

MATLAB TOOLBOX – GENETICKÉ ALGORITMY

Ivan Sekaj, Martin Foltin

Katedra ASR, Fakulta elektrotechniky a informatiky
Slovenská technická univerzita Bratislava, Ilkovičova 3, 812 19 Bratislava

1. Úvod

V súčasnosti sa čoraz častejšie na riešenie zložitých (technických ale aj netechnických) problémov používajú tzv. "evolučné výpočtové techniky" alebo evolučné algoritmy (evolutionary algorithms). Sú to metódy resp. algoritmy, ktoré vo svojej podstate napodobňujú mechanizmy biologickej evolúcie. Medzi evolučné algoritmy spravidla zahrňujeme viaceré prístupy ako Evolučné stratégie, Genetické algoritmy, Genetické programovanie, Umelý imunitný systém, prípadne iné. Všetky tieto prístupy majú veľa spoločných vlastností, ktorých základ tvorí optimalizácia na báze stochastických zmien a súťaženia jednotlivých potenciálnych riešení.

Cieľom tohto príspevku je predstaviť výpočtový prostriedok – Genetic Toolbox, ktorý bol vyvinutý na našom pracovisku, a ktorý dovoľuje aplikovať takéto formy výpočtov, predovšetkým však genetických algoritmov (GA) v programovom prostredí Matlab.

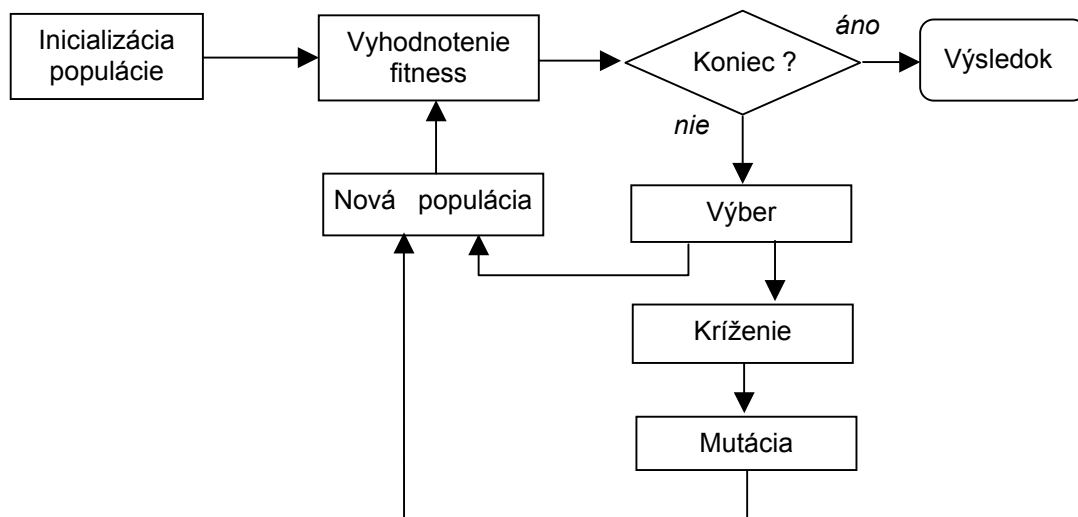
2. Princíp činnosti Genetických algoritmov

GA sú univerzálnym stochastickým prehľadávacím alebo optimalizačným prístupom, ktorý je v ohraničenom priestore prípustných riešení daného problému schopný nájsť alebo sa aspoň priblížiť ku globálnemu optimu [1],[2]. Uplatňuje sa pri tom z prírody vypozerovaný princíp prežitia najsilnejších resp. najprispôsobivejších jedincov a nevyhnutnosť zániku najslabších, neživotaschopných resp. neprispôsobivých.

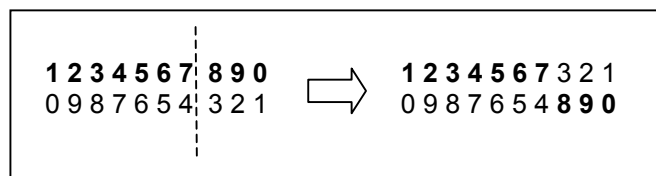
Principiálne vysvetlenie mechanizmu GA je zrejmé z obr.1. Algoritmus pracuje so skupinou potenciálnych riešení – s tzv. populáciou. Každé potenciálne riešenie je pritom reprezentované usporiadanou množinou parametrov alebo hodnôt, ktoré charakterizujú jeho vlastnosti. Prvky tejto množiny sa nazývajú *gény* a môžu byť binárneho, celočíselného, reálnečíselného, symbolového alebo kombinovaného typu v závislosti od charakteru daného problému. Sú usporiadané do postupnosti, ktorá sa nazýva *reťazec* alebo *chromozóm*. Počiatočná populácia reťazcov v prvom výpočtovom cykle (*v prvej generácii*) sa získa spravidla náhodným vygenerovaním ich génov v rámci uvažovaných ohraničení. Pre každé riešenie, ktoré sa dekoduje z reťazca do existujúceho počítačového modelu, sa vyčíslí výpočtom, počítačovou simuláciou atď. hodnota účelovej funkcie – *fitness*. Fitness je vlastne miera vhodnosti alebo úspešnosti daného reťazca alebo *jedinca*. Všetky jedince populácie sa navzájom porovnajú a potom sa vyberie skupina jedincov, ktoré sa nezmenené dostanú do novej populácie. Tiež sa vyberie druhá skupina jedincov, ktorá je určená na inováciu. V tejto skupine sa vytvoria náhodné páry reťazcov, s ktorými sa uskutoční genetická operácia *križenie*. Potom sa na tejto skupine realizuje ešte *mutácia*. Takto zmodifikované jedince dokompletujú novú populáciu, ktorá sa stane objektom rovnakého postupu v ďalšej generácii. Pritom je dôležité, že pri výbere do oboch

skupín majú najväčšiu pravdepodobnosť "prežitia" najúspešnejšie jedince, ale istú šancu majú aj tie menej úspešné. Ak sa uvedený postup počas mnohých generácií opakuje (napr. 100 x, 1000 x alebo 1 milión x), riešenie konverguje ku globálnemu optimu. Počet generácií závisí od povahy a zložitosti riešeného problému. Beh algoritmu sa môže ukončiť po dosiahnutí požadovaného resp. prijateľného riešenia alebo po ukončení určeného počtu generácií.

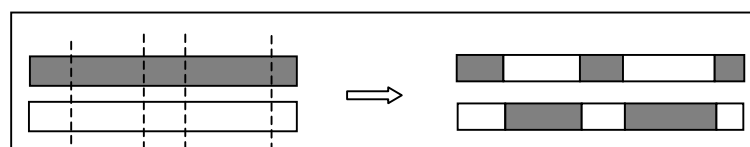
Spomínali sme tri typy operácií – kríženie, mutácia a výber. Operácia kríženia náhodne skombinuje gény dvoch rodičov do jedného alebo viacerých potomkov. Pri bežnom spôsobe kríženia sa dva rodičovské reťazce rozdelia na jednom alebo viacerých náhodných miestach (oba reťazce na rovnakých) a potomkovia získajú striedavo každú druhú časť takto oddelených podreťazcov od každého z rodičov (obr.2, obr.3). Operácia mutácie náhodne zmení náhodne zvolené gény náhodne vybraných jedincov (obr.4). Poznamenajme ešte, že existujú viaceré typy kríženia aj mutácie. Spôsobov výberu jedincov do nových populácií je tiež niekoľko druhov, líšia sa mierou a spôsobom preferovania najúspešnejších jedincov oproti náhodne vybraným jedincom. Voľba všetkých uvedených genetických operácií je ovplyvňovaná typom riešeného problému.



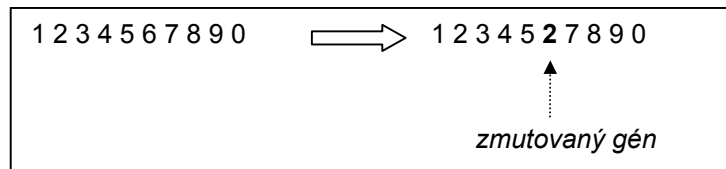
Obr. 1: Bloková schéma genetického algoritmu



Obr. 2: Príklad jednobodového kríženia dvoch celočíselných reťazcov



Obr. 3: Viacbodové kríženie dvoch reťazcov



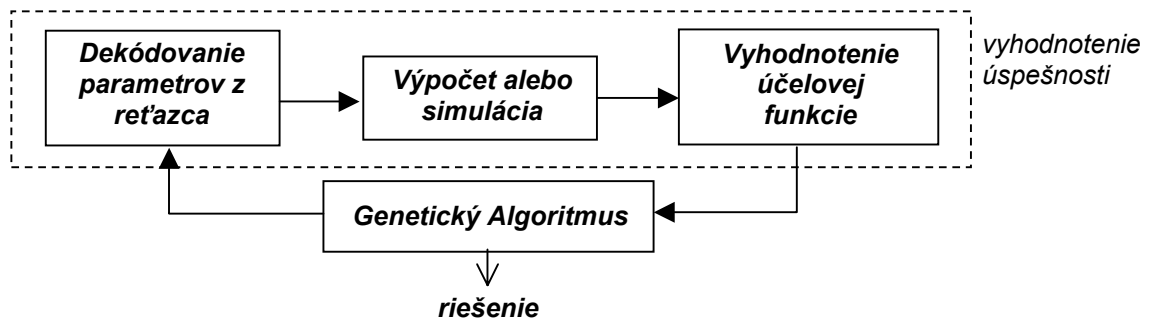
Obr. 4: Príklad mutácie celočíselného reťazca

GA sa od väčšiny konvenčných optimalizačných metód líšia niektorými znakmi:

- dokážu vyviaznúť z okolia lokálneho extrému a približovať sa ku globálnemu extrému,
- uskutočňujú paralelné prehľadávanie vo viacerých smeroch súčasne,
- nevyžadujú pomocné informácie o vývoji riešenia, ako je napr. gradient účelovej funkcie (na rozdiel od mnohých iných metód),
- intenzívne využívajú stochastické javy,
- sú schopné riešiť optimalizačné problémy s desiatkami až stovkami premenných,
- možno ich pomerne jednoducho aplikovať na riešenie širokého spektra rôznych typov problémov,
- patria k časovo resp. výpočtovo najnáročnejším prístupom.

3. Možnosti použitia GA

Ako bolo naznačené, genetické alebo evolučné algoritmy sa dajú použiť na riešenie veľmi širokého spektra úloh. Zopakujeme, že podmienkou je schopnosť sformulovať účelovú funkciu, ktorá sa má minimalizovať alebo maximalizovať. Pri minimalizácii ide obyčajne o minimalizáciu odchýlky (chyby) od požadovaného stavu, minimalizáciu spotreby energie, paliva, strát alebo nákladov, o minimalizáciu nežiaducich účinkov a podobne. Pri maximalizácii môže ísť o maximalizáciu účinnosti, výkonu, zisku atď. Ďalšou podmienkou ich použitia je existencia počítačovej reprezentácie optimalizovaného problému. To znamená, že pre ľubovoľný bod prehľadávaného priestoru, teda pre ľubovoľné potenciálne riešenie vieme počítačom vyčísliť hodnotu jeho účelovej funkcie – ohodnotiť ho z hľadiska úspešnosti, z hľadiska miery splnenia požadovaného cieľa. Pritom vôbec nezáleží na type daného procesu, ktorý môže byť fyzikálneho, chemického, biologického, ekonomického alebo hoci sociologického charakteru. Veľmi zjednodušene povedané: optimalizovaný problém z pohľadu evolučného algoritmu je čierna skrinka, ktorá poskytuje veľké množstvo viac alebo menej zmysluplných možností riešenia, pričom každú z nich vieme navonok ohodnotiť napríklad číselne alebo aspoň porovnať jej úspešnosť voči iným možnostiam. Pod optimalizáciou možno potom chápať hľadanie takej možnosti (sady parametrov, štruktúry vnútorných väzieb a pod.), ktorá najlepšie splňuje určené požiadavky. Aj keď to na prvý pohľad nie je vždy zrejmé, úlohu možno takto postaviť aj pre veľmi veľkú množinu problémov, ktoré vo svojej pôvodnej formulácii nie sú klasickými optimalizačnými úlohami. Často aj úplne bežné matematické a iné problémy sa dajú veľmi jednoducho transformovať na minimalizačné resp. maximalizačné úlohy. Schématické znázornenie riešenia pomocou GA je na obr.5



Obr. 5 Použitie GA pri riešení problému

Medzi problémy, ktoré sa použitím konvenčných optimalizačných prístupov riešia len ťažko alebo dokonca sa nedajú riešiť vôbec, a ktoré môžu byť typickými aplikáciami genetických, resp. evolučných algoritmov, patrí mnoho komplikovaných a/alebo rozsiahlych problémov. Významným poľom pôsobnosti sú inžinierske aplikácie. GA môžu byť silným nástrojom pri optimalizácii elektrických obvodov, pri optimalizácii prevádzky elektrizačnej sústavy, pri návrhu antén a filtrov, pri statickej optimalizácii technologických procesov alebo pri návrhu regulačných obvodov. Zo strojárskych aplikácií spomeňme napríklad návrh prevodoviek alebo návrh rezných plánov. Známe je využitie GA pri optimalizácii motorov Boeingu 777, keď sa zdanlivo malou konštrukčnou úpravou získala na dané pomery mimoriadne významná úspora paliva až 2,5 percenta. Toto predstavuje pri jednom lietadle za rok úsporu 2 miliónov dolárov. Výnimkou nie sú ani aplikácie v stavebníctve pri optimalizácii konštrukcie budov, cestných komunikácií alebo sietí. Veľmi široké uplatnenie nachádzajú evolučné techniky pri riešení ekonomických problémov, ktoré často predstavujú zložité úlohy lineárneho alebo nelineárneho programovania s mnohými obmedzeniami. Typickými aplikáciami sú dopravné alebo distribučné úlohy, ako napríklad hľadanie najkratšej alebo najlacnejšej cesty, iné grafovo orientované úlohy, ťažké kombinatorické problémy a podobne.

4. Toolbox Genetické Algoritmy

Toolbox - Genetické algoritmy pre Matlab je určený na riešenie predovšetkým reálnečíselných (prípadne celočíselných) optimalizačných problémov, čiže takých, ktoré používajú reálnečíselné kódovanie reťazcov. Jeho ďalším špecifikom je, že je minimalizujúci, čiže za lepšie jedince, ktoré sú preferované, považuje tie, ktoré dosahujú menších hodnôt účelovej funkcie resp. fitness funkcie. V prípade maximalizácie sa použije doplnková účelová funkcia (viď funkcia *invfit* alebo účelová funkcia pre násobená číslom -1). Účelová funkcia alebo fitness môže nadobúdať kladné aj záporné reálnečíselné hodnoty.

Toolbox [3] obsahuje 23 funkcií, ktoré je možné rozdeliť do niekoľkých skupín. Prvou skupinou sú genetické operácie na báze mutácie alebo kríženia, do ktorej patria funkcie:

mutx - obyčajná mutácia

muta - aditívna mutácia s rovnomerným rozdelením pravdepodobnosti

mutn - aditívna mutácia s normálnym rozdelením pravdepodobnosti

mutm - multiplikatívna mutácia

change - mutácia duplicitných reťazcov populácie
crossover - viacbodové kríženie dvoch reťazcov
crosdis - diskkrétne kríženie
around - medziľahlé kríženie.

Z metód výberu sú k dispozícii:

selbest - výber najúspešnejších reťazcov
selsort - výber a zotriedenie reťazcov podľa úspešnosti
selrand - náhodný výber reťazcov
seldiv - výber podľa maximálnej diverzity
selsus - stochastický rovnomerný výber
seltourn - turnajový výber
selwru - výber pomocou váhovaného ruletového kolesa.

Ďalšiu skupinu tvoria pomocné funkcie:

genrpop - vygenerovanie náhodnej reálnečíselnej populácie
invfit - doplnok účelovej funkcie
shake - náhodné premiešanie poradia reťazcov v populácii.

Špecifickú triedu tvoria funkcie, ktoré slúžia na optimalizáciu problémov tzv. "permutačného typu". Ide o úlohy hľadania optimálneho poradia, na ktoré sa používajú odlišné typy operácií, ako pri bežných úlohách parametrickej optimalizácie. Sú to funkcie:

crosgp - kríženie medzi viacerými rodičmi
crosord - kríženie permutačne kódovaných reťazcov
invord - inverzia poradia génov v subreťazci
swapgen - mutácia poradia génov v reťazci
swappart - vzájomná výmena poradia dvoch častí reťazcov.

Použitím toolboxu Genetické Algoritmy (Genetic Toolbox) je možné vytvoriť algoritmy prispôbené na mieru konkrétnym aplikáciám a požiadavkám používateľa. Je možné realizovať rôzne typy schém, ktoré sú odlišné od schémy na obr.1 a tiež použiť rôzne verzie genetických operácií. Popri jednoduchých typoch GA sú pomocou toolboxu realizovateľné rôzne typy paralelných GA štruktúr, adaptívne verzie GA a podobne.

5. Záver

Genetické algoritmy sú výkonným a perspektívnym optimalizačným nástrojom, ktorého frekvencia a sféry použitia sa v súčasnosti markantne rozširujú. Výpočtové prostredie Matlabu poskytuje na tieto účely veľmi dobrú a flexibilnú platformu. Toolbox Genetické Algoritmy bol na našom pracovisku mnohonásobne úspešne aplikovaný pri riešení rôzneho typu úloh. Predovšetkým sa jednalo o oblasť návrhu regulačných obvodov [4],[5], statickú optimalizáciu procesov [6], riešenie ekonomických úloh, dopravných úloh, manipulačných úloh, úloh s rôznymi typmi obmedzení, riešenie viackriteriálnej optimalizácie, riešenie celočíselných úloh a pod.

Literatúra

- [1] D.E.Goldberg: Genetic Algorithms in Search, Optimisation and Machine Learning, Addison-Wesley, 1989
- [2] Z.Michalewicz: Genetic Algorithms + Data Structures = Evolutionary Programs, Springer, 1996
- [3] I.Sekaj: Genetic Toolbox, používateľská príručka, KASR FEI STU, Bratislava, interný material, 1998
- [4] I.Sekaj: Genetic Algorithm-based Control System Design and System Identification, Proceedings on the Int. Conference Mendel '99, June 9-12, Brno, Czech Republic, pp.139-144, 1999
- [5] I.Sekaj, M.Foltin, M. Gonos: Genetic Algorithm Based Adaptive Control of an Electromechanical MIMO System, Proceedings of the GECCO 2002 Conference, July 9-13, New York, pp. 696
- [6] I.Sekaj: Genetic Algorithms with Changing Criterion functions, Intelligent Technologies – Theory and Applications, P.Synčák et al. (Eds), IOS Press, pp.183-188, 2002

E-mail: sekaj@kasr.elf.stuba.sk