amcs

THE IMPATIENCE MECHANISM AS A DIVERSITY MAINTAINING AND SADDLE CROSSING STRATEGY

Iwona KARCZ-DULEBA [a]

[a]Department of Control and Mechatronics, Faculty of Electronics
Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: `iwona.duleba@pwr.edu.pl`

The impatience mechanism diversifies the population and facilitates escaping from a local optima trap by modifying fitness values of poorly adapted individuals. In this paper, two versions of the impatience mechanism coupled with a phenotypic model of evolution are studied. A population subordinated to a basic version of the impatience mechanism polarizes itself and evolves as a dipole centered around an averaged individual. In the modified version, the impatience mechanism is supplied with extra knowledge about a currently found optimum. In this case, the behavior of a population is quite different than previously—considerable diversification is also observed, but the population is not polarized and evolves as a single cluster. The impatience mechanism allows crossing saddles relatively fast in different configurations of bimodal and multimodal fitness functions. Actions of impatience mechanisms are shown and compared with evolution without the impatience and with a fitness sharing. The efficiency of crossing saddles is experimentally examined for different fitness functions. Results presented in the paper confirm good properties of the impatience mechanism in diversity maintaining and saddle crossing.

Keywords: phenotypic evolution, impatience operator without and with extra knowledge, polarization of population, maintaining population diversity, saddle crossing.

1. Introduction

Although evolutionary computation based methods are widely successful in solving complex optimization problems, they still encounter some difficulties limiting their efficiency and applications. One of the most important among these is premature convergence to a local optimum. A population trapped at the optimum usually has limited diversity and limited ability to find another, possibly better, optimum. Many strategies of regaining population diversity have been proposed: crowding (DeJong, 1975; Mengshoel and Goldberg, 2008; Kowalczuk and Białaszewski, 2006), fitness sharing (Goldberg and Richardson, 1987; Sareni and Krähenbühl, 1998) and spatially structured populations (Tomassini, 2005; Dick and Whigham, 2006) are the most popular. In order to preserve diversity, hybrid methods combine different optimization strategies at various stages of the search process (Grosan and Abraham, 2007; Barabasz *et al.*, 2014).

In many practical applications, it is more important to quickly find a better solution than the current best one, rather than to maintain multiple solutions in a long-lasting process. We can name this task a "local global optimization". While solving the task, crossing saddles between optima by a population actually being in a state of a quasi-equilibrium around a local optimum becomes a key problem. Softening a selective pressure, taking a population of a small size or application of special operators may help escaping from the current optimum.

Selections with a low selective pressure allow the survival of worse adapted individuals. Thus a mean population fitness decreases and the chance of saddle crossing increases (Galar, 1989; Goldberg and Deb, 1991; Chorazyczewski and Galar, 1998; Chen *et al.*, 2010). Smaller populations are more mobile than bigger ones and they are also susceptible to a genetic drift that makes saddle crossing easier (Rogers and Prügel-Bennett, 1999; Karcz-Duleba, 2004; 2006). Mutations with heavy tails increase the probability of having a distant offspring that may cause some individuals to jump over the saddle (Obuchowicz and Prętki, 2004). Other ideas proposed to facilitate saddle crossing are based on modification

of either a fitness function or individual fitness values. *Penalty methods* discussed by Torn and Zilinskas (1989) or the *erosion* procedure proposed by Obuchowicz (1997) transformed a fitness function. Methods based on fitness sharing (Goldberg and Richardson, 1987) or exploiting the mechanism of impatience (Galar and Kopciuch, 1999; Karcz-Duleba, 2014) can be classified to the latter group.

The method of *fitness sharing* is based on the concept of finite resources and penalizes similar individuals by lowering their fitness to encourage the population to explore different regions of a search space. The impatience mechanism modifies the fitness of individuals of a population trapped at a local optimum and promotes worse adapted individuals. After spending some time wandering around a local optimum, a population becomes "impatient" and changes its preferences by rewarding individuals located on the periphery of a population cloud by increasing their fitness. If the impatience operator is referred to a population mean individual, a population is polarized and evolves as a dipole centered at the mean. If the operator is referred to an already found local optimum (a population has knowledge about the optimum), a population is not polarized and evolves as a whole but its diversity is still high.

Both versions of the impatience mechanism will be studied in detail in this paper from the viewpoint of maintaining population diversity and crossing saddles between optima. The paper is organized as follows. In Section 2 the impatience operators coupled with a simple model of a phenotypic evolution are depicted. Fitness functions used for analyzing properties and effectiveness of the mechanism are also presented in this section. The dynamics of populations evolving with impatience mechanisms and the population polarization phenomenon are described in Section 3. In this section the properties of both versions of impatience as well as their similarities and differences are discussed. In Section 4, comparison of an evolution with and without impatience and impatience with a fitness sharing method is provided. Results of simulation studies of crossing saddles for different bimodal and multimodal fitness functions are presented in Section 5. Section 6 concludes the paper.

2. Mechanism of impatience

2.1. Model of evolution with impatience.
Galar and Kopciuch (1999) introduced an impatience mechanism into a model of simple phenotypic generational asexual evolution (Galar, 1989). A population of m-elements $P = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is evolving in an n-dimensional unbounded real search space $\mathbb{R}^n$. Each member of the population is characterized by its position $\mathbf{x}_\bullet \in \mathbb{R}^n$ and a fitness value $q(\mathbf{x}_\bullet)$. An evolutionary process takes place by means of fitness-proportionate selection and a Gaussian mutation with a standard deviation σ. An

operator of impatience modifies the original fitness of an individual $q(\mathbf{x})$. The modification is based on the individuals' distance from the mean individual within a population and is given by

$$q_a(\mathbf{x}_i) = \left(\frac{d_i}{d_A} + c\right) q(\mathbf{x}_i), \qquad (1)$$

where $q(\mathbf{x}_i)$ is a fitness of the i-th individual, d_i and d_A are, respectively, the distance of the i-th individual and an average distance of all individuals from the mean $\overline{\mathbf{x}}$,

$$\overline{\mathbf{x}} = \frac{1}{m}\sum_{i=1}^m \mathbf{x}_i, \quad d_i = \|\mathbf{x}_i - \overline{\mathbf{x}}\|, \quad d_A = \frac{1}{m}\sum_{i=1}^m d_i,$$

where $\| \cdot \|$ denotes the Euclidean metric and c is a constant parameter taken from the interval $[0, 1]$. The value of parameter c will be set to $c = 1$ later on, as our earlier studies revealed its negligible influence on the performance of the impatience mechanism.

2.2. Impatience with knowledge.
The mechanism of impatience is switched on when a population "is impatient" remaining for a long time in the attraction basin of a local optimum without any further progress. A modification of the operator was proposed by Karcz-Duleba (2014) to include extra knowledge about a currently found optimum. Thus the mechanism of impatience with knowledge may start to act when a position of a local optimum is either known or can be reliably estimated. In the discussed version of the impatience mechanism, the operator (1) is changed by replacing the distances from the population mean with the distances from the (estimated) optimum. Consequently, a modified fitness of the i-th individual located at distance d_{io} from the local optimum is described by

$$q_K(\mathbf{x}_i) = \left(\frac{d_{io}}{d_{Ao}} + c\right) q(\mathbf{x}_i), \qquad (2)$$

where d_{Ao} denotes the average distance to the optimum for the whole population,

$$d_{Ao} = \frac{1}{m}\sum_{i=1}^m \|\mathbf{x}_i - \mathbf{x}^*\| = \frac{1}{m}\sum_{i=1}^m d_{io},$$

and $\mathbf{x}^*$ is the position of a local (possibly estimated) optimum.

2.3. Fitness landscapes.
As the main purpose of introducing the impatience mechanism was to maintain population diversity and to facilitate crossing saddles, the impact of the operator on the evolution process was studied in an adaptive landscape of an n-dimensional bimodal fitness function. The sum of two Gaussian functions with a local and a global optimum of

different heights forming a distinct saddle between peaks were chosen for tests,

$$q(\mathbf{x}) = \exp(-a_1 \mathbf{x}^T \mathbf{x}) \\ + h \exp\left(-a_2(\mathbf{x} - \mathbf{x_g})^T C^{-1}(\mathbf{x} - \mathbf{x_g})\right), \quad (3)$$

where $\mathbf{x}_g$ is the location of the global optimum and h denotes its height ($h > 1$), parameters a_1, a_2 influence slopes of both hills, and C is a covariance matrix. A local optimum with a height equal to one is located at $\mathbf{x}_l = (0, 0, \ldots, 0)$.

A majority of the previous research (Galar and Kopciuch, 1999; Karcz-Duleba, 2014) was carried out with a simplified version of the function (3),

$$q_1(x) = \exp\left(-a_1 \sum_{k=1}^{n} x_k^2\right) \\ + 2\exp\left(-a_2(x-1)^2 + a_2 \sum_{k=2}^{n} x_k^2\right), \quad (4)$$

presented in Fig. 1(a) for $n = 2$. The global optimum with the height $h = 2$ is shifted in the first coordinate $\mathbf{x}_g = (1, 0, \ldots, 0)$. In this paper, the values of the slopes of the local and global hills, resp. a_1 and a_2, were assumed to be constant and set to $a_1 = a_2 = 5$. In this case, the saddle is not very deep (its minimum height is about $h_s = 0.76$) and its width (defined as the shortest distance between the top of the local hill and a point at the global hill with the same value as at the local optimum) is $\Delta_s = 0.57$. The effect of a different saddle shape and thus the influence of a_2 on the efficiency of crossing the saddle were examined by Karcz-Duleba (2014).

As both the optima of the fitness function q_1 (4) are located along the ridge which stretches along the x-axis, the case is not very difficult for the analyzed evolutionary procedure, which is a path following process (Galar and Chorazyczewski, 2001). Later on, we will study impatience mechanisms also for more complex landscapes where the location of the second hill is shifted from the x-axis and/or the hill itself is transformed (squeezed, stretched, rotated). Three examples of a two-dimensional ($n = 2$) transformed global hill, depicted in Fig. 1(b)–(d) and called q_2, q_3, q_4 respectively, will be examined later on.

Finally, the efficiency of impatience operators will be verified in landscapes of fitness functions with more than two optima. In the paper, three multimodal two-dimensional benchmark functions were exploited (in the maximization version):

• Rastrigin,

$$q_R(\mathbf{x}) = 44.5 - \left[10n + \sum_{i=1}^{n}\left(x_i^2 - 10\cos(2\pi x_i)\right)\right], \quad (5)$$
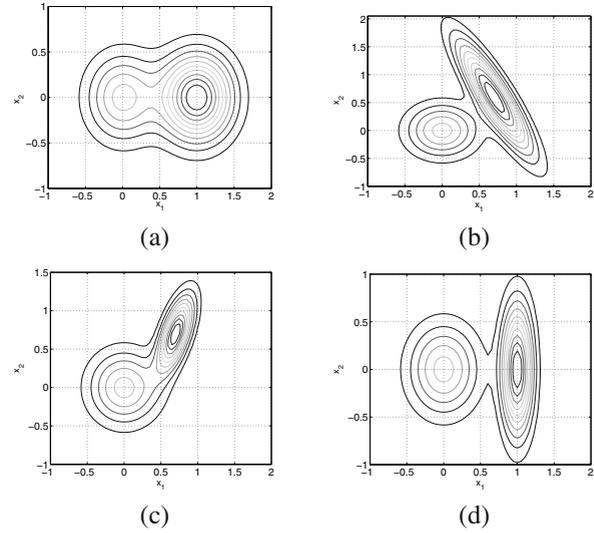
Fig. 1. Maps of the two-dimensional fitness functions q (2) with different locations and shapes of the global optimum. q_1 (4): $\mathbf{x}_g = [1; 0]$, $C = [1, 0; 0, 1]$ (a), q_2: $\mathbf{x}_g = [0.7; 0.7]$, $C = [1, -1.5; -2, 4]$ (b), q_3: $\mathbf{x}_g = [0.7; 0.7]$, $C = [0.25, 0.3; 0.3, 1]$ (c), q_4: $\mathbf{x}_g = [1; 0]$, $C = [0.2, 0; 0, 2]$ (d).

$-1.5 \le x_i \le 1.5$;

• Schwefel,

$$q_S(\mathbf{x}) = 840 + \sum_{i=1}^{n}\left(x_i \cdot \sin(x_i)\sqrt{|x_i|}\right), \quad (6)$$

$-500 \le x_i \le 500$;

• Ackley,

$$q_A(\mathbf{x}) = 20 \cdot \exp\left(-0.2 \cdot \sqrt{\frac{1}{2}\sum_{i=1}^{n} x_i^2}\right) \\ + \exp\left(\frac{1}{2}\sum_{i=1}^{n}\cos(2\pi \cdot x_i)\right) + e, \quad (7)$$

$-3 \le x_i \le 3$, where e signifies the base of the natural logarithm ($e = \exp(1)$).

The functions have optima of different heights and locations, thus the mechanism of crossing saddles and searching for global optima can be shown differently.

In the following sections, all presented computations and simulations were performed using the author's personal code written in the Matlab software and run on a PC computer (Intel Core 2, 2.66 GHz).

3. Dynamics of evolution with impatience

3.1. Impatience: Polarization of population. The operator of impatience decreases the quality of individuals

close to the population mean and increases the quality of those located at peripheries of the population. As the mechanism is applied when a population is located in the neighborhood of a local optimum, its influence on fitness is analyzed for a local component of the function (4) only. Sample transformations of one- and two-dimensional local optimum functions by the impatience operator (1) are presented in Fig. 2. Fitness values of individuals located far away from the population mean increase significantly and, which is more interesting, the function is not symmetrical: the farther the population mean from the optimum, the greater the values of the transformed fitness (see the plot of q_a for $d_A = 0.5$ in Fig. 2(a)).

(a)

(b)
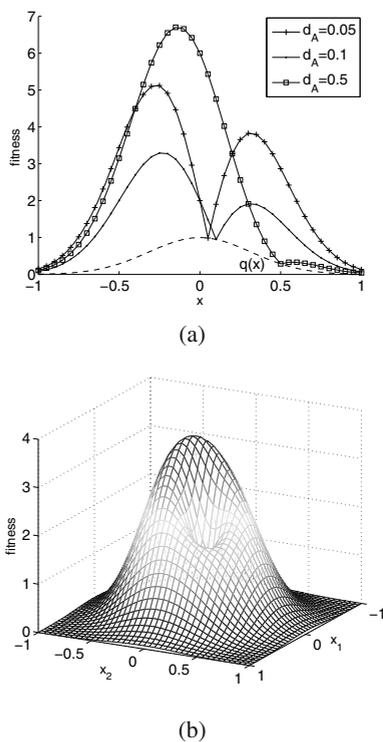
Fig. 2. Transformations of the local optimum component of the fitness q (4) by the operator of impatience (1): one-dimensional function for three population mean locations $d_A = 0.05, 0.1, 0.5$ (q is marked with a dotted line) (a), two-dimensional function for population mean location $d_A = 0.1$ (b).

In Fig. 3 the effect of the operator of impatience on evolution of a 32-element population in the landscape of the two-dimensional function q_1 (4) is illustrated. (Note that the fitness contour plots in Figs. 3, 5 and 7 correspond to original (unmodified) functions.) The simulation was started with an initial population of clones of locally optimal individuals. The operator of impatience was turned on after five generations, to allow preliminary diversification of clones. Such a population imitates the state of quasi-stability around the local optimum.

The distant (from the mean) individuals reproduce

(a)

(b)

(c)
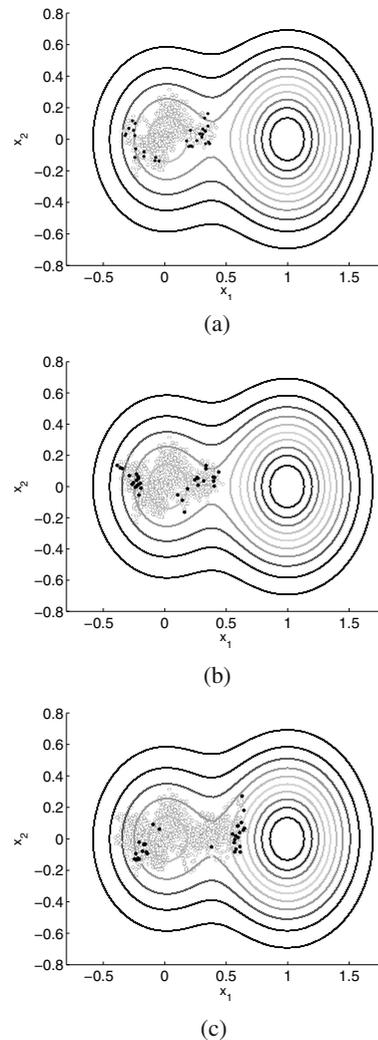
Fig. 3. Stages of saddle crossing by a population under the action of the impatience operator (1) in a landscape of the two-dimensional function q_1 (4): after generation no. 30 (a), after generation no. 40 (b), after generation no. 60 (c); $m = 32$, $\sigma = 0.025$. Polarization of population is clearly visible.

more intensively and the population diversity increases. A too large dispersion of a population is counteracted by a decrease in the fitness of far-away-from-optimum individuals subjected to selection. Consequently, a population is distributed on some kind of orbit around the population mean, where both the trends are balanced and the orbit acts as an attractor for a population fluctuating around an optimum. Additionally, some locations on the orbit are favored (two optima in Fig. 2(a)). As a result, the population polarizes and spontaneously breaks into two clustered sub-populations spawning a kind of dipole (Fig. 3(a)). The dipole fluctuates along an orbit. The number of individuals that belong to each sub-population changes dynamically.

When a local hill is separated from another optimum by a saddle, a dipole may be attracted by the saddle where an adaptation is slightly higher than on other parts of the orbit, and directed its axis along the saddle (Fig. 3(b)). In result, the saddle can be crossed faster by this part of the dipole situated closer to a saddle. Saddle crossing is done by one of the sub-populations, whereas the other remains somewhere around a local hill. Keeping one of sub-populations at a local optimum can be explained when looking at Fig. 2(a) and the function reflecting the population mean $d_A = 0.5$ (the saddle in Fig. 3 is about $\Delta_s \simeq 0.6$ wide). The transformed fitness of individuals on a local hill is very high compared with individuals just crossing saddles, so they stay at a local optimum (Fig. 3(c)). This sub-population slowly becomes extinct as individuals on the global hill reproduce themselves more often and as the impatience operator is switched off after crossing is done.

3.2. Impatience with knowledge. A modification of an individual fitness by the impatience mechanism with knowledge (2) is presented in Fig. 4 for $n = 1$ and $n = 2$. Similarly to Fig. 2, only the first (local) component of Eqn. (4) and, as the transformation Eqn. (2) is symmetrical, only the positive half-plane for the one-dimensional case were visualized. The operator of impatience with knowledge decreases the quality of near optimal individuals and increases that of those located at some distance from the optimum. When a population is closer to a local optimum (i.e., the average distance d_{Ao} is getting small), the transformed quality of some distant individuals highly increases (more than twice). (Note that for impatience without knowledge such an increase is much greater—about five times.) When the population is more dispersed, d_{Ao} is relatively large, the transformed quality of an individual is smaller but still exceeds the local hill quality.

In Fig. 5, the stages of saddle crossing by a population influenced by the operator of impatience with knowledge are presented. The main difference between the operator of impatience with and without knowledge is clearly visible: the population for which the operator of impatience is calculated according to Eqn. (2) is widely diversified but not polarized (cf. Fig. 3). The difference can be explained as follows. Populations that evolved without knowledge (1) have no information about the fitness function configuration. Their knowledge is restricted to information about the actual population position. When the operator with knowledge (2) is active, a population is located on an orbit around the local optimum. All locations on the orbit are equally likely (cf. Fig. 4). Additionally, in evolving populations there is a tendency to concentrate individuals (Karcz-Duleba, 2006), thus a population is not polarized and evolves as a cluster.
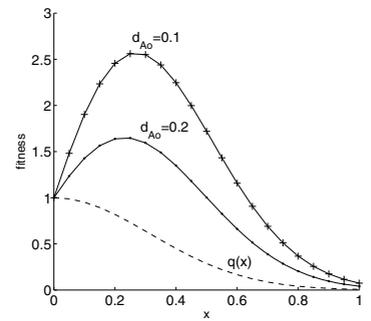
(a)
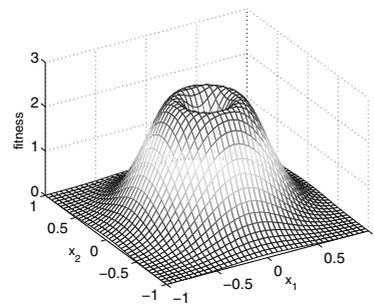
(b)

Fig. 4. Transformations of the local optimum component of the fitness q (4) by the operator of impatience with knowledge (2): one-dimensional fitness for two average distances $d_{Ao} = 0.1, 0.2$ (q is marked with a dotted line) (a), two-dimensional fitness for average distance $d_{Ao} = 0.1$ (b).

In order to understand better the actions of both operators, distances from the local optimum for both the operators were calculated for the mean individual and averaged for all population individuals. Results are presented in Fig. 6. The mean individual of a population evolved with the operator of impatience without knowledge is located close to the optimum until the population crosses a saddle. This means that while sub-populations wander around the local optimum, the center of the dipole is located very close to the optimum. Thus, although a population is diversified, information about the optimum is still present in the population, though not implicitly. When a populations already crossed a saddle, the distance to the local optimum increases and the center of the dipole moves towards the global optimum (after the 50th generation in Fig. 6). The average distance of individuals from the local optimum increases slowly as the population is diffused and one sub-population has crossed a saddle. For the operator of impatience with knowledge, both the distances are very similar as the population is not polarized and evolves as a whole. Thus, a mean individual does not stick to the local optimum but wanders with all individuals within a population.

(a)

(b)

(c)
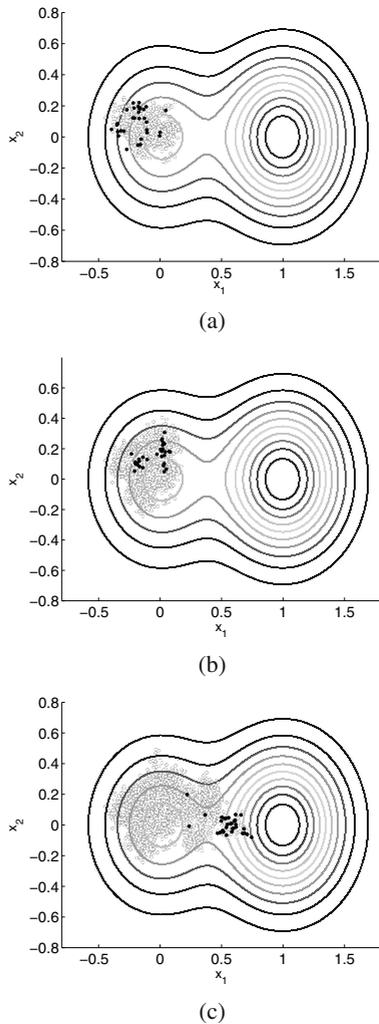
Fig. 5. Stages of saddle crossing by a population subjected to the operator of impatience with knowledge in a landscape of the two-dimensional function q_1 (4): generation no. 30 (a), generation no. 60 (b), generation no. 120 (c); $m = 32$, $n = 2$, $\sigma = 0.025$.

4. Evolution with impatience: A comparative study

4.1. Impatience vs. evolution without impatience.
The comparison of a phenotypic evolution with fitness proportionate selection and the Gaussian mutation with and without impatience mechanisms is presented in Fig. 7 for sample runs of 80 generations. Polarized populations evolved with the impatience mechanism are widely distributed. Furthermore, they crossed a saddle before a population without impatience and with impatience and knowledge. Populations with impatience and knowledge are diversified more than the population without the operator but less than the version without knowledge as a later population is polarized.

In Fig. 8, plots of an average fitness and

(a)

(b)
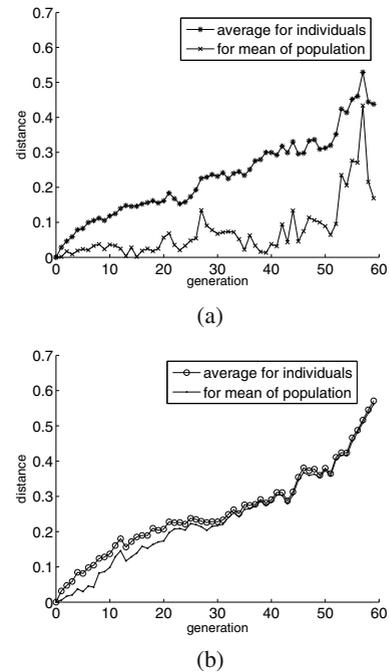
Fig. 6. Average distance of population individuals and the distance of the population mean from the local optimum of the function q_1 (4). Evolution with the operator of impatience: without knowledge (a), with knowledge (b); 60 generations, $m = 32$, $n = 2$, $\sigma = 0.025$.

diversification for populations from Fig. 7 are presented. The diversification of a population was defined as the average distance from the mean individual. In all figures, the operator of impatience and knowledge is marked as 'wK' (with knowledge). The average fitness of a population evolving without impatience mechanisms is changed slightly and remains on a similar level during the entire process, whereas the average fitness of the population with impatience is more reduced (which may help in faster saddle crossing) (Fig. 8(a)). Although the population evolving without knowledge is polarized, its average fitness is only slightly smaller than for the impatience with knowledge version.

In general, diversification of populations evolving with impatience operators is higher than for the population without it (Fig. 8(b)). The average distance of population individuals from the population mean for the operator of impatience without knowledge grows during the presented evolution runs and achieves quite high values. The knowledge incorporated in the impatience operator does not influence diversity as much as polarization. The diversity is higher than in the non-impatience case but less than the without-knowledge case at the beginning of evolution and decreases in later generations, reaching a level of the version without knowledge.

Results presented so far confirm that impatience operators may be classified into diversification preserving

(a)

(b)

(c)
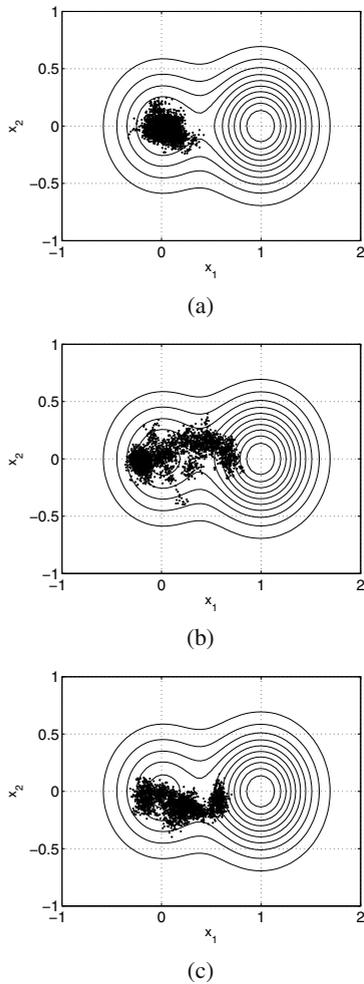
Fig. 7. Evolution without and with operators of impatience in a landscape of the two-dimensional function q_1 (4): without impatience (a), with impatience (b), with impatience and knowledge (c). Generations from no. 1 till the 80th are presented; $m = 32$, $n = 2$, $\sigma = 0.025$.

(a)

(b)

Fig. 8. Average fitness (a) and diversity of populations (b) of evolving populations from Fig. 7.

methods, particularly the version without knowledge about the location of a local optimum. As Figs. 7 and 8 presented only sample results of evolution with fitness proportionate selection and Gaussian mutation, we performed more statistically sound studies of population diversification. The average distance from the mean individual for populations evolving with both versions of the impatience mechanism were computed and compared with the version without impatience for two-, four- and eight-dimensional functions q_1 (4). Simulations were carried out by 100 generations or, if a population had already crossed a saddle, were stopped just after crossing a saddle by the population mean. Runs were repeated 100 times and averaged, and results are presented in Fig. 9. As some simulations could last less than 100 generations, this fact was taken into account when calculating the means.

Populations evolved with the operator of impatience are polarized, so they are much more dispersed than those with impatience and knowledge and populations without impatience (Fig. 9). Although populations in higher dimensions ($n = 4, 8$) are more dispersed, the impatience mechanism still maintained and enhanced the diversity (Fig. 9(b)–(c)). For the impatience operator with knowledge, the average distance remains on a similar level during evolution on a local hill and slightly surpasses diversity of populations without impatience.

So far, studies of the impatience mechanism have been conducted for fitness proportionate selection. As impatience is a universal operator, it may be also applied to other selection methods. In the paper we chose tournament selection as the one frequently used. Diversification of a population evolving with two-element (binary) tournament selection and the Gaussian mutation is given in Fig. 10. Results are very similar to those obtained for fitness proportionate selection: the biggest diversity is exhibited by the population evolving with the impatience without knowledge operator (as the population is polarized), the smallest one by the population evolving without impatience.

In Table 1, average generations of saddle crossing for 32-element populations evolving with impatience operators and without impatience in the landscape of the

(a)

(b)

(c)
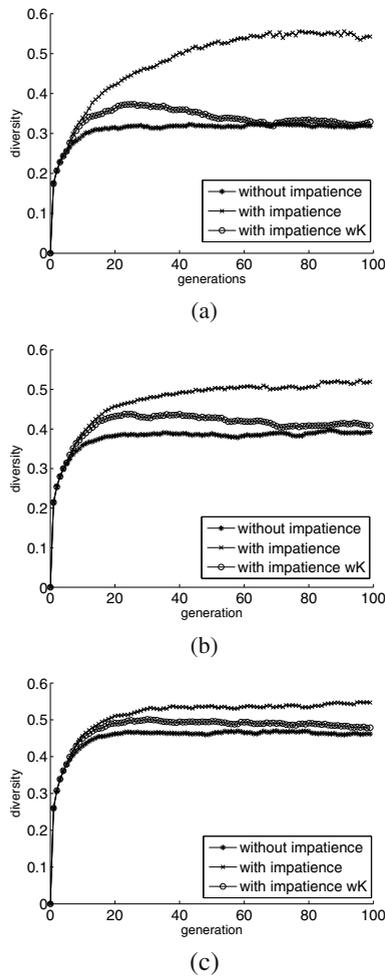
Fig. 9. Diversification of population for populations evolving without and with the operators of impatience (average for 100 runs). Fitness proportionate selection: $n = 2$ (a), $n = 4$ (b), $n = 8$ (c); $m = 32$, $\sigma = 0.025$.

Fig. 10. Diversification of population for populations evolving without and with the operators of impatience (average for 100 runs). Tournament selection (of size two): $n = 2$, $m = 32$, $\sigma = 0.025$.

Table 1. Average number of generations to cross a saddle for populations evolving with and without impatience in a landscape of the fitness function q_1 (4). Fitness proportionate and tournament selections, averaged from 1000 runs; $m = 32$, $n = 2$ (in brackets the number of unsuccessful runs).

fitness proportionate selection					
	without		impatience		impat. wK
σ	mean	std	mean	std	mean	std
0.025	364	324	87	41	81	55
0.05	113	76	40	22	32	18
0.1	28	19	17	7	17	9

tournament selection					
	without		impatience		impat. wK
σ	mean	std	mean	std	mean	std
0.025	4448 (44)	3322	54	17	35	12
0.05	3357 (8)	3082	29	9	22	6
0.1	63	48	19	7	16	5

fitness function q_1 (4) are presented for proportionate (see also Fig. 13(a)) and tournament selection. Simulations were performed for $\sigma = 0.025, 0.05$ and 0.1, and averaged over 1000 runs. In the case of tournament selection and $\sigma = 0.025, 0.05$, runs were terminated after 10000 generations without a crossing saddle. For $\sigma = 0.1$, the saddle is about six σ wide ($\Delta_s \simeq 0.6$) and crossing saddle is fast without any accelerating method, even though the impatience may yet speed it up. For smaller σ, the width of a saddle is about 11σ and 23σ, respectively, and the dominance of impatience operators in crossing the saddle is unquestionable. Populations evolving with tournament selection and without impatience crossed the saddle with some difficulty. Some runs do not escape from local optima even in 10000 generations (numbers in brackets). Note that fitness proportionate selection crosses such wide saddles much easier. Thus, acceleration caused by the impatience mechanism is significant. The efficiency in

crossing saddles of both impatience versions with both selection methods is very similar.

4.2. Impatience vs. fitness sharing.

In order to compare impatience operators with other diversity preserving methods, fitness sharing was added into the population evolving with fitness proportionate selection and the Gaussian mutation. We use the most popular version of fitness sharing (Sareni and Krähenbühl, 1998), with parameter $\alpha = 1$ and different niche radius ρ. The niche radius was chosen relative to mutation range σ: $\rho = (1, 5, 10) \cdot \sigma$.

In Fig. 11, the average diversity and fitness of populations evolving in the landscape of the function q_1 (4) are shown, averaged over 100 runs (only impatience without knowledge was regarded). The fitness sharing method increases population diversity by increasing a niche radius—for small ρ, diversity is much smaller than for a population with impatience. The diversity of

population for both the methods achieved the same level for the largest ρ. The average fitnesses have similar values for both the methods, the highest values are obtained for fitness sharing with the smallest $\rho = \sigma$ as the population is more compact.
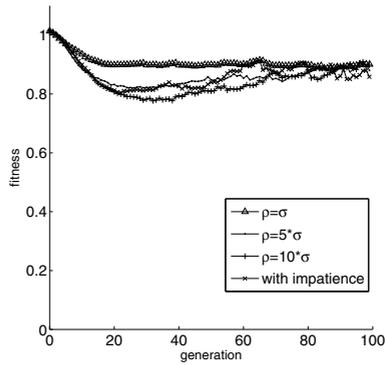
(a)
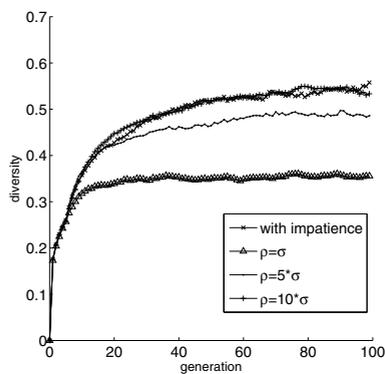
(b)

Fig. 11. Average fitness (a) and diversity (b) of populations evolving with impatience and with fitness sharing for different niche radii $\rho = (1, 5, 10) \cdot \sigma$ in the landscape of the function q_1 (4); $m = 32$, $n = 2$, $\sigma = 0.025$.

Although fitness sharing may increase diversity, it is not very useful for the main purpose we are interested in—searching for a global optimum by saddle crossing. The impatience operator greatly accelerates saddle crossing whereas fitness sharing, in some cases, even disturbs crossing saddles (for small ρ). In Table 2, the average numbers of generations to cross saddles for populations evolving with fitness sharing are collected for different niche radii ρ. As mechanisms supporting saddle crossing are needed when a population is trapped at a local optimum, only small values of mutation range $\sigma = 0.025, 0.05$ were taken into account. Fitness sharing accelerates crossing saddles mainly for larger value of ρ.

The main idea of fitness sharing consists in modifying the search landscape by reducing the payoff in densely populated regions. In contrast to the

Table 2. Average number of generations to cross a saddle for populations evolving with fitness sharing. Results averaged over 1000 runs; $m = 32$, $n = 2$ (in brackets the number of unsuccessful runs).

	σ			
	0.025		0.05	
ρ	mean	std	mean	std
σ	440 (377)	265	78	48
5σ	140	72	37	12
10σ	111	42	38	11

impatience operator, sharing decreases the fitness of similar individuals within the population. Sample fitness values of a randomly generated population in a landscape of the one-dimensional local optimum of the function q_1 (4) and under impatience and fitness sharing (with different ρ) mechanisms are presented in Fig. 12, together with the average fitness levels of the whole populations. According to the sharing idea, the niche radius ρ should have a rather small value compared with the whole population diameter, which is usually about several σ (Karcz-Duleba, 2006). In Fig. 12, when applying fitness sharing with $\rho = 0.025$, only a few (most crowded) individuals changed their fitness and the average population fitness is only slightly smaller than the average fitness of the original population. If ρ takes larger values, almost all population individuals may be regarded as similar and their fitness is greatly reduced. Fitness reduction may help in crossing saddles (which, in this case, have the minimum height of about 0.7), and this is confirmed by results presented in Table 2 for bigger ρ. Thus, fitness sharing may be comparable with an impatient operator (see Table 1) only when the values of parameter ρ are selected properly (whereas the impatience mechanism is parameter-less), and then saddle crossing occurs rather as a side effect of the average fitness decrease.

5. Efficiency of crossing saddles

5.1. Crossing saddles in bimodal fitness functions landscapes.
The efficiency of crossing a saddle by a population subjected to the operator of impatience with and without knowledge was studied by Galar and Kopciuch (1999) as well as Karcz-Duleba (2014) for the bimodal fitness function q_1 (4). The influence of the following parameters were tested: population size m, search space dimension n, saddle width and depth (depending on a slope, a_2, of the global optimum). Results were related to evolution with fitness proportionate selection and the Gaussian mutation without the impatience operator (see also Section 4.1).

An evolution exploiting the impatience operator outperformed its impatience-free version in a wide range
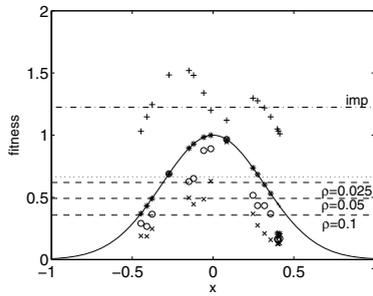
Fig. 12. Comparison of the fitness of individuals under impatience vs. fitness sharing mechanisms in a landscape of the one-dimensional local optimum of the function q_1 (4). Population of 32 individuals, randomly generated in $[-1, 1]$, marked as dots. Population under operators: impatience $(+)$, fitness sharing with $\rho = 0.025$ $(*)$, $\rho = 0.05$ $(\circ)$, $\rho = 0.1$ (x). Levels of the average fitness of a population: pointed line—original population, dot-point line—with impatience, dotted lines—with fitness sharing.

of parameters of the fitness function and the evolution model. The extra knowledge concerning the location of the local optimum accelerated saddle crossing only slightly. A bigger benefit was obtained for large populations ($m = 32$–128). Interestingly, an enlargement of a population size does not increase significantly the number of fitness function calculations needed to cross a saddle.

The efficiency of the three evolution models (without impatience, with impatience and with impatience and knowledge) in saddle crossing was similar for small search space dimensionalities (from $n = 2$ to 8), although the model without impatience is slightly worse. As the dimensionallity n grows, the efficiency of model with impatience and knowledge decreases and tends toward the one without impatience. The best performance is displayed by the model with impatience. Populations evolving with the impatience operator crossed effectively large and deep saddles.

Since a rotated global optimum of a bimodal fitness function usually poses some problems for evolution, the impact of the impatience mechanism on crossing more troublesome saddles will be studied now. Saddle crossing was analyzed for the two-dimensional functions q_2–q_4 and compared with the function q_1 (4). Different parameters were considered: population size $m = 8, 16, 32, 64$ and standard deviation of mutation $\sigma = 0.025, 0.05, 0.1$. Simulations were stopped after crossing the saddle by the average individuals (i.e., when the fitness of the average individual exceeds the value of the local optimum) or after 20000 generations without success. The numbers of fitness function evaluation Ev, averaged over 100 runs, are presented in Fig. 13. Let us note the logarithmic scale

on the y-axis.

The functions q_2–q_4 are indeed more difficult for optimization by evolution than q_1 (see different ranges of y-axis in Fig. 13(a) and others). The most troublesome are the functions q_2 and q_4 because their global optima are located crosswise to the local ones, saddles are deep and optima are not joined by a ridge, which generates an impediment to a path dependent search. In the case of the function q_4, larger populations evolved with small σ and without impatience operators crossed the saddle in a few trials over 100 runs only (in 20000 generations).

In all presented results evolution with impatience mechanisms surpasses evolution without it and populations crossed the saddle even in the most troublesome cases. The efficiency of both impatience mechanisms seems to be comparable for functions q_1 and q_3. Impatience with a knowledge operator is significantly better than without knowledge version for both troublesome functions q_2 and q_4, according to *Student's t-test* with the level of significance $\alpha = 0.05$. For the function q_3, impatience with knowledge is significantly better for bigger values of the standard deviation of mutation and larger populations ($m \geq 16$ and $\sigma \geq 0.05$). Thus, the presented results confirmed that impatience operators actually accelerate crossing for both easy and difficult saddle configurations. It is worth noting that small populations ($m = 8$) are the most effective.

5.2. Multioptima fitness functions.

The aforementioned examples of fitness functions exemplified two optima separated by distinct saddles. The efficiency of the impatient mechanism was also evaluated for fitness functions with more than two optima. In the work of Karcz-Duleba (2014) optimization of the multimodal Ursem 03 test function was presented. The results were promising; only about 0.07% of 100 runs failed to find a global optimum in 5000 generations. The impatience operator without knowledge was a bit more efficient than that with knowledge.

In this paper, multimodal two-dimensional Rastrigin (5), Schwefel (6) and Ackley (7) functions were selected to analyze the efficiency of impatience mechanisms. Operators of impatience were turned on when a population, evolving with fitness proportionate or tournament selection and Gaussian mutation, did not show any improvement in 10 or 15 (depending on the function) consecutive generations. This was regarded as an indicator that a population got stuck at a local optimum. The operators were turned off when a population found a better solution than the currently best. For the impatience with knowledge operator, local search procedures were not used for finding the exact value of the local optimum but the current best fitness was treated as the estimate of it.

Simulations were started with initial populations

Table 3. Average number of generations to find the global optimum for the Rastrigin function (5). Fitness proportionate and tournament selections; $\sigma = 0.06$ (in brackets the number of unsuccessful runs).

	fitness proportionate selection					
	without		impatience		impatience wK	
m	mean	std	mean	std	mean	std
8	638	699	384	395	581	603
16	630	563	343	397	793	1233
32	1625 (10)	2002	1326 (8)	2057	1841 (7)	2212

	tournament selection					
	without		impatience		impatience wK	
m	mean	std	mean	std	mean	std
8	1855	1702	826	721	2502 (19)	2625
16	4583 (54)	2704	710	768	1358 (32)	2059
32	3235 (97)	3613	852 (14)	1584	1307(41)	2017

Table 4. Average number of generations to find the global optimum for the Schwefel function (6). Fitness proportionate and tournament selections; $\sigma = 20$ (in brackets the number of unsuccessful runs).

	fitness proportionate selection					
	without		impatience		impatience wK	
m	mean	std	mean	std	mean	std
8	724	614	481	382	1005 (4)	909
16	881	678	443	224	826 (3)	904
32	2073 (2)	1454	552 (4)	1029	989 (7)	1210

	tournament selection					
	without		impatience		impatience wK	
m	mean	std	mean	std	mean	std
8	1464	1238	442	279	1079 (1)	1372
16	2884 (1)	1861	476	316	1249 (1)	1655
32	5140 (90)	2951	547 (1)	627	1977 (19)	2454

located on the selected local optima far away from the global one $((-1, -1)$ for Rastrigin, $(-300, -300)$ for Schwefel and $(-2, -2)$ for Ackley functions). Initial populations were normally distributed (with a small standard deviation) on the local hills so as to simulate a selection-mutation dynamic equilibrium around the local optimum. Computations were stopped when any individual found a global optimum (with the accuracy about 2–5%) or after 10000 generations without any success and repeated 100 times. The value of the standard deviation of mutation σ was related to the function ranges as 1/50 of their values. Although the Ackley function was included into very troublesome benchmark functions according to CEC05 outcomes (Hansen, 2005), the results obtained showed that it is not difficult for populations evolving with proportionate or tournament selection and the Gaussian mutation. In the case of tournament selection and the standard deviation of mutation equal to 1/50 of the function range ($\sigma = 0.12$), populations found the global optimum very fast, so there was no need for additional operators. Thus, as to check the effectiveness of impatience mechanisms, a more difficult case with a smaller standard deviation of mutation $\sigma = 0.06$

(1/100 of the function range) was considered. Results of a global optima search with and without impatience operators are presented in Tables 3–5. Only successful runs were averaged. The number of unsuccessful runs is given in brackets in the tables.

The presented results confirm very good properties of the impatience without knowledge operator in crossing saddles and finding the global optimum of multimodal functions. The mechanism accelerates the search significantly in almost all the cases studied. For Rastrigin and Schwefel functions, impatience gives the best results, with the efficiency decreasing as the population size increases. For Ackley function, which is, as we mention below, easy for the tested methods, the impatience operator can still accelerate the search. A considerable improvement is obtained in the case of tournament selection and a small standard deviation of mutation $\sigma = 0.06$ (for $m = 16$, impatience with knowledge gave the best results but it managed to obtain the global optimum only in 16% of runs). The efficiency of the operator of impatience with knowledge strongly depends on the shape of an optimized function. In most cases, it is more effective than evolution without impatience but worse than

(a)

(b)

(c)

(d)
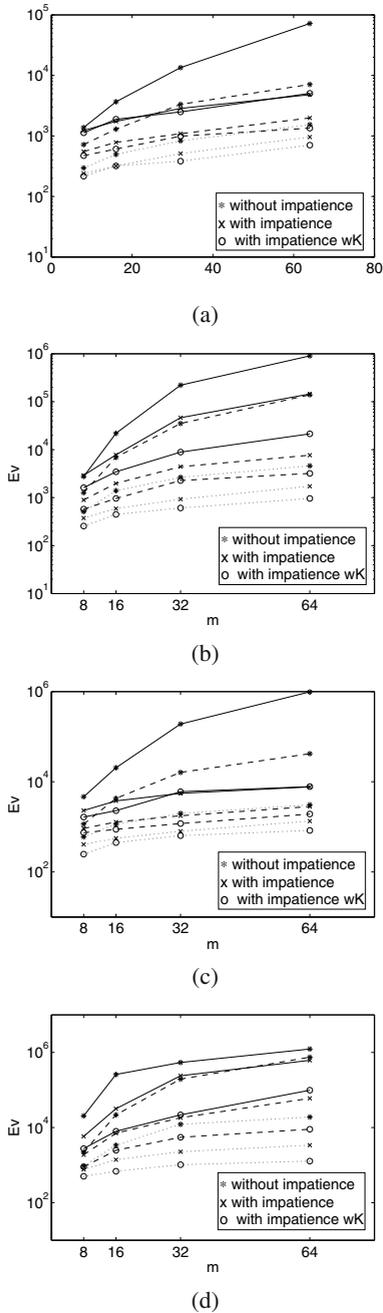
Fig. 13. Efficiency of crossing a saddle. Averaged over 100 runs, the number of fitness function evaluation Ev for the two-dimensional ($n = 2$) functions q_1–q_4 (see Fig. 1) and evolution with and without impatience; with regard to the population size of $m = 8, 16, 32, 64$ and the standard deviation of mutation: $\sigma = 0.025$ (solid line), $\sigma = 0.05$ (dashed line), $\sigma = 0.1$ (dotted lines).

the impatience version without knowledge.

For the Rastrigin function, the with-knowledge version works better than evolution without impatience

for a small population size ($m = 8$) when evolving with fitness proportionate selection and for larger populations ($m = 16, 32$) when evolving with tournament selection. For the Schwefel function, the with-knowledge version is worse than that without impatience only for a small population size ($m = 8$) and fitness proportionate selection. For the Ackley function and tournament selection, the operator is effective when populations are larger. The results may confirm our earlier observations that the impatience and knowledge mechanism behaves better for unsymmetrical functions, but it is less effective for separable ones, with small differences in optima values and/or quite deep saddles (as Rastrigin or Ackley). In all simulations, the standard deviations are high, so the results are spread out: some runs are very quick and some very slow. In general, small populations evolving with proportionate selection are the most effective. The larger the population, the more likely it is to get stuck in a local optimum, and 10000 generations would not be enough to find the global one. Tournament selection is less effective than fitness proportionate selection for Rastrigin and Schwefel functions and very efective for Ackley function.

6. Conclusions

Results presented in this paper confirmed very good properties of the impatience mechanism as a strategy of maintaining diversity and facilitating saddle crossing. The mechanism supports an effective search for global optima of both bimodal (of different topologies) and multimodal fitness functions. Impatience without knowledge significantly increases the diversity of a population by its polarization and accelerates finding global optima of multimodal functions. Impatience with knowledge about a local optimum shows good properties in landscapes of bimodal fitness functions with a difficult configuration of a global optimum, but it is less effective for some multimodal functions. Thus the latter version may be effective in demanding landscapes having a few optima not symmetrically and sparse in the search space.

A big advantage of the impatience mechanism is the lack of user defined parameters. The operator requires only slightly more additional computations and may supplement different optimization algorithms. In the reported research, the impatience mechanism supports simple evolutionary models (fitness proportionate or tournament selection and the Gaussian mutation). Preliminary attempts of applying the mechanism to other evolutionary procedures (differential evolution) are promising. Additionally, as the mechanism deals with a fitness function only (not interfering with an optimization algorithm itself), it can be useful when integrating the mechanism with existing optimization software supporting different algorithms (only a minor

Table 5. Average number of generations to find the global optimum for the Ackley function (7). Fitness proportionate and tournament selections (in brackets the number of unsuccessful runs).

fitness proportionate selection, $\sigma = 0.12$					
	without		impatience		impatience wK	
m	mean	std	mean	std	mean	std
8	229	193	225	199	1126	1436
16	127	85	137	84	612	793
32	94	43	87	37	456	657

tournament selection, $\sigma = 0.06$					
	without		impatience		impatience wK	
m	mean	std	mean	std	mean	std
8	528	299	315	277	2399 (72)	2461
16	1383	891	784	1242	713 (84)	1301
32	4748 (11)	2524	417 (19)	1102	2057 (69)	3188

modification of the fitness is required). The results confirmed also our earlier observations about the high efficiency of small populations in global optimization (Karcz-Duleba, 2006).

References

Barabasz, B., Gajda-Zagórska, E., Migórski, S., Paszyński, M., Schaefer, R. and Smołka, M. (2014). A hybrid algorithm for solving inverse problems in elasticity, *International Journal of Applied Mathematics and Computer Science* **24**(4): 865–886, DOI: 10.2478/amcs-2014-0064.

Chen, T., He, J., Chen, G. and Yao, X. (2010). Choosing selection pressure for wide-gap problems, *Theoretical Computer Science* **411**(6): 926–934.

Chorazyczewski, A. and Galar, R. (1998). Visualization of evolutionary adaptation in R^n, *in* V. Porto *et al.* (Eds.), *Evolutionary Programming VII*, Springer-Verlag, London, pp. 657–668.

DeJong, K. (1975). *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*, Ph.D. thesis, University of Michigan, Ann Arbour, MI.

Dick, G. and Whigham, P.A. (2006). Spatially-structured evolutionary algorithms and sharing: Do they mix?, *in* Wang *et al.* (Eds.), *SEAL 2006*, Lecture Notes in Computer Science, Vol. 4247, Springer-Verlag, Berlin/Heidelberg, pp. 457–464.

Galar, R. (1989). Evolutionary search with soft selection, *Biological Cybernetics* **60**(5): 357–364.

Galar, R. and Chorazyczewski, A. (2001). Evolutionary dynamics in space of states, *Congress on Evolutionary Computation, Seoul, Korea*, pp. 1366–1373.

Galar, R. and Kopciuch, P. (1999). Impatience and polarization in evolutionary processes, *Proceedings of the KAEiOG Conference, Potok Złoty, Poland*, pp. 115–122, (in Polish).

Goldberg, D. and Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms, *in* G. Rawlins (Ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann, San Mateo, CA, pp. 69–93.

Goldberg, D. and Richardson, J. (1987). Genetic algorithms with sharing for multi-modal function optimisation, *Proceedings of the 2nd International Conference on Genetic Algorithms and Their Applications, Hillsdale, NJ, USA*, pp. 41–49.

Grosan, C. and Abraham, A. (2007). Hybrid evolutionary algorithms: Methodologies, architectures, and reviews, *in* A. Abraham *et al.* (Eds.), *Hybrid Evolutionary Algorithms*, Studies in Computational Intelligence, Vol. 75, Springer, Berlin/Heidelberg, pp. 1–17.

Hansen, N. (2005). Compilation of results on the 2005 CEC benchmark function set, http:www.lri.fr/~hansen/cec2005compareresults.pdf.

Karcz-Dulęba, I. (2004). Asymptotic behavior of discrete dynamical system generated by simple evolutionary process, *International Journal of Applied Mathematics and Computer Science* **14**(1): 79–90.

Karcz-Duleba, I. (2006). Dynamics of two-element populations in the space of population states, *IEEE Transactions on Evolutionary Computation* **10**(2): 199–209.

Karcz-Duleba, I. (2014). Impatience mechanism in saddles' crossing, *Proceedings of the 6th International Conference ECTA, Rome, Italy*, Vol. 75 pp. 174–184.

Kowalczuk, Z. and Białaszewski, T. (2006). Niching mechanisms in evolutionary computations, *International Journal of Applied Mathematics and Computer Science* **16**(1): 59–84.

Mengshoel, O. and Goldberg, D. (2008). The crowding approach to niching in genetic algorithms, *Evolutionary Computation* **16**(3): 315–354.

Obuchowicz, A. (1997). The evolutionary search with soft selection and deterioration of the objective function, *Proceedings of the 6th International Conference on Intelligent Information Systems, Zakopane, Poland*, pp. 288–295.

Obuchowicz, A. and Prętki, P. (2004). Phenotypic evolution with a mutation based on symmetric α-stable distributions, *International Journal of Applied Mathematics and Computer Science* **14**(3): 289–316.

Rogers, A. and Prügel-Bennett, A. (1999). Genetic drift in genetic algorithm selection schemes, *IEEE Transactions on Evolutionary Computation* **3**(4): 298–303.

Sareni, B. and Krähenbühl, L. (1998). Fitness sharing and niching methods revisited, *IEEE Transactions on Evolutionary Computation* **2**(3): 97–106.

Tomassini, M. (2005). *Spatially Structured Evolutionary Algorithms*, Springer-Verlag New York, Inc., Secaucus, NJ.

Torn, A. and Zilinskas, A. (1989). *Global Optimization*, Springer-Verlag, New York, NY.

Iwona Karcz-Duleba received her M.Sc. and Ph.D. degrees from the Faculty of Electronics of the Wrocław University of Science and Technology in 1988 and 1992, respectively. In 2010 she received a D.Sc. degree. Currently, she is a professor at the Wrocław University of Science and Technology, Faculty of Electronics, where she is the head of the Control, Modelling and Mechatronics Group. Her scientific activities have been concerned with optimization, computational intelligence methods, modern heuristics, modeling and simulation of dynamical systems.