

## Obsah

1.	Úvod do genetických algoritmů (GA)	2
1.1	Základní informace	2
1.2	Výstupy z učení	2
1.3	Základní pojmy genetických algoritmů	2
1.3.1	Úvod	2
1.3.2	Základní pojmy	2
1.3.3	Operátor selekce	4
1.3.4	Křížení	6
1.3.5	Mutace	6
1.4	Úlohy GA	6
1.4.1	GA souhrn	6
1.4.2	Souvislost GA a NN	7
1.5	Seznam použité literatury	7

# 1. Úvod do genetických algoritmů (GA)

## 1.1 Základní informace

Následující text je součástí učebních textů předmětu Umělá inteligence a je určen hlavně pro studenty Matematické biologie. Tato kapitola se zabývá základními pojmy z oblasti genetických algoritmů a přibližuje postupy obecně používané při jejich realizaci.

## 1.2 Výstupy z učení

Zvládnutí učebního textu umožní studentům:

- Porozumět základním pojmům jako jedinec, populace, účelová funkce, fitness jedince
- Pochopit operace používané při realizaci genetických algoritmů - selekce, křížení, mutace
- Seznámit se s vybranými selekčními metodami, budou schopni je navzájem porovnat
- Pochopit souvislost GA s algoritmy používanými v NN

## 1.3 Základní pojmy genetických algoritmů

### 1.3.1 Úvod

Genetické algoritmy patří mezi nepříliš jasně omezenou, rozsáhlou třídu algoritmů umělé inteligence, která je inspirována biologií. GA obvykle slouží k řešení optimalizačních úloh. Podobně jako NN se uplatňují v takových oblastech, kde je nalezení řešení definovaného problému analyticky obtížné, či nedostupné. Velmi stručně a zjednodušeně poznamenejme, že vlastnosti všech vyšších živých organismů jsou zakódovány v genetické informaci uložené v sadě buněčných struktur – jaderných chromozomů. Chromozomy obsahují genetickou informaci vyjádřenou v řetězci DNA. Na tomto řetězci můžeme identifikovat jednotlivé úseky, logické funkční celky, které se nazývají geny. Opět s jistým zjednodušením můžeme říci, že konkrétní podoby genů (alely) a jejich kombinace, odpovídají za to, jaké vlastnosti bude mít konkrétní jedinec. Dle Darwinovy evoluční teorie pak jedinci, kteří jsou nejlépe přizpůsobení svými vlastnostmi prostředí (mají vhodný fenotyp), přežívají a rozmnožují se. V průběhu mnoha generací pak evoluční tlak vede k výběru v daném prostředí nejúspěšnější populace.

GA je rodina poměrně mladých algoritmů, která se uplatňuje cca od 60. let dvacátého století a využívá výše zmíněné poznatky biologie. Konkrétně, jak už z jejich názvu vyplývá, se jedná o oblast genetiky a evoluční teorie.

### 1.3.2 Základní pojmy

Mezi základní pojmy GA patří **jedinec**. Ten je v GA v souladu s biologickou analogií popsán sadou svých „chromosomů“ a konkrétním obsahem jednotlivých genů. Konkrétní reprezentace jedince může být v GA představována jediným (haploidní jedinec) řetězcem (polem) číselných hodnot, kde konkrétní pozice v tomto řetězci odpovídá konkrétnímu genu a hodnota v tomto poli pak zprostředkovaně určuje vlastnost daného jedince. Číselné hodnoty nejčastěji implementujeme z důvodu jednoduchosti práce s nimi jako binární, ale mohou obecně nabývat reálných hodnot. Jedince v GA můžeme tedy za použití binárních hodnot popsat jako řetězec

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k) \in \{0,1\}^k$$

(1)

**Populace** je pak tvořena sadou  $N$  jedinců

$$P = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$$

(2)

Aktuální populaci jedinců nazýváme **generací**.

**Účelová funkce**  $f(\alpha)$  provádí mapování z prostoru binárního genotypu do prostoru reálných čísel. Každý jedinec se tedy projevuje prostřednictvím hodnoty své účelové funkce, která modeluje prostředí. Její hodnota je analogií fenotypu a provádí transformaci

$$\{0,1\}^k \rightarrow R$$

(3)

Platí, že čím je hodnota účelové funkce nižší, tím je daný jedinec optimálnější v kontextu daného konkrétního řešeného problému. Jedinec s minimální hodnotou účelové funkce tedy představuje nejvhodnější aktuálně známé řešení daného problému. Účelová funkce je tedy kriteriem, které odhaduje blízkost daného jedince k řešení (jedinci) optimálnímu a které se snažíme minimalizovat. V NN je analogickým minimalizačním procesem hledání optimálních hodnot vah sítě v průběhu adaptační dynamiky – minimalizace chybové funkce sítě.

Každého jedince v populaci tedy můžeme ohodnotit dle jeho potenciálu, kondici k přežití vůči ostatním, dle jeho výhodnosti vzhledem k prostředí. **Fitness** jedince je zobrazení  $F(\alpha)$

$$\forall \alpha_n \in P : f(\alpha_1) \leq f(\alpha_2) \Rightarrow F(\alpha_1) \geq F(\alpha_2) \geq 0$$

(4)

Z praktických důvodů je vhodné zavést normalizaci tak, aby hodnota fitness byla v intervalu  $(0,1)$ , zavádíme tedy renormalizovanou fitness, kde platí:

$$F'(\alpha_i) = \frac{F(\alpha_i)}{\sum_{\alpha_j \in P} F(\alpha_j)}$$

(5)

$$\sum_{\alpha \in P} F'(\alpha) = 1$$

(6)

Nejjednodušším způsobem, jak vyčíslit fitness jedince je provést lineární interpolaci jednotlivých bodů fitness z účelové funkce  $f \rightarrow F$

$$F(\alpha) = \frac{1}{f_{\min} - f_{\max}} [(1 - \varepsilon)f(\alpha) + f_{\min}\varepsilon - f_{\max}]$$

, kde  $\varepsilon$  je hodnota blízká nule.

(7)

Řešení úlohy pomocí GA vypadá rámcově tak, že máme k dispozici počáteční populaci jedinců, představující několik prvků z prostoru řešení. Abychom byli schopni prozkoumat pokud možno co nejširší oblast řešení, musíme mít k dispozici možnost vytvářet jedince nové, necháme tedy stávající populaci rozmnožit a vygenerujeme tak populaci novou představující další prvky z prostoru řešení. Celý prostor možných řešení ale ve většině případů vzhledem k jeho velkým rozměrům vygenerovat a vyhodnotit nelze. Proto nenecháme rozmnožovat všechny jedince, ale pouze jedince nejnadějnější z pohledu fitness. Ostatní jedinci s nižší hodnotou fitness se zpravidla nerozmnožují a umírají. Je tak vytvořena populace nová, která je představována potomky z populace předcházející. Celková velikost populace zůstává zpravidla konstantní.

Řešení či skupina možných řešení úlohy pomocí GA je tedy představováno hledáním takových jedinců, kteří mají z pohledu daného problému optimální vlastnosti. Je zřejmé, že GA musí mít možnost generovat nové jedince a vybírat z nich ty nejlepší. K tomu slouží operátory (křížení, mutace, selekce), které jsou blíže popsány v dalších podkapitolách.

### 1.3.3 Operátor selekce

Operátor selekce slouží k výběru jedinců, kteří jsou v populaci zvoleni k dalšímu rozmnožování. Existuje celá řada algoritmů, které většinou pracují s fitness jedince. Hlavní metody pro výběr jedinců jsou uvedeny dále.

#### 1.3.3.1 Ruletová selekce

Pravděpodobnost výběru  $i$ -tého jedince z populace o  $N$  jedincích je

$$p(i) = \frac{F_i}{\sum_N F_j}$$

(8)

Očekávaná hodnota výběru jedince je rovna podílu kvality daného jedince vůči průměrnému jedinci v populaci (aritmetickému průměru)

$$EV(i) = \frac{F_i}{\frac{\sum_N F_j}{N}} = \frac{N * F_i}{\sum_N F_j}$$

(9)

Očekávaná hodnota výběru jedince je tedy úměrná jeho kvalitám vyjádřeným Fitness funkcí vůči zbytku populace. Tento algoritmus tedy při výběru jedinců od začátku silně preferuje elitnější jedince. Algoritmus rychle konverguje k řešení, ale preferováním elitních jedinců může algoritmus současně během své „přímé“ cesty prostorem možných řešení uváznout v lokálním minimu.

#### 1.3.3.2 Seřadovací metoda

Jedince v populaci seřadíme vzestupně podle hodnoty jejich fitness funkce. Posledního  $N$ -tého jedince ohodnotíme hodnotou  $H > 0$ .

$$H = EV(N)$$

(10)

Očekávanou hodnotu ostatních jedinců v populaci určíme dle vztahu

$$EV(i) = D + (H - D) \frac{i-1}{N-1}$$

, kde  $D = EV(1)$  prvního jedince v řadě.

(11)

Očekávaná hodnota výběru tak závisí pouze na pozici jedince v realizovaném seřazení. Snahou je zabránit předchozímu případu (ruletová selekce), kdy dochází s vysokou pravděpodobností k selekci pouze několika málo elitních jedinců a mohlo tak dojít k nedostatečnému prozkoumání prostoru řešení. Nevýhodné zde naopak může být, že uvedeným mapováním pouze na pozici v řadě ztrácíme informaci o skutečné kvalitě jedince.

### 1.3.3.3 Selektce lineárním a exponenciálním výběrem

Tyto způsoby selekce vyžadují populaci o  $N$  jedincích seřazených od nejhoršího jedince k nejlepšímu dle hodnoty fitness. Pravděpodobnost výběru  $i$ -tého jedince zde pak závisí pouze na pozici (indexu  $i$ ) jedince v daném seřazení. Tato závislost je buď lineární

$$p(i) = \frac{1}{N} (n^- + (n^+ - n^-)) \frac{i-1}{N-1}$$

, kde poměr  $n^-/N$  respektive  $n^+/N$  je pravděpodobnost výběru nejhoršího, respektive nejlepšího jedince.

(12)

nebo exponenciální

$$p(i) = \frac{c^{N-i}}{\sum_N c^{N-j}}$$

(13)

Základ exponentu  $c$  se volí z intervalu  $(0,1)$ . Změnou tohoto parametru lze zvyšovat či snižovat selektivitu algoritmu a vytvářet tak rozdílný populační tlak na výběr jedinců. Selektivní algoritmus s exponenciálním výběrem patří mezi algoritmy v praxi často používané.

### 1.3.3.4 Boltzmanův výběr

Boltzmanův výběr představuje analogii se simulovaným žíháním, se kterým jsme se setkali v Boltzmanově stroji. Pracuje tedy s proměnnou nazývanou teplotou  $T$ . Proměnná  $T$  se mění v čase, respektive se mění v průběhu jednotlivých generací populací jedinců. V počátku je opět teplota nastavena jako vysoká a postupně klesá k nule. Tím se postupně zvyšuje selektivní tlak a při nízké teplotě, tedy ke konci algoritmu, již necháváme rozmnožovat jen nejlepší jedince. Smysl algoritmu je stejný, jako v případě Boltzmanova stroje – tedy omezení pravděpodobnosti uvážnutí v lokálním minimu a větší pravděpodobnost dosažení minima globálního, zde ve smyslu účelové funkce.

$$(14) \quad EV(i) = \frac{Ne^{\frac{F_i}{T}}}{\sum_N F_j}$$

### 1.3.3.5 Další typy selekce

Předcházející metody selekce předpokládali, že je vždy v každé nové generaci vytvořena populace nová a rodičovská populace zaniká. Nemusí tomu tak ale vždy, zmiňme stručně alespoň některé z dalších metod. Metoda **turnajového výběru** má více variant, například náhodně vybereme dva jedince a lepší z nich postoupí do reprodukce. Není tedy nutné vyhodnocovat a porovnávat charakteristiky všech jedinců současně. Další možností metody selekce je **elitismus**, který zachovává nejlepší jedince. Nejúspěšnější jedinci dané generace neumírají, ale jsou přeneseni spolu s potomky do generace následující. Jinou metodou selekce je **ořezávání**, kdy je populace rozdělena na dvě či více částí a do reprodukce postupují pouze jedinci z úspěšnější skupiny. Další alternativou je **náhodný výběr**, kdy jsou jedinci vybíráni naprosto náhodně.

### 1.3.4 Křížení

Operátor selekce dokáže ve stávající generaci populace nalézt ty, kteří se mají rozmnožovat a vytvořit tak populaci novou. Máme tedy k dispozici jedince, představující rodiče. Operátor křížení vygeneruje jejich potomky. Nejčastější je jednobodové křížení dvou rodičů, kdy si rodiče od určitého náhodného genu (pozice v řetězci) zbývající část řetězce vymění,  $(1110011|1011) \times (1111111|1000) \Rightarrow (1110011|1000)$  a  $(1111111|1011)$ .

Dále obvykle rozhodneme, zda postoupí do další generace oba potomci, nebo například jen ten lepší. Operátor křížení opět existuje v celé řadě variant, kdy může docházet i k vícebodovému křížení i mezi více než dvěma rodiči. Nebo mohou někteří rodiče postoupit bez rozmnožování přímo do další generace, jak bylo zmíněno v selekci elitismem.

### 1.3.5 Mutace

Operátor mutace zavádí podobně jako v přírodě do rozmnožování prvek náhody a umožňuje tak uniknout z lokálního minima, pokud se mutace ukáže jako výhodná. Mutace je většinou realizována jako jednobodová, kdy je s určitou pravděpodobností náhodně provedena změna obsahu jednoho genu  $(1110011|1011) \times (1111111|1000) \Rightarrow (1110011|1000)$  a  $(1111111|1111)$ .

Mutace vlastně představuje pokus o útok stranou, kde se zmutovaný jedinec vyskytuje pravděpodobně mimo dosud prozkoumanou oblast prostoru řešení. Jsou tak prověřeny i nové možnosti a prostor řešení je prozkoumáván do větší šířky s ohledem na snahu nalezení globálního minima.

## 1.4 Úlohy GA

### 1.4.1 GA souhrn

Typické kroky GA lze shrnout následovně:

1. Inicializace populace, kódování a nastavení parametrů jedinců. První generace.
2. Ohodnocení populace, definice kritérií pro jedince. Výpočet fitness.
3. Selektce jedinců k reprodukci.
4. Křížení rodičovských jedinců, vytvoření potomků. Výběr potomků pro příští generaci.
5. Mutace některých jedinců v nové populaci.

6. Vyhodnocení jedinců nové populace, zda některý není řešením. Ano = konec.
7. Pokud není, pak pokračuj bodem 2 pro novou populaci.

Genetické algoritmy jsou vhodné pro úlohy, které jsou jinými prostředky obtížně řešitelné analyticky. Jejich implementace bývá relativně jednoduchá, nicméně často vyžadují mnoho strojového času a operační paměti. Velkou implementační výhodou GA je možnost jejich snadné paralelizace na více výpočetních jednotkách. GA vždy dosáhnou nějakého řešení, otázkou je, zda bude dostatečně kvalitní, záruka jeho optimálnosti neexistuje. Lze zde pozorovat jistý rozdíl od klasických algoritmů, které při hledání řešení postupují gradientními metodami, v zásadě tedy v jednom pravděpodobně nejsprávnějším směru dle hodnoty gradientu chybové funkce. GA oproti tomu prozkoumávají širší oblast řešení a operátor mutace umožňuje prozkoumat i řešení zcela mimo aktuální lokální minimum. GA zpravidla rychle konvergují k nalezení konkrétního suboptimálního řešení, ale další zpřesňování k nalezení optimálního řešení v této prohledávané oblasti bývá pomalé. Proto se často používají tzv. hybridní algoritmy, kdy je nejprve použit GA a optimální řešení je pak prohledáváno lokálně gradientní metodou.

#### 1.4.2 Souvislost GA a NN

GA lze s úspěchem použít i pro optimalizaci neuronových sítí. Nejčastěji se GA uplatňují v procesu organizační a adaptační dynamiky, kdy GA optimalizují topologii, respektive váhy neuronové sítě.

GA mohou být ve fázi organizační dynamiky NN používány při nalezení optimálního počtu neuronů, jejich parametrů, uspořádání do vrstev a spojů. Obvykle algoritmy vycházejí buď z minimální struktury sítě a postupně přidávají další neurony, nebo je postup opačný a počáteční komplexní síť je postupně při zachování výkonnostních parametrů zjednodušována. Realizace GA je ve fázi organizační dynamiky s ohledem na nutnost prověření velkého množství komplexních jedinců výpočetně velmi náročná a vyžaduje nemalé systémové prostředky. Každý jedinec je zde představován komplexní topologickou strukturou NN a i jeho samotné prověření a ohodnocení fitness není triviální.

Ve fázi adaptační dynamiky NN je topologie sítě neměnná, GA hledá optimální nastavení hodnot vah jednotlivých neuronů. Váhy mohou být reprezentovány jako reálné hodnoty, ale obvykle jsou kódovány jako binární, což vede na jednodušší realizaci GA. Nevýhodou je, že jsme schopni při dané délce řetězce popisujícího jedince zachytit pouze omezený počet hodnot vah, což nemusí být dostatečné. Prodlužováním řetězce lze jejich hodnoty zpřesňovat, nicméně jejich prodlužování významně zpomalí GA.

GA je možné použít i současně jak ve fázi adaptační, tak organizační dynamiky. Přínosnější se jeví použití GA i přes jejich náročnost ve fázi dynamiky organizační, protože GA jsou schopny nalézt při zachování výkonnostních parametrů sítí s co možná nejjednodušší a nejoptimálnější strukturou. Pro adaptaci vah sítě máme mimo GA k dispozici například gradientní backpropagation algoritmus, který je mnohem méně výpočetně náročný a přes nebezpečí uváznutí v lokálních minimech přináší většinou uspokojivé výsledky.

### 1.5 Seznam použité literatury

- [1] Šíma, J., Neruda, R.: Teoretické otázky neuronových sítí, MATFYZPRESS, 1996, ISBN 80-85863-18-9.
- [2] Volná, E.: Neuronové sítě 2. Skripta Ostravská universita v Ostravě, Ostrava, 2008.

- [3] V. Kvasnička, L. Beňušková, J. Pospíchal, I. Farkaš, P. Tiňo, and A. Král. Úvod do teórie neurónových sietí, IRIS, Bratislava, 1997, ISBN 80-88778-30-1.
- [4] PAVLOVIČ, Jan. Multikriteriální hybridní evoluční algoritmy pro výběr a optimalizaci dekontaminačních technologií, 2008. Rigorózní práce, Masarykova univerzita, [http://is.muni.cz/th/4035/fi\\_r/](http://is.muni.cz/th/4035/fi_r/).