



MATEMATICKO-FYZIKÁLNÍ
FAKULTA
Univerzita Karlova

Přírodou inspirované prohledávací algoritmy

Roman Neruda, Martin Pilát

4. září 2017

Před úvodem

Jak to je?

OSNOVA PREDNASKY TAK JAK JE TED

- Evoluční algoritmy
- Jednoduchý genetický algoritmus
- Teorie schémat
- Reprezentace a operátory v GA
- Evoluce kooperace
- Evoluční strategie
- Diferenciální evoluce
- Particle swarm optimization
- Evoluční strojové učení
- Vícekriteriální optimalizace
- Evoluční kombinatorická optimalizace
- Ladění, řízení, metaevoluce
- Teorie EVA podruhé
- Evoluční programování
- Genetické programování
- Neuroevoluce
- Memetické algoritmy
- Dynamické krajiny fitness
- Tabu search, scatter search
- Biologicky věrnější evoluce

Jak to bude?

Pok pok pokusy

SKORO TO TU MUZEME SMAZAT NE?

procedure EVOLUČNÍ ALGORITMUS $t \leftarrow 0$ *Inicializuj* populaci P_t náhodně vygenerovanými jedinci*Ohodnoť* jedince v populaci P_t **while** neplatí *kritérium ukončení* **do**vyber z P_t rodiče P'_{t+1} *Rodičovskou selekcí**Rekombinací* rodičů vzniknou potomci P'_{t+1} *Mutuj* potomky P'_{t+1} *Ohodnoť* potomky P'_{t+1} *Enviromentální selekcí* vyber P_{t+1} z P_t a P'_{t+1} $t \leftarrow t + 1$ **end while****end procedure**

Algoritmus 1: Schéma evolučního algoritmu

Evoluční algoritmy

Evoluce, geny a DNA

Obecné schéma evolučního algoritmu

Oblast evolučních výpočtů či algoritmů (v angličtině evolutionary computing) zastřešuje několik proudů, které se zpočátku vyvíjely samostatně. Za prehistorii této disciplíny lze považovat Turingovy návrhy na využití evolučního prohledávání z roku 1948 ¹ a Bremermannovy první pokusy o implementaci optimalizace pomocí evoluce a rekombinace z roku 1962 ². Během šedesátých let se objevily tři skupiny výzkumníků, kteří nezávisle na sobě vyvíjely a navrhly své varianty použití evolučních principů v informatice. Holland publikoval v roce 1975 svůj návrh genetických algoritmů ³, zatímco skupina Fogela a spolupracovníků vyvinula metodu nazvanou evoluční programování ⁴. Nezávisle na nich přišli Rechenberg a Schwefel v Německu na metodu nazvanou evoluční strategie ⁵. Až do přelomu osmdesátých a devadesátých let existovaly tyto směry bez výraznější interakce, ale poté se spojily do obecnější oblasti evolučních algoritmů. V té době Koza vytváří metodu genetického programování, Dorigo publikuje disertaci s návrhem mravenčích optimalizačních algoritmů a vznikají první pokusy o aplikaci evoluce na vývoj umělých neuronových sítí.

U zrodu různých variant evolučních algoritmů stála inspirace přírodními jevy, konkrétně jde o Darwinovu teorii přírodního výběru a zjednodušené principy genetiky, které poprvé načrtl Mendel. Z genetiky se evoluční algoritmy inspirovaly diskretní reprezentací genotypu, z biologické evoluční teorie používají Darwinovu myšlenku o výběru jedinců v prostředí s omezenými zdroji, který závisí na míře přizpůsobení se jedinců danému prostředí.

Základní obecnou myšlenku evolučních algoritmů lze vyjádřit následujícím způsobem. Mějme populaci jedinců v prostředí, které určuje jejich úspěšnost — fitness. Tito jedinci navzájem soupeří o možnost reprodukce a přežití, která závisí právě na hodnotě fitness. Jde tedy o množinu kandidátů na řešení problému definovaného prostředím. Způsoby reprezentace jedinců, jejich výběru a rekombinace závisí na konkrétním dialektu evolučních algoritmů, které probereme vzápětí.

Základní princip fungování evolučních algoritmů je tedy následující ⁶. Na začátku algoritmu vygenerujeme (nejčastěji náhodně) první iniciální populaci jedinců. Všechny jedince v populaci ohodnotíme ohodnocovací funkcí. Hodnota této funkce určuje šanci výběru jedinců během rodičovské selekce. Vybraní jedinci jsou potom rekombinováni pomocí rekombinačního operátoru, který typicky ze dvou jedinců vytváří jednoho či dva potomky, a

¹

Missing ref.

²

Missing ref.

³ John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. MIT Press, Cambridge, MA, USA, 1992

⁴ David B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, Piscataway, NJ, USA, 1995

⁵ Hans-Georg Beyer and Hans-Paul Schwefel. Evolution strategies – a comprehensive introduction. 1(1):3–52, May 2002

⁶ Agoston E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing*. Springer-Verlag, 2003

procedure EVOLUČNÍ ALGORITMUS

 $t \leftarrow 0$ *Inicializuj* populaci P_t náhodně vygenerovanými jedinci*Ohodnoť* jedince v populaci P_t **while** neplatí *kritérium ukončení* **do**vyber z P_t rodiče *Rodičovskou selekcí**Rekombinací* rodičů vzniknou potomci*Mutuj* potomky*Ohodnoť* potomky*Enviromentální selekcí* vyber P_{t+1} z P_t a potomků $t \leftarrow t + 1$ **end while****end procedure**

pomoci operátoru mutace, který typicky provádí drobné změny jednoho jedince. Tímto postupem si vytvoříme množinu nových kandidátů řešení, a tito noví jedinci potom soutěží s původními jedinci o místo v nové populaci. Výběr jedinců do nové populace (tedy jakési slití rodičů a potomků) má na starosti enviromentální selekce beroucí v úvahu fitness jedinců a případně další ukazatele jako je například stáří jedinců. Tím je vytvořena nová generace a tento cyklus pokračuje do splnění určitého kritéria ukončení, což je nejčastěji dostatečně dobrý nejlepší jedinec nebo předem určený počet generací.

Genetické algoritmy

Genetické algoritmy

Genetické algoritmy jsou asi nejznámější součástí evolučních výpočtů a v různých obměnách se používají hlavně při řešení optimalizačních úloh. Je zajímavé, že původní Hollandovou motivací při návrhu genetického algoritmu bylo studovat vlastnosti přírodou inspirované adaptace⁷. Velká část původní literatury byla věnována popisu principů, jak genetický algoritmus pracuje při hledání řešení úlohy. Zajímavé jsou paralely s matematickým problémem dvourukého bandity, který je příkladem na udržování optimální rovnováhy mezi explorací a exploatací.

Nejjednodušší varianta genetického algoritmu pracuje s binárními jedinci, to znamená, že parametry řešené úlohy je nutno vždy zakódovat jako binární řetězce. Tento přístup je výhodný z hlediska jednoduchosti použitých operátorů, ale binární zakódování na druhou stranu nemusí být nejvhodnější reprezentací problému. Způsob fungování jednoduchého genetického algoritmu je také poměrně jednoduchý. Algoritmus přechází mezi populacemi řešení tak, že nová populace zcela nahradí předchozí. Výběr rodičů je často realizován tzv. ruletovou selekcí, která vybírá jedince náhodně s pravděpodobností výběru úměrné jejich fitness. Rekombinačním operátorem je jednobodové křížení, které náhodně zvolí stejnou pozici v rodičích a vymění jejich části. Pravděpodobnost uskutečnění operace křížení je jedním z parametrů programu a obvykle je poměrně vysoká (0,5 i více). Mutace provádí drobné lokální změny tak, že prochází jednotlivé bity řetězce a každý bit s velmi malou pravděpodobností změní. Pravděpodobnost mutace je typicky nastavena, tak aby došlo průměrně ke změně jednoho bitu v populaci (oblíbená dolní mez) nebo v jedinci (horní mez).

Ruletovou selekci si dle metafory můžeme představit tak, že kolo rulety rozdělíme na výseče odpovídající velikostí hodnotám fitness jedinců a při výběru pak n krát vhodíme kuličku. Často používaným vylepšením ruletové selekce je tzv. stochastický univerzální výběr, který hodí kuličku do rulety jen jednou a další jedince vybírá deterministicky posunem pozice kuličky o $1/n$. Tento výběr pro malá n lépe aproximuje ideální počty zastoupení jedinců v další generaci. Dalšími varianty rodičovské selekce nepracují s absolutními hodnotami fitness, ale vybírají náhodně v závislosti na pořadí jedince v populaci setříděné podle fitness, což zanedbává absolutní rozdíly mezi hodnotami. Další variantou rodičovské selekce je tzv. k -turnaj, kdy nejprve vybereme k jedinců náhodně a z nich pak vybereme nejlepšího.

V současnosti se oblast genetických algoritmů neomezuje jen na binární kódování jedinců, časté je celočíselné, permutační nebo reálné kódování,

⁷ John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. MIT Press, Cambridge, MA, USA, 1992

procedure JEDNODUCHÝ GENETICKÝ ALGORITMUS

 $t \leftarrow 0$ *Inicializuj* populaci P_t N náhodně vygenerovanými binárními jedinci délky n *Ohodnoť* jedince v populaci P_t **while** neplatí *kritérium ukončení* **do****for** $i \leftarrow 1, \dots, N/2$ **do**vyber z P_t 2 rodiče *Ruletovou selekcí*S pravděpodobností p_C *Zkříž* rodičeS pravděpodobností p_M *Mutuj* potomky*Ohodnoť* potomkyPřidej potomky do P_{t+1} **end for**Zahoď P_t $t \leftarrow t + 1$ **end while****end procedure**

Algoritmus 3: Schéma Hollandova genetikého algoritmu

která ale vyžadují specifické operace křížení a mutace⁸. O některých se zmíníme dále.

⁸ Zbigniew Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs (3rd Ed.)*. Springer-Verlag, London, UK, UK, 1996; and Melanie Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA, USA, 1996

Evoluční strategie

Spojité optimalizace

Mnoho optimalizačních problémů z běžného života se dá definovat jako optimalizace funkce

$$f : D \rightarrow \mathcal{R},$$

kde $D \subseteq \mathcal{R}^d$. Je proto přirozené, že mnoho výzkumníků se zabývá právě evolučními algoritmy, kterou jsou schopné optimalizovat takové funkce. Problém optimalizace takových funkcí se v literatuře objevuje pod pojmem *spojitá optimalizace*⁹. Je důležité si uvědomit, že název vyjadřuje pouze to, že prostor, ve kterém se hledají řešení je spojitý (\mathcal{R}^n), samotná optimalizovaná funkce f být spojitá nemusí.

⁹ anglicky *continuous optimization*

Velmi často je definiční obor funkce D d -rozměrný interval

$$D = [l_1, u_1] \times \cdots \times [l_d, u_d],$$

kde l_i a u_i jsou dolní a horní meze pro i -tou proměnnou. Objevují se ale i obecnější problémy, kde je D dána pomocí množiny podmínek tvaru $g(\vec{x}) \leq 0$ a $h(\vec{x}) = 0$, pro $g, h : \mathcal{R}^d \rightarrow \mathcal{R}$.

V této kapitole napřed budeme uvažovat optimalizační problém tvaru

$$\begin{array}{ll} \min_{\vec{x}} & f(\vec{x}) \\ \text{za podmíněk} & \vec{x} \in [l_1, u_1] \times \cdots \times [l_d, u_d]. \end{array}$$

Vlastnosti funkcí

Je zřejmé, že některé typy funkcí budou pro evoluční algoritmy lehčí, než jiné. Velký vliv na efektivitu evolučního algoritmu mají především vlastnosti jako multi-modalita, separabilita a podmíněnost.

Funkce je *multi-modální*, pokud má velké množství lokálních optim. Je zřejmé, že v takovém případě může mít algoritmus problém s uváznutím v lokálním optimu a je potřeba tomu přizpůsobit operátory. Existuje i oblast multi-modální optimalizace, kde je cílem najít co nejvíce různorodých lokálních optim.

Separabilní funkce jsou naopak pro optimalizaci jednodušší. Jsou to takové funkce, které se dají zapsat pomocí funkcí jedné proměnné. Formálně, funkce $f(x_1, \dots, x_n) : \mathcal{R}^n \rightarrow \mathcal{R}$ je *aditivně separabilní*, pokud se dá zapsat jako součet funkcí $f_1(x_1), \dots, f_n(x_n)$, tj. $f(x_1, \dots, x_n) = \sum_{i=1}^n f_i(x_i)$. Obdobně můžeme zadefinovat i funkci multiplikativně separabilní. Z hlediska optimalizace je velkou výhodou separabilních funkcí, že se dají optimalizovat po jednotlivých složkách vektoru, tj. optimum můžeme najít tak, že vždy zafixujeme hodnoty $n - 1$ parametrů a optimalizujeme jen podle jednoho.

Další vlastností, které výrazně ovlivňuje evoluci je podmíněnost funkce. Ta vyjadřuje, jak moc se liší vliv jednotlivých proměnných na hodnotu funkce. Pro kvadratické funkce (u kterých vrstevnice vypadají jako elipsoidy), je jejich podmíněnost druhá odmocnina poměru mezi délkou nejdelší a nejkratší osy elipsoidu. Pokud je podmíněnost funkce velká, říká se, že je funkce špatně podmíněná. Z hlediska evolučního algoritmu je důležité, že by s různým vlivem proměnných na hodnotu funkce měly počítat operátory.

Příklady posledních dvou vlastností vidíme na obrázku 1, který ukazuje vrstevnice funkcí dvou proměnných. Horní funkce je jednoduchý dvoudimenzionální paraboloid, který je separabilní a dobře podmíněný. Na prostředním obrázku je paraboloid, který má jednu osu delší než druhou a znázorňuje tedy špatnou podmíněnost¹⁰. Poslední funkce potom kombinuje špatnou podmíněnost s neseparabilitou.

Výše uvedené vlastnosti jsou ty, které nejvíce ovlivňují efektivitu evolučních algoritmů při spojitě optimalizaci, nejsou to ale všechny. Některé algoritmy například mohou využívat různé symetrie dané funkce, naopak funkce, které mají své globální optimum jen ve velmi malé oblasti prohledávaného prostoru a jinak jsou konstantní jsou pro evoluci velmi těžké obecně.

Kódování pro spojitou optimalizaci

Jedno z prvních rozhodnutí, které je potřeba udělat při návrhu evolučního algoritmu je výběr kódování jedince. Vzhledem k tomu, že ve spojitě optimalizaci pracujeme s vektory reálných čísel, je otázka výběru kódování relativně snadná a jedinci jsou v naprosté většině případů kódování jako vektor typu float nebo double.

??? Existuje pro tohle ↓ nějaká reference?

Dalo by se uvažovat i o kódování jedince přímo po vzoru Hollandova genetického algoritmu, tj. binárním vektorem, a používat jednoduchá n -bodová křížení a bit-flip mutace, ale taková reprezentace trpí tím, že změna různých bitů v číslech vede k výrazně různým změnám hodnoty (např. změna bitu na konci mantisy vs. změna bitu na začátku exponentu).

Operátory pro spojitou optimalizaci

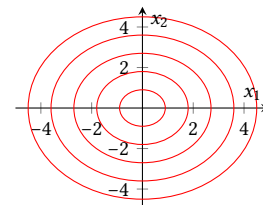
Pro spojitou optimalizaci můžeme samozřejmě použít stejné operátory jako pro každé jiné vektorové kódování jedince, tedy například n -bodové, případně uniformní křížení. Nicméně častěji se používají operátory specializované, které přímo využívají toho, že jedinci jsou vektory čísel.

Tzv. aritmetické křížení vektorů počítá vážený průměr dvou rodičů tak, aby vytvořilo potomka, potomci se tedy spočítají jako

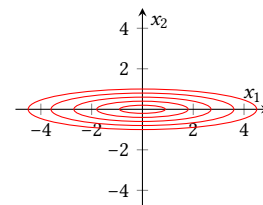
$$o_1 = w * p_1 + (1 - w) * p_2,$$

$$o_2 = (1 - w) * p_1 + w * p_2,$$

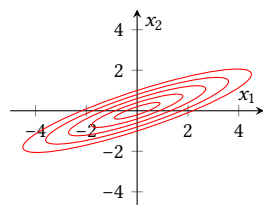
kde p_1 a p_2 jsou rodiče a $w \in (0, 1)$ je (případně náhodně) zvolená váha. Ačkoliv se takové křížení relativně často používá hlavně v jednodušších aplikacích, jeho velkou nevýhodou je, že výslední potomci se nikdy nemohou dostat z konvexního obalu počáteční populace.



(a) Dobře podmíněná separabilní funkce



(b) Špatně podmíněná funkce



(c) Špatně podmíněná a neseparabilní funkce

Obrázek 1: Příklady různých vlastností funkcí

¹⁰ podmíněnost zobrazené funkce jen cca 4.8, nedá se tedy považovat za špatně podmíněnou, objevují se i funkce s podmíněností 10^6

Mutace pro spojitou optimalizaci se občas rozdělují na dva typy – ovlivněné a neovlivněné. Neovlivněná mutace generuje novou hodnotu pro složku vektoru nezávisle na aktuální hodnotě, ovlivněná mutace naopak aktuální hodnotu používá. Typickým příkladem neovlivněné mutace je vygenerování nového čísla z daného rozsahu. Ovlivněná mutace typicky přičítá k dané složce vektoru číslo z normálního rozdělení $\mathcal{N}(\mu, \sigma^2)$.

Jednou z nevýhod výše uvedeného aritmetického křížení je, že potomci jsou velmi často relativně daleko od svých rodičů a nejsou jim tedy moc podobní. Právě z toho důvodu Deb a Agrawal navrhli simulované binární křížení (SBX)¹¹. Hlavní inspirací pro vytvoření SBX bylo jednobodové křížení binárních řetězců, které reprezentují čísla ve dvojkové soustavě. Při náhodné volbě bodu pro křížení velká část nových potomků je relativně blízko k jednomu z rodičů. Jednobodové křížení má navíc další pěknou vlastnost – oba potomci jsou stejně daleko od průměru rodičů.

Cílem SBX je právě simulovat velikosti změn, které se dějí při jednobodovém křížení binárních řetězců. Noví potomci se v SBX křížení spočítají jako

$$o_{1,2} = \frac{1}{2}(p_1 + p_2) \pm \frac{1}{2}\beta(p_2 - p_1),$$

kde p_1 a p_2 jsou rodiče a β je náhodné číslo s pravděpodobnostním rozdělením

$$P(\beta) = \begin{cases} \frac{1}{2}(n+1)\beta^n & \text{pro } \beta \leq 1 \\ \frac{1}{2}(n+1)\frac{1}{\beta^{n+2}} & \text{pro } \beta > 1, \end{cases}$$

kde n je parametr rozdělení, který určuje, jak často budou hodnoty β blízko 1. Pro vyšší hodnoty n je tato pravděpodobnost vyšší (viz Obrázek 2).

Podobnou motivaci jako SBX má i tzv. *polynomiální mutace* (PM)¹². V tomto případě se simuluje velikost změn v bit-flip mutaci binárně kódovaných čísel. To znamená, že nově vytvořený jedinec je s velkou pravděpodobností blízko svému rodiči. Gaussovská mutace se chová podobně, ale u PM je pravděpodobnost malých změn mnohem větší. Nový jedinec se v polynomiální mutaci vytvoří jako

$$o = p + \delta\Delta_{max}$$

kde Δ_{max} je maximální velikost mutace a δ je náhodné číslo z rozdělení

$$P(\delta) = \frac{1}{2}(n+1)(1-|\delta|)^n, \quad \delta \in (-1, 1).$$

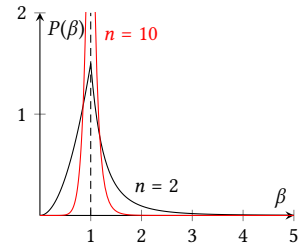
Proměnná n má v tomto případě podobný význam jako u SBX a určuje tvar distribuce. Opět větší hodnota n vede k větší pravděpodobnosti menších změn.

Diferenciální evoluce

Všechny výše popsané operátory mají jednu zásadní nevýhodu, a tou je, že operují s jedinci po složkách, jsou tedy vhodné především pro separabilní funkce a s neseparabilními si neporadí. *Diferenciální evoluce*¹³ přináší jiný operátor, který tímto problémem netrpí. Řeší ho tím, že se při mutaci k jedinci přičítá rozdíl dvou náhodných jedinců z populace. V prvním kroku je tedy vytvořen nový jedinec v diferenciální evoluci následujícím způsobem:

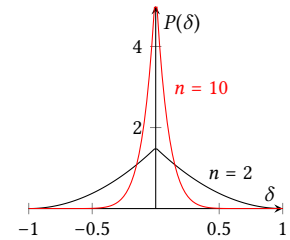
$$o = p_1 + F(p_2 - p_3),$$

¹¹ Kalyanmoy Deb and Ram Bhushan Agrawal. Simulated binary crossover for continuous search space. *Complex systems*, 9(2):115–148, 1995



Obrázek 2: Distribuce β v SBX křížení pro $n = 2$ a $n = 10$.

¹² Kalyanmoy Deb and Mayank Goyal. A combined genetic adaptive search (GeneAS) for engineering design. *Computer Science and Informatics*, 26:30–45, 1996



Obrázek 3: Pravděpodobnostní rozdělení δ v polynomiální mutaci pro $n = 2$ a $n = 10$.

¹³ Rainer Storn and Kenneth Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, Dec 1997

kde p_1 , p_2 a p_3 jsou rodiče a parametr F určuje, jak velké změny se v mutaci budou dělat. Jeho obvyklá hodnota je kolem $F = 0.8$, ale často se objevují hodnoty mezi 0.5 a 1.0, někdy dokonce až 2.0.

Kromě výše uvedené diferenciální mutace se v diferenciální evoluci používá i uniformní křížení s dalším náhodně zvoleným rodičem.

Doplnit popis diferenciální evoluce

Obsah

<i>Před úvodem</i>	3
<i>Jak to je?</i>	3
<i>Jak to bude?</i>	3
<i>Pok pok pokusy</i>	3
 <i>Evoluční algoritmy</i>	5
<i>Evoluce, geny a DNA</i>	5
<i>Obecné schéma evolučního algoritmu</i>	5
 <i>Genetické algoritmy</i>	7
<i>Genetické algoritmy</i>	7
 <i>Evoluční strategie</i>	9
 <i>Spojité optimalizace</i>	11
<i>Vlastnosti funkcí</i>	11
<i>Kódování pro spojitou optimalizaci</i>	12
<i>Operátory pro spojitou optimalizaci</i>	12
<i>Diferenciální evoluce</i>	13
 <i>Rejstřík</i>	23
 <i>Literatura</i>	27

Seznam obrázků

- 1 Příklady různých vlastností funkcí 12
- 2 Distribuce β v SBX křížení pro $n = 2$ a $n = 10$. 13
- 3 Pravděpodobnostní rozdělení δ v polynomiální mutaci pro $n = 2$ a $n = 10$. 13

Seznam tabulek

Seznam algoritmů

1	Schéma evolučního algoritmu	4
2	Schéma evolučního algoritmu	6
3	Schéma Hollandova gentického algoritmu	8







Rejstřík

Evoluční algoritmus, 5
Evoluční programování, 5

Evoluční strategie, 5

Genetický algoritmus, 3, 5, 7

Todo list

	OSNOVA PREDNASKY TAK JAK JE TED	3
	SKORO TO TU MUZEME SMAZAT NE?	3
	Missing ref.	5
	Missing ref.	5
	??? Existuje pro tohle ↓ nějaká reference?	12
	Doplnit popis diferenciální evoluce	14

Literatura

- [1] Hans-Georg Beyer and Hans-Paul Schwefel. Evolution strategies – a comprehensive introduction. 1(1):3–52, May 2002.
- [2] Kalyanmoy Deb and Ram Bhushan Agrawal. Simulated binary crossover for continuous search space. *Complex systems*, 9(2):115–148, 1995.
- [3] Kalyanmoy Deb and Mayank Goyal. A combined genetic adaptive search (GeneAS) for engineering design. *Computer Science and Informatics*, 26:30–45, 1996.
- [4] Agoston E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing*. Springer-Verlag, 2003.
- [5] David B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, Piscataway, NJ, USA, 1995.
- [6] Frédéric Gruau. *Neural Network Synthesis Using Cellular Encoding And The Genetic Algorithm*. PhD thesis, L’universite Claude Bernard-lyon I, 1994.
- [7] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. MIT Press, Cambridge, MA, USA, 1992.
- [8] James Kennedy, James F Kennedy, Russell C Eberhart, and Yuhui Shi. *Swarm intelligence*. Morgan Kaufmann, 2001.
- [9] John R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA, USA, 1992.
- [10] John R. Koza. *Genetic Programming II: Automatic Discovery of Reusable Programs*. MIT Press, Cambridge, MA, USA, 1994.
- [11] John R. Koza. *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*. Kluwer Academic Publishers, Norwell, MA, USA, 2003.
- [12] John R. Koza, David Andre, Forrest H. Bennett, and Martin A. Keane. *Genetic Programming III: Darwinian Invention & Problem Solving*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 1999.
- [13] Zbigniew Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs (3rd Ed.)*. Springer-Verlag, London, UK, UK, 1996.

- [14] Melanie Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA, USA, 1996.
- [15] Gene I. Sher. *Handbook of Neuroevolution Through Erlang*. Springer Publishing Company, Incorporated, 2012.
- [16] Kenneth O Stanley and Risto Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2):99–127, 2002.
- [17] Kenneth O Stanley and Risto Miikkulainen. A taxonomy for artificial embryogeny. *Artificial Life*, 9(2):93–130, 2003.
- [18] Rainer Storn and Kenneth Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, Dec 1997.