N_{d,d-1} \tag{30}$$

yields. For the case of d=2 we have

$$N_{2,2}\sum_{n}\sum_{k}H(n,k) = N_{2,2}\langle n\rangle = 2N_{2,1} = 2\sum_{n}\sum_{k\le n}e_1(n,k) \qquad (31)$$

where $N_{2,1}$ is the total number of edges (i.e. 1-dimensional faces) and ⟨n⟩ is the mean number of edges per cell in $\boldsymbol{R_2}$. For the case of d=3

$$N_{3,3}\sum_{n}\sum_{k}H(n,k) = N_{3,3}\langle n\rangle = 2N_{3,2} = 2\sum_{n}\sum_{k\le n}e_2(n,k) \qquad (32)$$

yields, where $N_{3,2}$ is the total number of 2-d faces and ⟨n⟩ is the mean number of faces per cell in $\boldsymbol{R_3}$. As an example of application of these ideas, we consider the 3-dimensional, periodic Weaire-Phelan cellular system shown in **Fig.1**. This 2-component and 4-valent polyhedral system (i.e. Φ=2, R=4) consisting of 12- and 14-sided polyhedra is characterized by the following topological quantities: p_{12}=1/4, p_{14}=3/4, ⟨n⟩=27/2=13.5, $\langle n^2\rangle$=183, φ=Φ/J=2/3=0.667, H(12,12)=1, H(12,14) = H(14,12)=2 and H(14,14)=17/2=8.5.

5. A fundamental property of LFPC and FTC systems

Neighborhood coefficients H(n,k) play a key role in the topological characterization of LFPC systems. In the following a fundamental property of LFPC and FTC systems will be presented which can be formulated in the following statements:

On the one hand

$$\langle n^{z+1}\rangle = \frac{\langle n\rangle}{2N_{d,d-1}}\sum_{n}\sum_{k\le n}\left(n^z + k^z\right)e_{d-1}(n,k) \qquad (33)$$

on the other hand

$$\langle n^{z+1}\rangle = \langle nm(z,n)\rangle \qquad (34)$$

where z is an arbitrary integer, and

$$m(z,n) = \frac{1}{np_n}\sum_{j=1}^{J(n)}p_{n,j}\left\{\sum_{k}k^z f_k^{(R_{n,j})}\right\} \qquad (35)$$

by definition. As it can be stated, Eq. (35) is the generalization of Eq.(20). As a special case, when z=1, from Eq.(35) we obtain the mean face coordination number m(n) of n-faceted cells which is defined by Eq.(20). Consequently, identity m(n)=m(1,n) is fulfilled.

Proof of Eq. (33) is based on the following concept. Starting with Eqs.(24 and 29), we have

$$
\begin{aligned}
&\langle n^{z+1}\rangle = \sum_{n}\sum_{k} k^{z}H(n,k) = \\
&\sum_{n} n^{z}H(n,n) + \sum_{n\neq k}\left\{\sum_{k} k^{z}H(n,k)\right\} = \\
&\frac{2}{N_{d,d}}\sum_{n} n^{z}e_{d-1}(n,n) + \frac{1}{N_{d,d}}\sum_{n\neq k}\left\{\sum_{k} k^{z}e_{d-1}(n,k)\right\} = \\
&\frac{1}{N_{d,d}}\sum_{n} e_{d-1}(n,n)\left\{n^{z} + n^{z}\right\} + \\
&\frac{1}{N_{d,d}}\left\{\sum_{\substack{n,k \\ n>k}} n^{z}e_{d-1}(n,k) + \sum_{\substack{n,k \\ n>k}} k^{z}e_{d-1}(n,k)\right\} = \\
&\frac{1}{N_{d,d}}\sum_{\substack{n,k \\ n\geq k}} (n^{z} + k^{z})e_{d-1}(n,k)
\end{aligned}
\tag{36}
$$

Additionally, from Eq. (30) it follows

$$
N_{d,d} = \frac{2N_{d,d-1}}{\langle n\rangle} \tag{37}
$$

Substituting Eq.(37) into Eq.(36) we have

$$
\langle n^{z+1}\rangle = \langle n\rangle \sum_{n}\sum_{k\leq n}\frac{e_{d-1}(n,k)}{N_{d,d-1}}\left(\frac{n^{z} + k^{z}}{2}\right) \tag{38}
$$

It is important to note that Eq.(38) can be interpreted geometrically as follows: Let us denote by $q(n,k)=e_{d-1}(n,k)/N_{d,d-1}$ the relative frequency of the common faces of n-faceted and k-faceted neighbor cells, for which $\sum q(n,k) = 1$. Additionally, let us define quantities denoted by $w(z,n,k)= (n^{z} + k^{z})/2$, which are considered as positive weights belonging to the (d-1) dimensional common face of n- and k-faceted neighbor cells. Now, Eq.(38) can be rewritten in the form

$$\langle n^{z+1}\rangle = \langle n\rangle \sum_{n}\sum_{k\leq n} q(n,k)w(z,n,k) \qquad (39)$$

It is worth noting, when z = -1, from Eqs.(37 and 38) the identity

$$N_{d,d} = \sum_{n}\sum_{k\leq n} e_{d-1}(n,k)\left(\frac{1}{n}+\frac{1}{k}\right) \qquad (40)$$

yields.

The second statement represented by Eq.(34) can be proved as follow: By using Eq.(35) and Eq.(23) incorporating the definitions of quantities m(z,n) and H(n,k), we have

$$\langle nm(z,n) = \sum_{n}\sum_{j=1}^{J(n)} p_{n,j}\left\{\sum_{k} k^{z} f_{k}^{(R_{n,j})}\right\} =$$

$$(41)$$

$$\sum_{n}\left\{\sum_{k} k^{z}\left\{\sum_{j=1}^{J(n)} p_{n,j} f_{k}^{(R_{n,j})}\right\}\right\} = \sum_{n}\left\{\sum_{k} k^{z} H(n,k)\right\} = \langle n^{z+1}\rangle$$

In particular cases, when z =1, from Eq.(41) we obtain the well-known identity formulated as $\langle nm(n)\rangle = \langle n^2\rangle$ which were considered by Weaire and Fortes [2,3] for random 2-d cellular systems, and by Fortes for random 3-d cellular systems [4].

In the following it will be demonstrated that the general concept used for the topological description of locally finite periodical cellular systems can be efficiently applicable to the structural characterization of fullerene solids represented by finite cellular surface systems.

6. Cellular fullerenes

The discovery of fullerene molecules and related forms of carbon such as nanotubes has generated an explosion of activity in physics, chemistry and material science. As it is known, the topological properties of fullerenes play a key role in a classification of possible fullerene structures and in predicting their various physical and chemical behaviors.

In chemistry, the traditional definition is that a fullerene is an all-carbon molecule in which the atoms are arranged on a pseudospherical framework made up entirely of hexagons and pentagons. Based on the concept outlined in Refs. [12-14], define a fullerene in the wider sense as follows: A fullerene is considered as a simple finite cellular system (SFC system) defined on an unbounded, closed and oriented surface, and composed of a finite set of combinatorial polygons (called cells), where cells are simply connected regions and all common edges are shared only by two different neighbor cells. Fullerenes of such types will be referred to as cellular fullerenes. According to this general definition, the closed nonotubes with negative curvature and the so-called onion-like structures are also considered as fullerenes [13,14]. Taking into consideration the decisive role of the Euler characteristic (χ) in the topological analysis of unbounded, closed and oriented surfaces, cellular fullerenes with χ=2 are called spherical, while cellular fullerenes with χ=0 are called toroidal fullerenes.

6.1 Non-polyhedral spherical fullerenes

Cellular spherical fullerenes generated by the tessellation of the surface of the unit sphere can be classified into two classes. Spherical fullerenes represented by convex polyhedra are called polyhedral fullerenes, while the others, which cannot be represented by convex polyhedra, are called non-polyhedral fullerenes. It should be noted that, the Schlegel diagram of a polyhedral fullerene is considered as a polyhedral graph. According to the Steinitz's theorem, a finite graph is polyhedral if and only if it is planar and 3-connected [16]. This implies that the Schlegel diagram of a non-polyhedral spherical fullerene is represented by a 2-connected graph. It is important to note, that among the spherical fullerenes there exist several topologically distinct isomers, which can be of polyhedral and non-polyhedral types.

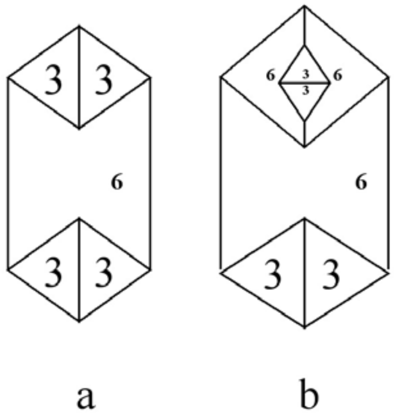


Fig.5 Schlegel diagrams of non-polyhedral, trivalent fullerenes: (a) a generalized triangular $\mathbf{T}_{8,is}$ fullerene with 8 vertices, (b) a generalized triangular $\mathbf{T}_{12,is}$ fullerene with 12 vertices

Deza et al. investigated some special types of polyhedral fullerenes, namely, the so-called triangular fullerenes composed of triangles and hexagons only [15]. A triangular fullerene is defined as a simple polyhedron with trivalent vertices, for which the k vertices are arranged in 4 triangles and (k/2-2) hexagons (and 3k/2 edges). Triangular fullerenes denoted by $\mathbf{T}_k$ can be constructed for all k≡0 (mod 4) except k=8. Examples of such polyhedra are the tetrahedron $\mathbf{T}_4$, the truncated tetrahedron $\mathbf{T}_{12}$ and the chamfered tetrahedron $\mathbf{T}_{16}$ [15]. By extending the definition of triangular fullerenes we can construct spherical and trivalent isomers, which are of non-polyhedral types. In **Fig.5** the corresponding Schlegel diagrams of two non-polyhedral triangular fullerenes denoted by $\mathbf{T}_{8,is}$ and $\mathbf{T}_{12,is}$ are shown. It is worth noting that $\mathbf{T}_{8,is}$ is the smallest non-polyhedral trivalent fullerene because $\mathbf{T}_{8,is}$ has no isomers of polyhedral types.

6.2 Global topological properties of cellular fullerenes

In a SFC system, the total number E of edges is related to the total number N_t of cells, the number V of vertices, the average valency [r] and the mean number of sides per cell ⟨n⟩

$$2E = \sum_n nN_n = \langle n\rangle N_t = \sum_r rV_r = [r]V \qquad (42)$$

where N_n is the number of n-sided polygons. The number V of vertices is $V = \sum V_r$ where V_r is the number of r-valent vertices, the total number N_t of cells (polygons) is N$_t$= $\sum N_n$ where n=2,3,... n$_{max}$, and the fraction p$_n$ of n-sided cells is p$_n$=N$_n$/N$_t$, for p$_n$>0.

Assuming that all the cells are simply connected regions, Euler's equation can be formulated as

$$\chi = N_t - E + V = 2 - 2g \tag{43}$$

where χ is the Euler-characteristic, g is the genus of the surface [16]. The genus of the surface, which can be an arbitrary non-negative integer, is identical to the number of handles that are attached to the sphere to obtain a surface. For the sphere χ=2 and g=0, for the torus (donut) χ= 0 and g=1, for the double torus, χ = - 2 and g=2, for the triple torus, χ = -4 and g=3, respectively. It is worth noting that identity (43) is a possible generalization of Eq.(11), because for the torus (where equalities χ=0 and g=1 are fulfilled), from Eq.(43) we obtain Eq.(11) as a special case. Taking into consideration, that for a SFC system equalities $\langle n\rangle$ = 2E/N$_t$ and [r]= 2E/V are fulfilled, from Eqs.(42 and 43), we have

$$\frac{1}{[r]} + \frac{1}{\langle n\rangle} = \frac{1}{2} + \frac{1-g}{E} = \frac{1}{2} + \frac{\chi}{2E}\,. \tag{44}$$

Additionally, from Eqs. (42 - 44) it follows

$$\langle n\rangle = \frac{2[r]}{[r]-2}\left\{1 - \frac{\chi}{N_t}\right\} = 2\frac{1-\dfrac{\chi}{N_t}}{1-\dfrac{V}{E}} \tag{45}$$

$$N_t = \chi + V([r]-2)/2 \tag{46}$$

and

$$\frac{1}{V}\sum_r (r-2)V_r = \sum_r (r-2)u_r = 2\frac{N_t-\chi}{V} = 2\frac{E-V}{V} \tag{47}$$

where u$_r$ =V$_r$/V is the fraction of r-valent vertices, for which $\sum u_r = 1$ and $\sum ru_r = [r]$ are fulfilled. An immediate consequence of Eq.(45) is that for trivalent SFC systems, equality $\langle n\rangle$=6(1-χ/N$_t$) = 6-12χ/(2χ+V) is valid. Depending on the particular choice of the Euler–characteristic χ, as particular cases, we get $\langle n\rangle$<6 for a sphere (case χ=2), $\langle n\rangle$=6 for a torus (case χ=0) and $\langle n\rangle$>6 for a double torus (case χ=-2), respectively.

6.3 Local topological properties of cellular fullerenes

First of all, it is important to emphasize that general formulae derived for d-dimensional LFPC systems (see formulae given by Eqs.(6 - 41)) can be applied to arbitrary fullerenes represented by simple finite cellular systems. It is easy to see that the fundamental identities given by Eqs.(33 and 34) remain valid for any SFC system (i.e. for cellular fullerenes).

In the following, by introducing the notion of the so-called vertex corona, we will demonstrate that the vertex corona distribution can be efficiently used to the local topological characterization of regular (R-valent) cellular fullerenes represented by SFC systems.

In a SFC system, cells A and B are called diagonally adjacent (diagonal neighbors) if they have a common vertex X. Vertex corona $C_V(X)$ of an arbitrary vertex X is defined as a union of diagonal neighbor cells having a common vertex X. For a finite cellular system characterized by the sequence of vertices X_k (k =1,2, …V)

$$C_V(X_k) = \bigcup_j A_j(X_k) \tag{48}$$

where X_k is a common vertex of cells $A_j(X_k)$.

Vertex corona distribution of a cellular fullerene represented by a regular SFC system has an interesting property. Consider a finite cellular system, where $X_{r,k}$ denotes the kth and r-valent vertex, and define the topological quantity

$$M_v = \frac{1}{V}\sum_{k=1}^{V} M(r,k) \tag{49}$$

where M(r,k) stands for the mean number of sides of cells in $C_V(X_{r,k})$ for k = 1,2,…V. The local topological parameter M(r,k) is called the vertex coordination number of $X_{r,k}$ while M_v is said to be the total vertex coordination number of the FTC system.

Now, we will verify that for regular, R-valent FTC systems (i.e. cellular fullerenes), identity

$$\langle n^2 \rangle = \langle n \rangle M_v \tag{50}$$

is valid. Proof of Eq. (50) is based on the following considerations: Let us denote by $n(B_j^{(k)})$ the side number of cell $B_j^{(k)}$ belonging to the vertex corona $C_V(X_{R,k})$, where k = 1,2,…V and j=1,2,…R. It follows that

$$M(R,k) = \frac{1}{R}\left\{n(B_1^{(k)}) + n(B_2^{(k)}) + ...n(B_j^{(k)}),... + n(B_R^{(k)})\right\} \tag{51}$$

Starting with Eq.(51), we have

$$M_v = \frac{1}{V}\sum_{k=1}^{V} M(R,k) = \frac{1}{VR}\sum_{k=1}^{V}\sum_{j=1}^{R} n(B_j^{(k)}) = \frac{1}{VR}\sum_{n} n^2 N_n =$$

$$= \frac{N_t}{VR}\sum_{n} n^2 p_n = \frac{N_t}{2E}\langle n^2\rangle = \frac{\langle n^2\rangle}{\langle n\rangle} \qquad (52)$$

It is easy to see that formula (50) can be generalized as follows: If z is an arbitrary integer, then identity

$$\langle n^{z+1}\rangle = \langle n\rangle M_v(z) \qquad (53)$$

is valid for regular SFC systems, where

$$M_v(z) = \frac{1}{V}\sum_{k=1}^{V} M(R,k,z) \qquad (54)$$

and

$$M(R,k,z) = \frac{1}{R}\left\{n^z(B_1^{(k)}) + n^z(B_2^{(k)})+,... + n^z(B_R^{(k)})\right\} \qquad (55)$$

by definition. As it can be stated, when z = 1, this implies that Eq. (53) is simplified to Eq. (50). This is due to the fact, in the case of z=1, it follows that M(R,k,1)= M(R,k) and $M_V(1) = M_V$, respectively.

For trivalent and 2-component cellular fullerenes, (i.e. for the case of Φ=2 and R=3), which include α-sided and β-sided cells with frequencies p_α and p_β = 1- p_α (where α<β), there exist only four possible types of vertex coronas denoted by $C_{\alpha,\alpha,\alpha}$, $C_{\alpha,\alpha,\beta}$, $C_{\alpha,\beta,\beta}$ and $C_{\beta,\beta,\beta}$ respectively. This implies that in this particular case, Eq.(50) can be reduced to the form

$$M_v = s_1M_1 + s_2M_2 + s_3M_3 + s_4M_4 = \frac{\langle n^2\rangle}{\langle n\rangle} \qquad (56)$$

where M_1=(α+α+α)/3, M_2=(α+α+β)/3, M_3=(α+β+β)/3 and M_4 =(β+β+β)/3 are the vertex coordination numbers of the four possible types of vertex coronas, and s_i (i=1,2,3,4) are the corresponding relative fractions of the vertex coronas, for which $\sum s_i = 1$ holds. Using identity (56) facilitates the computation of vertex fractions s_i (i=1,2,3,4), which are topological invariants. (For example, if quantities s_1 and s_2 are known, then s_3 and s_4 can be directly calculated.)

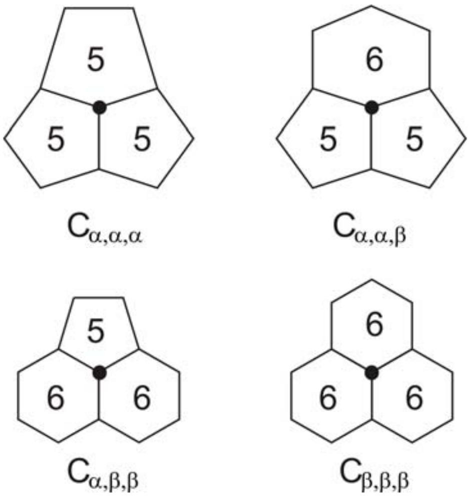


Fig.6 The four possible vertex coronas in a 2-component, trivalent fullerene (case α=5 and β=6).

6.4 Application: Topological characterization of C_{60} fullerenes

As we have mentioned previously, the analysis of the distribution of vertex coronas of different types plays a significant role in algorithms for the perception and classification of topological properties of fullerenes. Balaban et al. have used this concept in order to classify the C_{60} isomers on the basis of topological properties of their vertex coronas [17]. Since C_{60} fullerene isomers are composed of 12 pentagons and 20 hexagons, in this particular case, we have: α=5, β=6, and p_5=12/32, p_6=20/32, $\langle n\rangle$=45/8=5.625, $\langle n^2\rangle$=255/8 and M_V=17/3=5.667. The corresponding vertex coordination numbers are: M_1=15/3, M_2=16/3, M_3=17/3 and M_4 =18/3. As Balaban et al. [17] pointed out, the 1812 structural isomers of C_{60} fullerenes could be partitioned into 42 equivalence classes (subclasses) on the basis of the four types of vertex coronas $C_{5,5,5}$, $C_{5,5,6}$, $C_{5,6,6}$ and $C_{6,6,6}$ which are illustrated in **Fig. 6**.

To characterize the local topological structure of cellular fullerenes, we defined the topological descriptor IS calculated on the basis of the neighborhood coefficients:

$$IS = \sum_{n} H(n,n) \qquad (57)$$

The topological descriptor IS which is called the isolation index can be simply computed for C_{60} isomers

$$IS = \langle n \rangle - 2H(5,6) = \langle n \rangle - \frac{2V}{N_t}\left(1 - s_1 - s_4\right) = \frac{45}{8} - \frac{15}{4}\left(s_2 + s_3\right) \geq \frac{15}{8} \quad (58)$$

where s_i (i=1,2,3,4) are the relative fractions of the corresponding vertex coronas. (See **Fig.6**.) Based on the calculated results the following conclusions can be drawn:

By using the isolation index the 1812 C_{60} isomers can be partitioned into 18 subclasses. Calculated values of isolation index IS are in the interval 1.875 to 4.375.

The buckminster-fullerene denoted by C60B (containing 12 isolated pentagons) is the sole isomer which is characterized by the minimum value of IS (namely IS=15/8 =1.875). The computed neighborhood coefficients are: H(5,5) = 0, H(5,6) = H(6,5) = H(6,6) =15/8). It should be noted that it is supposed that fullerene structures with isolated pentagons are likely to be more stable than structures containing fused five-membered rings [18].

On the other hand, we found that the maximum value of IS belongs to C60W isomer (IS= 4.375). (See the corresponding Schlegel diagram of C60W shown in Fig.1 in Ref.[17]). It is important to emphasize that C60W is judged to be the least stable C_{60} isomer [17], for which the corresponding neighborhood coefficients are: H(5,5) = 10/8, H(5,6) = H(6,5) = 5/8 and H(6,6) =25/8.

We have also observed that the discriminating performance of the topological index IS is determined (and limited) primarily by the local neighborhood structure of the cellular system. For cellular systems characterized by a topologically similar first neighbor structure, the neighborhood dependent isolation index has only a limited ability for discrimination. The main advantage of using the isolation index lies in the fact that IS can be generally applied to the topological characterization of any cellular system, not only fullerene-like but also arbitrary infinite periodical cellular structures.

7. Summary and conclusions

A general method has been developed to characterize and compare infinite and finite cellular systems on the basis of quantitative topological criteria. First, we analyzed the global and local topological properties of infinite periodic cellular structures, and then the theoretical results obtained have been adapted to the local topological characterization of 2-dimensional finite cellular surface systems. The general concept of this new approach is based on the use of the so-called double toroidal embedding (DT embedding) by which a finite cellular system defined on a torus can be generated from an infinite periodic cellular system.

As a result of performing a DT embedding, so-called neighborhood coefficients can be generated. The neighborhood coefficients H(n,k) are simple scalar topological invariants, by which the local topological structure of cellular systems can be quantitatively evaluated and compared. Moreover, by investigating the relationship between the neighborhood coefficients and other local topological quantities, we have verified that the validity of the Weaire-Fortes identity (playing a key role in the topological description of 2-dimensional random cellular patterns), could be extended to infinite periodic cellular systems and 2-d finite cellular surface systems (i.e. generalized fullerene-like structures). It has been also shown that the traditional definition of fullerenes can be generalized by introducing the notion of the cellular fullerene, which is considered as a finite cellular system defined on a 2-d unbounded, closed and orientable surface.

From the previous considerations it follows that the fundamental Eqs. (33 and 34) remain valid not only for cellular systems consisting of combinatorial polyhedra (which are topologically equivalent to a d-dimensional ball), but

- for finite cellular systems defined on an unbounded, closed and orientable surface (sphere, torus, double torus , etc.),
- for infinite triply periodic 3-d surface systems, in which the internal surface represented by “infinite tunnels” is composed of polygons [19]. (Typical examples are the so-called zeolitic structures [20]),
- for all “pseudo-random” cellular systems which are artificially generated by the tessellation of the d-dimensional unit cube using periodic boundary conditions. Due to the periodic boundary extension, these pseudo-random structures are also considered as infinite periodic cellular systems. A well-known example is the computer simulation of the Poisson Voronoi cells where the periodic boundary condition is used to avoid edge effects [1, 21].

Finally, it should be emphasized that Eqs. (33 and 34) remain valid for such cases when the space-filling polyhedra are not equivalent topologically to d-dimensional balls, provided that the cellular system is generated from a finite set of d-dimensional cells with (d-1)-dimensional faces in such a way that all common faces are shared by two different neighboring cells.

Acknowledgements

We would like to thank Prof. A. Fortes (Instituto Superior Técnico, Lisboa) for interesting discussions, and I. Vissai (Budapest Polytechnic, Budapest) for extensive help with computer graphics.

References

[1] D. Weaire and M. Rivier: Soap, Cells and Statistics – Random Pattern in Two Dimensions, Contemp. Phys. Vol. 25 (1984) p. 59-99.

[2] D. Weaire: Some Remarks on the Arrangement of Grains in a Polycrystal, Metallography, Vol.7 (1974) p. 157-160.

[3] M.A. Fortes and P.N. Andrade: The arrangement of cells in 3- and 4-regular planar networks formed by random straight lines, J. Phys. France Vol.50 (1989) p. 717-724.

[4] M.A. Fortes: Applicability of Aboav's rule to a three-dimensional Voronoi partition, Phil. Mag. Lett. Vol. 68, (1993) p. 69-71.

[5] L. Oger, A. Gervois, J.P. Tradec and N. Rivier: Voronoi tessellation of packings of sphere: topological correlation and statistics, Phil. Mag, B, Vol.74 (1996) p. 177-197.

[6] A.G Evans, J.W. Hutchinson, N.A Fleck, M.F. Ashby and H.N.G. Wadley: The topological design of multifunctional cellular metals, Progress in Materials Science, Vol. 46, (2001) p. 309-327.

[7] B. Grünbaum and G.C. Shephard: Tiling and Patterns, Freeman, New York, 1985.

[8] D. Weaire and R. Phelan: A Counter-Example to Kelvin's Conjecture on Minimal Surfaces. Phil. Mag. Let., Vol.69, (1994) p. 107-110.

[9] H.S.M. Coxeter: Regular Polytopes, Macmillen, New York, 1963, p. 72-73.

[10] L.C. Kinsey: Topology of Surfaces, Springer Verlag, New York, 1991, p. 217-219.

[11] E.H. Spanier: Algebraic Topology, McGraw-Hill Series in Higher Mathematics. New York, 1966, p. 205.

[12] M. Deza, P.W. Fowler, A. Rassat and K.M. Rogers: Fullerenes as Tiling of Surfaces, J. Chem. Inf. Comput. Sci., Vol. 40, (2000) p. 550-558.

[13] H. Terrones and M. Terrones: Fullerenes and Nanotubes with Non-positive Gaussian Curvature, Carbon, Vol. 36, (1998) p. 725-730.

[14] R.B. King: Novel highly symmetrical trivalent graphs which lead to negative curvature carbon and boron nitride chemical structures, Discrete Mathematics, Vol. 244, (2002) p. 203-210.

[15] A. Deza, M. Deza and V. Grishukhin: Fullernes and coordination polyhedra versus half-cube embeddings, Discrete Mathematics, Vol. 192, (1998) p. 41-80.

[16] A. T. White and L.W. Beineke: Topological Graph Theory, in Selected Topics in Graph Theory, Ed. by L.W. Beineke and R.J. Wilson, Academic Press, London, 1978.

[17] A. T. Balaban, X. Liu, D.J. Klein, D. Babics, T.G. Schmalz, W.A. Seitz and M. Randic: Graph Invariants for Fullerenes, J. Chem. Inf. Comput. Sci., Vol. 35, (1995) p. 396-404.

[18] D. E. Manolopoulos and P. W. Fowler: Molecular graphs, point groups, and fullerenes, J. Chem. Phys. Vol. 96, (1992) p. 7603-7614.

[19] S.T. Hyde: The topology and geometry of infinite periodic surfaces, Z. Kristallogr. Vol. 187 (1989) p. 165-185.

[20] A.F. Wells: Structural Inorganic Chemistry, Calendron Press, Oxford, 1975, p. 828-831.

[21] A. Okabe, B. Boots, K. Sugihara and S.N. Chou: Spatial Tessellations Concepts and Applications to Voronoi diagrams, 2nd Ed., John Wiley and Sons, Chichester, 2000, p. 306-311.

Grafting of Industrial Cellulose Pulp with Vinyl acetate Monomer by Ceric Ion Redox System as Initiator

Éva Borbély,* József Erdélyi**

* associated professor of Budapest Polytechnic, Dept. of Packaging and Paper Technology
**professor of Budapest Polytechnic, Dept. of Packaging and Paper Technology
H-1034 Budapest, Doberdo 6.

Abstract: The Ce^{4+}/ Ce^{3+} redox system was studied to initiate the grafting of idustrial cellulose pulp with vinyl acetate monomer. The parameters of the copolymerization reaction (reaction time, temperature, monomer and initiator contentration, freeness, chemical composition of the cellulose) were investigated and their effects are discussed.

*Keywords**:** cellulose, copolymerization, vinyl acetate, ceric ion redox system, parameters.*

1. Introduction

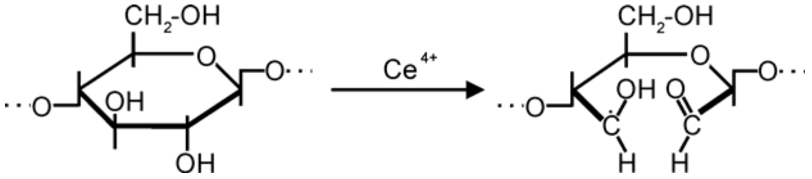
The chemical modification of cellulose by graft copolymerization has generated interest among researches because few comonomer molecules change significantly a number of characteristics of the original natural polymer. Thus new areas of the application might be opened for the modified cellulose-type materials.

Being a polymer itself, cellulose can be copolymerized only with block or grafting procedures. The block copolymerization of cellulose essentially modifies its physical structure, and therefore cannot be used. A graft copolymer is a system comprised of a backbone material to which a second polymer is attached at intervals along the chain. The pulp cellulose modified by grafting copolymerization does not change its fibrous structure, which is very advantageous for further use.

Most of graft copolymerization examined so far [1,2,3,4,5,6] described the use of cotton or regenerated cellulose as the substrate, and there are only few papers about the grafting of cellulose pulp used in the paper industry [7]

Recently we examined the grafting of industrial cellulose with vinyl acetate monomer initiated with cerium(IV)salts [8]

The rather complex action of Ce^{4+}on cellulose can be formulated so:



(1)

2. Experimental

The graft copolymerization was carried out in a 500 cm^3, three-necked, round flask equipped with a mechanical stirrer, a reflux condenser and a thermometer, which was immersed in a thermostat water bath. First the definite amount of cellulose was pulped for 30 minutes, then the required concentrations of cerium-ammonium-sulphate initiator solution and vinyl acetate monomer were added sucessively to the reaction system. At the end of the reaction time we stopped the reaction by L-ascorbic acid.

After the completion of the reaction, the rough products were first precipiated in an excess of acetone and then separated by filtration. To obtain the pure graft copolymer, we used carbon-tetrachloride to extract the homopolymer that might be produced during the polymerization. Extracting for 6 hours was sufficient to remove the polyvinyl acetate homopolymer.

Then the graft copolymer was dried in the air to a constant weight. On the basis of gravimetric measurements, the grafting parameters were determined as follows:

$$G = \frac{W_p - W_o}{W_o} \cdot 100 \quad (\%) \qquad (2)$$

$$GM = \frac{W_p - W_o}{M} \cdot 100 \quad (\%) \qquad (3)$$

where *G* is the grafting percentage (2), and *GM* the grafting conversion of the monomer (3). $W_{p,}$ W_o are the weights of the purified graft copolymer and the cellulose, *M* is the weight of the vinyl acetate monomer.

The applicability of grafted copolymers in the paper industry depends strongly on the percent grafting (G%) reached. If this G% is too small, the properties of the paper are not improved sufficiently, if it is too big, the pulp becomes unsuitable for paper making.

2.1. Effect of temperature and reaction time on grafting efficiency

As the rate of chemical reactions can be regulated with variation of temperature, experiments indicated that an inccrease in the temperature had a very strong effect on G % (Fig.1) At 20 °C (293 K) there was practically no reaction. Increase the temperature improved G % dramatically, but the difference between 50 and 60 °C (323 and 333 K) was only a few percent. This latter can be explaned by the low boiling point (72,3 °C) of the monomer. From these results, 50 °C (323 K) can be proposed for the grafting temperature of this system.

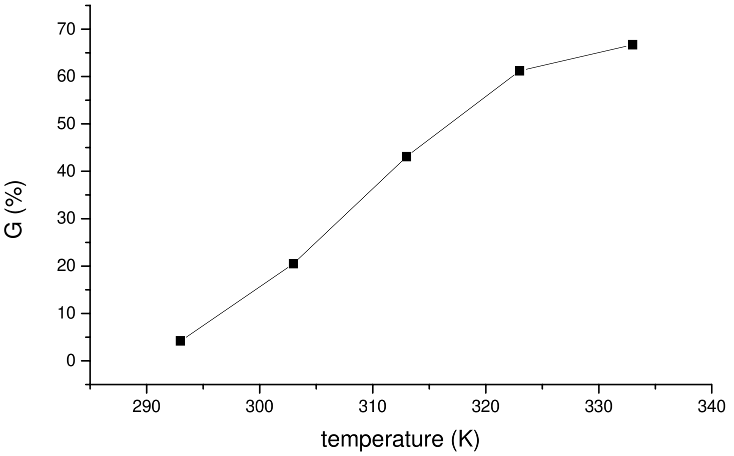


Figure 1. Effect of temperature on grafting efficiency

As to the time dependence of the grafting reaction, G % increased rapidly in the first 40 minutes of the process (Fig.2.). Longer durations did not significantly improve the efficiency of grafting.

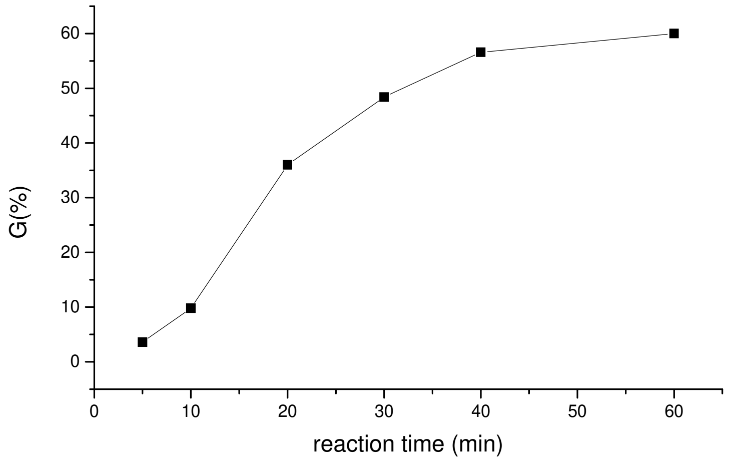


Figure 2. Effect of reaction time on grafting efficiency

2.2. Effect of initiator concentration on grafting efficiency

Graft copolymerization was studied at various cerium-ammonium-sulphate initiator concentration (10^{-3} - 5.10^{-3} mól/dm^3) (Fig.3.).

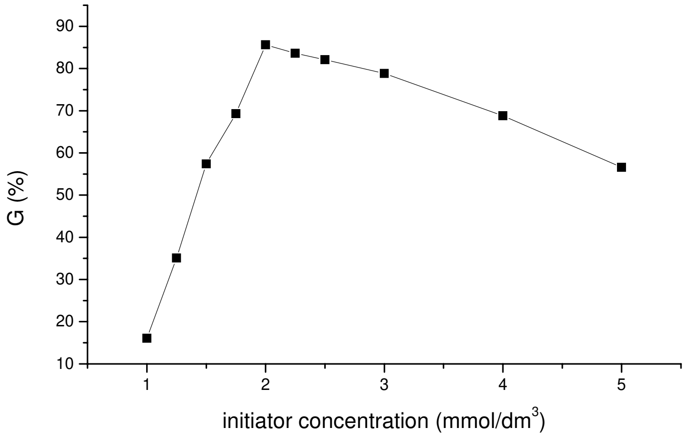


Figure 3. Effect of initiator concentration on grafting efficiency

It has been observed that G% increases on increasing the initiator concentration up to 2.10^{-3} mól/dm^3, beyond which it decreases. The increase of percent grafting with increasing initiator concentration may be ascribed to the increase of active sites on the backbone of the cellulose fiber. The retarding effect of G% with initiator concentration beyond 2.10^{-3} mól/dm^3 may be due to predominancy of homopolymerization over grafting and termination of growing grafted chains by excess of primary radicals formed from the initiator. From these results 2.10^{-3} mól/dm^3 cerium-ammonium-sulphate can be proposed for the optimal grafting efficiency.

2.3. Effect of monomer concentration on grafting efficiency

G% depends, first of all, on the applied monomer concentration. The change in G% as the function of the monomer amount at a fixed initiator concentration (2.10^{-3} mol/dm^3) is shown by a curve with a maximum peak (Fig.4.). This maximum occured because over a given concentration the termination is preferred reaction among other reactions. 1 mol/dm^3 vinyl-acetate concentration can be proposed for the adequate grafting efficiency.

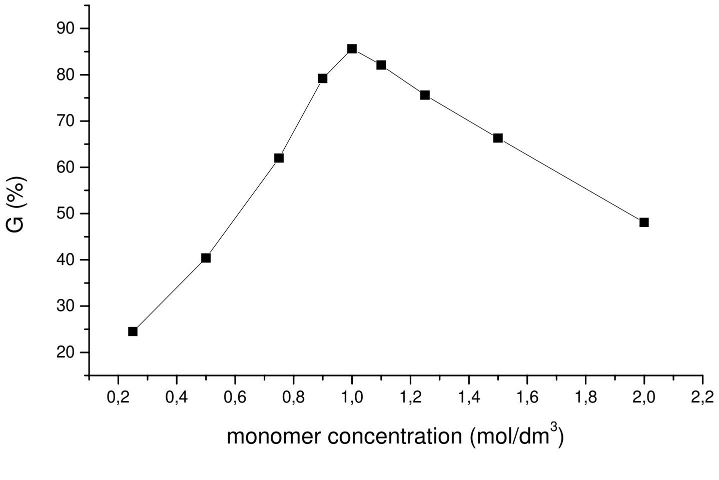


Figure 4. Effect of monomer concentration on grafting efficiency

2.4. Effect of freeness on grafting efficiency

Graft copolymerization occured on the surface of the pulp, so the efficiency of grafting depends strongly on this surface. Increasing the pulping time the freeness and the externel specific surface of the cellulose is increased (Table I. and Fig.5.)

Pulping time (min)	Freeness (SR^o)	External specific surface (m^2/g)
0	15	1,056
10	20	1,472
20	22	2,112
30	27	2,752
40	34	5,632
50	44	8,512
60	56	11,232

Table I. Effect of the pulping time and freeness on the specific surface of the cellulose

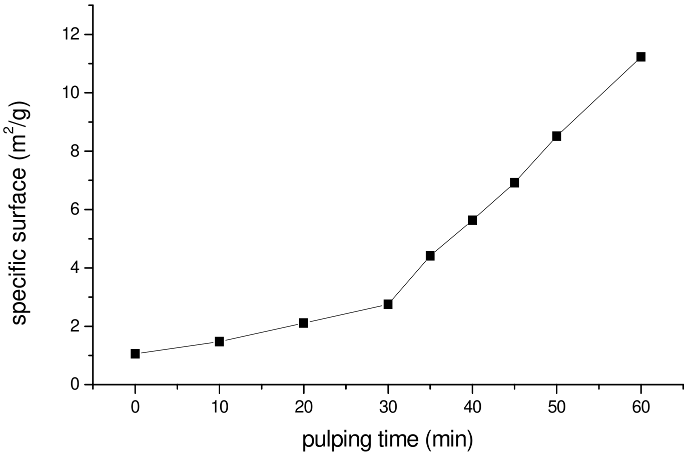


Figure 5. Effect of pulping time on the specific surface of the cellulose

The change in G% with increasing the specific surface of the cellulose is shown by a curve with a maximum peak (Fig.6.).

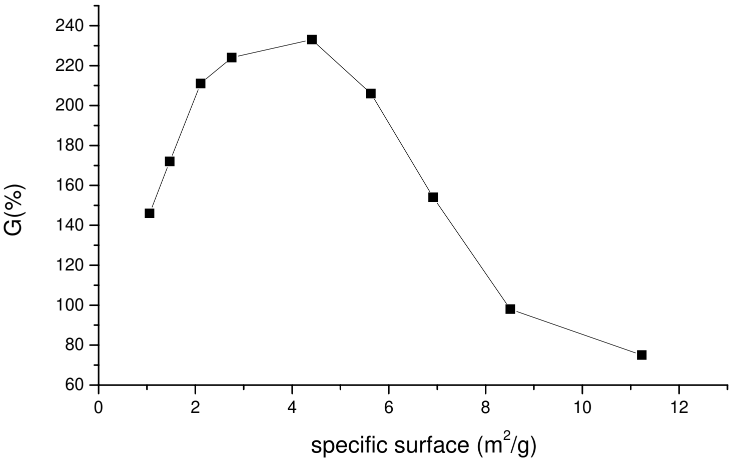


Figure 6. Effect of specific surface on grafting efficiency

Increasing the specific surface of the cellulose increased G %, but after 5 m^2/g grafting efficiency decreased. This latter can be explaned by the decrease of the amount of radicals formed on the same surface at a fixed initiator concentration. So the best grafting efficiency can be reached with 35 min pulping time, when the the specific surface of the cellulose was 4,4 m^2/g.

2.5. Effect of chemical composition of the cellulose

The chemical composition of idustrial cellulose pulp has also a strong effect on grafting efficiency. Although G% depends, first of all, on the lignin content of the applied cellulose pulp only a few papers can be found about the investigation of this effect [9].

To decrease the lignin content of an unbleached cellulose we treated the samples in 5 steps with sodium hypochlorite for 5, 10, 30, 60 and 90 minutes. The change of lignin content during the bleaching time - measured with Kőnig-Komarov method - is shown in Table II.

Bleaching time(min)	Lignin content(%)
0	12,0
5	10,3
15	5,8
30	3,5

60	2,1
90	0,5

Table II. Effect of bleaching time on lignin content of the cellulose

After the bleaching treatment the cellulose samples were grafted with vinyl acetate for 40 minutes at three different temperatures applying the adequate monomer and initiator concentration (Fig 7.).

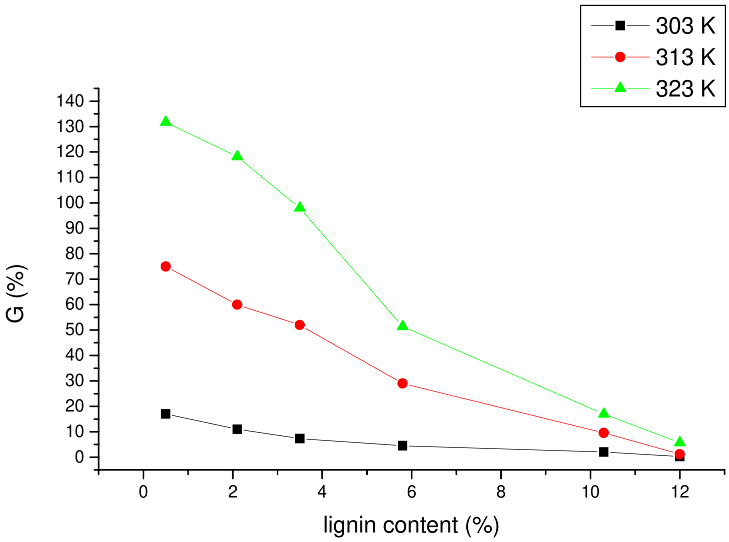


Figure 7. Effect of lignin content of the cellulose on grafting efficiency

Increasing the lignin content of the cellulose decreased G % and after 12 % lignin content the grafting efficiency becomes 0 %. This latter can be explaned by the inhibitor role of lignin in the grafting reaction /9/. The industrial cellulose which is suitable for grafting copolimerization may contain maximum 2 % of lignin.

To investigate the induction period of grafting reaction causing by the presence of lignin the cellulose sample which has the maximum lignin content (12%) was grafted with vinyl acetate for 20, 40, 60, 70, 80, 100 and 120 minutes (Fig. 8.)

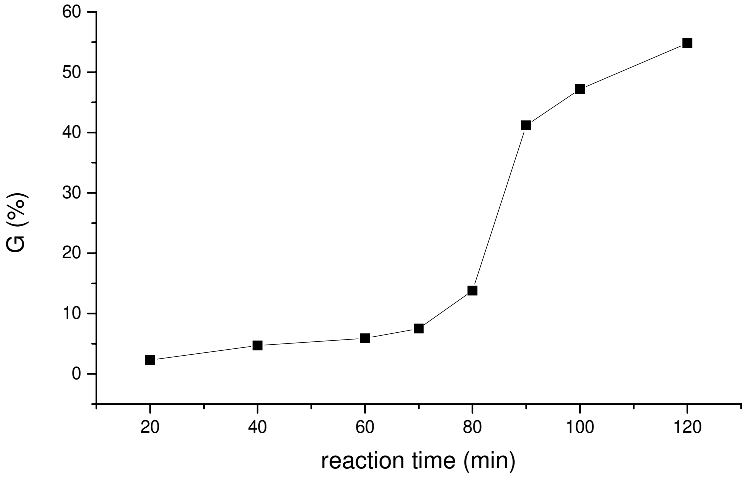


Figure 8. The induction period of grafting reaction causing by lignin

Until 70 minutes there was practically no reaction, but beyond 80 minutes reaction time increased G % dramatically, so the induction period of grafting reaction in this case is seen to be between 70-80 minutes. This is in agreement with the results of other authors investigated the same problem [7].

The effect of hemicellulose content of the cellulose pulp was also investigated, but as it is shown in Figure 9. the change in grafting efficiency causing by the different hemicellulose content of the samples not so significant as the effest of lignin content of the same sample.

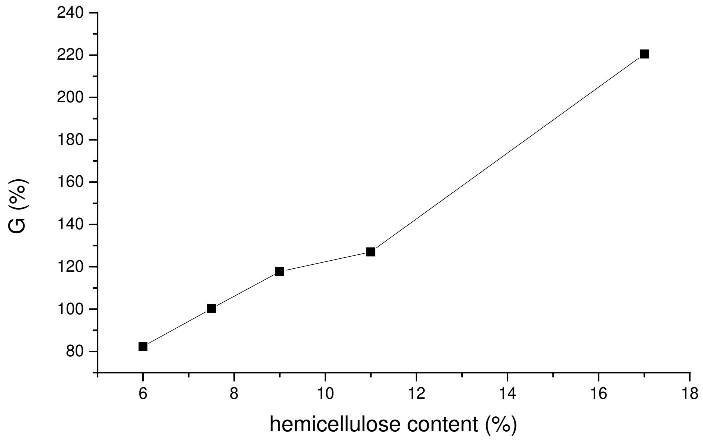


Figure 9. Effect of hemicellulose content of the cellulose on grafting efficiency

At the end of our research work the unmodified and grafted cellulose samples were analyzed by IR spectroscopy, thermal analysis and scanning electron microscopy.

Conslusions

In this work the graft copolimerization of vinyl acetate onto industrial cellulose pulp by Ce^{4+}/ Ce^{3+} redox system is characterized by maximum graft yield at varying temperature, reaction time, specific surface and chemical compositions of the cellulose and at different concentrations of monomer and initiator. This method is suitable for producing binder fibers applied in special synthetic papers.

References

[1] See, E..G.,Bains, M. S.: J. Poly. Sci. 1972. 37. p 125 – 182.

[2] Gupta, K. C., Sahoo, S.: J. Appl. Polym. Sci. 2001. 79. p 767 – 789.

[3] Flaque, C., Rodrigo, L.C., Ribes-greus, A.: J. Appl. Polym. Sci. 2000. 76. p 326 – 337.

[4] Abdel-Razik, E. A.: J. Photochem Photobio. 1993. 73. p 53 – 72.

[5] Zahran, M. K., Mahmoud, R. I.: J. Appl. Polym. Sci. 2003. 87. p 1879 – 1888.

[6] Okiemen, F.E., Idehen, K. I.: J. Appl. Polym. Sci. 1989. 37. p 1253 – 1276.

[7] Phillips, R. B., Quere, J., Guiroy, G. Tappi. 1972. 55. p 858 – 879.

[8] Borbely, E. Thesis, Budapest, 1984. p 25 - 67.

[9] Erdelyi, J. Thesis, Budapest, 1993. p 37 – 98.é

Decision Supporting Model for Highway Maintenance

András I. Bakó[*] Zoltán Horváth[**]

[*]Professor of Budapest Politechnik
[**]Adviser, Hungarian Development Bank
H-1034, Budapest, 6, Doberdo str.

Abstract

A decision model for pavement management has been developed herein based on linear programming formulation. Markov transition probability matrices are introduced to model the deterioration process of the road sections. To every type of road surfaces and class of traffic amount belongs a certain Markov matrix. The presented model and methodology is used to determine the optimal rehabilitation and maintenance policy in network level. Depending on the objective function two types of problems could be solved by the model : the necessary funds calculation and the optimal budget allocation for the entire network. We have developed the computer program on microcomputer and it has been used by the Ministry of Transport who is responsible for the 30000 km road network of Hungary.

Keywords: Markov process, Decision Support System, Pavement Management System, Network - Level Model, Linear Programming,

1. INTRODUCTION

The current budget condition in the Eastern European Countries needs an effective economical politics. It is true about Hungary too. We try to use the most powerful optimization models in every possible field.

In this paper we present an optimal decision supporting model that is used to maintain our highway network. Its length is 30000 kilometres. The works began some years ago. First a large scale Road Data Bank was developed ([1]). The second step was to develop a network level Pavement Management System, PMS ([2], [7], [8], [13]).

Concerning the PMS problem several types of models are known. One of them is the network level model, the other is a project level one. The network level model deals with the whole network. Its aim is to determine the most advantageous maintenance technique for every subset of the road having the same type of surface, the same condition parameters and the same traffic category. This type of model is a budget planning tool capable of estimating the total lengths and costs of works required on the network for pavement rehabilitation, resurfacing and routine maintenance. One type of financial planning is generally connected with the determination of the level funding needed to maintain the health of the pavement network at a desirable level. In the other type of model the available budget is given and we have to determine the maintenance politics that fulfil the required constraint of conditions and optimize the total benefit of the society. In the project level model a maintenance and rehabilitation program are determined for each pavement section. We usually use this model in a district [6].

Several types of solution algorithms can be used depending on the given task, the available data, the budget constraints, etc.([3], [4], [5]). Two main types are the heuristic and the optimization algorithms. The heuristic technique is usually used in project level, but it could be used in network level too. The optimization models are solved by the traditional optimization algorithms. Depending on the problem to be solved we use integer ([14]), a linear ([11]) or a dynamic programming algorithm([7]).

Our model is a linear programming one which has some stochastic elements. Namely the road deterioration process is described by the Markov transition probability matrices. In the second chapter we describe this probability supposition. The third chapter deals with the model formulation. In the last chapter we summarise the applied model itself. The engineering part of the model was developed by Gaspar ([9]), the program system was written by Szantai, ([13]). Similar model was proposed in Arizona and in Finland ([14]).

2. Markov transition probability matrix

In the model we will use the theory of Markov chains (Prekopa, [12]). To demonstrate this let us suppose that the pavement conditions are described by a certain discrete state. This state contains a discrete set. Let us denote these states by numbers 1,2,..The change of the system condition in time is probabilistic, and we fix these states in the time t=1,2,...

The probabilistic variables x_0,x_1,... are defined in the following way : x_n=i when the system is in state i in time period t=n. The system conditions are described by the x_0 ,x_1,...variables.

We can suppose, that the initial state e.g. x_o is fixed. The set of $x=(x_o,x_1, ...)$ is called a Markov chain, when any integer time set $t_o < t_1 < ... < t_{n+1}$ and states $k_1, k_2, ..., k_{n+1}$ the following condition is satisfied:

$$P(x_{t_{n+1}}=k_{n+1} \mid x_{t_1}=k_1, x_{t_2}=k_2, ... x_{t_n}=k_n)=P(x_{t_{n+1}}=k_{n+1} \mid x_{t_n}=k_n) \qquad (1).$$

This condition means that the probability that the system in time t_{n+1} is in state k_{n+1} depends on only the previous state, and independent from the earlier states. Now we define the Markov transition matrix. The r-step homogenous transition probability is defined by

$$q_{ik}^{(r)} = P\ (x_{n+r}=k \mid x_n=i)$$

The $q_{ik}^{(r)}$ values are ordered into a matrix Q_r which is called the transition probability matrix.

$$Q_r = \begin{pmatrix} q_{11}^{(r)} & q_{12}^{(r)} & \dots \\ q_{21}^{(r)} & q_{22}^{(r)} & \dots \\ \dots & \dots & \dots \end{pmatrix}. \qquad (2)$$

This matrix is a stochastic matrix because it is quadratic its element are non negative and the sum of the columns is equal to 1. It could be shown that the product of two stochastic matrices is also stochastic. We will use this result later.

It can be proved the r-step transition probability matrix equal to the rth power of the one-step transition probability matrix:

$$Q_r=Q^r \qquad (3)$$

The system is ergodic, we can reach every state by positive probability. On the basis of this theorem we build up the matrix which is used in our model. In this case a state corresponds to a certain condition of a set of sections which has the same type of surface, amount of traffic and quality. The number of rows (and columns) is equal to the number of discrete road states. The $q_{ij} \in Q$ is the probability that the road being in state j will be in state i at the end of the planning period.

Let us suppose that the initial distribution $X=(X_{01}, X_{02}, ..., X_{0m})$ is known. We compute the distribution X_1 at the end of the planning period using the Markov matrix Q:

$$X_1=X_oQ \qquad (4)$$

If there is m planning period, t=1, 2, ..., m, then the corresponding distributions X_1, X_2, ..., X_m are determined by a recursive procedure:

$$\begin{aligned}
&X_1=QX_{0,}\\
&X_2=QX_1=QQX_0=Q^2X_{0,}\\
&X_3=QX_2=QQ^2X_0=Q^3X_{0,} \qquad (5)\\
&\ldots \quad \ldots \quad \ldots \quad \ldots\\
&X_m=QX_{m-1}=Q^mX_{0.}
\end{aligned}$$

3. Model formulation

The Markov matrix depends on the pavement type, the volume of traffic and the maintenance actions.

In the model we suppose that there are s different type of pavement, f class of traffic volume and t type of maintenance politics. In this case we have s∗f∗t different Q matrix. Let us denote the Markov matrix by Q_{sft}, which belongs to the pavement type s, traffic class f and maintenance politics t.

There are several constraints to be fulfilled. We will denote the unknown variable by x_{ijk} which belongs to the pavement type i, to the traffic volume j and to the maintenance politics. The solution have to be Markov stabile. The Markovian stability constraint is

$$\sum_{i=1}^{s}\sum_{j=1}^{f}\sum_{k=1}^{t}\left(Q_{ijk}-E\right)x_{ijk}=0, \qquad (6)$$

where E is a unit matrix.

Because the equality is usually not fulfilled or not desirable, we use ≤ or ≥ relation instead of equality in (6). There are several further constraints which are connected with other suppositions. We suppose that the traffic volume will not change during the planning period:

$$\sum_{k=1}^{t}x_{ijk}=b_{ij}, i=1,2,\ldots.,s, \qquad (7)$$

j=1,2,....,f,

where b_{ij} belongs to the pavement type i and the traffic volume j.

The total area of the road surface type i will remain the same at the end of the planning period

$$\sum_{j=1}^{f}\sum_{k=1}^{t} x_{ijk} = d_i \ , \qquad \text{i=1, 2, ... s,} \qquad (8)$$

where d_i belongs to the pavement type i and $\sum_{i=1}^{s} d_i = \underline{1}$

We have to apply one of the maintenance politics on every road section

$$\sum_{i=1}^{s}\sum_{j=1}^{f}\sum_{k=1}^{t} x_{ijk} = \underline{1} \qquad (9)$$

We divide the segments into 3 groups: acceptable (good), unacceptable (bad) and the rest. Let us denote the tree set by J (good) by R (bad) and by E (rest of the segments) and by H the whole set of segments. The relations for these sets are given by

$$J\cap R=\emptyset, J\cap E=0,$$
$$R\cap E=\emptyset, \quad J\cup R\cup E=H.$$

The following conditions are related to these sets

$$\sum_{i,j,k\in J} x_{ijk} \geq v_J ,$$
$$\sum_{i,j,k\in R} x_{ijk} \leq v_R , \qquad (10)$$
$$\underline{v}_E \leq \sum x_{ijk} \leq \overline{v}_E ,$$

where J, R, E are given above, and

v_j the total length of the good road after the planning period

v_R the total length of the bad road after the planning period

$\underline{v}_E$ the lower bound of the other road

$\overline{v}_E$ the upper bound of the other road

The meaning of the first condition is that the amount of good segment have to be greater than or equal to a given value. The second relation does not allow more bad roads than it is fixed in advance. The third relation gives an upper and lover limit to the amount of the rest road.

Let us denote by c_{ijk} the unit cost of the maintenance politic k on the pavement type i and traffic volume j. Our objective is to choose such an X which:

- fulfils the conditions given above,
- with minimal rehabilitation cost.

The objective is

$$\sum_{i=1}^{s} \sum_{j=1}^{f} \sum_{k=1}^{t} x_{ijk} c_{ijk} \to \min!$$

Let us denote this value by C. The budget C* which is available for the maintenance purpose is usually less than C, so C*<C. In this case we modify our model: the above mentioned rehabilitation cost function becomes constrained:

$$\sum x_{ijk} c_{ijk} \leq C^*, \qquad (11)$$

and we use another objective. Let us denote the benefit by h_{ijk} where this is the benefit of the societies when we apply on the pavement type i and with the traffic volume j the maintenance politics k.

Our aim is to determine such a solution X which fulfils the constraints (6)-(10) and (11) and maximises the total benefit of the society.

The objective in this case is

$$\sum x_{ijk} h_{ijk} \to \max! \qquad (12)$$

4. Two types of optimization models

We could build up two different types of models using the element given above. One of them is the Necessary Funds Model (NFM), the other is the Budget Bound Model (BBM). In the NFM model we determine the necessary funds needed for ensuring a given condition level of roads with minimal cost. The BBM model is used to distribute a certain amount of money with the given constraints and maximises the benefit of the road users.

The NFM model

Let us determine the unknown variable matrix $X=(x_{ijk})$ that fulfils the following conditions

$$\sum_{i,j,k} (Q_{ijk} - E) x_{ijk} = 0,$$

$$\sum_{k} x_{ijk} = b_{ij}, \qquad i=1,2,.....,s,$$

$$j=1,2,.....,f,$$

$$\sum_{jk} x_{ijk} = d_i, \qquad i=1,2,.....,s, \qquad (13)$$

$$\sum_{i,j,k} x_{ijk} = \underline{1},$$

$$\sum_{i,j,k \in J} x_{ijk} \geq v_J,$$

$$\sum x_{ijk} \leq v_R,$$

$$\underline{v}_E \leq \sum_{i,j,k} x_{ijk} \leq \overline{v_E},$$

and
$$\sum_{i,j,k} x_{ijk} c_{ijk} \rightarrow \min!$$

The Budget Bound Model

Determining the unknown matrix X=(x_{ijk}) which fulfils the condition (13) and the following budget limit condition:

$$\sum_{i,j,k} x_{ijk} c_{ijk} \leq C$$

and

$$\sum_{i,j,k} x_{ijk} h_{ijk} \rightarrow \max!$$

5. Application

The two models have been applied for solving the Hungarian network level Pavement Management System. The road network is divided into smaller groups which depend on the pavement type, the traffic volume and the maintenance action. Two pavement types were taken into consideration, the asphalt concrete and the asphalt macadam. Three traffic classes were chosen. These are low, medium and high traffic category. In our model we use three type maintenance actions. Theoretically 2x3x3=18 different categories were formed but two of them are unrealistic. So the aim was to elaborate the 16 categories. One Markov matrix belongs to each category.

The condition of a road section is described by 3 parameters: bearing capacity (5 classes), longitudinal unevenness note (3 classes), pavement surface quality note (5 classes). The number of the condition states are 3x5x5=75. For practical reason and simplification we reduce this number to 41.

The NFM model was used to determine the necessary funds needed to held the road network a desired condition level. The available budget for that purpose was lower, that is why we use the BBM model with a fixed budget limit. Instead of the benefit h_{ijk}, we apply the vehicle operating cost in the objective.

Firstly we distribute the available budget country-wide according to the maintenance actions, pavement types, and traffic categories. There after we distribute the result among regional traffic agencies. This distribution was based on the area shares of sections with given characteristics (traffic volume, pavement type, pavement condition).

Both problems can be solved by a linear program package. This package consists of two parts: data generation and optimization. The data generation uses the Hungarian Road Data Bank. Depending on the constraints a selection and a data aggregation is used to generate the proper data to the model.

The size of the matrix is quite large

- the number of columns in both models is 734 (18x41),
- the number of rows in

  NFM model is 91

  BBM model is 92.

The computer solves the problem in 1-3 minutes (on IBM PC PENTIUM) depending on the output and the structure of the matrix. The LP code was written in FORTRAN by Szantai(1990).

For the funds need calculation optimization model two strategies have been tested first.

Strategy 1. The proportion of 20 pavement surfaces with wrong condition level can not be increased.

Strategy 2. As Strategy 1. and the proportion of 4 pavement surfaces with the best condition levels should be increased.

The following table shows results according to the two strategies:

	Strategy 1.	*Strategy 2.*
routine maintenance	610 million HUF (29.3 %)	422 million HUF (2.8 %)
surface dressing	646 million HUF (31.0 %)	264 million HUF (1.7 %)
new asphalt layer(s)	826 million HUF (39.7 %)	14.410 million HUF (95.5 %)
total funds need	2.082 million HUF	15.096 million HUF

It can be seen from the above table that if we want to preserve the proportion of the pavement surfaces with the best conditin levels, the cost of the new asphalt overlay will extremely increase and the total funds need becomes also very high. On the base of these results it was decided to develop five new strategies. These are:

Strategy 1. Proportion of 20 wrong variants can not be

increased.

Strategy 2. Proportion of 16 wrong variants can not be

increased and the 4 worst ones should be

decreased by 5 %.

Strategy 3. As *Strategy 2* with 10 %.

Strategy 4. As *Strategy 2* with 15 %.

Strategy 5. As *Strategy 2* with 20 %.

In these cases the total funds need went from 2.000 million HUF and they were acceptable for the administration.

When applying the funds split model it was first solved for the whole country, then for several regions (counties) separately.

When the optimal ratios (proportions) of various maintenance techniques in case of selected funds available (2000 million HUF, 3000 million HUF, 4000 million HUF, 5000 million HUF, 6000 million HUF, 7000 million HUF) the following main results were obtained:

- in case of the allocation of 3.0×10^9 HUF funds only one-third of the financial means was used for asphalt overlays, the highest share is spent for surface dressings,

- increasing the funds available, the financial means allocated to asphalt overlay considerably grow while the shares of other two intervention types, evidently, decrease;

- among the areas of various intervention types not so high percentage changes can be observed since the unit costs of routine maintenance and surface dressing gradually decrease accordingly, as - together with the increase of total funds - asphalt overlay is applied on the worst sections that obtained earlier only patching or surface dressing.

When comparing the splits according to the previous and the optimized model (various levels of funds were considered here) it could be seen that a significant changing in shares of some counties presented itself as a consequence of the use of the new model. The funds available had a minor effect on the shares destinated to the counties.

It was also analysed how the funds, increased by $1.0\text{x}10^9$ HUF steps, influence the vehicle operating costs. There was a definite tendency that the "savings" (reduced fuel costs) werw smaller and smaller as the total funds grow. This statement was, naturally, not surprising at all, because the extra funds permited to repair not only the very poor but also the less bad sections. In the latter case, evidently, a lower fuel costs reduction could be attained by the interventions.

References

[1] Bako, A. - Gyulai, L. - Erben ,P.: Structure of the Road Data BaProceedings of the Pavement Management System, Budapest, (1989), pp. 43 - 47.

[2]Bakó, A., Klafszki , E., SzántayT., Gáspár, L.: Optimization Techniques for Planning Highway Pavement Improvements, Annal of Operations Research 58(1995) 55-66.

[3] Bakó, A., Ambrusné S.K., Horváth, L.: Development a Highway PMS, Proceedings of the 1st European PMS Conference, 2000, pp. 12.

[4] Chua, K.H. - Der Kiureghian,A. - Monishmith,C.L.: Stochastic Model for Pavement Design,J. of Transportation Engineering,118(1992) ,pp. 769-786.

[5] Cook,W.D. - Lytton,R.L.: Recent Development and Potential Future Directions
in Ranking and Optimisation Procedures for Pavement Management,N.A.C. on Managing Pavement, V.2., 2.213-2.157.

[6] Feighan, K.J. - Shahin, M.Y. - Sinha, K.C. - White, T.D.: An Application

of Dynamic Programming and Other Mathematical Techniques to Pavement Management Systems, Transportation Research Record 1200, 1988.

[7] Feighan, K.J. - Shahin, M.Y. - Sinha, K.C.: A Dynamic Programming Approach to Optimisation for Pavement Management Systems", Proceedings of N.A.C. on Managing Pavement 1988, V.2.,2.195-2.206.

[8] Gaspar L. - Bako A.: Compilation of the Hungarian Network-level Pavement Management System, Revue Generale des Routes et des Aerodromes 170(1993), 34-37.

[9] Gáspár, L.: Network level use of FWD in Hungary. First European FWD User's
Group Meeting, 1-2 February 2001, Delft. Information binder Presentation No.12. 9pp.

[10] Gáspár, L.: Highway pavement performance models. 9th International Conference
on Asphalt Pavements. Copenhagen, August 17-22, 2002. CD-ROM Proceedings.

[11] Markow,M. J. - Brademeyer, B. D. - Sherwood, G. M. - Kenis, W. J.: The Economic Optimisation of Pavement and Maintenance and Rehabilitation Policy,S.N.A.C. on Managing Pavement, (1988), pp. 2.169- 2.182.

[12] Prekopa, A.: Probability Theory, Technical P.C., 1972, pp. 440 (in Hungarian).

[13] Szantai, T.: Computer Programming System for Solving the Hungarian Network - Level PMS, Research Report , 1990, pp. 32 (in Hungarian).

[14] Talvitie, A. - Osen, R.: Selecting Asphalt Concrete Condition States Finland's Highways", 67th Annual Meeting of the TRB, Washington, DC., 1988, pp.37.

Microtopography Changes in Wear Process

Béla Palásti-Kovács*, Zoltán Néder, Árpád Czifra**, Károly Váradi****

* Budapest Polytechnic, H-1081 Budapest, E-Mail:palasti.bela@bgk.bmf.hu
** Budapest University of Technology and Economics, H-1111 Budapest
E-Mail:varadik@eik.bme.hu

Abstract
Wear experiments and measurements were performed to study surface microtopography changes. Investigations extended to wear in the course of the non-lubricated sliding friction of ground bronze-steel sliding pairs. In the knowledge of 3D microtopography, asperities were statistically processed. Asperities were replaced by paraboloid and pyramidal surfaces, in order to determine the distribution of the direction angle of asperities, the height distribution of the peak points, the radius distribution of the peak curvatures, and slope angle distribution. These can be properly used for characterizing microtopography changes in the course of the wear process. It was also examined what additional information was provided by SEM recordings of surfaces on surface structure, with particular regard to tribological phenomena.

Key words: microtopography, asperities, wear, statistical analysis

1. Introduction

The operation, reliability, and lifetime of parts produced in different ways greatly depend on the quality of machined surfaces as well. Higher quality criteria require adequate accuracy of manufacturing as well as a deeper analysis of surface microtopography. Surface quality includes surface microgeometry discrepancies, such as roughness and waviness as well as the physical and chemical conditions of the surface layer, the latter including plastic deformation in the course of machining, hardness of the surface layer, residual stress, texture, and chemical composition [1].

The relationship between surface quality and the wear process has been studied by Whitehouse [2], Hirst and Hollander [3] and others; however, no general correlation has been managed to be established. Nevertheless, practical research is characterized by investigation of the wear of particular material pairs under certain conditions. However, in the literature available [4, 5, 6], these studies trace the changes of only few surface roughness parameters (ex. average roughness). In most cases only the impact of initial surfaces of varying roughness is investigated there as well.

This study involves surface microgeometry investigations. The aim is to trace microgeometry changes on surfaces in the course of the wear process. In our work, the results of tests by stylus instrument and scanning electron microscope were processed to characterize the wear process on surfaces. For evaluation, surface asperities were replaced by paraboloids and pyramids and the distribution function changes of the surfaces received this way were studied.

2. Test and evaluation procedures for studying the wear process

2.1. Wear experiments

The aim of the experiments performed was to model the friction / wear process of machine elements under certain circumstances by wear experiments simple to be performed by which the changes of surface microtopography can be obtained.

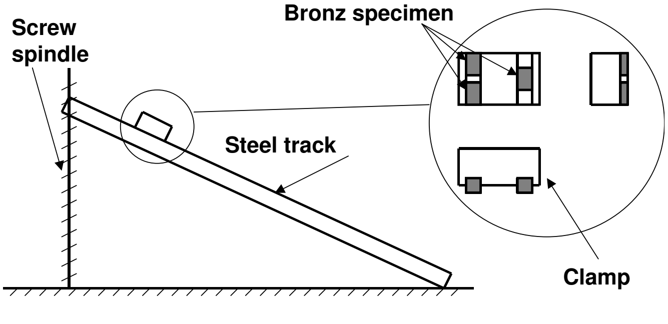


Figure 1. Arrangement of the experiment

A steel sliding track was used for the tests, where bronze specimens were slid (Figure 1). Both materials had ground surfaces. On the steel track, there were grooves of grinding in the direction of relative displacement, while those on the specimens were perpendicular thereto. The length of the sliding track was 700 mm and the nominal load of the specimens examined was 0.0125 MPa (10 N). Specimens were managed to be slid down by setting the slope using a screwed spindle. The environment temperature was 21 °C and there was no considerable warming between the surfaces in the course of the tests. This can be explained by the small number of cycles, low load, and low sliding speed. No lubrication was applied; the only layer of lubrication was produced by the atmosphere of the lab.

2.2. Measurement technique

Specimen microtopography was recorded at 2 different places using a Perthen Concept type stylus instrument. The size of the surface measured was 1x1 mm, with a sampling distance of 0.5 and 2 μm (2000x500 points), respectively. There was no use of higher resolution due to the 1 μm tip radius of the stylus needle and the error probability of lateral shift. An inaccuracy of the stylus instrument is that it ”flattens” real surfaces, therefore it provides a ”filtered” image. However, the characteristics of scanned surfaces can be quantified in the form of various parameters. Measurements were performed on identical surface section in each phase of the wear process, therefore the changes of a given surface section can be traced accurately, not only statistically, in the course of the wear process.

Surfaces were also recorded by a JEOL JSM 5310 type scanning electron microscope. Electron microscope recordings can present the smallest details of a surface, therefore certain phenomena can be explained and the “microscopic world” of surfaces can be explored.

2.3. Surface microtopography evaluation

One of the most frequently applied evaluation techniques is the use of statistical functions and parameters. It is obvious to use due to the statistical nature of surface topography data. Initially, only the scalar parameters known from mathematical statistics were used; by now, however, further parameters have been defined. Surfaces can be characterized by functions well-known from statistics: the distribution function, showing the distribution of measurement points around the mean-plane and the density function, yielding the value of material partition at a given height level.

Additionally to the statistical parameters known from the literature, regular surfaces substituting asperities were used for processing the cluster of points yielded by measurements. A significant benefit of this processing technology is that well-defined mathematical functions need to be examined this way instead of processing a ”disordered” cluster of points. Obviously, this processing technology has an inherent modeling error, namely the kind of surface to replace asperities. Paraboloid and pyramid surfaces proved to be the most suitable ones. They were used for defining the distribution of the direction angle of asperities, the height distribution of the peak points, the radius distribution of the peak curvatures and to determine slope angle distribution.

3. Results

3.1. Visual characterization of the wear process

Figure 2 illustrates the changes of bronze sliding surfaces in various phases of the wear process. It can be observed that in the initial phase of the wear process (wear cycle 110) the higher asperities (crests) partly disappear, being somehow ”stumped”, and the asperities of the countering surface are ”finely” copied to this surface in accordance with the direction of relative movement.

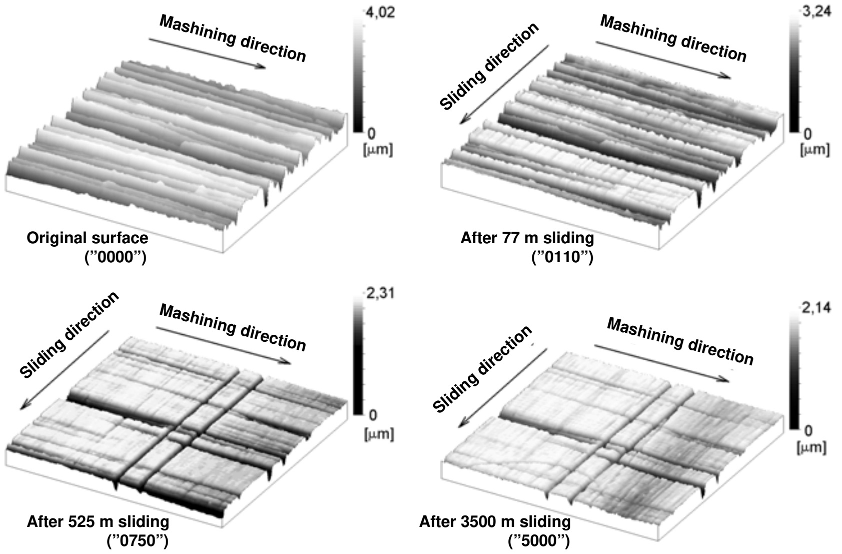


Figure 2. Wear process of a bronze specimen

This copying is partial since not a perfect copy of the track surface is generated but only grooves formed by the highest crests and peaks, the size of which probably strongly depends on the load. This phenomenon can be interpreted as a deformation process, where ”deformation resistance” depends on the depth of the grooves and the material, while the deforming force is in proportion with the surface load. The new pattern on the surface is generated as a balance between these two quantities. In the next phase of the process (wear cycle 750), some deeper scratches appear on the surface in the direction of sliding. This may be explained by the fact that some particles separated from the bronze surface and embedded to the steel surface have left their grooves on the countering surface. In this phase of the wear process, minor bronze deposits became visible on the steel surface, and a small amount of wear particles were to be observed at the bottom of the slope. In the meantime, the size of the contour contact area reached the size of the nominal contact area and the wear process considerably decreased. If these wear groves are compared to the wear grooves in the figure representing to wear cycle 110, it can be established that a finer pattern is generated in the later state. This corresponds to the deformation theory outlined above as the deformation force mentioned there is constant in the course of the process, but ”deformation resistance” may only remain constant as the number of grooves increases if the depth thereof is reduced. At wear cycle 5000, only some very deep grooves remained from the original grinding: an almost completely new pattern was formed on the surface.

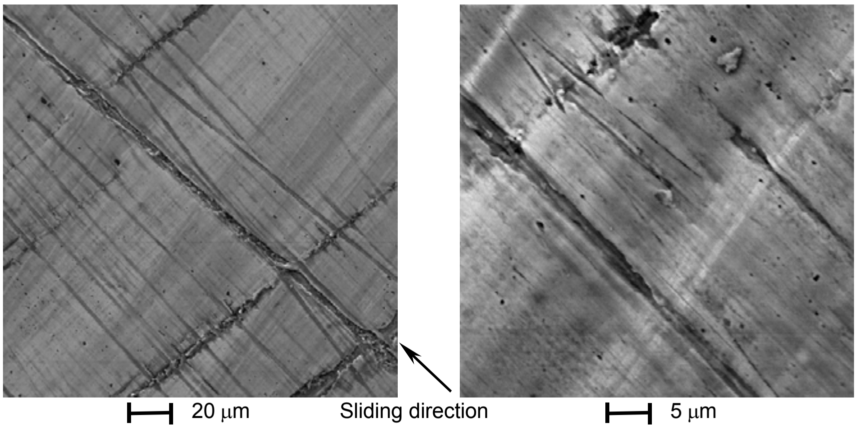


Figure 3. SEM recordings of the bronze surface after 5000 wear cycles

In order to further observe the wear process, SEM photos were also made of the surfaces to illustrate the final worn condition (Figure 3). The scratches in the direction of sliding can be properly observed on the ”new” surface mentioned above, with the original grooves of machining almost completely disappeared. There is a rough scratch along the middle of the image, probably formed by a hard particle. The rest of the scratches, however, are really fine. It can also be observed that the material is smearing into the grooves. This considerable strain can be explained by a high contact pressure. Since the size of the real contact area took up only a small part of the nominal contact area, the real contact pressure was a much greater then the nominal contact pressure. It can be seen that the grooves of machining are partly covered with worn particles, and even deeper are formed in the course of the wear process. In the highly enlarged image, two small wear particles can be observed, deposited on the surface and almost entirely embedded. ”Having escaped” later on, they can give rise to new deep scratches.

3.2. Characterization of the wear process based on the characterization of asperities

The quantitative evaluation of experimental results was performed by characterization with distribution functions rather than by the traditionally applied parametric evaluation technique. The benefit thereof lies in its broader information content as surface characteristics are defined by their respective distribution functions rather than by an average scalar parameter.

3.2.1. The high distribution curve and the bearing area curve

A number of functions and parameters can be used for characterizing surface topography [7]. Figure 4 shows the height distribution of the original and worn surface. In the course of the analysis of height distribution curves, the problem of adjustment arose, namely how to place the curves pertaining to particular wear phases into the same diagram to refer to actual changes.

It would be an error to perform adjustments according to mean-plane as the wear process brings about more dominant changes in the upper layers of the surface, in the proximity of asperities. The formation of observable plateaux entails a ”downward” displacement of mean-plane.

The most realistic option for adjustment seems to be adjustment according to the lowermost point. However, lowermost points cannot be considered to be constant, either. This is explained by the following reasons:

- In the course of the wear process, new deep grooves may be formed, the depth of which may exceed the depth of the lowermost point of the original (or previous) surface. Therefore lowermost points go lower.
- Major plastic deformations frequently occur in the course of the wear process. This may mean that the asperities wearing off are not removed in the form of wear particles but they are ”smeared” into the valleys, or the valleys themselves are filled up by wear particles; thereby considerably affecting the lower part of the height distribution curve.

Opportunities provided by visual display (particularly SEM), answering questions on the nature of the wear process, may offer guidelines in clarifying the problem. In our investigations, we came to the conclusion that the depth of new scratches does not exceed the depth of the original pattern, and ”smearing” into the valleys is not of such degree that it would exert a dominant influence on results, therefore adjustment according to the lowermost point can be accepted.

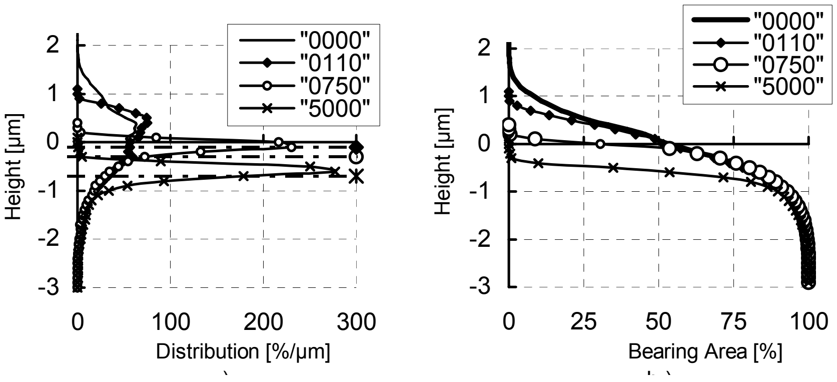


Figure 4. The height distribution curves (a) and the bearing area curves (b) fitted by the lowermost points

In the course of adjustment, an increasing adjustment displacement was necessary as the wear process progressed. Compared to the mean-plane of the original surface, the mean-plane was displaced by 0.1, 0.3, and 0.7 µm at wear cycles 110, 750, and 5000, respectively, assuming that the lowermost point of the curves was physically the same point. The mean-planes pertaining to each distribution function are indicated in Figure 4.

It can be established that peaks go lower as the wear process progresses, with the highest peaks wearing away. In the course of the process, the height distribution curve becomes ever more “acute” and asymmetrical. This means that measurement points are agglomerated at a given height level, where there are many surface points, therefore the ”plateaux” observed earlier are formed. The fact that the asymmetry of the curve is also increased indicates that primarily the ”upper” layers were changed in the course of the formation of the new surface, with asperities disappearing from there and replaced by a new pattern consisting of lower peaks. At the end of the wear process studied the new surface is almost entirely below the mean of the initial surface. This obviously results in a refinement of the surface as well as in the fact that a considerable part of the original pattern completely disappears, therefore the surface is not only transformed but an entirely new microtopography is generated.

In Figure 4b the bearing area curve was also fitted by the lowest point. It can be observed that the curve not only moves down, but the gradient of it becomes smaller. That is why the load bearing ability of the formed plateaux-like surfaces is better.

3.2.2. Asperity analysis

A number of methods have spread for defining and examining asperities (8-point method [8], contour mapping [9, 10], etc.), indicating that the experts involved in this subject have not yet found an advisable method which would be generally applicable and acceptable. In the present case, our attention is directed to wear processes; this is the reason why investigations are focused on the characterization of asperities. As regards wear processes in the event of dry friction, dominant changes from the viewpoint of surface quality occur in the proximity of asperities.

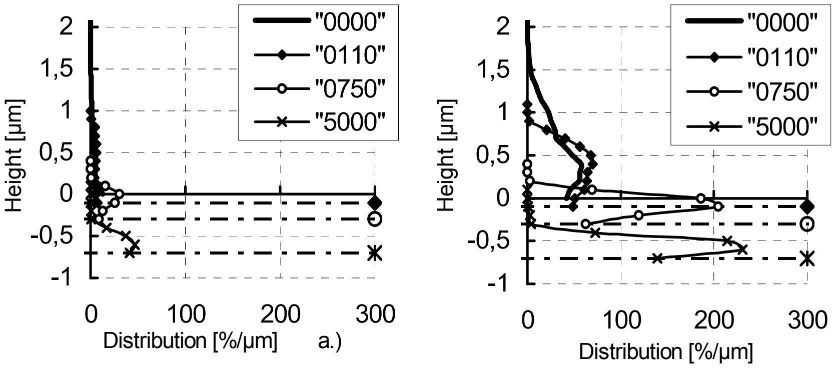


Figure 5. Height distribution of
a.) smaller asperities (area less than 40 μm^2)
b.) larger asperities (area greater than 40 μm^2)

Studying the measured microtopography, a new procedure was applied instead of filtering in the traditional sense. Essentially, the ”filtering” process is based on the principle that only asperities larger than a certain size have a significant impact on the wear process. Results are distorted by minor peaks on the surface of these asperities. Therefore asperities with a smaller area than the one specified were jointed by the algorithm into dominant asperities and evaluation was performed using those. Thus, each “elevation” located over the mean-plane, with a local maximum and an area exceeding a certain size was defined as an asperity. Figure 5 shows that the significance of asperities with an area lower than 10 cells (40 μm^2) is negligible: they do not have a dominant impact on the microtopography, particularly in the initial phases of the wear process. By studying the height distribution curve, it can also be established that the consideration of peaks smaller than 10 cell sizes may yield to false results because their number is comparable to the number of major peaks, while their area is much smaller.

Figure 6 shows distribution of peak points of asperities. The peak point of asperities are continuously reduced as the wear process progresses, and the ”even” distribution, characteristic of the initial phase, is eliminated. It can also be observed that upper peaks disappear and many new peaks are formed at a lower level. Also taking into consideration the adjustment by height of the curves, it can be established that the asperities of the worn surface generated have been lowered below the mean-plane of the original surface. Therefore the asperities originally defined have completely disappeared, with a new surface formed in their place which can be characterized by different asperities. As the topmost point of asperities does not only go lower but is approaching to the mean-plane, this also indicates that asperities themselves have got lower.

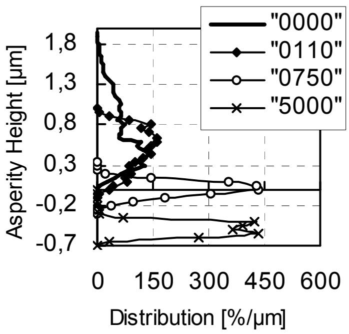


Figure 6. Distribution of peak points of asperities fitted by the lowermost point

The distribution curve in Figure 7 describes the direction of the major axis of the paraboloids substituting asperities. In the initial state, the curve reflects anisotropy characterizing ground surfaces. As the wear process is started (as a result of transversal sliding), this orientation discontinues. Although some orientation is shown in the direction of sliding as the process progresses, but this cannot be taken as dominant. The curves rather refer to a non-oriented isotropic surface.

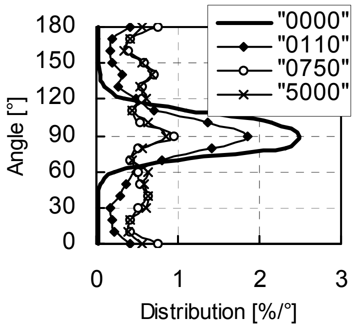


Figure 7. Asperity orientation distribution curve

It can also be observed that while considerable discrepancies were experienced in terms of wear cycle numbers in height distribution curves, the curves pertaining to wear cycles 750 and 5000, respectively, are almost completely identical in the case of orientational distribution curves, therefore the direction of asperities – the new surface – is formed earlier on.

Figure 8 shows the peak curvature radius of the substitute paraboloids to be defined according to major and minor axes. It can be observed that the radius of the peak curvatures pertaining to the major axis does not change considerably as the wear process progresses. On the other hand, the peak curvature radius of the minor axis increases in the course of the wear process. The shape of asperities changes. This is in relation with the fact that the original peaks caused by grinding are replaced by ”smeared” peaks.

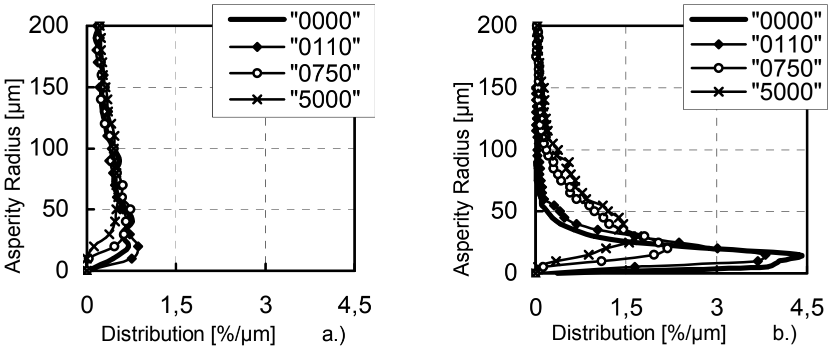


Figure 8. Radius distribution curve of the major (a.) and minor (b.) peak curvatures of the paraboloids substituting asperities

Initial asperities – with well-defined orientation and elongated in the direction of orientation – are replaced by less elongated asperities with larger radii. Larger radius represents better contact conditions. Therefore, in the course of operation, the surface changes in a way that the new surface formed will have more favorable contact conditions.

When examining slope angles (Figure 9) using pyramid substitution, two arrays of curves can be produced, which characterize the longitudinal and transversal directions peak by peak. It can be observed that the slope angle is always an obtuse angle. In displaying the microtopography the scale in height direction is usually greater than the one defined in the other two directions, however, it is misleading as regards the slope angle as this latter is considerably distorted. This fact may cause mistakes in the interpretation and explanation of wear processes.

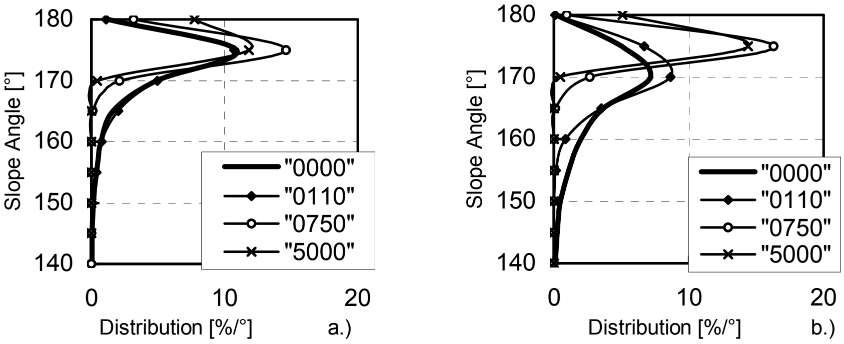


Figure 9. Slope angles pertaining to the major (a.) and minor (b.) axes

This set of curves is similar to the previous one characterizing the radius of the peak curvatures in that the curves indicate increased obtuseness and stumpiness of peaks here as well.

Conclusions, experiences

As a result of wear tests, the surface microtopography undergoes two fundamental changes. One of them is the increasingly disappearance of the original pattern, simultaneously with the formation of a new surface texture. Formation of the new pattern is considerably influenced by the direction of sliding, "deposits" developing on the surface, and wear particles between the surfaces.

The technique using distribution curves developed for characterizing asperities is suitable for studying the wear behavior of surfaces in the course of dry friction as well as for tracing surface changes.

In the wear experiments presented, an originally anisotropic surface was converted into an isotropic surface. By adjusting distribution curves, it was made possible to establish that initially existing asperities disappeared almost completely and were replaced by a new pattern.

In the course of the wear process, the surface changed in a way that contact and load bearing characteristics improved, with their slope radius increasing, their peak angle becoming more obtuse; therefore the surface change developed in the direction of reducing the effects generating such surface change.

Acknowledgements

The present study was supported by OTKA T 043151 research program.

References

[1] Palásti K. B., Czifra Á., Kovács K.: *Microtopography of machined surfaces, tribological aspects of surface and operation,* DMC 2002, Kassa 2002. 50-57

[2] D.J Whitehouse: Handbook of surface metrology, Inside of Phisics Publ., Bristol (1994), 988

[3] Hirst W., Hollander A. E.: Surface finish and damage in sliding, Proc. R. Soc. A 337 379-94 (1974)

[4] H. H. Gatzen, M. Beck: Wear of single silicon as a function of surface roughness, Wear 254 (2003), 907-910.

[5] S. Akkurt: On the effect of surface roughness on wear of acetal-metal gear pairs, Wear 184 (1995), 107-109.

[6] Y. A.-H. Mashal: The role of fracture mechanics parameters in glass/ceramic wear, Engineering Fracture Mechanics, Vol.52. No.1. (1995), 43-50.

[7] Kovács K. - Palásti Kovács B., Műszaki felületek mikro-topográfiájának jellemzése háromdimenziós paraméterekkel. I. A háromdimenziós topográfiai paraméterek áttekintése, Gépgyártástechnológia, 1999/8. 19-24.

[8] K. J. Stout, P. J. Sullivan, W. P. Dong, E. Mainsah, N. Luo, T. Mathia, H. Zahouani: The development of methods for the characterisation of roughness in three dimension, University of Birmingham Edgbaston, Birmingham (1993), 358

[9] K. Yanagi, S. Hara, T. Endoh: Summit identification of anisotropic surface texture and directionally assessment based on asperity tip geometry, International Journal of Machine Tools & Manufacture 41 (2001), 1863-1871.

[10] Z. Néder, K. Váradi, K. Friedrich: Characterisation of real polymer composite-steel surfaces by asperity-analysis and –substitution, ASME/STLE International Tribology Conference, Cancum, Mexico oct. 27-30, 2002

Measurement Methods in the field of benchmarking

Istvan Szuts, Dr.,

Budapest Polytechnic Institute for Entrepreneurship Management
6 Doberdo street, Budapest, 1034
szuts@bmf.hu

Abstract: In benchmarking[1] we often come across with parameters being difficult to measure while executing comparisons or analyzing performance, yet they have to be compared and measured so as to be able to choose the best practices. The situation is similar in the case of complex, multidimensional evaluation as well, when the relative importance and order of different dimensions, parameters to be evaluated have to be determined or when the range of similar performance indicators have to be decreased with regard to simpler comparisons. In such cases we can use the ordinal or interval scales of measurement elaborated by S.S. Stevens.

1. Ordinal scale

In case of ordinal scale the entities can be compared by the desired features, by means of it we can achieve relations like for example: better than ..., more useful than ..., bigger than ..., or their opposites: worse, less useful, smaller, etc. When measuring with ordinal scale entities must be comparable and transitive by one common criterion.

To indicate comparison the so called preference relation is used. Preference relation shows the rank order of the entities. A prefers to B ($A \succ B$) means that A can be regarded as better with regard to the actual criterion. For indication the natural numbers are used in an increasing or decreasing order. By means of it the above transitive criterion can be formulated as follows.

If $A \succ B$ and $B \succ C$, entity A is better than B and B is better than C, then A is also better than C.

[1] Comptetitive benchmarking involves analyzing the performance and practices of best-in-class companies. Their performance becomes a benchmark to which a firm can compare its own performance and their practices are used to improve that firm's practices.

Assignment of rank numbers can be carried out in an increasing sequence in the following way. We allocate 3 to A, 2 to B, 1 to C.

A→3
B→2
C→1

Only the ranking is important while allocating. We can not state that entity A is twice as good as B or C is three times better. The allocation of numbers is totally arbitrary with the exception of sticking to the sequence, it is just a matter of decision that the first three natural numbers have been used. Differences and ratios between the values of the scale tell nothing about real differences and proportions, they only establish the sequence. The scale will remain untouched by any sequence preserving transformation.

Measuring by ordinal scale can be carried out in two ways in accordance with practical requirements. In the example presented so far no ecriterion has been permitted, the preference between the two entities had to be decided. In case of such rules we deal with a so called strong preference. In case we permit ecriterion between the entities, the sequence can be regarded as a so called weak preference. A sequence can significantly limit the applicable evaluating statistic methods. On the ordinal scale natural numbers are usually used. The entities on the scale are not at identical intervals, they are not of the same magnitude. In this case only those operations can be accomplished which do not presume the identical magnitude of intervals. With regard to ordinal measuring we have to speak about the problems of multidimensional comparisons as it is important from the point of view of performance evaluation. As mentioned, comparisons can be accomplished only if entities have at least one common criterion. If comparison is executed only by one well defined criterion we have to do with a one dimensional comparison. However, we are well aware that a product, a process or a company may have theoretically infinite number of criterians. With regard to benchmarking [5] analysis not all the criterians can be taken into consideration, so we choose only some important ones. The selected set of criterians serve also as a base for the evaluation of products, processes and companies, so the criterians can be termed as evaluation factors as well. If comparison is accomplished with several criterians and evaluation factors we are faced with a multidimensional comparison. In this case specific problems arise that are not easy to solve. Some evaluation factors can be measured by ordinal, others by interval scales. The problem is how to compare the different dimensions.

In product comparison investigations in the so called “criterion range” provide solutions in case of ratio scales, but their practical applications are significantly limited by extremely difficult mathematical operations. Some criterians to be evaluated can not be measured on ratio scales. Besides solving the dimensional and measuring problems of comparisons, priorities between evaluation factors have also to be established. For this we have to be familiar with the importance of evaluation factors with regard to each other.

2. Interval scale

The interval scale possesses the characteristics of the ordinal scale and the range between two numbers is known and has a definite value. This scale can be regarded as a measuring scale in its traditional sense. The differences in a numerical sense show equal differences in reality as well. Measuring units, the 0 point can be determined arbitrarily. So linear transformation (x' = ax + b) can be permitted. Adding up proportions and quantities make no sense as both change according to the position of 0 point. But if the differences of the entities are considered, the entities have additive properties and as such they are suitable for comparisons. In this case the "b" constant of the relation "ax+b" is eliminated and as a consequence the 0 point is also eliminated so the different entities can be regarded as an absolute quantity.

So the proportion of any two intervals is independent from measurement units and 0 point on the interval scale. Several criterians can be measured on the interval scale in benchmarking, but this is much more difficult than on the ordinal scale. The elaboration of a proper interval scale for measuring a criterion of a phenomenon often means a complicated scale designing technique. Data gained from ordinal scales can be transformed into interval scales by using specific measuring methods.

The axioms of ecriterion (1., 2., 3.) and the axioms of rank ordering are valid (4., 5.,) on ordinal scales. Newer axioms can not be provided for the interval scales in the above described axiom system, but it must be noted that the axioms of additivity (6., 7., 8., 9.,) are valid for the differences of values on the scales. [1]

1. Or a=b or a≠b
2. If a = b, then b = a ecriterion
3. If a = b and b = c, then a = c
4. If a > b, then $b \triangleright a$
5. If a > b and b > a, then a > c (ordinal axioms)
6. If a = p and b > 0, then a + b > p
7. a + b = b + a additivity
8. If a = p and b = q, then a + b = p + q
9. (a + b) + c = a + (b + c)

3. Pairwise Comparison

The comparative method of pairs means comparing the criterians of the entities in pairs and establishing the preference between two things (or dimensions of multidimensional comparison). The comparison, as stated earlier , may be one or multidimensional. Of course usually not two but several entities are compared and the limits of our ability to process the information may cause problems as the number of comparisons increases.

If we meet the requirements of transitivity (see ordinal scale) a consistent opinion or priority order is formulated. In the opposite case inconsistent result is obtained that can be described as follows: If $A \succ B$ and $B \succ C$, that $C \succ A$.

An inconsistent decision impairs or makes impossible the transformation from the ordinal scale into the interval scale.

As a great number of pairs must be compared in the course of analyzing performance in benchmarking, the control of consistency, the determination of non parametric factors for evaluation is an essential condition. Indicators for consistency can be formulated by establishing the inconsistent decision:

$$C = 1 - \frac{24d}{n^2 - n}, \qquad \text{if n is odd number}$$

$$C = 1 - \frac{24d}{n^3 - 4n}, \qquad \text{if n is even number;}$$

Where:

n= the number of compared entities

d= the number of inconsistent triads.

The inconsistent triads is shown by Figure.

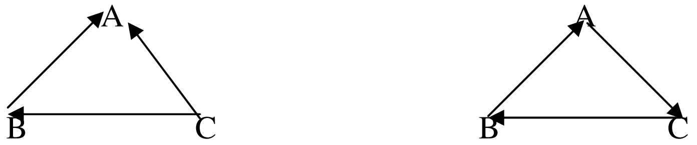


Figure 1.: Inconsistent triads

Results obtained by pairwise comparisons are usually demonstrated in a preference matrix, where lines and bars mean all compared entities. Such a preference matrix is shown in Table 1. with regard to seven compared entities with large evaluation factor. Number 1 in row A and bar B means that is entity. A was found preferable with regard to entity B ($A \succ B$). Table 1. contains the result of comparing in pairs with the final result of the following:

$$A \succ B \succ D \succ E \succ F \succ G \succ C$$

pairwise comparison provide possibility only for creating ordinal scales. The preference numbers obtained on the basis of preference matrix containing the results of the comparisons supply only ordinal scales without determining the magnitudes of intervals on the scales.

	A	B	C	D	E	F	G	Preference numbers
A	-	1	1	1	1	1	1	6
B	0	-	1	1	1	1	1	5
C	0	0	-	0	0	0	0	0
D	0	0	1	-	1	1	1	4
E	0	0	1	0	-	1	1	3
F	0	0	1	0	0	-	1	2
G	0	0	1	0	0	0	-	1

Table 1.: contains the result of comparing in pairs with the final result

If ordinal scales are obtained by putting them in priority order end not by comparing in pairs, then numbers can be allocated directly to the entities in an increasing or decreasing order. Of course in this way no information is given about the differences of scale values in reality.

In case of consistent decisions the preference matrix contains consistent partial square matrixes at all diagonals. For example: if 3,4,5, etc. square matrixes are chosen along the diagonal of the matrix in case of a ten-element matrix, the partial matrixes will be consistent. Let us see that in the case of a consistent matrix how many relations are needed between two elements to get a matrix. For example in a matrix where N=6, we have fifteen possible pairwise comparisons. This is the maximal number of comparisons but supposing the matrix is being consistent 5 pairwise comparisons will determine the remaining 10 preferences and the matrix can be filled in. In general it can be stated that a consistent matrix containing N elements can be determined by N-1 definite, not arbitrary comparisons.

An important problem of pairwise comparisons from the point of view of our investigation can be regarded as quantitative. If the number (N) of element to be compared exceed 8-10 the following problems may arise. The consistency of the decision makers will impair, matrixes will contain inconsistent decisions leading to ecriterians on interval scales and it will decrease the efficiency of the analyses. pairwise comparisons to the performed will drastically increase with the proliferation of entities.

While in case of N=4 the number of pairwise comparisons is 6, then in case of 10 elements it is 45, in case of 20 elements is as many as 190 [2]. The priority statistic and graphic methods provide possibility to determine the priority order of evaluation factors the basis of limited comparisons or to choose the ones belonging to the same set of criterians, this way the multidimensional evaluation can be simplified. This method can be very useful when establishing consumer preferences or when analyzing performance with benchmarking.

Let us consider the following example of purchasing a car, it is easily understandable for everyone, namely what comparisons occur most frequently.

Among the criterians mentioned we can find both expert and consumer preferences. By means of pairwise comparisons the following preference priority order of the pairs were obtained.

Preferences are marked by: $\succ$

Final Speed $\succ$ Comfort

Braking Distance $\succ$ Acceleration

Design $\succ$ Image

Consumption $\succ$ Final Speed

Final speed $\succ$ Function

Comfort $\succ$ Image

Function $\succ$ Design

Consumption $\succ$ Braking Distance

Safety $\succ$ Slowing Down

Comfort $\succ$ Design

Function $\succ$ Appearance

Acceleration $\succ$ Comfort

Safety $\succ$ Consumption

Comfort $\succ$ Appearance

Slowing Down $\succ$ Final Speed

Appearance $\succ$ Image

Safety $\succ$ Acceleration

Maximal performance $\succ$ Function

Slowing Down $\succ$ Max. performance

In graphic theory the problem can be formulated as follows. Let us consider criterians, features as a set where certain criterians end features are more important than others. We would like to set up a priority order so as to meet all preferences.

We would like to note that in case of 12 factors we ought to perform 66 pairwise comparisons.

T1 $\succ$ T5, T1 $\succ$ T8, T2 $\succ$ T9, T2 $\succ$ T10, T2 $\succ$ T11

T4 $\succ$ T11, T5 $\succ$ T3, T5 $\succ$ T7, T5 $\succ$ T12, T6 $\succ$ T8

T7 $\succ$ T3, T8 $\succ$ T7, T8 $\succ$ T12, T9 $\succ$ T1, T9 $\succ$ T6

T10 $\succ$ T1, T10 $\succ$ T4, T11 $\succ$ T5, T12 $\succ$ T3

These 12 criterians and their 19 relevant requirements can be represented by a graph where peaks substitute criterians and (TI,TJ) curve is shown in the graph if TI ≺ TJ [3].

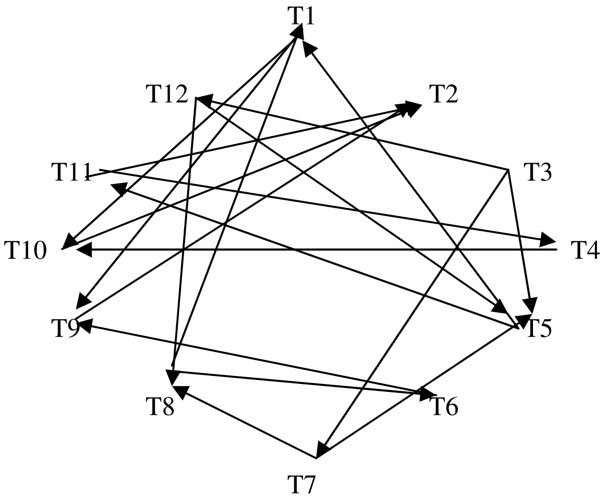


Figure 2.: 12 criterians and their 19 relevant requirements

It must be noted that this graph can not contain a cycle as in that case one event would precede itself and this can not happen because of the nature of the investigated matter.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12		V0	V1	V2	V3	V4	V5	V6
T1					1			1						2	2	2	0	x	x	x
T2									1	1	1			3	3	3	3	2	1	0
T3														0	x	x	x	x	x	x
T4											1			1	1	1	1	0	x	x
T5			1				1					1		3	2	0	x	x	x	x
T6								1						1	1	1	0	x	x	x
T7			1											1	0	x	x	x	x	x
T8							1					1		2	2	0	x	x	x	x
T9	1					1								2	2	2	2	0	x	x
T1	1			1										2	2	2	2	1	0	x
T1					1									1	1	1	0	x	x	x
T1			1											1	0	x	x	x	x	x
														T3 0	T7 T12 1	T5 T8 2	T1 T6 T11 3	T4 T9 4	T10 5	T2 6

Table 2.: Matrix

Let us take now the matrix representing the earlier graph. This matrix has been enlarged with a certain number of bars, will speak about it later. For better transparency deliberately no 0-s were written in the matrix where according to the definition we ought to have written them. Let us mark the column vectors of the matrix by V_{T1}, V_{T2}, …, V_{T12}.

First we can calculate vector

$$V_0 = V_{T1} + V_{T2} + ... + V_{T12}$$

and the result is written in bar V_0. In this vector 0 can be found in the place corresponding to row T_3 indicating that this peak (i.e. criterion) is not followed by any other.

So we can state that the level of T_3 is 0.

Now let us calculate vector

$$V_1 = V_0 - V_{T3}$$

and let's write x in the T3 row of vector V_1. We find a new 0 in rows of T7 and T12 of bar V_1; so if T3 is omitted T7 and T12 are not followed by any peaks. Therefore, we state that the level of T7 and T12 is one. Now we calculate vector

$$V_2 = V_1 - V_{T3} - V_{T12}$$

and we write x in vector V_2 in all places where 0 was to be found in the earlier vector, etc.

At least the 12 peaks were divided into 7 levels: N0, N1, N2, …., N6. These levels define what we call: graph priority function free of circular triads [4] [6].

The next figure (Figure 3.) shows the representation of this level determination: numbering was started at T2, we also could have made it in the reverse direction starting from T3. The figure shows the priority order of criterians and characteristics: they are shown not only in relation to each other but with regard to the whole structure.

It is obvious that operation T3 is the last, it is preceded by T7 and T12 (their priority between each other is not important), then they are preceded by T5 and T8 (again priority can be neglected) and so on.

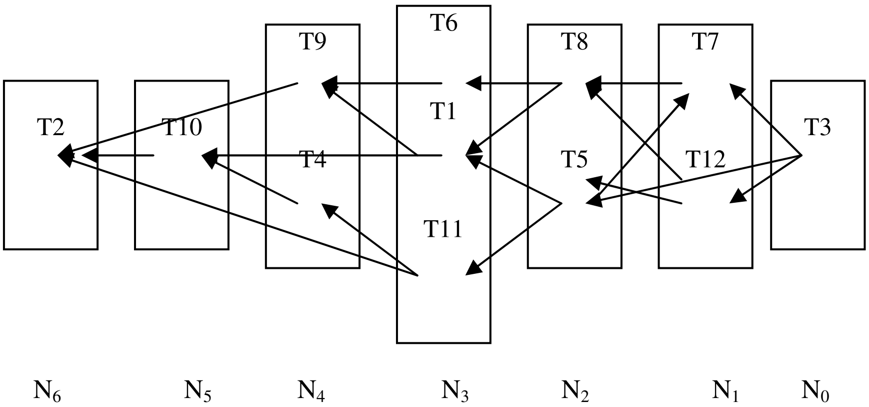


Figure 3.: Representation of the whole structure

Let us be aware that an other priority function would have been obtained if we had worked with the inverse relation "TI follows TJ". In this case we would have dealt with the row vectors of the matrix and we would have gained a priority function different from the previous one. However, all peaks have an arrangement that is compatible with all priority functions. On the basis of above said we can formulate the preference groups (consumer or benchmarking expert) determining the choice or comparison.

The previously obtained 19 preferencies are the following (Figure 4.):

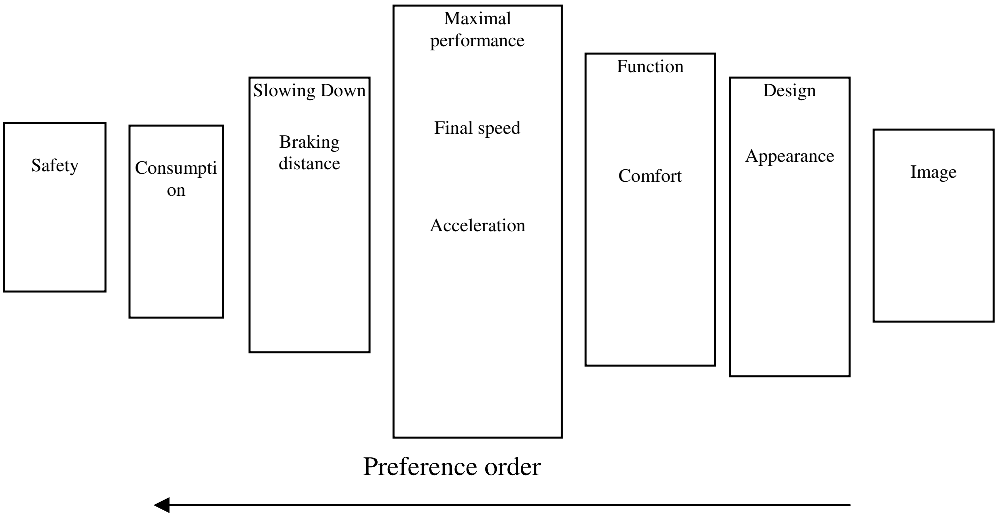


Figure 4.: The previously obtained 19 preferencies

References:

[1.] Stevens, S. S.: Handbook of Experimental Psychology
Wiley, New York 1951.

[2.] Dr. Szûts István: Methods for comparative analysis of company efficiency
KJK Budapest 1983 (In hungarian)

[3.] Istvan Szuts dr.: An operation research model for the credit system
In-Tech-Ed '96, Budapest, 1996.

[4.] Istvan Szuts dr.: Potential algorithms for the Credit System Analysis
KMF Budapest, 1999.

[5.] Robert C. Camp: Business Process Benchmarking
ASQC Criterion Press, Milwaukee 1996.

[6.] Dr. Bakó András, Dr. Szûts István: Gráfelméleti algoritmusok az oktatás korszerûsítésére. KMF Tudományos Közlemények, 1999. 19-29. o.