

## POROVNÁNÍ HEURISTIK V ALGORITMECH GLOBÁLNÍ OPTIMALIZACE

JOSEF TVRDÍK

ABSTRACT. The paper deals with the experimental comparison of heuristics used in the evolutionary algorithms (EAs) of the global optimization over continuous search space. Results of experiment with fifty heuristics of four classes in two instances of the EAs on three test functions were scanned and analysed by statistical tools. The aim of the paper is to propose a classification of heuristics according to their properties when the heuristics is used alone in the EAs. This classification is needed for the design of the EAs with alternating, competing or cooperating heuristics. Some comments on the process of statistical data analysis in the case when the analyst and the client is the identical person are added.

Резюме. В этой работе рассматриваются эвристики которые используются в эволюционных алгоритмах глобальной оптимизации. Данные из экспериментов, в которых пятьдесят эвристик четырех классов применено в двух алгоритмах на трех функциях, были статистически анализированы. В работе предложена классификация нужна для эволюционных алгоритмов с соревнованием или кооперацией эвристик. Записки о взаимодействии аналитика и клиента в процессе обработки данных когда они одно лицо тоже добавлены.

### 1. ÚVOD

Problém globální optimalizace lze stručně formulovat takto: pro danou účelovou funkci  $f : D \rightarrow \mathcal{R}$ ,  $D \subset \mathcal{R}^d$ , má být nalezen bod  $\mathbf{x}^*$  tak, aby  $\mathbf{x}^* = \arg \min_{\mathbf{x} \in D} f(\mathbf{x})$ . Bod  $\mathbf{x}^*$  se nazývá globální minimum a  $D$  je oblast prohledávání. K nalezení minima multimodálních funkcí se v posledních desetiletích poměrně úspěšně užívají evoluční algoritmy [1], [6], [13]. Omezíme se na úlohy, kdy hledáme globální minimum v souvislé oblasti  $D = \prod_{i=1}^d \langle a_i, b_i \rangle$ ,  $a_i < b_i$ ,  $i = 1, 2, \dots, d$  a účelovou funkci  $f(\mathbf{x})$  umíme vyhodnotit s požadovanou přesností v každém bodě  $\mathbf{x} \in D$ . Heuristikou v evolučním algoritmu rozumíme nějaké nedeterministické pravidlo, které z bodů populace  $\mathcal{P}$  vytvoří bod  $\mathbf{y} \in D$ . Pokud tento bod  $\mathbf{y}$  „zlepšuje“ populaci, je zařazen do vyhledávacího procesu.

Při návrhu algoritmů s více heuristikami [16], [17], [18] se zjistilo, že je potřeba znát vlastnosti heuristik, aby v těchto algoritmech bylo možno nabídnout vhodnou směs odlišných heuristik a implementovat vhodná pravidla pro jejich soutěž či spolupráci. K nalezení vlastností heuristik byl navržen experiment, v němž se porovnává chování heuristik ve dvou algoritmech třídy popsané v [9] na třech funkcích různé obtížnosti. Vzhledem k tomu, že klient byl sám sobě statistikem, nehrozily obecně známé komunikační problémy mezi statistikem a klientem zmiňované např.

---

2000 *Mathematics Subject Classification.* 65K99 92D99 90C27.

*Klíčová slova.* Globální optimalizace, evoluční algoritmy, řízené náhodné prohledávání, srovnání heuristik, analýza experimentálních dat.

Práce byla podporována z grantu 402/00/1165 Grantové agentury ČR a z institucionálního záměru CEZ: J09/98:179000002.

v [7], [14]. Jak experimentálně zjištěné vlastnosti heuristik, tak i průběh analýzy dat byly však velmi odlišné od představ v době navrhování experimentu. Stručné postřehy o „spolupráci“ klienta a statistika jsou psány v textu *kurzivou*.

*Návrh experimentu připadal zpočátku zcela jasný. Chování heuristik se očekávalo natolik podobné, že k odlišení heuristik budou potřebné konfirmační metody statistické analýzy (lineární model, ANOVA). Polovina experimentálních dat byla k dispozici více než půl roku před dokončením práce, takže ani obvyklá termínová krize nehrozila. Ale později se ukázalo, že Parkinsonovy zákony platí a práce zabere všechnen čas, který je k dispozici.*

## 2. EVOLUČNÍ ALGORITMUS BEZ EXPLICITNÍ MUTACE

V této práci byl použit evoluční algoritmus patřící do třídy algoritmů popsané ve sborníku minulého Robustu [9], kde jsou i uvedeny podmínky asymptotické konvergence této třídy algoritmů. Zde užíváme variantu algoritmu bez explicitní mutace, kterou můžeme zapsat následujícím způsobem:

### procedure EA

generate  $\mathcal{P}$  (an old population of  $N$  points taken at random from  $D$ )

### repeat

find the worst point in  $\mathcal{P}$ ,  $\mathbf{x}_{\text{worst}}$ , with the highest value of  $f$   
copy  $M$  best points of  $\mathcal{P}$  into new population  $\mathcal{Q}$ ,  $1 \leq M < N$

### repeat

#### repeat

generate a new trial point  $\mathbf{y}$  by some heuristics applied to  $\mathcal{P}$

until  $f(\mathbf{y}) < f(\mathbf{x}_{\text{worst}})$

insert the next trial point into  $\mathcal{Q}$

until  $\mathcal{Q}$  is completed to  $N$  points

replace  $\mathcal{P}$  by  $\mathcal{Q}$

until stopping condition is true

Heuristika užitá ve vnitřním cyklu tohoto algoritmu je obecně libovolné nedeterministické pravidlo generující nový bod  $\mathbf{y} \in D$ , většinou tím míníme postup užívající evoluční operátory [1], který z bodů v generaci  $\mathcal{P}$  vygeneruje nový bod  $\mathbf{y} \in D$ . Pokud nastavíme vstupní parametr  $M = N - 1$ , dostaneme zobecnění algoritmu řízeného náhodného prohledávání (controlled random search, CRS), který navrhl Price [11]. Jako druhá varianta testovaného algoritmu bylo zvoleno nastavení parametru  $M = N/2$ . Tato varianta je v dalším textu označena jako HLF.

## 3. HEURISTIKY

Heuristiky porovnávané v tomto experimentu jsou převzaty nebo odvozeny ze známých stochastických algoritmů globální optimalizace, a to z evoluční strategie (ES) [1], [6], reflexe simplexu (REFL) [11], [5], diferenciální evoluce (DE) [12], [15] a šlechtitelských genetických algoritmů, (breeder genetic algorithm, BGA) [2], [10].

Heuristiky inspirované evoluční strategií generují nový bod  $\mathbf{y}$  podle následujícího pravidla

$$y_i = x_i^{\text{best}} + Y, \quad i = 1, 2, \dots, d,$$

kde  $x_i^{\text{best}}$  je  $i$ -tá souřadnice bodu  $\mathbf{x}^{\text{best}}$ , tj. bodu s nejmenší funkční hodnotou v populaci  $\mathcal{P}$ , a  $Y$  je náhodná veličina,  $Y \sim N(0, \sigma_i^2)$ . Zatímco v evoluční strategii je hodnota parametru  $\sigma_i^2$  adaptována v průběhu procesu vyhledávání podle tzv. pravidla 1/5 úspěchu [6], zde je hodnota tohoto parametru odvozována z aktuální populace.

V heuristice dále označované jako *esbest-pop* je hodnota směrodatné odchylky  $\sigma_i$  určována z celé populace jako

$$\sigma_i = c \left( \max_{\mathbf{x} \in \mathcal{P}} x_i - \min_{\mathbf{x} \in \mathcal{P}} x_i \right) + \varepsilon, \quad i = 1, 2, \dots, d,$$

kde  $\varepsilon$  (v implementaci  $\varepsilon = 1 \times 10^{-4}$ ) zabezpečuje, že hodnota  $\sigma_i$  je kladná,  $c$  je vstupní řídicí parametr heuristiky,  $c > 0$ , v tabulce 1 je uveden ve sloupci  $p_1$ . V heuristice označované *esbest-2pts* je směrodatná odchylka  $\sigma_i$  určována ze dvou náhodně vybraných bodů  $\mathbf{r}$ ,  $\mathbf{s}$  populace  $\mathcal{P}$  při vynechání nejlepšího bodu  $\mathbf{x}^{\text{best}}$

$$\sigma_i = c |r_i - s_i| + \varepsilon, \quad i = 1, 2, \dots, d,$$

kde  $c$  je opět vstupní parametr heuristiky, ostatní symboly mají stejný význam jako výše.

Znáhodněná reflexe simplexu [5] náhodně vybraného z populace  $P$  je popsána vztahem

$$\mathbf{y} = \mathbf{g} + U(\mathbf{g} - \mathbf{x}),$$

kde  $\mathbf{x}$  je bod simplexu s nejvyšší funkční hodnotou (v heuristice označované zkratkou *refl-worst*) nebo náhodně vybraný bod simplexu (v heuristice *refl-rand*),  $\mathbf{g}$  je těžiště zbývajících  $d$  bodů simplexu a  $U$  je spojitá náhodná veličina s rovnoměrným rozdělením na intervalu  $(0, \alpha)$ ,  $\alpha$  je vstupní parametr heuristiky, jeho hodnota je v tabulce 1 ve sloupci  $p_1$ . Heuristika označovaná *refl-fix* je reflexe simplexu, kdy  $U$  je konstantní [11].

Z diferenciální evoluce vycházejí dva typy heuristik. Heuristika označovaná dále zkratkou *de-rand* generuje bod  $\mathbf{u}$  vztahem

$$\mathbf{u} = \mathbf{r}_1 + F(\mathbf{r}_2 - \mathbf{r}_3),$$

v němž  $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$  jsou navzájem různé body náhodně vybrané z populace  $\mathcal{P}$ ,  $F$  je vstupní parametr heuristiky,  $F > 0$ , sloupec  $p_1$  v tabulce 1. Heuristika *de-best* je založena na modifikaci bodu  $\mathbf{x}^{\text{best}}$  (s nejmenší funkční hodnotou v aktuální populaci) podle následujícího pravidla

$$\mathbf{u} = \mathbf{x}^{\text{best}} + F(\mathbf{r}_1 + \mathbf{r}_2 - \mathbf{r}_3 - \mathbf{r}_4),$$

$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4$  jsou opět navzájem různé body náhodně vybrané z populace  $\mathcal{P}$  (kromě  $\mathbf{x}^{\text{best}}$ ) a  $F$  je vstupní parametr. Nový vektor  $\mathbf{y}$  vznikne „křížením“ vektoru  $\mathbf{u}$  a náhodně vybraného vektoru  $\mathbf{x}$  tak, že každá složka  $x_i$  je přepsána hodnotou  $u_i$  s pravděpodobností  $C$ ,  $C \in (0, 1)$  je další vstupní parametr heuristiky, jeho hodnoty užití v pokusech jsou v tabulce 1 ve sloupci  $p_2$ .

$$y_i = \begin{cases} u_i & \text{když } U < C \text{ nebo } i = j \\ x_i & \text{jinak} \end{cases}, \quad i = 1, 2, \dots, d,$$

kde  $U$  je náhodná veličina rovnoměrně rozdělená na intervalu  $(0, 1)$  a  $j$  je náhodně vybraný index složky vektoru zabezpečující přepis alespoň jedné složky vektoru  $\mathbf{x}$  při jakékoliv volbě hodnoty  $C$ .

<i>i</i>	Heuristika	$p_1$ $p_2$		DeJong1				Rosenbrock				Ackley			
				CRS		HLF		CRS		HLF		CRS		HLF	
				<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>	<i>R</i>	<i>P</i>
1	<i>bga-box</i>	0.2	0.5	1	0	38	0	1	4	2	2	0	70	0	74
2	<i>bga-box</i>	0.2	1	51	0	86	0	8	1	5	0	0	86	0	70
3	<i>bga-box</i>	0.2	2	100	0	100	0	24	0	44	0	0	84	0	80
4	<i>bga-box</i>	0.5	0.5	44	0	84	0	1	3	1	3	0	68	0	60
5	<i>bga-box</i>	0.5	1	100	0	100	0	6	1	7	1	0	60	0	47
6	<i>bga-box</i>	0.5	2	100	0	100	0	93	0	97	0	0	61	0	50
7	<i>bga-box</i>	0.8	0.5	82	0	98	0	2	6	4	1	0	61	0	47
8	<i>bga-box</i>	0.8	1	100	0	100	0	19	0	12	1	0	41	0	27
9	<i>bga-box</i>	0.8	2	100	0	100	0	96	0	99	0	0	47	0	32
10	<i>bga-line</i>	0.2	0.5	0	42	0	22	0	87	0	82	0	92	0	81
11	<i>bga-line</i>	0.2	1	0	58	0	25	1	83	0	85	0	96	0	96
12	<i>bga-line</i>	0.2	2	0	100	0	94	0	100	0	100	0	100	0	100
13	<i>bga-line</i>	0.5	0.5	0	19	0	9	0	82	0	86	0	84	2	76
14	<i>bga-line</i>	0.5	1	0	54	4	12	1	93	3	80	3	95	6	72
15	<i>bga-line</i>	0.5	2	37	52	72	13	31	69	71	29	4	92	10	78
16	<i>bga-line</i>	0.8	0.5	1	13	4	6	0	87	8	68	5	76	3	56
17	<i>bga-line</i>	0.8	1	10	22	23	9	13	81	33	48	3	81	4	61
18	<i>bga-line</i>	0.8	2	92	5	100	0	100	0	100	0	4	66	8	51
19	<i>de-best</i>	0.5	0	98	2	100	0	1	30	9	33	87	13	95	5
20	<i>de-best</i>	0.5	0.5	100	0	100	0	100	0	91	0	93	7	96	4
21	<i>de-best</i>	0.5	1	100	0	100	0	100	0	99	1	93	7	95	5
22	<i>de-best</i>	0.9	0	97	3	100	0	4	18	22	20	88	12	93	7
23	<i>de-best</i>	0.9	0.5	100	0	100	0	100	0	100	0	96	4	100	0
24	<i>de-best</i>	0.9	1	100	0	100	0	100	0	100	0	100	0	100	0
25	<i>de-rand</i>	0.5	0	57	43	83	17	0	62	0	66	29	71	69	31
26	<i>de-rand</i>	0.5	0.5	96	4	100	0	39	61	42	58	60	40	84	16
27	<i>de-rand</i>	0.5	1	93	7	100	0	57	43	71	29	68	32	92	8
28	<i>de-rand</i>	0.9	0	80	20	77	23	0	67	1	64	42	58	68	32
29	<i>de-rand</i>	0.9	0.5	100	0	100	0	82	18	92	3	75	25	90	10
30	<i>de-rand</i>	0.9	1	100	0	100	0	100	0	100	0	98	2	100	0
31	<i>esbest-2pts</i>	0.2		100	0	61	0	16	0	3	0	0	96	0	92
32	<i>esbest-2pts</i>	0.5		100	0	100	0	72	0	10	0	0	78	0	81
33	<i>esbest-2pts</i>	1		100	0	100	0	100	0	80	0	0	54	0	47
34	<i>esbest-2pts</i>	2		100	0	100	0	100	0	100	0	0	38	0	31
35	<i>esbest-pop</i>	0.1		100	0	100	0	25	0	12	1	0	97	0	92
36	<i>esbest-pop</i>	0.2		100	0	100	0	100	0	53	0	0	32	0	40
37	<i>esbest-pop</i>	0.5		100	0	100	0	48	0	0	0	0	0	0	0
38	<i>esbest-pop</i>	1		61	0	0	0	0	0	0	0	0	0	0	0
39	<i>refl-fix</i>	1		100	0	100	0	82	18	100	0	36	64	75	25
40	<i>refl-fix</i>	2		0	0	0	0	100	0	99	1	14	32	1	4
41	<i>refl-fix</i>	3		0	0	0	0	57	17	1	14	0	15	0	11
42	<i>refl-fix</i>	4		0	0	0	0	3	30	0	20	0	15	0	14
43	<i>refl-rand</i>	2		100	0	100	0	100	0	100	0	99	1	100	0
44	<i>refl-rand</i>	4		100	0	100	0	100	0	100	0	100	0	100	0
45	<i>refl-rand</i>	6		100	0	100	0	100	0	100	0	99	1	100	0
46	<i>refl-rand</i>	8		100	0	100	0	100	0	100	0	100	0	100	0
47	<i>refl-worst</i>	2		100	0	100	0	100	0	100	0	99	1	96	4
48	<i>refl-worst</i>	4		100	0	100	0	100	0	100	0	99	1	98	2
49	<i>refl-worst</i>	6		100	0	100	0	100	0	100	0	98	2	100	0
50	<i>refl-worst</i>	8		100	0	100	0	100	0	100	0	100	0	99	1

Tab. 1 Heuristiky, jejich parametry, spolehlivost a předčasná konvergence.

V heuristikách BGA se z populace  $N$  bodů pro generování nového bodu užívá pouze subpopulace  $\mathcal{P}^*$  nejlepších  $\tau N$  bodů,  $\tau$  je vstupní parametr, jeho hodnoty jsou ve sloupci  $p_1$  v tabulce 1. Ze subpopulace  $\mathcal{P}^*$  se pak náhodně vyberou dva body  $\mathbf{x}$  a  $\mathbf{z}$ . Varianta *bga-line* vytváří nový bod  $\mathbf{y}$  podle vztahu

$$\mathbf{y} = \mathbf{x} + \theta(\mathbf{z} - \mathbf{x}),$$

varianta *bga-box* užívá vztah

$$y_i = x_i + \theta_i(z_i - x_i),$$

kde  $\theta$  a  $\theta_i$  ( $i = 1, 2, \dots, d$ ) jsou rovnoměrně rozděleny na  $\langle -\Delta, 1 + \Delta \rangle$ ,  $\Delta$  je další vstupní parametr, jeho hodnoty jsou ve sloupci  $p_2$  v tabulce 1.

Žádná z uvedených heuristik nezaručuje, že jí vygenerovaný bod  $\mathbf{y}$  nutně leží v  $D$ . Pokud nastane, že  $\mathbf{y} \notin D$ , je aplikována tzv. perturbace [6], která zrcadlově překlopí souřadnici  $y_i \notin \langle a_i, b_i \rangle$  dovnitř tohoto intervalu.

#### 4. TESTOVACÍ FUNKCE

Pro pokusy byly zvoleny tři v testech evolučních algoritmů často užívané funkce různého stupně obtížnosti [12], [15].

První De Jongova funkce,

$$f(\mathbf{x}) = \sum_{i=1}^d x_i^2,$$

$\langle a_i, b_i \rangle = \langle -5.12, 5.12 \rangle$ , jednodílná funkce, minimum je  $f(\mathbf{0}) = 0$ , je považována za velmi snadnou úlohu.

Druhá De Jongova funkce (Rosenbrockovo sedlo, známá i pod názvem banánové údolí),

$$f(\mathbf{x}) = \sum_{i=1}^{d-1} [100(x_i^2 - x_{i+1})^2 + (1 - x_i)^2],$$

$\langle a_i, b_i \rangle = \langle -2.048, 2.048 \rangle$ , jednodílná funkce, minimum je  $f(\mathbf{1}) = 0$ , středně obtížná pro algoritmy globální optimalizace.

Ackleyho funkce,

$$f(\mathbf{x}) = -20 \exp \left( -0.02 \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2} \right) - \exp \left( \frac{1}{d} \sum_{i=1}^d \cos 2\pi x_i \right) + 20 + \exp(1),$$

$\langle a_i, b_i \rangle = \langle -30, 30 \rangle$ , multimodální funkce s několika stovkami lokálních minim v  $D$ , globální minimum je  $f(\mathbf{0}) = 0$ .

#### NÁVRH EXPERIMENTU A SLEDOVANÉ VELIČINY

Padesát heuristik z tabulky 1 bylo postupně ověřováno ve dvou variantách algoritmu (CRS a HLF) na třech testovacích funkcích, první De Jongově (označena DeJong1,  $d = 3$ ), druhé De Jongově funkci (Rosenbrock,  $d = 2$ ) a Ackleyově funkci (Ackley,  $d = 2$ ). Testování tedy znamenalo 300 pokusů, každý byl opakován stokrát. Algoritmy pro testování byly implementovány v Matlabu [8]. Společné vstupní parametry algoritmů byly nastaveny takto:  $N = 10d$ , podmínka ukončení je  $f_{(N/2)} - f_{(1)} \leq 1 \times 10^{-7}$ , když hodnoty účelové funkce v jednotlivých bodech populace jsou uspořádány  $f_{(1)} \leq f_{(2)} \leq \dots \leq f_{(N)}$ . Hledání globálního minima je považováno za úspěšné, když  $f_{(1)} \leq f_{VTR}$ . Hodnoty  $f_{VTR}$  byly zvoleny v souladu s [12]  $1 \times 10^{-3}$  pro Ackleyovu funkci a  $1 \times 10^{-6}$  pro ostatní dvě funkce. Hledání bylo ukončeno před dosažením podmínky ukončení, když počet vyhodnocení cílové funkce přesáhl hodnotu  $5000d$ .

Sledovanými veličinami v každém opakování byly:

- TYP\_RES, typ ukončení, kategoriální, možné hodnoty
  - 1 – úspěšné hledání ( $f_{(1)} < f_{VTR}$ ) s dosažením podmínky ukončení,
  - 2 – úspěšné hledání ( $f_{(1)} < f_{VTR}$ ), podmínka ukončení nedosažena, počet vyhodnocení cílové funkce přesáhl  $N = 10d$ ,
  - 3 – předčasná konvergence, ( $f_{(1)} > f_{VTR}$ ), podmínka ukončení dosažena,
  - 4 – neúspěšné hledání ( $f_{(1)} > f_{VTR}$ ), počet vyhodnocení cílové funkce přesáhl  $N = 10d$ ;
- NE, celkový počet vyhodnocení cílové funkce;
- U, počet úspěšných vyhodnocení cílové funkce, kdy  $f(\mathbf{y}) < f(\mathbf{x}_{\text{worst}})$ ;
- U0, počet velmi úspěšných vyhodnocení cílové funkce, kdy  $f(\mathbf{y}) < f(\mathbf{x}_{\text{best}})$  (nový bod je lepší než dosud nejlepší);
- NE\_VTR, U\_VTR, U0\_VTR, počty vyhodnocení cílové funkce k dosažení podmínky ( $f_{(1)} < f_{VTR}$ ), pokud tato podmínka nebyla dosažena, hodnoty těchto veličin chybí;
- NE\_5, U\_5, U0\_5, počty vyhodnocení cílové funkce k dosažení podmínky  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$ , pokud tato podmínka nebyla dosažena, hodnoty těchto veličin chybí.

V analýze dat byly dále užity odvozené veličiny:

- U\_R = U / NE, relativní četnost úspěšných vyhodnocení cílové funkce;
- U0\_R = U0 / U, relativní četnost velmi úspěšných vyhodnocení cílové funkce;
- NE\_VTR\_R = NE\_VTR / NE, relativní četnost velmi úspěšných vyhodnocení cílové funkce, definována, pokud má NE\_VTR zjištěnou hodnotu;
- U\_VTR\_R = U\_VTR / NE\_VTR, relativní četnost úspěšných vyhodnocení cílové funkce potřebných k dosažení  $f_{VTR}$ , definována, pokud má NE\_VTR zjištěnou hodnotu;
- U0\_VTR\_R = U0\_VTR / U\_VTR, relativní četnost velmi úspěšných vyhodnocení cílové funkce potřebných k dosažení  $f_{VTR}$ , definována, pokud má U\_VTR zjištěnou hodnotu;
- NE\_5\_R = NE\_5 / NE, relativní četnost vyhodnocení cílové funkce potřebných k dosažení  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$ , definována, pokud má NE\_5 zjištěnou hodnotu;
- U\_5\_R = U\_5 / NE\_5, relativní četnost úspěšných vyhodnocení cílové funkce potřebných k dosažení  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$ , definována, pokud má NE\_5 zjištěnou hodnotu;
- U0\_5\_R = U0\_5 / U\_5, relativní četnost velmi úspěšných vyhodnocení cílové funkce potřebných k dosažení  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$ , definována, pokud má NE\_5 zjištěnou hodnotu;
- DIF\_NE\_5 = NE - NE\_5, počet vyhodnocení cílové funkce potřebných od dosažení  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$  do dosažení podmínky ukočení, charakterizuje rychlost konvergence v závěru vyhledávání, definována, pokud NE\_5 má zjištěnou hodnotu;
- DIF\_5\_VTR = NE\_5 - NE\_VTR, počet vyhodnocení cílové funkce potřebných od dosažení  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$  do dosažení podmínky ukočení, charakterizuje rychlost konvergence uprostřed vyhledávání, definována, pokud NE\_VTR má zjištěnou hodnotu a  $NE_5 > NE_VTR$ ;

- IRREG = missing, pokud NE\_VTR nemá zjištěnou hodnotu,  
IRREG = 0, pokud NE\_5 > NE\_VTR,  
IRREG = 1, pokud NE\_5 ≤ NE\_VTR.

Tato veličina byla zavedena, když v přeběžné analýze dat se ukázalo, že některé heuristiky v některých pokusech se chovají poněkud nečekaně a podmínku  $f_{(5)} - f_{(1)} \leq 1 \times 10^{-7}$  dosáhnou dříve než dosáhnou hodnotu  $f_{VTR}$ . Tato veličina tedy znamená příznak nepravidelného chování heuristiky.

Tato data, ve kterých je objektem (řádkem datové matice) opakování, budeme označovat jako původní soubor mající 30 000 řádků.

Pro každý pokus (tj. 100 opakování při stejných podmínkách) byly spočítány souhrné charakteristiky. Spolehlivost hledání  $R$  je hodnocena jako četnost dosažení TYP\_RES=1 ve sto opakováních, předčasná konvergence  $P$  jako četnost dosažení TYP\_RES=3 ve sto opakováních. Tyto hodnoty pro jednotlivé pokusy jsou uvedeny v tabulce 1. V souhrnném souboru (300 řádků) jsou i další výběrové charakteristiky pro jednotlivé pokusy a typy výsledku. Pro TYP\_RES=1 jsou jména veličin obsahující průměry shodná se jmény veličin, ze kterých jsou průměry počítány, směrodatné odchylky mají předponu SD\_, pro TYP\_RES=3 mají jména výběrových charakteristik příponu 3.

## 5. ANALÝZA DAT

Už předběžná analýza dat ukázala, že vlastnosti heuristik se liší daleko více, než se očekávalo. Podobně jako se liší spolehlivost v tabulce 1, liší se podstatně i další změřené charakteristiky. Spolehlivost nalezení globálního minima se nezvyšuje s počtem vyhodnocení účelové funkce (tedy s důkladností prohledávání oblasti  $D$ ), jak bylo očekáváno. Místo předpokládného užití statistických konfirmačních metod bylo nutno užít metody explorační a hledat alespoň hrubou klasifikaci heuristik.

*Hledání pravidel pro klasifikaci je obtížná úloha, v jejímž řešení je nutno činit kompromisy, někdy i nesnadno zdůvodnitelné. Výsledek je závislý na subjektu řešitele. Zde byla situace komplikovanější tím, že shodnost klienta a řešitele znamenala vlastně neustálé hledání takových kompromisů, bez existence větších oddělených kroků, které jsou vymezeny, spolupracuje-li více lidí a v komunikaci mezi klientem a řešitelem jsou užívána formalizovanější pravidla. Postup řešení této úlohy občas připomínal příběh o Achillovi dohánějícím želvu.*

Pro klasifikaci lze užít řadu postupů explorativní analýzy dat od metody GUHA po celou paletu metod shlukové analýzy. Vzhledem k tomu, že většinu veličin v této úloze můžeme považovat za spojitě a jejich variabilita je stochastická, byla dána přednost postupům statistickým. Pomocí regresní analýzy byly vyhledávány veličiny vysvětlující úspěšnost nalezení globálního minima. V těchto výpočtech byl užít paket Stata [4], ve kterém je dobře podporována transformace kategoriálních regresorů a zadávání úloh je pohodlné. Postupnou logistickou regresí v původním souboru (vysvětlovaná veličina  $Y = 1$ , když TYP\_RES=1,  $Y = 0$ , když TYP\_RES=2, jinak je případ vypuštěn) po adjustaci na funkci a algoritmus bylo zjištěno, že významně od nuly odlišné jsou odhady regresních koeficientů u veličin NE, U\_R, U0\_R a DIF\_NE\_5, tj. u všech veličin, kde jsou k dispozici naměřené hodnoty a tedy mohly být nabídnuty do procedury postupného výběru. Všechny odhady regresních koeficientů byly záporné, u všech je významnost  $p < 0.0005$ .

Postupnou lineární regresí ze souhrnných dat, kdy vysvětlovanou veličinou byla spolehlivost  $R$ , byly nalezeny regresory NE, IRREG, DIF\_NE\_5, U\_VTR\_R, NE\_5\_R

a SD\_U\_R. Další veličiny potenciálně rozlišující heuristiky byly vybrány poněkud od-  
vážnou extrapolací výsledků lineární regrese vysvětlující předčasnou konvergenci  $P$   
veličinami DIF\_NE\_53, U\_R3, U0\_R3 a SD\_U\_R3 a úvahami opřenými o předběžnou  
analýzu.

Takto vybraných 16 veličin (viz tabulka 2) bylo využito v nehierarchické shlukové  
analýze, v proceduře k-means. K výpočtům byl užít paket [3]. Objekty (tzn. pokusy)  
byly klasifikovány do čtyř shluků, neboť se očekávalo, že heuristiky se budou seskupovat  
do shluků, které bude možno charakterizovat jako nejméně tři kategorie: rychlé,  
spolehlivé, jiné. Výtah z výsledků této klasifikace je uveden v tabulce 2. V řádku  
Kategorie je uvedeno ad hoc hodnocení skupiny heuristik podle průměrů veličin  
v daném shluku, R znamená rychlá, S spolehlivá, N nevhodná. U všech veličin ve  
shlukové analýze na všech třech funkcích bylo zjištěno F-testem, že variabilita mezi  
shluky je významně větší než variabilita uvnitř shluků, významnost  $p < 0.00005$ .  
Jedinou výjimkou byla veličina IRREG v případě Ackleyho funkce, kde rozdíl mezi  
shluky nebyl významný, neboť u této funkce se zřejmě heuristiky s nepravidelným  
chováním mezi klasifikované vůbec nepropracovaly (vzhledem k tomu, že bylo nutno  
dosáhnout nejméně dvou úspěšných nalezení globálního minima ze sta opakování,  
aby veličiny neměly chybějící hodnoty).

Pokud nízká spolehlivost nalezení globálního minima a vysoký počet vyhodnocení  
cílové funkce potřebný k dosažení podmínky ukončení – viz tab. 1 a 2 – vzbuzují  
dojem, že užité heuristiky nejsou lepší než náhodné prohledávání, není tento do-  
jem správný. V experimentech, kdy nový bod  $y$  byl rovnoměrně rozložený na  $D$  se  
u všech třech funkcí ani v jednom ze sta opakování nepodařilo před dosažením 5000d  
vyhodnocení cílové funkce nalézt bod, ve kterém by  $f_{VTR} \leq f_{(1)}$ .

Veličina	DeJong1				Rosenbrock				Ackley			
	1	2	3	4	1	2	3	4	1	2	3	4
R	98.2	96.0	17.6	62.4	50.1	91.0	44.7	27.8	5.0	89.7	80.8	2.0
NE	4372	1453	6519	4565	950	1834	5205	4774	5457	4160	1333	3117
IRREG	0.000	0.047	0.882	0.345	0.107	0.002	0.121	0.449	0.000	0.000	0.008	0.000
DIF_5_VT	801	170	113	42	154	367	1062	186	2167	1528	578	535
DIF_NE_5	348	112	1894	92	65	120	534	575	2139	163	59	62
U_R	0.336	0.792	0.872	0.953	0.685	0.310	0.232	0.544	0.355	0.272	0.682	0.922
U0_R	0.033	0.054	0.084	0.266	0.057	0.057	0.043	0.205	0.084	0.045	0.051	0.173
NE_VTR_R	0.743	0.783	0.945	0.898	0.775	0.743	0.700	0.856	0.307	0.589	0.529	0.808
U_VTR_R	0.342	0.787	0.871	0.942	0.685	0.312	0.266	0.581	0.731	0.210	0.584	0.946
U0_VTR_R	0.035	0.057	0.084	0.275	0.060	0.061	0.046	0.209	0.087	0.048	0.054	0.173
NE_5_R	0.921	0.919	0.691	0.941	0.922	0.937	0.895	0.853	0.680	0.963	0.955	0.980
U_5_R	0.338	0.792	0.874	0.949	0.691	0.311	0.245	0.584	0.458	0.268	0.676	0.932
U0_5_R	0.010	0.044	0.068	0.258	0.040	0.017	0.011	0.130	0.045	0.011	0.034	0.162
SD_NE	266	439	3927	3031	375	269	1453	2651	2324	425	139	49
SD_U_R	0.012	0.016	0.012	0.030	0.043	0.030	0.041	0.119	0.232	0.023	0.033	0.033
P	0.0	2.9	16.3	0.0	26.6	6.3	9.6	5.9	72.8	2.5	19.2	76.0
Počet	22	48	7	7	22	30	18	15	10	19	25	1
Kategorie	S	R	N	N	R	S	N	N	N	S	R	N

Tab. 2 Předběžná klasifikace procedurou k-means, průměry shluků.

Po této předběžné klasifikaci byly nalezeny heuristiky, které patřily u všech funkcí  
do kategorie R (celkem 24 pokusů pro oba testované algoritmy), heuristiky vždy pa-  
třící do kategorie S (celkem 45). Žádná z heuristik nebyla v předběžné klasifikaci  
nalezena jako nevhodná pro všechny tři funkce. Proto kategorie nevhodných heuris-  
tik (N) potřebná pro další klasifikaci byla vytvořena z heuristik, které patřily mezi



nevhodné alespoň u dvou funkcí (celkem 36 pokusů) a z kategorie heuristik rychlých v případě snadné De Jongovy funkce, nevhodných pro Rosenbrokovu funkci a zcela neúspěšných v případě Ackleyho funkce (48 pokusů).

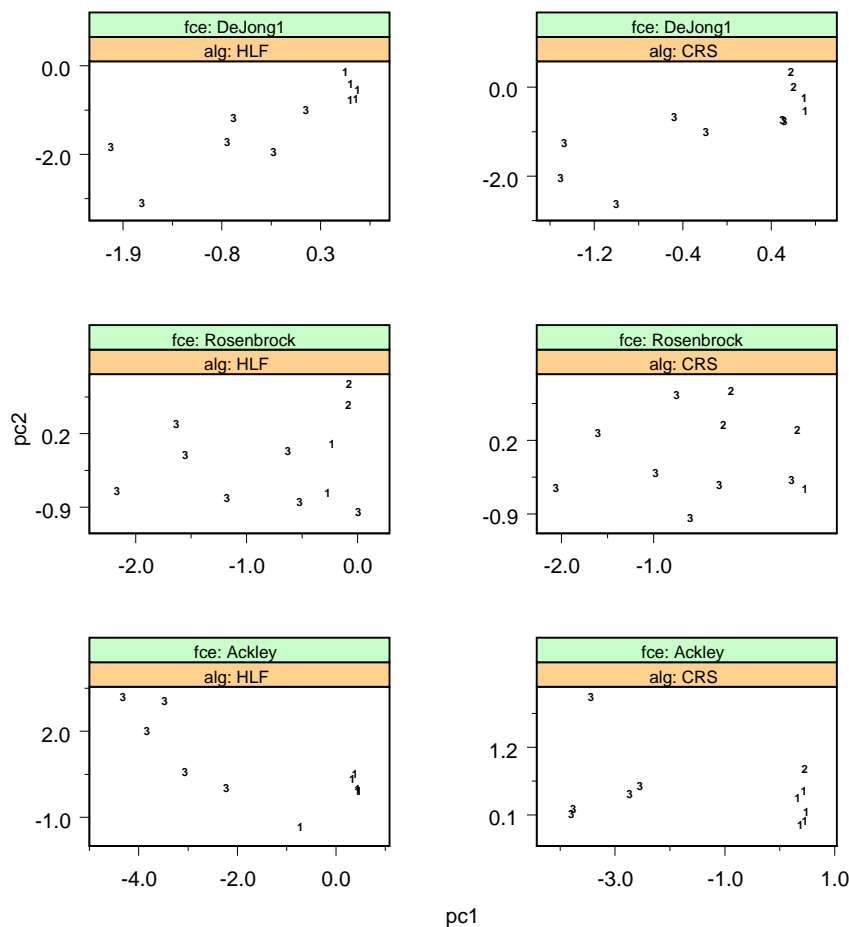
Takto vytvořená kategorizace byla využita pro klasifikaci zbylých heuristik pomocí diskriminační analýzy. Pro postupné vytváření lineární diskriminační funkce byly užity veličiny z tabulky 2 kromě veličin U0\_VTR\_R a U\_5\_R, které byly silně korelovány s veličinami U0\_R, resp. s U\_R, koeficient korelace byl v obou případech větší než 0,99. Pomocí nalezené diskriminační funkce byly klasifikovány i některé heuristiky, které nebylo možno předběžně kategorizovat. Výsledky této klasifikace jsou pro testované funkce a algoritmy jsou na obr. 1 nakresleny v rovině prvních dvou hlavních komponent korelační matice veličin z tab. 2 kromě U0\_VTR\_R a U\_5\_R. První dvě hlavní komponenty vysvětlují 59 procent celkového rozptylu. Symbol 1 v grafech označuje heuristiky rychlé, 2 spolehlivé a 3 heuristiky klasifikované jako nevhodné.

Všech 24 pokusů apriorně klasifikovaných jako R bylo správně zařazeno do kategorie 1, ze 45 heuristik apriorně klasifikovaných jako spolehlivé byla jedna reklasifikována do kategorie 1, ostatní byly zařazeny shodně s apriorní klasifikací a z 56 apriorně nevhodných heuristik, u kterých byly k dispozici údaje umožňující predikovat klasifikaci pomocí lineární diskriminační funkce, bylo 44 zařazeno do kategorie 3, 4 mezi spolehlivé a 8 mezi rychlé. Shoda predikované a apriorní klasifikace je tedy celkem uspokojivá. Základní charakteristiky nalezených kategorií heuristik jsou uvedeny v tabulce 3.

Jak z obr. 1, tak z tabulky 3 je zjevné, že nalezená klasifikace heuristik není ostrá a je nutno ji brát s jistou opatrností. Přesto základní obraz o chování heuristik je na základě této klasifikace možný. Podstatné rozdíly mezi kategoriemi heuristik jsou především ve veličinách IRREG a U\_R. U všech heuristik typu BGA a ES byla nepravidelnost charakterizovaná veličinou IRREG nejméně ve třech ze 6 pokusů a tyto heuristiky jsou také nejhůře klasifikovány, z celkem 180 pokusů není 87 klasifikováno pro naprosté selhání, z 93 klasifikovaných je 61 v kategorii nevhodných. Z heuristik typu REFL byla nepravidelnost zjištěna jen u REFL\_FIX a tato heuristika je mezi spolehlivé či rychlé zařazena jen velmi zřídka. Z heuristik typu DE se nepravidelnost zjistila pouze u heuristik 22, 25, 26 a 28 a ze 72 pokusů byly tyto heuristiky 46-krát klasifikovány jako rychlé a 19-krát jako spolehlivé. Ve všech šesti úlohách byly mezi spolehlivé zařazeny pouze heuristiky 44-46 a 48-50, mezi rychlé 21 a 27, buď jako rychlé nebo spolehlivé byly klasifikovány heuristiky 20, 22-24, 26, 29, 30, 39, 43 a 47.

Veličina	Predikovaná kategorie					
	1 (rychlá)		2 (spolehlivá)		3 (nevhodná)	
	průměr	sm. odch.	průměr	sm. odch.	průměr	sm. odch.
R	85.7	25.6	90.2	23.8	51.8	40.5
NE	1344	808	3576	2577	3474	2380
IRREG	0.0040	0.0127	0.0102	0.0487	0.2112	0.2819
DIF_5_VT	357	268	893	762	506	937
DIF_NE_5	72	51	231	222	640	886
U_R	0.735	0.134	0.292	0.129	0.628	0.252
U0_R	0.0461	0.0234	0.0435	0.0188	0.1213	0.0838
P	11.2	18.8	4.5	13.2	18.3	30.5

**Tab. 3** Základní charakteristiky kategorií heuristik.



Obr. 1 Klasifikace heuristik v rovině prvních dvou hlavních komponent.

## 6. ZÁVĚR

Výsledky z experimentálního porovnání heuristik a navržené klasifikace lze shrnout stručně shrnout: Heuristiky třídy BGA jsou velmi citlivé na volbu vstupních parametrů, variabilita jejich chování vysoká, v převážné většině byly zařazeny do kategorie nevhodných a jejich využití v evolučních algoritmech se soutěží nebo spoluprací heuristik je problematické. Podobně i heuristiky ES jsou citlivé na volbu vstupního parametru a mají tendenci velmi rychle zhušťovat populaci k předčasné konvergenci. Proto zcela selhávaly v nalezení globálního minima vícemodální Ackleyho funkce. V algoritmech se soutěží nebo spoluprací heuristik je lze využít jen tehdy, když pravidla střídání heuristik budou omezovat příliš časté užití těchto heuristik v procesu vyhledávání. Heuristiky DE jsou převážně klasifikovány v kategorii rychlých, při vyšších hodnotách parametru  $C$  (heuristika 23, 24, 30) přecházejí většinou do kategorie spolehlivých. Podobně se chovají i heuristiky *refl-worst* a *refl-rand* při nízkých hodnotách parametru  $\alpha$ . Heuristiky *refl-fix* dosahovaly většinou horších výsledků než

jejich znáhodněná obdoba a zřejmě není důvod pro jejich využívání v algoritmech s více heuristikami. Heuristiky *refl-worst* a *refl-rand* při hodnotách parametru  $\alpha \geq 4$  ve všech pokusech patřily do kategorie spolehlivých a jejich využití v algoritmech s více heuristikami je nepochybně žádoucí.

Dále je zřejmé, že spolehlivost hledání je ovlivňována relativní úspěšností (veličina U\_R), a proto by tato relativní úspěšnost měla být jednou z veličin, která bude využívána pro řízení algoritmů s více heuristikami. Nepravidelnost (IRREG) je příznakem „rizikového“ chování a heuristiky vykazující tuto nepravidelnost by neměly být v algoritmech s více heuristikami užívány jako spolehlivé.

Ačkoliv je zřejmé, že nalezená klasifikace je neostrá a její zobecnění na jiné funkce problematické, evoluční algoritmy se soutěžícími heuristikami, které byly navrženy na základě předběžné analýzy dat, byly úspěšné na testovacích funkcích [18] i v řešení některých problémů výpočetní statistiky (odhady parametrů nelineárních regresních modelů, robustní CM odhady parametrů lineárního regresního modelu) [19]. Důkladnější využití výsledků této práce v návrhu pravidel spolupráce heuristik by mělo přispět k nalezení evolučních algoritmů s vysokou samoadaptací a tím i jejich jednoduchému využívání pro širší třídu optimalizačních problémů.

*Oproti všem předpokladům se ve „spolupráci“ statistika a klienta objevily snad všechny už tradiční potíže, od nedodržování termínů a dohod, přes vyhýbání se zodpovědnosti až po nesmířitelné sváry. Nevím, zda to bylo způsobeno povahou úlohy nebo osobními vlastnostmi, ale ke svým skutečným klientům po těchto zkušenostech budu mít více trpělivosti a porozumění než dosud.*

#### LITERATURA

- [1] Bäck, T. (1996): *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, New York.
- [2] Haslinger, J., Jedelský, D., Kozubek, T. and Tvrdík, J. (2000): Genetic and Random Search Methods in Optimal Shape Design Problems, *J. Global Optimization* **16**, 109–131.
- [3] Hintze, J. (2001): *NCSS and PASS, Number Cruncher Statistical Systems*. Kaysville, Utah: WWW.NCSS.COM.
- [4] Intercooled Stata 7.0 for Windows 98/95/NT. (2000). Stata Corporation, College Station, Texas, USA.
- [5] Křivý, I. and Tvrdík, J. (1995): The Controlled Random Search Algorithm in Optimizing Regression Models. *Comput. Statist. and Data Anal.* **20**, 229–234.
- [6] Kvasnička, V., Pospíchal, J. a Tiňo, P. (2000): *Evoluční algoritmy*. Slovenská technická univerzita, Bratislava.
- [7] Malý, M. a Roth, Z. (2001): Otázky komunikace statistika s lékařem, In: *PRASTAN 2001*, 98–103, Slovenská štatistická a demografická spoločnosť, Bratislava.
- [8] MATLAB, version 6, The MathWorks, Inc., (2000).
- [9] Mišík, L., Tvrdík, J. and Křivý, I. (2001): On Convergence of a Class of Stochastic Algorithms, In: Antoch J. and Dohnal G. (eds): *Proceedings of ROBUST 2000*, JČMF, Praha, 198 - 209.
- [10] Mühlenbein, H. and Schlierkamp-Voosen, D. (1993): Predictive Models for Breeder Genetic Algorithm, I. Continuous Parameter Optimization, *Evolutionary Computation* **1**, 25–49.
- [11] Price, W.L. (1976): A Controlled Random Search Procedure for Global Optimization. *Computer J.* **20**, 367-370.
- [12] Storn R. and Price K. (1997): Differential Evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, *J. Global Optimization* **11**, 341–359.
- [13] Törn, A. and Žilinskas., A. (1989): *Global Optimization*, Lecture Notes in Computer Science 350, Springer-Verlag, Berlin Heidelberg.

- [14] Tvrđík, J. (1990): Statistické programy a jejich uživatelé, In: Antoch, J. (ed.): *RO-BUST'90*, 155–159, JČMF Praha,
- [15] Tvrđík, J. and Křivý, I. (1999): Simple Evolutionary Heuristics for Global Optimization. *Comput. Statist. and Data Anal.* **30**, 345–352
- [16] Tvrđík, J., Křivý, I. and Mišík, L. (2001): Evolutionary Algorithm with Competing Heuristics. In: Ošmera, P. (ed.): *MENDEL 2001, 7th International Conference on Soft Computing*. Technical University, Brno, 58–64.
- [17] Tvrđík, J. (2002): Algoritmus řízeného náhodného prohledávání a alternující heuristiky. *Automa* **8**, 54–57.
- [18] Tvrđík, J., Mišík, L. and Křivý, I. (2002): Competing Heuristics in Evolutionary Algorithms. Accepted to the 2nd Euro-ISCI (Košice, July 16-19, 2002), to appear in *Studies in Computational Intelligence*. Springer-Verlag, Berlin.
- [19] Tvrđík, J., Křivý, I. and Mišík, L. (2002): Evolutionary Algorithms with Competing Heuristics in Computational Statistics. Accepted to the COMPSTAT 2002 (Berlin, August 24-28, 2002) as contributed paper

KATEDRA INFORMATIKY, PŘÍRODOVĚDECKÁ FAKULTA, OSTRAVSKÁ UNIVERSITA, 30. DUBNA 22,  
701 03 OSTRAVA  
E-MAIL: [tvrdik@osu.cz](mailto:tvrdik@osu.cz)