XI. FIATAL MŰSZAKIAK TUDOMÁNYOS ÜLÉSSZAKA

Kolozsvár, 2006. március 24-25.

GÉNSEBÉSZET MINT OPTIMALIZÁLÁSI PARADIGMA

Iclănzan David

Abstract

Genetic algorithms (GA) have several important key features, like parallel processing, the ability to work with highly nonlinear, noisy systems, tolerance for imprecision, uncertainty and partial truth etc. that makes them a useful tool when it comes to solve “real life” optimization problems, found in many technological areas. But they suffer from a tendency to converge prematurely or stagnate without fitness progress for highly dimensioned complex optimization problems which leads to an excessively long run-time that is needed to deliver satisfactory results. This may be due to the loss of the capability to discover new genetic material before the increased selection pressure has limited the algorithm’s ability to search widely. In this paper the Genetic Engineering Algorithm (GEA) is proposed, where the selection pressure is deliberately loosened to induce an enhanced exploratory power. On the other hand good individuals are strongly exploited by performing local search in their vicinity and by introduction of a deterministic crossover operator. We show that by maintaining a low selection pressure while fostering the exploitation of high-fitness individuals, the new method is able to significantly outperform GA and other heuristic methods on benchmark tests.

Összefoglalás

A Genetikus Algoritmusoknak (GA) számos olyan előnye van (*többpontos, párhuzamos keresést tesz lehetővé; erősen nemlineáris rendszerekkel, bizonytalan, pontatlan és hiányos ismeretekkel szemben is robusztusan viselkedik stb.*), amely ezt a lágy számítási módszert alkalmassá teszi a műszaki tudományok által felvetett nehéz optimalizálási feladatok megoldására. De sokszor e módszereknek a konvergencia kimondottan lassú magas dimenziójú komplex optimalizálási feladatok esetén. Ez talán annak nyilvánítható, hogy a túl gyorsan megnőtt szelekciós nyomás nem engedi meg új értékes génállományok felkutatását, így a módszer korai konvergenciáját, szuboptimális pontokban - hosszú ideig - való leragadását eredményezi. A dolgozatban egy új módszert tagolok, a Génsebészeti Algoritmust (GSA), melyben a megfelelő felfedező hatást a szelekciós nyomás lényeges csökkentése biztosítja. A gyors konvergenciát a jó egyedek potenciáljának intenzív kiaknázásával éri el, lokális keresést végezve szomszédságukban és alávétve őket egy új determinisztikus, keresztezési operátornak. A teszt eredmények arra utalnak, hogy alacsony szelekciós nyomást alkalmazva és a legjobb egyedek potenciálját maximálisan kihasználva az új módszer képes egyértelműen felülmúlni a GA-t és más optimalizálási eljárásokat.

1. Bevezetés

Az 1950-60-as években merült fel először az a gondolat, hogy a biológiai evolúció a mérnöki problémákban is felhasználható lehet optimalizációs eszközként. Az azóta kidolgozott, tanulmányozott módszerek egyik fontos pillérét a genetikai algoritmusok (GA) alkotják [1].

A GA egy probléma független heurisztika, mely kevés tudást igénylő optimalizálást képes végezni általános keresési terekben. A darwini evolúciós mechanizmusokat utánozó technika bizonytalan és pontatlan információs környezetben is eredményesen alkalmazható, és hatékonyságában sokszor felülmúlja a „hagyományos” optimalizálási eljárásokat, elsősorban flexibilitásának, a párhuzamos – többpontos – keresésnek és a módszerben rejlő nagy szabadságnak köszönhetően.

A sok előny mellett a módszer sajnos számos hátránnyal is rendelkezik: a konvergencia általánosan nem bizonyított és kimondottan lassú, az algoritmus jó paraméterezésére (korai konvergencia elkerülése vs. rövid futási idő) nincs elméleti háttér, a kapott eredményről sokat nem tudunk mondani (az eredmény jó, de mennyire?). Ezek nagymértékben annak köszönhetőek, hogy a szelekciós nyomás túl hamar elér olyan szintet, mely már nem engedi meg új értékes génállományok felkutatását. Ez okozza a módszer korai konvergenciáját, lokális optimumokban - hosszú ideig - való leragadását.

E dolgozatban bemutatásra kerül egy új, eredeti optimalizálási paradigma, miszerint a gyors evolúció nem a magas szelekciós nyomás által van elérve, hanem a legjobb egyedet „génsebészetnek” alávetése által. Mivel az alacsony szelekciós nyomás megengedi a hasznos génállomány felkutatását, míg a génsebészet biztosítja a legjobb egyedek gyors evolúcióját, az új módszer nagymértékben képes áthidalni a GA számos előnytelenységét.

2. Új paradigma

Két alapvető tényező van a genetikai keresés folyamatában [2]: a populáció változatossága (szórása a keresési térben) és a szelekciós nyomás. Ezek funkcionálisan függnek, nagymértékben befolyásolják egymást: a szelekciós nyomás növekedése csökkenti a populáció szórását és fordítva. Ezért egy túl nagy szelekciós nyomás korai konvergenciához vezethet, míg egy túl kicsi a keresést hatékonyságát csökkenti drasztikusan. Whitley szerint [3], ez nem más mint Holland [1] és mások által tárgyalt „*exploration vs. exploitation*” működési elv variációja. Megmutatta, hogy a sok paraméter mely használatának a célja a genetikai keresés „tuningolása” igazából csak indirekt eszköz, mely a *populáció szórására*, illetve a *szelekciós nyomásra* hat. A szelekciós nyomás növekedésével a keresés a populáció legjobb egyedire fókuszál, ez intenzív „kikhasználás” (*exploitation*) mellékhatása a genetikai állomány változatosságának elvesztése. Csökkentve a nyomást növeljük a „bejárást, felfedezést” (*exploration*) mivel a keresésben több genotípust vonunk be, implicite több skémát, hipersíkpartíciót.

A tudományos közvélemény [1, 3-5] hamar úgy vélte, hogy az algoritmus jó működése érdekében nagyon fontos egy egyensúlyt kialakítani e két alapvető tényező között. Ezért a kutatások ezen a téren az utóbbi 25 évben arra koncentráltak, hogy egy általános mintavételező mechanizmust alakítsanak ki, mely biztosítaná ezt a vitális egyensúlyt. Ezt a célt mai napig nem sikerült teljes mértékben teljesíteni.

E dolgozatban egy paradigma cserét helyezek kilátásba, miszerint ne próbáljunk egy ideális szelekciós mechanizmust kidolgozni mely biztosítja az egyensúlyt a gyors konvergencia és a megbízhatóság között. A fentebb bemutatott funkcionális függőség megszüntethető oly módon, hogy az algoritmus alapvető mechanizmusai ne sérüljenek meg, ez által kapható egy új optimalizációs módszer, mely egyben gyors konvergenciával és nagy megbízhatósággal rendelkezik. Kicsi szelekciós nyomással dolgozva biztosíthatjuk a nagy megbízhatóságot. A gyors konvergencia elérésében nem támaszkodunk a genetikai operátorokra, ezáltal feloldva a függőséget, hanem más mechanizmusokat, operátorokat használunk az úgynevezett génsebészeti (GS) eljárást. Ezek az új operátorok csak a populáció kicsi

százalékán (kb. 3%), a legrepresentatívabb egyedeken, a génsebészethez hasonló hatást fejtenek ki: a kívánatos jellegzetességek gyors kifejlődnek és a rosszak elnyomódnak.

A GSA működési elve a következő: az algoritmus a keresési tér bejárásával felfedezi a lokális optimumokat, ezek közelébe egyedeket helyez el. Ezeket a GS átveszi, és szuperegyedeket képez, azaz az optimumot nagy pontossággal, gyorsan megközelíti.

3. A Génsebészeti eljárás és az algoritmus

A genetikai operátorok, a jobb egyedek kifejlesztésénél, véletlent felhasználó, perturbációs jelleget mutatnak. Ezért a legkézenfekvőbb helyettesítésnek a szimulált hűtést találtam - egy egyeddel hatékonyan dolgozik, nem igényli a populáció csoportosulását, lokális minimumokat is képes átugrani. A GS eljárás, mivel nem hipersíkpartíciókkal dolgozik, értékesítheti a lokális optimum körüli adatokat direkt módon. Ha az atraktor medencéje az optimumoknak elég kicsi, akkor egy kimerítő keresést is alkalmazhatunk. Másik használt heurisztikus eljárás az úgynevezett „*gén infúzió*”. Vesszük a legjobb egyedet és a legtávolabb álló, de mégis magas fitnessszértékkel rendelkező reprezentánst. A második egyed génállományának elemeit sorra behelyezzük a legjobb egyed génállományának a megfelelő pozíciójára. Ha így jobb fitnessz értéket kapunk, a változtatást véglegesítjük, ha nem, eldobjuk. A motiváció e módszere az, hogy egy távol álló egyed, mivel neki is magas fitnessz értéke van, valószínű, hogy értékes génekkel rendelkezik, melyek különböznek a legjobb egyedtől. Ezek sikeres felkutatása nagyon felgyorsíthatja a konvergenciát.

A GSA is természetesen ciklusos működés során fejt ki hatását. A ciklus lépései sorrendben (1-4):

- (0) a kezdeti populáció véletlenszerű kiválasztása után;
- (1) a következő generáció létrehozásának előkészítése, azaz az aktuális generációt alkotó egyedek „jószágának” kiszámítása és az így megkapott alkalmasságuktól függően a következő generációs egyedek szülőinek a kiválasztása, gyenge szelekciós nyomás alatt;
- (2) az új egyedek létrehozása reprodukciós műveletek, mutáció, keresztezés végrehajtása révén;
- (3) ha adott generációszámnál vagyunk a populáció reprezentatív egyedeit alávétjük a GS eljárásnak;
- (4) az új generáció meghatározása, kicsi szelekciós nyomás alatt de részben elitista módon mely védi a GS eljárás által kifejlesztet „szuper” egyedeket.

4. Teszt eredmények

Két nehéz minimalizálási feladatot - erősen multimodális szintetikus tesztkörnyezet választottam a De Jong, illetve Schewefel mesterséges fitnesszdomborzatokból a GA, GSA és szimulált hűtés teljesítmény összehasonlítására. A függvények részletes ismertetése [6]-ban megtalálható. Statisztikai megbízhatóság érdekében 10 futatás eredményét átlagoltam.

Ahogy az 1. ábrán látható a GSA képes gyors és pontos konvergenciát biztosítani a tesztfeladatokra, míg a másik két módszer esetén a konvergencia lassú és/vagy pontatlan a sok lokális optimum – erős multimoduságnak köszönhetően.

1. ábra. A módszerek konvergenciája 10000 függvény kiértékelés esetén. A minimalizálandó célfüggvények: egy 30 dimenziós Ackley illetve egy 10 dimenziós Fletcher Powell mesterséges fitnesszomborzat. Minden lépésben a módszerek által addig megtalált legjobb megoldás van kirajzolva.

5. Következtetések

A GSA érdekes és új optimalizálási paradigmát helyez kilátásba, miszerint nem használ különböző mechanizmusokat a lokális optimumokból való kiszabadulásra, hanem a bennük való leragadást előzi meg gyenge szelekciós nyomást gyakorolva az egyedekre. Másfelől a legjobb egyedeket ösztönzi gyors és pontos konvergenciára intenzív génebeszeti eljárásnak vetve őket alá. A teszteredmények arra utalnak, hogy ezt a filozófiát implementáló eljárás hatékonyabb lehet más módszereknél, komplex optimalizálási feladatok esetén.

Irodalom

- [1] Holland, J.H, *Adaption in Natural Artificial Systems*, Michigan Press, Ann Arbor, 1975.
- [2] Zbignev, M, *GA + Data Structures = Evolution Programs*, 3rd Edition, Springer, 1997.
- [3] Whitley, D, *The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best*, in Proceedings of the RMCAI, Denver. 1988, pp. 116-121.
- [4] Brindle, A, *GA for Function Optimization*, PhD thesis, university of Alberta, Edmonton, 1981.
- [5] Baker, J.E, *Adaptive Selection Methods for GA*, *Proceedings of the I International Conference on Genetic Algorithms*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1985, pp. 101-111.
- [6] Iclănzan, D. *Genetic Engineering Algorithm*, Proceedings of the 7th International Scientific Student Conference on Technical Sciences, Timișoara, 2005. ISBN: 973-578-195-1.

Iclănzan David, abszolvens

SAPIENTIA EMTE Műszaki és Humántudományok Kar Marosvásárhely

Cím: 540053, ROMANIA, Târgu Mures, P-ta Trandafirilor (Rózsák tere) 61.

Tel: +40265-208170, +40365-403030, Fax: +40265-206211

E-mail: iclanzand@yahoo.com