Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 1–6

An Application of Pescar's Univalence Criterion

Nicoleta Ularu[a,*], Daniel Breaz[b]

[a] *University of Piteşti, Department of Mathematics and Computer Science, Târgul din Vale Str., No.1, 110040 Piteşti, Argeş, Romania.*
[b] *"1 Decembrie 1918" University of Alba Iulia, Department of Mathematics and Informatics, N. Iorga Str., No. 11-13, 510009, Alba Iulia, Alba, Romania.*

Abstract

For the operator $F_\alpha(z) = \left(\alpha \int\limits_0^z t^{\alpha-1} f'(t)dt\right)^{\frac{1}{\alpha}}$, Pescar has obtained a generalization of Ahlfor's and Becker's criterion of univalence. In this paper we generalize the Pescar's univalence criterion for other two operators $G_{\alpha_1,\ldots,\alpha_n,n}(z)$ and $J_{\gamma_1,\ldots,\gamma_n}(z)$ and we obtain new univalence conditions of analytic functions in the unit disk $\mathcal{U} = \{z \in \mathbb{C} : |z| < 1\}$.

Keywords: Analytic, univalent, unit disk, Ahlfor's and Becker's criterion of univalence.
2000 MSC: 30C45.

1. Introduction and preliminaries

Let $\mathcal{U} = \{z : |z| < 1\}$ the unit disk and $\mathcal{A}$ the class of all functions of the form:

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n \tag{1.1}$$

which are analytic in $\mathcal{U}$ and satisfy the condition

$$f(0) = f'(0) - 1 = 0.$$

Theorem 1.1. *(Mocanu et al., 2009) (Maximum Modulus Principle) Let f be a nonconstant analytic function on a connected open set U. Then $|f|$ cannot attains maximum in U, i.e. there exists $\alpha \in U$ such that $|f(a)| \geq |f(z)|$ for all $z \in U$.*

The next lemma is a result given by J. Becker (Becker, 1972):

*Corresponding author
Email addresses: nicoletaularu@yahoo.com (Nicoleta Ularu), dbreaz@uab.ro (Daniel Breaz)

Lemma 1.2. *(Becker, 1972) If $f(z) = z + a_2z^2 + \dots$ is analytic in $\mathcal{U}$ and*

$$(1-|z|^2)\left|\frac{zf''(z)}{f'(z)}\right| \leq 1$$

for all $z \in \mathcal{U}$, then the function $f(z)$ is univalent in $\mathcal{U}$.

L. V. Ahlfors (Ahlfors, 1973) and J. Becker (Becker, 1973) has obtain the next univalence criterion:

Theorem 1.3. *((Ahlfors, 1973) and (Becker, 1973)) Let c be a complex number, $|c| \leq 1, c \neq -1$. If $f(z) = z + a_2z^2 + \dots$ is a regular function in $\mathcal{U}$ and*

$$\left|c|z|^2 + (1-|z|^2)\frac{zf''(z)}{f'(z)}\right| \leq 1$$

for all $z \in \mathcal{U}$, then the function f is regular and univalent in $\mathcal{U}$.

V. Pescar in (Pescar, 1996) obtain an univalence criterion which is a generalization of Ahlfor's and Becker's criterion of univalence and is given in next theorem.

Theorem 1.4. *(Pescar, 1996) Let α and c be complex numbers, $\operatorname{Re}\alpha > 0, |c| \leq 1, c \neq -1$. If $f(z) = z + a_2z^2 + \dots$ is a regular function in $\mathcal{U}$ and*

$$\left|c|z|^{2\alpha} + (1-|z|^{2\alpha})\frac{zf''(z)}{f'(z)}\right| \leq 1$$

for $z \in \mathcal{U}$, then the function

$$F_\alpha(z) = \left(\alpha\int_0^z t^{\alpha-1}f'(t)dt\right)^{\frac{1}{\alpha}} = z + \dots$$

is regular and univalent in $\mathcal{U}$.

In (Pascu & Radomir, 1989) N.N. Pascu and I. Radomir has obtain:

Theorem 1.5. *(Pascu & Radomir, 1989) Let β and c complex numbers, $\operatorname{Re}\beta > 0, |c| \leq 1, c \neq -1$ and $f(z) = z + a_2z^2 + \dots$ be a regular function in $\mathcal{U}$. If*

$$\left|ce^{-2t\beta} + (1-e^{-2t\beta})\frac{e^{-t}zf''(e^{-t}z)}{\beta f'(e^{-t}z)}\right| \leq 1$$

holds for every $z \in \mathcal{U}$ and $t \geq 0$, then the function

$$F_\beta(z) = \left(\beta\int_0^z t^{\beta-1}f'(t)dt\right)^{\frac{1}{\beta}} = z + \dots$$

is regular and univalent in $\mathcal{U}$.

We define the operators

$$G_{\alpha_1,\alpha_2,\ldots,\alpha_n,n}(z) = \left(\left(\sum_{i=1}^{n} \alpha_i - n + 1\right)\int_0^z \prod_{i=1}^{n} (g_i(t))^{\alpha_i - 1} dt\right)^{\frac{1}{\sum_{i=1}^{n}\alpha_i - n + 1}} \tag{1.2}$$

for $g_i \in \mathcal{A}, i = \overline{1,n}$ and

$$J_{\gamma_1,\gamma_2,\ldots,\gamma_n}(z) = \left(\left(\sum_{j=1}^{n} \frac{1}{\gamma_j}\right)\int_0^z t^{-1}\prod_{j=1}^{n} (f_j(t))^{\frac{1}{\gamma_j}} dt\right)^{\frac{1}{\sum_{j=1}^{n}\frac{1}{\gamma_j}}}. \tag{1.3}$$

for $f_j \in \mathcal{A}, j = \overline{1,n}$.
The operator $G_{\alpha_1,\alpha_2,\ldots,\alpha_n,n}(z)$ is a generalization of an operator defined by Breaz et all in (Breaz *et al.*, 2009).

2. Main results

Theorem 2.1. *Let α_i and c complex numbers, $n \in \mathbb{N}, n \geq 1, i = \overline{1,n}$, $\mathrm{Re}\left(\sum_{i=1}^{n}\alpha_i - n + 1\right) > 0, |c| < 1, c \neq -1$. We suppose that the function f defined by (1.1) is analytic in $\mathcal{U}$. If*

$$\left| c|z|^{2(\sum_{i=1}^{n}\alpha_i - n+1)} + \left(1 - |z|^{2(\sum_{i=1}^{n}\alpha_i - n+1)}\right)\frac{zf''(z)}{(\sum_{i=1}^{n}\alpha_i - n + 1)f'(z)}\right| \leq 1 \tag{2.1}$$

for all $z \in \mathcal{U}$, then the function $G_{\alpha_1,\alpha_2,\ldots,\alpha_n,n}(z)$ defined by (1.2) is analytic and univalent in $\mathcal{U}$.

Proof. For $z \in \mathcal{U}$ from (2.1) we have that $f'(z) \neq 0$ and from here the function

$$w(z,t) = c \cdot e^{-2t(\sum_{i=1}^{n}\alpha_i - n+1)} + (1 - e^{-2t(\sum_{i=1}^{n}\alpha_i - n+1)})\frac{e^{-t}zf''(e^{-t}z)}{(\sum_{i=1}^{n}\alpha_i - n + 1)f'(e^{-t}z)} \tag{2.2}$$

is analytic in $\overline{\mathcal{U}} = \{z : |z| \leq 1\}$, for $t > 0$. To the function $w(z,t)$ we apply the maximum modulus principle and we have that

$$\begin{aligned} |w(z,t)| &< \max_{|z|=1} |w(z,t)| = \left|w(e^{i\theta},t)\right| \\ &= \left| ce^{-2t(\sum_{i=1}^{n}\alpha_i - n+1)} + (1 - e^{-2t(\sum_{i=1}^{n}\alpha_i - n+1)})\frac{e^{-t+i\theta}f''(e^{-t+i\theta})}{(\sum_{i=1}^{n}\alpha_i - n + 1)f'(e^{-t+i\theta})}\right| \end{aligned} \tag{2.3}$$

where $\theta = \theta(t) \in \mathbb{R}$.
We note $\phi = e^{-t+i\theta}$. From here we have that $|\phi| = e^{-t}$.
If in relation (2.3) we replace $e^{-t+i\theta}$ with ϕ we obtain:

$$\left|w(e^{i\theta},t)\right| = \left| c \cdot |\phi|^{2(\sum_{i=1}^{n}\alpha_i - n+1)} + (1 - |\phi|^{2(\sum_{i=1}^{n}\alpha_i - n+1)})\frac{\phi f''(\phi)}{(\sum_{i=1}^{n}\alpha_i - n + 1)f'(\phi)}\right| \tag{2.4}$$

Because $|\phi| = e^{-t}$, for all $t > 0$ it results that $\phi \in \mathcal{U}$.
For $z = \phi$ using the relations (2.1) and (2.4) we obtain that

$$\left|w(e^{i\theta}, t)\right| \leq 1 \tag{2.5}$$

From (2.3) and (2.5) we have $|w(z, t)| < 1$, for $z \in \mathcal{U}, t > 0$.
For $t = 0$ we have that $w(z, 0) = c$. Using the hypothesis we obtain $|w(z, 0)| < 1, z \in \mathcal{U}$. So, $|w(z, t)| < 1, z \in \mathcal{U}, t \geq 0$ and from here and using Theorem 1.5 for $\beta = \sum_{i=1}^{n} \alpha_i - n + 1$ results that $G_{\alpha_1,\alpha_2,\ldots,\alpha_n,n}(z)$ is a analytic and univalent function in $\mathcal{U}$. □

Corollary 2.2. *Let $\alpha \in \mathbb{C}, n \in \mathbb{N}, n \geq 1, \mathrm{Re}[\alpha - n + 1] > 0$ and $c \in \mathbb{C}, |c| < 1, c \neq -1$. We suppose that the function f given by (1.1) is analytic in $\mathcal{U}$. If*

$$\left|c|z|^{2(\alpha-n+1)} + (1 - |z|^{2(\alpha-n+1)}\frac{zf''(z)}{(\alpha - n + 1)f'(z)})\right| \leq 1$$

for all $z \in \mathcal{U}$, then the function

$$G_{n,\alpha}(z) = \left((\alpha - n + 1)\int_0^z (g_1(t))^{\alpha-1}\ldots(g_n(t))^{\alpha-1}dt\right)^{\frac{1}{\alpha-n+1}}$$

is analytic and univalent in $\mathcal{U}$.

Proof. Similar with the proof of previous theorem for $\alpha_1 = \alpha_2 = \cdots = \alpha_n = \alpha$ □

Remark. For $n = 1$ in Theorem 2.1 we obtain the Pescar's criterion of univalence.

Remark. For $n = 1$ and $\alpha = 1$ in Theorem 2.1 we obtain Ahlfor's and Becker's univalence criterion.

Theorem 2.3. *Let $\gamma_j \in \mathbb{C}, j = \overline{1,n}, \mathrm{Re}\left(\sum_{j=1}^{n}\frac{1}{\gamma_i}\right) > 0$ and $c \in \mathbb{C}, |c| \leq 1,\ c \neq -1$. We suppose that the function f defined by (1.1) is analytic in $\mathcal{U}$. If*

$$\left|c|z|^{2\sum_{j=1}^{n}\frac{1}{\gamma_j}} + (1 - |z|^{2\sum_{j=1}^{n}\frac{1}{\gamma_j}})\frac{zf''(z)}{\sum_{j=1}^{n}\frac{1}{\gamma_j}f'(z)}\right| \leq 1 \tag{2.6}$$

for all $z \in \mathcal{U}$, then the function $J_{\gamma_1,\gamma_2,\ldots,\gamma_n}(z)$ defined by (1.3) is analytic and univalent in $\mathcal{U}$.

Proof. From (2.6) we have that $f'(z) \neq 0$, for $z \in \mathcal{U}$. We have that the function

$$v(z, t) = c \cdot e^{-2t\sum_{j=1}^{n}\frac{1}{\gamma_j}} + \left(1 - e^{-2t\sum_{j=1}^{n}\frac{1}{\gamma_j}}\right)\frac{e^{-t}zf''(e^{-t}z)}{\sum_{j=1}^{n}\frac{1}{\gamma_j}f'(e^{-t}z)} \tag{2.7}$$

is analytic in $\overline{\mathcal{U}}$, for $t > 0$. □

For the function $v(z,t)$ we apply the maximum modulus principle and we obtain

$$\begin{aligned} |v(z,t)| < \max_{|z|=1} |v(z,t)| &= \left|v(e^{j\theta},t)\right| \\ &= \left| ce^{-2t\sum_{j=1}^{n}\frac{1}{\gamma_j}} + \left(1 - e^{-2t\sum_{j=1}^{n}\frac{1}{\gamma_j}}\right) \frac{e^{-t+j\theta} f''(e^{-t+j\theta})}{\sum_{j=1}^{n}\frac{1}{\gamma_j} f'(e^{-t+j\theta})} \right| \end{aligned} \tag{2.8}$$

where $\theta = \theta(t) \in \mathbb{R}$.
We note with $\psi = e^{-t+j\theta}$ and we have that $|\psi| = e^{-t}, \forall t > 0$.
If in (2.8)we replace $e^{-t}e^{j\theta}$ with ψ we obtain

$$\left|v(e^{j\theta},t)\right| = \left| c|\psi|^{2\sum_{j=1}^{n}\frac{1}{\gamma_j}} + \left(1 - |\psi|^{2\sum_{j=1}^{n}\frac{1}{\gamma_j}}\right) \frac{\psi \cdot f''(\psi)}{\sum_{j=1}^{n}\frac{1}{\gamma_j} f'(\psi)} \right| \tag{2.9}$$

But $|\psi| = e^{-t} < 1$ for $t > 0$ implies that $\psi \in \mathcal{U}$.
Using (2.6) and (2.9) for $z = \psi$ we obtain:

$$\left|v(e^{j\theta},t)\right| \le 1 \tag{2.10}$$

From (2.8) and (2.10) we have that $|v(z,t)| < 1$, for all $z \in \mathcal{U}, t > 0$. For $t = 0$ we obtain $v(z,0) = c$. Using the hypothesis we obtain that $|v(z,0)| < 1$ for all $z \in \mathcal{U}$. So, $|v(z,t)| < 1$ for all $z \in \mathcal{U}$ and $t \ge 0$. Hence and from Theorem 1.5 for $\beta = \sum_{j=1}^{n}\frac{1}{\gamma_j}$ we obtain that $J_{\gamma_1,\gamma_2,\ldots,\gamma_n}(z)$ is univalent and analytic in $\mathcal{U}$.

Corollary 2.4. *Let $\gamma \in \mathbb{C}, \mathrm{Re}\left(\frac{1}{\gamma}\right) > 0$ and $c \in \mathbb{C}, |c| \le 1$, $c \ne -1$. We suppose that the function f defined by (1.1) is analytic in $\mathcal{U}$. If*

$$\left| c|z|^{\frac{2}{\gamma}} + (1 - |z|^{\frac{2}{\gamma}}) \frac{zf''(z)}{\frac{1}{\gamma} f'(z)} \right| \le 1$$

for all $z \in \mathcal{U}$, then the function

$$J_\gamma(z) = \left(\frac{1}{\gamma} \int_0^z t^{-1} (f(t))^{\frac{1}{\gamma}} dt \right)^{\frac{1}{\gamma}}$$

is analytic and univalent in $\mathcal{U}$.

Remark. For $\gamma_1 = \gamma_2 = \cdots = \gamma_n = \gamma = 1$ in Theorem 2.3 we obtain Ahlfor's and Becker's criterion of univalence.

Acknowledge: This work was partially supported by the strategic project POSDRU 107/1.5/S/77265, inside POSDRU Romania 2007-2013 co-financed by the European Social Fund-Investing in People.

References

Ahlfors, L.V. (1973). Sufficient conditions for quasiconformal extension. In: *Proc. Conf., Univ. Maryland, College Park, Md., 1973. Ann. of Math. Studies, No. 79, Princeton Univ. Press, Princeton, N.J., 1974*. pp. 23–29.

Becker, J. (1972). Löwnersche differentialgleichung und quasikonform fortsetzbare schlichte funktionen. *J. Reine Angew. Math.* **255**, 23–43.

Becker, J. (1973). Löwnersche differentialgleichung und schlichtheits-kriterion. *Math. Ann.* **202**(4), 321–335.

Breaz, D., N. Breaz and H.M. Srivastava (2009). An extension of the univalent condition for a family of integral operators. *Applied Mathematics Letters* **22**(1), 41–44.

Mocanu, P.T., D. Breaz, G.I. Oros and Gh. Oros (2009). *Analiză complexă*. Editura Aeternitas, Alba Iulia.

Pascu, N.N. and I. Radomir (1989). A new generalization of ahlfor's and becker's criterion of univalence. *Seminar on Geometric function theory, Cluj-Napoca, Preprint Nr.5*.

Pescar, V. (1996). A new generalization of ahlfor's and becker's criterion of univalence. *Bull. Malaysian Math. Soc.* **19**(Second Series), 53–54.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 7–15

An Extension of Kuttner's Theorem

Vakeel A. Khan[a,*]

[a] *Department of Mathematics, A. M. U. Aligarh-202002, India.*

Abstract

If $0 < p < 1$ and X is a locally convex FK - space, then $X \supset l_\infty$ whenever $X \supset w_0(p)$ (Kuttner's theorem see (B.Thorpe, 1981)).The aim of this paper is to give some extensions of this theorem by replacing $w_0(p)$ by $[c_A, M]_0$.

Keywords: Sequence spaces, strong summability, Orlicz functions.
2000 MSC: 40F05, 46A45, 40A05.

1. Introduction

A real function g on a linear space X is called an F - norm if

[i] $g(x) = 0$ if and only if $x = 0$,
[ii] $|\alpha| \leq 1 (\alpha \in K) \Rightarrow g(\alpha x) \leq g(x)$ for all $x \in X$,
[iii] $g(x + y) \leq g(x) + g(y)$ for all $x, y \in X$,
[iv] $\lim_n \alpha_n = 0 \quad (\alpha_n \in K), x \in X \quad \Rightarrow \lim_n g(\alpha_n x) = 0$.

An F -norm g in a sequence space X is called absolutely monotone if $|x_k| \leq |y_k|, k \in \mathbb{N} \Rightarrow g(x) \leq g(y)$, for all $x = (x_k), y = (y_k) \in X$.

An F -space is defined as a complete F - normed space. If a sequence space X is an F - space on which the coordinate functionals $\pi_k(x) = x_k$ are continuous, then X is called an FK- space. An FK - space with normable topology is called a BK - space . Some authors include local convexity in the definition of a Fréchet Space and of an FK - space . We do not and we follow the definition used by Maddox and by Wilansky (Wilansky, 1964).

Let ϕ be the space of all finite sequences. An F- space X containing ϕ is called an AK- space if $x = \lim_n \sum_{k=1}^{n} x_k e_k$, for all $x = (x_k) \in X$.

For a sequence space X we denote by X^α and X^β the Köthe - Toeplitz duals of X , i.e.

$$X^\alpha = \left\{ \alpha = (\alpha_k) : \quad \sum_{k=1}^{n} |\alpha_k x_k| < \infty \text{ for all } (x_k) \in X \right\} \tag{1.1}$$

*Corresponding author
Email address: vakhan@math.com (Vakeel A. Khan)

and

$$X^{\beta} = \left\{ \alpha = (\alpha_k) : \sum_{k=1}^{n} \alpha_k x_k \text{ converges for all } (x_k) \in X \right\}. \tag{1.2}$$

For an F - normed sequence space X we denote by X' the topological dual of X and in the case $\phi \subset X$, we use the notation

$$X^{\phi} = \{ f(e_k)) : f \in X' \}. \tag{1.3}$$

A sequence space X is called solid (or normal), if $(\alpha_k x_k) \in X$, whenever $(x_k) \in X$ for all sequences of scalars (α_k) with $|\alpha_k| \leq 1, k \in \mathbb{N}$.

A sequence space X is called monotone, if X contains the canonical pre-images of all its step spaces.

Lemma 1.1. *If a sequence space X is solid then X is monotone.*

Let X and Y be any two sequence spaces and $A = (a_{nk})_{n,k=1}^{\infty}$ an infinite matrix. We say that the matrix A maps X into Y if for each $x \in X$, the sequence $Ax = (A_n(x)) \in Y$, where

$$A_n(x) = \sum_{k=1}^{\infty} a_{nk} x_k, \quad n = 1, 2, \cdots \tag{1.4}$$

provided that the series on the right converges for each n. We denote by (X, Y) the class of all matrices A which map X into Y.

Let S be a subset of a real linear space.

[a] The set S is called convex if for all $x, y \in S$

$$\lambda x + (1 - \lambda) y \in S \quad \forall\, \lambda \in [0, 1], \tag{1.5}$$

[b] If S is a nonempty and convex set , we say that a functional $f : S \to \mathbb{R}$ is convex if

$$f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y), \quad \forall\, \lambda \in [0, 1] \text{ and } \forall\; x, y \in S. \tag{1.6}$$

An Orlicz function is a function $M : [0, \infty) \to [0, \infty)$, which is continuous, non-decreasing and convex with $M(0) = 0$, $M(x) > 0$, for $x > 0$ and $M(x) \to \infty$ as $x \to \infty$. Let $0 < p \leq 1$. A function $M : [0, \infty) \to [0, \infty)$ is called $p - convex$ if

$$M(\alpha x + \beta y) \leq \alpha^p M(x) + \beta^p M(y) \tag{1.7}$$

for all $x, y \geq 0$ and $\alpha^p + \beta^p = 1$.

In this paper we consider $p - convex$ $(0 < p < 1)$ Orlicz functions. Note that the notion of $1 - convex$ functions coincides with the notion of convex functions.

Example. The function $M(t) = t^p$, $0 < p < 1$ is $p - convex$ and it is not $r - convex$ if $r > p$.

If convexity of an Orlicz function M is replaced by

$$M(x + y) \leq M(x) + M(y) \tag{1.8}$$

then this function is called a modulus function, defined and discussed by Nakano (Nakano, 1953), Ruckle (Ruckle, 1973), Maddox (Maddox, 1986) and others. Lindenstrauss and Tzafriri (Lindenstrauss & Tzafriri, 1971) used the idea of an Orlicz function to construct the sequence space

$$\ell_M = \left\{ x = (x_k) : \sum_{k=1}^{\infty} M\left(\frac{|x_k|}{\rho}\right) < \infty \text{ for some } \rho > 0 \right\}. \tag{1.9}$$

The space ℓ_M with the norm

$$\|x\| = \inf\{\rho > 0 : \sum_{k=1}^{\infty} M\left(\frac{|x_k|}{\rho}\right) \leq 1\}$$

becomes a Banach space which is called an Orlicz sequence space. For $M(x) = x^p$, $1 \leq p < \infty$, the space ℓ_M coincide with the classical sequence space l_p.

An Orlicz function M is said to satisfy the $\triangle_2$ - condition for all values x, if there exists a constant $K > 0$, such that

$$M(2x) \leq KM(x) \text{ for all } x \geq 0. \tag{1.10}$$

The $\triangle_2$ - condition is equivalent to

$$M(Lx) \leq KLM(x), \text{ for all values of } x \geq 0, \text{ and for } L > 1. \tag{1.11}$$

An Orlicz function M can always be represented in the following integral form

$$M(x) = \int_0^x \eta(t)dt, \tag{1.12}$$

where η is known as the kernel of M, is right differentiable for $t \geq 0$, $\eta(0) = 0$, $\eta(t) > 0$, η is non-decreasing and $\eta(t) \to \infty$ as $t \to \infty$. Note that an Orlicz function satisfies the inequality

$$M(\lambda x) \leq \lambda M(x) \text{ for all } \lambda \text{ with } 0 < \lambda < 1. \tag{1.13}$$

Let $A = (a_{nk})$ be an infinite matrix with $a_{nk} \geq 0$ and let c_A be the summability field of matrix method A (see Virge Soomer (Soomer, 2003))i.e.

$$c_A = \left\{ x = (x_k) : A(x) = \lim_n \sum_k a_{nk} x_k \ \ exists \right\}. \tag{1.14}$$

Then, passing to strong summability,

$$[c_A] = \left\{ x = (x_k) : \text{ there exists } L, \ \lim_n \sum_k a_{nk}|x_k - L| = 0 \right\} \tag{1.15}$$

and

$$[c_A]_0 = \left\{ x = (x_k) : \ \lim_n \sum_k a_{nk}|x_k| = 0 \right\} \tag{1.16}$$

are the spaces of strongly A - summable and strongly A - summable to zero sequences , respectively.

Thorpe (B.Thorpe, 1981) gave the following generalization of Kuttner's theorem.

Theorem 1.2. *If $0 < p < 1$ and X is a locally convex FK - space, then $X \supset l_\infty$ whenever $X \supset w_0(p)$).*

Kuttner (Kuttner, 1946) proved this result in the case $X = c_A$, where A is a regular matrix method (Kuttner's theorem).

If the matrix $A = (a_{nk})$ satisfies the condition

$$\sup_n a_{nk} > 0 \quad \text{for each} \quad k \in \mathbb{N}, \tag{1.17}$$

then $[c_A]_0$ is a solid $AK - BK$ - space with the norm

$$\|x\| = \sup_n \sum_{k=1}^{\infty} a_{nk}|x_k|. \tag{1.18}$$

Since for every solid $AK - BK$ - space X we have

$$X^\alpha = X^\beta = X^\phi, \tag{1.19}$$

this is also true for $X = [c_A, M]_0$.

For a positive matrix method $A = (a_{nk})$ Virge Soomer (Soomer, 2003) defined

$$D(A,p) = \left\{ x = (x_k) : \quad \lim_n \sum_{k=1}^{\infty} a_{nk}^{1/p}|x_k| = 0 \right\}. \tag{1.20}$$

A sequence of positive integers $\theta = (k_r)$ is called "lacunary" if $k_0 = 0$, $0 < k_r < k_{r+1}$ and $h_r = k_r - k_{r-1} \to \infty$ as $r \to \infty$. The intervals determined by θ will be denoted by $I_r = (k_{r-1}, k_r]$ and $q_r = \frac{k_r}{k_{r-1}}$. The space of lacunary strongly convergent sequences L_θ was defined by Freedman et al (Freedman *et al.*, 1978) as :

$$L_\theta = \left\{ x = (x_k) : \lim_r \frac{1}{h_r} \sum_{k \in I_r} |x_k - l| = 0 \ \text{ for some } \ 1 \right\}. \tag{1.21}$$

The space L_θ is a solid AK-BK - space with the norm

$$\|x\|_\theta = \sup_r \frac{1}{h_r} \sum_{k \in I_r} |x_k|. \tag{1.22}$$

L_θ^0 denotes the subset of L_θ those sequences for which $l = 0$ in the definition of L_θ. Then L_θ^0 is the strong null summability field of the matrix method $A_\theta = (a_{rk}^\theta)$ where

$$a_{rk}^\theta = \begin{cases} \frac{1}{h_r} & (k_r \le k \le k_{r+1} - 1, \quad r,k \in \mathbb{N}) \\ 0 & \text{otherwise} \end{cases} \tag{1.23}$$

For $\theta = (2^r)$ we have $L_\theta^0 = w_0(1)$ and the norm $\|x\|_\theta$ is equivalent to the usual norm

$$\|x\| = \sup_n \frac{1}{n+1} \sum_{k=0}^{n} |x_k| \ \text{ in } \ w_0(1) \text{ (see (Maddox, 1970))}. \tag{1.24}$$

Note that $L_\theta^0(M)$ is a solid $AK - FK$ - space with the F - norm

$$g_M(x) = \sup_r \frac{1}{h_r} \sum_{k \in I_r} M\left(\frac{|x_k|}{\rho}\right), \text{for some } \rho > 0. \tag{1.25}$$

T. Bilgin (Bilgin, 2003) defined the following sequence spaces :

$$L_\theta^0(M,p)_\triangle = \left\{ x=(x_k) : \lim_r \frac{1}{h_r} \sum_{k\in I_r} M\left(\frac{|\triangle x_k|}{\rho}\right)^{p_k} = 0, \text{ for some } \rho > 0 \right\}. \tag{1.26}$$

$$L_\theta(M,p)_\triangle = \left\{ x=(x_k) : \lim_r \frac{1}{h_r} \sum_{k\in I_r} M\left(\frac{|\triangle x_k - l|}{\rho}\right)^{p_k} = 0, \text{ for some l and } \rho > 0 \right\}. \tag{1.27}$$

$$L_\theta^\infty(M,p)_\triangle = \left\{ x=(x_k) : \sup_r \frac{1}{h_r} \sum_{k\in I_r} M\left(\frac{|\triangle x_k|}{\rho}\right)^{p_k} < \infty, \text{ for some } \rho > 0 \right\}. \tag{1.28}$$

If we take x_k instead of $\triangle x_k$ and $p_k = 1$ for all k, then we have the following sequence spaces :

$$L_\theta^0(M) = \left\{ x=(x_k) : \lim_r \frac{1}{h_r} \sum_{k\in I_r} M\left(\frac{|x_k|}{\rho}\right) = 0, \text{ for some } \rho > 0 \right\}. \tag{1.29}$$

$$L_\theta(M) = \left\{ x=(x_k) : \lim_r \frac{1}{h_r} \sum_{k\in I_r} M\left(\frac{|x_k - l|}{\rho}\right) = 0, \text{ for some l and } \rho > 0 \right\}. \tag{1.30}$$

$$L_\theta^\infty(M) = \left\{ x=(x_k) : \sup_r \frac{1}{h_r} \sum_{k\in I_r} M\left(\frac{|x_k|}{\rho}\right) < \infty, \text{ for some } \rho > 0 \right\}. \tag{1.31}$$

Virge Soomer (Soomer, 2003) defined the sequence space

$$H_\theta(p) = \left\{ \alpha = (\alpha_k) : \sum_{r=0}^{\infty} h_r^{1/p} \max_{k\in I_r} |\alpha_k| < \infty \right\}. \tag{1.32}$$

The aim of this paper is to give some extensions of Theorem 1.2 by replacing $w_0(p)$ by $[c_A, M]_0$.

2. Main results

In this paper we define the sequence space :

$$[c_A, M]_0 = \left\{ x=(x_k) : \lim_n \sum_k a_{nk} M\left(\frac{|x_k|}{\rho}\right) = 0, \text{ for some } \rho > 0 \right\}. \tag{2.1}$$

If $M(x) = x^p$, $p \geq 1$, then we have $[c_A, M]_0 = [c_A]_0^p$, the space of the sequences that are strongly A -summable to zero with index p. By taking $A = (C, 1)$, the Cesàro matrix , and for $0 < p < \infty$ the space $[c_A]_0^p$ is usually denoted by $w_0(p)$, i.e.

$$w_0(p) = \left\{ x=(x_k) : \quad \lim_n \frac{1}{n+1} \sum_{k=0}^{n} |x_k|^p = 0 \right\}. \tag{2.2}$$

Theorem 2.1. *Let M be an Orlicz function and let $A = (a_{nk})$ be positive regular matrix method with finite rows satisfying the conditions*

$$\sup_n a_{nk} > 0 \quad \textit{for each} \;\; k \in \mathbb{N}, \tag{2.3}$$

and

$$\sum_{k=1}^{\infty} a_{nk} = 1 \quad \textit{for each} \;\; n \in \mathbb{N}. \tag{2.4}$$

Then the following statements hold :

[i] $D(A, p)$ is a solid $AK - BK$ - space with the norm

$$q(x) = \sup_n \sum_{k=1}^{\infty} a_{nk}^{1/p} |x_k|. \tag{2.5}$$

[ii] If M is p - convex , then $[c_A, M]_0 \subset D(A, p)$.

[iii] $l_\infty \subset D(A, p)$ if and only if $\lim_n \sum_{k=1}^{\infty} a_{nk}^{1/p} = 0$.

Proof.

[i] The proof is straightforward.

[ii] Since M is p - convex and $\alpha_k \geq 0$, $\sum_{k=1}^{n} \alpha_k^p = 1$, $t_k \geq 0$, then

$$M\left(\sum_{k=1}^{n} \alpha_k t_k\right) \leq \sum_{k=1}^{n} \alpha_k^p M(t_k). \tag{2.6}$$

Putting $\alpha_k = a_{nk}^{1/p}$ and $t_k = \frac{|x_k|}{\rho}$ we get (note that the matrix A has finite rows and satisfies $\sum_{k=1}^{\infty} a_{nk} = 1$ for each $n \in \mathbb{N}$)

$$M\left(\sum_{k=1}^{\infty} a_{nk}^{1/p} \frac{|x_k|}{\rho}\right) \leq \sum_{k=1}^{\infty} a_{nk} M\left(\frac{|x_k|}{\rho}\right). \tag{2.7}$$

Then [ii] follows by the properties of Orlicz functions.

[iii] It is clear that (see (Boos, 2000), Theorem 2.4.1(of Schur)) that the matrix method $A_p = (a_{nk}^{1/p})$ sums all bounded sequences if and only if

$$\lim_n \sum_{k=1}^{\infty} a_{nk}^{1/p} = 0. \tag{2.8}$$

Theorem 2.2. *Let X be a locally convex FK - space. If the matrix method A and the Orlicz function M satisfy conditions of Theorem 2.1 and*

$$([c_A, M]_0)^\phi \subset (D(A, p))^\phi, \quad \textit{then the condition} \tag{2.9}$$

$$\lim_n \sum_{k=1}^{\infty} a_{nk}^{1/p} = 0 \quad \textit{is sufficient for} \tag{2.10}$$

$$X \supset [c_A, M]_0 \Longrightarrow X \supset l_\infty. \tag{2.11}$$

Proof. Let

$$X \supset [c_A, M]_0, \text{ then } X^\phi \subset ([c_A, M]_0)^\phi \tag{2.12}$$

and

$$X^\phi \subset (D(A,p))^\phi, \ (\text{since } ([c_A, M]_0)^\phi \subset (D(A,p))^\phi). \tag{2.13}$$

Since the BK - space $D(A, p)$ is an AK - space and hence also an AD - space. (i.e. ϕ is dence in $D(A, p)$), $X \supset D(A, p)$ follows from Theorem 4 of (Snyder & Wilansky, 1972).Thus , by Theorem 1[ii], we get $X \supset l_\infty$.

Theorem 2.3. *Let M be an unbounded p - convex Orlicz function satisfying the condition*

$$M(t^{\frac{1}{p}}) = O(t), t \to \infty. \tag{2.14}$$

Then

$$(L_\theta^0(M))^\alpha = H_\theta(p). \tag{2.15}$$

Proof.

Suppose that $x = (x_k) \in L_\theta^0(M)$,$\alpha = (\alpha_k) \in H_\theta(p)$ and let M^{-1} be the inverse function of M. Let $A_{rk} = |\alpha_k| h_r^{1/p}$ $(r, k \in \mathbb{N})$. Then

$$\sum_{k\in I_r} |\alpha_k x_k| \le \max_{k\in I_r} A_{rk} (h_r^{1/p})^{-1} \sum_{k\in I_r} |x_k| = \rho \max_{k\in I_r} M^{-1}\left[M\left((h_r^{1/p})^{-1} \sum_{k\in I_r} \frac{|x_k|}{\rho}\right)\right]. \tag{2.16}$$

By applying p-convexity of M we have

$$\sum_{k\in I_r} |\alpha_k x_k| \le \max_{k\in I_r} A_{rk} M^{-1}\left[(h_r)^{-1} \sum_{k\in I_r} M\left(\frac{|x_k|}{\rho}\right)\right] = \max_{k\in I_r} A_{rk} M^{-1}[g_M(x)]. \tag{2.17}$$

and

$$\sum_{r=0}^{\infty} |\alpha_r x_r| = \sum_{r=0}^{\infty} \sum_{k\in I_r} |\alpha_k x_k| \le M^{-1}[g_M(x)] \sum_{r=0}^{\infty} (h_r^{1/p}) \max_{k\in I_r} |\alpha_k| < \infty. \tag{2.18}$$

Hence $\alpha = (\alpha_k) \in (L_\theta^0(M))^\alpha$ and thus $H_\theta(p) \subset (L_\theta^0(M))^\alpha$.

Now suppose that $\alpha = (\alpha_k) \notin H_\theta(p)$. Then the series in $H_\theta(p)$ is divergent , and therefore there exists a sequence (c_r), $0 < c_r \to 0$, $r \to \infty$ such that

$$\sum_{r=0}^{\infty} c_r (h_r^{1/p}) \max_{k\in I_r} |\alpha_k| = \infty. \tag{2.19}$$

Let $\max_{k\in I_r} |\alpha_k| = |\alpha_{k_r}|$ and let $\bar{x} = (\bar{x}_k)$ be defined by

$$\bar{x}_k = \begin{cases} \rho c_r h_r^{1/p} & \text{for } k = k_r \\ 0 & \text{for } k \ne k_r \quad r, k \in \mathbb{N} \end{cases} \tag{2.20}$$

Since $c_r \to 0$, $r \to \infty$, we have $c_r < 1$ for sufficiently large r. Now by convexity of M, by the definition $M(0) = 0$ and by the given condition (2.14) we have

$$(h_r)^{-1} \sum_{k\in I_r} M\left(\frac{|x_k|}{\rho}\right) = (h_r)^{-1} M(c_r h_r^{1/p}) \le \frac{c_r^p M(h_r^{1/p})}{h_r} = O(1), \text{ as } r \to \infty. \tag{2.21}$$

Hence $x \in L^0_\theta(M)$.

But

$$\sum_{k\in I_r} |\alpha_k \bar{x}_k| = |\alpha_{k_r}| c_r h_r^{1/p} \tag{2.22}$$

so that by (3.2) the series $\sum_{k=0}^{\infty} |\alpha_k \bar{x}_k|$ diverges and therefore $(\alpha_k) \notin (L^0_\theta(M))^\alpha$.This completes the proof.

Theorem 2.4. *Let X be a locally convex FK - space and let M be an unbounded p - convex Orlicz function satisfying the condition*

$$H_\theta(p) = \left\{ \alpha = (\alpha_k) : \lim_r \frac{1}{h_r^{1/p}} \max_{k\in I_r} |\alpha_k| < \infty \right\}. \tag{2.23}$$

Then the following statements holds :

[i] $(L^0_\theta(M))^\alpha \subset (D(A_\theta, p))^\alpha$,

[ii] $X \supset L^0_\theta(M) \Longrightarrow X \supset l_\infty$.

Proof.

[i] Since $L^0_\theta(M)$ and $D(A_\theta, p)$ are solid $AK - FK$ spaces. This implies that their α - duals and ϕ - duals are equal and so it is sufficient to prove $(L^0_\theta(M))^\phi \subset (D(A_\theta, p))^\phi$,. By Theorem 2.3 it is sufficient to show $H_\theta(p) \subset (D(A_\theta, p))^\alpha$.

Suppose that $\alpha = (\alpha_k) \in H_\theta(p)$, then for each $x = (x_k) \in D(A_\theta, p)$ we have

$$\sum_{r=0}^{\infty} |\alpha_r x_r| = \sum_{r=0}^{\infty} \sum_{k\in I_r} |\alpha_k x_k| \tag{2.24}$$

$$\leq \sum_{r=0}^{\infty} h_r^{1/p} \max_{k\in I_r} |\alpha_k| \frac{1}{h_r^{1/p}} \sum_{k\in I_r} |x_k| \leq q(x) \sum_{r=0}^{\infty} h_r^{1/p} \max_{k\in I_r} |\alpha_k| < \infty. \tag{2.25}$$

This implies that $(\alpha_k) \in (D(A_\theta, p))^\phi$.

[ii] The matrix $A_\theta = (a^\theta_{nk})$ satisfies conditions of Theorem 2.1, $(L^0_\theta(M))^\phi \subset (D(A_\theta, p))^\phi$ by [i] and

$$\lim_r (a^\theta_{nk})^{1/p} = \lim_r h_r^{1-1/p} = 0. \tag{2.26}$$

Consequently proof of [ii] follows immediately by Theorem 2.2.

Acknowledge: The author would like to record their gratitude to the reviewer for his careful reading and making some useful corrections which improved the presentation of the paper.

References

Bilgin, T. (2003). Some new difference sequence spaces defined by an Orlicz function. *Filomat* **17**, 1–8.

Boos, J. (2000). *Classical and Modern Methods in Summability*. Oxford Unoversity Press.

B.Thorpe (1981). An extension of Kuttner's theorem. *Bull. London Math. Soc.* (13), 301–302.

Freedman, A.R., J.J. Sember and M. Raphael (1978). Some Cesaro - type summability spaces. In: *Proc. London Math.Soc.(37(3))*. pp. 508–520.

Kuttner, B. (1946). Note on strong summability. *J. London Math. Soc.* (21), 118–122.

Lindenstrauss, J. and L. Tzafriri (1971). On Orlicz sequence spaces. *Israel J. Math.* (10), 379–390).

Maddox, I.J. (1970). *Elements of functional analysis*. Cambridge university press, London - New York.

Maddox, I.J. (1986). Sequence spaces defined by a modulus. *Math. Camb. Phil. Soc.* (100), 161–166.

Nakano, H. (1953). Concave modulus. *J. Math Soc. Japan* **5**(1), 29–49.

Ruckle, W.H. (1973). FK spaces in which the sequence of coordinate vectors is bounded. *Canad. J.Math.* (25), 973–975.

Snyder, A. K. and A. Wilansky (1972). Inclusion theorems and semiconservative FK spaces. *Rocky Mountain J. Math.* (2), 595–603.

Soomer, V. (2003). Strong summability defined by p - convex modulus functions and Kuttner's theorem. *Acta Comment. Univ. Tartu* (7), 71–78.

Wilansky, A. (1964). *Functional Analysis*. Blaisdail Publishing Co. Ginn and Co., New York - Toronto -London.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 16–21

Univalence Conditions for a New Integral Operator

Laura Stanciu[a,*]

[a] *University of Piteşti, Department of Mathematics, Târgul din Vale Str., No.1, 110040, Piteşti, Argeş, România.*

Abstract

In this paper, we study the univalence conditions for a new integral operator defined by Al-Oboudi differential operator. Many known univalence conditions are written to prove our main results.

Keywords: Analytic functions, general Schwarz Lemma, differential operator.
2000 MSC: 30C45, 30C75.

1. Introduction and Preliminaries

Let $\mathcal{A}$ denote the class of all functions of the form

$$f(z) = z + \sum_{k=2}^{\infty} a_k z^k, \tag{1.1}$$

which are analytic in the open unit disk

$$\mathbb{U} = \{z \in \mathbb{C} : |z| < 1\} \tag{1.2}$$

and satisfy the following usual normalization condition

$$f(0) = f'(0) - 1 = 0. \tag{1.3}$$

We denote by $\mathcal{S}$ the subclass of $\mathcal{A}$ consisting of functions f which are univalent in $\mathbb{U}$.
For $f \in \mathcal{A}$, Al-Oboudi (Al-Oboudi, 2004) introduced the following operator

$$D^0 f(z) = f(z), \tag{1.4}$$

$$D^1 f(z) = (1 - \delta) f(z) + \delta z f'(z), \qquad \delta \geq 0 \tag{1.5}$$

*Corresponding author
Email address: laura_stanciu_30@yahoo.com (Laura Stanciu)

...

$$D^n f(z) = D_\delta\left(D^{n-1} f(z)\right), \qquad (n \in \mathbb{N} = \{1, 2, ...\}), \tag{1.6}$$

If f is given by (1.1), then from (1.5) and (1.6) we see that

$$D^n f(z) = z + \sum_{k=2}^{\infty} [1 + (k-1)\,\delta]^n a_k z^k, \qquad (n \in \mathbb{N}_0 = \mathbb{N} \cup 0), \tag{1.7}$$

with $D^n f(0) = 0$.

Remark. When $\delta = 1$, we get Sălăgean differential operator (Sălăgean, 1983).

Here, in our present investigation, we introduce a new general integral operator by means of the Al-Oboudi differential operator as follows

$$F_{n,m,\beta}(z) = \left(\beta \int_0^z t^{\beta-1} \prod_{i=1}^{n} \left(\frac{D^m f_i(t)}{t}\right)^{\alpha_i} \left(e^{g_i(t)}\right)^{\gamma_i} dt\right)^{\frac{1}{\beta}} \tag{1.8}$$

$\alpha_i, \gamma_i \in \mathbb{C}, \beta \in \mathbb{C} - \{0\}$, $f_i, g_i \in \mathcal{A}$ for all $i \in \{1, 2, ..., n\}$, D^m is the Al-Oboudi differential operator, $m \in \mathbb{N}_0$ and $\frac{D^m f_i(z)}{z} \neq 0$.

In this paper, we study the univalence conditions involving the general integral operator defined by (1.8).

In the proof of our main results (Theorem 2.1) we need the following univalence criterion. The univalence criterion, asserted by Theorem 1.1, is a generalization of Ahlfors's and Becker's univalence criterion; it was proven by Pescar (Pescar, 1996).

Theorem 1.1. (Pescar, 1996) *Let β be a complex number, $Re\beta > 0$, and c a complex number, $|c| \leq 1$, $c \neq -1$ and $f(z) = z + ...$ a regular function in U. If*

$$\left|c\,|z|^{2\beta} + \left(1 - |z|^{2\beta}\right) \frac{zf''(z)}{\beta f'(z)}\right| \leq 1, \tag{1.9}$$

for all $z \in U$, then the function

$$F_\beta(z) = \left(\beta \int_0^z t^{\beta-1} f'(t) dt\right)^{\frac{1}{\beta}} = z + ... \tag{1.10}$$

is regular and univalent in U.

In (Yang & Liu, 1999) is defined the class $\mathcal{S}(p)$. For $0 < p \leq 2$, let $\mathcal{S}(p)$ denote the class of functions $f \in \mathcal{A}$ which satisfies the conditions $f(z) \neq 0, (0 < |z| < 1)$ and $\left|\left(\frac{z}{f(z)}\right)''\right| \leq p, (z \in \mathbb{U})$. Also, if $f \in \mathcal{S}(p)$ then the following property is true

$$\left|\frac{z^2 f'(z)}{[f(z)]^2} - 1\right| \leq p\,|z|^2, (z \in \mathbb{U}) \tag{1.11}$$

relation proved in (Singh, 2000).

Finally, in our present investigation, we shall also need the familiar Schwarz Lemma (see, for details, (Nehari, 1952)).

Lemma 1.2. (Nehari, 1952) *Let the function f be regular in the disk $\mathbb{U}_R = \{z \in \mathbb{C} : |z| < R\}$, with $|f(z)| < M$ for fixed M. If f has one zero with multiplicity order bigger or equal to m for $z = 0$, then*

$$|f(z)| \leq \frac{M}{R^m} |z|^m , (z \in \mathbb{U}_R) . \tag{1.12}$$

The equality can hold only if

$$f(z) = e^{i\theta} \frac{M}{R^m} z^m, \tag{1.13}$$

where θ is constant.

2. Main Results

Theorem 2.1. *Let the functions $f_i \in \mathcal{A}$ satisfy the conditions*

$$\left| \frac{z^2 (D^m f_i(z))'}{[D^m f_i(z)]^2} - 1 \right| \leq p_i |z|^2 , (z \in \mathbb{U}; 0 < p_i \leq 2) , \tag{2.1}$$

$$\frac{D^m f_i(z)}{z} \neq 0, (z \in \mathbb{U}; m \in \mathbb{N}_0) \tag{2.2}$$

and $g_i \in \mathcal{A}$ with

$$\left| \frac{z g_i'(z)}{g_i(z)} - 1 \right| \leq 1, (z \in \mathbb{U}) \tag{2.3}$$

for all $i \in \{1, 2, ..., n\}$. Also, let $\alpha_i, \gamma_i, \beta$ be complex numbers with the property

$$\mathrm{Re}\beta \geq \sum_{i=1}^{n} [\, |\alpha_i| ((1 + p_i) M_i + 1) + 2 |\gamma_i| N_i] > 0, (i \in \{1, 2, ..., n\}) . \tag{2.4}$$

If for all $i \in \{1, 2, ..., n\}$

$$\left| D^m f_i(z) \right| \leq M_i, (z \in \mathbb{U}; M_i \geq 1; m \in \mathbb{N}_0) , \tag{2.5}$$

$$|g_i(z)| \leq N_i \, (z \in \mathbb{U}; N_i \geq 1) \tag{2.6}$$

and

$$|c| \leq 1 - \frac{1}{\mathrm{Re}\beta} \sum_{i=1}^{n} [\, |\alpha_i| ((p_i + 1) M_i + 1) + 2 |\gamma_i| N_i] \tag{2.7}$$

then the integral operator $F_{n,m,\beta}(z)$ defined by (1.8) is in the class $\mathcal{S}$.

Proof. By (1.7), we have

$$\frac{D^m f_i(z)}{z} = 1 + \sum_{k=2}^{\infty} [1 + (k-1)\, \delta]^m a_{k,i} \, z^{k-1}, (m \in \mathrm{N}_0) \tag{2.8}$$

for all $i \in \{1, 2, ..., n\}$. Define a function

$$h(z) = \int_0^z \prod_{i=1}^{n} \left(\frac{D^m f_i(t)}{t} \right)^{\alpha_i} \left(e^{g_i(t)} \right)^{\gamma_i} dt, \tag{2.9}$$

then we have $h(0) = h'(0) - 1 = 0$. Also a simple computation yields

$$h'(z) = \prod_{i=1}^{n} \left(\frac{D^m f_i(z)}{z} \right)^{\alpha_i} \left(e^{g_i(z)} \right)^{\gamma_i} \tag{2.10}$$

and

$$\frac{zh''(z)}{h'(z)} = \sum_{i=1}^{n} \left[\alpha_i \left(\frac{z\left(D^m f_i(z)\right)'}{D^m f_i(z)} - 1 \right) + \gamma_i z g_i'(z) \right]. \tag{2.11}$$

From the equation (2.11), we have

$$\begin{aligned}
\left| c\,|z|^{2\beta} + \left(1 - |z|^{2\beta}\right) \frac{zh''(z)}{\beta h'(z)} \right| &= \left| c\,|z|^{2\beta} + \left(1 - |z|^{2\beta}\right) \frac{1}{\beta} \sum_{i=1}^{n} \left(\alpha_i \left(\frac{z\left(D^m f_i(z)\right)'}{D^m f_i(z)} - 1 \right) + \gamma_i z g_i'(z) \right) \right| \\
&\le |c| + \frac{1}{|\beta|} \sum_{i=1}^{n} \left(|\alpha_i| \left(\left| \frac{z\left(D^m f_i(z)\right)'}{D^m f_i(z)} \right| + 1 \right) + |\gamma_i| \left| z g_i'(z) \right| \right) \\
&\le |c| + \frac{1}{|\beta|} \sum_{i=1}^{n} \left[|\alpha_i| \left(\left| \frac{z^2\left(D^m f_i(z)\right)'}{\left[D^m f_i(z)\right]^2} \right| \left| \frac{D^m f_i(z)}{z} \right| + 1 \right) \right. \\
&\qquad \left. + |\gamma_i| \left| \frac{z g_i'(z)}{g_i(z)} \right| |g_i(z)| \right].
\end{aligned} \tag{2.12}$$

From the hypothesis, we have

$$\left| D^m f_i(z) \right| \le M_i, \, (z \in \mathbb{U}) \, , |g_i(z)| \le N_i, \, (z \in \mathbb{U}) \, ,$$

then by the General Schwarz Lemma for the functions f_i $(i \in \{1, 2, ..., n\})$, we obtain

$$\left| D^m f_i(z) \right| \le M_i \,|z| \, , (z \in \mathbb{U}; i \in \{1, 2, ..., n\}) \, .$$

We apply this result in the inequality (2.12) and from (2.1), (2.3) we obtain

$$\begin{aligned}
\left| c\,|z|^{2\beta} + \left(1 - |z|^{2\beta}\right) \frac{zh''(z)}{\beta h'(z)} \right| &\le |c| + \frac{1}{|\beta|} \sum_{i=1}^{n} \left[|\alpha_i| \left(\left(\left| \frac{z^2\left(D^m f_i(z)\right)'}{\left[D^m f_i(z)\right]^2} - 1 \right| + 1 \right) M_i + 1 \right) \right. \\
\left. + |\gamma_i| \left(\left| \frac{z g_i'(z)}{g_i(z)} - 1 \right| + 1 \right) N_i \right] &\le |c| + \frac{1}{|\beta|} \sum_{i=1}^{n} \left[\, |\alpha_i| \left(\left(p_i \,|z|^2 + 1 \right) M_i + 1 \right) + 2\,|\gamma_i|\, N_i \right] \\
&\le |c| + \frac{1}{\mathrm{Re}\beta} \sum_{i=1}^{n} \left[\, |\alpha_i| \left((p_i + 1) \, M_i + 1 \right) + 2\,|\gamma_i|\, N_i \right].
\end{aligned} \tag{2.13}$$

So, from (2.7) we have

$$\left| c\,|z|^{2\beta} + \left(1 - |z|^{2\beta}\right) \frac{zh''(z)}{\beta h'(z)} \right| \le 1. \tag{2.14}$$

Applying Theorem 1.1, we obtain that $F_{n,m,\beta}(z)$ is in the class $\mathcal{S}$. □

If we set $m = 0$ in Theorem 2.1, we can obtain the following interesting consequence of this theorem.

Corollary 2.2. *Let $f_i \in \mathcal{A}$ satisfy the condition*

$$\left|\frac{z^2 f_i'(z)}{[f_i(z)]^2} - 1\right| \le p_i |z|^2, (z \in \mathbb{U}; 0 < p_i \le 2) \tag{2.15}$$

and $g_i \in \mathcal{A}$ with

$$\left|\frac{z g_i'(z)}{g_i(z)} - 1\right| \le 1, (z \in \mathbb{U}) \tag{2.16}$$

for all $i \in \{1, 2, ..., n\}$. Also, let $\alpha_i, \gamma_i, \beta$ be complex numbers with the property

$$\mathrm{Re}\beta \ge \sum_{i=1}^{n} [\,|\alpha_i|\,((1 + p_i)\,M_i + 1) + 2\,|\gamma_i|\,N_i] > 0, (i \in \{1, 2, ..., n\}). \tag{2.17}$$

If for all $i \in \{1, 2, ..., n\}$

$$|f_i(z)| \le M_i\,(z \in \mathbb{U}; M_i \ge 1), \tag{2.18}$$

$$|g_i(z)| \le N_i\,(z \in \mathbb{U}; N_i \ge 1) \tag{2.19}$$

and

$$|c| \le 1 - \frac{1}{\mathrm{Re}\beta} \sum_{i=1}^{n} [\,|\alpha_i|\,((p_i + 1)\,M_i + 1) + 2\,|\gamma_i|\,N_i] \tag{2.20}$$

then the integral operator

$$F_{n,\beta}(z) = \left(\beta \int_0^z t^{\beta-1} \prod_{i=1}^{n} \left(\frac{f_i(t)}{t}\right)^{\alpha_i} \left(e^{g_i(t)}\right)^{\gamma_i} dt\right)^{\frac{1}{\beta}} \tag{2.21}$$

is in the class S.

Setting $n = 1$ in Theorem 2.1 we have:

Corollary 2.3. *Let $f \in \mathcal{A}$ satisfies the conditions*

$$\left|\frac{z^2 (D^m f(z))'}{[D^m f(z)]^2} - 1\right| \le p\,|z|^2, (z \in \mathbb{U}; 0 < p \le 2), \tag{2.22}$$

$$\frac{D^m f(z)}{z} \ne 0, (z \in \mathbb{U}; m \in \mathbb{N}_0) \tag{2.23}$$

and $g \in \mathcal{A}$ with

$$\left|\frac{z g'(z)}{g(z)} - 1\right| \le 1, (z \in \mathbb{U}). \tag{2.24}$$

Also, let α, γ, β be complex numbers with the property

$$\mathrm{Re}\beta \ge [\,|\alpha|\,((1 + p)\,M + 1) + 2\,|\gamma|\,N] > 0. \tag{2.25}$$

If

$$\left|D^m f(z)\right| \le M, (z \in \mathbb{U}; M \ge 1; m \in \mathbb{N}_0), \tag{2.26}$$

$$|g(z)| \leq N, (z \in \mathbb{U}; N \geq 1)\,, \tag{2.27}$$

and

$$|c| \leq 1 - \frac{1}{\mathrm{Re}\beta}[\,|\alpha|\,((p+1)\,M+1) + 2\,|\gamma|\,N] \tag{2.28}$$

then the integral operator

$$F_{m,\beta}(z) = \left(\beta \int_0^z t^{\beta-1} \left(\frac{D^m f(t)}{t}\right)^{\alpha} \left(e^{g(t)}\right)^{\gamma} dt\right)^{\frac{1}{\beta}} \tag{2.29}$$

is in the class S.

If we set $m = 0$ in Corollary 2.3 we have another interesting consequence:

Corollary 2.4. *Let $f \in \mathcal{A}$ satisfies the condition (1.11) and $g \in \mathcal{A}$ with*

$$\left|\frac{zg'(z)}{g(z)} - 1\right| \leq 1 \qquad (z \in \mathbb{U})\,. \tag{2.30}$$

Also, let α, γ, β be complex numbers with the property

$$\mathrm{Re}\beta \geq [\,|\alpha|\,((1+p)\,M+1) + 2\,|\gamma|\,N] > 0. \tag{2.31}$$

If

$$|f(z)| \leq M, (z \in \mathbb{U}; M \geq 1)\,, \tag{2.32}$$

$$|g(z)| \leq N, (z \in \mathbb{U}; N \geq 1) \tag{2.33}$$

and

$$|c| \leq 1 - \frac{1}{\mathrm{Re}\beta}[\,|\alpha|\,((p+1)\,M+1) + 2\,|\gamma|\,N] \tag{2.34}$$

then the integral operator

$$F_{\beta}(z) = \left(\beta \int_0^z t^{\beta-1} \left(\frac{f(t)}{t}\right)^{\alpha} \left(e^{g(t)}\right)^{\gamma} dt\right)^{\frac{1}{\beta}} \tag{2.35}$$

is in the class $\mathcal{S}$.

Acknowledgments: This work was partially supported by the strategic project POSDRU 107/1.5/S/77265, inside POSDRU Romania 2007-2013 co-financed by the European Social Fund-Investing in People.

References

Al-Oboudi, F. M. (2004). On univalent functions defined by generalized Sălăgean. *Int. J. Math. Math. Sci.* (25–28), 1429–1436.

Nehari, Z. (1952). *Conformal mapping*. McGraw-Hill Book Comp.

Pescar, V. (1996). A new generalization of Ahlfors's and Becker's criterion of univalence. *Bull. Malaysian Math. Soc.* (19), 53–54.

Singh, V. (2000). On a class of univalent functions. *Internat. J. Math. Math. Sci.* **23**, 855–857.

Sălăgean, G.S. (1983). Subclasses of univalent functions. *Lecture Notes in Math., Springer Verlage* **1013**, 362–372.

Yang, D. and J. Liu (1999). On a class of univalent functions. *IJMMS* **22**(3), 605–610.

Theory and Applications of
Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 22–30

On Some I-Convergent Sequence Spaces Defined by a Modulus Function

Vakeel A. Khan[a,*], Khalid Ebadullah[a]

[a]*Department of Mathematics, A. M. U. Aligarh-202002, India.*

Abstract

In this article we introduce the sequence spaces $c_0^I(f)$, $c^I(f)$ and $l_\infty^I(f)$ for a modulus function f and study some of the properties of these spaces.

Keywords: Ideal, filter,modulus function, Lipschitz function, I-convergence field, I-convergent, monotone and solid spaces.
2000 MSC: 40A05, 40A35, 40C05, 46A45.

---

1. Introduction

Throughout the article $\mathbb{N}, \mathbb{R}, \mathbb{C}$ and ω denotes the set of natural,real,complex numbers and the class of all sequences respectively. The notion of the statistical convergence was introduced by (Fast, 1951). Later on it was studied by (Fridy, 1985) and (Fridy, 1993) from the sequence space point of view and linked it with the summability theory. The notion of I-convergence is a generalization of the statistical convergence. At the initial stage it was studied by Kostyrko, Salat and Wilezynski in (Kostyrko *et al.*, 2000). Later on it was studied by Salat, Tripathy and Ziman in (Šalát *et al.*, 2004) and Demirci in (Demirci, 2001).

Here we give some preliminaries about the notion of I-convergence. Let X be a non empty set. A set $I \subseteq 2^X$ (2^X denoting the power set of X) is said to be an ideal if I is additive i.e $A, B \in I \Rightarrow A \cup B \in I$ and hereditary i.e $A \in I, B \subseteq A \Rightarrow B \in I$.

A non-empty family of sets $\pounds(I) \subseteq 2^X$ is said to be filter on X if and only if $\Phi \notin \pounds(I)$,for $A, B \in \pounds(I)$ we have $A \cap B \in \pounds(I)$ and for each $A \in \pounds(I)$ and $A \subseteq B$ implies $B \in \pounds(I)$.

An Ideal I$\subseteq 2^X$ is called non-trivial if $I \neq 2^X$.

A non-trivial ideal $I \subseteq 2^X$ is called admissible if $\{\{x\} : x \in X\} \subseteq I$.

A non-trivial ideal I is maximal if there cannot exist any non-trivial ideal $J \neq I$ containing I as a subset.

For each ideal I, there is a filter $\pounds(I)$ corresponding to I. i.e $\pounds(I) = \{K \subseteq N : K^c \in I\}$,where $K^c = N - K$.

The idea of modulus was structured in 1953 by Nakano. (See (Nakano, 1953)).

A function $f : [0, \infty) \longrightarrow [0, \infty)$ is called a modulus if:

(1) $f(t) = 0$ if and only if $t = 0$,

---

*Corresponding author
Email addresses: `vakhan@math.com` (Vakeel A. Khan), `khalidebadullah@gmail.com` (Khalid Ebadullah)

(2) $f(t+u) \leq f(t) + f(u)$ for all $t, u \geq 0$,
(3) f is nondecreasing, and,
(4) f is continuous from the right at zero.

Ruckle in (Ruckle, 1968) used the idea of a modulus function f to construct the sequence space

$$X(f) = \{x = (x_k) : (f(|x_k|)) \in X\} \tag{1.1}$$

This space is an FK space, and (Ruckle, 1967) proved that the intersecton of all such $X(f)$ spaces is ϕ, the space of all finite sequences.

The space $X(f)$ is closely related to the space l_1 which is an $X(f)$ space with $f(x) = x$ for all real $x \geq 0$. Thus (Ruckle, 1973) proved that, for any modulus f.

$$X(f) \subset l_1 \text{ and } X(f)^{\alpha} = l_{\infty} \tag{1.2}$$

Where

$$X(f)^{\alpha} = \{y = (y_k) \in \omega : \sum_{k=1}^{\infty} f(|y_k x_k|) < \infty\} \tag{1.3}$$

The space $X(f)$ is a Banach space with respect to the norm

$$\|x\| = \sum_{k=1}^{\infty} f(|x_k|) < \infty.\text{(See (Ruckle, 1973)).} \tag{1.4}$$

Spaces of the type $X(f)$ are a special case of the spaces structured by B.Gramsch in (Gramsch, n.d.). From the point of view of local convexity, spaces of the type $X(f)$ are quite pathological. Therefore symmetric sequence spaces, which are locally convex have been frequently studied by D. J. H Garling in (Garling, 1966) and (Garling, 1968), G. Kothe in (Köthe, 1970)and W. H. Ruckle in (Ruckle, 1968) and (Ruckle, 1967).

Definition 1.1. A sequence space E is said to be solid or normal if $(x_k) \in$ E implies $(\alpha_k x_k) \in E$ for all sequence of scalars (α_k) with $|\alpha_k| < 1$ for all $k \in \mathbb{N}$.

Definition 1.2. A sequence space E is said to be monotone if it contains the cannonical preimages of all its stepspaces.

Definition 1.3. A sequence space E is said to be covergence free if $(y_k) \in E$ whenever $(x_k) \in E$ and $x_k = 0$ implies $y_k = 0$.

Definition 1.4. A sequence space E is said to be a sequence algebra if $(x_k y_k) \in E$ whenever $(x_k) \in E$, $(y_k) \in E$.

Definition 1.5. A sequence space E is said to be symmetric if $(x_{\pi(k)}) \in E$ whenever $(x_k) \in E$ where $\pi(k)$ is a permutation on $\mathbb{N}$.

Definition 1.6. A sequence $(x_k) \in \omega$ is said to be I-convergent to a number L if for every $\epsilon > 0$. $\{k \in N : |x_k - L| \geq \epsilon\} \in I$. In this case we write $I - \lim x_k = L$.

The space c^I of all I-convergent sequences to L is given by

$$c^I = \{(x_k) \in \omega : \{k \in \ : |x_k - L| \geq \epsilon\} \in I, \text{for some } L \in \mathbb{C}\}. \tag{1.5}$$

Definition 1.7. A sequence $(x_k) \in \omega$ is said to be I-null if $L = 0$. In this case we write $I - \lim x_k = 0$.

Definition 1.8. A sequence $(x_k) \in \omega$ is said to be I-cauchy if for every $\epsilon > 0$ there exists a number $m = m(\epsilon)$ such that $\{k \in N : |x_k - x_m| \geq \epsilon\} \in I$.

Definition 1.9. A sequence$(x_k) \in \omega$ is said to be I-bounded if there exists M >0 such that $\{k \in N : |x_k| > M\} \in I$

Definition 1.10. A modulus function f is said to satisfy $\triangle_2$ condition if for all values of u there exists a constant $K > 0$ such that $f(Lu) \leq KLf(u)$ for all values of $L > 1$.

Definition 1.11. Take for I the class I_f of all finite subsets of $\mathbb{N}$. Then I_f is a non-trivial admissible ideal and I_f convergence coincides with the usual convergence with respect to the metric in X.(see (Kostyrko *et al.*, 2000)).

Definition 1.12. For $I = I_\delta$ and $A \subset \mathbb{N}$ with $\delta(A) = 0$ respectively. I_δ is a non-trivial admissible ideal, I_δ-convergence is said to be logarithmic statistical covergence.(see (Kostyrko *et al.*, 2000)).

Definition 1.13. A map $\hbar$ defined on a domain $D \subset X$ i.e $\hbar : D \subset X \to \mathbb{R}$ is said to satisfy Lipschitz condition if $|\hbar(x) - \hbar(y)| \leq K|x - y|$ where K is known as the Lipschitz constant. The class of K-Lipschitz functions defined on D is denoted by $\hbar \in (D, K)$ (see (Šalát *et al.*, 2004)).

Definition 1.14. A convergence field of I-covergence is a set

$$F(I) = \{x = (x_k) \in l_\infty : \text{there exists } I - \lim x \in \mathbb{R}\}. \tag{1.6}$$

The convergence field $F(I)$ is a closed linear subspace of l_∞ with respect to the supremum norm, $F(I) = l_\infty \cap c^I$ (See (Šalát *et al.*, 2004)).

Define a function $\hbar : F(I) \to \mathbb{R}$ such that $\hbar(x) = I - \lim x$, for all $x \in F(I)$, then the function $\hbar : F(I) \to \mathbb{R}$ is a Lipschitz function (see (Šalát *et al.*, 2004)).
(c.f (Connor & Kline, 1996), (Dems, 2005), (Gurdal, 2004), (Jones & Retherford, 1967), (Kamthan & Gupta, 1980), (Maddox, 1970), (Maddox, 1986), (Maddox, 1969), (Šalát, 1980), (Singer, 1970), (Wilansky, 1964))

Throughout the article $l_\infty, c^I, c_0^I, m^I$ and m_0^I represent the bounded , I-convergent, I-null, bounded I-convergent and bounded I-null sequence spaces respectively.

In this article we introduce the following classes of sequence spaces:

$$c^I(f) = \{(x_k) \in \omega : I - \lim f(|x_k|) = L \text{ for some L}\} \in I \tag{1.7}$$

$$c_0^I(f) = \{(x_k) \in \omega : I - \lim f(|x_k|) = 0\} \in I \tag{1.8}$$

$$l_\infty^I(f) = \{(x_k) \in \omega : \sup_k f(|x_k|) < \infty\} \in I \tag{1.9}$$

We also denote by

$$m^I(f) = c^I(f) \cap l_\infty(f) \tag{1.10}$$

and

$$m_0^I(f) = c_0^I(f) \cap l_\infty(f) \tag{1.11}$$

The following Lemmas will be used for establishing some results of this article:

Lemma 1.1. *Let E be a sequence space. If E is solid then E is monotone.*

Lemma 1.2. *Let $K \in £(I)$ and $M \subseteq N$. If $M \notin I$, then $M \cap N \notin I$.*

Lemma 1.3. *If $I \subset 2^N$ and $M \subseteq N$. If $M \notin I$, then $M \cap N \notin I$.*

2. Main results

Theorem 2.1. *For any modulus function f, the classes of sequences $c^I(f), c_0^I(f), m^I(f)$ and $m_0^I(f)$ are linear spaces.*

Proof. We shall prove the result for the space $c^I(f)$. The proof for the other spaces will follow similarly. Let $(x_k), (y_k) \in c^I(f)$ and let α, β be scalars. Then

$$I - \lim f(|x_k - L_1|) = 0, \text{ for some } L_1 \in c; \tag{2.1}$$

$$I - \lim f(|y_k - L_2|) = 0, \text{ for some } L_2 \in c; \tag{2.2}$$

That is for a given $\epsilon > 0$, we have

$$A_1 = \{k \in N : f(|x_k - L_1|) > \frac{\epsilon}{2}\} \in I, \tag{2.3}$$

$$A_2 = \{k \in N : f(|y_k - L_2|) > \frac{\epsilon}{2}\} \in I. \tag{2.4}$$

Since f is a modulus function, we have

$$f(|(\alpha x_k + \beta y_k) - (\alpha L_1 + \beta L_2)|) \leq f(|\alpha||x_k - L_1|) + f(|\beta||y_k - L_2|) \leq f(|x_k - L_1|) + f(|y_k - L_2|) \tag{2.5}$$

Now, by (2.3) and (2.4),

$$\{k \in N : f(|(\alpha x_k + \beta y_k) - (\alpha L_1 + \beta L_2)|) > \epsilon\} \subset A_1 \cup A_2. \tag{2.6}$$

Therefore

$$(\alpha x_k + \beta y_k) \in c^I(f). \tag{2.7}$$

Hence $c^I(f)$ is a linear space. □

Theorem 2.2. *A sequence $x = (x_k) \in m^I(f)$ I-converges if and only if for every $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that*

$$\{k \in \mathbb{N} : f(|x_k - x_{N_\epsilon}|) < \epsilon\} \in m^I(f) \tag{2.8}$$

Proof. Suppose that$L = I - \lim x$. Then

$$B_\epsilon = \{k \in \mathbb{N} : |x_k - L| < \frac{\epsilon}{2}\} \in m^I(f). \text{ For all} \epsilon > 0. \tag{2.9}$$

Fix an $N_\epsilon \in B_\epsilon$.Then we have

$$|x_{N_\epsilon} - x_k| \leq |x_{N_\epsilon} - L| + |L - x_k| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{2.10}$$

which holds for all $k \in B_\epsilon$. Hence

$$\{k \in \mathbb{N} : f(|x_k - x_{N_\epsilon}|) < \epsilon\} \in m^I(f). \tag{2.11}$$

Conversely, suppose that $\{k \in \mathbb{N} : f(|x_k - x_{N_\epsilon}|) < \epsilon\} \in m^I(f)$.
That is $\{k \in \mathbb{N} : (|x_k - x_{N_\epsilon}|) < \epsilon\} \in m^I(f)$ for all $\epsilon > 0$. Then the set

$$C_\epsilon = \{k \in \mathbb{N} : x_k \in [x_{N_\epsilon} - \epsilon, x_{N_\epsilon} + \epsilon]\} \in m^I(f) \text{for all} \epsilon > 0. \tag{2.12}$$

Let $J_\epsilon = [x_{N_\epsilon} - \epsilon, x_{N_\epsilon} + \epsilon]$. If we fix an $\epsilon > 0$ then we have $C_\epsilon \in m^I(f)$ as well as $C_{\frac{\epsilon}{2}} \in m^I(f)$. Hence $C_\epsilon \cap C_{\frac{\epsilon}{2}} \in m^I(f)$. This implies that

$$J_\epsilon \cap J_{\frac{\epsilon}{2}} \neq \phi \tag{2.13}$$

that is

$$\{k \in \mathbb{N} : x_k \in J\} \in m^I(f) \tag{2.14}$$

that is

$$diam J \leq diam J_\epsilon \tag{2.15}$$

where the diam of J denotes the length of interval J.
In this way, by induction we get the sequence of closed intervals

$$J_\epsilon = I_0 \supseteq I_1 \supseteq ..... \supseteq I_k \supseteq .... \tag{2.16}$$

with the property that $diam I_k \leq \frac{1}{2} diam I_{k-1}$ for (k=2,3,4,.....) and $\{k \in \mathbb{N} : x_k \in I_k\} \in m^I(f)$ for $(k = 1, 2, 3, 4, ...)$.
Then there exists a $\xi \in \cap I_k$ where $k \in \mathbb{N}$ such that $\xi = I - \lim x$. So that $f(\xi) = I - \lim f(x)$, that is $L = I - \lim f(x)$.

□

Theorem 2.3. *Let f and g be modulus functions that satisfy the $\triangle_2$-condition.If X is any of the spaces c^I, c_0^I, m^I and m_0^I etc, then the following assertions hold.*

(1) $X(g) \subseteq X(f.g)$,
(4) $X(f) \cap X(g) \subseteq X(f + g)$.

Proof. (1) Let $(x_k) \in c_0^I(g)$.Then

$$I - \lim_k g(|x_k|) = 0 \tag{2.17}$$

Let $\epsilon > 0$ and choose δ with $0 < \delta < 1$ such that $f(t) < \epsilon$ for $0 < t < \delta$. Write $y_k = g(|x_k|)$ and consider $\lim_k f(y_k) = \lim_k f(y_k)_{y_k<\delta} + \lim_k f(y_k)_{y_k>\delta}$. We have

$$\lim_k f(y_k) \leq f(2) \lim_k (y_k). \tag{2.18}$$

For $y_k > \delta$, we have $y_k < \frac{y_k}{\delta} < 1 + \frac{y_k}{\delta}$. Since f is non-decreasing, it follows that

$$f(y_k) < f(1 + \frac{y_k}{\delta}) < \frac{1}{2} f(2) + \frac{1}{2} f(\frac{2y_k}{\delta}). \tag{2.19}$$

Since f satisfies the $\triangle_2$-condition, we have

$$f(y_k) < \frac{1}{2} K \frac{y_k}{\delta} f(2) + \frac{1}{2} K \frac{y_k}{\delta} f(2) = K \frac{y_k}{\delta} f(2). \tag{2.20}$$

Hence

$$\lim_k f(y_k) \leq max(1,K)\delta^{-1} f(2) \lim_k (y_k). \tag{2.21}$$

From (2.17), (2.18) and (2.21), we have $(x_k) \in c_0^I(f.g)$.

Thus $c_0^I(g) \subseteq c_0^I(f.g)$. The other cases can be proved similarly.

(2) Let $(x_k) \in c_0^I(f) \cap c_0^I(\mathrm{g})$. Then

$$I - \lim_k f(|x_k|) = 0 \tag{2.22}$$

and

$$I - \lim_k g(|x_k|) = 0 \tag{2.23}$$

The rest of the proof follows from the following equality

$$\lim_k (f+g)(|x_k|) = \lim_k f(|x_k|) + \lim_k g(|x_k|). \tag{2.24}$$

□

Corollary 2.4. *$X \subseteq X(f)$ for $X = c^I, c_0^I, m^I$ and m_0^I.*

Theorem 2.5. *The spaces $c_0^I(f)$ and $m_0^I(f)$ are solid and monotone .*

Proof. We shall prove the result for $c_0^I(f)$. Let $x_k \in c_0^I(f)$.Then

$$I - \lim_k f(|x_k|) = 0 \tag{2.25}$$

Let (α_k) be a sequence of scalars with $|\alpha_k| \leq 1$ for all $k \in \mathbb{N}$. Then the result follows from (2.25) and the following inequality

$$f(|\alpha_k x_k|) \leq |\alpha_k| f(|x_k|) \leq f(|x_k|) \text{ for all } k \in \mathbb{N}. \tag{2.26}$$

That the space c_0^I(f) is monotone follows from the Lemma 1.1.
For $m_0^I(f)$ the result can be proved similarly. □

Theorem 2.6. *The spaces $c^I(f)$ and $m^I(f)$ are neither solid nor monotone in general.*

Proof. Here we give a counter example.
Let $I = I_\delta$ and $f(x) = x^2$ for all $x \in [0,\infty)$. Consider the K-step space $X_K(f)$ of X defined as follows.
Let $(x_k) \in X$ and let $(y_k) \in X_K$ be such that

$$(y_k) = \begin{cases} (x_k), \text{if } k \text{ is even}, \\ 0, otherwise. \end{cases} \tag{2.27}$$

Consider the sequence (x_k) defined by $(x_k) = 1$ for all $k \in \mathbb{N}$. Then $(x_k) \in c^I(f)$ but its K-stepspace preimage does not belong to $c^I(f)$. Thus $c^I(f)$ is not monotone. Hence $c^I(f)$ is not solid. □

Theorem 2.7. *The spaces $c^I(f)$ and $c_0^I(f)$ are sequence algebras.*

Proof. We prove that $c_0^I(f)$ is a sequence algebra. Let $(x_k), (y_k) \in c_0^I(f)$. Then

$$I - \lim f(|x_k|) = 0$$

and

$$I - \lim f(|y_k|) = 0. \tag{2.28}$$

Then we have

$$I - \lim f(|(x_k.y_k)|) = 0. \tag{2.29}$$

Thus $(x_k.y_k) \in c_0^I(f)$ is a sequence algebra. For the space $c^I(f)$, the result can be proved similarly. □

Theorem 2.8. *The spaces $c^I(f)$ and $c_0^I(f)$ are not convergence free in general.*

Proof. Here we give a counter example.
Let $I = I_f$ and $f(x) = x^3$ for all $x \in [0, \infty)$. Consider the sequence (x_k) and (y_k) defined by

$$x_k = \frac{1}{k} \quad and \quad y_k = k \text{ for all } k \in N$$

Then $(x_k) \in c^I(f)$ and $c_0^I(f)$, but $(y_k) \notin c^I(f)$ and $c_0^I(f)$.
Hence the spaces $c^I(f)$ and $c_0^I(f)$ are not convergence free.

□

Theorem 2.9. *If I is not maximal and $I \neq I_f$, then the spaces $c^I(f)$ and $c_0^I(f)$ are not symmetric.*

Proof. Let $A \in I$ be infinite and $f(x) = x$ for all x$\in [0, \infty)$.
If

$$x_k = \begin{cases} 1, \text{for } k \in A, \\ 0, otherwise. \end{cases}$$

Then by lemma 1.3 $x_k \in c_0^I(f) \subset c^I(f)$
let $K \subset \mathbb{N}$ be such that $K \notin I$ and $\mathbb{N} - K \notin I$. Let $\phi : K \to A$ and $\psi : \mathbb{N} - K \to \mathbb{N} - A$ be bijections, then the map $\pi : \mathbb{N} \to \mathbb{N}$ defined by

$$\pi(k) = \begin{cases} \phi(k), \text{for } k \in K, \\ \psi(k), otherwise. \end{cases}$$

is a permutation on $\mathbb{N}$, but $x_{\pi(k)} \notin c^I(f)$ and $x_{\pi(k)} \notin c_0^I(f)$.
Hence $c_0^I(f)$ and $c^I(f)$are not symmetric. □

Theorem 2.10. *Let f be a modulus function. Then $c_0^I(f) \subset c^I(f) \subset l_\infty^I(f)$ and the inclusions are proper.*

Proof. Let $x_k \in c^I(f)$. Then there exists $L \in C$ such that

$$I - \lim f(|x_k - L|) = 0. \tag{2.30}$$

We have $f(|x_k|) \leq \frac{1}{2} f(|x_k - L|) + f \frac{1}{2}(|L|)$.
Taking the supremum over k on both sides we get $x_k \in l_\infty(f)$.
The inclusion $c_0^I(f) \subset c^I(f)$ is obvious. □

Theorem 2.11. *The function $\hbar : m^I(f) \to \mathbb{R}$ is the Lipschitz function,where $m^I(f) = c^I(f) \cap l_\infty(f)$, and hence uniformly cotinuous.*

Proof. Let $x, y \in m^I(f)$, $x \neq y$. Then the sets

$$A_x = \{k \in \mathbb{N} : |x_k - \hbar(x)| \geq \|x - y\|\} \in I, \tag{2.31}$$

$$A_y = \{k \in \mathbb{N} : |y_k - \hbar(y)| \geq \|x - y\|\} \in I. \tag{2.32}$$

Thus the sets,

$$B_x = \{k \in \mathbb{N} : |x_k - \hbar(x)| < \|x - y\|\} \in m^I(f), \tag{2.33}$$

$$B_y = \{k \in \mathbb{N} : |y_k - \hbar(y)| < \|x - y\|\} \in m^I(f). \tag{2.34}$$

Hence also $B = B_x \cap B_y \in m^I(f)$, so that $B \neq \phi$. Now taking k in B,

$$|\hbar(x) - \hbar(y)| \leq |\hbar(x) - x_k| + |x_k - y_k| + |y_k - \hbar(y)| \leq 3\|x - y\|. \tag{2.35}$$

Thus $\hbar$ is a Lipschitz function. For $m_0^I(f)$ the result can be proved similarly. □

Theorem 2.12. *If $x, y \in m^I(f)$, then $(x.y) \in m^I(f)$ and $\hbar(xy) = \hbar(x)\hbar(y)$.*

Proof. For $\epsilon > 0$

$$B_x = \{k \in \mathbb{N} : |x_k - \hbar(x)| < \epsilon\} \in m^I(f), \tag{2.36}$$

$$B_y = \{k \in \mathbb{N} : |y_k - \hbar(y)| < \epsilon\} \in m^I(f). \tag{2.37}$$

Now,

$$|x_k y_k - \hbar(x)\hbar(y)| = |x_k y_k - x_k\hbar(y) + x_k\hbar(y) - \hbar(x)\hbar(y)|$$

$$\leq |x_k||y_k - \hbar(y)| + |\hbar(y)||x_k - \hbar(x)| \tag{2.38}$$

As $m^I(f) \subseteq l_\infty(f)$, there exists an $M \in \mathbb{R}$ such that $|x_k| < M$ and $|\hbar(y)| < M$. Using (2.38) we get

$$|x_k y_k - \hbar(x)\hbar(y)| \leq M\epsilon + M\epsilon = 2M\epsilon.$$

For all $k \in B_x \cap B_y \in m^I(f)$. Hence $(x.y) \in m^I(f)$ and $\hbar(xy) = \hbar(x)\hbar(y)$. For $m_0^I(f)$ the result can be proved similarly. □

Acknowledgement: The authors would like to record their gratitude to the reviewer for his careful reading and making some useful corrections which improved the presentation of the paper.

References

Connor, J. and J. Kline (1996). On statistical limit points and the consistency of statistical convergence. *J. Math. Anal. Appl.* **197**, 392–399.

Demirci, K. (2001). I-limit superior and limit inferior. *Mathematical Communications* (6), 165–172.

Dems, K. (2005). On I-Cauchy sequences. *Real Analysis Exchange* (30), 123–128).

Fast, H. (1951). Sur la convergence statistique. *Colloquium Mathematicum* (2), 241–244.

Fridy, J. A. (1985). On statistical convergence. *Analysis* (5), 301–313.

Fridy, J. A. (1993). Statistical limit points. *Proceedings of the American Mathematical Society* (11), 1187–1192.

Garling, D. J. H. (1966). On symmetric sequence spaces. *Proc. London. Math. Soc.* **16**, 85–106.

Garling, D. J. H. (1968). Symmetric bases of locally convex spaces. *Studia Math. Soc.* **30**, 163–181.

Gramsch, B. (n.d.). Die Klasse metrisher linearer Raume L(ϕ).

Gurdal, M. (2004). *Some types of convergence*. Doctoral Dissertation, S. Demirel Univ., Isparta.

Jones, O. T. and J. R. Retherford (1967). On similar bases in barrelled spaces. *Proc. Amer. Math. Soc.* (18), 677–680.

Köthe, G. (1970). *Topological vector spaces. (Vol. 1)*. Springer, Berlin.

Kamthan, P.K. and M. Gupta (1980). *Sequence spaces and series*. Marcel Dekker Inc, New York.

Kostyrko, P., T. Šalát and W. Wilczyński (2000). I-convergence. *Real Analysis Exchange* **26**(2), 669–686.

Maddox, I.J. (1969). Some properties of paranormed sequence spaces. *J. London Math. Soc* (1), 316–322.

Maddox, I.J. (1970). *Elements of functional analysis*. Cambridge university press, London - New York.

Maddox, I.J. (1986). Sequence spaces defined by a modulus. *Math. Camb. Phil. Soc.* (100), 161–166.

Nakano, H. (1953). Concave modulars. *J. Math Soc. Japan* (5), 29–49.

Ruckle, W.H. (1967). Symmetric coordinate spaces and symmetric bases. *Canad. J. Math.* (19), 828–838.

Ruckle, W.H. (1968). On perfect symmetric BK spaces. *Mathematische Annalen* **175**, 121–126. 10.1007/BF01418767.

Ruckle, W.H. (1973). FK-spaces in which the sequence of coordinate vectors is bounded. *Canad. J. Math.* **25**(5), 973–975.

Šalát, T. (1980). On statistically convergent sequences of real numbers. *Mathematica Slovaca* **30**(2), 139–150.

Šalát, T., B.C. Tripathy and M. Ziman (2004). On some properties of I-convergence. *Tatra Mt. Math. Publ.* (28), 279–286.

Singer, J. (1970). *Bases in Banach spaces. Vol. 1*. Springer, Berlin.

Wilansky, A. (1964). *Functional Analysis*. Blaisdell, New York.

Theory and Applications of
Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 31–41

Wrapper Approach for Feature Selections in RBF Network Classifier

Jasmina Novakovic[a,*]

[a]*Faculty of Public Administration, Megatrend University, Belgrade, Serbia.*

Abstract

In this paper we investigate the impact of wrapper approach on classification accuracy and performance of RBF network. Wrapper approach used six rule induction algorithms for evaluators on supervised learning algorithms RBF network and tested using eight real and three artificial benchmark data sets. Classification accuracy and performance of RBF network depends on evaluators. Our experimental results indicate that every rule induction algorithms in wrapper approach maintains or improves the accuracy of RBF network for more than half data sets. Evaluation of selecting features with wrappers approach is not so fast compare with filters approach.

Keywords: classification accuracy, feature selection, RBF network, rule induction algorithm, wrapper approach.
2000 MSC: 68T01, 97P20, 97R40.

1. Introduction

Feature selection has been a fertile field of research and development since 1970's in statistical pattern recognition, machine learning and data mining. It is a fundamental problem in many different areas, especially in forecasting, document classification, bioinformatics, object recognition or in modelling of complex technological processes. Datasets with thousands of features are not uncommon in such applications. For some problems all features may be important, but for some target concept only a small subset of features is usually relevant.

Feature selection reduces the dimensionality of feature space, removes redundant, irrelevant, or noisy data. It brings the immediate effects for application: speeding up a data mining algorithm, improving the data quality and thereof the performance of data mining, and increasing the comprehensibility of the mining results.

The main idea of feature selection is to choose a subset of input variables by eliminating features with little or no predictive information. We try to avoid selecting too many or too few features than necessary. If insuficient features are selected, the information content to keep the concept of the data is degraded. If too many features are selected, including redundant or irrelevant features, the classification accuracies may be lower due to the interference of irrelevant information. The goal of feature selection is to find those features

*Corresponding author
Email address: jnovakovic@megatrend.edu.rs (Jasmina Novakovic)

that may neither affect the target in any way (called irrelevant features) nor add anything new to the target (called redundant features) and exclude them.

Feature selection can be defined as a process that chooses a minimum subset of M features from the original set of N features, so that the feature space is optimally reduced according to a certain evaluation criterion. As the dimensionality of a domain expands, the number of feature N increases. Finding the best feature subset is usually intractable (Kohavi & John, 1997) and many problem related to feature selection have been shown to be NP-hard (Blum & Rivest, 1992).

Algorithms for feature selection may be divided into filters (Almuallim & Dietterich, 1991), (Kira & Rendell, 1992), wrappers (Kohavi & John, 1997) and embedded approaches. Filters methods evaluate quality of selected features, independently from the classification algorithm, while wrapper methods require application of a classifier (which should be trained on a given feature subset) to evaluate this quality. The weakness of the filter approach lies in that the selected feature subset may not lead to high performance in induction systems, such as the classification system. The wrapper approach combines data dimension reduction with induction algorithms, but high computational cost is a heavy burden. Embedded methods perform feature selection during learning of optimal parameters (for example, neural network weights between the input and the hidden layer).

Some classification algorithms have inherited ability to focus on relevant features and ignore irrelevant ones. Decision trees are primary example of a class of such algorithms (Breiman *et al.*, 1984), (Quinlan, 1993), but also multi-layer perceptron (MLP), neural networks with strong regularization of the input layer may exclude the irrelevant features in an automatic way (Duch *et al.*, 2001). Such methods may also benefit from independent feature selection. On the other hand, some algorithms have no provisions for feature selection. The k-nearest neighbor algorithm is one family of such methods that classify novel examples by retrieving the nearest training example, strongly relaying on feature selection methods to remove noisy features.

Our research interest includes wrapper approaches for feature selections. Wrapper methods require application of a classifier; in our experiment we use rule induction algorithms. We chose the radial basis function (RBF) network, because it cannot deal effectively with irrelevant features. A disadvantage of RBF network is that it gives every feature the same weight because all are treated equally in the distance computation. The main aim of this paper is to experimentally verify, on benchmark data sets, the impact on performance and classification accuracy on RBF network with wrapper approaches.

This paper is organized as follows. In the next section we briefly describe the wrapper approach. Section 3 gives a brief overview of RBF network as classification algorithm that we use in our experiment. Section 4 presents experimental evaluation. Final section contains discussion of the obtained results and some closing remarks.

2. Wrapper Approach for Feature Selections

Wrapper approach uses the method of classification itself to measure the importance of features set; hence the feature selected depends on the classifier model used. Wrapper methods generally result in better performance than filter methods because the feature selection process is optimized for the classification algorithm to be used. The fact is, wrapper methods are too expensive for large dimensional database in terms of computational complexity and time since each feature set considered must be evaluated with the classifier algorithm used (Dash & Liu, 1997),(Doraisamy *et al.*, 2008), (Saeys *et al.*, 2007).

Kohavi et al. were the first to introduce the wrapper approach to the mainstream data mining community. They successfully used the wrapper approach to search for an optimal feature subset customized to a specific induction learning algorithm and domain. The idea behind the wrapper approach is very simple and is shown

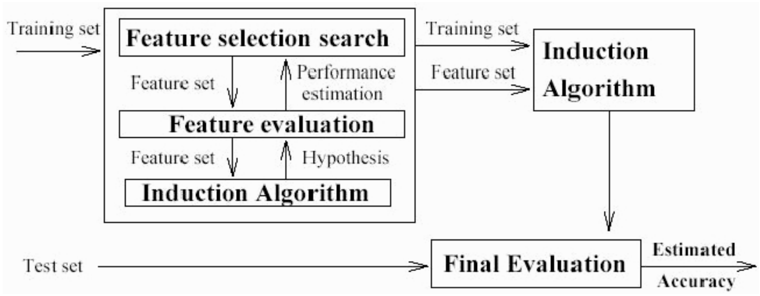


Figure 1: Wrapper approach to variable subset selection based on the incorporation of the learning algorithm (Almuallim & Dietterich, 1991).

in Figure 1. Some performance measure is used to evaluate the classifier built on each feature subset using a set aside distinct portion of the dataset, and the feature subset with the highest evaluation is used as the final set to build the final classifier on all the data instances in the training set. The resulting classifier can then be evaluated on an independent test set that is not used during the search process to assess the efficacy of the wrapper approach in selecting the feature subset.

After Kohavi et al. many researchers experimented with the wrapper approach in various contexts. The wrapper approach in selecting the features for a Nave-Bayes classifier was used by Langley and Sage (Langley & Sage, 1994). Pazzani (Pazzani, 1995) created super-features by combining the base features for a Nave-Bayes classifier by using the wrapper approach and demonstrated that it really was able to find the correct combination of features when they interacted. Singh and Provan (Singh & Provan, 1995) significantly improved the original K2 algorithm when they selected the features for Bayesian networks using the wrapper approach. Kohavi and John (Kohavi & John, 1997) again demonstrated the use of the wrapper with other search methods using probabilistic estimates for feature subset selection.

The wrapper approach has been used for many other problems except than feature selection. The wrapper approach for tweaking the parameters of C4.5 for maximal performance were applied by Kohavi and John (Kohavi & John, 1995). Skalak (Skalak, 1994) used the wrapper approach in an interesting fashion to select the training instances instead of the features in connection with nearest-neighbor classifiers.

3. RBF Network Classifier

RBF network as supervised learning algorithms is adopted here to build models. This section gives a brief overview of this algorithm. This network emerged as variant of artificial neural network in late 80's. RBF network is a popular artificial neural network architecture that has found wide applications in diverse fields of engineering. It is used in function approximation, time series prediction, and control.

RBF network is an artificial neural network that uses radial basis functions as activation functions. This network is a special class of neural networks in which the activation of a hidden neuron is determined by the distance between the input vector and a prototype vector. Prototype vectors refer to centers of clusters obtained during RBF training. Usually, in RBF network three kinds of distance metrics can be used: Euclidean, Manhattan, and Mahalanobis distances.

In Figure 2 is presented architecture of RBF network. An input vector x is used as input to all radial basis functions, each with different parameters. The output of the network is a linear combination of the outputs from radial basis functions.

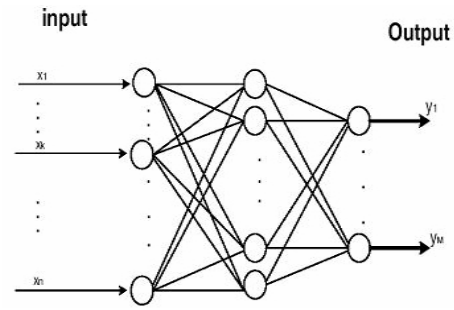


Figure 2. Architecture of RBF network.

RBF network has three layers: an input layer, a hidden layer with a non-linear RBF activation function and a linear output layer. The output,$\phi : \mathbb{R}^n \to \mathbb{R}$, of the network is thus

$$\phi(\mathbf{x}) = \sum_{i=1}^{N} a_i \rho(||\mathbf{x} - \mathbf{c}_i||) \tag{3.1}$$

where N is the number of neurons in the hidden layer, $\mathbf{c}_i$ is the center vector for neuron i, and a_i are the weights of the linear output neuron. In RBF network, in the basic form, all inputs are connected to each hidden neuron. Typically, the norm is taken to be the Euclidean distance and the basis function is taken to be Gaussian

$$\rho(||\mathbf{x} - \mathbf{c}_i||) = \exp\left[-\beta\, ||\mathbf{x} - \mathbf{c}_i||^2\right]. \tag{3.2}$$

In RBF network the Gaussian basis functions are local in the sense that

$$\lim_{||\mathbf{x}|| \to \infty} \rho(||\mathbf{x} - \mathbf{c}_i||) = 0 \tag{3.3}$$

i.e. changing parameters of one neuron has only a small effect for input values that are far away from the center of that neuron.

RBF network is universal approximators on a compact subset of $\mathbb{R}^n$, which means that a RBF network with enough hidden neurons can approximate any continuous function with arbitrary precision. The weights a_i, c_i, and β are determined in a manner that optimizes the fit between ϕ and the data.

In training phase, there are three types of parameters that need to be chosen to adapt RBF network for a particular task: the center vectors c_i, the output weights w_i, and the RBF width parameters β_i. In the sequential training of the weights are updated at each time step as data streams in. In RBF network for some tasks it makes sense to define an objective function and select the parameter values that minimize its value. The least squares function is the most common objective function

$$K(\mathbf{w}) \stackrel{def}{=} \sum_{t=1}^{\infty} K_t(\mathbf{w}) \tag{3.4}$$

where

$$K_t(\mathbf{w}) \stackrel{def}{=} [\mathbf{y}(t) - \phi(\mathbf{x}(t), \mathbf{w})]^2 \tag{3.5}$$

The dependence on the weights is explicitly included. Minimization of the least squares objective function by optimal choice of weights optimizes accuracy of fit.

In some cases, multiple objectives, such as smoothness as well as accuracy, must be optimized. If it is, it is useful to optimize a regularized objective function such as

$$H(\mathbf{w}) \stackrel{def}{=} K(\mathbf{w}) + \lambda S(\mathbf{w}) \stackrel{def}{=} \sum_{t=1}^{\infty} H_t(\mathbf{w}) \tag{3.6}$$

where

$$S(\mathbf{w}) \stackrel{def}{=} \sum_{t=1}^{\infty} S_t(\mathbf{w}) \tag{3.7}$$

and

$$H_t(\mathbf{w}) \stackrel{def}{=} K_t(\mathbf{w}) + \lambda S_t(\mathbf{w}) \tag{3.8}$$

where optimization of S maximizes smoothness and λ is known as a regularization parameter.

Three training schemes for RBF network are: one-stage training, two-stage training and three-stage training. In one-stage training procedure, only the weights connecting the hidden layer and the output layer are adjusted through some kind of supervised methods, e.g., minimizing the squared difference between the RBF network's output and the target output. The centers of hidden neurons are subsampled from the set of input vectors (or all data points are used as centers) and, typically, all scaling parameters of hidden neurons are fixed at a predefined real value. Usually, two-stage training is used for constructing RBF network. At the first stage, the hidden layer is constructed by selecting the center and the width for each hidden neuron using various clustering algorithms. At the second stage, the weights between hidden neurons and output neurons are determined, for example by using the linear least square method. In a three-stage training procedure RBF network is adjusted through a further optimization after being trained using a two-stage learning scheme.

In function approximation and classification tasks, generalization and the learning abilities are important issues. If RBF network has as many hidden neurons as the training patterns, RBF network can attain no errors for a given training data set. In that case, the size of the network may be too large when tackling large data sets and the generalization ability of such a large RBF network may be poor. Smaller RBF networks may have better generalization ability; but, too small RBF network will perform poorly on both training and test data sets. It is recommendable to determine a training method which takes the learning ability and the generalization ability into consideration at the same time.

The problem is how to optimally determine the key parameters of RBF classifier. Determination the so-called 'sufficient number of hidden units' is required prior knowledge. Though the number of the training patterns is known in advance, it is not the only element which affects the number of hidden units. The data distribution is another element affecting the architecture of RBF network.

The performance of RBF network depends heavily on the network structure especially the input and hidden neurons. Incorrect input neurons or poorly located RBF centers will induce bias to the fitted network model. The main difficulty in operating RBF network concerns the optimization of the hidden layer. When a dynamic architecture is preferred to a predefined one, the optimization method often consists in a gradient descent algorithm which can get trapped in local minima. Furthermore, the activation function has an influence on the final state of the optimization process.

4. Experimental Results

Natural and artificial domains were used for evaluating wrapper approach with RBF network, taken from the UCI repository of machine learning databases. These domains were chosen because of: (a) their predominance in the literature, and (b) the prevalence of nominal features, thus reducing the need to discretize feature values. Table 1 is shown the characteristics of these domains.

Table 1. Domain characteristics.

Domain	Instances	Features	% Missing	Average # Feature Vals	Max/Min # Feature Vals	Default Class Vals	Accuracy
mu	8124	22	1.3	5.3	12/1	2	51.8
vote	435	16	5.3	2.0	2/2	2	61.4
cr	690	15	0.6	4.4	14/2	2	55.5
ly	148	18	0.0	2.9	8/2	4	54.7
pt	339	17	3.7	2.2	3/2	22	24.8
bc	286	9	0.3	4.6	11/2	2	70.3
au	226	69	2.0	2.2	6/2	24	25.2
sb	683	35	9.5	2.8	7/2	19	13.5
M1	432	6	0.0	2.8	4/2	2	50.0
M2	432	6	0.0	2.8	4/2	2	67.1
M3	432	6	0.0	2.8	4/2	2	52.8

On Table 1 data sets above the horizontal line are natural domains, those below are artificial. The default accuracy is the accuracy of always predicting the majority class on the whole data set. The % Missing column shows what percentage of the data set's entries (number of features X number of instances) have missing values. Average # Feature Vals and Max/Min # Feature Vals are calculated from the nominal features present in the data sets. The following is a brief description of the data sets.

Mushroom (mu) is a large data set containing 8124 instances which includes descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family. The task is to distinguish edible from poisonous mushrooms on the basis of 22 nominal attributes describing characteristics of the mushrooms such as the shape of the cap, odour, and gill spacing.

Vote data set includes votes for each of the U.S. House of Representatives Congressmen on the 16 key issues such as education spending and immigration. In the original data, there are lists with nine different types of votes. There are 435 (267 democrats, 168 republicans) instances and all features are binary.

Australian credit screening (cr) data set concerns credit card applications. The task is to distinguish credit-worthy from non credit-worthy customers. Data set characteristics is multivariate; feature characteristics are categorical, integer and real. Number of instances is 690, number of features is 15, and there are missing values.

Lymphography (ly) is a small medical data set containing 148 instances with 18 nominal features. The task is to distinguish healthy patients from those with metastases or malignant lymphoma. The values for class attribute are normal find, metastases, malign lymph and fibrosis.

Primary Tumor (pt) data set involves predicting the location of a tumor in the body of a patient on the basis of 17 nominal features. There are 339 instances. There are 22 values for class attribute corresponding to body locations: lung, head & neck, esophasus, thyroid, stomach, duoden & sm.int, colon, rectum, anus, salivary glands, pancreas, gallblader, liver, kidney, bladder, testis, prostate, ovary, corpus uteri, cervix uteri, vagina and breast.

Breast Cancer (bc) data set involves predicting whether cancer will recur in patients. There are 9 nominal attributes describing characteristics such as tumor size and location with 286 examples.

Audiology (au) data set containing 226 instances described by 69 nominal features. The task is to diagnose ear dysfunctions. There are 24 values for classe attribure.

Soybean-large (sb) data set containing 683 instances described by 35 nominal features. The task is to diagnose diseases in soybean plants. Features measure properties of leaves and various plant abnormalities. There are 19 values for class attribute (diseases).

Monk's problems domains are the same for all Monk's problems with 432 instances. There are three Monk's problems. Monk's domains contain instances of robots described by six nominal features:

$Head - shape \in \{round, square, octagon\}$

$Body - shape \in \{round, square, octagon\}$

$Is - smiling \in \{yes, no\}$

$Holding \in \{sword, balloon, flag\}$

$Jacket - colour \in \{red, yellow, green, blue\}$

$Has - tie \in \{yes, no\}$

The concept of Monk1 (M1) is: (head-shape = body-shape) or (jacket-colour = red)

This problem is difficult due to the interaction between the first two features. But, only one value of the jacket-colour feature is useful.

The concept of Monk2 (M2) is: Exactly two of the features have their first value.

This is a hard problem due to the pairwise feature interactions and the fact that only one value of each feature is useful. Note that all six features are relevant to the concept.

The concept of Monk3 (M3) is:

(jacket-colour = green and holding = sword) or

(jacket-colour ≠ blue and body-shape ≠ octagon)

In M3 5% class noise added to the training set. This is the only Monk's problem that is with noise. It is possible to achieve approximately 97% accuracy using only the (jacket-colour ≠ blue and body-shape ≠ octagon) disjunct.

The typical goal of supervised learning algorithms is to maximize classification accuracy on unseen test set, so we have adopted this as our goal in guiding the feature subset selection.

In our experiment a normalized Gaussian radial basis function network is used. It uses the k-means clustering algorithm to provide the basis functions and learns either a logistic regression (discrete class problems) or linear regression (numeric class problems) on top of that. In RBF network symmetric multivariate Gaussians are fit to the data from each cluster. If the class is nominal it uses the given number of clusters per class. RBF network standardizes all numeric attributes to zero mean and unit variance. We set RBF network in following way:

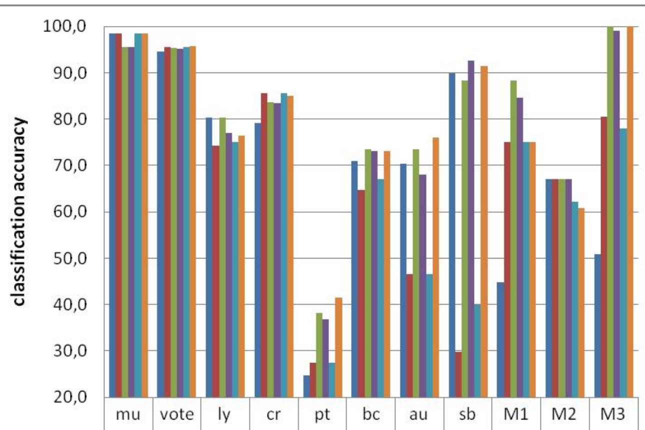
- The random seed to pass on to k-means is set on value 1.
- Maximum number of iterations for the logistic regression to perform is set on value -1.
- The minimum standard deviation for the clusters is set on value 0.1.

- The number of clusters for k-means to generate is set on value 2.
- The ridge value for the logistic or linear regression is set on value 1.0E-8.

Wrapper approach evaluates attribute sets by using a learning scheme. Cross validation is used to estimate the accuracy of the learning scheme for a set of attributes. Five rule induction algorithms are used as classifiers for estimating the accuracy of subsets. These algorithms are: ConjuctiveRule (CR), DecisionTable (DT), JRip, OneR and PART.

CR algorithm implements a single conjunctive rule learner that can predict for numeric and nominal class labels. A rule consists of antecedents "AND"ed together and the consequent (class value) for the classification/regression. In this case, the consequent is the distribution of the available classes (or mean for a numeric value) in the dataset. DT algorithm builds a decision rule using a simple decision table majority classifier. It summarizes the dataset with a 'decision table' which contains the same number of attributes as the original dataset. JRip implements a propositional rule learner - Repeated Incremental Pruning to Produce Error Reduction (Ripper). Ripper builds a ruleset by repeatedly adding rules to an empty ruleset until all positive examples are covered. Rules are formed by greedily adding conditions to the antecedent of a rule (starting with empty antecendent) until no negative examples are covered. After a ruleset is constructed, an optimization postpass massages the ruleset so as to reduce its size and improve its fit to the training data. OneR is a simple algorithm, it builds one rule for each attribute in the training data and then selects the rule with the smallest error rate as its 'one rule'. PART is a separate-and-conquer rule learner, producing sets of rules called 'decision lists' which are ordered set of rules. This algorithm builds a partial C4.5 decision tree in each iteration and makes the "best" leaf into a rule. PART is a combination of C4.5 and Ripper rule learning.

In our experiment, rule induction algorithms are used following settings: number of xval folds to use when estimating subset accuracy is five; seed to use for randomly generating xval splits is set on 1; threshold is set on 0.01.



	mu	vote	ly	cr	pt	bc	au	sb	M1	M2	M3
full	98,5	94,5	80,4	79,1	24,8	71,0	70,3	89,9	44,9	67,1	50,9
CR	98,5	95,6	74,3	85,5	27,4	64,7	46,5	29,9	75,0	67,1	80,6
DT	95,6	95,4	80,4	83,6	38,1	73,4	73,4	88,4	88,4	67,1	100,0
JRip	95,6	95,2	77,0	83,5	36,9	73,1	68,1	92,7	84,7	67,1	99,1
OneR	98,5	95,6	75,0	85,5	27,4	67,1	46,5	40,0	75,0	62,1	77,9
PART	98,5	95,8	76,4	85,1	41,6	73,1	76,1	91,4	75,0	60,9	100,0

Figure 3. Classification acccuracy of RBF network with wrapper approach.

We chose these values of parameters for algorithms, on the basis of those parameters that generated the best results in most cases. But, in some cases, we can get better results with different values of these parameters. If we change these parameters, classification accuracy are changed.

The results of testing wrapper approach on eight natural domains and three artificial domains are described in this section. The purpose of the experiments described in this section is to empirically test the claim that wrapper approach can improve the accuracy of RBF network. The performance of learning algorithm with and without feature selection is taken as an indication of wrapper approach success in selecting useful features, because the relevant features are often not known in advance for natural domains. Classification accuracy was estimated using ten-fold cross validation on each data set.

Table 2. Features selections by wrapper approach.

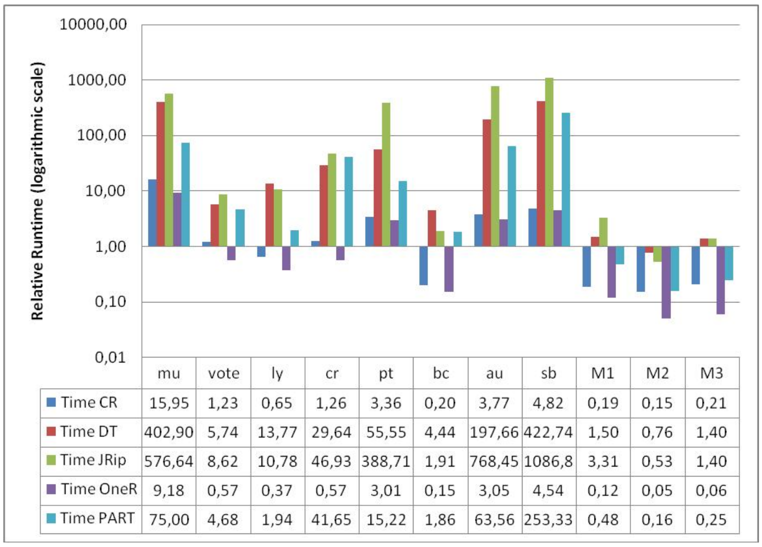
Data set	*CR*	*DT*	*JRip*	*OneR*	*PART*
mu	5	2,3,5,12,20	2,3,5,12,20	5	3,5,8,12,20
vote	4	4	3,4,7,8,16	4	3,4,7,9,11
ly	13	7,12,13,17	1,7,9,13,17	13	7,13,17
cr	9	4,9,10,12	3,9,10,11,14,15	9	2,3,4,5,7,9,10,11,13
pt	15	2,3,5,7,10,13,15	1,2,7,9,10,13,14,15	15	2,3,5,6,8,9,13,15,16
bc	-	5,6	5,6	5	1,5,6
au	1	1,11,15,17,66	1,2,6,7,8,10,11,13,14,15,17,26,27, 39,40,54,55,57,58,63,65,66	1	1,2,6,7,11,15,51,65,66
sb	13	3,12,14,15,16,17, 18,22,28,29	1,3,9,15,17,18,19,22,23,24,26,28, 29,30,31,32,35	29	1,3,9,10,13,14,15,17,18,20,22, 23,26,28,29,35
M1	6	2,3,6	2,3,6,7	6	3
M2	-	-	-	-	-
M3	3	3,5,6	2,3,5,6,7	3	3,5,6

Wrapper approach with CR maintains or improves the accuracy of RBF network for seven data sets and degrades its accuracy for four. For DT wrapper approach maintains or improves accuracy for none data sets and degrades for two. For JRip wrapper approach maintains or improves accuracy for eight data sets and degrades for three. For OneR wrapper approach maintains or improves accuracy for six data sets and degrades for five. For PART wrapper approach maintains or improves accuracy for nine data sets and degrades for two. The best results we have with DT and PART. Wrapper approach is able to improve the accuracy of RBF network dramatically on M1 and M3.

Table 2 shows features selected by wrapper approach for each data set. Rule induction algorithms CR and OneR reduced the number of features the most, for each data set they selected one feature. Only for two natural and two artificial domains some of the implemented rule induction algorithms have not reduced the number of features by more than half.

The experiments presented in this article show that wrapper approach's ability to select useful features does carry over from artificial to natural domains. Analysis of the results on the natural domains has revealed a weakness with RBF network. RBF network with huge data set size, many instances and number of classes, required the long time for processing, with or without feature selection.

The computational efficiency of different wrapper approach is measured in seconds of running time. The experiment was run on AMD Phenom (tm) 9650 Quard-Core Pocessor 2.31 GHz with 4GB RAM. Each configuration is run 10 times, and the average running time is used as a result. The Figure 4 shows the relative runtime (logarithmic scale). As we can see, RBF network had difficulty on domains with the highest number of classes (especially pt, au and sb). For four data sets, with huge data set size, many instances and number of classes (au, sb, pt and mu), very long time was nessesery to calculate the results, with or without feature selection.



	mu	vote	ly	cr	pt	bc	au	sb	M1	M2	M3
Time CR	15,95	1,23	0,65	1,26	3,36	0,20	3,77	4,82	0,19	0,15	0,21
Time DT	402,90	5,74	13,77	29,64	55,55	4,44	197,66	422,74	1,50	0,76	1,40
Time JRip	576,64	8,62	10,78	46,93	388,71	1,91	768,45	1086,8	3,31	0,53	1,40
Time OneR	9,18	0,57	0,37	0,57	3,01	0,15	3,05	4,54	0,12	0,05	0,06
Time PART	75,00	4,68	1,94	41,65	15,22	1,86	63,56	253,33	0,48	0,16	0,25

Figure 4. Time taken to build model (seconds).

5. Conclusions

A disadvantage of RBF networks is that they cannot deal effectively with irrelevant features. Wrapper approach may filter features leading to reduce dimensionality of the feature space. Wrapper approach with different rule induction algorithms have been used for feature selection, evaluated and compared using RBF network as classifiers on eight real and three artificial benchmark data.

Every rule induction algorithms in wrapper approach maintains or improves the accuracy of RBF network for more than half data sets. The best results we have with DT and PART, they maintains or improves the accuracy of RBF network for nine data sets, and only degrades for two. Wrapper approach is able to improve the accuracy of RBF network dramatically on M1 and M3. But, evaluation of selecting features is not fast, compare to filter approach.

There are many questions and issues that remain to be addressed and that we intend to investigate in future work. Some improvements of the selecting methods presented here are possible. The algorithms and data sets will be selected according to precise criteria: classify algorithms and several data sets, either real or artificial, with nominal, binary and continuous features. These conclusions and recommendations will be tested on larger data sets using various classification algorithms in the near future.

References

Almuallim, Hussein and Thomas G. Dietterich (1991). Learning with many irrelevant features. In: *In Proceedings of the Ninth National Conference on Artificial Intelligence*. AAAI Press. pp. 547–552.

Blum, A. L. and R. L. Rivest (1992). Training a 3-node neural networks is np-complete. *Neural Networks* **5**, 117–127.

Breiman, L., J. Friedman, R. Olshen and C. Stone (1984). *Classification and Regression Trees*. Wadsworth and Brooks. Monterey, CA.

Dash, M. and H. Liu (1997). Feature selection for classification. *Intelligent Data Analysis* **1**, 131–156.

Doraisamy, Shyamala, Shahram Golzari, Noris Mohd. Norowi, Md Nasir Sulaiman and Nur Izura Udzir (2008). A study on feature selection and classification techniques for automatic genre classification of traditional malay music.. In: *ISMIR* (Juan Pablo Bello, Elaine Chew and Douglas Turnbull, Eds.). pp. 331–336.

Duch, W., R. Adamczak and K. Grabczewski (2001). A new methodology of extraction, optimization and application of crisp and fuzzy logical rules. *IEEE Transactions on Neural Networks* **12**(2), 277 – 306.

Kira, Kenji and Larry A. Rendell (1992). The Feature Selection Problem: Traditional Methods and a New Algorithm.. In: *AAAI*. AAAI Press and MIT Press. Cambridge, MA, USA. pp. 129–134.

Kohavi, R. and G. H. John (1997). Wrappers for feature subset selection. *Artificial Intelligence* **97**, 273–324.

Kohavi, Ron and George H. John (1995). Automatic parameter selection by minimizing estimated error. In: *In Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann. pp. 304–312.

Langley, Pat and Stephanie Sage (1994). Induction of selective bayesian classifiers. In: *Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann. pp. 399–406.

Pazzani, M. (1995). Searching for attribute dependencies in bayesian classifiers. In: *Proceedings of the Fifth International Workshop on Artificial Intelligence and Statistics*. pp. 112–128.

Quinlan, J. Ross (1993). *C4.5: Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning)*. 1 ed.. Morgan Kaufmann.

Saeys, Yvan, Iñaki Inza and Pedro Larrañaga (2007). A review of feature selection techniques in bioinformatics.. *Bioinformatics (Oxford, England)* **23**(19), 2507–2517.

Singh, Moninder and Gregory M. Provan (1995). A comparison of induction algorithms for selective and non-selective bayesian classifiers. In: *Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann. pp. 497–505.

Skalak, David B. (1994). Prototype and feature selection by sampling and random mutation hill climbing algorithms. In: *Machine Learning: Proceedings of the Eleventh International Conference*. Morgan Kaufmann. pp. 293–301.

Theory and Applications of
Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 42–55

Symmetry, Hierarchy, Analogy Versus Embedded Systems

Tudor-Răzvan Niculiu[a,*], Anton Manolescu[a]

[a]*Politehnica University of Bucureşti, Faculty of Electronics, Telecommunications and Information Technology Faculty, Splaiul Independenţei 313, 060042 Bucureşti, Romania.*

Abstract

Intelligence = Consciousness × Adaptability × Intention and *Faith* = Intuition × Inspiration × Imagination, are the complementary parts of the human mind; the link between is *Conscience* = Consciousness × Inspiration. Simulation is the relation between function and structure. Conscience simulation demands transcending from computability to simulability, by an intensive effort on extensive research to integrate essential mathematical and physical knowledge guided by philosophical goals. A way to begin is hierarchical simulation. Coexistent interdependent hierarchical types structure the universe of models for complex systems. They belong to different types of hierarchy. Symmetry between simulation and theory can model the conscience (Fig. 1).

Keywords: Faith, Conscience, and Intelligence are ☯ in our Life.

1. Hierarchic approach

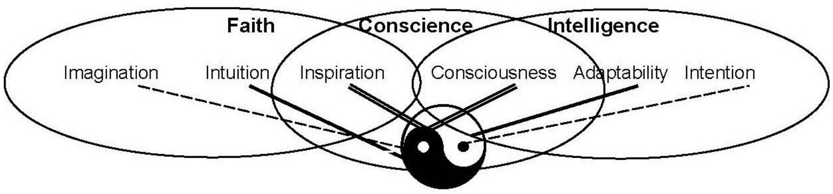


Figure 1. A possible model of the human mind.

The algorithmic approach is equivalent to the formal one. Algorithms, designs, artificial systems can be computer simulated so they represent computability, top-down (construction, design, plan) or bottom-up (understanding, verification, learning). Knowledge and construction hierarchies can cooperate to integrate design and verification into simulation: structural object-oriented concepts handle data and operations symbolically. Hierarchy types open the way to simulate intelligence as intentioned adaptable consciousness by extending the present limits of computability.

*Corresponding author
Email addresses: `tudor-razvan@ieee.org` (Tudor-Răzvan Niculiu), `tony@golana.pub.ro` (Anton Manolescu)

We enrich the template concept to structures and create a theoretical kernel, for self-organizing systems, based on a hierarchical formalism. This permits theoretical development as well as efficient application to different cosimulation types of reconfigurable systems. Coexistent interdependent hierarchies structure the universe of models for complex systems, e.g., hardware-software. They belong to different hierarchy types defined by modules, symbols, and classes. Hierarchy is the syntax of abstraction-simulation and knowledge abstraction:

- *Class* hierarchy: ↑concepts-virtual framework to represent any kind of hierarchy, based on form-contents, modularity, inheritance, polymorphism.
- *Symbol* hierarchy: ↑mathematics-stepwise formalism for all types, including hierarchy types.
- *Module* hierarchy: ↑managing-stepwise managing of all (other hierarchy) types on different levels by recursive autonomous block decomposition, following the principle *Divide et Impera et Intellige*.
- *Construction* hierarchy: ↑simulation-design/ verification/ optimization framework of autonomous levels for different abstraction grades.
- *Knowledge* hierarchy: ↑theories-reflexive abstraction, aiming that each level has knowledge of its inferior levels, including itself; this kind of abstraction enables consciousness.

Knowledge and construction have correspondent hierarchy types: their syntax relies on classes, their meaning on symbols and their use/ action on modules. Together with the construction hierarchy type, the knowledge hierarchy type hierarchically realizes the consciousness.

The hierarchy types can be formalized in the theory of categories (Ageron, 2001). Hierarchical types are objects of equivalent categories-functorial isomorphic-that formally represent hierarchy types. The consciousness hierarchy type communicates to the other hierarchy types by countervariant functors, while the others intercommunicate by covariant ones. Hierarchy consists of a net that can represent any type of mathematical structure-algebraic, topological, order-the first step to model the Conscience.

Constructive type theory permits formal simulation by generating an object satisfying the specification. Applying similar abstractions to hardware and software, representations and operations based on object-orientation, symbolization and structural abstraction can be extended from soft to hard (Keutzer *et al.*, 2000). A generic type (form of polymorphism) assures the ability to parameterize with types a hard/ soft element. Recurrence is confined to discrete worlds, while abstraction is not. This suggests searching for understanding mathematical structures that order algebra into topology.

The alternative ways followed to extension of the computability concept to concentrate respectively on the mental world of the good managed by engineering, the physical world of the truth researched by science, and Plato's world of the beautiful abstractions discovered by arts, as mathematics (Bacalu, 2004). We follow the mathematical paradigm of intelligent simulation by functionally modeling the self-aware adaptable behavior to simulate intelligence.

$$\begin{gathered}\text{Simulation} \in \text{Behavior} \times \text{Structure} \Leftarrow \text{Knowledge}\\ \text{Knowledge} \Leftarrow \text{Intelligence} :: \text{information}\\ \text{Imagination} \Leftarrow |\text{Intuition} - \text{Consciousness}|\\ \text{Intention} \Leftarrow \left|\text{Inspiration} - \text{Adaptability}\right|\end{gathered}$$

$$\text{Adaptability} \Leftarrow \text{simplifying\_Abstraction (Imagination)}$$
$$\text{Consciousness} \Leftarrow \text{reflexive\_Abstraction (Intention)}$$

2. Symmetry. Hierarchy. Analogy

The integration between discrete and analog is needed for a softer adaptability and for consciousness simulation as analog reaction. Recurrence of structures and operations enables approximate self-knowledge, with more precision on the higher knowledge levels.

We oversimplify to move towards intelligent simulation: First, we neglect the essential but hard to understand intuition and inspiration, formalizing reflexive abstraction by knowledge hierarchy type and simplifying abstraction by construction hierarchy type:

$$\mathit{Consciousness} = \mathit{knowledge} \circ \mathit{simulation}\ (\mathit{Consciousness})$$

This fixed-point relation suggests modeling the consciousness by associating to any hierarchical level of the construction process a knowledge level. We need a metric space where knowledge construction is a contraction, i.e., elements implied in the construction get closer to one another in the formal understanding of the formal construct. General functional relations between the essential parts of the faith-assisted intelligence imply:

$$\mathit{Consciousness} = \mathit{knowledge}\ (\mathit{intention}\ (\mathit{Inspiration},\ \mathit{simulation}\ (\mathit{imagination}\ (\mathit{Intuition},\ \mathit{Consciousness}))))$$

A continuous model for superior hierarchy levels offers a better model for consciousness part of intelligence. Representation for design and verification is common, algebraic structures on which the different hierarchy types are based on are extended to topologic structures; the different simulation entities are symbolic (Hennig & Sommer, 2001).

Reality is beyond Nature ($\mathbb{N} \subset \mathbb{R}$)

The hierarchical principle is applied to the object of knowledge as to the knowledge structure itself: it mediates the action of a paradigm on an environment. An intelligent system is capable of reflexive abstraction, controlled by problem specification and solving strategies, derived from higher knowledge levels of approach principles that are structured by an even higher level containing abstract types.

Simulability is computability to the power of continuum

Applying the hierarchic principle both at environment and simulation level ensures flexibility of the framework, by defining it precisely only in the neighborhood of solved cases.

For representation, the hierarchic principle offers the advantage of open modeling, which enables reconfigurable realization. Formalizing hierarchical descriptions in continuous spaces we come closer to self-control, -organization, -awareness, i.e., (intention, adaptability, consciousness), hence to intelligence.

There are enough positive signs for this from analog electronics, control systems, mechatronics (Traub, 1999). Real progress towards real simulation instead of natural computation calls for unrestricted mathematics, integrated physics, and thinking by analogies. Knowledge is based on morphisms between the real system and the simulator. An intelligent simulator learns generating and validating models.

Mathematics contains appropriate structures for self-referent models: the richest domain is functional analysis: it integrates algebra, topology and order, e.g., contractions and fixed points in metric spaces, reflexive normed vector spaces, self-adjoint operators of *Hilbert* spaces, inductive limits of locally convex spaces, reversible operators in *Banach* algebra.

God's ways are uncountable
His plans may be hierarchic

E.g., let U be a universe that is structured by different hierarchies. U is a category, e.g., containing *Hilbert* spaces with almost everywhere-continuous functions as morphisms, enabling different ways to simulate self-organizing, -control, and -awareness. A hierarchic formal system is defined by: hierarchic universe, functional objects (global functions, level structures, simplifying and knowledge abstractions), initial functions, and transformation rules. We consider the self-adjoint operators as objects on the higher levels of the knowledge hierarchy.

These levels strive then for self-knowledge, whose degree rises as the knowledge abstraction, in the context of the inferior level knowledge, and of superior level qualitative knowledge. Functorial morphisms on the functors of different hierarchy types solve the correspondence problem, i.e., the association of a knowledge hierarchy to the simulation one. Intention results by human-system dialog, completing the simulation of the intelligence.

Further than modeling consciousness to simulate intelligence is the search to comprehend inspiration. First we use *Lebesgue* measure on differentiable manifolds of (non)separable spaces. Even mathematics has to develop more philosophy-oriented to approach intuition. Evolution needs separation of faith and intelligence, understanding and using consciously more of faith's domain, and integrating them to human wisdom to be divided further to get more human. Metaphorically phrased: *Hierarchy* is a functional/ structural concept that fulfils mathematically/ physically the concept of abstraction. Symmetry is on the basic hierarchy level the complete correspondence among syntax, semantics, and pragmatics, i.e., among class, symbol, and structure hierarchies; Analogy is the correspondence among hierarchic levels.

3. Function. Structure. Architecture

To begin was the word. Words enable us to express ourselves, to be humans among humans. The expression is complex, so it has to be hierarchical in order to be comprehensible. Words are sequences of letters, sentences are sequences of words, and texts are sequences of sentences. Phrases, paragraphs, (sub)chapters, volumes, etc can enrich the levels of the textual hierarchy.

The basic hierarchical type is tree-like, to optimally represent the generic strategy of *Divide et Impera et Intellige*, or even graph-like, in order not to constrain the links between levels. Class, concept, term are aspects (syntax, semantics, pragmatics) of the expression.

Class is a primitive notion. Set is a class that belongs to another class. The set operations are paradigmatic: serial ($\cup$), parallel ($\times$), and hierarchic ($\wp$ -set of all parts). The possible expressions form a language. Syntax, semantics, and pragmatics define any language; the rules of each of the former defining components refer, respectively, to correct construction, interpretation, and application.

The syntax is determined grammatically: grammars are of different types that can build a hierarchy that corresponds to the reciprocal inclusion of the defined languages. Grammar is a language that refers to the language that grammar defines, i.e., is beyond the defined language-a metalanguage. This is another hierarchy type than modularization (of a text) of inclusion (of the languages) due to the stronger rules of the defining grammar. Its definition is based upon the principle that each level is a metalevel of its inferior ones. This idea is realized symbolically.

Further, the language can be symbolic, and symbols can symbolize other symbols, what reveals another hierarchy type. We classified, we symbolized, we divided into modules, and we reflected an inferior level (language) on a higher one (grammar). Grammar is a language, so it has a grammar, which, if isomorphic to the initial grammar or to the language itself, would mean that we obtained a reflexive language, i.e., capable to express itself.

Classes, symbols, and modules permit the construction of a system that structurally implements a function expressed in a language, i.e., behavior. The same way, with classes, symbols, and modules, the behavior of a structurally described system can be determined.

Another hierarchy type, simulation hierarchy, orders the variety of languages that describe function and structure. It assists the passing from the goal function, constrained by functional parameters, to the structural form, and inversely, to determine the mathematical function/ physical behavior of a structurally defined system.

Architecture is the level on any hierarchy of any type that sees both ends.

Researching intelligence by simulating it, to enable intelligent simulation, demands the study of combined essential mathematical structures to understand the different hierarchy/ abstraction types. Being a hierarchical relation between static/ structural and dynamic/ functional structures simulation contributes essentially to understand the human mind. We model the consciousness for simulating the intelligence, then to reach for intelligent simulation.

A hierarchical type expressing reflexive abstraction can represent the conscious knowledge. The aspects of the Reality, and of the human mind reflecting it, have not to be neglected, although they are neither constructive nor intuitive. A way from Reason to Intelligence is to integrate Consciousness and Intention, then further to relate Intelligence and Faith to become Reality-aware.

We could consider just the simplifying types of hierarchy (classes, symbols, modules) and then express the construction, hoping to aim the absolute liberty, if we considered God as the simplest, totally unconstrained, essence of the Reality. However, we can simulate/ construct/ work/ live, associating knowledge hierarchies to all our activities, aiming to constructive understanding of the most complex absolute necessity, by this defining *God*.

Abstraction is the human gift for going beyond natural limits, meanwhile extending pure reason to real intelligence. Faith and intelligence can converge to integration, or destroy each other if not linked philosophically by Conscience. The power to abstract is the crucial difference between human and other natural living beings.

Divide et Impera et Intellige applies the hierarchical expressed abstraction. *Intelligence* and *faith*, like any dichotomy, can converge to integration or can destroy one another if not associated by *Conscience*. Metaphor is a popular instance: we detail the metaphorical thesis:

God is the absolute abstraction,
the evolution goal for faith-assisted intelligence

Function is a transformation that can be mathematically formalized, or physically instantiated as temporal behavior. *Structure* is a set of properties that characterize a mathematical/ physical space. *Structured set* = (Set, structure). *Simulation* is the relation between function and structure.

Language/ system is a generic form of a mathematical/ physical *model* resulting of an inversion-able representation of the simulation object. *Abstraction* is a human defining capacity that enables him to think. Simplifying abstraction concentrates on a superior level the information that is considered essential for the current simulation approach. Reducing informational complexity has in view to clear the operation and to ease its formalism; it can be only quantitative, but also qualitative.

The reflexive abstraction-expressed as knowledge hierarchy type-tries to understand itself better at higher levels, understanding more of the inferior ones.

God is in us-faith belongs to our definition, with us-by the others, and for us-spiritual evolution, conditioned, then assisted, to be followed by the social one. Against the danger of dichotomy, we concentrate in three different ways on the unique Reality- *Plato*:

- Art for the art-to look for the essential Way,
- Science with God's fear-to search for the existential Truth,
- Engineering-to understand the Being and to concentrate more on the Spirit in our Life.

Reason is an extension of the nature. Nature is not an ephemeral context, but the matter we are built of in order to develop spiritually. The integration experiments for the spirit-matter dichotomy failed because of their extremism. To go further, thinking while advancing, we divide twofold, as we cannot yet Intellige the dichotomies:

Spirit-Matter (force substrate, software-hardware) ⇒
real-natural (continuous-discrete, analog-digital),
form-contents (category-functor, representation-simulation, class/structure-function),
real-possible, perspective-profoundness,
beauty-truth (arts-science, mathematics-physics)
⇐ *Space-Time* (evolution)

Clearly, balance should not be in most dichotomies. Yin-Yang can represent any dichotomy, and its extension to a triad: *faith, conscience, intelligence*.

Society is only the memory of the past, manager of the present problems, assurance for a right future. We have to live together in respect of the others on the way to understand each other, in order to evolve toward essential beings for an integrated existence.

The present society is extremely materialistic, and tries to destroy every trace of ideal. The method is analog to *embedded systems*: adapt the human to the socialistic society-society is the master and human the slave.

We have to surpass the limits imposed by the essential dichotomy by a unique ideal, named God that should be constructive by continuous intelligent reconfiguration. Human among humans should reflect a strategic equilibrium, without hiding or even violating, as happens nowadays, the principle that the society has to assist unconditioned the individual, with correct continuous education, and assistance by an intelligent faith to search and research the *unknown*.

The unknown can be interpreted as a *unique God*: the absolute freedom by understanding all the necessities, and the absolute unity by closing all the *Divide et Impera et Intellige* necessary for the Way to look for the Truth along the Life. Extending reconfigurability to the simulation itself, by a self-aware simulation, we get self-control of the simulation process.

We build a knowledge hierarchy corresponding to the simulation hierarchy, and by expressing both hierarchy types in the reference system of the basic hierarchy types (classes, symbols, modules) we create the context for a self-organization of the simulation. The triad of the basic hierarchy types corresponds to the fundamental partition of the real life (beauty-arts, truth-science, good-engineering) that has to be continuously integrated by philosophy (essence, existence, being).

The absolute functionality is symbolized by yin-yang, while the waves suggest hierarchical levels that are increasingly structured for simulation and knowledge. Each of the nondeterministic separated complementary pairs of the *yin-yang* model is functionally structured like (interface, kernel, ambassador of the complement). The model was not randomly selected: it is formed of three tangent circles emphasizing the centers of the inner ones. It retains only the essence of a dichotomy symbol that suggests a complete integration of the parts without loss of autonomy, realized by vicinity and pointing one to another. The Chinese symbol reflects the importance of something else, reminding of creation as love for something else.

Three circles, each tangent to the others, models a partition of something to be understood in order to get further, say the center of Europe. Circle is *cerc* only in our mother tongue, a perfect expression: *Cer* (sky) is the infinite, *cerc* is the finite representation of the infinite, by the permanent link from the (never)begin to the (never)end. π is the most famous real number-*Pythagoras*. *Cerc* means perfection, which we permanently desire, therefore there exist integer numbers, having a perfect and beautiful theory, but not forgetting to continue the evolution searching and researching further-*cercetare*. The western Europeans attain research/ *rechercher* by recursive search/ *chercher*. The essential difference between analog and digital simulation paradigm is induced by that between the mathematical structures their models are based on algebraic-for digital, analytical-for analog. In view of intelligent simulation the whole intelligence has to be simulated, i.e., conscience and intention along with adaptability.

The discrete nature of simulation-design is a sequence of decisions and verification implies a sequence of stimuli-does not easily match the continuity of analog properties or parameters. The difficulty of analog simulation is avoided by defining an auxiliary representation domain that intermediates between the behavior and the structure.

At this circuit level, the problem is decomposed into topology selection and dimension computation. The first process is discrete and the second one is continuous (Hofstadter, 2000) over a restricted problem space.

Object-oriented representation lends itself for this instance of complementary form-contents. However, topology selection is more systematic if continuous modifications of the form are possible, and dimension computing is more efficient when symbolic algebraic methods are used.

We searched for a compromise between simulation algebra and analog analysis on three ways: defining for the analog domain, upper abstraction levels that are governed by algebraic laws; modeling analog simulation for algebraic-analytical structures from the functional analysis, (Rudin, 1973) associating an analytical syntax to the analog simulation process.

$$\textit{Simulability is computability}^{\textit{continuum}}$$

Intelligence = Consciousness × Adaptability × Intention, and Faith = Inspiration × Intuition × Imagination, are complementary parts of the human mind, separated by the Conscience = Consciousness × Inspiration. Conscience demands continuous feedback.

Both intelligent simulation and simulation of intelligence demand transcending the present limits of computability toward simulability.

The historical experiment of the pure reason should have ended long time ago. Human thoughts cannot be explained or handled by our adaptability-based reason, even if non-deterministic or parallel. Therefore, reason has to extend to intelligence in the context of faith. An obvious way is to integrate consciousness, then intention and imagination to intelligence, then to extend this to inspiration and intuition.

4. Discrete to continuous

Mathematics develops the countable natural numbers to the uncountable real numbers closing to the inverse, on its three integrated ways: algebra, order, and analysis. Physics uses particles or fields in various chapters. All other sciences are chapters of physics, inheriting and developing the inheritance. At the limits of reasonable understanding, quantum physics tries to balance the knowledge and the unknown, without success. Engineers have always considered digital a mere ingenious abstraction of analog. Presently, we talk about electronic computers, but the nowadays trend is to copy from the living Nature, i.e., the emulation of the advantages the living beings show to achieve unconsciously complex duties.

Vanguard domains are biotechnology and computational intelligence. Neither intelligence nor life is well understood; remember *Goethe's Zauberlehrling*. More important is that emulation is less human than simulation-they should always develop in parallel, permanently exchanging experience. Reality does not reduce to Nature, as cardinal ($\mathbb{N}$) is strictly inferior to cardinal ($\mathbb{R}$)-*Georg Cantor*. Reason is the closure of the Nature relative to the primary operations, as $\mathbb{Q}$ results from the closure of $\mathbb{N}$ to the inverse operations of addition and multiplication. However, the Reason is dense in Reality-as the real numbers are the analytical closure of the rational numbers, $\mathbb{R} = \left\{ \lim_{n\to\infty} (q_n) \,|(q_n) \in \mathbb{N} \to \mathbb{Q} \right\}$.

Reality extends beyond Nature and Reason, not just for the quality of the quantity, but also regarding the power of transforming operations. $\mathbb{R}$ closes $\mathbb{Q}$ to the inverse of exponent-the last arithmetic operation resulted by recurrence of the prior one, which can be pursued by Reason, e.g., algorithmically.

Reason closes Nature 2 the inverse of natural functions.

Further, closing to the inclusion order, the set of all subsets of countable sets is the uncountable IR, the power of continuum. To get to complex numbers is a matter of imagination.

Example: *Transfer Function Singularities:*
We presented a related work that compared two methods to determine the poles and zeros of a transfer function, based on state-equations, respectively on node-equations (Mărculescu & Niculiu, 1987). Complexity of the set-up actions of the first was balanced by weak convergence of the second. This is a typical case to try heuristics together with expert systems; hence we presented (Niculiu & Manolescu, 2009) a knowledge-based object-oriented analog simulation system. The Newton-Raphson method was used in circuit simulation for 40 years, and the interest for its optimization has not decreased (Zhu *et al.*, 2007).

The graphical or numerical results of a circuit simulator are the primary information that has to be sampled with a variable rate appropriate to the simulator output variation. Knowing the dominant singularities is decisive for simulation, as they reflect the stability of the circuit, (Niculiu *et al.*, 2008) or can represent primary data in formal simulation, e.g., root locus method. The transfer function of a linear-linearized around a static operation point-circuit is a ratio between real coefficient polynoms with complex roots, functionally describing the frequency behavior.

A pattern-matching search decides which rule applies, and at the end, the transfer function results as a two polynoms ratio (Manolescu *et al.*, 2009). The search is bottom-up while determining the singularities, and top-down to find recursively the dominant ones. The function of our program is threefold:

- *classification* - to recognize the type of singularity from the transfer function or Nyquist diagram;
- *control* - for stability;
- *anticipation* - to link the results to possible alternatives for improved behavior.

The is object-oriented, and written in Java. The main classes are *Element, Rule, Match,* and *Act*. The input is a circuit simulator .AC result-numerical or graphic, the output a rational function representing the approximate transfer function that describes the essential behavior.

5. Revolution by opening to evolution

The human has to enlarge, not to tear, the bands of the Reason, and to apply them to the society. The Reason has to transform into the consciously recognized limits of the Intelligence in front of the Faith that offers to the human the way to evolve beyond any limits.

A reasonable society is hierarchical. Its essential architecture contains three tree-like structures for the same set of humans, therefore, interdependent: arts, science, and engineering-technology.

The social hierarchies reflect only a temporary order, generated by humans, to help them concentrate on the spiritual evolution, without neglecting the material problems. The hierarchical social structure can assure an optimal organization of humans among humans. The interdependence of the three social classes is assured functional, not only structural. Without giving up anything essentially human: different cultures and social or natural togetherness, humans among humans have a lot in common: philosophic desire, comprehension of the own hierarchy in the context of the other two, free life based on understanding the necessities, constructive fear of the unknown, and especially love for creation/ discovery. Except the three cultural ways, that permanently *Divide et Impera et Intellige*, there is no other.

We need Consciousness to return intelligently to Faith.

People of one choice exist, in all senses of the word. They either comprehend all the alternative ways and their convergence, or, in the context of natural love for philosophy and interest for the other selectable directions, put more passion in one direction. Of the first category are temporary elected, in different convergent hierarchical modes, the social leaders, of the second, the institutional directors. Both kinds of leaders are more philosophical than their cohabitants, even if the ones master the strategic perspective given by an attained peak, while the others have the joy of the courage to climb into profoundness. The elected artists permanently reconfigure a system of laws, to be beautiful by intelligibility, true by consistence, and good by human understanding. The elected physicists, pure or of different correlated scientific domains all collaborating with mathematics and engineering, govern by research strategies with Gods Fear. The elected engineers critically construct and criticize constructively. For any social role, the elected concentrate, respectively, on Faith-mathematicians, Intelligence- physicists, and Conscience-engineers. There always exists a human, called No.1 or the Philosopher, depending on the stability of the times, cloudy or clear Sky. He will always lead directly the elected or the philosophers, who will know to educate and optimally learn the humans of all ages, including themselves. We have to start. Otherwise, it is no hurry (Penrose, 1996).

Intellige is to link, to understand, and to be aware. In Latin: *intellego* = to understand, to feel, to master, to gather in mind. Artificial has a derogatory sense; however, the root of the word is art. Arts remind of liberty, as *Arts for arts*. Artificial is at first sight the complement of natural. Our ideas transfer us to places that are neither natural nor artificial. Maybe artificial means something natural created by the human being and Nature is an extension of our body.

However, we feel to be superior to Nature, as to our body: we can think. Why are only humans creating arts, why do they need to know more, and why do they construct other and other natural things they have not found in the Nature? We learned the arts have to discover the Beauty that science looks for the Truth, and that engineering invents things to help us, caring for the Good. *Goethe* wrote on the theatre in Frankfurt:

Das schöne wahre Gute

because the three wonderful scopes have to be always together. He stretched the Good that is important to all natural beings, whereby for beautiful or true cares only the human being. Arts and science demand a distinct power for both development as understanding, and possibly for usefulness. The power of abstraction distinguishes us among the natural beings. Engineering is to be ingenious, not only to design engines. Any human choice to surpass the Nature by arts, to know it better by science, or to enrich it by ingenious construction, is as noble and legitimate, because to follow any selected way demands intelligence and faith.

Artificial intelligence has an initial sense of enriching natural domains by natural extensions. Reason is an extension of the Nature. The natural language whispers to us: as the rationales are a straight extension of the naturals, if we neglect the integers, however, you remain in a countable world as the Nature initially is. We should not be ashamed if someone that we only understand by proper preparation and that is at least so powerful as the Nature; let's remember the beautiful mother language. *Cer* (sky) suggests the infinite, and we desire to see it and to link the its begin to its end, or better the never begin to the never end, and we find the *cerc* (circle). The language whispers to us again: is not rational, it is more than this, and it is as if we listen to a symphony by *Beethoven*. We understand that the Reality of our Existence is more than the Nature of our Being, therefore, we should know them better, because only Nature can open us the way to Reality. We wonder whether any of the alternative ways demands the same kind of intelligence, and if not, which of them should we first research (*cerceta*) in order to simulate.

- Arts are free, and even when they return to Reason, as mathematics, they bring results, that could before just be seen by Intuition, to be sent by Inspiration and Imagination to Intelligence;
- Physics reaches and gets conscious of Reasons limits, both by the quantum theory and by the too complex phenomena, e.g., society and human;
- No difference seems to be for the intelligence that is useful to one of the ways. An example, that confirms that they simply represent different approaches to understand and develop the-presently natural-Reality, is *architecture* that we cite for each of them.

Therefore, there is something else in the Intelligence, which allows us to consider ourselves humans, human groups, peoples, beings on the Earth, or conscious beings in the physical Universe. We also feel that there is something essential beyond the physical-the metaphysical-*Plato*. More, there is something exterior to the human intelligence, without that we could not fight the Time to evolve. We have to feel complete, even if we need education and permanent work in communication with the other humans, of the past, the present, and the future.

We need Conscience to link Faith to Intelligence

You see now why we neglected the integers when we showed that the rational numbers are countable, i.e., they are as much as the naturals. This way, we divided the problem into two others that we do not forget to reintegrate after we have solved them - *Divide et Impera et Intellige*. We count the positive rational numbers x/y along the secondary diagonals in an odd quadrant of the coordinate system - $x0y$. Then we repeat this counting for the negative ones in an even quadrant. Finally, we count them together by jumping between quadrants for every current number. We come to the idea how to count the $\mathbb{IQ}$ is without using *Divide et Impera et Intellige*, that we have to keep in mind for harder problems, as Life, Truth, and Way.

We have to remember the abstractions that assisted us to go further. We said complete human to someone complete in a context, what implicitly supposes the power to go beyond the context. This is the story of the integers-*integer* = perfect, complete-they have a beautiful complete theory, however, do not forget to build the rational numbers to feel as close as needed to any real number. Nevertheless, they realize this is not enough, rewarded by the conscience of the continuous Reality-infinitely more powerful than the discrete/ countable one.

To $\mathbb{IR}$, we get by the perfect circle that is beyond the power of Reason. Another way to the same scope is by the boring perfection of the square, when computing its diagonal ($\sqrt{2}$). Again and not fortuitous this alternative is due to Pythagoras, the godfather of π . The beautiful natural induction tells us that the equilateral triangle and the square are but the pioneers of the regular polygon sequence that converges to the circle. Encouraged, we turn an equilateral Δ or a square about itself, obtaining the area of the circumscribed circle when the number of sides $n \to \infty$, from the areas of the n-sided polygons. However ... we wanted to approach π by a sequence of rational numbers, but the example is wrong. Again, we hear like a sweet wind from the sea: *Alle guten Dinge sind drei* and intuitively sense that we have to know how mathematics masters the infinite.

For long time, we knew nothing of sets, but we knew too well to play the role of a calculator. We should not forget what Intuition said to Intelligence, by Imagination: we just had imagined a sequence of *algebraic irrational numbers* converging to the *transcendent* number π.

We scare to be further taught rather what a discrete computer, instead of what an intelligent human, has to know. For example, we plan to realize artificial intelligence, to have a friend that is conscious of the problems to solve together. For the moment, there is no artificial intelligence. However, we learn to be conscious of the computer limit to process only rational numbers. This means it uses a sequence $(x_n)\, n \in \mathbb{Q}$ that converges to $\sqrt[n]{a}$ (Newton), what reminds us of the density of $\mathbb{Q}$ in $\mathbb{R}$.

Perhaps not practice has to push us into evolution, but Gods fear, i.e., the scientific desire for further ascending encountered on any reached level of knowledge. Conscience attaches us to science and unfastens us of the false eternity, arrogated by some level of the evolution to freedom. To be free we have to understand all the necessities in the Reality, or metaphorically: to escape God of any fear. Intelligent systems need a cosimulation of the parts that belong to different domains, e.g., hardware + software, in the context of representation unifying for simulation-design/ verification/ optimization.

Unified simulation of the hardware-software systems is imposed by the incompatibility or the lack of optimality that results of the initial partition of the system. The design-verification cycle is not efficiently processed for a fixed partition. This disadvantage is eliminated when the simulation methodologies are unified, e.g., by categorical strategies, what implies planning and learning, i.e., the possibility for interlevel communication in the knowledge hierarchy. An intelligent simulator learns by recursive generation and validation-possibly interactive-of models. The objective of human-machine dialog is to advance toward simulated intelligence by knowledge communication in a common language between human and his mental/ physical extensions.

We pleaded that abstraction is the handiest tool for the human among other beings. Let us use it to simulate the present situation. Neglecting the conscience, let us see what remains of the human.

$$\textit{What should I do? What you want. What do I want?}$$

$$\textit{What you like. What do I like? What you should do.}$$

This is a cyclic definition only at first sight, because most probably that what he should do has changed while crossing the cycle by what he wants or likes. We sketched a minimal intelligent system: it has to be adaptable, self- and context-aware, and to communicate with the exterior by signals and actions. It is most disputable that Consciousness can be extracted from Conscience. However, we try that the system is fashionable. Consequently, we also abstract from the fact that a discrete processing is not capable of self-consciousness- *Gödel*. To avoid any discussions we abstain from any hypothesis on the class of the processor, discrete or continuous.

What is the Conscience: it is the link, in our mind, between what we are conscious of and what we are not. Presently, only the extended to Reason adaptability, and the unjustified Intention, are conscious. We can imagine an intelligent machine that looks like a human-robot is labor, Slavonic. It accumulates knowledge and behavior rules by preprocessing the senses, and it can change the interior defining rules (reconfigurable) corresponding to the behavioral (professional, ethical) knowledge that is considered most important, e.g., most recent or most decent.

Therefore, it can consciously filter the actions that determine a new state of the context, what also means new knowledge to accumulate and to be conscious of (adaptability). It means, the dialog with the external environment determines the intentions. If the system had conscience, the external dialog would be more complex and interesting. Consciousness only makes the adaptability more efficient, what, among others, transforms the human into the most powerful animal.

Why do we compare the system without conscience with an animal, not to a human? It is true that we could compare it to an animal, if we had attributed intuition to it. However, what for should we do this, when the human just adapted to a consumption society? The built artificial objects and the socially useful natural objects send him the necessary messages to adapt consciously at the rising efficiency of the society. He neglects both the warnings from the superfluous Conscience and the unnecessary Intuition. If sometimes the two beasts shout too loudly, it is just unpleasant. To be useful Intuition should be linked by Conscience to Intelligence, and intelligently bridled by Imagination.

More, Intuition should also know to bridle by Intention the Adaptability. Whether he is human or animal, the human is anyway a machine, a social machine. His use is to contribute at the eternity, on an arbitrary level of evolution, of a materialistic consumption society. The evolution is for the human among humans, assisted by a reasonably organized society developing by the human, for the human towards the Human. We said arbitrary level as, if the educated and encouraged consumption were not strictly materialistic, the human himself would escape from the vicious circle together with the others. And more, the present level is artificial in the human evolution. The essential limit of discrete computability, inherited by the computational intelligence, is the necessity of self-reference to integrate the knowledge of the levels to that of the metalevels for modeling the Conscience.

A hierarchical type representing reflexive abstraction can model the conscious knowledge and the knowing Consciousness, if it categorically collaborates with a simulation hierarchical type. We have to search and research for the aspects of the Reality, and of the human mind that reflects it, even if they are neither constructively nor intuitively expressible.

The humans that have consciously chosen the eternity have transformed into dolphins. This also taught the English how to conserve a stabile existence. The desire to stop the human evolution on arbitrary stages has no real argument.

To conclude: Intelligence is more than Reason, to make us feel as beings superior to Nature, what also means that we have to respect Nature more:

Spiritus Sanus in Mens sana in corpore sano

For the present, the evolution is forced to halt on an inhuman level, a consumption society transforming the society into a beehive without interest for Conscience and Faith, which most probably has been realized by destabilization of all revolutionary forms.

We need intelligent Faith to develop to freedom as humans among humans

6. Conclusion

The religion had to learn us about God's existence in our being. The philosophy has to learn us about essence, existence, and being. Our conscience is our representation of the essence of our existence as being, i.e., God is in us, for ourselves, and among ourselves.

We have to be to search our essence researching our existence. *Divide et Impera et Intellige* has three parts as *alle guten Dinge sind drei* of the most philosophic European people.

Neither intelligence nor life is well understood, remember *Goethe's Zauberlehrling*. More important is that emulation is less human than simulation, remember *Mozart′s Zauberflöte*; they should always develop in parallel, permanently exchanging, remember *Thomas Mann's Zauberberg*.

References

Ageron, P. (2001). Limites inductives point par point dans les categories accessibles. *Theory and Applications of Categories* **8**(11), 313–323.

Bacalu, I. (2004). *Differential Equations. Stability*. Matrix Publishers, Bucureşti.

Hennig, E. and R. Sommer (2001). A reliable iterative error tracking method for approximate symbolic pole/zero analysis. In: *Proceedings of ECCTD'01 - European Conference on Circuit Theory and Design, August 28-31, 2001, Espoo, Finland*. pp. I.193 – I.196.

Hofstadter, D.R. (2000). *Godel, Escher, Bach: An Eternal Golden Braid*. Penguin; 20-th Anniversary edition. 20-th anniversary edition with a new preface by the author.

Keutzer, Kurt, Sharad Malik, A. Richard Newton, Jan M. Rabaey and Alberto Sangiovanni-Vincentelli (2000). System-level design: Orthogonalization of concerns and platform-based design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits & Systems* **19**(12), 1523–43.

Manolescu, A., A-M. Manolescu and T. Niculiu (2009). A simple algorithm for approximating the transfer function of analog integrated circuits and systems. In: *Proceedings of the 32nd IEEE International Semiconductor Conference, Sinaia, Oct. 12-14, 2009*. pp. 503 – 506.

Mărculescu, R. and T. Niculiu (1987). Transfer function singularities determination. In: *Proceedings of the IEEE International Semiconductors Conference, Sinaia, 1987*. pp. 61 – 64.

Niculiu, T., A. Copilău, C. Lupu and A. Manolescu (2008). Hierarchical intelligent analog simulation. In: *Proceedings of the I3M European Modeling and Simulation Symposium, Campora San Giovanni, 2008*. pp. 323 – 332.

Niculiu, T. and A. Manolescu (2009). Symmetry for intelligent analog simulation. *Romanian Academy Noema J.* **VIII**, 230–250.

Penrose, R. (1996). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press, USA.

Rudin, W. (1973). *Functional Analysis*. McGraw-Hill. 1st edition, McGraw-Hill Series in Higher Mathematics.

Traub, J.F. (1999). A continuous model of computation. *Physics Today* **52**(5), 39–43.

Zhu, Zhenyong, He Peng, Chung-Kuan Cheng, Khosro Rouz, Manjit Borah and Ernest S. Kuh (2007). Two-stage Newton-Raphson method for transistor-level simulation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **26**(5), 881–895.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 56–70

Development of fuzzy supplier-rating by trapeze fuzzy membership functions with trigonometrical legs

Tamás Portik[a,*], Tamás Varga[b]

[a]*University of Debrecen, Doctoral School of Informatics, Debrecen, Kassai út 26, HU-4028, Hungary*
[b]*IHI Charging System International GmbH, Leimen, Im Schussgarten 5, DE-69181, Germany*

Abstract

In every fields of industry, suppliers create the elements for the final products of OEM (Original Equipment Manufacturer), which means, suppliers "define" the quality of the final products via the quality of elements. Due to this, the continuous evaluation of supplier performance can be one of the most efficient risk assessment tools to identify weakness in early stages, make the possibility to implement corrective actions in time. The paper offers new development of supplier-rating based on fuzzy set theory where trapeze fuzzy membership functions are used with trigonometrical legs. This investigation is made for providing better rating by the authors. This interpretation is more effective and correct of course in point of view of manufacturing.

Keywords: Supplier evaluation, fuzzy set theory, quality management, set-transition.
2000 MSC: 03E72, 03E75.

1. Introduction

In industrial field there are two types of companies. One of them is OEM (Original Equipment Manufacturer) and the other ones are suppliers. The continuous evaluation of supplier performance should be one of the most efficient risk assessment tools to identify weakness in early stages, make the possibility to implement corrective actions in time. One of the authors is a practical expert, who provided this problem. The practical experts always think with notions and the transitions between the notions is described well with fuzzy mathematics. In manufacturing it is indispensable to rate the suppliers because the quality of the product(s) mainly depends on supplier according to (Esse, 2008),(Humphreys *et al.*, 2007) and (Krause & Ellram, 1997). They have been studied the supply chain and what kind of condition should be to have high quality of product(s). The authors have been studied the supplier rating in their earlier papers like (Portik *et al.*, 2011) and achieved some results which will be shown in section two. To have base for fuzzy set theory and logic it is a good introduction (Ross, 2010) and (Retter, 2006).

Liu, Martínez, Wang, Rodríguez and Novozhilov made a very deep overview about application of fuzzy mathematic which cover wide range of fields of applications to show possible utilizations of risk assessment activities, which were created to identify risk, what could be potential source of harm, or reason of quality

*Corresponding author
Email addresses: `portik@eng.unideb.hu` (Tamás Portik), `t.varga@ihi-csi.de` (Tamás Varga)

problems. Risk assessment is the most efficient tool to sufficiently minimize the relevant risk in whole life cycle stages including design, implementation and operation (Liu *et al.*, 2010). Even in military science it can be used fuzzy logic for risk assessment according to (Pokorádi, 2002). Pokorádi gives complex description of methods applied for fuzzy logic based decision making (Pokorádi, 2008). Another Pokorádi's paper used fuzzy logic to model inaccuracy and uncertainty of human thinking and proposed fuzzy logic application for risk assessment of hazards connected with hydraulic system (Pokorádi, 2009).

The "classical" way of supplier evaluation is to create groups with crisp boundaries based on measurement values like DPPM (Defect Part Per Million). The main disadvantage of crisp boundaries, that small difference in the input can cause big difference in the output.

The aim of this paper is to present this issue and show a new method in supplier rating which is better rating on transient phases. It has been stated that evaluated model can be used by the practicing engineers who do not have knowledge and expert on fuzzy mathematics.

The outline of the paper is as follows: The Section 2. shows authors' earlier results and a proof that the classical trapeze fuzzy membership functions are not good for rating on transient phases. what it means a fuzzy membership function with trigonometrical legs. The Section 3., what it means a fuzzy membership function with trigonometrical legs, explains the constant set-transition in rating-points similarly presents the proportional set-transition as well. The Section 4. provides the application of the constant and the rate set-transition. The Section 5. gives the conclusion and future work.

2. Earlier results

In this introduction the authors present some earlier results which was achieved in some earlier papers. This examination is based in point of view of the failure rate evaluation. It should be introduced a standard number in equation (2.1) which is called DPPM (**D**efected **P**arts **P**er **M**illion).

Definition 2.1. DPPM number is defined with the following equation:

$$DPPM = \frac{NNCE}{NAEFS} \times 1000000, \tag{2.1}$$

where:

NNCE	–	**N**umber of **N**on-**C**onform **E**lements,
NAEFS	–	**N**umber of **A**ll **E**lements **F**rom **S**upplier (Portik *et al.*, 2011).

On DPPM numbers are created some groups which is shown on Table 1. This groups have crisp boundaries which should be fuzzyfied to make more sensible rating for suppliers. Our fuzzy mathematic based on (Ross, 2010) and (Retter, 2006). Our problem is graphically represented on Fig. 1 and Fig. 2.

Table 1
Groups based on DPPM

DPPM	%	Rating-Point (*mark*)
0—2000	0.2 %	20
2001—4000	0.4 %	17
4001—7500	0.75 %	15
7501—10000	1.00 %	10
10001—15000	1.15 %	5
> 1500	≥ 1.5 %	0

For further investigation the authors introduced so-called *set-transition*.

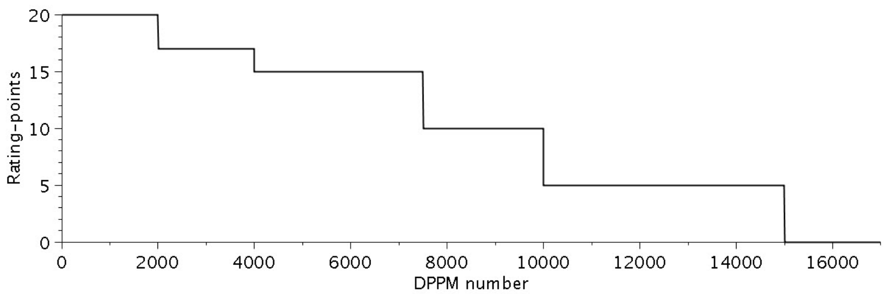


Figure 1. The graphical representation of the basic problem.

Definition 2.2. *Set-transition* is called as the length of projection of trapeze legs to the abscissa. It is signed with H (Portik *et al.*, 2011).

By definition of set-transition, it was investigated with line trapeze-legs fuzzy membership functions if the H was constant or proportional. In case of H is constant, the set-size is not to take into consider. If H is rate then the set-size is taken into consideration. The problem was that the transient phases come somewhere concave somewhere else convex and this is not allowed because it takes different methods between the rating transient phases e.g. it is shown on Fig. 3 and Fig. 4. The calculation method was the next:

$$M = \frac{\sum_{i=1}^{n} P_i \times \mu_i(DP)}{\sum_{i=1}^{n} \mu_i(DP)} \tag{2.2}$$

where:

DP	–	DPPM number,
M	–	rating point which belongs to the given DP,
P_i	–	mark which belongs to the i^{th} fuzzy membership function,
$\mu_i(DP)$	–	the value of the i^{th} fuzzy membership function to given DP,
n	–	the number of fuzzy membership functions (Portik *et al.*, 2011).

A classical trapeze fuzzy membership function is given by equation (2.3) and the short sign for it is $\mu(x)_{classic} = [a, b, c, d]_{classic}$.

$$\mu(x)_{classic} = \begin{cases} \frac{x-a}{b-a} & ; \quad a \le x < b \\ 1 & ; \quad b \le x \le c \\ \frac{d-x}{d-c} & ; \quad c < x \le d \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{2.3}$$

This was a short overview about earlier results.

Proposition 1. *There exist such transient phases by classical trapeze fuzzy membership functions which are given by proportional set transition, that some of them are convex and the other ones are concave and the computation is made by two trapeze legs only.*

Proof. The formulation is for proving it is the next according to equation (2.2):

$$M(x)_{i,i+1} = \frac{[a_i, b_i, c_i, d_i]_{classic} \cdot p_i + [a_{i+1}, b_{i+1}, c_{i+1}, d_{i+1}]_{classic} \cdot p_{i+1}}{[a_i, b_i, c_i, d_i]_{classic} + [a_{i+1}, b_{i+1}, c_{i+1}, d_{i+1}]_{classic}} = \frac{\frac{d_i-x}{d_i-c_i} \cdot p_i + \frac{x-a_{i+1}}{b_{i+1}-a_{i+1}} \cdot p_{i+1}}{\frac{d_i-x}{d_i-c_i} + \frac{x-a_{i+1}}{b_{i+1}-a_{i+1}}} =$$

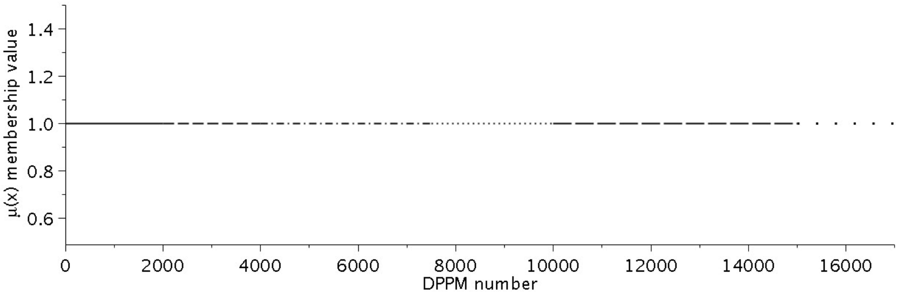


Figure 2. The fuzzy membership functions which belong to the basic problem.

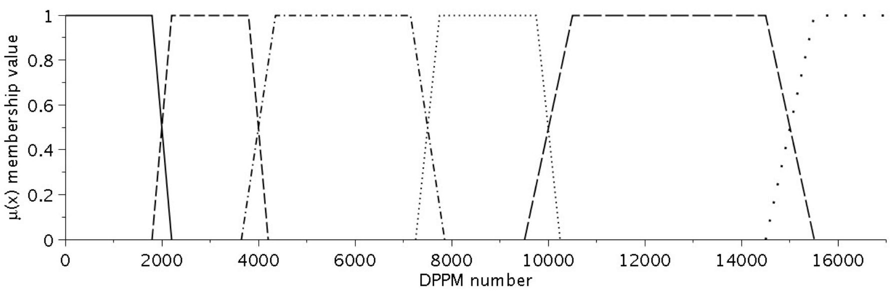


Figure 3. The trapeze fuzzy membership functions at $H = 40$ % by line legs.

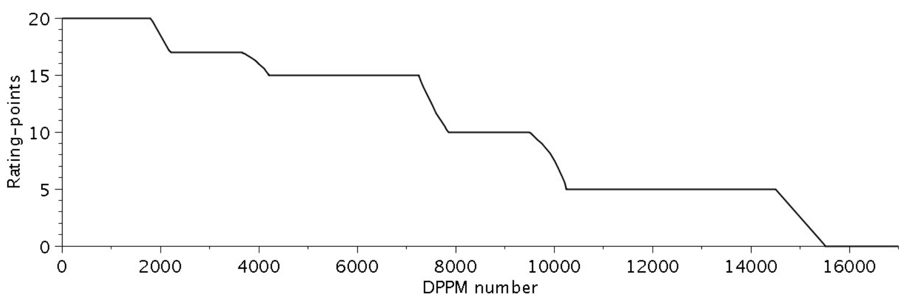


Figure 4. The rating-points at $H = 40$ % used line legs trapeze fuzzy membership function.

$$= \frac{(-p_i b_{i+1} + p_i a_{i+1} + p_{i+1} d_i - p_{i+1} c_i)\, x + p_i d_i b_{i+1} - p_i d_i a_{i+1} - p_{i+1} d_i a_{i+1} + p_{i+1} a_{i+1} c_i}{(-b_{i+1} + a_{i+1} + d_i - c_i)\, x + d_i b_{i+1} - 2\, d_i a_{i+1} + a_{i+1} c_i}$$

if $x \in [c_i, d_i] \cap [a_{i+1}, b_{i+1}]$ and $i \in \{1, 2, 3, 4, 5\}$. The next step is to give the fuzzy membership functions which are in Table 2.

Table 2
The classical trapeze fuzzy membership functions with rate set-transition $H \in]0,1]$

i	Fuzzy membership function	Mark (p_i)
1.	$[0, 0, 2000 - 500 \cdot H, 2000 + 500 \cdot H]_{classic}$	20
2.	$[2000 - 500 \cdot H, 2000 + 500 \cdot H, 4000 - 500 \cdot H, 4000 + 500 \cdot H]_{classic}$	17
3.	$[4000 - 875 \cdot H, 4000 + 875 \cdot H, 7500 - 875 \cdot H, 7500 + 875 \cdot H]_{classic}$	15
4.	$[7500 - 625 \cdot H, 7500 + 625 \cdot H, 10000 - 625 \cdot H, 10000 + 625 \cdot H]_{classic}$	10
5.	$[10000 - 1250 \cdot H, 10000 + 1250 \cdot H, 15000 - 1250 \cdot H, 15000 + 1250 \cdot H]_{classic}$	5
6.	$[15000 - 1250 \cdot H, 15000 + 1250 \cdot H, 0, 0]_{classic}$	0

The best way to prove the statement is to use computer algebra system like Maple 14. This is good for symbolical and numerical computation. To prove convexity or concavity the second derivative of $M(x)_{i,i+1}$ is used. Generally the first derivative of $M(x)_{i,i+1}$ is

$$M(x)'_{i,i+1} = \frac{(-b_{i+1} + a_{i+1})(-p_{i+1} + p_i)(-d_i + c_i)(-d_i + a_{i+1})}{((-b_{i+1} + a_{i+1} + d_i - c_i)\,x + d_i b_{i+1} - 2\,d_i a_{i+1} + a_{i+1} c_i)^2}$$

and the second derivative of $M(x)_{i,i+1}$ is

$$M(x)''_{i,i+1} = -2\,\frac{(-b_{i+1} + a_{i+1})(-p_{i+1} + p_i)(-d_i + c_i)(-d_i + a_{i+1})(-b_{i+1} + a_{i+1} + d_i - c_i)}{((-b_{i+1} + a_{i+1} + d_i - c_i)\,x + d_i b_{i+1} - 2\,d_i a_{i+1} + a_{i+1} c_i)^3}.$$

Now, the Rating-point should be provided which is presented in Table 3. according to the classical fuzzy membership functions on Table 2.

Table 3
The Rating-point with function $\mu(x)_{classic}$ and proportional set-transition $H \in]0,1]$

$M(x)_{i,i+1}$	$M(x)''_{i,i+1}$	Interval
$M(x)_{1,2} = \frac{6000+18500\,H-3\,x}{1000\,H}$	$M(x)''_{1,2} = 0$	$[2000 - 500 \cdot H, 2000 + 500 \cdot H]$
$M(x)_{2,3} = \frac{236000+112000\,H-59\,x}{12000+7000\,H-3\,x}$	$M(x)''_{2,3} = \frac{-462000\,H}{(12000+7000\,H-3\,x)^3}$	$[4000 - 500 \cdot H, 4000 + 500 \cdot H]$
$M(x)_{3,4} = \frac{5\,(7500+21875\,H-x)}{2\,(-7500+4375\,H+x)}$	$M(x)''_{3,4} = \frac{131250\,H}{(-7500+4375\,H+x)^3}$	$[7500 - 625 \cdot H, 7500 + 625 \cdot H]$
$M(x)_{4,5} = \frac{15\,(10000+1250\,H-x)}{-x+2500\,H+10000}$	$M(x)''_{4,5} = \frac{-37500\,H}{(-x+2500\,H+10000)^3}$	$[10000 - 625 \cdot H, 10000 + 625 \cdot H]$
$M(x)_{5,6} = \frac{-45000+6875\,H+3\,x}{625\,H}$	$M(x)''_{5,6} = 0$	$[15000 - 1250 \cdot H, 15000 + 1250 \cdot H]$

So to prove now the existence, it is enough to choose transient phases.
Let it be $H \in]0,1]$, $M(x)_{2,3}$, $M(x)_{3,4}$ and $M(x)_{4,5}$, their second derivate are strictly monotone and continues functions on the given intervals. Therefore $M(4000 - 500 \cdot H)''_{2,3} = \frac{-462000}{(8500 \cdot H)^3} < 0\ \forall H \in]0,1]$ *and* $M(4000 + 500 \cdot H)''_{2,3} = \frac{-462000}{(5500 \cdot H)^3} < 0\ \forall H \in]0,1]$, so the curve is concave because the second derivative is negative on the given interval. *Hence* $M(7500 - 625 \cdot H)''_{3,4} = \frac{131250 \cdot H}{(3750 \cdot H)^3} > 0\ \forall H \in]0,1]$ *and* $M(7500 + 625 \cdot H)''_{3,4} = \frac{131250 \cdot H}{(5000 \cdot H)^3} > 0\ \forall H \in]0,1]$, so the curve is convex because the second derivative is positive on the given interval. For $M(10000 - 625 \cdot H)''_{4,5} = \frac{-37500 \cdot H}{(3125 \cdot H)^3} < 0\ \forall H \in]0,1]$ *and*

$M(10000 + 625 \cdot H)''_{4,5} = \frac{-37500 \cdot H}{(1875 \cdot H)^3} < 0 \; \forall H \in]0, 1]$, so the curve is concave because the second derivative is negative on the given interval. Therefore the existence is proved. □

So the classical trapeze fuzzy membership functions provide three different ratings such as line, convex and concave curves on transient phases. Therefore other kinds of membership functions should be used.

3. Trapeze fuzzy membership function with trigonometrical legs

In this section it will be given the general equation for trapeze fuzzy membership function with trigonometrical legs, it is shown in equation (3.1).

Definition 3.1. The next equation is called trapeze fuzzy membership function with trigonometrical legs:

$$\mu(x)_{trig} = \begin{cases} \frac{1}{2} + \frac{1}{2} \cdot sin(a \cdot x + b) & \text{if } x_1 \leq x < x_2, \quad a = \frac{\pi}{x_1 - x_2}, \quad b = \pi \cdot \left(\frac{3}{2} - \frac{x_1}{x_1 - x_2}\right) \\ 1 & \text{if } x_2 \leq x \leq x_3 \\ \frac{1}{2} + \frac{1}{2} \cdot cos(a \cdot x + b) & \text{if } x_3 < x \leq x_4, \quad a = \frac{-\pi}{x_3 - x_4}, \quad b = \frac{\pi \cdot x_3}{x_3 - x_4} \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

Here the x_1, x_2, x_3 and x_4 are real points on the abscissa ordered increasingly which have next values:

$$\mu(x_1) = 0, \; \mu(x_2) = 1, \; \mu(x_3) = 1 \text{ and } \mu(x_4) = 0. \tag{3.2}$$

The short sign for it is $\mu(x)_{trig} = [x_1, x_2, x_3, x_4]_{trig}$.

In every case the set-transition is chosen as symmetric one. The first and the last fuzzy membership functions were chosen as half of set-transition H — it is shown e.g. on Fig. 5. and Fig. 6, because of the principle of symmetric division is close enough to real application.

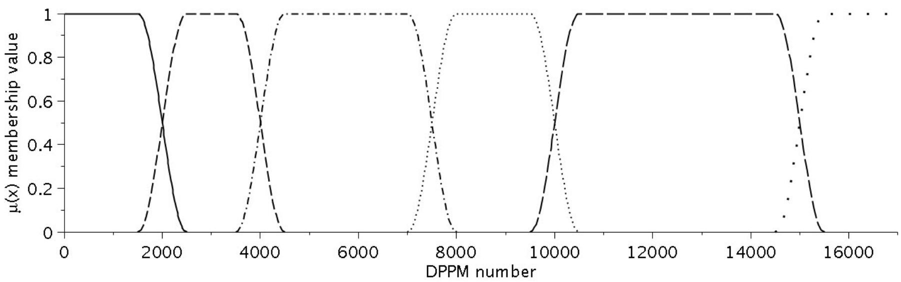


Figure 5. The fuzzy membership functions at $H = 2000$ by trigonometrical legs.

3.1. Constant set-transition in rating-points

In this part next set-transition are chosen $H = 200$ on Fig. 7 and Fig. 8., $H = 400$ on Fig. 9 and Fig. 10., $H = 800$ on Fig. 11 and Fig. 12., $H = 2000$ on Fig. 5 and Fig. 6..

In this case the set-transition is the same for all fuzzy membership function. Therefore set-sizes are not to taken into consideration. Let it be an example here to fuzzyfy in Table 1. the second group: if the set-transition $H = 2000$ $DPPM$ then e.g. on Fig. 5 and Fig. 6. for the second membership function the H should be divided into two part so it will be symmetric. For the left side it is $x_1 = 1500$ $DPPM$, $x_2 = 2500$ $DPPM$, $x_3 = 3500$ $DPPM$ and $x_4 = 4500$ $DPPM$ so it is provided the second fuzzy membership function according to equation (3.1). Of course to have the rating-points to the given number DPPM it should be taken the equation (2.2).

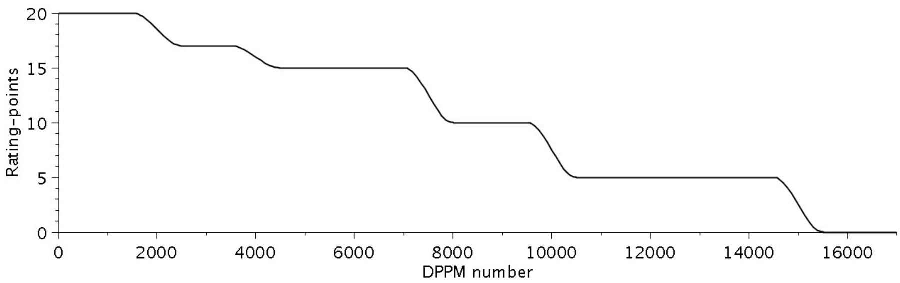


Figure 6. The Rating-point at $H = 2000$.

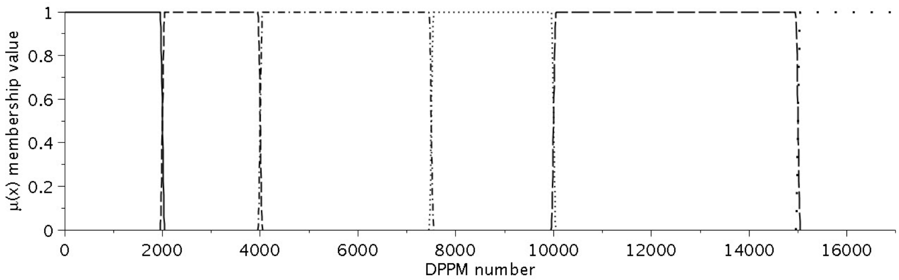


Figure 7. The fuzzy membership functions at $H = 200$ by trigonometrical legs.

3.2. *Proportion (rate) set-transition in rating-points*

In this part the set-transition takes into consideration the set-size. The next set-transition were chosen by authors: $H = 10$ % on Fig. 13. and Fig. 14., $H = 20$ % on Fig. 15. and Fig. 16., $H = 100$ % on Fig. 19. and Fig. 20. Let it be again an example. Take the third group from the Table 1. and $H = 20\%$ it is shown on Fig. 15 and Fig. 16. the third membership function. So the set-size takes 3500 $DPPM$ therefore it is $H = 0.2 \cdot 3500 = 700$. After this point the method is according to the previous section, so the fuzzy membership function is given with points $x_1 = 3825$ $DPPM$, $x_2 = 4175$ $DPPM$, $x_3 = 7325$ $DPPM$ and $x_4 = 7675$ $DPPM$ according to equation (3.1). Of course the rating-point is evaluated with equation (2.2).

In Table 4, Table 5 and Table 6 it can be seen, there is always an inflexion point according to numerical computation on the given interval and it is always turned from concave to convex as the signum column shows. The proof that there exists inflexion point on every transient phases with changing from concave to convex using proportional set-transition by trigonometrical legs computing with two legs only, is to

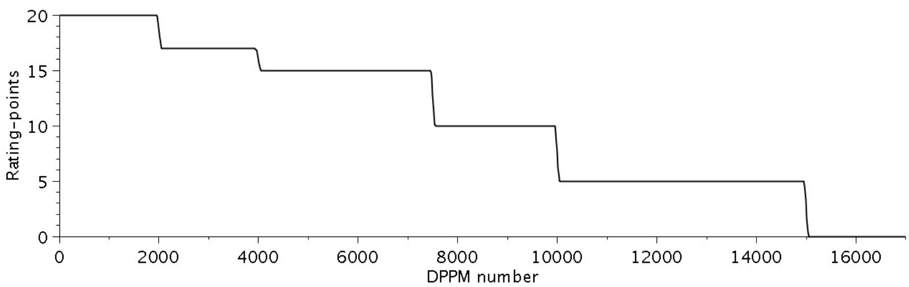


Figure 8. The Rating-point at $H = 200$.

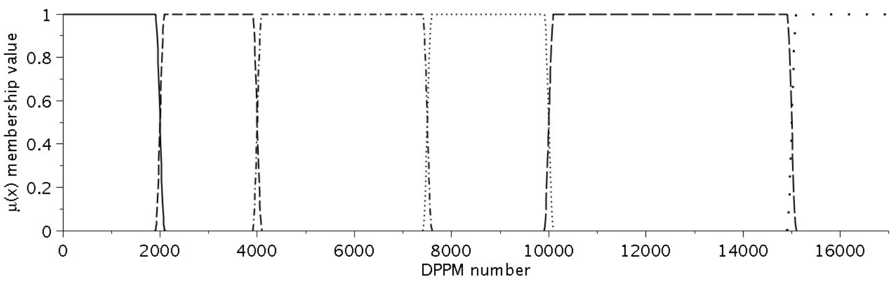


Figure 9. The fuzzy membership functions at $H = 400$ by trigonometrical legs.

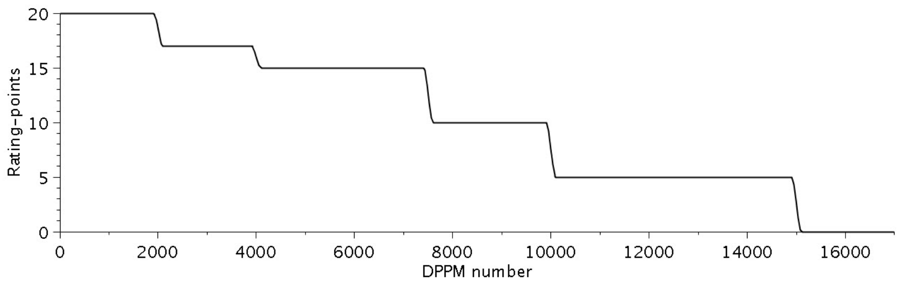


Figure 10. The Rating-point at $H = 400$.

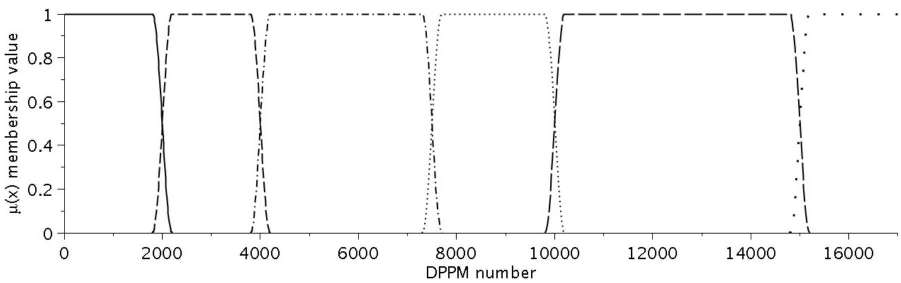


Figure 11. The fuzzy membership functions at $H = 800$ by trigonometrical legs.

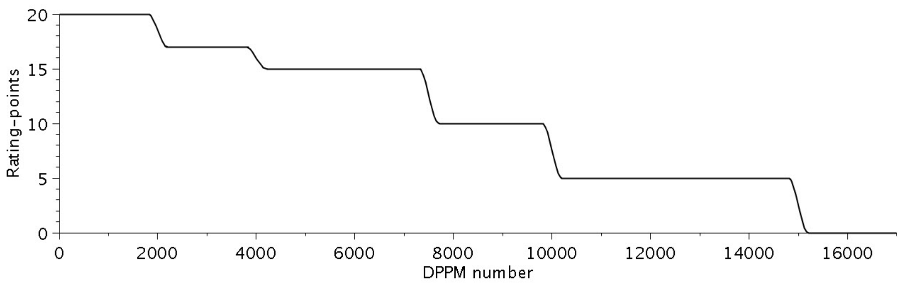


Figure 12. The Rating-point at $H = 800$.

be difficult because the formula is too complicate therefore it is investigated numerically. Generally the

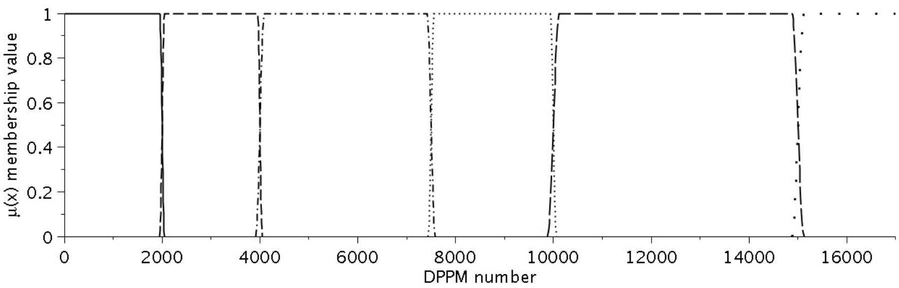


Figure 13. The fuzzy membership functions at $H = 10\%$ by trigonometrical legs.

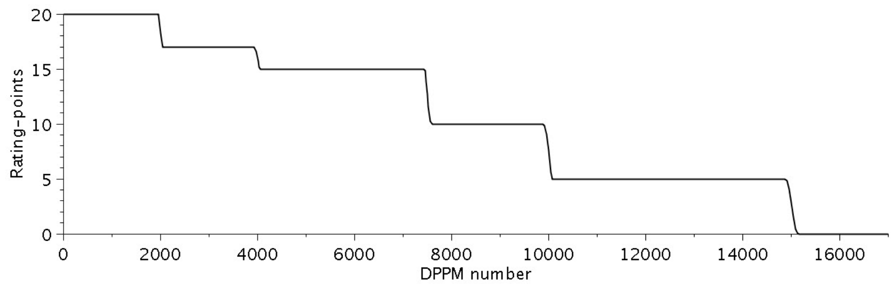


Figure 14. The Rating-point at $H = 10\%$.

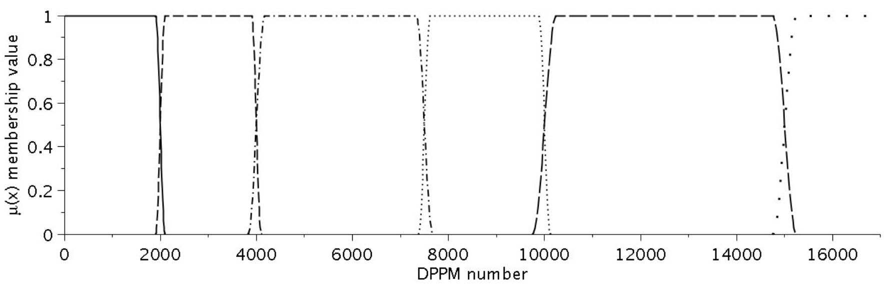


Figure 15. The fuzzy membership functions at $H = 20\%$ by trigonometrical legs.

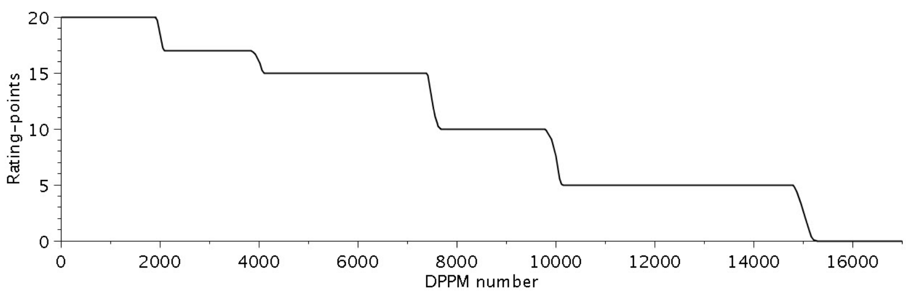


Figure 16. The Rating-point at $H = 20\%$.

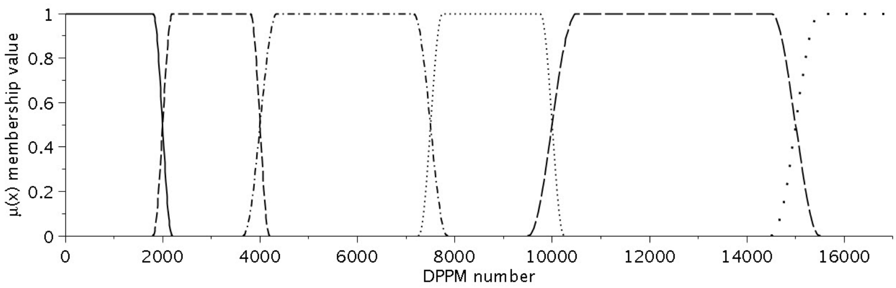


Figure 17. The fuzzy membership functions at $H = 40\%$ by trigonometrical legs.

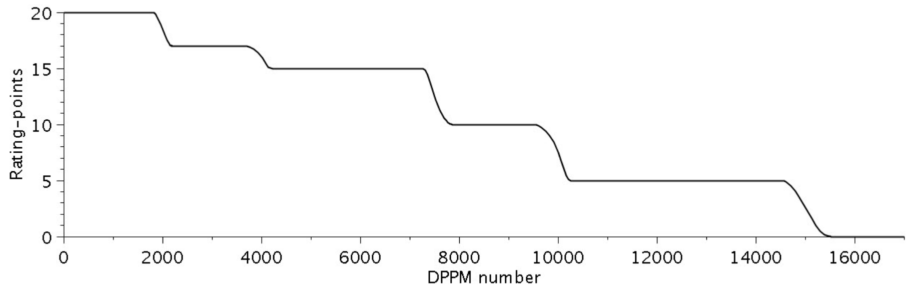


Figure 18. The Rating-point at $H = 40\%$.

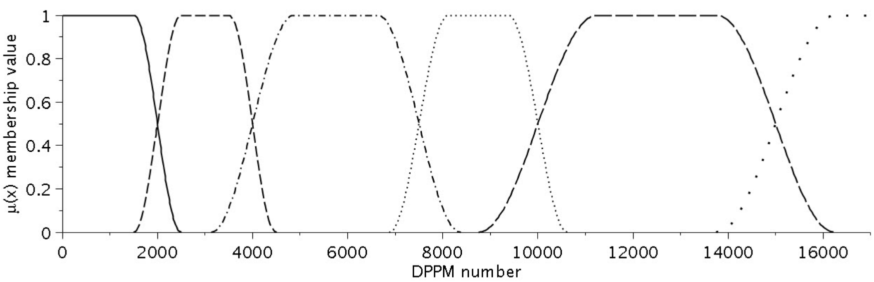


Figure 19. The fuzzy membership functions at $H = 100\%$ by trigonometrical legs.

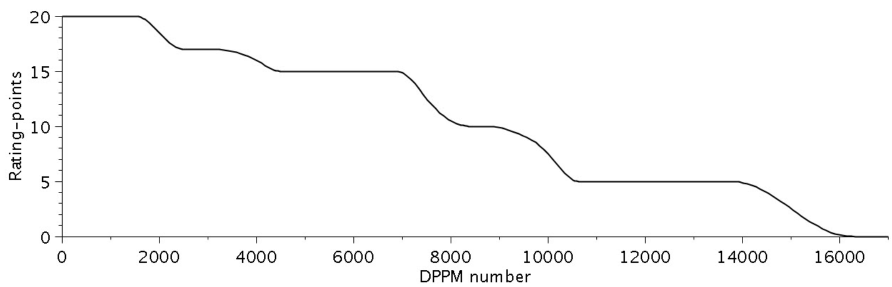


Figure 20. The Rating-point at $H = 100\%$.

equation is on transient phase according to equation (2.2):

$$M_{i,i+1}(x) = \frac{p_i + p_i \cos\left(\frac{\pi(x-c_i)}{c_i-d_i}\right) + p_{i+1} + p_{i+1} \sin\left(\frac{1}{2}\frac{\pi(2x+a_{i+1}-3b_{i+1})}{a_{i+1}-b_{i+1}}\right)}{2 + \cos\left(\frac{\pi(x-c_i)}{c_i-d_i}\right) + \sin\left(\frac{1}{2}\frac{\pi(2x+a_{i+1}-3b_{i+1})}{a_{i+1}-b_{i+1}}\right)} \tag{3.3}$$

if $x \in [c_i, d_i] \cap [a_{i+1}, b_{i+1}]$, $i \in \{1,2,3,4,5\}$ and the fuzzy membership functions are given with $[a_i, b_i, c_i, d_i]_{trig}$ and $[a_{i+1}, b_{i+1}, c_{i+1}, d_{i+1}]_{trig}$. The inflexion point is always $x = 2000$ DPPM for $M_{1,2}(x)$ by different $H \in]0,1]$ and for $M_{5,6}(x)$ the inflexion point is always at $x = 15000$ DPPM by different $H \in]0,1]$ on the given interval. The other results are shown in Table 4 and Table 5. The Table 4, the Table 5 and the Table 6 contain the next columns: first column is the set-transition H, the second one is inflexion point for $M_{i,i+1}$ for $i \in \{2,3,4\}$, the third one is the Signum column which contain the sign changes of area of inflexion point and the forth one is the given interval for each H. All calculation was made with Maple 14^{th} on a DELL Vostro laptop with 4 cores and 4 GB memory using Windows 7 operation system. To evaluate one rating point it was not measurable the computation time. For example to compute the Table 6 it takes 4.898 seconds.

Table 4
Numerical computation for inflexion point by different $H \in]0,1]$ for $M_{2,3}$

H	$M_{2,3}$ Inf.p.	Signum	Interval
0.05	4006.842812	[-1,1]	[3975.0, 4025.0]
0.1	4013.685624	[-1,1]	[3950.0, 4050.0]
0.15	4020.528438	[-1,1]	[3925.0, 4075.0]
0.2	4027.371250	[-1,1]	[3900.0, 4100.0]
0.25	4034.214061	[-1,1]	[3875.0, 4125.0]
0.3	4041.056878	[-1,1]	[3850.0, 4150.0]
0.35	4047.899688	[-1,1]	[3825.0, 4175.0]
0.4	4054.742502	[-1,1]	[3800.0, 4200.0]
0.45	4061.585315	[-1,1]	[3775.0, 4225.0]
0.5	4068.428126	[-1,1]	[3750.0, 4250.0]
0.55	4075.270940	[-1,1]	[3725.0, 4275.0]
0.6	4082.113750	[-1,1]	[3700.0, 4300.0]
0.65	4088.956563	[-1,1]	[3675.0, 4325.0]
0.7	4095.799377	[-1,1]	[3650.0, 4350.0]
0.75	4102.642190	[-1,1]	[3625.0, 4375.0]
0.8	4109.485001	[-1,1]	[3600.0, 4400.0]
0.85	4116.327815	[-1,1]	[3575.0, 4425.0]
0.9	4123.170628	[-1,1]	[3550.0, 4450.0]
0.95	4130.013442	[-1,1]	[3525.0, 4475.0]
1	4136.856253	[-1,1]	[3500,4500]

Table 5

Numerical computation for inflexion point by different $H \in]0,1]$ for $M_{3,4}$

H	$M_{3,4}$ Inf.p.	Signum	Interval
0.05	7493.944539	[-1,1]	[7468.75, 7531.25]
0.1	7487.889079	[-1,1]	[7437.5, 7562.5]
0.15	7481.833616	[-1,1]	[7406.25, 7593.75]
0.2	7475.778155	[-1,1]	[7375.0, 7625.0]
0.25	7469.722696	[-1,1]	[7343.75, 7656.25]
0.3	7463.667234	[-1,1]	[7312.5, 7687.5]
0.35	7457.611773	[-1,1]	[7281.25, 7718.75]
0.4	7451.556308	[-1,1]	[7250.0, 7750.0]
0.45	7445.500850	[-1,1]	[7218.75, 7781.25]
0.5	7439.445389	[-1,1]	[7187.5, 7812.5]
0.55	7433.389928	[-1,1]	[7156.25, 7843.75]
0.6	7427.334467	[-1,1]	[7125.0, 7875.0]
0.65	7421.279003	[-1,1]	[7093.75, 7906.25]
0.7	7415.223545	[-1,1]	[7062.5, 7937.5]
0.75	7409.168084	[-1,1]	[7031.25, 7968.75]
0.8	7403.112621	[-1,1]	[7000.0, 8000.0]
0.85	7397.057160	[-1,1]	[6968.75, 8031.25]
0.9	7391.001700	[-1,1]	[6937.5, 8062.5]
0.95	7384.946237	[-1,1]	[6906.25, 8093.75]
1	7378.890776	[-1,1]	[6875,8125]

Table 6
Numerical computation for inflexion point by different $H \in]0, 1]$ for $M_{4,5}$

H	$M_{4,5}$ Inf.p.	Signum	Interval
0.05	10009.65208	[-1,1]	[9968.75, 10031.25]
0.1	10019.30417	[-1,1]	[9937.5, 10062.5]
0.15	10028.95624	[-1,1]	[9906.25, 10093.75]
0.2	10038.60833	[-1,1]	[9875.0, 10125.0]
0.25	10048.26041	[-1,1]	[9843.75, 10156.25]
0.3	10057.91249	[-1,1]	[9812.5, 10187.5]
0.35	10067.56458	[-1,1]	[9781.25, 10218.75]
0.4	10077.21666	[-1,1]	[9750.0, 10250.0]
0.45	10086.86874	[-1,1]	[9718.75, 10281.25]
0.5	10096.52083	[-1,1]	[9687.5, 10312.5]
0.55	10106.17291	[-1,1]	[9656.25, 10343.75]
0.6	10115.82499	[-1,1]	[9625.0, 10375.0]
0.65	10125.47707	[-1,1]	[9593.75, 10406.25]
0.7	10135.12916	[-1,1]	[9562.5, 10437.5]
0.75	10144.78123	[-1,1]	[9531.25, 10468.75]
0.8	10154.43332	[-1,1]	[9500.0, 10500.0]
0.85	10164.08541	[-1,1]	[9468.75, 10531.25]
0.9	10173.73749	[-1,1]	[9437.5, 10562.5]
0.95	10183.38957	[-1,1]	[9406.25, 10593.75]
1	10193.04165	[-1,1]	[9375, 10625]

4. Application of methods

In this section authors provide an example therefore next DPPM numbers are chosen 7500, 7501 and 7510. The evaluation results are in the Table 7. To take into consideration, the proportional set-transition should be always better if the set-sizes are different. Of course to choose the correct DPPM number, one should take opinion and experiences of experts or managers. This method and principal can be applied for any other cases where the structure allowed that.

On the other hand, in the paper (Portik *et al.*, 2011) the authors examined and presented the differences among the conventional, constant and proportional set-transitions via three concrete samples in a transient phase in order to verify the possibility to get a method, which provide more similar results like human thinking. Three samples were chosen, value at 7500 PPM, 7501 PPM and 7510 PPM. On the classical way, the outcome were 15 points at 7500 DPPM and 10 points at 7501 and 7510 DPPM in the evaluation, while the constant and proportional set-transitions calculations result effected 12,5 points at 7500 DPPM in the second two cases, which means adjustment comparing the classical way.

In this paper the authors have done further evaluation in order to provide further adjustments, with changing set-transition. The rate of the change in the output value is more harmonized with the rate of the change in the input parameters as presented in Table 7. Moreover, in case of 7501 DPPM the supplier gets less then 33.33 % by classical method despite of increasing of failure part was only one piece from one million parts. The result was 12.47 points at 7501 DPPM for $H = 20$ % in Table 7, the rate of change was 0.24 % compared to 12.5 rating points despite of 33.33 % when the DPPM changed only with 0.0001 % . The calculation is 12.23 points at 7510 DPPM for $H = 20$ %, the rate of change was 2.16 %. The authors have achieved their aim namely to have a good rate on transient phases.

Table 7
Results of application of methods

	Point belongs to the given number of DPPM		
Methods	7500 *DPPM*	7501 *DPPM*	7510 *DPPM*
Conventional	15	10	10
Constant set-transition			
$H = 200$ *DPPM*	12.5	12.42	11.72
$H = 400$ *DPPM*	12.5	12.46	12.11
$H = 800$ *DPPM*	12.5	12.48	12.3
$H = 2000$ *DPPM*	12.5	12.49	12.42
Proportional set-transition			
$H = 10$ %	12.5	12.44	11.98
$H = 20$ %	12.5	12.47	12.23
$H = 40$ %	12.5	12.48	12.36
$H = 100$ %	12.5	12.49	12.44

5. Conclusion

In this paper the authors provided new results which fulfill the principle of good rating on transient phases. In earlier results of (Portik *et al.*, 2011) it was not good rating on transient phases because on some transient phases there are somewhere convex and somewhere else concave, it is shown on Fig. 4 and Fig. 3. Also in this paper, the authors have achieved the next result: to have a good rating on all transient phases which means the convex and concave transient phases are disappeared. Also it is a good principle for rating suppliers by trapeze fuzzy membership function with trigonometrical legs and with proportional set-transition.

The authors emphasize, the results of this paper can be used in every field, which follows the group-based classification with crisp boundaries for the evaluation. This means wide range of applications in quality assurance, risk assessment, audit results evaluation etc.

One of the authors working in the industry as a quality engineer for years. Previously in electronic, now in automotive industry. The original problem was raised up by him. After the authors search in the scientific literature in order to find any solution to develop this issue, they have started to work out a method to improve. The calculation with sample values illustrated the better distribution of values at the group borders thanks to the new calculation concept. Due to this, the change of the output will be more similar to the change of the input providing a more realistic evaluation. The authors' theoretical and industrial experiences confirm the results.

The authors' plan for the future is to examine other applications to get a comprehensive overview of the efficiency of the method shown in this paper.

References

Esse, B. (2008). Aspect of the decision of supplier-selection. In: *Study Workshop, University of Corvinus, Institute of Business Economics,* Volume 90,Budapest, Hungary (in Hungarian).

Humphreys, P. K., W. L. Li and L.Y. Chan (2007). The impact of specific supplier development efforts on buyer competitive advantage: an empirical model. *International Journal of Production Economics,* Volume 106, Issue 1, 2007/3.

Krause, D. R. and L. M. Ellram (1997). Critical elements of supplier development. *European Journal of Purchasing & Supply Management,* Volume 3, Issue 1, 1997/3 pp. 21–31.

Liu, J., L. Martínez, H. Wang, R. M. Rodríguez and V. Novozhilov (2010). Computing with words in risk assessment. *International Journal of Computational Intelligence Systems* **3**(4 (October 2010)), 396–419.

Pokorádi, L. (2002). Fuzzy logic-based risk assessment. *AARMS, Academic and Applied Research in Military Science* **1**(1), 63–73.

Pokorádi, L. (2008). *Systems and Processes Modeling*. Campus Kiadó, Debrecen(in Hungarian).

Pokorádi, L. (2009). Risk assessment based upon fuzzy set theory. In: 15^{th} *Building Services, Mechanical and Building Industry Days, Debrecen, Hungary*. pp. 311–318.

Portik, T., T. Varga and L. Pokorádi (2011). Development of supplier-rating based on fuzzy set theory. In: *Annals of MTEM for 2011 & Proceeding of the* 10^{th} *International Conference Modern Technologies in Manufacturing* $6^{th}-8^{th}$ *October 2011, Cluj-Napoca, Romania*. pp. 279–282.

Retter, Gy. (2006). *Fuzzy, neural-genetics, chaotic systems*. Akadémiai Kiadó, Budapest, Hungay (in Hungarian).

Ross, T. J. (2010). *Fuzzy logic with engineering applications*. John Wiley & Sons Ltd., USA.

Theory and Applications of
Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 71–88

Design and Scheduling of Chemical Batch Processes: Generalizing a Deterministic to a Stochastic Model

João Luís de Miranda[a,*], Miguel Casquilho[b]

[a]*Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Portalegre: Lugar da Abadessa, Apt. 148; 7301-999 Portalegre; Portugal.*
[b]*Instituto Superior Técnico, Universidade Técnica de Lisboa (IST/UTL) Ave. Rovisco Pais, IST; 1049-001 Lisboa; Portugal, Centre for Chemical Processes at IST/UTL.*

Abstract

A stochastic optimization model for the design and scheduling of batch chemical processes is developed in a Two-Stage Stochastic Programming framework, with the uncertainty formulated through a number of discrete scenarios. The sparse model presents binary variables in the first stage and systematically generalizes a deterministic model chosen from the literature, in an approach based on computational complexity. The combination of single product campaign (SPC) with multiple machines was found to be the most promising from a computational standpoint, and it is here generalized toward a stochastic environment within the relaxation of the soft demand constraints. Numerical examples are presented, and the results point to a significant reduction of 8-20 % of the investment costs in comparison to the SPC non-relaxed case, without real losses if the multiple product campaign (MPC) policy is adopted.

Keywords: robust optimization, two-stage stochastic programming, design, scheduling, batch processes.
2000 MSC: 90B36.

---

1. Introduction

The number of industrial cases published on the design and scheduling of batch processes is very small, and the industrial works published in the open literature are also illustrative of the difficulties to conjugate different time ranges with efficiency and detail, as referred in the reviews of (Floudas & Lin, 2004) and (Barbosa-Póvoa, 2007).

This paper addresses the design of batch chemical processes and simultaneously considers the scheduling of operations. A generalization of literature models (Voudouris & Grossmann, 1992) is proposed, from a deterministic Mixed Integer Linear Programming (MILP) model to a Two-Stage Stochastic Programming (2SSP) one. The generalization to a stochastic, multiperiod, and robust model is based on computational complexity studies (Miranda, 2011*a*). When reducing the multiperiod model into one single time period, the optimality study shows that significant reduction of investment costs is possible.

*Corresponding author
Email addresses: `jlmiranda@estgp.pt` (João Luís de Miranda), `mcasquilho@ist.utl.pt` (Miguel Casquilho)

The generalized model treats different time ranges, namely, the investment and scheduling horizons. Furthermore, the 2SSP framework allows robustness promotion: i) in the solution, by penalizing the deviations; and ii) in the model, with relaxation of the integrality constraints in the second phase variables.

In the design of batch processes for multiproduct units (flowshop batch plant), the implementation of difficult MILP/MINLP models is common: it is required to design and enumerate the equipment, simple or distributed in parallel, or to consider single product (SPC) or multiple products (MPC) campaigns. This usually corresponds to the first phase of the 2SSP framework. (The word ”phase” is used here instead of ”stage” because of the traditional use of ”stage” in chemical engineering scheduling problems).

In order to better select the equipments, the optimal production policy must also be found since it directly affects the equipment sizing. However, it involves the detailed solution of scheduling subproblems where decomposition schemes are pertinent. These subproblems are focused in the second phase of 2SSP, where the control variables (recourse) occur. The integer and binary variables related to the scheduling and precedence constraints are disregarded as control variables, as they would make very hard the treatment of the recourse problem. Consequently, the second phase variables are assumed continuous (for example, the number of batches) and binary variables occur only in the 2SSP first phase.

The generalization approach is not like the one proposed by (Moreno *et al.*, 2007), who treat a similar problem of design and scheduling of batch processes. This approach coincides with Ahmed and Sahinidis (Ahmed & Sahinidis, 2000), (Ahmed & Sahinidis, 2003) and (Liu & Sahinidis, 1997) in the sense that analytical studies of computational complexity can yield significant improvements in terms of algorithms and problem structures, with good and realistic solutions.

In this study, the generalization to a 2SSP framework within a multiperiod and robustness environment points to a significant reduction (8%-20 %) in the investment costs. And the results motivate to pursue this generalization approach, with foreseen developments.

The paper sequentially addresses: in Section 2, the design and scheduling of batch processes, and the deterministic model that is enlarged here; in Section 3, the presentation of the generalized model on a stochastic and robust 2SSP framework; in Section 4, the optimality study, which is directed to promote robustness; finally, the main conclusions are presented in Section 5.

2. The Design and Scheduling of Batch Processes

In this section, the issue of scheduling of batch processes is integrated with process design (sizing), combining the short term decisions with the long term investment planning. The models that were systematically studied in (Miranda, 2007) are from the open literature and the state-of-art at that time is described.

The study of existing models in the literature induces the enlargement of models and related applications (Miranda, 2007), and this generalization of models simultaneously causes increasing complexity and difficulties. A design and scheduling, deterministic, and single time period model (Voudouris & Grossmann, 1992) that seems to have no improvements for more than a decade is addressed in this paper.

The models studied were those in (Voudouris & Grossmann, 1992), and appear to belong to a research line initiated in (Birewar & Grossmann, 1989) , featuring a quick resolution, and finished in (Voudouris & Grossmann, 1993), realizing the approach impracticability. The approach option was changed to the jobshop framework in (Voudouris & Grossmann, 1996), but flowshop continues to be widely used in chemical industry.

The options set and the successive generalizations are the main criteria to the models selection: i) single machine vs. multiple parallel machines in each stage; ii) single product campaigns (SPC) or multiple products (MPC); and iii) assumption of some storage policy or zero wait (ZW) operations. The complexity of the problems leads to the adoption of alternative methodologies, such as the evolutionary procedures of Xia and Macchietto in (Xia & Macchietto, 1997) and Tan and Mah in (Tan & Mah, 1998). Pekny and Miller in (Pekny & Miller, 1991) recommended the utilization of heuristics whenever a satisfactory method is not available for the problem at hand. The latter authors treated the scheduling flowshop problem with a ZW policy, both exactly, through a branch-and-bound algorithm, and heuristically, through a permutations procedure and simulated annealing. Jayaraman et al. in (Jayaraman *et al.*, 2000) addressed the design and scheduling of batch processes using a heuristic method (ant colony) and, they obtained exact results for short sized instances. Cavin et al. in (Cavin *et al.*, 2004) used tabu search to address the combinatorial and multiobjective optimization in the design and scheduling of a multipurpose batch plant problem.

A different path was adopted by Liu and Sahinidis in (Liu & Sahinidis, 1997) and Ahmed and Sahinidis in (Ahmed & Sahinidis, 2000) and (Ahmed & Sahinidis, 2003) when assessing the possibility to develop exact and efficient algorithms. They developed analytical investigations to verify that their problems are NP-hard, both the static version and the dynamic version of their planning process models. Using computational complexity techniques, they verified that it is not possible to develop exact polynomial algorithms and pointed the need to systematically build good procedures.

Moreno et al. in (Moreno *et al.*, 2007) addressed multiproduct batch plant and considered SPC policy. They presented a multiperiod model aiming at the sizing and planning in batch multiproduct plants, considering: return values, and operation and investment costs; assignment of discrete dimensions to the processes, batch or semicontinuous; implementation of intermediate storage; and variations on demands and component prices, due both to seasonal and structural effects. The approach presented here is different from the one of (Moreno *et al.*, 2007), but the kind of problem focused is similar.

This work uses analytical results and applies computational complexity techniques at the deterministic model *MS* (Miranda, 2011*a*), which feature multiple machines per stage and SPC (Multiple machine, SPC). This model, was selected because (Miranda, 2011*b*):

- for industrial applications with realistic number and quantities of products, the option of the multiple processes in parallel at each stage should be considered: otherwise unfeasibility will certainly occur;
- the option for the SPC mode arises from the current difficulties to apply MPC in a *multiple machine* environment, just due to insufficiency of the related model;
- the investment cost assuming SPC is estimated to exceed in near 5 % the cost of the more efficient MPC policy; this surplus results from a selection of the next discrete dimension on nearly half of the stages (values derived from comparable instances); that is, the SPC sizing is *a priori* overdesigned, and this will permit to introduce new products, or even to accommodate un-forecast growth on product demands.

3. Development of a Generalized Robust Model

In this section, a robust model is presented for the design and scheduling of batch processes, generalizing the deterministic model *MS* to a stochastic context in two phases (2SSP), with promotion of robustness. This permits the treatment of the risk associated to medium and long term investments.

When a high investment is needed, usually, a long return term is associated. The generalized model includes the optimization of long term investment and also considers the short term scheduling of batch processes. Deterministic models do not conveniently address the risk of a wider planning horizon, and scheduling models often deal with certain data in a single time horizon. Thus, difficulty increases when the combinatorial scheduling problem is integrated with the uncertainty of the design problem.

The objective of the 2SSP model is the maximization of a robust measure of the Net Present Value (NPV), by selecting the discrete dimensions for the batch processes and the number of processes operating in parallel (out-of-phase operation). In addition, multiple time periods are supposed in the NPV evaluation. Given the uncertainty of the quantities and unit returns of each product, then returns are evaluated in a probabilistic way.

The stochastic model aims to maximize the robust NPV, Φ , considering the expected return minus the investment costs, with the latter occurring only in the first period. Robustness is promoted by penalizing the expected values or estimators of: *i)* the variability of the discrete scenarios solutions, *dvtn*; *ii)* the non-satisfied product demands, *Qns*; and *iii)* the capacity slacks, *slk*. That is:

$$\begin{aligned}[\max]\,\Phi = &\sum_{r=1}^{NR} prob_r\,\xi_r - \lambda dsv \sum_{r=1}^{NR} prob_r\, dvtn_r - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT}\left(\sum_{j=1}^{NC}\sum_{t=1}^{NT} Qns_{jtr}\right) \\ &-\lambda slk \sum_{r=1}^{NR} \frac{prob_r}{M.NC.NT}\left(\sum_{i=1}^{M}\sum_{j=1}^{NC}\sum_{t=1}^{NT} slk_{ijtr}\right).\end{aligned} \tag{3.1}$$

The objective function uses technical estimators (Appendix A) that are built to assess the quality of the generalized model. Each probabilistic component, ξ_r, corresponds to the NPV obtained at each discrete scenario *r*, and this component is obtained from: *i)* the present amount of sales return, obtained in the second phase of the 2SSP (probabilistic net values, related to materials purchases and operations costs); minus *ii)* the investment costs, defined in the first phase of the 2SSP (deterministic and discrete costs, accordingly with the discrete dimensions of equipments at each stage). This means:

$$\xi_r = \sum_{j=1}^{NC}\sum_{t=1}^{NT} ret_{jtr} W_{jtr} - \sum_{i=1}^{M}\sum_{s=1}^{NS(i)}\sum_{p=1}^{NP(i)} c_{isp} y_{isp} \ , \quad \forall r. \tag{3.2}$$

A linear measure is used for penalizing variability. The negative deviation of NPV, $dvtn_r$, is evaluated at each discrete scenario to measure only the deviations below the NPV's expected value, and provided that each probabilistic component, ξ_r, is computed from relations 3.2,

$$dvtn_r \geq \sum_{r'=1}^{NR}(prob_{r'}\xi_{r'}) - \xi_r \ \geq 0\ , \quad \forall r. \tag{3.3}$$

Assuming ZW policy and the non-existence of intermediate storage, the non-satisfied demand of each product *j*, Qns_{jtr}, is defined at each time period, *t*, and for each scenario, *r*, by the related definition constraint slack. That is, by the difference between the quantities of product demands, Q_{jtr}, and quantities produced, W_{jtr}. The following constraints sets are employed:

$$W_{jtr} + Qns_{jtr} = Q_{jtr}\ , \quad \forall j, t, r. \tag{3.4}$$

$$Qns_{jtr} \geq 0, \quad \forall j, t, r. \tag{3.5}$$

The global quantities produced for each product, W_{jtr}, at each discrete scenario and each time period are related to the aggregated batches number, nc_{ijsptr},

$$S_{ij} W_{jtr} \leq \sum_{s=1}^{NS(j)}\sum_{p=1}^{NP(j)} dv_{is} nc_{ijsptr}, \quad \forall i, j, t, r. \tag{3.6}$$

Thus, the global excess on the implemented production capacities (slk_{ijtr}) results directly from

$$S_{ij}W_{jtr} + slk_{ijtr} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is} nc_{ijsptr}, \quad \forall i, j, t, r. \tag{3.7}$$

The disaggregated number of batches, nc_{ijsptr} (further details in Appendix A), corresponds to the product-aggregation of variables (n_{jtr} . y_{isp}). Three logical sets of constraints are required: upper bounds; only one value is selected; and the definition of the selected value. Respectively:

$$nc_{ijsptr} - nc_{ijsp}^{Upp} y_{isp} \leq 0 \,, \ \forall\, i, j, s, p, t, r, \tag{3.8}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} y_{isp} = 1, \quad \forall i, \tag{3.9}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} nc_{ijsptr} - n_{jtr} = 0, \quad \forall i, j, t, r. \tag{3.10}$$

The campaign times must be determined, $tcamp_{jtr}$, either in relation to the disaggregated number of batches, nc_{ijsptr}, and by satisfying the time horizon, H:

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} \left(\frac{\tau_{ij}}{p(i)} nc_{ijsptr} \right) - tcamp_{jtr} \ \leq 0, \forall i, j, t, r, \tag{3.11}$$

$$\sum_{j=1}^{NC} tcamp_{jtr} \leq H, \quad \forall t, r. \tag{3.12}$$

Conjugating the described relations, the stochastic model *spbatch_ms* for the design and scheduling of batch chemical processes aims at the maximization of NPV, promotes robustness in solution and model, and it assumes *flowshop* configuration, several processes in parallel at each stage (*multiple machine*), single product campaigns (SPC), and ZW policy:

Model *spbatch_ms*:

$$\begin{aligned}[\max]\,\Phi = \sum_{r=1}^{NR} prob_r \xi_r - \lambda dsv \sum_{r=1}^{NR} prob_r.dvtn_r - \lambda qns \sum_{r=1}^{NR} \tfrac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right) \\ -\lambda slk \sum_{r=1}^{NR} \tfrac{prob_r}{M.NC.NT} \left(\sum_{i=1}^{M} \sum_{j=1}^{NC} \sum_{t=1}^{NT} slk_{ijtr} \right),\end{aligned} \tag{3.13}$$

subject to,

Relations 3.2 to 3.4
Relations 3.7 to 3.12

$$\xi_r, dvtn_r, \ slk_{ijtr}, nc_{ijsptr}, \ n_{jtr}\,, \ Qns_{jtr}, \ tcamp_{jtr}, W_{jtr} \geq 0, \quad \forall i, j, s, p, t, r \tag{3.14}$$

$$y_{isp} \in \{0;1\}, \quad \forall i, s, p. \tag{3.15}$$

The application of this generalized model for the design and scheduling of batch processes is illustrated through numerical examples: the design of batch processes is satisfying uncertain demands on a unique time period (”static”), and the minimization of investment costs considers a stochastic and robust formulation. Appendix B contains further details of the development of the generalized 2SSP model, while Appendix C describes how the 2SSP model relates to the deterministic model *MS*.

4. Illustrative Examples

The problem of sizing batch chemical processes is now discussed to illustrate the characteristics of the stochastic model *spbatch_ms*. The former deterministic model, *MS* (theoretically focused in Appendix C), is targeted as reference model. Through the usual reasoning of polynomial reduction of problem instances, the following is assumed: ***i)*** only one time period; and ***ii)*** zero value of return in products.

Given the single time period considered, the unsuitability of NPV maximization must be noticed: NPV is usually addressed in a multiperiod horizon (dynamic problem) due to high investment costs that do not allow payback on the first time period. Then, assuming $ret_{jtr} = 0$, means not to account for cash flows returning, and the NPV variables ξ are representing only the investment costs.

$$\xi = \xi_r = -\sum_{i=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp}, \quad \forall r. \tag{4.1}$$

The NPV variables ξ are scenario independent, so no variability is presented, and a null deviation, $dvtn_r = 0$, will be observed in all scenarios. One mainly wishes to satisfy the uncertain product demands, and the penalization of capacity slack will not be considered ($slk = 0$). The objective function in equation 3.1 is thus reduced to the robust minimization of investment costs, assuming only the penalization of non-satisfied demand:

$$\begin{aligned} [\max]\, \Phi &= \xi \,.\, \sum_{r=1}^{NR} prob_r - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right) \\ &= -\sum_{i=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp} - \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right), \end{aligned} \tag{4.2}$$

or,

$$[\min]\, \Psi = \sum_{j=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp} + \lambda qns \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right). \tag{4.3}$$

The model defined in the relations set 3.13 is restricted, and the following can be neglected: the time period index, *t*, because only one time period is considered; the constraint sets concerning the definition of the probabilistic variables, ξ , which will have a constant and scenario-independent value; and, the same for the deviation definitions, *dvtn*, which consequently will be null and useless.

The main characteristics and the average execution times for the instances treated in the various examples (EX1 to EX6) are described in Table 1. A laptop ASUS-F3JC (Intel Core2 T5500, 1.55GHz and 2GB of RAM) and GAMS/OSL are used (data generation in Appendix D).

Table 1: Numbers of parameters, variables and constraints corresponding to examples solved assuming: $NC = 4; M = 3; NS = 5; NP = 3; NT = 1$.

Numerical examples	**Parameters** NR	**Binary variables**	**Continuous variables**	**Constraints**	**Execution times(s)**
EX1	1	45	209	226	0.33
EX2	3	45	603	672	1.68
EX3	7	45	1391	1564	7.48
EX4	15	45	2967	3348	21.46
EX5	30	45	5922	6693	91.60
EX6	100	45	19712	22303	940.31

4.1. Robustness and Number of Scenarios

The effect related to the utilization of distinct numbers of discrete scenarios is analyzed in the generalized and stochastic model, which conceptually reduces to the deterministic one when considering a single scenario. Although the number of binary variables is kept constant for the various examples, the number of continuous variables and the number of constraints vary linearly with the number of scenarios, *NR*. Instead, the evolution of the average execution times is approximately quadratic.

Graphical representations are shown for the variation of the different estimators (robust cost, ψ; expected value of the non-satisfied demand, *Ensd*; expected value of the capacity slacks, *Eslk*; and non-robust cost, *Ecsi*), with the increasing scenarios number, NR, and penalization for non-satisfied demand, λqns. The subsequent analysis is presented.

Due to the near coincidence of the different lines represented, from NR = 1 (EX1) to NR = 100 (EX6), six lines may not be presented in some of the graphs. A general and descriptive approach is intended for the various examples, because they are targeted: the estimators' sensitivity to the variation of penalization parameter: and to support model and solution robustness.

From Figure 1, the values of the robust cost (NPV, ψ) vary significantly in the range of λqns from 1 to 4, which reveals the most sensitive zone and to which attention will be drawn. This evolution pattern is similar in the following figures that successively present *Ensd*, *Eslk*, and *Ecsi* evolutions, and a general range of stability for the estimator values is also observed for λqns greater than 5. Although not represented in *Figure 1*, when λqns increases in the range of 20 to 40, the robust costs ($NR > 1$) tend to the deterministic cost: the expected values of the non-satisfied demand, *Ensd*, are already small in that range, so the penalized component shall present a significant value only after a strong increment of λqns.

In relation to Figure 2, the expected value in *Ensd* for $\lambda qns = 0$ (not represented) is the result of the minimum configuration: a single process in each stage and with minimum dimension is selected, to which corresponds a virtual value of high non-satisfied demand (about $83 \times 10^3 kg$). For the first value represented, $\lambda qns = 1$, *Ensd* varies between 30×10^3 and $40 \times 10^3 kg$, representing about 23 % of the average demand ($150 \times 10^3 kg$) or a reduction to less than half in relation to the null penalization parameter value.

The non-satisfied demand concentrates in quite specific situations, worthy of a directed analysis: in one or several products, for one or several instances. From the observation of the optimum values for the variables Qns_{jtr}, it becomes clear that these variables are null in almost all the examples and for almost all the products. It will then be important to verify if the non-satisfaction of the demand of a specific product will have other consequences (accomplishment of trade agreements, customer service, etc.). In this case, penalization parameters associating the product in analysis may be introduced.

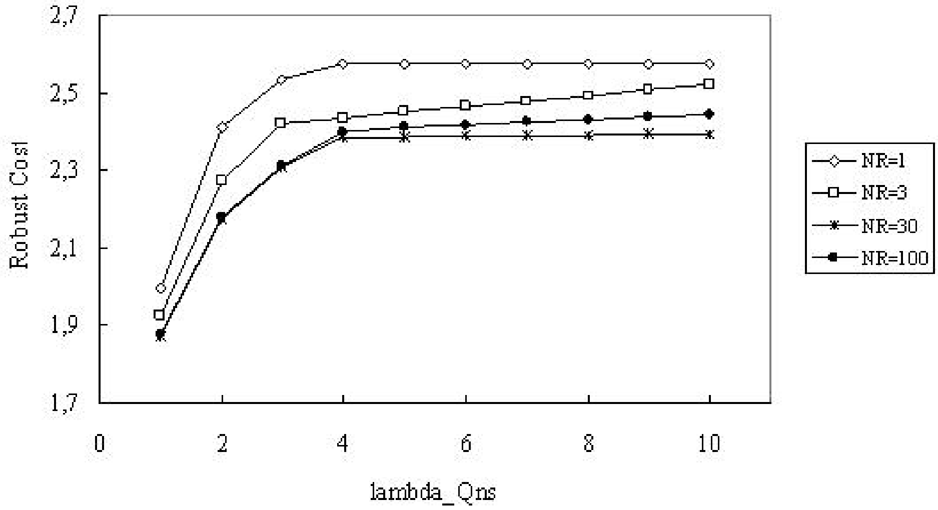


Figure 1: Variation of robust costs (10^5 €) versus the penalty values on non-satisfied demand, λqns, for various numerical examples.

In Figure 3, $Eslk$ tends to stabilize on an upper limit (about 1400 L) and, supposing the maximum configuration of discrete volumes, the plateau represents a percent underutilization (aggregating the discrete capacities of all the stages, 10929 L are obtained) of up to 13 %. Anyway, $Eslk$ represents a value of relative importance, as it depends on the instance at hand.

If, however, attention is paid to the percentage of Eslk as a function of the total capacity, stabilized values in the range 11-13 % are observed (as in Table 3, for $NR = 7$) in all the examples. In fact, for $\lambda qns = 0$, the minimum configuration corresponds to about 3451 L, in which a percent underutilization is verified (aggregated for all the stages) of 6.6 %. For the first values represented ($\lambda qns = 1$), *Eslk* is about 712 L, or 6.5 % of underutilization of the maximum capacity (10929), but corresponding to 11.7 % of the optimum capacity (6087 L) for this example (Table 2 and Table 3). Therefore, when λqns increases, *Eslk* also increases rapidly until the upper limit mentioned (about 1400 L), but the percentages of equipment underutilization in each example stay around between 11 and 13 %.

Figure 4 also yields a strong initial growth of non-robust cost, *Ecsi*, which then tends to stabilize in an upper value. This upper limit of *Ecsi* corresponds to the complete satisfaction of all the instances of the uncertain demand, and it also corresponds to the deterministic cost of investment of about 2.6×10^5 €. This value is significantly greater (about 8 %) than the remaining ones that are around 2.4×10^5€. Nevertheless, *Ecsi* presents a "permanent" value (for $\lambda qns = 0$, the minimum configuration) of about 1.25×10^5€, which represents a fraction of 48 %, relatively to the maximum configuration. For the first values ($\lambda qns = 1$ or 2) presented, *Ecsi* is about 1.6×10^5€, that is, 62 % of the cost of the maximum configuration. Then such configuration would decrease the complementary expected cost by about 38 %.

In Figure 5, the difference between the robust cost (ψ, or $NPVrob$ in the graph legend) and the non-robust cost (*Ecsi*) significantly comes from the consideration of the penalization on the non-satisfied demands, λqns: the variation of robust cost is linearly related to the evolution of this penalization parameter, when λqns is greater than 5 and non-satisfied demand is stabilized on a low value.

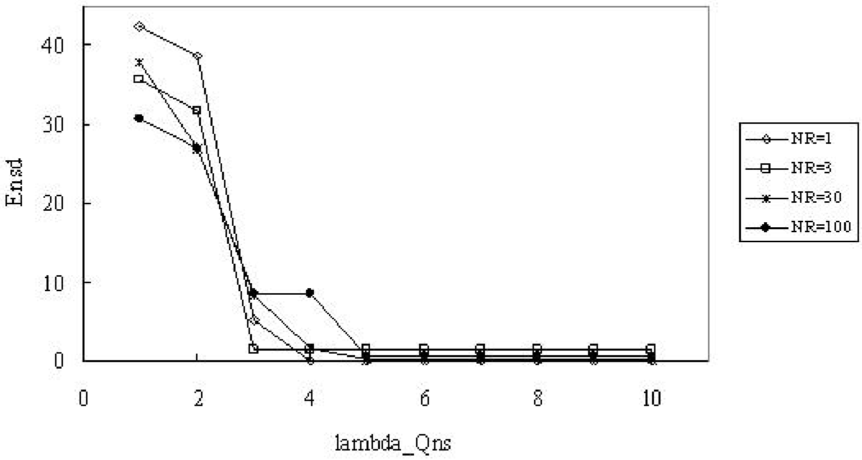


Figure 2: Variation of the expected value for non-satisfied demand, $Ensd(10^3 kg)$, with the evolution of λqns, for various numerical examples.

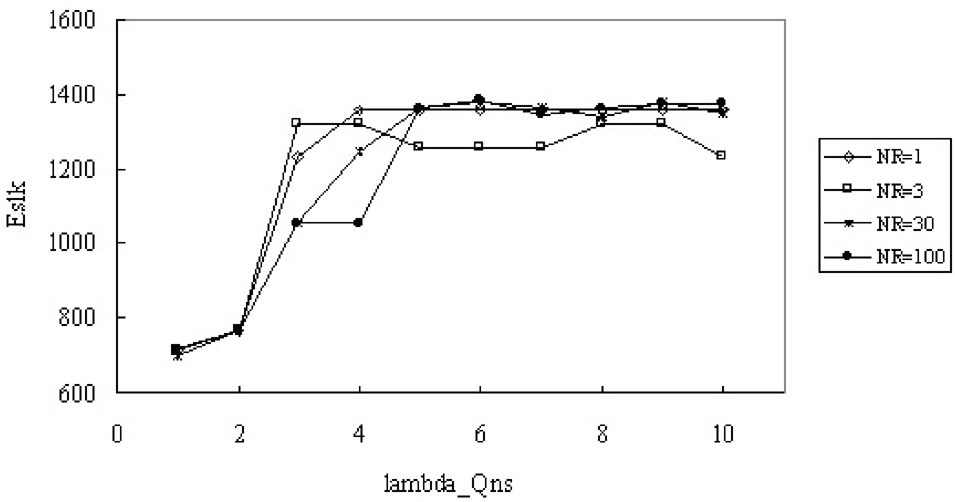


Figure 3: Variation of the expected value for capacity slacks, $Eslk(L)$, with the evolution of penalty values, λqns, for various numerical examples.

Accordingly with the figures 1 to 5, two key subjects are noticed:

i) the similarity of the behavior of the numerical instances when different number of scenarios is considered, *NR* from 1 to 100; and

ii) the model robustness, with adequate sensitivity of technical estimators to the evolution of the non-satisfied demand penalization parameter, λqns.

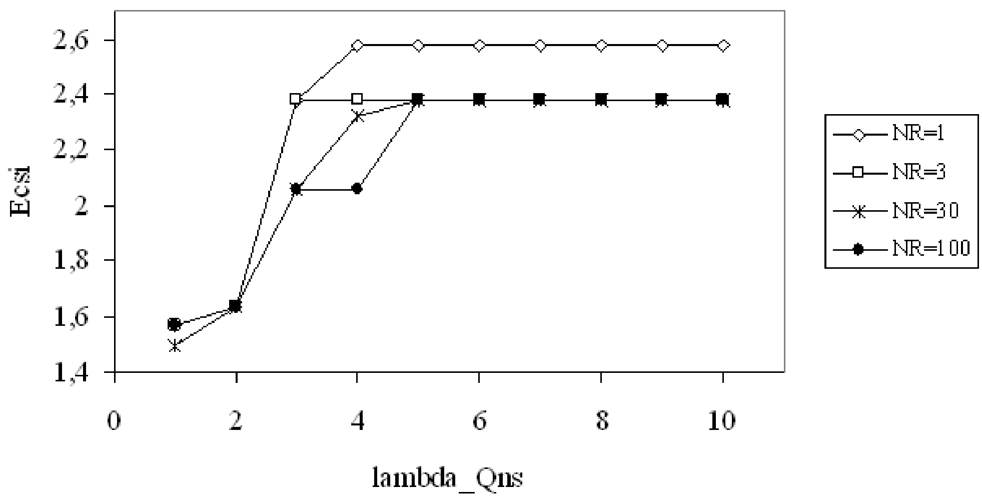


Figure 4: Variation of the expected value for investment costs, *Ecsi* (10^5 €), with the evolution of penalty values, λqns, for various numerical examples.

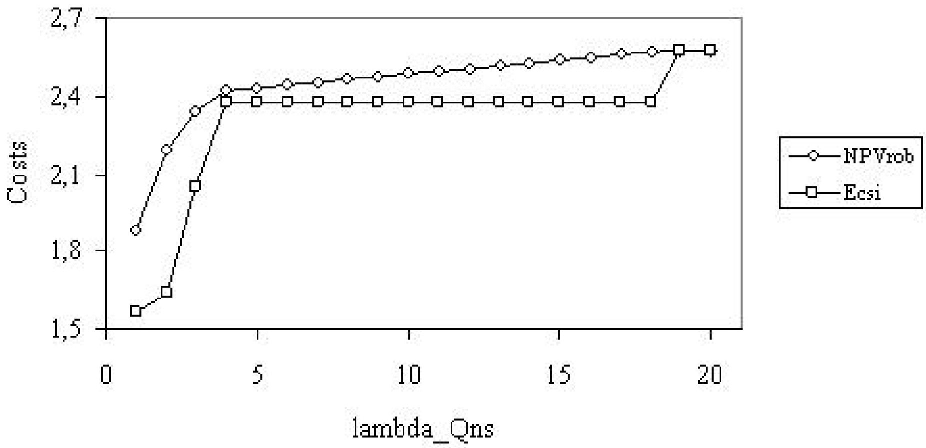


Figure 5: Variation of robust ($NPVrob$) and non-robust (*Ecsi*) costs (x10^5), with the evolution of penalty values, λqns (for $NR = 15$).

4.2. Focusing Robustness in Capacity Slack

In the following tables, Table 2 and Table 3, significant values are shown for instances of EX3, in which instances are considered with seven scenarios, but whose variation with the growth of the penalization on the non-satisfied demand, λqns, is similar to all the other examples, with different numbers of scenarios. These significant values are the values that are associated to alterations in the optimum configuration.

In Table 2 the expected values (robust and non-robust costs, and *Edvt*, *Ensd*, *Eslk*) of interest are shown, proving numerically the type of the evolutions observed in the previous graphs. Similarly to the expected value for the variability deviations, *Edvt*, the theoretical prediction is verified numerically: null return values, $ret_{jtr} = 0$, imply that the NPV to be maximized corresponds to the minimization of investment costs, which depends only on the binary variables.

Such order of results does not prove the robustness in the solution, but invariability is verified in this case: the equipment dimensions of the processes being defined in the first phase, the quantities produced will always be as great as possible, as the non-satisfied demand is penalized, which is equivalent to promoting production.

The prompt variation in the expected values of the non-satisfied demand, *Ensd*, with the increase of the respective penalty parameter, even making null this estimator, permits to state the model robustness.

Table 2: Significant values of the stochastic optimization, considering the evolution of penalty values and distinct instances of numerical example EX3 (for $NR = 7$).

λqns	**Costs (rob)**	**Costs**	***Edvt***	***Ensd***	***Eslk***
0.	124596.08	124596.08	0.0	153052.69	725.64
1.	188885.36	156737.60	0.0	32147.76	712.72
2.	220447.78	163661.91	0.0	28392.93	767.74
3.	234553.01	205467.00	0.0	9695.34	1042.93
4.	241546.56	237689.36	0.0	964.30	1404.58
5.	242510.86	237689.36	0.0	964.30	1233.89
	(...)		(...)		(...)
10.	247332.35	237689.36	0.0	964.30	1359.57
	(...)		(...)		
20.	256975.34	237689.36	0.0	964.30	1305.95
21.	257332.14	257332.14	0.0	0.0	1448.17
	(...)		(...)		(...)
40.	257332.14	257332.14	0.0	0.0	1400.87

Similarly to the capacity slacks, it should be noted that the *Eslk* estimator contains a character of permanence, as the equipment underutilization is underlying this type of problem: the underutilization or a slack would not exist, in the context of the ZW policy, if and only if all the products would present equal values for the technical parameters (namely, the dimension factors, S_{ij}, in L/kg) in the different stages. This numerical information on the expected values is complemented with the values in Table 3, relating the penalty parameters λqns and the (non-robust) costs, showing:

- the order of the discrete dimension (size, *s*) selected in each stage, *Ord(s)*; for example, ”1/ 3/ 4”, indicates that the first dimension in the first stage was chosen, the third dimension in the second stage, and the fourth dimension in the third stage; these values come directly from the binary solution;
- the sum of the discrete dimensions or equipment volumes selected, *Sum(dv)*, being opportune for the analysis of the slacks, given the interest of estimating *Eslk* in relative terms for each example;
- the percentage of the expected value of the non-satisfied demand, *%Ensd*, while 100 is the expected value of the uncertain demand in each example;
- the percentage of the expected value of the capacity slacks, *%Eslk*, the sum of volumes being the basis of the calculation of this estimator, *Sum(dv)*, in each example.

5. Conclusions

The model *spbatch_ms* is generalized from the deterministic model *MS*, featuring a two-stage stochastic framework with promotion of robustness. The generalized model simultaneously treats the scheduling of the production cycle embedded in the problem of design of batch processes, considering multiproduct environment (flowshop), multiple processes in each stage, SPC production and ZW storage policies.

The combination of SPC with multiple machines was found to be the most promising from a computational standpoint (Miranda, 2011*b*), and the model *MS* was here generalized toward a stochastic environment within the relaxation of the soft constraints regarding uncertain demand. Even for a reduced number of scenarios, results point to a significant reduction (8-20 %) on the investment costs in comparison to the deterministic non-relaxed case. If the MPC policy is adopted or if a slight relaxation is made to the impositions on the uncertain demands (respectively, of 1 % to 6 %), the demand relaxation does not cause real losses.

Beyond this significant decrease in terms of investment costs, other conclusions are obtained. Analyzing the variation of the number of scenarios, a similar behavior was observed for the various estimators. This similarity for the different numbers of scenarios (from 1 to 100) is to be remarked due to their impact in the execution times.

The robustness of the model *spbatch_ms* was also observed, with estimators responding adequately to the variation of the penalty parameter for non-satisfied demand. Also, the invariability of the configurations is mainly due to the realistic presupposition of discrete volumes. These results motivate to further continue this approach, which is integrating other developments under way.

Acknowledgements

This work was supported in part by CPQ ("Centro de Processos Químicos", Centre for Chemical Processes), IST, Technical University of Lisbon, Portugal. We thank the College of Technology and Management at the Polytechnic Institute of Portalegre and CIIST (IST Computing Centre).

Nomenclature

Index and sets

M	– number of stages *i*;
NC	– number of components or products *j*;
NP(*i*)	– (cardinal) number of processes *p*(*i*) per stage;
NR	– number of discrete scenarios *r*;
NS(*i*)	– (cardinal) number of discrete dimensions *s*(*i*) in the process of stage *i*;
NT	– number of time periods *t*;

Parameters

τ	– processing times (h), for each product j in stage i;
λdvt	– Negative deviation on NPV penalization parameter;
λqns	– non-satisfied demand penalization parameter;
λslk	– capacity slack penalization parameter;
c	– equipment cost related to process $p(i)$ and size $s(i)$ selected in stage i;
dv	– discrete equipment volume in each stage;
H	– time horizon;
$ncUpp$	– upper limit for disaggregated number of batches;
$prob$	– probability of scenario r;
$p(i)$	– (ordinal) number of processes in stage i;
Q	– demand quantities (uncertain) for each product i;
ret	– unit (uncertain) values of return (net values) of the products j, in period t and scenario r;
$s(i)$	– (ordinal) number of process discrete dimensions in stage i;
S	– dimension factor (L/kg), for each product j in stage i;
V	– equipment volume (continuous value) in each stage;

Variables

ξ	– NPV value in scenario r;
$dvtn$	– negative deviation on the value of NPV in scenario r;
n	– number of batches of product j, in period t and scenario r;
nc	– number of batches of product j, in period t and scenario r, disaggregated by process $p(i)$ and size $s(i)$ in each stage i;
Qns	– non-satisfied demand quantities of product j, in period t and scenario r;
slk	– capacity slacks in each stage i, concerning totality of the *batches* of each product j, in period t and scenario r;
$tcamp$	– campaign times (SPC) relative to each product j;
W	– global quantities produced of product j, in period t and scenario r;
y	– binary decision related to process $p(i)$ and size $s(i)$ selected in stage i;

Glossary

MPC	– multiple product campaign;
MS	– deterministic model focusing multiple processes and SPC;
spbatch_ms	– stochastic MILP model focusing multiple processes and SPC;
SPC	– single product campaign;
ZW	– zero wait storage policy;

6. Appendix A: Technical estimators

Non-robust NPV expected value:

$$Ecsi = \sum_{r=1}^{NR} prob_r \xi_r \tag{6.1}$$

Negative deviation expected value:

$$Edvt = \sum_{r=1}^{NR} prob_r.dvtn_r \tag{6.2}$$

Non-satisfied demand expected value:

$$Ensd = \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Qns_{jtr} \right) \tag{6.3}$$

Capacity slack expected value:

$$Eslk = \sum_{r=1}^{NR} \frac{prob_r}{M.NC.NT} \sum_{j=1}^{NC} \sum_{t=1}^{NT} \left\{ \sum_{i=1}^{M} \sum_{s=1}^{NS} \sum_{p=1}^{NP} p(i).y_{isp}. \left(dv_{js} - S_{ij}. \frac{W_{jtr}}{n_{jtr}} \right) \right\} \tag{6.4}$$

Percent non-satisfied demand expected value:

$$\%Ensd = \frac{Ensd}{Qmed}.100, \text{ with } \; Qmed = \sum_{r=1}^{NR} \frac{prob_r}{NC.NT} \left(\sum_{j=1}^{NC} \sum_{t=1}^{NT} Q_{jtr} \right) \tag{6.5}$$

Percent capacity slack expected value:

$$\%Eslk = \frac{Eslk}{Vtotal}.100, \text{ with } Vtotal = \sum_{i=1}^{M} \sum_{s=1}^{NS} \sum_{p=1}^{NP} \left(y_{isp}.dv_{is} \right) \tag{6.6}$$

7. Appendix B: Aspects of the generalized model

The generalization approach at hand foresees a dynamic treatment of the multiperiod horizon, within the uncertainty considered in the NPV maximization. The uncertainty in product demands, Q_{jtr}, is modeled through discrete scenarios. If full satisfaction of demand is required, the equipment sizing may be directed for scenarios of low probability but requiring large dimensions. Thus, the production flows, W_{jtr}, are defined through soft constraints, simultaneously with the definition of non-satisfied demand, Qns_{jtr}, which is penalized in the robust objective function:

$$W_{jtr} \leq Q_{jtr} \Rightarrow W_{jtr} + Qns_{jtr} = Q_{jtr}, \quad \forall j, t, r. \tag{7.1}$$

When the production flows variable, W_{jtr}, is used instead of the parameter on uncertain demand, Q_{jtr}, non-linearities occur in the constraints defining the number of batches, n_{jtr},

$$W_{jtr}.y_{isp} = Waux_{ijsptr}, \quad \forall i, j, s, p, t, r, \tag{7.2}$$

would lead to the simultaneous consideration of the following three sets of constraints:

$$Waux_{ijsptr} \leq Waux_{ijsp}^{Upp} y_{isp}\,, \; \forall\, i, j, s, p, t, r, \tag{7.3}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} y_{isp} = 1, \quad \forall i, \tag{7.4}$$

$$\sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} Waux_{ijsptr} = W_{jtr}, \quad \forall i, j, t, r. \tag{7.5}$$

Thus, the number of variables *(M.N.NS.NP.NT.NR)* and the number of constraints *(M.N.NS.NP.NT.NR + M.N.NT.NR)* would be highly increased, and a coincidence would also occur with the usual aggregated variable:

$$n_{jtr}.y_{isp} = nc_{ijsptr}, \quad \forall i, j, s, p, t, r. \tag{7.6}$$

Now, using the original formulation (Kocis & Grossmann, 1988), which had batch size, B_{jtr}, associated with the equipment volume dimension, V_i, and if production flows, W_{jtr}, are further formulated in association with discrete volumes, dv_{is}, it follows:

$$\begin{aligned}
&S_{ij}B_{jtr} \leq V_i, \quad \forall i, j, t, r \quad \Rightarrow S_{ij}B_{jtr} \leq \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp} \Rightarrow \\
&S_{ij}\frac{W_{jtr}}{n_{jtr}} \leq \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp} \Rightarrow S_{ij}W_{jtr} \leq \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp}n_{jtr} \Rightarrow \\
&S_{ij}W_{jtr} \leq \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}nc_{ijsptr}, \quad \forall i, j, t, r.
\end{aligned} \tag{7.7}$$

Consequently, the semicontinuous variables, nc_{ijsptr}, are required, and the auxiliary variables, $Waux_{ijsptr}$, and the inherent constraints are avoided.

Furthermore, the capacity slacks must be penalized in the robust model. The constraint slacks, slk_{ijtr}, occur in each stage and they are related to the global quantity produced, W_{jtr}, in all time periods and discrete scenarios,

$$S_{ij}W_{jtr} + slk_{ijtr} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}nc_{ijsptr}, \quad \forall i, j, t, r. \tag{7.8}$$

These constraint slacks, slk_{ijtr}, correspond to the aggregated value of the slacks that are occurring in all of the production cycles. Such variables can be related to the effective slack that occurs in each one of the started batches, Zp_{ijtr},

$$\begin{aligned}
&S_{ij}\frac{W_{jtr}}{n_{jtr}} + \frac{slk_{ijtr}}{n_{jtr}} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp} \Rightarrow \\
&S_{ij}B_{jtr} + Zp_{ijtr} = \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} dv_{is}y_{isp}.
\end{aligned} \tag{7.9}$$

The effective slack verified in the selected equipment, Zp_{ijtr}, can be obtained through the disaggregation of the global slack, slk_{ijtr}, in association with the number of started batches, n_{jtr}:

$$Zp_{ijtr} = \frac{slk_{ijtr}}{n_{jtr}}, \quad \forall i, j, t, r. \tag{7.10}$$

This non-linear relation between the two slack variables forbids the consideration of the effective capacity slack, Zp_{ijtr}, in the robust objective function. Nevertheless, this variable can be used to assess the slacks verified in the selected configuration of equipment volumes.

8. Appendix C: Reduction to Static *MS* Problem

The computational complexity of the model *spbatch_ms* is studied, as follows:

- if one supposes $NT = NR = 1$, with $prob(1) = 1$, then the time t and random r index can be neglected;
- provided that $NR = 1$, the linear negative deviation is null,

$$dvtn_r = \sum_{r'=1}^{1} (prob_{r'}\xi_{r'}) - \xi_r = 0; \tag{8.1}$$

- assuming $\lambda qns = bigM$, the optimization resolution will lead to

$$Qns_i = 0. \Rightarrow W_i = Q_i \ , \quad \forall i; \tag{8.2}$$

- assuming $\lambda slk = 0.$ (i.e., do not penalize capacity slack) and $ret_{jtr} = 0.$ (do not account return or cash flows), the objective function presents the following pattern of minimization of the investment costs:

$$\max \ \Phi = \max \ \xi_r = \max \left\{ - \sum_{j=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp} \right\} = \min \left\{ \sum_{j=1}^{M} \sum_{s=1}^{NS(i)} \sum_{p=1}^{NP(i)} c_{isp} y_{isp} \right\}; \tag{8.3}$$

- the remaining constraint sets are equivalent to the constraint sets of *MS*; it follows, by instance reduction, that if it is proved that *MS* problem belongs to NP-Hard class, it also drives the computational complexity of *spbatch_ms* problem.

Consequently, the study of the deterministic and static *MS* problem (a single time period) presented characteristics that remain valid in all of the discrete scenarios and time periods considered in the stochastic and multiperiod *spbatch_ms* problem:

- from constraint set 3.10, in all stages i, one and only one of discrete dimension *s'(i)* and number of machines *p'(i)* can be selected,

$$\begin{cases} y_{isp} = 0, & \forall \quad s \neq s'(i), \quad p \neq p'(i) \\ y_{i,s'(i),p'(i)} = 1, & \exists^1 \ (s'(i), p'(i)) \end{cases} \tag{8.4}$$

- by conjugation of constraint sets 3.8 and 3.9, the nc_{ijsp} variables present zero value, with the exception of the ones associated to the selected discrete dimension *s'(i)* and number of machines *p'(i)*, thus the value presented by n_j is equal in all stages; further, the trivial satisfaction of the upper limit in logic constraint 3.7 ($nc^{Upp} = bigM$) is supposed and it follows that

$$\begin{cases} nc_{ijsp} = 0, & \forall \quad s \neq s'(i), \quad p \neq p'(i) \\ nc_{ij,s'(i),p'(i)} = n_j, & \exists^1 \ (s'(i), p'(i)) \end{cases} \tag{8.5}$$

- from constraint set 3.7 it is observed that, for each product j, the number of batch n_j definition must be satisfied in all stages i; thus, the n_j value obtained corresponds to the maximum of the values assessed through the i index,

$$S_{ij} W_j \leq p'(i) dv_{i,s'(i)} n_j, \quad \forall i, j \Rightarrow n_j = \max_i \left\{ \frac{S_{ij} W_j}{p'(i) dv_{i,s'(i)}} \right\}, \ \forall i \tag{8.6}$$

- in constraint set 3.11, the time campaign $tcamp_j$ for each product j must be satisfied in each stage i; it thus corresponds to the maximum value of the sum formulated in this constraint set; in the sum, the number of production cycles nc_{ijsp} is limited by the nj value in all stages i, as expressed by constraint sets 3.8 and 3.10,

$$tcamp_j = \max_i \left\{ \frac{\tau_{ij}}{p'(i)} n_j \right\} = \max_i \left\{ \frac{\tau_{ij} S_{ij} W_j}{p'(i) dv_{i,s'(i)}} \right\}, \quad \forall j \tag{8.7}$$

- in constraint set 12, the time campaign $tcamp_j$ sum must satisfy the time horizon H. Then it follows that

$$\sum_{j=1}^{NC} \max_i \left\{ \frac{\tau_{ij} S_{ij} W_j}{p'(i) dv_{i,s'(i)}} \right\} \leq H \tag{8.8}$$

In conclusion, the robust objective of maximizing NPV must be observed. Thus the equipment configuration (s', p') must be selected satisfying the constraint sets in all stages i and for all products j, given all discrete scenarios in all time periods. This way, the production flow W_{jtr} is associated to the return (cash) flows. The decrease of solution variability is foreseen through the penalization of the negative deviation on NPV, and model robustness is promoted through the penalization of capacity slacks.

9. Appendix D: Data generation for numerical examples

Comprehensively and in GAMS environment, the generation of random data for numerical examples is specified through the following lines of code.

```
OPTION SEED = 08012007
H = UNIFORM (6000, 8000);
alpha(j)= UNIFORM (300, 700); /*cost function coefficient */
beta(j) = UNIFORM (0.5, 0.7); /*cost function exponent */
LOOP (j,
    dv(j,”1”) = UNIFORM(800,1300);
    dv(j,”2”) = UNIFORM(1500,2000);
    dv(j,”3”) = UNIFORM(2200,2700);
    dv(j,”4”) = UNIFORM(2800,3300);
    dv(j,”5”) = UNIFORM(3500,4000);
);
c(j,s) = alpha(j)*dv(j,s)**beta(j);
S(i,j) = UNIFORM (1,5);
τ(i,j) = UNIFORM (2,9);
Q_med = 150000; Q_dsv = Q_med / 6.;
Q(i,t,r) = NORMAL (Q_med, Q_dsv);
```

References

Ahmed, S. and N. V. Sahinidis (2000). Analytical investigations of the process planning problem. *Computers and Chemical Engineering* **23**, 1605– 1625.

Ahmed, S. and N. V. Sahinidis (2003). An approximation scheme for stochastic integer programs arising in capacity expansion. *Operations Research* **53**(3), 461–471.

Barbosa-Póvoa, A.P. (2007). A critical review on the design and retrofit of batch plants. *Computers and Chemical Engineering* (31), 833–855.

Birewar, D.B. and I. E. Grossmann (1989). Efficient optimization algorithms for zero-wait scheduling of multiproduct batch plants. *Ind. Eng. Chem. Res.* (28), 1333–1345.

Cavin, L., U. Fischer, F. Glover and K. Hungerbhler (2004). Multi-objective process design in multi-purpose batch plants using a tabu search optimization algorithm. *Computers and Chemical Engineering* (28), 459–478.

Floudas, C.A. and X. Lin (2004). Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. *Computers and Chemical Engineering* (28), 2109–2129.

Jayaraman, V. K., B. D. Kulkarni, S. Karale and P. Shelokar (2000). Ant colony framework for optimal design and scheduling of batch plants. *Computers and Chemical Engineering* (24), 1901–1912.

Kocis, G.R. and I.E. Grossmann (1988). Global optimization of nonconvex mixed-integer nonlinear programming (MINLP) problems in process synthesis. *Ind. Eng. Chem. Res.* (27), 1407–1421.

Liu, M.L. and N.V. Sahinidis (1997). Bridging the gap between heuristics and optimization: capacity expansion case. *AIChE J.* (43), 2289–2299.

Miranda, J. L. (2007). Optimização em sistemas de processos químicos: Generalização de modelos com planeamento e sequenciamento. PhD dissertation, Instituto Superior Técnico, Technical University of Lisbon, Lisboa.

Miranda, J. L. (2011*a*). Computational complexity studies in the design and scheduling of batch processes.. In: *Book of Abstracts of IO2011 - 15th Congress of APDIO*. University of Coimbra, Coimbra, Portugal.

Miranda, J.L. (2011*b*). The design and scheduling of chemical batch processes: building heuristics and probabilistic analysis. *Theory and Applications of Mathematics & Computer Science* **1**(1), 45–62.

Moreno, M.S., J.M. Montagna and O.A. Iribarren (2007). Multiperiod optimization for the design and planning of multiproduct batch plants. *Computers and Chemical Engineering* (31), 1159–1173.

Pekny, J. F. and D. L. Miller (1991). Exact solution of the no-wait flowshop scheduling problem with a comparison to heuristic methods. *Computers and Chemical Engineering* **15**(11), 741–748.

Tan, S. and R. S. H. Mah (1998). Evolutionary design of noncontinuous plants. *Computers and Chemical Engineering* **22**(1/2), 69–85.

Voudouris, V. T. and I. E. Grossmann (1992). Mixed-integer linear programming reformulations for batch processes design with discrete equipment sizes. *Industrial and Engineering Chemistry Research* (31), 1315–1325.

Voudouris, V. T. and I. E. Grossmann (1993). Optimal synthesis of multiproduct batch plants with cyclic scheduling and inventory considerations. *Industrial and Engineering Chemistry Research* (32), 1962–1980.

Voudouris, V. T. and I. E. Grossmann (1996). Milp model for scheduling and design of a special class of multipurpose batch plants. *Computers and Chemical Engineering* (20), 1335–1360.

Xia, Q. and S. Macchietto (1997). Design and synthesis of batch plants - minlp solution based on a stochastic method. *Computers and Chemical Engineering* **Suppl.**(21), S697–S702.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 1 (2) (2011) 89–92

Distortion Estimate and the Radius of Starlikeness of Janowski Close-to-Star Functions

Arzu Sen[a], Melike Aydoğan[b,*], Yaşar Polatoğlu[a]

[a]İstanbul Kültür University, Department of Mathematics and Computer Sciences İstanbul, Turkey.
[b]Isik University, Sile Kampusu, Department of Mathematics, İstanbul, Turkey.

Abstract

Let F be the class of all analytic functions in the open unit disc $\mathbb{D} = \{z \quad |z| < 1\}$ of the form $f(z) = z + a_2z^2 + a_3z^3 + \cdots$. Let $g(z)$ be an element of F such that $g(z)$ satisfies the condition

$$(z\frac{g'(z)}{g(z)}) = \frac{1 + A\phi(z)}{1 + B\phi(z)},$$

for all $z \in \mathbb{D}$, where $\phi(z)$ is analytic in $\mathbb{D}$ and satisfying the conditions $\phi(0) = 0$, $|\phi(z)| < 1$, and $-1 \leq B < A \leq 1$, then $g(z)$ is called Janowski Starlike function in $\mathbb{D}$. The class of such functions is denoted by $\mathcal{S}^*(A, B)$ (Janowski, 1973).
The aim of this paper is to give a distortion estimation and the radius of starlikeness of the class $\mathcal{S}^*K(A, B)$.

Keywords: Close-to-star function, the radius of starlikeness, distortion estimate.
2000 MSC: 30C45.

1. Introduction

Let Ω be the family of functions $\phi(z)$ which are regular in $\mathbb{D}$ and satisfy the condition $\phi(0) = 0$, $|\phi(z)| < 1$ for all $z \in \mathbb{D}$. The family of functions $p(z) = 1 + p_1z + p_2z^2 + \cdots$ analytic in $\mathbb{D}$, and satisfying the conditions $p(0) = 1$, $Re\ p(z) > 0$ is denoted by $\mathcal{P}$ such that $p(z)$ in $\mathcal{P}$ if and only if

$$p(z) = \frac{1 + \phi(z)}{1 - \phi(z)}, \tag{1.1}$$

for some $\phi(z) \in \Omega$, and every $z \in \mathbb{D}$. The class $\mathcal{P}$ is the Caratheodory class (Nehari, 1952).

Next, for arbitrary fixed real numbers A and B which satisfy $-1 \leq B < A \leq 1$, we say $p(z)$ belongs to the class $P(A, B)$ if $p(z) = 1 + p_2z^2 + p_3z^3 + \cdots$ is analytic in $\mathbb{D}$ and $p(z)$ is given by

$$p(z) = \frac{1 + A\phi(z)}{1 + B\phi(z)}, \tag{1.2}$$

*Corresponding author
Email addresses: a.sen@iku.edu.tr (Arzu Sen), melikeaydogan43@gmail.com (Melike Aydoğan), y.polatoglu@iku.edu.tr (Yaşar Polatoğlu)

for some $\phi(z) \in \Omega$ and every $z \in \mathbb{D}$. The class $P(A, B)$ was introduced by (Janowski, 1973).

Let $f(z)$ be an element of F and $g(z)$ be an element of $\mathcal{S}^*(A, B)$, if the condition

$$Re(\frac{f(z)}{g(z)}) > 0 \tag{1.3}$$

is satisfied, then $f(z)$ is called Janowski close-to-star function in $\mathbb{D}$. The class of such functions is denoted by $\mathcal{S}^*K(A, B)$. The class of $\mathcal{S}^*K(A, B)$ is not empty because,

$$f(z) = \begin{cases} z(1 + Bz)^{\frac{A-B}{B}} . \frac{1+z}{1-z}; & B \neq 0, \\ ze^{Az} \frac{1+z}{1-z}; & B = 0. \end{cases} \tag{1.4}$$

Finally, let $F(z) = z + \alpha_2 z^2 + \alpha_3 z^3 + \cdots$ and $G(z) = z + \beta_2 z^2 + \beta_3 z^3 + \cdots$ be analytic functions in $\mathbb{D}$, if there exist a function $\phi(z) \in \Omega$ such that $F(z) = G(\phi(z))$ for every $z \in \mathbb{D}$, then we say that $F(z)$ is subordinate to $G(z)$, and we write $F(z) \prec G(z)$. We also note that if $F(z) \prec G(z)$, then $F(\mathbb{D}) \subset G(\mathbb{D})$.

2. Main Results

Lemma 2.1. *Let $f(z)$ be an element of $\mathcal{S}^*K(A, B)$, then*

$$Re[(1 + Bz)^{-\frac{A-B}{B}} \frac{f(z)}{z}] > 0, B \neq 0, \tag{2.1}$$

$$Re[e^{-Az} \frac{f(z)}{z}] > 0, B = 0. \tag{2.2}$$

Proof. Since the function,

$$g(z) = \begin{cases} z(1 + Bz)^{\frac{A-B}{B}}; & B \neq 0, \\ ze^{Az}; & B = 0. \end{cases} \tag{2.3}$$

belongs to the class $S^*(A, B)$ (Janowski, 1973), then using the definition of the class $\mathcal{S}^*K(A, B)$,

$$Re(\frac{f(z)}{g(z)}) = Re[\frac{f(z)}{z(1 + Bz)^{\frac{A-B}{B}}}] = Re[(1 + Bz)^{-\frac{A-B}{B}} . \frac{f(z)}{z}] > 0, B \neq 0, \tag{2.4}$$

$$Re(\frac{f(z)}{g(z)}) = Re[\frac{f(z)}{ze^{Az}}] = Re[e^{-Az} \frac{f(z)}{z}] > 0, B = 0. \tag{2.5}$$

□

Theorem 2.2. *Let $f(z)$ be an element of $\mathcal{S}^*K(A, B)$, then for $r = |z|$*

$$F(-A, -B, r) \leq |f(z)| \leq F(A, B, r), B \neq 0, \tag{2.6}$$

$$G(-A, r) \leq |f(z)| \leq G(A, r), B = 0, \tag{2.7}$$

where

$$F(A,B,r)=\frac{r(1+r)(1+Br)^{\frac{A-B}{B}}}{(1-r)} \tag{2.8}$$

and

$$G(A,r)=\frac{r(1+r)e^{Ar}}{(1-r)}. \tag{2.9}$$

Proof. Let $g(z)$ be an element of $\mathcal{S}^*(A,B)$, then

$$F_1(-A,-B,r)\leq|g(z)|\leq F_1(A,B,r), \tag{2.10}$$

where

$$F_1(A,B,r)=\begin{cases} r(1+Br)^{\frac{A-B}{B}}; & B\neq 0,\\ re^{Ar}; & B=0.\end{cases} \tag{2.11}$$

On the other hand, if $p(z)\in\mathcal{P}$ then we have,

$$\frac{1-r}{1+r}\leq|p(z)|\leq\frac{1+r}{1-r} \tag{2.12}$$

(Goodman, 1983).

Considering 2.10 and 2.12 together and after the straightforward calculations, we get 2.6 and 2.7. We also note that the inequalities 2.6 and 2.7 are sharp because the extremal functions are;

$$(1+Bz)^{-\frac{A-B}{B}}\frac{f(z)}{z}=p(z)=\frac{1+z}{1-z}\Rightarrow f(z)=\frac{z(1+z)(1+Bz)^{\frac{A-B}{B}}}{1-z}, B\neq 0, \tag{2.13}$$

$$e^{-Az}\frac{f(z)}{z}=p(z)=\frac{1+z}{1-z}\Rightarrow f(z)=\frac{z(1+z)e^{Az}}{1-z}, B=0. \tag{2.14}$$

□

Remark. If we give the special values to A and B, we obtain that new inequalities and new growth theorems for the subclass of $S^*K(A,B)$. The special values of A and B can be ordered in the following manner:

i. $A=1$, $B=-1$;
ii. $A=1-2\alpha$, $B=-1$, $0\leq\alpha<1$;
iii. $A=1$, $B=0$;
iv. $A=\alpha$, $B=0$, $0<\alpha<1$;
v. $A=1$, $B=-1+\frac{1}{M}$, $M>\frac{1}{2}$;
vi. $A=\alpha$, $B=-\alpha$, $0<\alpha<1$.

Theorem 2.3. *The radius of starlikeness of the class $\mathcal{S}^*K(A,B)$ is the smallest positive roots r_0 of the equations,*

$$\begin{cases} Q_1(r)=-2r(1-Br)+(1-r^2)(1-Ar); & B\neq 0,\\ Q_2(r)=-2r+(1-Ar)(1-r^2); & B=0.\end{cases} \tag{2.15}$$

Proof. Using Lemma 2.1, then we obtain

$$(1+Bz)^{-\frac{A-B}{B}}\frac{f(z)}{z}=p(z), B\neq 0, \tag{2.16}$$

$$e^{-Az}\frac{f(z)}{z}=p(z), B=0, \tag{2.17}$$

$$\begin{cases} z\frac{f'(z)}{f(z)}=\frac{1+Az}{1+Bz}+z\frac{p'(z)}{p(z)}; & B\neq 0,\\ z\frac{f'(z)}{f(z)}=(1+Az)+z\frac{p'(z)}{p(z)}; & B=0. \end{cases} \tag{2.18}$$

Thus,

$$\begin{cases} Re(z\frac{f'(z)}{f(z)})\geq Min_{|z|=r}Re(\frac{1+Az}{1+Bz})+Min_{|z|=r,p(z)\in\mathcal{P}}\ Re(z\frac{p'(z)}{p(z)}); & B\neq 0,\\ Re(z\frac{f'(z)}{f(z)})\geq Min_{|z|=r}Re(1+Az)+Min_{|z|=r,p(z)\in\mathcal{P}}\ Re(z\frac{p'(z)}{p(z)}); & B=0. \end{cases} \tag{2.19}$$

$$Re(z\frac{f'(z)}{f(z)})\geq Min_{|z|=r}Rez\frac{p'(z)}{p(z)}+Min_{|z|=r}Rez\frac{(A-B)}{1+Bz}+Min_{|z|=r}Re(1). \tag{2.20}$$

Therefore we have:

$$\begin{cases} Re(z\frac{f'(z)}{f(z)})\geq\frac{-2r(1-Br)+(1-r^2)(1-Ar)}{(1-r^2)(1-Br)}; & B\neq 0,\\ Re(z\frac{f'(z)}{f(z)})\geq\frac{-2r+(1-Ar)(1-r^2)}{(1-r^2)}; & B=0. \end{cases} \tag{2.21}$$

The denominator of the expression on the right hand sides of the inequalities 2.20 is positive for $0\leq r<1$, $Q_1(0)=1$, $Q(1)=-2(1-B)<0$, $Q_2(0)=1$, $Q_2(1)=-2<0$.

Thus using intermediate value theorem and mean value theorem, the smallest positive roots of the equations 2.15 lies between 0 and 1.

Therefore the inequality

$$Re\left(z\frac{f'(z)}{f(z)}\right)>0, \tag{2.22}$$

is valid for $r=|z|<r_0$. Hence the radius of starlikeness for $\mathcal{S}^*K(A,B)$ is not less than r_0. The theorem is proved. □

References

Goodman, A. W. (1983). *Univalent functions, Volume I and II*. Mariner publishing company, Tampa Florida, USA.

Janowski, W. (1973). Some extremal problems for certain families of analytic functions. I. *Annales Polonici Mathematici* **28**, 297–326.

Nehari, Z. (1952). *Conformal Mapping*. McGraw-Hill, New York.