

Univerzita Hradec Králové
Fakulta informatiky a managementu
Katedra informatiky a kvantitativních metod

**Vliv selekčních metod na tvorbu nové populace a chování
genetického algoritmu**
Bakalářská práce

Autor: Ondřej Tázlar
Studijní obor: Aplikovaná informatika (ai3)

Vedoucí práce: prof. RNDr. Josef Hynek, MBA, Ph.D.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně a s použitím uvedené literatury.

V Hradci Králové dne 20.1.2017

Ondřej Tázlar

Poděkování:

Děkuji vedoucímu bakalářské práce prof. RNDr. Josefu Hynkovi, MBA, Ph.D. za cenné rady a podmětné připomínky v průběhu jejího vypracování.

Anotace

Cílem této bakalářské práce je představit základní selekční metody a popsat jejich vlastnosti a vliv na tvorbu nové populace a genetický algoritmus. Práce je rozdělena do tří hlavních částí. První dvě jsou teoretické a poslední je praktická.

První část obsahuje úvod do genetických algoritmů (od historie evolučních algoritmů a popisu principů jejich funkce až po detailní vysvětlení jednotlivých kroků jednoduchého genetického algoritmu).

Druhá část se zaměřuje pouze na podrobný popis vybraných selekčních metod včetně jejich vzájemného srovnání.

Třetí část této bakalářské práce popisuje testovací úlohu, její implementaci a výsledky testů jednotlivých selekčních metod.

Výsledky pokusů popsaných v praktické části odpovídají závěrům z teoretické části práce.

Annotation

Title: Effects of selection methods on creating new population and the behaviour of the genetic algorithm

The aim of this Bachelor Thesis is to introduce common selection methods and describe their properties and effects on creating new population and genetic algorithm. The thesis is divided into three main parts. The first two are theoretical and the last one is practical.

The first part contains an introduction to the genetic algorithms (from history of evolutionary algorithms and description of their principles to the detailed explanation of the various steps of simple genetic algorithm).

The second part focuses only on detailed description of the chosen selection methods including peer comparisons.

The third part of this bachelor thesis describes a testing task, its implementation and the results of tests of various selection methods.

Results of experiments described in the practical part are consistent with the findings from the theoretical part.

Obsah

1	Úvod.....	1
2	Cíl práce.....	2
3	Úvod do Genetických algoritmů	3
3.1	Úvod.....	3
3.2	Historie.....	3
3.3	Evoluční algoritmy.....	4
3.4	Techniky evolučních algoritmů	4
3.5	Genetický algoritmus	6
3.5.1	Schéma GA.....	7
3.5.2	Počáteční populace	8
3.5.3	Ohodnocení jedinců	9
3.5.4	Metody výběru jedinců (selekce)	10
3.5.5	Křížení.....	11
3.5.6	Mutace.....	12
3.5.7	Nová populace.....	13
3.5.8	Ukončení běhu algoritmu	13
3.6	Závěr.....	13
4	Selekční metody.....	14
4.1	Úvod.....	14
4.2	Proporcionální metody	15
4.2.1	Ruletová selekce.....	15
4.2.2	Stochastické univerzální vzorkování	18
4.3	Pořadová selekce	19
4.3.1	Lineární pořadová selekce	19
4.3.2	Nelineární pořadová selekce.....	20

4.3.3	Srovnání lineární a nelineární pořadové metody	20
4.3.4	Analýza lineární pořadové metody.....	22
4.4	Turnajová selekce	23
4.4.1	Analýza turnajové metody	24
4.5	Ořezová selekce.....	25
4.5.1	Analýza ořezové metody.....	25
4.6	Genitor / Steady-State.....	27
4.7	Srovnání selekčních metod.....	29
4.7.1	Parametry selekce a její intenzita	29
4.7.2	Ztráta diverzity a intenzita selekce.....	30
4.7.3	Rozptyl selekce a její intenzita	31
4.8	Závěr.....	31
5	Praktická část.....	32
5.1	Testovací úloha	32
5.2	Implementace úlohy – obsluha programu.....	33
5.3	Výsledky testování.....	35
5.3.1	Proporcionální metody.....	35
5.3.2	Lineární pořadová selekce	37
5.3.3	Turnajová selekce.....	39
5.3.4	Ořezová selekce.....	42
5.4	Shrnutí výsledků testování.....	44
6	Závěry a doporučení	45
7	Seznam použité literatury.....	46
8	Přílohy	48

Seznam obrázků

Obr. 1 Klasifikace metaheuristik.....	5
Obr. 2 Struktura jedince	6
Obr. 3 Vliv velikosti populace GA.....	9
Obr. 4 Křížení.....	12
Obr. 5 Mutace	12
Obr. 6 Ruletová selekce (RWS).....	17
Obr. 7 Stochastické univerzální vzorkování (SUS)	18
Obr. 8 Přiřazení fitness pro lineární a nelineární pořadovou selekci.....	21
Obr. 9 Vlastnosti lineární pořadové selekce.....	22
Obr. 10 Vlastnosti turnajové selekce.....	24
Obr. 11 Vlastnosti ořezové selekce	26
Obr. 12 Random genetic drift chart	28
Obr. 13 Závislost parametru selekce na intenzitě selekce.....	29
Obr. 14 Závislost ztráty diverzity na intenzitě selekce	30
Obr. 15 Závislost rozptylu selekce na intenzitě selekce	31
Obr. 16 Grafické uživatelské rozhraní programu.....	34
Obr. 17 Ruletová selekce.....	36
Obr. 18 Stochastické univerzální vzorkování	37
Obr. 19 Lineární pořadová selekce, $SP = 1,1$	38
Obr. 20 Lineární pořadová selekce, $SP = 1,5$	38
Obr. 21 Lineární pořadová selekce, $SP = 2$	39
Obr. 22 Turnajová selekce, $Tour = 2$	40
Obr. 23 Turnajová selekce, $Tour = 5$	40
Obr. 24 Turnajová selekce, $Tour = 25$	41
Obr. 25 Ořezová selekce, $Trunc = 0,1$	42
Obr. 26 Ořezová selekce, $Trunc = 0,5$	42
Obr. 27 Ořezová selekce, $Trunc = 0,8$	43

Seznam tabulek

Tabulka 1 Příklad populace	16
Tabulka 2 Pravděpodobnost selekce a hodnota fitness.....	17
Tabulka 3 Závislost hodnoty fitness na tlaku selekce	20
Tabulka 4 Vztah mezi velikostí turnaje a intenzitou selekce.....	23
Tabulka 5 Vztah mezi prahem ořezu a intenzitou selekce	25
Tabulka 6 Souhrn výsledků testování.....	44

1 Úvod

V dnešní době je možno pozorovat neustále se rozvíjející aplikaci optimalizačních metod využívaných v běžném životě, aniž by o nich normální člověk nějak přemýšlel, nebo o nich vůbec věděl. Příkladem může být plánování firemních procesů i jednotlivých úkonů ve výrobě – jako je například co nejefektivnější využití materiálu (rozmístění dílů s cílem minimalizovat množství odpadu) atd. Další široce využívanou funkcí je navrhování tras v dopravě – snad každý ví, k čemu slouží auto-navigace, nebo GPS v mobilním telefonu, ale kolik lidí opravdu zajímalo, jak takové zařízení funguje? Uživatel ví, že jeho poloha se určuje pomocí družic. Málokdo ovšem dál zkoumal, jak se vybírá vhodná trasa do cílového bodu. Právě zde lze najít aplikované optimalizační metody, jejichž rozvíjející se skupinou jsou tzv. evoluční výpočetní techniky (EVT) a potažmo pak genetické algoritmy (GA) – jako jedny z prvních zástupců těchto technik.

Jsem si vědom toho, že genetické algoritmy a jejich problematika je natolik rozsáhlé téma, že není možné, vzhledem k rozsahu této práce, obsáhnout je v celé šíři a složitosti. Proto jsem se v práci zaměřil na jeden ze základních prvků činnosti genetických algoritmů, tj. výběr vhodných řešení pro další zpracování. V praxi existuje nepřeberně variant selekčních metod uzpůsobených pro řešení konkrétních úloh. Tato práce se bude zabývat právě těmito metodami, jejichž činnost a vlastnosti se pokusí podrobněji vysvětlit.

2 Cíl práce

Cílem práce je představit čtenáři základní a zároveň nejrozšířenější selekční metody využívané v genetických algoritmech a dále prozkoumat a porovnat jejich vlastnosti a vliv na další činnost algoritmu – konkrétně tedy, jak ovlivňuje selekční metoda jedince v nové populaci potenciálních řešení.

Za tímto účelem bude čtenáři představen v teoretické části práce popis činnosti těchto vybraných metod a dále několik používaných objektivních parametrů, které nám poslouží k jejich následnému srovnání.

Závěry z tohoto vyvozené budou následně ověřeny experimentálně v praktické části této práce.

3 Úvod do Genetických algoritmů

3.1 Úvod

Cílem této kapitoly je posloužit jako velmi stručný úvod do rozsáhlé, avšak zajímavé oblasti genetických algoritmů.

Genetické algoritmy jsou jednou z technik spadajících do skupiny evolučních algoritmů. Jedná se o (meta)heuristické metody napodobující procesy biologické evoluce, jejichž cílem je najít co nejlepší řešení úloh globální optimalizace v rámci daných parametrů.

Začneme krátkým pohledem do jejich historie. Přes vlastnosti evolučních algoritmů a jejich techniky se dostaneme k algoritmům genetickým, jejichž průběh si následně podrobněji rozebereme.

3.2 Historie

Počátky evolučních algoritmů lze umístit do padesátých až šedesátých let dvacátého století, kdy se začaly objevovat rozmanité pokusy simulovat a využívat principů evoluční teorie definované roku 1859 Charlesem Darwinem [15]. Mezi první experimentátory patřili logicky biologové – zmiňme například genetiky Alexe Frasera, Donalda Burnella a Jacka Crosbyho, jejichž práce obsahovaly všechny základní prvky moderních genetických algoritmů. Některé rané práce lze nalézt ve sborníku Davida Fogela *Evolutionary Computation: The Fossil Record* (1998) [6].

Mezi další směry výzkumu patřily techniky evolučního programování – např. Lawrence Fogel a evolučních strategií – Ingo Rechenberg a Hans-Paul Schwefel.

Genetické algoritmy získaly svou popularitu v sedmdesátých letech díky práci Johna Hollanda zabývající se buněčnými automaty na Michiganské univerzitě a jeho publikaci *Adaptation in Natural and Artificial Systems* (1975) [9]. Nicméně jejich výzkum pokračoval hlavně v teoretické rovině až do první mezinárodní konference zabývající se genetickými algoritmy konané v Pittsburghu a současnému nárůstu výpočetní síly soudobých počítačů poloviny osmdesátých let, což umožnilo první prakticky využitelné aplikace této metody.

3.3 Evoluční algoritmy

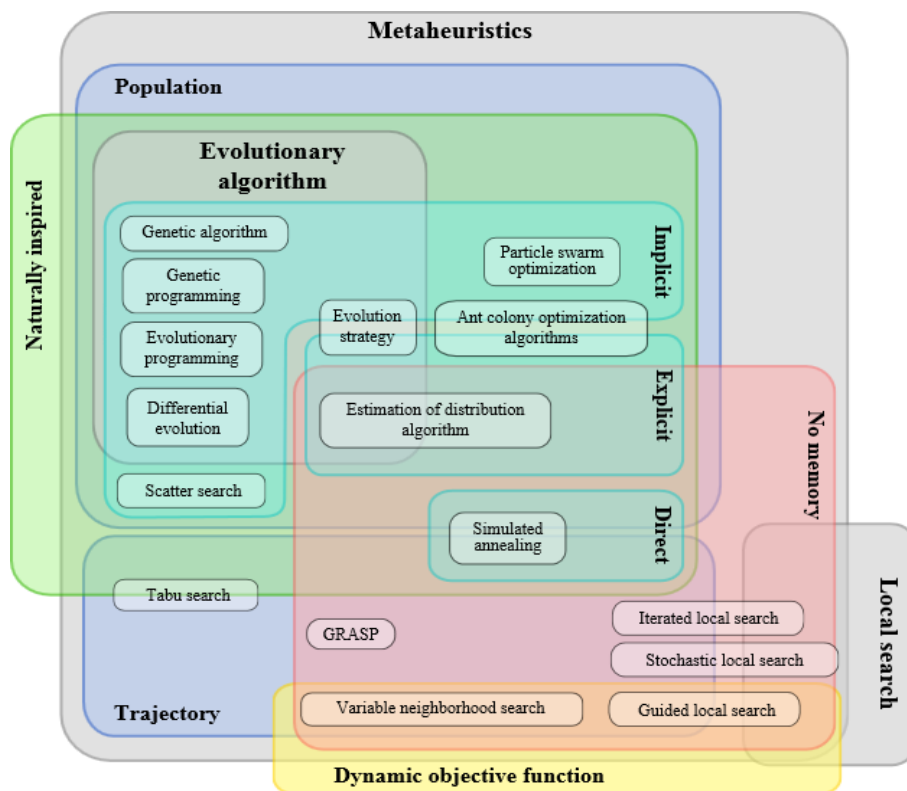
Jak již bylo řečeno v úvodu, genetické algoritmy jsou pouze jednou z metod spadající do skupiny algoritmů evolučních – evolučních proto, že napodobují biologickou „Darwinovu“ evoluci. Podle Darwina [15] mezi sebou jedinci daného druhu (i zástupci druhů rozdílných) neustále soupeří o přežití – o potravu a jiné přírodní zdroje a právo se rozmnožovat. Zaměřením těchto metod jsou obecně různé optimalizační úlohy, které jsou tradičními technikami velmi obtížně řešitelné, či přímo neřešitelné nebo úlohy, jež jsou obtížně definovatelné, ale máme představu o požadovaném výsledku.

Klasické evoluční algoritmy pracují s populací (množinou) možných řešení dané úlohy – obsahují tedy jakýsi vnitřní paralelismus. Populace řešení se postupně vyvíjí – tzn. horší jedinci jsou nahrazováni lepšími (z pohledu žádaného řešení), vznikajícími užitím evolučních operátorů křížení, mutace, selekce apod. na jedince stávající populace. Tyto operátory jsou aplikovány opakovaně až do nalezení vyhovujícího řešení, či splnění jiných ukončovacích podmínek.

Kvalitu jednotlivých řešení určuje tzv. ohodnocující (fitness) funkce, která přezkoumá a ohodnotí všechny jedince nově vzniklé populace v závislosti na tom, jak se blíží žádanému (optimálnímu, či suboptimálnímu) řešení. Vzhledem k tomu, že vlastní implementace této funkce může být i velice složitá, je logické, že výpočetní čas požadovaný evolučním algoritmem je v přímé korelaci s touto funkcí a velikostí populace, na kterou je aplikována. Významný vliv má také vhodná reprezentace jedinců a vnitřní implementace aplikovaných operátorů.

3.4 Techniky evolučních algoritmů

Evoluční algoritmy lze řadit mezi tzv. (meta)heuristiky (viz Obr. 1), což jsou procedury pro vyhledávání řešení obvykle různých optimalizačních úloh. Tyto procedury negarantují nalezení optimálního řešení, ale nám často stačí nalézt řešení pouze „dostatečně dobré“.



Obr. 1 Klasifikace metaheuristik

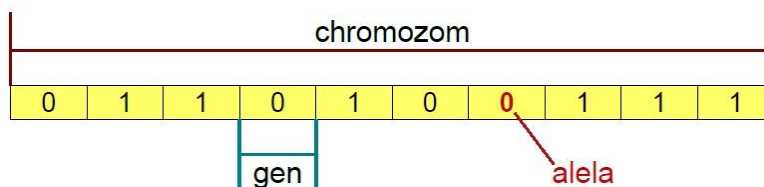
Zdroj: http://en.wikipedia.org/wiki/File:Metaheuristics_classification.svg

Mezi základní techniky evolučních algoritmů (EA) řadíme [5]:

- *Genetické algoritmy (Genetic algorithm)*
 - Asi nejpopulárnější z technik EA užívaná k řešení optimalizačních úloh. Touto technikou se podrobněji zabývám v navazujících kapitolách.
- *Genetické programování (Genetic programming)*
 - Jedinec reprezentuje počítačový program/algoritmus, jehož ohodnocení je určeno jeho schopností řešit daný problém a je uchovávan ve formě stromu.
- *Evoluční programování (Evolutionary programming)*
 - Podobné GP, ale struktura programu je pevně dána a mění se hodnoty jednotlivých proměnných – více využívá operátoru mutace.
- *Evoluční strategie (Evolution strategy)*
 - Jedinec je typicky reprezentován formou vektoru reálných čísel. Časté je užití proměnné míry mutace.
- *Diferenciální evoluce (Differential evolution)*
 - Řeší problémy číselné optimalizace. Navrženo k hledání extrémů multimodálních funkcí.

3.5 Genetický algoritmus

Genetický algoritmus, jakožto typický zástupce algoritmů evolučních, pracuje s populací potenciálních řešení – jedinců. Jedinec, označovaný také jako „chromozom“ se skládá z řetězce „genů“, jejichž hodnota se pak nazývá „alela“. Pořadí genů v chromozomu je pevně dáno – každý má svůj vlastní význam, poněvadž přípuštěním jejich permutace by byla prakticky znemožněna jakákoliv smysluplná interpretace řešení. Jedince tedy označujeme za shodné, pokud jsou jejich alely v odpovídajících genech shodné. Odtud plyne i to, že všechna potenciální řešení jsou stejné délky.



Obr. 2 Struktura jedince
Zdroj: vlastní zpracování

Jednotlivá řešení bývají nejčastěji reprezentována jako řetězec hodnot – alel, potažmo genů, jejichž hodnoty jsou nejčastěji číselné – pak označujeme tento řetězec za „vektor“. Nicméně reprezentace řešení by měla co nejvíce odpovídat danému problému. Proto se setkáváme i s jinými reprezentacemi, nejčastěji v podobě grafů – stromů (genetické programování).

Populace je definována jako množina těchto jedinců. Cílem GA je „šlechtit“ taková řešení, která se přibližují řešení optimálnímu, dokud nejsou splněny ukončovací podmínky algoritmu. Toho se docílí užitím různých výběrových strategií (selekce), jež určují, jaká potenciální řešení vybrat pro tvorbu nových a snad i lepších řešení. Aby GA mohl využívat tyto výběrové strategie, musí být každému řešení přiřazeno ohodnocení (fitness) tzv. ohodnocující funkcí dle jeho kvality (jak moc splňuje podmínky na něj kladené – tj. podmínky, které chceme, aby plnilo řešení optimální).

Za základní vlastnosti genetických algoritmů lze tedy (mimo jiných) považovat:

- Genetická reprezentace řešení
- Populace řešení
- Ohodnocující funkce řešení
- Evoluční operátory
- Generační vývoj

3.5.1 Schéma GA

Podle způsobu práce s populací potenciálních řešení rozpoznáváme dva standardní modely genetických algoritmů: GGA a SGA. Oba typy zde stručně popíši, ale v dalších částech se budu zabývat obvyklejším schématem GGA.

3.5.1.1 Steady-state GA (SGA)

Je nutné zmínit, že stejnou zkratkou – SGA bývá také často označován tzv. jednoduchý GA (Simple GA).

SGA vytváří obvykle pouze jednoho, nebo dva potomky z každé generace (průchodu algoritmu), kteří nahradí stávající jedince populace (jednoho až dva). Je zřejmé, že v populaci se tedy nekonají rozsáhlé změny a je v ní udržován přibližně stálý stav -> „Steady-state“ GA.

Díky rozdílné funkci algoritmu se SGA jeví rychlejší než GGA, ale při větším počtu iterací může vykazovat horší výsledky, jelikož neprohledává tak důkladně prostor potenciálních řešení [3, s. 138].

Základní kroky SGA [3, s. 138]:

1. Vyber dva jedince (rodiče) z populace P .
2. Vytvoř potomka užitím evolučních operátorů (křížení a mutace).
3. Ohodnoť potomka.
4. Vyber jedince z populace P , který má být nahrazen potomkem.
5. Rozhodni, bude-li tento jedinec skutečně nahrazen.

3.5.1.2 Generational GA (GGA)

GGA obměňuje při každém průchodu algoritmu většinou část až celou populaci nově vytvořenými jedinci, čímž vzniká nová generace odlišná od té předchozí -> "Generační" GA.

Tři základní operace GGA [3, s. 136]:

1. Ohodnot' jedince v populaci.
2. Vyber jedince vhodné pro tvorbu nové populace (selektce).
3. Aplikuj evoluční operátory (křížení a mutace) na vybrané jedince.

Obecné schéma GGA [10, s. 20]:

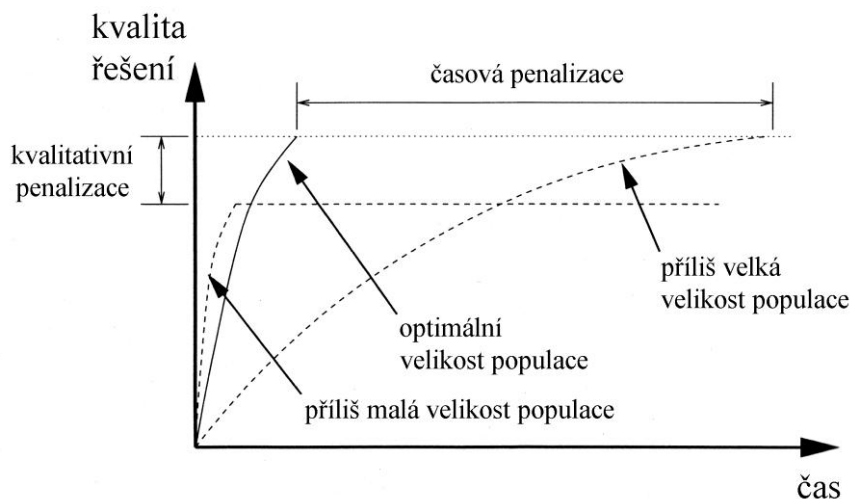
1. Vynuluj hodnotu počítadla generací $t=0$.
2. Náhodně vygeneruj počáteční populaci $P(0)$.
3. Vypočítej ohodnocení (fitness) každého individua v počáteční populaci $P(0)$.
4. Vyber dvojice individuí z populace $P(t)$ a vytvoř jejich potomky $P'(t)$.
5. Ohodnot' nově vytvořená individua v $P'(t)$.
6. Vytvoř novou populaci $P(t+1)$ z původní populace $P(t)$ a množiny potomků $P'(t)$.
7. Zvětši hodnotu počítadla generací o jedničku ($t=t+1$).
8. Vypočítej ohodnocení (fitness) každého individua v populaci $P(t)$.
9. Pokud t je rovno maximálnímu počtu generací nebo je splněno jiné ukončovací kritérium, vrať jako výsledek populaci $P(t)$; jinak pokračuj krokem číslo 4.

3.5.2 Počáteční populace

Prvotní populace bývá obvykle vytvořena zcela náhodně – aby byla zajištěna její dostatečná diverzita a my si tak neomezili prohledávaný prostor řešení. Lze se ovšem setkat i s generováním jedinců za pomoci nějaké heuristické metody, jež ale musí být vystavěna na znalostech daného konkrétního problému.

Běh algoritmu může významně ovlivnit už samotná velikost populace, ale odhad její správné velikosti, aby GA poskytoval co nejlepší výsledky při řešení daného problému, je velmi náročný.

Při příliš malé velikosti populace hrozí nebezpečí uváznutí v lokálním extrému a naopak, při příliš velkém počtu jedinců bude běh algoritmu výrazně delší, ale nemusí přinést žádné zlepšení kvality řešení. (viz Obr. 3) [11, s. 222]



Obr. 3 Vliv velikosti populace GA
Zdroj: upraveno z [13, s. 222]

„Bylo prokázáno, teoreticky i empiricky, že optimální velikost je něco, co se odlišuje problém od problému. Navíc, někteří vědci pozorovali, že v různých stádiích jednoho běhu (GA) může být optimální rozdílná velikost populace.“ [12, s. 185]

Z tohoto důvodu se setkáváme s aplikacemi tzv. schémat s adaptivní velikostí populace (Adaptive Population Sizing Schema), které se pokouší určit správnou velikost populace za vlastního běhu genetického algoritmu.

3.5.3 Ohodnocení jedinců

Jak již bylo vícekrát zmiňováno, každé řešení má svou přiřazenou hodnotu fitness, jež určuje jeho kvalitu. Toto jejich ohodnocení pak hraje významnou úlohu v dalším rozhodování o tom, jestli se bude jedinec podílet na tvorbě nové populace. Toto ohodnocování provádí tzv. ohodnocující či fitness funkce, jež je navržena tak, aby odrážela specifika řešeného problému.

3.5.4 Metody výběru jedinců (selekce)

Existuje celá řada metod napodobujících různými způsoby tzv. „přírodní výběr“ a dále nespočet jejich variant a modifikací. To, jakým mechanismem vykonávají svou funkci, se může značně odlišovat – od prostého porovnávání jedinců, přes pseudonáhodný výběr, po složité a komplexní heuristiky. Lehce se tak můžeme setkat s metodami kombinujícími prvky dvou, či více základních metod.

Přehled často používaných metod k výběru jedinců přežívajících do nové generace [13]:

- *Proporcionální selekce (Fitness-proportionate selection)*
 - Šance jedince na výběr je úměrná velikosti jeho ohodnocení vůči ohodnocení ostatních jedinců. Jedinci s větším ohodnocením mají tedy větší pravděpodobnost vybrání.
 - *Ruletová selekce (Roulette-wheel selection)*
 - Zástupce proporcionálního typu selekce.
 - Často bývá reprezentována jako „ruletové“ kolo dělené na díly úměrné velikosti ohodnocení jedinců v rámci populace. Následně je náhodně zvolena pozice na ruletě a tím i jí odpovídající jedinec.
- *Pořadová selekce (Rank selection)*
 - Jedincům je přiřazeno číslo reprezentující jejich pořadí v soutěži o nejlepší ohodnocení v rámci dané populace. Následná selekce probíhá na základě tohoto pořadí, nikoli velikosti rozdílů v ohodnocení jedinců.
 - Výhodou je zamezení případů, kdy vysoce hodnocený jedinec dominuje ve výběru na úkor hůře hodnoceným a tím omezuje diverzitu dalších populací.
 - Bývá také reprezentováno ve formě modifikovaného ruletového kola.
- *Turnajová selekce (Tournament selection)*
 - Populace je dělena na skupiny jedinců, kteří mezi sebou „soutěží“ v rámci daných skupin. Vybírání jsou vítězové skupin.

Často používaným prvkem selekce je [13]:

- *Elitismus (Elitist selection)*
 - Jedinci s největším ohodnocením přecházejí do následující generace v nezměněné podobě. Jejich počet bývá omezen.
 - Zabezpečuje přežití nejkvalitnějších jedinců stávající populace.

Selekční metoda může být vnitřně upravena i například takto [13]:

- *Hiearchická selekce (Hierarchical selection)*
 - Každá generace prochází několikanásobným výběrem. Jedinci, kteří projdou prvotními stadii výběru (rychlý a zběžný), jsou hodnoceni důkladněji.
 - Výhodou této metody je redukce časové náročnosti výpočtu díky rychlému odstranění zjevně nevyhovujících jedinců a následnému se zaměření na jedince slibnější, kteří prošli úvodním výběrem.
- *Selekce se vzrůstajícím „přísností“ ohodnocování (Scaling selection)*
 - Se vzrůstajícím průměrným ohodnocením populace roste i tzv. „selektivní tlak“ a ohodnocující funkce se stává přísnější.
 - Je užitečné u pozdějších generací, kde má více jedinců podobně vysoké ohodnocení.

3.5.5 Křížení

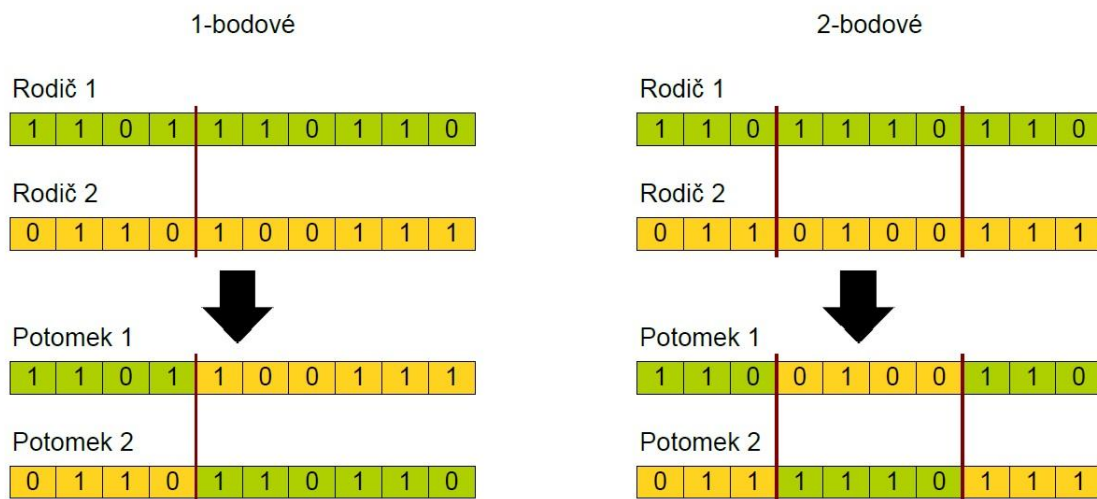
Křížení (spolu s mutací) představuje jeden z nejrozšířenějších evolučních operátorů. Proto se můžeme setkat s nepřeberným množstvím jeho modifikací vzniklých za účelem optimalizace jeho činnosti vzhledem k řešené úloze a reprezentaci jedinců v populaci. Přesto zůstává jeho základní princip činnosti zachován.

Operátor křížení pracuje s „rodiči“ (vybranými jedinci populace) a křížením jejich chromozomů vytváří nové jedince – „potomky“. Obvykle se uvádějí příklady se dvěma rodiči, jejichž zkřížením vznikají dva potomci, ale není těžké upravit operátor pro práci s více jedinci najednou.

Základní verzí tohoto operátoru je tzv. jedno-bodové křížení. Náhodně se zvolí pozice v chromozomu a od ní se vymění jeho zbývající část mezi oběma rodiči, čímž dostaneme potomky složené z genů obou rodičovských jedinců. Těchto dělicích pozic je možno zvolit i více (např. dvou-bodové křížení) (viz Obr. 4).

Operátor křížení se obecně aplikuje na vybrané jedince s vysokou mírou pravděpodobnosti 0,75 až 0,95 [17].

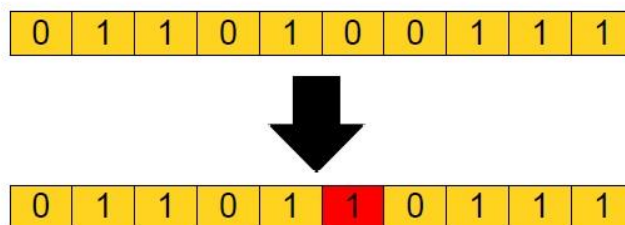
V jiných modifikacích se nad zvolenými úseky provádějí různé matematické či logické operace.



Obr. 4 Křížení
Zdroj: vlastní zpracování

3.5.6 Mutace

Dalším evolučním operátorem je mutace. Principiálně jde o velmi jednoduchou operaci: změnu jednoho, či více náhodně vybraných genů v chromozomu. Používání tohoto operátoru pomáhá diverzifikovat populaci a brání tak uváznutí v lokálním optimu.



Obr. 5 Mutace
Zdroj: vlastní zpracování

V zájmu zachování kvalitních jedinců je mutace prováděna pouze u malého počtu jedinců – míra pravděpodobnosti od 0,001 do 0,05 [17]. Setkáváme se také s proměnlivou mírou mutace – např. metoda simulovaného žhání (*Simulated Annealing*), jestliže hrozí uváznutí v lokálním extrému, nebo je malá diverzita populace, zvýší se přechodně míra pravděpodobnosti tohoto evolučního operátoru. Tak jako u křížení, se můžeme setkat se složitějšími operacemi nad geny v závislosti na typu jejich obsahu.

3.5.7 Nová populace

Aplikací různých operátorů, jako je křížení a mutace, dostáváme jedince tvořící novou populaci. V závislosti na užití selekční strategii mohou přecházet do nové populace i jedinci s vysokým ohodnocením z předcházející populace – tzv. elitismus.

3.5.8 Ukončení běhu algoritmu

Proces generování nové populace a následného ohodnocování jejích jedinců je opakován až do okamžiku splnění předem definované ukončovací podmínky.

Časté formulace těchto podmínek jsou:

- Nalezení řešení splňující požadovaná kritéria
- Dosažení zadaného limitu (maximálního povoleného počtu) iterací běhu algoritmu (generací)
- Nedostatečné zlepšení ohodnocení jedinců v posledních k generacích
- Vyčerpání přiděleného strojového času
- Stop za účelem kontroly (mezi)výsledku - v případě, že tuto provádí člověk, mluvíme o tzv. interaktivní evoluci
- Kombinace předchozích podmínek

3.6 Závěr

Cílem této kapitoly bylo nenásilně uvést čtenáře do problematiky genetických algoritmů. Vzhledem k tomuto určení nezabíhá do zbytečných podrobností a snaží se provést čtenáře touto zajímavou oblastí algoritmizace stručně a přehledně.

4 Selekční metody

4.1 Úvod

Cílem této části je popis různých základních selekčních metod (schémat) a jejich vzájemné srovnání.

Selekční metody, jak už bylo uvedeno v předchozích kapitolách této práce, jsou různé verze mechanismu výběru rodičovských jedinců zastávající funkci tzv. přirozeného výběru.

Přestože existuje mnoho různých modifikací těchto metod, všechny mají společný jeden požadavek – pro svou funkci potřebují ohodnocení jedinců (fitness). Podle způsobu práce výběrové metody může tato klást požadavky (omezení) na hodnoty, jichž může ohodnocení nabývat. Častý bývá požadavek například na nezáporné hodnoty. Jedinci jsou metodou vybíráni s pravděpodobností odvozenou od jejich ohodnocení.

Nejprve vždy rozeberu jednotlivé metody z hlediska principu jejich vnitřního fungování a následně uvedu jejich porovnání v několika možných charakteristikách vystihujících vliv na jimi vytvářenou populaci.

Pro popis selekčních metod bylo a je stále používáno vícero vzájemně více či méně podobných veličin – termínů. V následujícím textu používám tyto termíny tak, jak jsou definovány a používány v [1] a [2] a následně shrnuty v [16]:

- rozpětí (spread)
 - rozsah možných hodnot počtu potomků jedince
- zkreslení (bias)
 - absolutní rozdíl mezi normalizovaným ohodnocením jedince a očekávanou pravděpodobností jeho reprodukce
- tlak selekce (selective pressure)
 - pravděpodobnost výběru nejlepšího jedince ve srovnání s průměrnou pravděpodobností výběru všech jedinců
- ztráta diverzity (loss of diversity)
 - podíl jedinců v populaci, kteří nejsou vybráni během fáze selekce

- intenzita selekce (selection intensity)
 - očekávaná průměrná hodnota ohodnocení populace po použití výběrové metody na normalizované Gaussovo rozdělení
- rozptyl selekce (selection variance)
 - očekávaný rozptyl rozdělení ohodnocení populace po použití metody výběru na normalizované Gaussovo rozdělení

4.2 Proporcionální metody

Pro metody této kategorie, v anglickém jazyce označované jako „*Fitness-proportionate selection*“, platí, že pravděpodobnost výběru jedince je přímo úměrná velikosti jeho ohodnocení. Jedinec s větším ohodnocením tedy bude vybrán s větší pravděpodobností, než jiný jedinec hůře hodnocený.

Protože selekční metody z této skupiny pracují přímo s ohodnocením jedinců, nevyžadují jejich seřazení podle velikosti jejich ohodnocení.

U proporcionálních metod se můžeme setkat s pojmem: „problém měřítka“ (scaling problem). Jedná se o stagnaci při příliš nízké hodnotě selektivního tlaku, nebo předčasnou konvergenci při příliš rychle zúženém výběru.

4.2.1 Ruletová selekce

Název této metody je doslovným překladem jejího anglického označení „*Roulette-wheel selection*“. Odtud také pochází frekventovaná zkratka RWS. Vzácněji se vyskytuje pod názvem „stochastické vzorkování s náhradou (stochastic sampling with replacement) [1].

Tato výběrová metoda je zřejmě nejčastěji používanou formou implementace přirozeného výběru vůbec. [10, s. 22] [14, s. 124]

Přes svou oblíbenost není ruletová selekce vhodná pro každou úlohu. Její hlavní nevýhoda plyne přímo z principu její funkce – totiž přímé úměrnosti pravděpodobnosti výběru vůči ohodnocení jedince. Získá-li tedy jeden, či více jedinců vysoce nadprůměrné ohodnocení, budou slabší jedinci při selekci ve značné nevýhodě a mohou být zcela vytlačeni z dalších generací.

Pravděpodobnost výběru i -tého jedince v populaci o velikosti $N > 0$ s ohodnocením $f_i \geq 0$, pro $i = 1, \dots, N$, a současně i velikost jemu přiřazené výseče je dána vztahem:

$$p_i = \frac{f_i}{\sum_{j=1}^N f_j}, \text{ kde } i \in \{1, \dots, N\}$$

Pro názornost uvedeme jednoduchý příklad. V tabulce číslo 1, je pět jedinců. Znamená to tedy, že ruleta bude tvořena pěti výsečemi, jejichž plocha bude korespondovat s velikostí ohodnocení příslušných jedinců. Dále můžeme vyčíst pravděpodobnost jejich vybrání (p_i). Použijeme-li dále kumulované ohodnocení, tak jak je v tabulce, pak stačí pro výběr jedince vygenerovat náhodné číslo r z intervalu $\langle 0;1 \rangle$. Poslední jedinec, jehož kumulované ohodnocení je menší, nebo rovno r je vybrán. Na příklad pro $r = 0,348$ je vybrán jedinec číslo 2.

Tabulka 1 Příklad populace

Jedinec (i)	Ohodnocení (f_i)	% z celkového ohodnocení (p_i)	Kumulované ohodnocení
1	3	15	0,15
2	6	30	0,45
3	2	10	0,55
4	5	25	0,80
5	4	20	1,00
Celkové ohodnocení populace			20
Průměrné ohodnocení populace			4

Zdroj: vlastní zpracování

Selekční mechanismus ovšem nemusí nutně používat pro výběr jedince hodnoty z intervalu $\langle 0;1 \rangle$, stačí změnit horní hranici intervalu na hodnotu celkového ohodnocení populace a hodnoty generovat v tomto širším rozmezí.

Následující algoritmus platí pro obě varianty [20, s. 175]:

1. *Suma všech účelových hodnot v populaci – S*
2. *Vygenerování náhodného čísla z intervalu $(0,S)$ – r*
3. *Průchod populací a počítání sumy účelových hodnot od hodnoty nula – s*

Jakmile se hodnota s stane větší než hodnota r , je vybrán jedinec do reprodukčního pole.

V nezměněné podobě se jedná o neefektivní implementaci vyhledávání konkrétního jedince v populaci. Efektivitu můžeme zvýšit použitím už výše zmiňovaného kumulovaného ohodnocení, čímž zabráníme opakovanému počítání sumy ohodnocení jedinců, pro každého hledaného jedince zvlášť. [10, s. 22]

Přestože je metoda známá pod názvem „Ruletová selekce“ a proto se běžně hovoří o „výseči“, je velmi snadné provést rozvinutí rulety-kružnice na úsečku a jednotlivé výseče převést na její úseky, jak je předvedeno na příkladu níže (převzatém z [16]):

Tabulka 2 Pravděpodobnost selekce a hodnota fitness

Číslo jedince (i)	1	2	3	4	5	6	7	8	9	10
Ohodnocení (f_i)	2,0	1,8	1,6	1,4	1,2	1,0	0,8	0,6	0,4	0,2
Pravděpodobnost výběru (p_i)	0,18	0,16	0,15	0,13	0,11	0,09	0,07	0,06	0,03	0,02
Kumulativní ohodnocení	0,18	0,34	0,49	0,62	0,73	0,82	0,89	0,95	0,98	1,00

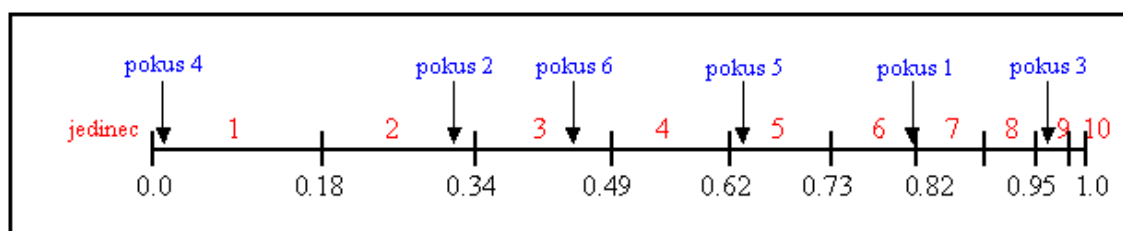
Zdroj: upraveno a rozšířeno z [16]

Výběr jedinců určených k tvorbě potomstva provedeme nezávislým, náhodným vygenerováním čísel z intervalu od nuly do jedné.

Příklad šesti náhodných čísel:

0,81; 0,32; 0,96; 0,01; 0,65; 0,42

Obrázek ukazuje proces výběru jedinců na „jednotkové úsečce“ (znázornění intervalu $<0;1>$) pro šest výše zvolených čísel.



Obr. 6 Ruletová selekce (RWS)

Zdroj: upraveno z [16]

Už z tabulky a následně i obrázku je lehce patrné, že byli vybráni jedinci označení číslem:

6, 2, 9, 1, 5, 3

Tato výběrová metoda má nulové zkreslení (bias), ale současně už negarantuje minimální rozpětí (spread). [16]

4.2.2 Stochastické univerzální vzorkování

Anglický název této metody – „Stochastic universal sampling“ bývá často reprezentován zkratkou SUS.

Jedná se o selekční metodu odvozenou od ruletové metody (RWS) navrženou Bakerem [1], jenž má kromě nulového zkreslení (bias) i minimální rozpětí (spread). [16]

Tak jako u RWS, jsou i zde jedincům přiřazeny části intervalu od nuly do jedné (jednotkové úsečky) přímo úměrné velikosti jejich ohodnocení.

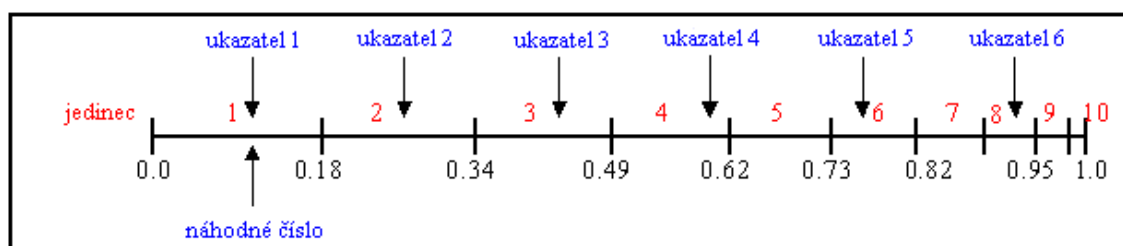
Rozdíl spočívá v tom, že se výběr neprovádí jednotlivě – nezávisle na sobě, nýbrž hromadně. Výběr požadovaného počtu jedinců se provede v jediném kroku. To znamená, že pro každého jedince máme jeden (hypotetický) ukazatel na naší jednotkové úsečce. Díky tomu, že jejich vzájemná vzdálenost je konstantní, máme jistotu rovnoměrného pokrytí celé úsečky – populace.

Rozestup mezi ukazateli dostaneme jako převrácenou hodnotu jejich počtu – tj. počtu vybíraných jedinců. Pozici prvního ukazatele určíme jakožto náhodně vygenerované číslo z intervalu od nuly do hodnoty předtím spočítané převrácené hodnoty. Viz následující příklad (převzato z [16]):

Pro výběr požadovaných šesti jedinců, je vzdálenost mezi ukazateli dána takto:

$$\frac{1}{6} = 0,167$$

Náhodně vygenerované číslo z intervalu $\langle 0; 0,167 \rangle$ je 0,1.



Obr. 7 Stochastické univerzální vzorkování (SUS)

Zdroj: upraveno z [16]

Vybrání byli jedinci označení číslem:

1, 2, 3, 4, 6, 8

4.3 Pořadová selekce

Je dalším rozšířeným zástupcem výběrových metod, známým ve světě pod svým anglickým názvem „Rank selection“.

Pořadová selekce se od proporčních metod liší ve své funkci v jednom významném prvku – v kritériu podle něhož jsou vybíráni rodičovští jedinci. Seřazení populace podle kvality ohodnocení jejích jedinců umožňuje přidělit jedinci číslo určující jeho pořadí v rámci populace. Díky tomu, že následný výběr neprobíhá podle ohodnocení, ale pořadí, odpadá u této metody možnost výrazně nevyvážené pravděpodobnosti selekce jedince, kdy jeden či několik kvalitních jedinců ovládne výběr na úkor méně kvalitních individuů.

Metoda může být tak jako zástupci proporční selekce implementována v podobě upraveného ruletového kola, či výběru na (jednotkové) úsečce.

Pořadová selekce umožňuje jednoduchou kontrolu selekčního tlaku a zároveň řeší problém měřítka (scaling problem).

4.3.1 Lineární pořadová selekce

Ohodnocení jedince je dáno následujícím vztahem (upraveno z [16]), kde:

- N – velikost populace (počet jedinců)
- Pos – pozice jedince v populaci
(nejhorší jedinec má $Pos = 1$, nejlepší $Pos = N$)
- SP – tlak selekce (selective pressure) v intervalu $\langle 1,0; 2,0 \rangle$

$$Fitness(Pos) = 2 - SP + \frac{2(SP - 1)(Pos - 1)}{N - 1}$$

4.3.2 Nelineární pořadová selekce

Metoda nelineární pořadové selekce umožňuje vyšší hodnoty selektivního tlaku oproti metodě lineární.

Ohodnocení jedince pro tuto metodu je dáno vztahem (upraveno z [16]):

$$Fitness(Pos) = \frac{N * X^{Pos-1}}{\sum_{i=1}^N X^{i-1}}$$

X je spočítáno jako kořen polynomu:

$$(SP - N) * X^{N-1} + SP * X^{N-2} + \dots + SP * X + SP = 0$$

SP nabývá hodnot v intervalu $\langle 1,0; N-2,0 \rangle$.

4.3.3 Srovnání lineární a nelineární pořadové metody

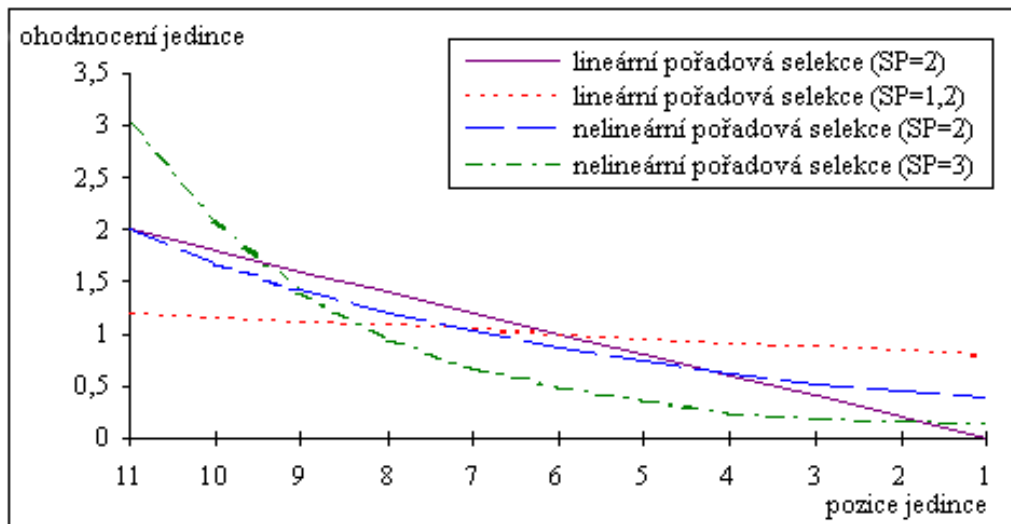
Tabulka číslo 3 obsahuje hodnoty ohodnocení jedinců v populaci o velikosti $N = 11$ při různých hodnotách selektivního tlaku pro lineární i nelineární pořadovou selekci. Z hodnot uvedených v tabulce a dále i v grafu níže (Obr. 8), je na první pohled patrné oproštění se od přímé závislosti na objektivním ohodnocení jedinců. Přestože jsou tyto hodnoty velmi rozdílné (v rozsahu od 1 do 95), je ohodnocení přiřazené těmto jedincům ve výrazně menším intervalu a s pravidelnými rozestupy.

Tabulka 3 Závislost hodnoty fitness na tlaku selekce

objektivní ohodnocení	ohodnocení v závislosti na parametru: tlak selekce				
	lineární řazení		bez řazení	nelineární řazení	
	2,0	1,1	1,0	3,0	2,0
95	2,0	1,10	1,0	3,00	2,00
30	1,8	1,08	1,0	2,21	1,69
20	1,6	1,06	1,0	1,62	1,43
15	1,4	1,04	1,0	1,99	1,21
10	1,2	1,02	1,0	0,88	1,03
9	1,0	1,00	1,0	0,65	0,87
8	0,8	0,98	1,0	0,48	0,74
7	0,6	0,96	1,0	0,35	0,62
4	0,4	0,94	1,0	0,26	0,53
3	0,2	0,92	1,0	0,19	0,45
1	0,0	0,90	1,0	0,14	0,38

Zdroj: upraveno z [16]

Za pozornost stojí určitě prostřední sloupec tabulky „bez řazení“, jenž udává pro hodnotu selektivního tlaku rovnu 1,0 všem jedincům ohodnocení totožné hodnoty. Přestože tedy patří hodnota $SP = 1,0$ do intervalu pro lineární pořadovou selekci, jejím použitím se z této metody stává prostý náhodný výběr.



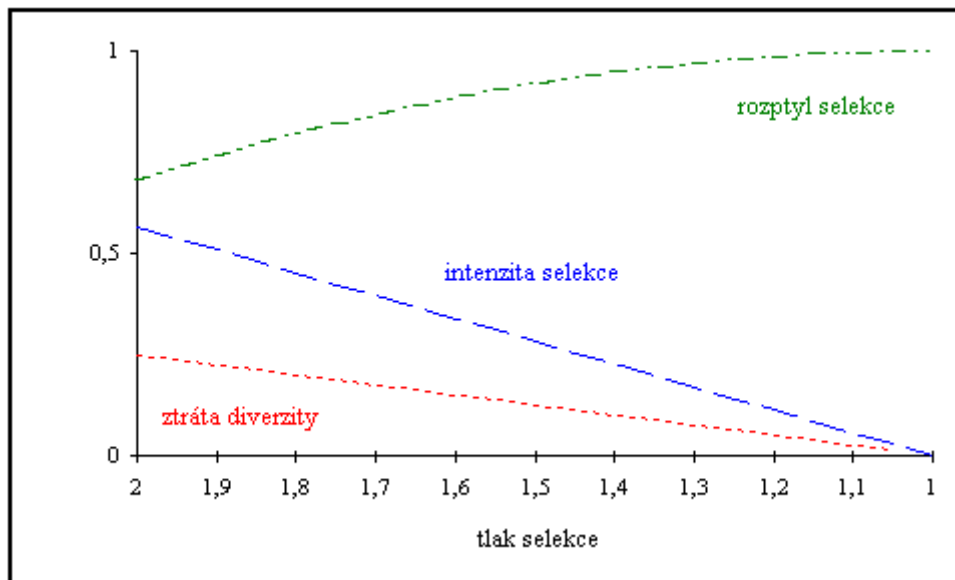
Obr. 8 Přiřazení fitness pro lineární a nelineární pořadovou selekci

Zdroj: upraveno z [16]

Graf zachycuje průběh ohodnocujících funkcí obou typů pořadové selekce – lineární i nelineární pro konkrétní hodnoty tlaku selekce – tak jak jsou uvedeny v legendě pravém horním rohu grafu.

4.3.4 Analýza lineární pořadové metody

Graf zachycuje závislost třech veličin používaných pro popis výběrových metod v závislosti na tlaku selekce. Jejich konkrétní hodnotu lze odečíst na svislé ose grafu.



Obr. 9 Vlastnosti lineární pořadové selekce

Zdroj: upraveno z [16]

Opět je zde vidět, že pro tlak selekce $SP = 1$ je intenzita spolu s diverzitou selekce rovna 0 a zároveň pak rozptyl roven 1. Logicky tedy platí: čím je menší intenzita, tím je větší rozptyl.

Rovnice převzaty a upraveny z [16]:

Intenzita selekce

$$SelInt_{LinRank}(SP) = \frac{SP - 1}{\sqrt{\pi}}$$

Ztráta diverzity

$$LossDiv_{LinRank}(SP) = \frac{SP - 1}{4}$$

Rozptyl selekce

$$SelVar_{LinRank}(SP) = 1 - SelInt_{LinRank}(SP)^2$$

$$SelVar_{LinRank}(SP) = 1 - \frac{(SP-1)^2}{\pi}$$

4.4 Turnajová selekce

V anglickém jazyce je označována jako „Tournament selection“.

Jak je patrné už z názvu, tato metoda je založena na principu turnajového výběru – lidově označovaného termínem „pavouk“.

Z populace je zcela náhodně vybrána skupina jedinců, ze kterých je vybrán jedinec (rodič) s nejlepším ohodnocením. Tento postup opakujeme dokud nemáme požadovaný počet rodičovských jedinců pro „tvorbu“ potomstva do nové populace. Stěžejním parametrem této metody je právě velikost vybíraných skupin – „velikost turnaje“ (tournament size) dále označovaná *Tour*. Velikost tohoto parametru lze teoreticky volit v intervalu od 1 do N (velikost populace), ale tyto krajní hodnoty nejsou vhodné pro správnou funkci metody. Při $Tour = 1$ se totiž jedná o zcela náhodný výběr a při $Tour = N$ dojde vždy k výběru nejlépe hodnoceného jedince z celé populace. Odtud je snadné nahlédnout, jak lze přímo ovlivnit intenzitu výběru změnou velikosti parametru *Tour*. Čím větší bude tato hodnota, tím větší bude pravděpodobnost výběru vysoce hodnoceného jedince, a naopak.

Tabulka 4 Vztah mezi velikostí turnaje a intenzitou selekce

velikost turnaje	1	2	3	4	5	7	10	20	30	40
intenzita selekce	0	0,56	0,84	1,03	1,16	1,35	1,54	1,87	2,04	2,16

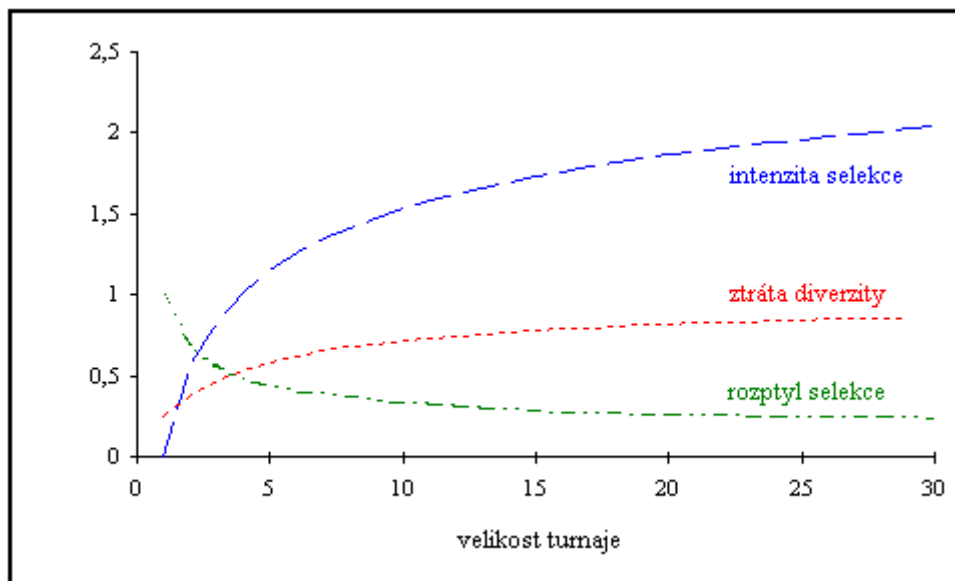
Zdroj: upraveno z [2, s. 46] a [16]

Tabulka ukazuje přehledně závislost mezi velikostí turnajové skupiny a intenzitou selekce. Čím je skupina větší, tím větší je konkurence v ní, a tedy i počet „poražených“ jedinců.

Za zmínku jistě stojí Goldberg a Deb [7, s. 79], kteří navrhli úpravu binární metody ($Tour = 2$) přidáním nového parametru, určujícího pevně danou pravděpodobnost p vítězství silnějšího jedince v rozsahu $0.5 < p \leq 1$.

4.4.1 Analýza turnajové metody

Na grafu je zachycena výše popsaná závislost intenzity selekce na velikosti parametru $Tour$. Z grafického znázornění je zřejmé, že intenzita selekce roste současně se ztrátou diverzity při zvyšování velikosti turnaje, přičemž rozptyl naopak klesá. Pro $Tour = 1$ (náhodný výběr) je intenzita selekce rovna 0 a rozptyl má hodnotu 1.



Obr. 10 Vlastnosti turnajové selekce

Zdroj: upraveno z [16]

Rovnice převzaty a upraveny z [16]:

Intenzita selekce

$$SelInt_{Tournament}(Tour) \approx \sqrt{2 * (\ln Tour - \ln \sqrt{4,14 * \ln Tour})}$$

Ztráta diverzity

$$LossDiv_{Tournament}(Tour) = Tour^{\frac{-1}{Tour-1}} - Tour^{\frac{-Tour}{Tour-1}}$$

Při turnaji o velikosti $Tour = 5$ je ztraceno přibližně 50 % populace.

Rozptyl selekce

$$SelVar_{Tournament}(Tour) \approx \frac{0,918}{\ln(1,186+1,328*Tour)}$$

4.5 Ořezová selekce

Pro účely této práce byl originální název „Truncation selection“ této výběrové metody přeložen jako „ořezová“ selekce či výběr.

V porovnání s předešlými selekčními metodami, modelujícími přirozený výběr, se jedná o umělou výběrovou metodu, používanou například v oboru populační genetiky – tedy u velkých populací a masové selekce.

Ořezová selekce vyžaduje setřídění populace podle hodnot ohodnocení jejích jedinců. Pro rozmnožování jsou vybrána pouze nejlepší individua. Parametrem metody je tak zvaný „ořezový práh“ (truncation treshold) označovaný jako *Trunc*, jehož hodnota určuje velikost části populace, jež bude vybrána. Běžně používané hodnoty *Trunc* se pohybují v rozmezí 10 % až 50 %. Jedinci s ohodnocením spadajícím pod tuto hranici (práh) nemohou už do dalšího běhu algoritmu zasáhnout.

Vybraní jedinci produkují potomky náhodně – tj. jedinci, jenž prošli ořezem, jsou vybírání se stejnou pravděpodobností, bez ohledu na ohodnocení.

Tabulka 5 Vztah mezi prahem ořezu a intenzitou selekce

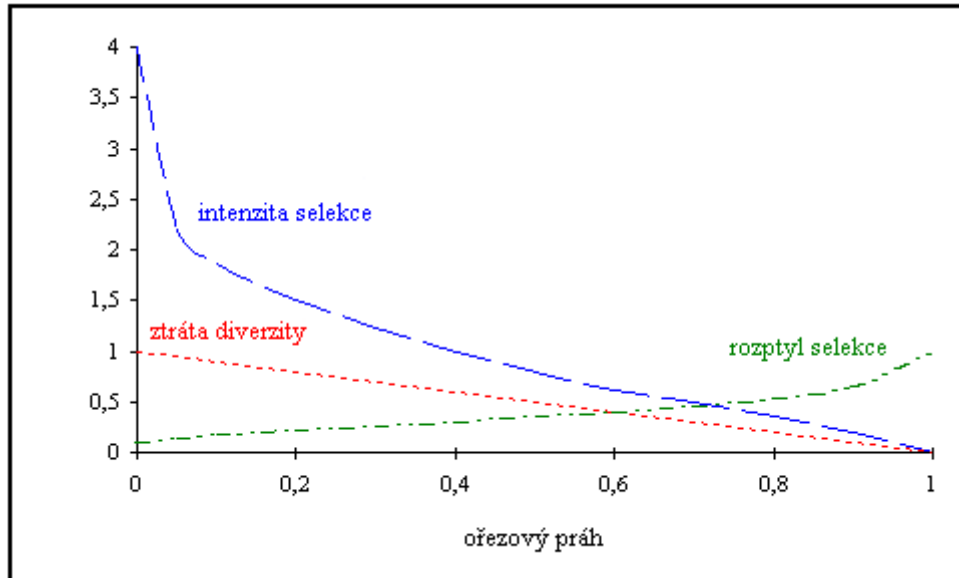
práh ořezu [%]	1	10	20	30	40	50	80
intenzita selekce	2.66	1.76	1.2	1,16	0.97	0.8	0.34

Zdroj: upraveno z [2, s. 46] a [16]

Z tabulky je patrné, že čím je větší práh ořezu, tím menší je intenzita selekce, protože zachováváme více jedinců v populaci.

4.5.1 Analýza ořezové metody

Graf (Obr. 11) zachycuje již popsanou závislost intenzity selekce na hodnotě ořezového prahu, a dále už obvyklou protichůdnou tendenci dalších dvou ukazatelů, když s rostoucí intenzitou (klesající práh) roste i ztráta diverzity a současně tedy klesá rozptyl selekce.



Obr. 11 Vlastnosti ořezové selekce

Zdroj: upraveno z [16]

Rovnice převzaty a upraveny z [16]:

Intenzita selekce

$$SelInt_{Truncation}(Trunc) = \frac{1}{Trunc * \sqrt{2\pi}} * e^{-\frac{f_c^2}{2}}$$

Kde f_c je určeno [2, s. 25]:

$$Trunc = \int_{f_c}^{\infty} \frac{1}{\sqrt{2\pi}} * e^{-\frac{f^2}{2}} df$$

Ztráta diverzity

$$LossDiv_{Truncation}(Trunc) = 1 - Trunc$$

Rozptyl selekce

$$SelVar_{Truncation}(Trunc) = 1 - SelInt_{Truncation}(Trunc) * (SelInt_{Truncation}(Trunc) - f_c)$$

Podrobnější analýzu této metody lze nalézt např. v [2]. Stejné výsledky byly odvozeny jiným způsobem také v [4].

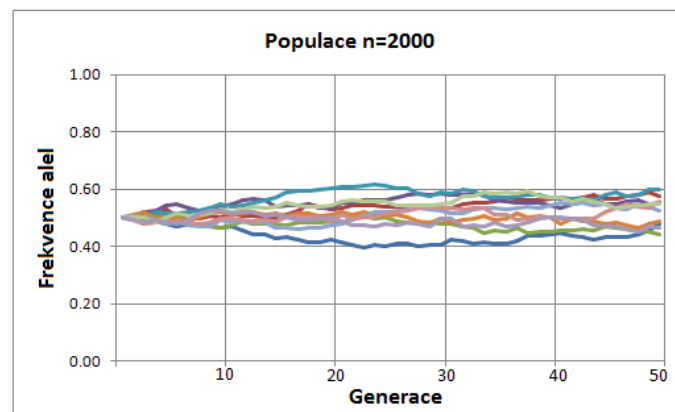
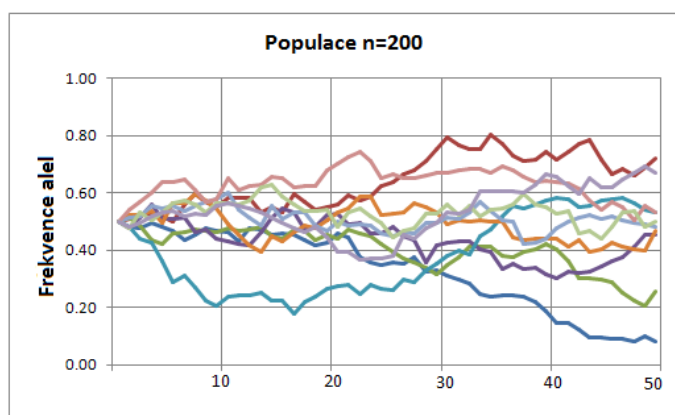
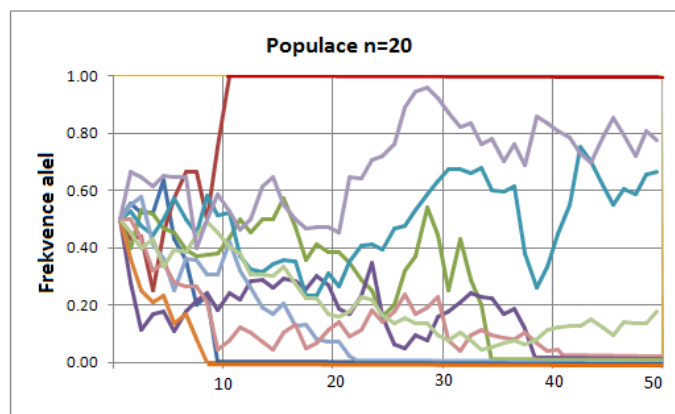
4.6 Genitor / Steady-State

Tato metoda byla představena v úvodní teoretické části této práce, jakožto jedno ze dvou základních schémat genetického algoritmu pod názvem SGA, za nějž je někdy považována [3, s. 136] z důvodu rozdílného způsobu činnosti.

SGA pracuje tzv. jedinec po jedinci, vybírá je podle lineárního pořadí a jejich potomkem nahrazuje momentálně nejhoršího jedince v populaci. Nenajdeme zde tedy klasické výrazně odlišné generace potomstva, ale jednu opakovaně „vylepšovanou“. Z tohoto důvodu je těžké provést porovnání s klasickými selekčními metodami GGA.

„Hlavním rozdílem mezi SGA a GGA je, že pro všech P jedinců populace generovaných v GGA je $2 \cdot P$ výběrů. V důsledku toho, je síla selekce a genetický posun (genetic drift) SGA dvojnásobný GGA. SGA se tudíž jeví dvojnásobně rychlý, ačkoli může ztrácet v dlouhodobém výhledu, protože neprohledává krajinu (prostor potencionálních řešení) tak důkladně jako GGA.“ [3, s. 138]

Genetický posun je jev, při němž dochází k náhodným změnám (posunu) frekvence alel (hodnota genu chromozomu) v populaci. Tyto změny tedy nejsou způsobeny přirozeným výběrem či v našem případě činností umělých selekčních metod a nesouvisí ani s vlastní hodnotou alel (a tedy ani kvalitou celého chromozomu – jedince). Významným se tento jev stává právě u malých populací, kde může dojít až k takzvané „fixaci“ hodnoty alely, při níž bude konkrétní alela shodná ve všech chromozomech populace. Příklad tohoto jevu lze vidět na následujícím obrázku (Obr. 12). Čím je velikost populace menší, tím větší jsou výkyvy frekvence jednotlivých alel. Dokud nedojde k jejich fixaci jak je vidět na grafu s populací velikosti $n=20$.



Obr. 12 Random genetic drift chart

Zdroj: upraveno z http://en.wikipedia.org/wiki/File:Random_genetic_drift_chart.png

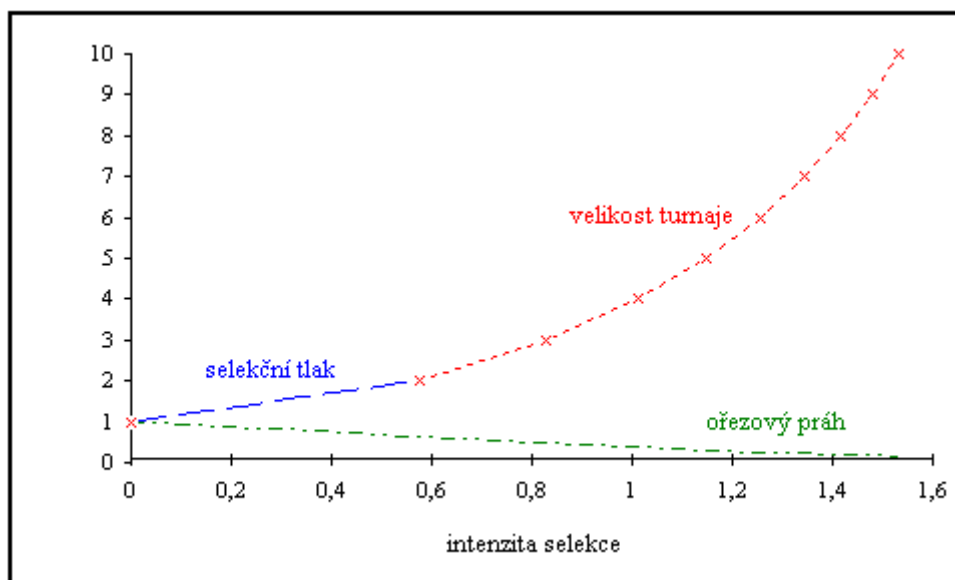
4.7 Srovnání selekčních metod

Z výše uvedeného je vidět, že lze zvolit hodnoty parametrů výběrových metod tak, abychom docílili stejné, či alespoň podobné intenzity selekce. Výsledkem bude často vzájemně podobné chování těchto metod, jak je ukázáno dále.

4.7.1 Parametry selekce a její intenzita

Graf zachycuje vztah mezi intenzitou selekce a následujícími příslušnými parametry výběrových metod:

- Selekční tlak (*SP*) <1,0; 2,0>
- Ořezový práh (*Trunc*) <0,0; 1,0>
- Velikost turnaje (*Tour*) <1; *N*>, kde *N* je velikost populace



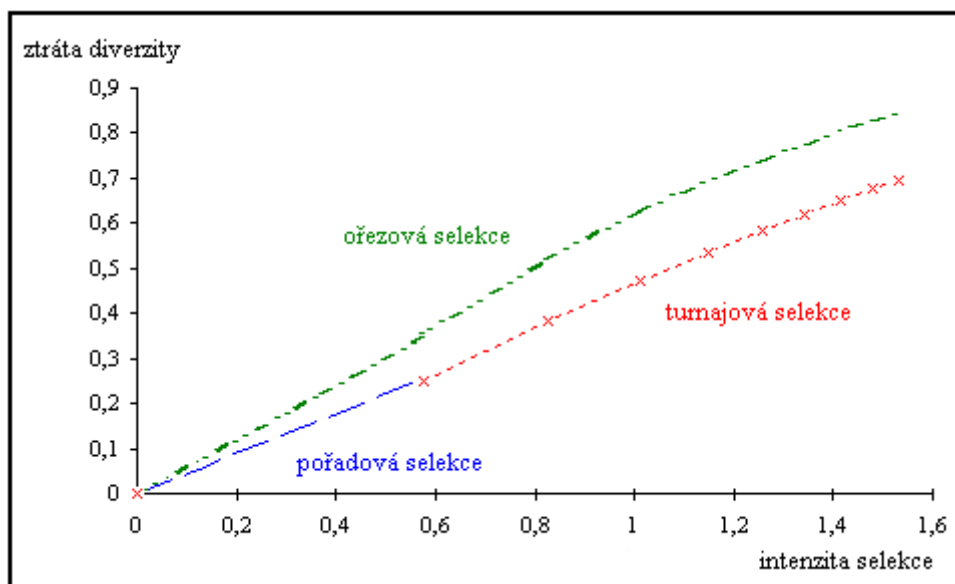
Obr. 13 Závislost parametru selekce na intenzitě selekce

Zdroj: upraveno z [16]

Přestože graf zobrazuje parametry daných metod, v pro ně celém přípustném intervalu hodnot, tak jak jsou uvedeny ve výčtu výše, je nutno mít na paměti, že pro správnou funkci genetického algoritmu využívajícího tyto metody, nejsou mezní hodnoty těchto intervalů, s výjimkou $SP = 2,0$, použitelné.

Turnajová metoda pracuje výhradně s diskrétními hodnotami (přirozená čísla), zatímco lineární pořadová selekce plně využívá menšího intervalu hodnot selekčního tlaku.

4.7.2 Ztráta diverzity a intenzita selekce

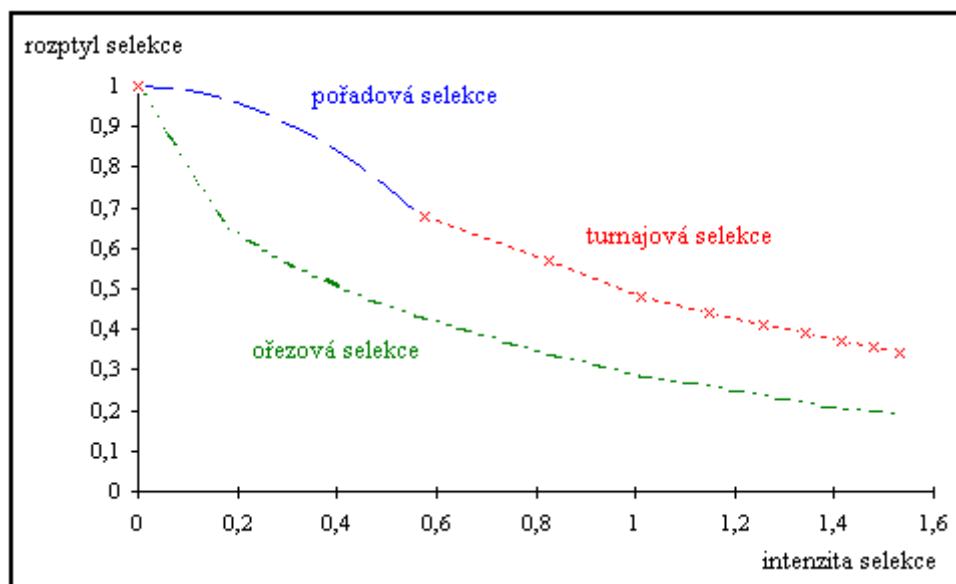


Obr. 14 Závislost ztráty diverzity na intenzitě selekce

Zdroj: upraveno z [16]

Z grafu je patrné, že se všechny tři výběrové metody chovají podobně. Jak již bylo zmíněno v předcházející kapitole, lineární pořadová metoda pracuje v intervalu hodnot nepřístupném metodě turnajové díky jejímu diskrétnímu charakteru. Ořezová metoda má oproti pořadové a turnajové vyšší ztrátu diverzity při stejné intenzitě selekce. To je zapříčiněno způsobem její práce, kdy jsou všichni jedinci s tak nízkým ohodnocením, že se nevešly do prahu ořezu, nahrazeni potomky kvalitnějších řešení.

4.7.3 Rozptyl selekce a její intenzita



Obr. 15 Závislost rozptylu selekce na intenzitě selekce

Zdroj: upraveno z [16]

Jak je vidět na obrázku, ořezová metoda má při stejné intenzitě znatelně menší rozptyl než pořadová nebo turnajová metoda výběru. (Logicky tedy měla pro změnu největší ztrátu diverzity. Viz Obr. 14 v předchozí kapitole.)

Lineární pořadová metoda se chová podobně jako turnajová, přičemž stále platí, že pracuje v oblasti, kde turnajová nefunguje kvůli svému diskrétnímu charakteru.

Distribuce ohodnocení pořadové a turnajové selekce je pro $SP = 2$ a $Tour = 2$ ($Sellnt = \frac{1}{\sqrt{\pi}}$) shodná. [2]

4.8 Závěr

Cílem této části bylo provést přehledné srovnání selekčních metod – potažmo jejich vlastností. Pro zájemce o další srovnávací práce doporučuji např. [2] a [7].

5 Praktická část

Cílem této části je provést experimentální ověření znalostí a poznatků shrnutých v předchozích kapitolách této práce, tj. vliv výběrových metod na genetický algoritmus a jejich funkčnost vůbec.

Za tímto účelem jsem implementoval jednoduchý generační genetický algoritmus (GGA) s možností volby aktuálně užívané selekční metody pro výběr jedinců z populace potenciálních řešení.

5.1 Testovací úloha

Pro ověření výsledků teoretické části jsem použil vzorový příklad evoluce řetězců. Zvolil jsem ho pro svou názornost (porovnávání řetězců) a zároveň skutečnost, že jde snad o nejčastěji používaný způsob kódování potenciálních řešení v genetických algoritmech v praxi.

Aby bylo možno považovat výsledky za prokazatelné, byly provedeny testy pro zkoumané metody se shodnými parametry pokusu:

- Typ alel = binární (0/1)
- Délka chromozomu = 50
- Počet jedinců v populaci = 50
- Počet opakování * = 50
- Požadované ohodnocení jedince = 50
- Maximální počet generací = 1000
- Pravděpodobnost křížení = 0,85
- Pravděpodobnost mutace = 0,025

*) implementováno s hromadným výpisem do jednoho souboru formátu CSV.

Hodnoty těchto parametrů byly zvoleny tak, aby se mohla dostatečně projevit činnost genetického algoritmu (dostatečně velké). Požadované ohodnocení jedince je shodné s délkou jeho chromozomu – chceme tedy, aby se shodovaly všechny geny či alely jedince s cílovým chromozomem. Ohodnocující funkce je velmi jednoduchá – kolik alel je shodných, takové je hodnocení. Protože předmětem této práce není zkoumání vlivu evolučních operátorů křížení a mutace, byla jim přiřazena, pro všechny testy, hodnota míry pravděpodobnosti odpovídající středu rozmezí hodnot, jichž se nejčastěji používá dle [17].

5.2 Implementace úlohy – obsluha programu

Program je vytvořen v jazyce Java – ve formě spustitelného souboru JAR. Pro svůj běh tedy vyžaduje, aby měl příslušný počítač nainstalovány podporu tohoto programovacího jazyka. (Použito bylo vývojové prostředí NetBeans.)

Ovládání programu je velmi jednoduché a intuitivní, neboť celé grafické uživatelské rozhraní je tvořeno jediným oknem (viz Obr. 16 dále). Uživatel postupuje podle předepsaných pokynů a vyplní tak postupně všechny potřebné vstupní parametry pro běh programu.

Chce-li uživatel změnit cílené optimální řešení, stačí jej nahradit v textovém souboru „*cilovyChromozom.txt*“ v tomtéž adresáři jako se nachází vlastní program. Pro běh programu je struktura cílového chromozomu – binárního řetězce nepodstatná, pracuje totiž správně pro jakoukoliv kombinaci alel rovných 0, nebo 1. Odtud vyplývá, že jde o změnu čistě estetickou – lze sledovat, jak se jedinci přibližují (začínají se podobat) zadanému řešení. Nutné je dbát pouze na to, aby byl cílový chromozom minimálně stejně dlouhý nebo delší, než je délka chromozomu a požadované ohodnocení, jež nastavujeme jedincům v parametrech pro spuštění programu.

Uživatel zadá do připravených textových polí hodnotu těchto parametrů:

- Délka chromozomu
- Počet jedinců v populaci
- Počet populací paralelně
- Požadované ohodnocení jedince
- Maximální počet generací
- Pravděpodobnost křížení
- Pravděpodobnost mutace

Dále provede volbu selekční metody, a je-li to třeba, zadá hodnotu jejího parametru. Všechny tyto metody jsou popsány v předchozí teoreticky zaměřené části práce, včetně běžně používaných parametrů pro jejich vzájemné srovnání.

Jedná se o metody užívané v genetických algoritmech generačního typu:

- Proporcionální metody
 - Ruletová selekce
 - Univerzální Stochastické Vzorkování
- Pořadová selekce
 - Lineární selekce
- Turnajová selekce
- Ořezová selekce

V poslední volbě nastaví formát výstupu výsledků – zaškrtně, jaké hodnoty mají být zachyceny ve výsledném souboru:

- Číslo generace (první číslo na řádku, pro celou populaci shodné)
- Chromozom jedince
- Ohodnocení jedince

Ke spuštění běhu programu slouží jediné tlačítko „START“, pod nímž je umístěn ukazatel stavu programu. Kromě oznámení konce běhu programu, upozorní také uživatele, v jakém textovém poli jsou zadány nevyhovující vstupní hodnoty.

Obr. 16 Grafické uživatelské rozhraní programu

Zdroj: vlastní zpracování

Počáteční populace je vždy náhodně generována při spuštění běhu aplikace. Výsledky jsou exportovány v podobě CSV souboru do adresáře programu. Název vygenerovaného souboru obsahuje všechna aktuální nastavení programu, kromě voleb formátu tisku, v takovém pořadí, jak jsou uvedena a zadávána do okna programu. Například soubor obsahující výsledky běhu programu, nastaveném jako na Obr. 16, budou vygenerovány a uloženy do souboru s názvem:

50,50,50,50,1000,0.85,0.025,RWS.csv

V případě, že by uživatel zvolil při zachování stejných hodnot parametrů lineární pořadovou metodu, bude výsledný soubor vypadat takto:

50,50,50,50,1000,0.85,0.025,Linear,1.5.csv

Kromě jiného názvu metody, přibyla na konec hodnota jejího parametru.

5.3 Výsledky testování

Výstupy z aplikace ve formě souborů CSV byly dále zpracovány v tabulkovém procesoru Microsoft Excel 2010 a 2016. Všechny tyto soubory lze nalézt v příloze.

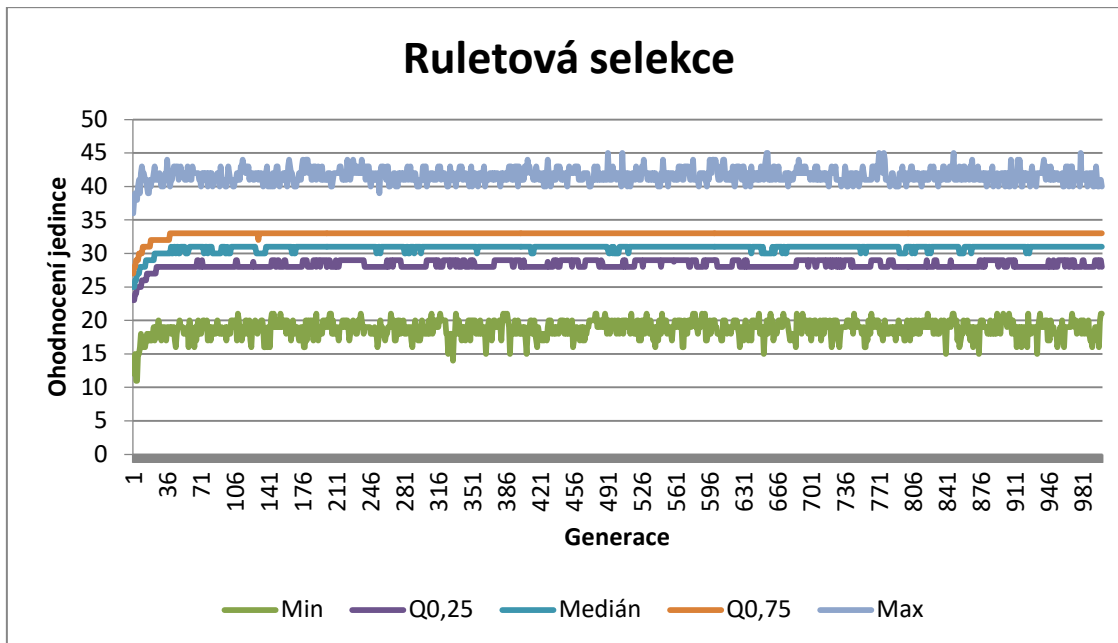
5.3.1 Proporcionální metody

Ani jedna z proporcionálních selekčních metod nedokázala nalézt námi požadované řešení v přiděleném kvantu 1000 generací pro 50 samostatných populací.

Jako možný způsob zlepšení jejich výkonu lze uvažovat implementaci jiné ohodnocující funkce, jež by více zohlednila rozdíly v kvalitě jedinců - tzn. na příklad funkce na exponenciálním základu.

5.3.1.1 Ruletová selekce

V počáteční fázi je vidět celkový vzestup ohodnocení jedinců, který se však záhy zastavil a už neopustil tuto hladinu.



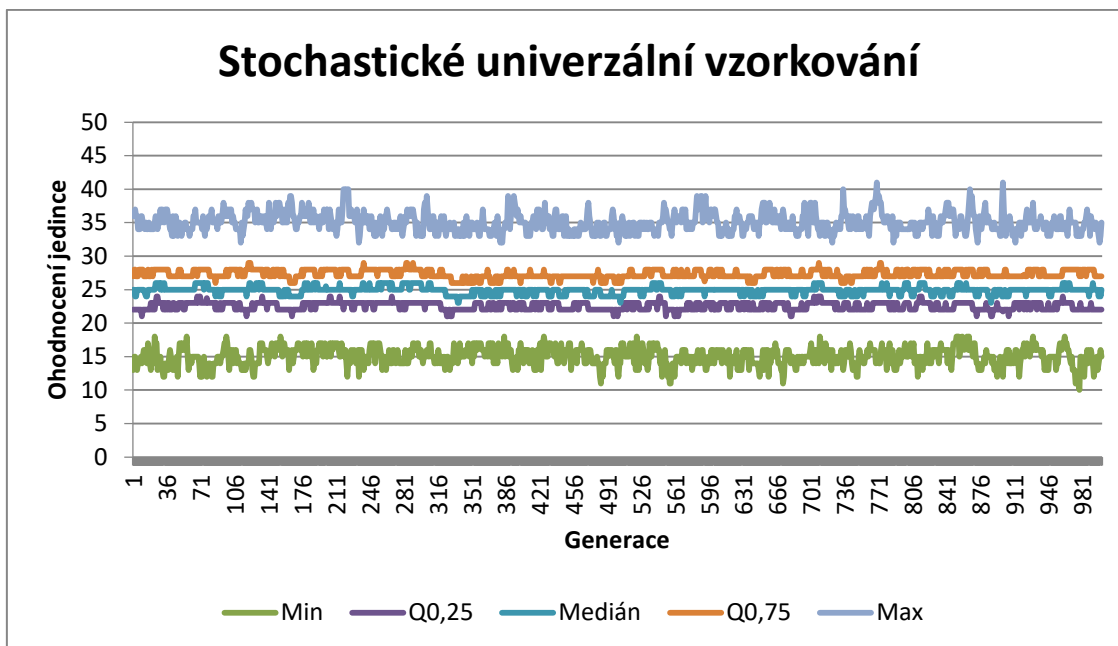
Obr. 17 Ruletová selekce

Zdroj: vlastní zpracování

Nejlepší jedinec má ohodnocení rovno 45.

5.3.1.2 Stochastické univerzální vzorkování

Jak lze nahlédnout (viz Obr. 18), v případě této metody byly výsledky horší, než u metody ruletové selekce, z níž je tato odvozena. U této úlohy se její mechanismus výběru rodičovských jedinců jeví kontraproduktivním, neboť udržuje až přílišnou diverzitu populace a brání tak v evoluci silnějších jedinců.



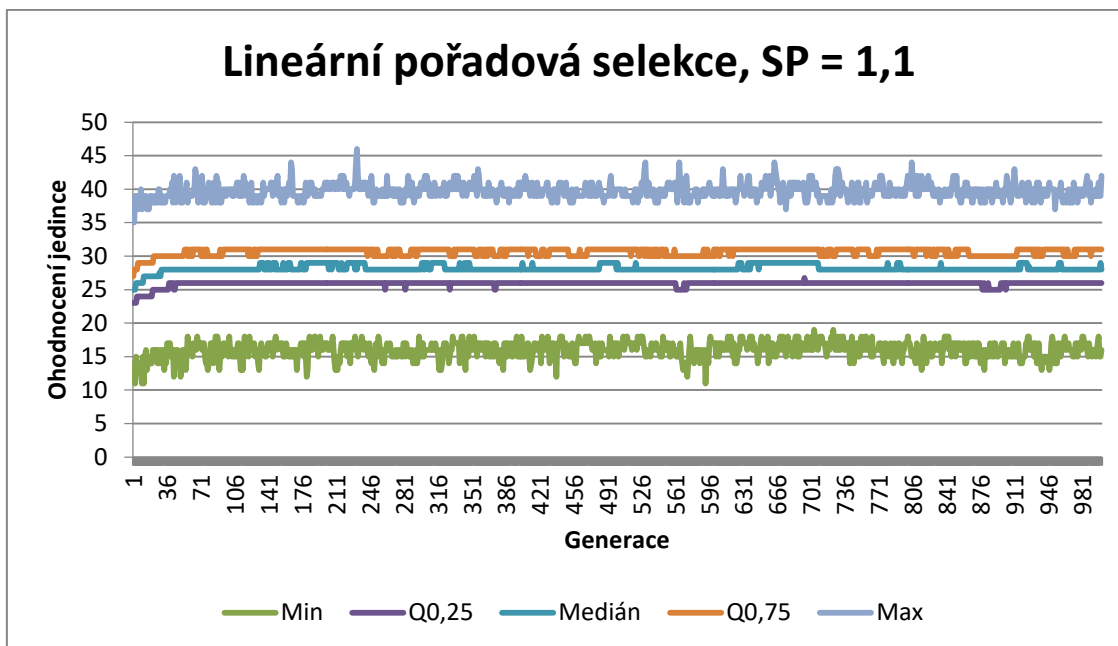
Obr. 18 Stochastické univerzální vzorkování
Zdroj: vlastní zpracování

Nejvyšší dosažené ohodnocení jedince činí 41.

5.3.2 Lineární pořadová selekce

Pro tuto a následující metody s parametrem, byly zahrnuty ukázky pro více možných hodnot nastavení těchto parametrů, abychom získali názorný přehled jejich vlivu na funkci dané metody a dále potažmo na celý genetický algoritmus.

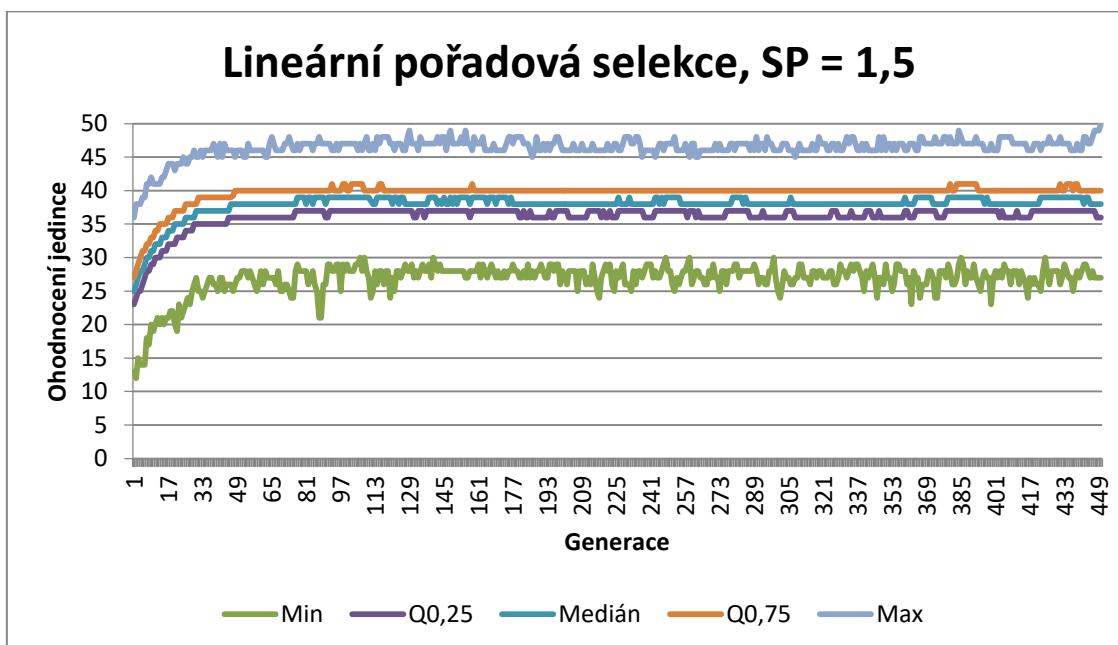
Parametr *SP* (selekční tlak) může nabývat hodnot z intervalu $\langle 1,0; 2,0 \rangle$, ale jeho spodní hranice se nepoužívá, neboť mají potom všichni jedinci v populaci ohodnocení stejné – a to právě 1. Vlastní výběr jedinců je pak otázkou prostého náhodného výběru.



Obr. 19 Lineární pořadová selekce, SP = 1,1
Zdroj: vlastní zpracování

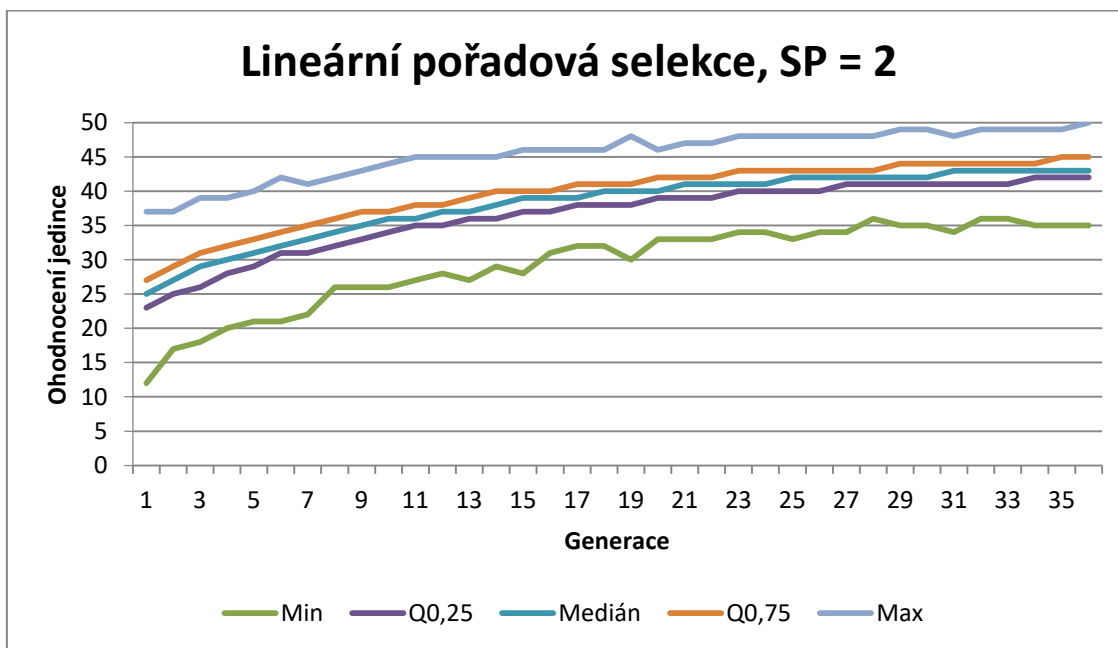
V grafu metody pro $SP = 1,1$ lze vidět menší vzestup hladiny ohodnocení jedinců, nikoli však dostatečně velký pro potřeby naší testovací úlohy.

Dosaženo bylo nejlepšího výsledku ohodnocení 46.



Obr. 20 Lineární pořadová selekce, SP = 1,5
Zdroj: vlastní zpracování

Při použití parametru $SP = 1,5$ vidíme výrazné zlepšení, přičemž se podařilo nalézt řešení o požadované kvalitě (ohodnocení = 50) v generaci číslo 499.

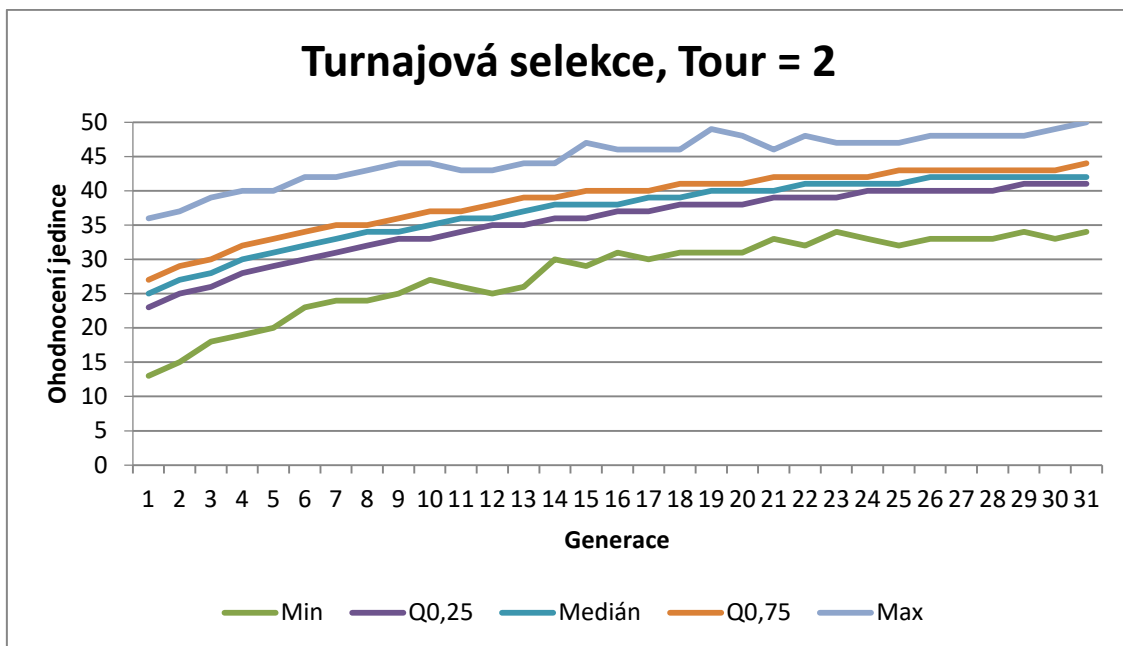


Obr. 21 Lineární pořadová selekce, $SP = 2$
Zdroj: vlastní zpracování

Pro $SP = 2$ jsou výsledky zatím bezkonkurenčně nejlepší. Požadované řešení (ohodnocení = 50) bylo nalezeno už v generaci číslo 39.

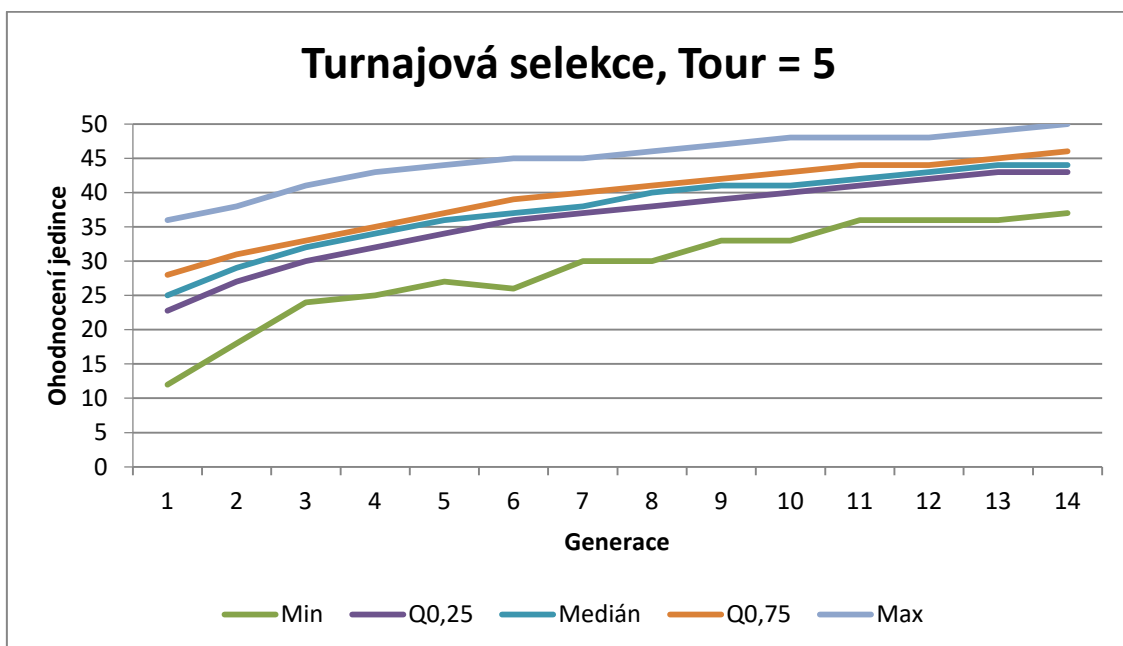
5.3.3 Turnajová selekce

Parametr *Tour* této metody může nabývat hodnot z intervalu $\langle 1; N \rangle$, kde N je počet jedinců v populaci. Jednoduše lze odvodit, že ani jedna z těchto mezních hodnot není žádoucí. Pro $Tour = 1$ se jedná o náhodný výběr a pro $Tour = N$ bude vždy vybrán pouze ten nejlépe hodnocený jedinec v celé populaci.



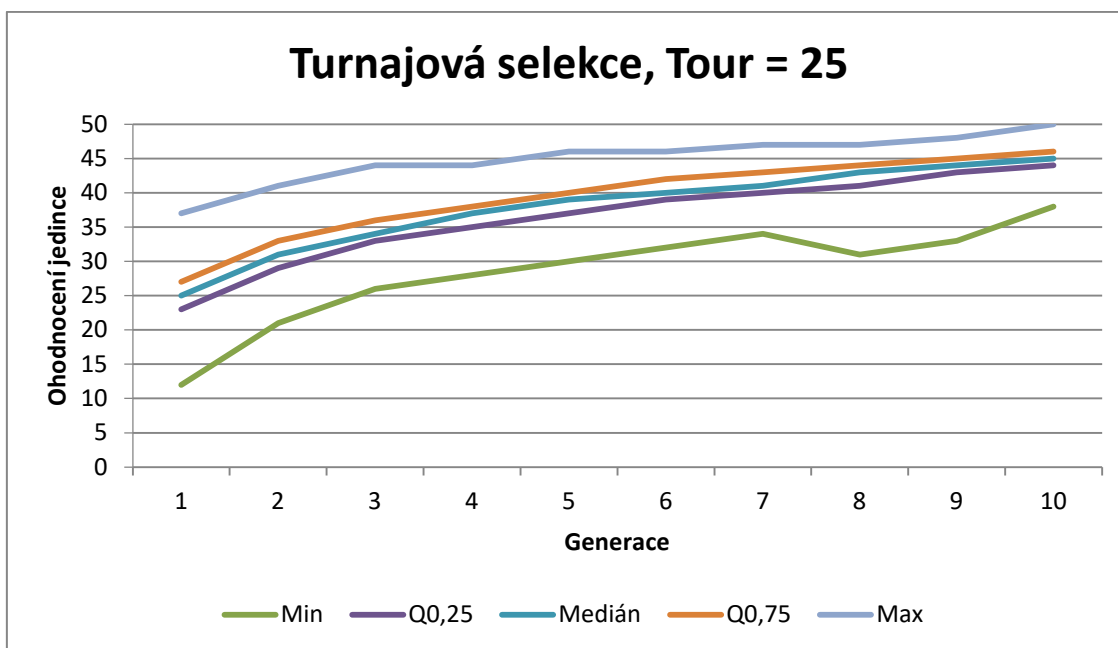
Obr. 22 Turnajová selekce, Tour = 2
Zdroj: vlastní zpracování

Pro *Tour* = 2 bylo nalezeno řešení (ohodnocení = 50) už v generaci číslo 30.



Obr. 23 Turnajová selekce, Tour = 5
Zdroj: vlastní zpracování

Pro *Tour* = 5 bylo nalezeno řešení (ohodnocení = 50) v generaci číslo 13.



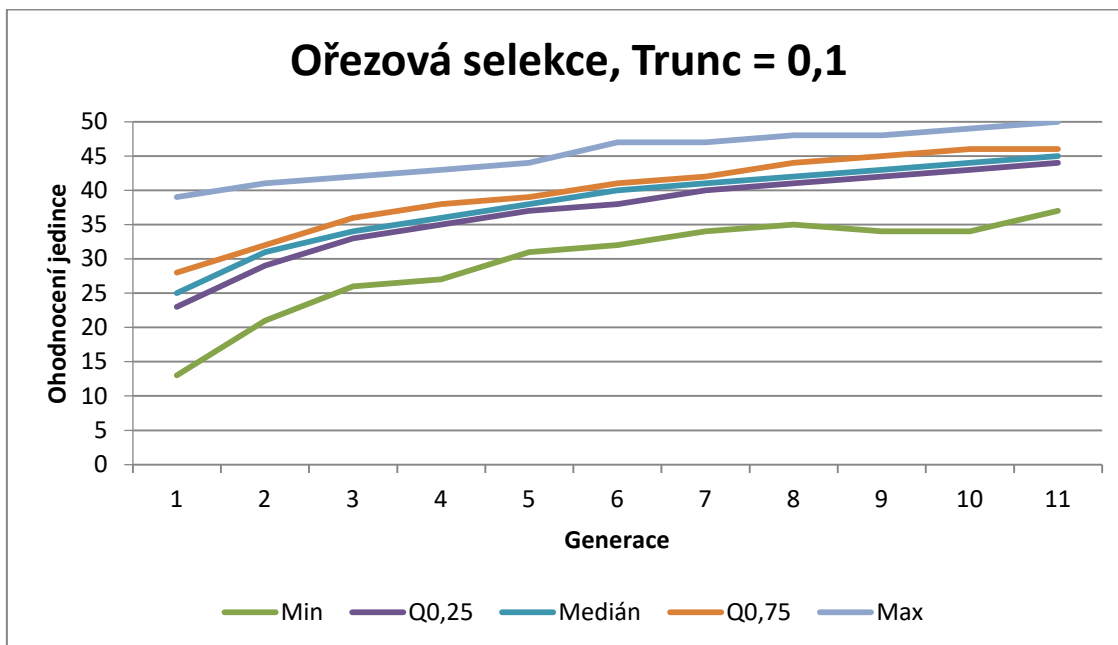
Obr. 24 Turnajová selekce, Tour = 25
Zdroj: vlastní zpracování

A pro $Tour = 25$ (tedy polovině velikosti populace) bylo nalezeno řešení (ohodnocení = 50) už v generaci číslo 9.

Zde už není s přibývajícím hodnotou parametru tak výrazné zlepšení výsledků, přestože roste intenzita selekce.

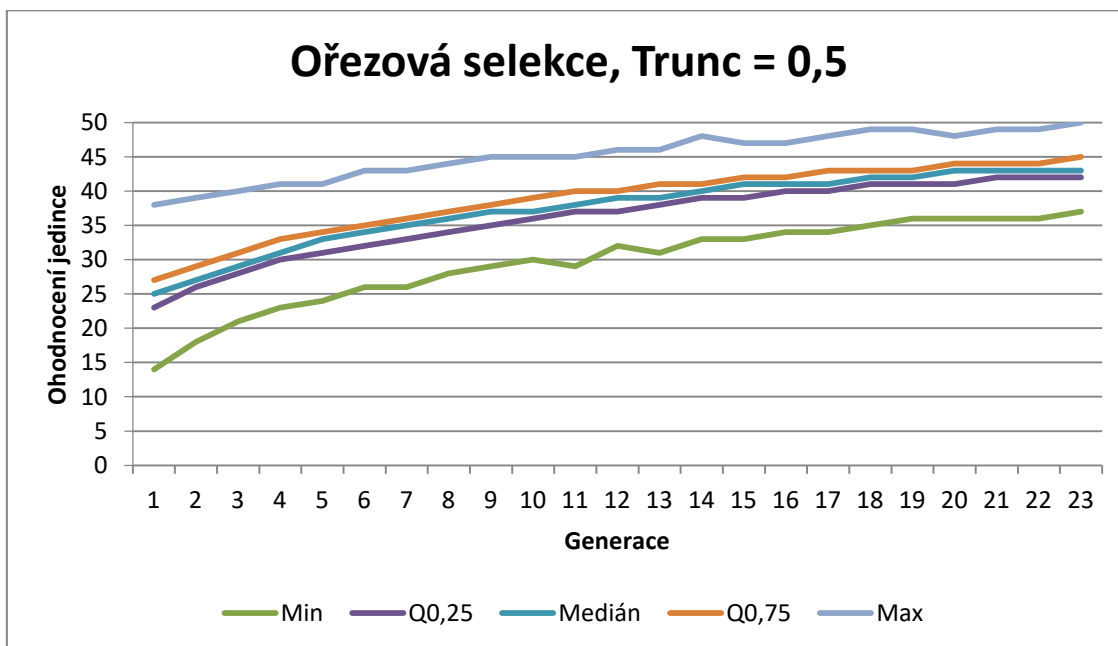
5.3.4 Ořezová selekce

U této metody naopak intenzita selekce klesá, se vzrůstající hodnotou parametru *Trunc*. Často se používají hodnoty od 0,1 po 0,5.



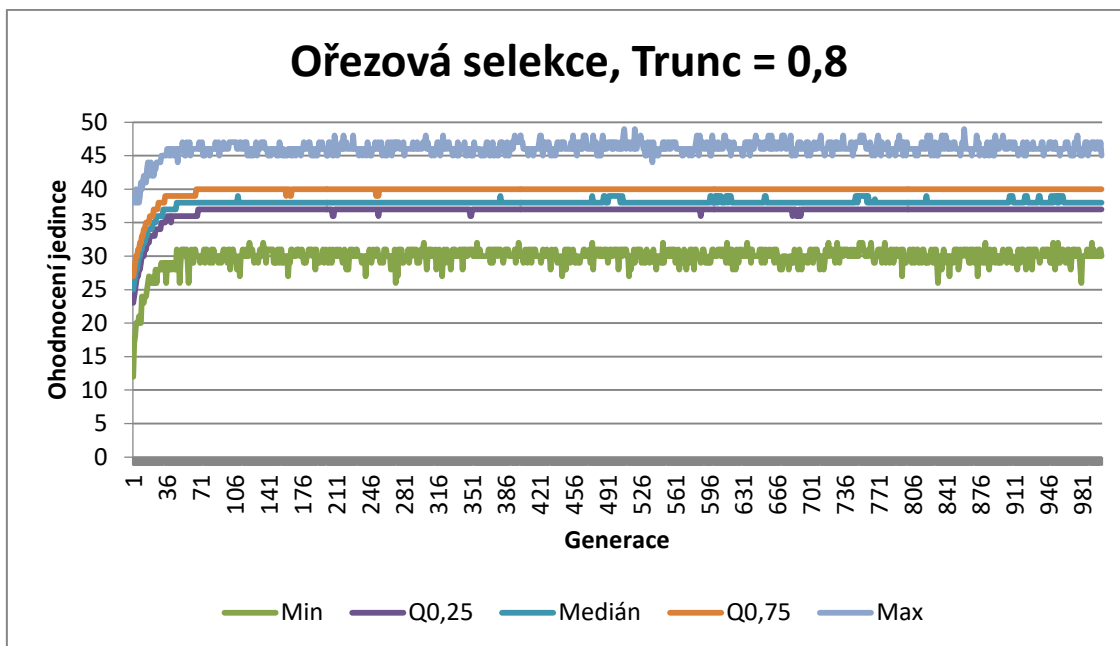
Obr. 25 Ořezová selekce, Trunc = 0,1
Zdroj: vlastní zpracování

Pro *Trunc* = 0,1 bylo řešení (ohodnocení = 50) nalezeno už v generaci číslo 10.



Obr. 26 Ořezová selekce, Trunc = 0,5
Zdroj: vlastní zpracování

Pro *Trunc* = 0,5 bylo řešení (ohodnocení = 50) nalezeno až v generaci číslo 22.



Obr. 27 Ořezová selekce, Trunc = 0,8
Zdroj: vlastní zpracování

A pro $Trunc = 0,8$ nebylo řešení nalezeno vůbec, i když těsně. Nejlepší jedinec měl ohodnocení rovno 49.

5.4 Shrnutí výsledků testování

V testu je zahrnuto pět nejrozšířenějších selekčních metod generačního typu (GGA). Pokud má daná metoda vlastní parametr, pak byly zvoleny tři hodnoty z jeho intervalu tak, aby byl názorně předveden jeho vliv na chování metody, jak byl popsán v teoretické části této práce.

V tabulce níže jsou uvedeny všechny získané výsledky testované úlohy.

Tabulka 6 Souhrn výsledků testování

Metoda selekce	Parametr	Počet vytvořených generací	Nejlépe ohodnocený jedinec
Ruletová		1000	45
Stochastické univerzální vzorkování		1000	41
Lineární pořadová	SP = 1,1	1000	46
	SP = 1,5	449	50
	SP = 2,0	35	50
Turnajová	Tour = 2	30	50
	Tour = 5	13	50
	Tour = 25	9	50
Ořezová	Trunc = 0,1	10	50
	Trunc = 0,5	22	50
	Trunc = 0,8	1000	49

Zdroj: vlastní zpracování

Nejllepších výsledků dosáhla metoda turnajová, následována ořezovou a lineární metodou. Toto pořadí ovšem nelze brát jakožto absolutní, vzhledem ke stochastickému základu těchto metod – tzn. při opakovaném pokusu, se mohou výsledky lišit.

Jak už bylo zmíněno u jednotlivých výsledků v předchozí části této práce, proporcionální metody se v této implementaci neosvědčily. Zřejmě by mohl pomoci jiný způsob ohodnocování jedinců.

Ze získaných výsledků je zřetelné, jak velký je vliv parametrů jednotlivých metod a i proto, nelze obecně prohlásit jednu metodu za lepší, či horší, než druhou. Vždy bude výrazně záležet na konkrétní implementaci a parametrech algoritmu metody.

Za zmínění stojí i tzv. „No Free Lunch (NFL)“ teorém, jenž říká [19, s. 67]: „Pro jakýkoliv algoritmus platí, že je jeho vysoká účinnost v jedné skupině úloh, vykoupena horším výkonem v jiné skupině.“

6 Závěry a doporučení

Předmětem zkoumání této bakalářské práce jsou metody selekce jedinců používané v genetických algoritmech – konkrétněji jejich vlastnosti a vliv na tvorbu nové populace potenciálních řešení.

První částí práce je stručný úvod do genetických algoritmů – od historie jejich vzniku, až po popis jejich funkce.

Druhá teoretická část se zaměřuje už výhradně na selekční metody a obsahuje jejich podrobnější popis, spolu s jejich vzájemným porovnáním pomocí objektivních parametrů užívaných právě pro popis výběru jedinců, přičemž vysvětluje, jakým způsobem vybírají metody jedince do nové populace, čímž současně ovlivňují její skladbu (diverzitu).

Třetí část se zabývá praktickým ověřením dříve popsanych teoretických poznatků. K tomuto účelu slouží testovací aplikace implementující zkoumané selekční metody na zkušební úloze. Cílem byla především názornost, nikoliv výkon aplikace. Jeho zvýšení by měla pomoci například úprava exportu výsledků nebo paralelní běh programu.

Zde lze konstatovat, že se při praktických pokusech nevyskytly výsledky, jenž by odporovaly dříve předložené teoretické části této práce, a je možné prohlásit, že při správných podmínkách – tj. nastavení parametrů a typu úlohy nám mohou genetické algoritmy poskytnout velmi dobré výsledky.

Je nutno podotknout, že mnou provedené testování bylo v mnoha ohledech svým záběrem omezené, vzhledem k variabilitě možností nastavení parametrů genetického algoritmu či samotné testovací úlohy, a nelze tedy rezolutně prohlásit jednu z výběrových metod za univerzálně nejlepší.

Dalším stádiem při zkoumání uvedené problematiky by bylo právě provedení testů na různorodých zkušebních úlohách (nikoli pouze evoluce binárních řetězců) s odzkoušením odlišných hodnot parametrů selektivních metod i genetického algoritmu samého.

Ale i v tomto případě by výsledky nebylo možno považovat za všeobecně reprezentativní, protože je každá implementace na novou úlohu individuální a pro co nejlepší výsledky požaduje důkladné přizpůsobení všech parametrů i vlastní implementace genetického algoritmu.

7 Seznam použité literatury

- [1] BAKER, James E. Reducing bias and inefficiency in the selection algorithm. In: GREFENSTETTE, John J. (editor). *Genetic algorithms and their applications: proceedings of the second International Conference on Genetic Algorithms : July 28-31, 1987 at the Massachusetts Institute of Technology, Cambridge, MA*. 1. print. Hillsdale, N.J: L. Erlbaum Associates, 1987, pages 14-21. ISBN 0805801588.
- [2] BLICKLE, Tobias and Lothar THIELE. *A Comparison of Selection Schemes Used in Genetic Algorithms*. 2. Edition. TIK-Report No.11 Zurich: Computer Engineering and Networks Lab (TIK), Swiss Federal Institute of Technology (ETH), January 1995 [cit. 2017-01-20]. Dostupné z: http://www.tik.ee.ethz.ch/db/public/tik/?db=publications&form=report_single_publication&publication_id=492
- [3] CANO, José R., Francisco HERRERA and Manuel LOZANO. Instance Selection Using Evolutionary Algorithms: An Experimental Study. PAL, Nikhil R. and Lakhmi JAIN. *Advanced Techniques in Knowledge Discovery and Data Mining*. New York: Springer-Verlag, 2004, pages 127-152. ISBN 1852338679.
- [4] CROW, James F and Motoo KIMURA. *An Introduction to Population Genetics Theory*. Caldwell, NJ: Blackburn Press, 2005, 608 p. ISBN 19-328-4612-3.
- [5] Evolutionary algorithm. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2017-01-20]. Dostupné z: http://en.wikipedia.org/wiki/Evolutionary_algorithm
- [6] FOGEL, David B. *Evolutionary Computation: The Fossil Record*. New York: IEEE Press, c1998, xiii, 641 p. ISBN 0780334817.
- [7] GOLDBERG, David E. and Kalyanmoy DEB. A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. In: RAWLINS, Gregory J. E. (Ed.). *Foundations of genetic algorithms*. 1st ed. San Mateo, California: M. Kaufmann Publishers, 1991, pages 69-93. ISBN 1558601708.
- [8] HAUPT, Randy L a Sue Ellen HAUPT. *Practical Genetics Algorithms*. New York: John Wiley, 1998, 177 p. ISBN 04-711-8873-5.
- [9] HOLLAND, John H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. 1st MIT Press ed. Cambridge, Mass.: MIT Press, 1992, xiv, 211 p. ISBN 9780262581110.
- [10] HYNEK, Josef. *Genetické algoritmy a genetické programování*. 1. vyd. Praha: Grada, 2008, 182 s. ISBN 978-80-247-2695-3.

- [11] LOBO, Fernando G. and David E. GOLDBERG. The parameter-less genetic algorithm in practice. *Information Sciences*. 2004, volume 167, issues 1-4, pages 217-232. DOI: 10.1016/j.ins.2003.03.029 [cit. 2017-01-20]. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0020025504000441>
- [12] LOBO, Fernando G. and Cláudio F. LIMA. Adaptive Population Sizing Schemes in Genetic Algorithms. In: *Parameter Settings in Evolutionary Algorithms*. Fernando G. Lobo, Cláudio F. Lima and Zbigniew Michalewicz (Eds.). Berlin: Springer, c2007, pages 185-204. Studies in Computational Intelligence, Vol. 54. ISBN 978-3-540-69431-1.
- [13] MARCZYK, Adam. Genetic Algorithms and Evolutionary Computation. *TalkOrigins Archive* [online]. 2004 [cit. 2017-01-20]. Dostupné z: <http://www.talkorigins.org/faqs/genalg/genalg.html>
- [14] MITCHELL, Melanie. *Introduction to genetic algorithms*. Vyd. 1. Massachusetts: MIT Press, 1997, 209 p. ISBN 02-621-3316-4.
- [15] On the Origin of Species by Means of Natural Selection (1859). DARWIN, Charles. *Evolutionary Writings: Including the Autobiographies*. 1st ed. James A. Secord. New York: Oxford University Press, 2008, pages 107-211. ISBN 978-0-19-920863-0.
- [16] POHLHEIM, Hartmut. *GEATbx: Introduction: Evolutionary Algorithms: Overview, Methods and Operators* [online]. 3.8. 2006, 95 s. [cit. 2017-01-20]. Dostupné z: <http://www.geatbx.com/>
- [17] SCHAFFER, J. D., R. A. CARUANA, L. J. ESHELMAN and R. DAS. A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization. In: *Schaffer, J. D. (ed.): Proceedings of the Third International Conference on Genetic Algorithms*. San Mateo, California: Morgan Kaufmann, 1989, pages 51-60. ISBN 1558600663.
- [18] VOSE, Michael D. *The Simple Genetic Algorithm: Foundations and Theory*. Cambridge, Mass.: MIT Press, c1999, ix, 251 p. ISBN 02-622-2058-X.
- [19] WOLPERT, David H. and William G. MACREADY. No Free Lunch Theorems for Optimization. In: *IEEE Transactions on Evolutionary Computation*. New York: IEEE Press, c1997, pages 67-82. Vol. 1. No. 1.
- [20] ZELINKA, Ivan, Zuzana OPLATKOVÁ, Miloš ŠEDA, Pavel OŠMERA a František VČELAŘ. *Evoluční výpočetní techniky: Principy a aplikace*. 1. české vyd. Praha: BEN, 2009, 534 s. ISBN 978-80-7300-218-3.

8 Přílohy

Na přiloženém nosiči CD jsou uloženy následující přílohy:

- 1) Elektronická verze této práce (pdf)
- 2) Testovací aplikace vzorového příkladu (jar)
- 3) Zdrojové kódy aplikace (java, form)
- 4) Zpracované výstupy testů (xlsx)



UNIVERZITA HRADEC KRÁLOVÉ
Fakulta informatiky a managementu
Rokitanského 62, 500 03 Hradec Králové, tel: 493 331 111, fax: 493 332 235

Zadání k závěrečné práci

Jméno a příjmení studenta:

Ondřej Tázlar

Obor studia:

Aplikovaná informatika

Jméno a příjmení vedoucího práce:

Josef Hynek

Název práce:

Vliv selekčních metod na tvorbu nové populace a chování genetického algoritmu

Název práce v AJ:

Effects of selection methods on creating new population and the behaviour of the genetic algorithm

Podtitul práce:

Podtitul práce v AJ:

Cíl práce: Cílem práce je vzájemné porovnání vlivu vybraných selekčních metod na tvorbu nové populace potenciálních řešení a na chování jednoduchého genetického algoritmu.

Osnova práce:

Úvod

Cíl práce

Úvod do Genetických algoritmů

Podrobný popis vybraných selekčních metod

Vlivy selekčních metod a metodika jejich zkoumání

Jednotlivé pokusy a jejich výsledky

Shrnutí výsledků výzkumu

Závěr

Seznam použité literatury

Přílohy

Projednáno dne: 25.10.2013

Podpis studenta

Podpis vedoucího práce