Sveučilište u Zagrebu
Fakultet elektrotehnike i računarstva
Zavod za elektroniku, mikroelektroniku, računalne i inteligentne sustave

Decepcijski problemi

Stjepan Picek

Kolegij: Algoritmi za ugrađene računalne sustave

Nastavnici: doc. dr. sc. Leo Budin, doc.dr.sc. Marin Golub

2008./2009.

Zagreb, 11.09.2008.

1. Sadržaj

2. Uvod

Genetski algoritam (GA) je robusna adaptivna metoda koja je uspješno primjenjena na teške optimizacijske probleme, kako umjetne tako i one iz stvarnog svijeta. Unatoč tome GA griješe. Kada i zašto?

Pitanje što čini problem teškim za GA je dobilo puno pažnje. Ako želimo razumjeti koliko težak problem GA može riješiti, koliko brzo i s kojom pouzdanošću, moramo najprije odgovoriti na pitanje što je „teško“. Goldberg (1993.) predlaže nekoliko odrednica kvazi - odvojivih dimenzija teškoća problema GA. To su:

a) izolacija
b) navođenje na krivi trag
c) šum
d) multimodalnost
e) preslušavanje

Kombinacija izolacije rješenja i navođenja na krivi trag (neoptimalna rješenja) čine decepciju.
Decepcijski problem je optimizacijski problem koji navodi genetski algoritam na krivo, ne uvijek suboptimalno rješenje.

3. Teorijske osnove GA

Kako bismo mogli razumjeti decepciju i odrediti kada ona predstavlja problem, moramo dobro razumjeti teoriju genetskih algoritama: teorem sheme i hipotezu građevnih blokova.

3.1.Teorem sheme

Teorem sheme je poseban primjer analize forme. Prvi ga je izrazio Holland 1975. godine.

3.1.1. Shema i maska

Pretpostavimo da su genotipovi g u prostoru pretrage G genetskih algoritama nizovi definirane dužine l nad alfabetom $\sum$, tj. da vrijedi G = $\sum^l$. Uobičajeno, $\sum$ je binaran alfabet $\sum = \{0, 1\}$. Za genome (nizove) definirane dužine možemo smatrati vrijednost neke pozicije (na nekom *loci*-ju) kao svojstvo genotipa. Postoje dva osnovna principa za definiranje takvih svojstava: maske i „bilo što“ simboli.

Definicija 1: (Maska). Za genom definirane dužine G = $\sum^l$, definiramo skup svih maski genotipa M_l kao skup ispravnog *loci*- a $M_l = P(\{1,...,l\})$. Svaka maska $m_i \varepsilon M_l$ definira svojstvo φ_i i relaciju ekvivalencije:

$$g \approx \phi_i h \Leftrightarrow g[j] = h[j] \cdot \forall j \varepsilon m_i$$

gdje je red „red (m_i)“ maske m_i broj pozicija definiranih redom:

$$\text{red}(m_i) = |m_i|$$

Definirajuća dužina $\delta(m_i)$ maske m_i je udaljenost između prve i zadnje nepromjenjive (fiksne) znamenke u maski:

$$\delta(m_i) = \max\{|j-k| \; \forall j, k \; \varepsilon \; m_i\}$$

Maska sadrži indekse svih elemenata u nizu koji su zanimljivi u pogledu svojstva koje definira. Pretpostavimo bit niz dužine l = 3 kao genotip ($G=B^3$). Skup važećih maski M_3 je tada

$M_3 = \{\{1\} , \{2\} , \{3\} , \{1, 2\} , \{1, 3\} , \{2, 3\} , \{1, 2, 3\}\}$

Maska $m_1 = \{1, 2\}$, na primjer, određuje da vrijednost na poziciji 1 i 2 genotipa određuje vrijednost svojstava φ_1 te da je vrijednost bita na poziciji 3 nevažna. Stoga su definirane 4 forme:

$A\varphi 1=(0,0) = \{(0, 0, 0) , (0, 0, 1)\}$,
$A\varphi 1=(0,1) = \{(0, 1, 0) , (0, 1, 1)\}$,

Aφ1=(1,0) = {(1, 0, 0) , (1, 0, 0)},
Aφ1=(1,1) = {(1, 1, 0) , (1, 1, 1)}.

Definicija 2. (Shema). Shema je uzorak ili skup rješenja koja su međusobno slična. Sličnost se odnosi na jednakost gena na pojedinim mjestima kromosoma.

3.1.2. „bilo što" znakovi

Druga metoda određivanja shema je upotreba „bilo što" znakova kako bi se stvorili nacrti H članova sheme. Stavljamo simbol „ * " na sve nevažne pozicije i karakterizirajuće vrijednosti svojstva na ostala mjesta.

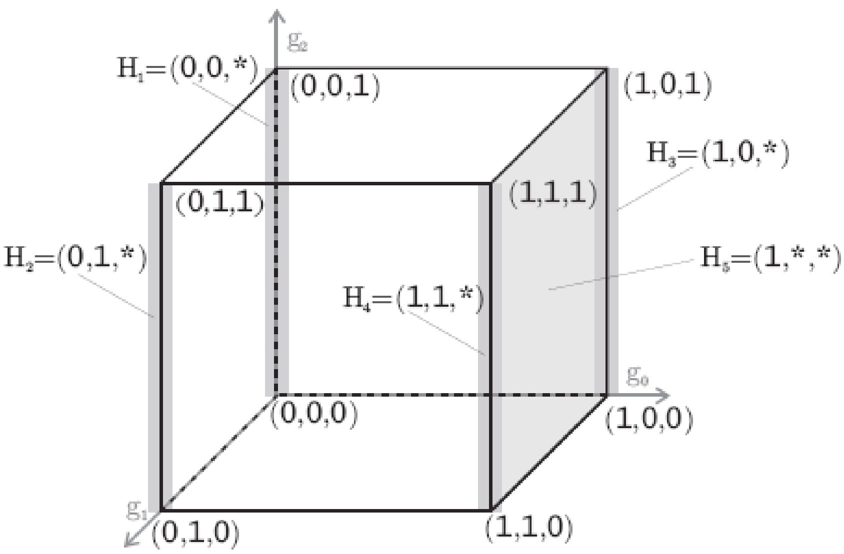
$$\forall \varepsilon 1...l \Rightarrow H[j] = \begin{cases} g[j] ako\ je\ j\ \varepsilon\ m_4 \\ * \, inace \end{cases}$$

$$\mathrm{H}[\mathrm{j}]\,\varepsilon \sum \cup \{*\} \forall j\, \varepsilon 1..\mathrm{l}$$

Sada možemo redefinirati navedenu shemu kao:

Aφ1=(0,0) ≡ H1 = (0, 0, *),
Aφ1=(0,1) ≡ H1 = (0, 1, *),
Aφ1=(1,0) ≡ H1 = (1, 0, *),
Aφ1=(1,1) ≡ H1 = (1, 1, *)

Ove sheme označavaju hiperravnine u prostoru pretrage G za genom od 3 bita, kako je prikazano na slici ispod:



Slika 1. Shema sa genomom dužine 3 bita

3.1.3. Hollandov teorem sheme

Teorem sheme je definirao Holland za genetske algoritme koji koriste dobrota - proporcionalnu selekciju i gdje je dobrota predmet optimizacije:

$$N(S,t+1) \geq N(S,\mathrm{t})\frac{\overline{\mathrm{D}_s}}{\overline{\mathrm{D}}}\left[1-\frac{\delta(S)}{n-1}p_c - o(S)p_m\right]$$

gdje su:

N(S,t)	– očekivani broj jedinki koje su podskup sheme S u generaciji t
$\overline{D_s}$	– prosječna dobrota sheme
$\overline{D}$	– prosječna dobrota populacije
$\delta(S)$	– definirana dužina sheme S
o(S)	– red sheme S
p_c	– vjerojatnost križanja
p_m	– vjerojatnost mutacije
n	– dužina kromosoma

Iz ove se formule može zaključiti da će genetski algoritam stvarati kratke, iznadprosječne sheme eksponencijalno rastući broj puta. To se dešava zato jer se one multipliciraju za određeni faktor u svakoj generaciji a samo će mali broj njih biti uništen reproduktivnim operacijama.

Teorem sheme je direktno primjenjiv za jedan generacijski ciklus, ali se i pomoću njega može dobiti osjećaj o dinamici GA razmatrajući što se događa hiperravnini koja konzistentno ima višu opaženu dobrotu od prosječne dobrote populacije. Taj slučaj se često opisuje kao: „Uobičajena interpretacija teorema sheme je da će podprostori s višom od prosječne isplativosti biti alocirani eksponencijalno više pokušaja u nekom vremenu, dok će podprostori s ispodprosječnom isplativošću biti alocirani eksponencijalno manje puta. To predpostavlja da ... efekt križanja i mutacije nije razarajuć." (Spears i De Jong, 1991.).

Strogo govoreći, riječ opaženo bi trebala biti stavljena ispred svakog spominjanja riječi isplativost, ali značenje je uobičajeno jasno i kraća fraza je često zgodnija za upotrebu.

3.1.4. Kritika teorema sheme

Kako bi teorem sheme bio važeći postavljeni su strogi preduvjeti i on vrijedi samo za jednostavni GA. Zaključak da će se dobra shema proširiti eksponencijalno je samo vrlo optimistična pretpostavka i generalno nije istinita. Ako vrlo dobra shema ima mnogo potomaka s velikom dobrotom, to će također popraviti ukupnu dobrotu populacije. Tako će se vjerojatnost u navedenoj formuli promijeniti tijekom vremena. Općenito govoreći, teorem sheme predstavlja donju granicu koja će vrijediti samo za jednu generaciju. Pokušaj predviđanja rezultata za više od jedne ili dvije generacije dovesti će do decepcije. Nadalje, populacija genetskog algoritma predstavlja samo uzorak ograničene veličine prostora pretrage G. To ograničava reprodukciju shema, ali i čini tvrdnje o vjerojatnosti kompliciranijima. Pošto imamo samo uzorke shema H ne možemo biti sigurni da prosječna vrijednost dobrote populacije u vremenu t stvarno predstavlja prosječnu dobrotu svih članova shema. Tako čak i operatori reprodukcije koji čuvaju integritet shema mogu voditi do smanjenja prosječne vrijednosti dobrote populacije tokom vremena.

Moguće je da su dijelovi populacije već konvergirali i da se drugi članovi sheme više neće istraživati, pa tako ne dobivamo daljnje informacije o stvarnoj korisnosti.

Također ne možemo znati da li je dobro da se pojedina shema širi jako brzo, čak i ako ima jako visoku dobrotu.

Postavlja se i pitanje da li su većina shema kompatibilne i mogu li se kombinirati, tj. da postoji niska razina interakcije između raznih gena što također generalno gledajući ne vrijedi.

3.2. Teorem građevnih blokova

Hipoteza građevnih blokova (BBH) predložena od strane Goldberga i Hollanda je temeljena na dvije pretpostavke:

- kada GA rješava problem, postoje neke sheme niskog reda i kratke dužine sa nadprosječnom dobrotom (takozvani građevni blokovi).
- te sheme su kombinirane u svakom koraku od strane GA kako bi stvorile bolje i duže nizove. Koristeći građevne blokove umjesto testirajući bilo koju moguću binarnu kombinaciju, GA efikasno smanjuje složenost problema.

Mada se čini da je hipoteza građevnih blokova podržana teoremom sheme, to se ne može lako potvrditi. Globalno govoreći, postoje mnoge kritike hipoteze građevnih blokova i ne može se smatrati da je ona dovoljno dokazana.

Grefenstette je u svojim radovima iz 1989. i 1991. izrazio sumnje u dotadašnje definicije decepcije smatrajući da su bile temeljene na pogrešnim pretpostavkama o dinamici GA.
Koristeći isto skraćivanje kao i u teoremu sheme (nema riječi opaženo ispred svakog spominjanja riječi isplativost), ukupna dinamika GA je često izražena kroz hipotezu građevnih blokova.
Genetski algoritam traži skoro optimalnu izvedbu kroz rekombinaciju kratkih shema koje su niskog reda i visoke izvedbe, ili građevnih blokova.

Upotreba hipoteze građevnih blokova koja je definirana na takav skraćeni način može dovesti do ozbiljnih zabluda o teoriji djelovanja GA.

Jedna takva verzija hipoteze građevnih blokova je takozvana Statička hipoteza građevnih blokova.

3.2.1. Statička hipoteza građevnih blokova

Statička hipoteza građevnih blokova (SBBH) glasi:
uzimajući bilo koju kratku particiju hiperravnine niskog reda, GA će očekivano skrenuti ka hiperravnini sa najboljom prosječnom statičnom dobrotom („očekivani pobjednik“).
Na primjer, uzmimo dio hiperravnine reda-2:

	$f(H)$
H_a: 00######	1
H_b: 01######	3
H_c: 10######	10
H_d: 11######	7

Vrijednost f(H) označava prosječnu statičnu dobrotu sheme H, drugim riječima, to je srednja vrijednost dobrote svake točke opisane shemom. Taj statičan prosjek je nezavisan o točkama koje se nalaze u populaciji u bilo kojem trenutku. Prema SBBH, „očekivani pobjednik" natjecanja shema u primjeru iznad je hiperravnina Hc.

SBBH implicira da funkcije za koje sheme nižeg reda povezane s optimumom imaju višu statičnu prosječnu dobrotu nego li suparničke sheme trebaju biti GA-lake. Na primjer, neka je globalni optimum 00...0,
i neka je

f(0#...#) > f(1#...#)
f(00#...#) > f(01#...#)
f(00#...#) > f(10#...#)
f(00#...#) > f(11#...#)

i tako dalje za svaku hiperravninu u prostoru pretrage. Prema SBBH, za GA bi funkcija f trebala biti lagana za optimiranje. U stvari, takve funkcije su uobičajeno zvane GA-lake (Wilson, 1991.). Nadalje, SBBH implicira da bi funkcije koje imaju sheme nižeg reda povezane s optimumom koji ima nižu prosječnu statičku dobrotu nego suparničke sheme trebale biti teške za GA. Na primjer, neka je globalni optimum 00...0, i neka je

f(0#...#) < f(1#...#)
f(00#...#) < f(01#...#)
f(00#...#) < f(10#...#)
f(00#...#) < f(11#...#).

Tada bismo takvu funkciju zvali decepcijskom.

Čini se da SBBH potkopava velik dio radova o teoriji GA, posebno o decepciji. SBBH je sam rijetko citiran kao pretpostavka u literaturi o decepciji. Umjesto toga, govori se o teoremu sheme, kao i o neformalnoj hipotezi građevnih blokova. Na primjer,
„Konstruirajmo najjednostavniji problem koji će uzrokovati skretanje GA od globalnog optimuma...Kako bi to napravili, želimo ekstremno narušiti hipotezu građevnih blokova" (Goldberg, uvodni pojam MDP, 1989.).
ili
„Pokazano je da genetski algoritmi rade dobro kada se građevni blokovi, kratki, niskog reda ... sheme sa nadprosječnom vrijednošću dobrote udružuju kako bi formirali optimalno ili skoro optimalno rješenje" (Homaifar i suradnici, 1991.).

Mada je SBBH dobro intuitivno objašnjenje kako GA radi, nikad nije eksplicitno dokazan, i ne slijedi iz teorema sheme. Teorem sheme opisuje očekivani rast hiperravnine za jednu generaciju na temelju opažene prosječne dobrote, tj. na temelju prosječne dobrote trenutnih uzoraka hiperavnine u populaciji. Nakon nekog broja generacija, opažena prosječna dobrota hiperavnine ne mora nužno odražavati prosječnu statičnu dobrotu hiperravnine. SBBH nastaje kada ignoriramo ključnu razliku između opažene prosječne dobrote i statične prosječne dobrote. Točnije, SBBH dobivamo ako izbacimo vrijeme (t) iz izraza f(H, t)/f(t) u teoremu sheme. Tada SBBH predstavlja samo aproksimaciju dinamike opisane teoremom sheme.

Postoji mnogo načina na koje bi bilo očekivano da takva aproksimacija skreće od dinamike opisane u punom teoremu sheme. Navedena su dva načina kada SBBH ne opisuje točno dinamiku GA.

1. Neuspjeh uračunavanja ukupne konvergencije
2. Neuspjeh uračunavanja dobrote variance unutar shema.

Efekt ukupne konvergencije je taj da jednom kada populacija počne konvergirati, mada i malo, nije više moguće procijeniti statičnu prosječnu dobrotu shema koristeći podatke prisutne u trenutačnoj populaciji.

Efekt dobrote variance unutar sheme je taj da u populacijama realne veličine, opažena dobrota sheme može biti razmjerno daleko od statične prosječne dobrote, čak i u početnoj populaciji.

4. Decepcija

4.1. Uvod

Problem je decepcijski ako stanovite hiperravnine vode potragu prema nekim rješenjima genetskih građevnih blokova koji nisu globalno kompetativni. Najraniji radovi u ovom području su se koncentrirali na konstrukciju teških ili decepcijskih problema (Goldberg). U zadnje vrijeme se taj posao proširio uzimajući u obzir odnos između decepcije varijance hiperravnina, dominancije hiperravnina i genetskih algoritama kao dinamičkih sustava (Grefenstette). Hiperravnine se pokazuju kao kombinacije bitova poznatih kao sheme. Razmotrimo red-3 problem koji uključuje prostor potrage prikazan sa tri bita. Ovaj jednostavan prostor može biti predstavljen kao kocka i označen tako da se susjedni vrhovi u kocki razlikuju samo u jednom bitu. Koristeći * kao „bilo što“ simbol, postavimo da shema 0** predstavlja lice ove kocke. 0** je shema reda-1, s obzirom da definira dio prostora pretrage koji sadrži sve nizove sa jednim bitom (0) na naznačenoj poziciji. Vrijednost hiperravnine (ili u ovom slučaju obične ravnine) prikazane kao 0** se obilježava sa f(0**) i izračunava se kao prosječna vrijednost dobrote svih nizova koji odgovaraju predlošku 0** (prosječna vrijednost 4 niza koji oformljuju odgovarajuću stranicu kocke). Red hiperravnine se odnosi na broj 0 ili 1 bitova koji odgovaraju shemi koja predstavlja tu hiperravninu. Red također daje informacije o broju nizova sadržanih unutar hiperravnine koja je prikazana: hiperravnina reda 1 sadrži 50 % nizova prostora pretrage. Tako je shema *** 1****0*...* reda-2 i odgovarajuća hiperravnina koja sadrži (tj. njena shema odgovara) 25% svih nizova u prostoru pretrage.

U slučaju potpuno decepcijske red-3 funkcije hiperravnina, informacija prikazana sa red-1 i red-2 shemama u prostoru pretrage vodi pretragu od globalnog optimuma prema onome što ćemo zvati decepcijski atraktor. Pretpostavlja se da bitovi 111 predstavljaju globalni optimum i da je decepcijski atraktor 000, te potpuna red-3 decepcija implicira da se slijedeće veze drže nad shemama niskog reda:

f(0**) > f(1**)	f(00*) > f(11*), f(01*) , f(10*)
f(*0*) > f(*1*)	f(0*0) > f(1*1), f(0*1) , f(1*0)
f(**0) > f(**1)	f(*00) > f(*11), f(*01) , f(*10)

gdje je f(S) dobrota ili niza ili sheme S, te se računa uprosječujući sve nizove sadržane u hiperravnini predstavljene tom shemom. Drugim riječima, sve informacije o hiperravninama nižeg reda povezane sa shemama čiji se prikaz sastoji od nekih podnizova tih triju bitova vodi pretragu od globalnog optimuma 111. Potpuni decepcijski red-3 problem koji zadovoljava gore navedene nejednakosti je definiran od strane Goldberga, Korba i Deba (1989.) gdje nizovi bitova imaju slijedeće vrijednosti.

Decepcijska funkcija 1:

f(000) = 28	f(001) = 26
f(010) = 22	f(100) = 14
f(110) = 0	f(011) = 0
f(101) = 0	f(111) = 30

Dio motivacije za izgradnju decepcijskih problema je bolje razumijevanje koje vrste uvjeta mogu spriječiti genetsku pretragu od nalaženja globalno optimalnog ili blizu optimalnog rješenja. Problemi koji imaju sposobnost ozbiljnog zavaravanja genetske pretrage su nazvani GA teški problemi. Potpuno decepcijski podproblemi u paru sa lošom povezanošću mogu rezultirati GA teškim problemima.

4.2. Pojmovi u teoriji decepcije

U pokušaju objašnjavanja prirode decepcije potrebno je razlikovati različite vrste decepcijskih situacija i koristiti izraze s nekom preciznošću.

Prvo ćemo razmatrati natjecanja hiperravnina u kojima su neka natjecanja posebno važna. Primarno natjecanje hiperravnina reda N sadrži potpuni skup primarnih natjecatelja gdje je primarni natjecatelj u hiperravnini reda N skup od 2^N hiperravnina koje sadržavaju shemu sa N bit vrijednosti na istim mjestima, npr. *0***0, *0**1, *1**0 i *1**1 su natjecatelji u primarnom natjecanju hiperravnina reda-2.

Globalni pobjednik primarnog natjecanja hiperravnina je ona hiperravnina koja sadrži najveću vrijednost dobrote u skupu natjecatelja, pri čemu je vrijednost dobrote te hiperravnine prosječna dobrota svih nizova koji su sadržani u toj hiperravnini. Ovo ne znači nužno da ta hiperravnina točno vodi ka ili ima bitove u skladu s globalno optimalnim rješenjem.

Još su dva osnovna koncepta korisna kod rasprave o decepciji. Hiperravnina X sadrži hiperravninu Y kada je X hiperravnina nižeg reda i Y hiperravnina višeg reda takva da Y sadrži točno iste bit vrijednosti na točno istim pozicijama kao X, ali također ima dodatne bit vrijednosti na pozicijama koje sadrže „bilo što“ (*) simbol u X-u. Npr. 0** sadrži 01*; primjećujemo da 0** također sadrži iste nizove kao 01* ali sadrži i druge nizove. To vodi jasnije određenoj vezi između natjecanja hiperravnina koje je važno za decepciju.

Dva primarna natjecanja hiperravnina reda N i reda K (gdje je K< N) su „u odnosu“ jedan prema drugome kada kolektivno natjecatelji hiperravnine reda K sadrže natjecatelje hiperravnine reda N. Tako primarno natjecanje hiperravnine između 11***, 10***, 01*** i 00*** sadrži skup hiperravnina koje kolektivno sadrže primarno natjecanje hiperravnina između 111**, 110**, 101**, 100**, 011**, 010**, 001** i 000** .Gledajući natjecanja hiperravnine reda N „važne“ hiperravnine reda K (gdje je K < N) su predstavljene shemama koje imaju bit vrijednosti (ili 1 ili 0) na ispravnim podskupovima lokacija na kojima se bitovi pojavljuju u natjecanju hiperravnine reda N.

Decepcija implicira da globalni pobjednik nekog natjecanja hiperravnina reda N ima uzorak bitova koji je različit od uzorka bitova globalnog pobjednika nekog relevantnog natjecanja hiperravnina nižeg reda. Drugim riječima, decepcija implicira da ta natjecanja nižeg reda imaju vlastite globalne pobjednike koji nemaju iste bit vrijednosti kao globalni pobjednik relevantnog natjecanja hiperravnina reda N.

Decepcijski problem je bilo koji problem reda N koji sadrži decepciju u jednom ili više relevantnih natjecanja hiperravnina nižeg reda. Decepcija implicira da postoji jedno ili više natjecanje hiperravnina nižeg reda koje je relevantno i potencijalno može odvesti genetsku pretragu od globalnog pobjednika natjecanja hiperravnina reda N. To ne implicira da je problem potpuno decepcijski, ili da postoji dovoljno decepcije da zavara GA. Ta upotreba je u skladu sa Goldbergovom raspravom na temu minimalnog decepcijskog problema, koji sadrži decepciju ali nije potpuno decepcijski.

Navedena upotreba je dosta široka: većina problema može sadržavati neki stupanj decepcije, s obzirom da općenito ne bismo očekivali da sva natjecanja hiperravnina nižeg reda budu u skladu sa relevantnim natjecanjima hiperravnina višeg reda. Ali, ova definicija razdvaja optimizacijske zadatke u dvije grupe, oni koji nisu decepcijski (i tako su teoretski laki) i oni koji su decepcijski.

Potpuno decepcijski problem (ili podproblem) reda N je decepcijski kada sve relevantne hiperravnine nižeg reda vode ka decepcijskom atraktoru. Decepcijski atraktor je hiperravnina reda N, različita od pravog globalnog pobjednika koji je podržan sa relevantnim natjecanjima hiperrravnina nižeg nivoa. Za potpuno decepcijski problem može se pokazati da decepcijski atraktor može biti samo hiperravnina prikazana shemom s uzorkom bita koji je komplementaran onome kod globalnog pobjednika natjecanja hiperravnine reda N (ako je globalni pobjednik natjecanja hiperravnina reda-3 **1**1*1**, tada je komplement i potencijalni decepcijski atraktor **0**0*0**). Ispostavlja se da je korisno razlikovati različite stupnjeve decepcije u skladu s drukčijim kriterijem od reda kada se opisuje i definira decepcijski atraktor.

Konzistentno decepcijski problem (ili podproblem) reda N je onaj u kojem niti jedna od relevantnih hiperravnina nižeg reda ne vodi ka globalnom pobjedniku primarnog natjecanja hiperravnina reda N za koje smo zainteresirani, ali sve relevantne hiperravnine nižeg reda ne vode nužno ka decepcijskom atraktoru - osim hiperravnine reda-1 koja može voditi pretragu samo ka globalnom pobjedniku ili decepcijskom atraktoru natjecanja hiperravnina reda N. Može se pokazati da je izbor decepcijskog atraktora kao komplementa globalnog pobjednika natjecanja reda N također nužan uvjet za postojanje konzistentno decepcijskog problema zbog hiperravnina reda-1 koje su sadržane. Važno je napomenuti da je potpuno decepcijski problem uvijek konzistentno decepcijski problem, ali obrnuto ne vrijedi.

Pojam decepcijska funkcija se odnosi na konzistentno decepcijski problem gdje broj bitova korištenih za prikaz prostora rješenja odgovara redu decepcije. Decepcijska funkcija je uvijek konzistentno decepcijska i može biti potpuno decepcijska. Tako konzistentno decepcijski problem reda-3 sa prikazom dužine 3 bita tvori decepcijsku funkciju.

Decepcijski građevni blok reda N se odnosi na situaciju gdje neke hiperravnine reda H imaju veću vrijednost dobrote nego njeni primarni natjecatelji, ali su sva relevantna natjecanja hiperravnina nižeg reda između shema sastavljenih od podskupa bitova na istim lokacijama decepcijska. Decepcijski građevni blok je sličan konceptu decepcijske funkcije osim što decepcijski građevni blok može odgovarati određenom natjecanju hiperravnine koji je dio većeg problema optimizacije funkcije. Tako je decepcijski građevni blok u nekom smislu podproblem neke veće funkcije. Koncept građevnog bloka se koristi u GA kada se govori o shemama niskog reda (i često karakteriziranih kratkom dužinom), te dobrotom iznad prosjeka. Smatra se da su građevni blokovi važni u procesu genetske pretrage. Koncentrirajući se na natjecanja hiperravnina koje sadrže građevne blokove (što treba biti dobro predstavljeno u populaciji rješenja i treba biti relativno stabilno pod rekombinacijom) GA je u stanju stvoriti dobra djelomična rješenja i, uz pomoć rekombinacije, sagraditi sheme višeg reda i potpune nizove. Empirijski rezultati pokazuju da je narušena sposobnost GA da dovede do globalnog rješenja, ako nema građevnih blokova na raspolaganju. Decepcijski građevni blok je u nekom smislu „anti - građevni - blok“ s obzirom da predstavlja natjecanje hiperravnine koje može zavarati GA. Decepcijski građevni blok nema nužno kratku dužinu; u stvari decepcijski građevni blok velike dužine je situacija koja čini decepciju još problematičnijom. Pritom, decepcijski građevni blok sadrži natjecanja između hiperravnina sa nadprosječnom dobrotom.

Decepcijska hiperravnina je hiperravnina koja netočno vodi pretragu od nekih hiperravnina višeg reda koje su u biti superiorne svojim suparnicima. Slijedi da su hiperravnine

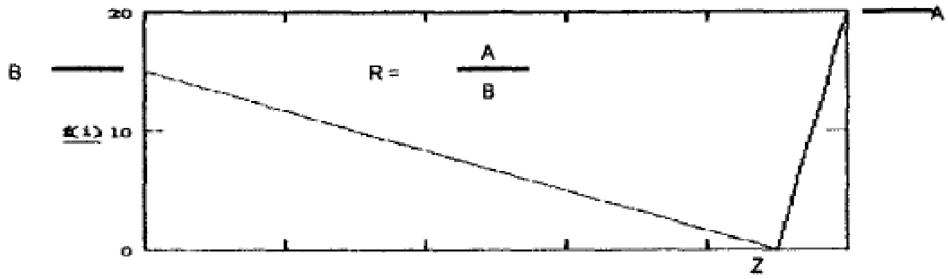
nižeg reda koje stvaraju decepcijske građevne blokove „decepcijske hiperravnine“ u odnosu na taj određeni građevni blok.
Pojam decepcijske hiperravnine je definiran u odnosu na određeni decepcijski građevni blok jer bitovi koji sačinjavaju globalnog pobjednika natjecanja hiperravnine u decepcijskom građevnom bloku ne moraju biti isti bitovi koji se pojavljuju u drugom natjecanju hiperravnine ili u nizu koji odgovara globalnom rješenju problema. Npr. u potpunoj decepcijskoj funkciji reda-4, hiperravnine reda-1 su decepcijske hiperravnine u odnosu na građevni blok reda-4 koji tvori globalno rješenje, ali hiperravnine reda-1 nisu decepcijske u odnosu na bilo koje natjecanje reda-2 ili reda-3 s obzirom da sva ta natjecanja konzistentno vode ka decepcijskom atraktoru. Kada bi se definirala „decepcijska hiperravnina“ u odnosu na samo globalno rješenje to bi bilo previše restriktivno s obzirom da bi to ignoriralo sile koje vode genetsku pretragu. Ali definicija upotrijebljena ovdje nas ne priječi u odnošenju ka hiperravnini kao decepcijskoj u odnosu na globalno rješenje.

Krajolik dobrote predstavlja moćnu metaforu u globalnoj optimizaciji. On predstavlja vezu između pojedinaca populacije i njihove dobrote.

Decepcija postaje praktičniji alat za GA teoriju kada je uključena u krajolik dobrote kao parcijalna decepcija. Trenutačno, parcijalna decepcija se može definirati samo kao ona koja je manje decepcijska od potpune decepcije i više decepcijska od GA lakog problema. Pokušaj sortiranja parcijalnih decepcijskih funkcija prema nekoj skalarnoj ljestvici decepcije je problematično. Goldberg je prepoznao potrebu za generalizacijom potpune decepcije te je definirao potpunu decepciju reda k. Homafair, Qi i Frost su nazvali ograničenu decepciju reduciranom decepcijom reda k. Mada je poznata i pod drugim imenima ova vrsta decepcije je najpoznatija pod imenom ograničena decepcija. Problem ima ograničenu decepciju reda k ako i samo ako su sve particije reda (k – 1) ili manje decepcijske. Particije reda => k mogu i ne moraju biti decepcijske. Tako je za k = l ograničena decepcija ujedno i potpuna decepcija. Goldberg, Deb i Korb (1991.) su konstruirali primjere problema sa ograničenom decepcijom na način da su spojili neki broj m potpuno decepcijskih podfunkcija reda l_s kako bi dobili ($m * l_s$) - bitnu funkciju ograničene decepcije reda l_s. Takve funkcije imaju 2^m lokalnih optimuma, od kojih je jedan globalni.

U definiranju parcijalne decepcije se mora biti oprezan. Lako je oslabiti zahtjeve za potpunu decepciju tako da GA može lako naći globalni optimum neke funkcije koja ispunjava te zahtjeve. Parcijalno decepcijski problemi se mogu podijeliti prema primarnim natjecanjima shema na red-k kronološki decepcijski problem i red-k osnovno decepcijski problem.

Decepcijske funkcije zamke su linearne funkcije broja bitova l u ulazu te su korištene kao decepcijske funkcije za testiranje GA. Deb i Goldberg (1991.) su napravili detaljnu analizu klasa svih funkcija zamki čije uvjete izvode iz uvjeta za potpunu decepciju. Različiti stupnjevi decepcije se pojavljuju kao funkcija dva parametra koji opisuju funkciju zamke i ukupni broj bitova ulaza L. Ti parametri su omjer globalnog prema najvećem lokalnom optimumu, R = A/B, i pozicija globalnog minimuma Z kao što je prikazano na slici 2.



Slika 2. Parametri funkcije zamke R i Z potpuno određuju funkciju dužine L

U R, Z prostoru za datu veličinu L, funkcije zamke su potpuno definirane. Dio funkcija koje su potpuno decepcijske, u ograničenju velikog L, dali su Deb i Goldberg:

$$d(L) = \frac{\ln(\mathrm{L}/2)}{2 \cdot \mathrm{L}}, \text{za L} >> 1$$

4.3. Samo zahtjevni problemi su decepcijski

Kao što je rečeno ranije, većina problema sadrži neki stupanj decepcije pošto se općenito ne može očekivati da će sva natjecanja hiperravnina nižeg reda biti konzistentna sa relevantnim natjecanjima hiperravnina viših redova. Slijedeći teorem podržava tvrdnju da su samo izazovni optimizacijski zadaci problemi koji sadrže neki stupanj decepcije:

Teorem 1:

Imajući funkciju dobrote za problem koji prikazuje neki optimizacijski zadatak sa binarnim prikazom dužine L, ako

1) se ne javi decepcija niti na jednoj hiperravnini povezanoj s tim binarnim prikazom

i

2) pobjednici L hiperravnina reda 1 mogu biti ispravno određeni

tada je globalni optimum funkcije određen jednim nizom sadržanim u presjeku L pobjednika natjecanja hiperravnina reda 1.

Dokaz kontradikcijom:

Pretpostavimo da se decepcija ne pojavljuje niti na jednoj hiperravnini funkcije dobrote i određujemo ispravne pobjednike svake hiperravnine reda-1. Presjek shema reda-1 će identificirati točno jedan niz kao kandidata za globalni optimum. Ako kandidat nije globalni optimum tada je barem jedna hiperravnina reda-1 decepcijska , što stvara kontradikciju.

QED

Dok ovaj teorem ima neke idealističke pretpostavke može biti operacionaliziran. Ako se ne pojavljuje decepcija tada možemo riješiti problem na slijedeći način. Za svaku bit lokaciju možemo direktno riješiti dvo-ručni bandit problem koji sadrži dvije sheme *...*0*...* i *...*1*...*. Alternativno natjecanje može biti riješeno upotrebom statističkih metoda. Upotreba statističkih metoda potencijalno osigurava rješenje problema šuma i time problema određivanja prikladnih pobjednika natjecanja hiperravnina. Ako nema decepcije takve metode će odrediti (do nekog stupnja pouzdanosti) vjerojatnu vrijednost bitova koji pripadaju odgovarajućim pozicijama u binarnom prikazu niza globalnog rješenja problema. Stanovite teškoće se mogu pojaviti: što ako dvije hiperravnine reda-1 koje se natječu imaju istu ili skoro istu vrijednost? Treba primjetiti da ako nema decepcije, sve podparticije hiperravnine reda-1 moraju voditi ka globalnom rješenju (suprotno bi impliciralo decepciju). Tako, ako možemo riješiti bilo koju hiperravninu reda-1, možemo iskoristiti to rješenje da suzimo prostor pretrage. Ako nema decepcije nema niti veze u kojem redoslijedu su hiperravnine reda-1 riješene. To znači da možemo razbiti veze i razriješiti natjecanja ako pobjednici natjecateljskih hiperravnina reda-1 mogu biti izabrani u bilo kojoj particiji reduciranog prostora pretrage. Mada možda ponekad neće biti moguće iskoristiti ovu ideju, često je to moguće.

Tako se osnovna teorija decepcije čini skoro trivijalna. Istraživači koji rade sa GA već dugo znaju da hiperravnine na brojnim razinama igraju važnu ulogu u genetskoj pretrazi. Ipak, ovo znanje nije bilo direktno povezano sa decepcijom i time što to implicira o Hollandovoj analogiji o dvo-ručnom ili k-ručnom problemu bandita. Analogija za dvo-ručnog bandita za hiperravnine reda-1 ide na slijedeći način: pretpostavimo da imamo aparat za igre na sreću sa dvije ruke. Svaka ruka ima drugu isplativost sa povezanom srednjom varijancom. Problem je odrediti koja ruka ima veću očekivanu isplativost. Želimo ispitati dvije ruke, ali u isto vrijeme minimalizirati očekivani gubitak. Da bismo to napravili moramo imati eksponencijalno više pokušaja za ruku s boljim rezultatom (opaženim). U stvarnosti to može uključivati početno ispitivanje da ustanovimo koja je ruka bolja. Ako radimo sa natjecanjima hiperravnina reda-1 tada *1*** i *0*** predstavljaju te dvije ruke. Ako radimo sa hiperravninama većim od reda-1 tada to postaje k-ručno natjecanje. Holland je uspio pokazati da genetski algoritam potencijalno može alocirati pokušaje ka hiperravninama na skoro optimalan način, pridružujući eksponencijalno više pokušaja ka opaženoj boljoj u višestrukom natjecanju hiperravnina. Činjenica da se to dešava simultano za mnogo različitih natjecanja hiperravnina različitih redova se naziva implicitni paralelizam.

Decepcija i analogija dvo-ručnog bandita su direktno povezani s pitanjima koja su postavili Grefenstette i Baker o važnosti implicitnog paralelizma u genetskoj pretrazi. Oni su naznačili da genetski algoritam ne rješava dvo-ručno (ili k-ručno) natjecanje bandita među hiperravninama zbog činjenice da „genetski algoritam ne provodi uniformno uzorkovanje među hiperravninama u populaciji, barem ne nakon početne generacije“ (Grefenstette i Baker). Kao primjer oni predlažu slijedeće pridruživanje vrijednosti dobrote za neku optimizacijsku funkciju:

$$f\,(\mathrm{x}) = \begin{cases} 2 & \text{ako je x } \varepsilon\, 111*\ldots* \\ 1 & \text{ako je x } \varepsilon\, 0***\ldots* \\ 0 & \text{inace} \end{cases}$$

Uniformni uzorak niza pokazuje da f (0*...*) > f (1*...*). Ipak, kako genetska pretraga napreduje, shema 111*...* dominira nad natjecateljima hiperravnina i tako će 1*...* predstavljati veći udio nizova u populaciji nego 0*...* zato jer populacija više ne uzima uniformne uzorke iz prostora svih nizova.

Grefenstette i Baker smatraju da to predstavlja grešku u analogiji dvo-ručnog bandita jer se natjecanje između 0*...* i 1*...* ne rješava točno. Također, oni napominju da GA implicitno ne rješava sva moguća natjecanja hiperravnina točno, što pokazuje da implicitni paralelizam ne djeluje sveprisutno. Ali Holland nije nikad ni pretpostavio da se u implicitnom paralelizmu sva

natjecanja simultano rješavaju pomoću genetske pretrage nego samo da se mnoga natjecanja događaju simultano.
Nakon što smo ustanovili da su samo izazovni problemi decepcijski problemi postoji dodatni razlog za oprez kada se ocrtava priroda implicitnog paralelizma. Grefenstette i Baker ističu ograničenje implicitnog paralelizma i njihove ozbiljne implikacije. Da vidimo zašto je to točno prvo napomenimo da je Grafenstetteov i Bakerov problem decepcijski problem. Hiperravnina 0*...* je decepcijska u odnosu na natjecanje hiperravnine koje uključuje 111*...*. To nije potpuno decepcijski problem (ili podproblem): f(*1*...*) > f (*0*...*) i f (**1*...*) > f(**0*...*). Nadalje, problem nije težak s obzirom da Grefenstette i Baker ispravno napominju da će GA brzo konvergirati ka nizovima koji se nalaze u hiperravnini predstavljenoj sa 111*...*.
Pojavljivanje decepcijskog problema će uvijek stvoriti barem dva primarna natjecanja hiperravnina (što je analogno dvama različitima k-bandit problemima), koja imaju rješenja s uključenim različitim uzorcima bitova. Ako GA efektivno i konzistentno alocira eksponencijalno više reproduktivnih pokušaja ka pravom globalnom pobjedniku u jednom od ovih dvaju natjecanja hiperravnina, takve da se shema koja dobija eksponencijalno više pokušaja postaje sveprisutna u populaciji, tada će GA neuspješno riješiti alternativno natjecanje hiperravnina. GA ne može nastaviti alocirati eksponencijalno više pokušaja prema opaženoj boljoj (pretpostavljajući da je opaženi bolji u stvari globalni pobjednik) u obim natjecanjima hiperravnina, s obzirom da su dva globalna pobjednika nekompatibilna i samo jedan može postati sveprisutan u populaciji.

Konjunktura 1:

Pretpostavljajući da šum ne kvari procjenjivanja hiperravnina, dvo-ručna (ili k-ručna) analogija bandita ne uspijeva se primjeniti na pojedinom natjecanju hiperravnina onda i samo onda ako je decepcija uključena: GA može točno riješiti samo jedno od dva natjecanja koja sadržavaju skup natjecatelja hiperravnine reda K i nekih relevantnih natjecanja hiperravnina reda N, gdje je K < N i gdje je globalni pobjednik natjecanja reda K decepcijski u odnosu na globalnog pobjednika natjecanja reda N.
Ova konjunktura se temelji na slijedećim opažanjima. Ako se ne pojavi decepcija u zadatku optimiziranja funkcije tada implicitni paralelizam ima priliku točno riješiti sva natjecanja hiperravnina i analogija dvo-ručnog bandita se potencijalno održava. S druge strane, decepcija će uvijek stvoriti sukobljena natjecanja hiperravnina čiji globalni pobjednici imaju konfliktne uzorke bitova. Moguće je da se sukobljeni pobjednici predstave u populaciji na takav način da mogu koegzistirati. Takav raspored se čini da će biti nekonzistentan sa kritičnim dijelom analogije dvo-ručnog bandita koja pokazuje da opaženi bolji treba biti alociran eksponencijalno veći broj reproduktivnih pokušaja. Čini se vjerojatno kako bi bilo razumno očekivati rješavanje svih natjecanja hiperravnina ako ne postoji decepcija i tada bi takav problem teoretski mogao biti riješen bez GA. Uznemiravajuće je da ista decepcija koja čini globalnu optimizaciju netrivijalnom također potkopava implicitni paralelizam na točno istim natjecanjima hiperravnina koje su najkritičnije. Utjecaj decepcije na kritična natjecanja hiperravnina udara u samo srce teoretskih osnova GA.
Empirijski dokazi nalažu da rezultati sukobljenih natjecanja mogu ići u oba smjera. Ako shema višeg reda koja predstavlja globalnog pobjednika u potpuno decepcijskom građevnom bloku ima veliku dužinu, hiperravnine nižeg reda će pobijediti u njihovim k-ručnim bandit natjecanjima a hiperravnine višeg reda neće. S druge strane, ako shema višeg reda koja predstavlja globalnog pobjednika potpuno decepcijskog građevnog bloka ima čvrstu povezanost i ako je dobro

predstavljena u populaciji, globalni pobjednik natjecanja hiperravnine višeg reda će pobijediti u svojem natjecanju hiperravnine, a decepcijska hiperravnina nižeg reda neće.

Ova opažanja vode ka više ograničenom pogledu na implicitni paralelizam. Mnoga natjecanja hiperravnina se rješavaju simultano tako da je pretraga usmjerena ka globalnom rješenju, ali to ne mora implicirati da su sva natjecanja hiperravnina ispravno riješena pomoću GA. Tako, mogli bismo očekivati da stanovit stupanj decepcije nije dovoljan da zavara GA ako je većina natjecanja hiperravnina razrješena tako da vodi ka globalnom rješenju. Međutim, mnogo bolje razumijevanje implicitnog paralelizma je potrebno kako bi se teoretski usidrilo GA.

4.4. Konstrukcija potpuno decepcijske funkcije

4.4.1.Konstrukcija potpuno decepcijske funkcije na temelju rada Liepinsa i Vosea

Liepins i Vose (1990.) su pokazali da sve potpuno decepcijske funkcije mogu biti prikazane kao suma potpuno laganih funkcija i funkcije šiljka u optimalnoj točki. Nadalje, konstruirali su potpuno decepcijske funkcije reda d > 2 (za nizove čija je dužina n > 2).

$$f(x) = \begin{cases} 1 - 2^{-n} & \text{ako je } o(x) = 0 \\ 1 - (1 + o(z))/n & \text{ako je } 0 < o(x) < n \\ 1 & \text{ako je } o(x) = n \end{cases}$$

Liepins i Vose su dokazali da se ova klasa potpuno decepcijskih funkcija može transformirati u potpuno lagane funkcije pomoću transformacije g(k) = f ○ M(k), gdje je M reverzibilno linearno preraspoređivanje nad skupom Z_2. Smatrajući binarne nizove kao stupac vektore, njihova n x n matrica M glasi

$$m_{ij} = \begin{cases} 0 \text{ za } i = j \neq n \\ 1 \text{ inace} \end{cases}$$

4.4.2. Konstrukcija potpuno decepcijskih funkcija pomoću Walshove transformacije

4.4.2.1. Walsh – shema analiza

Bethke je razvio Walsh-shema transformaciju, u kojoj je diskretna verzija Walshove funkcije upotrijebljena za izračun prosječne dobrote sheme. Tada je upotrijebio ovu transformaciju kako bi opisao funkciju kao laku ili tešku za optimizaciju pomoću GA.

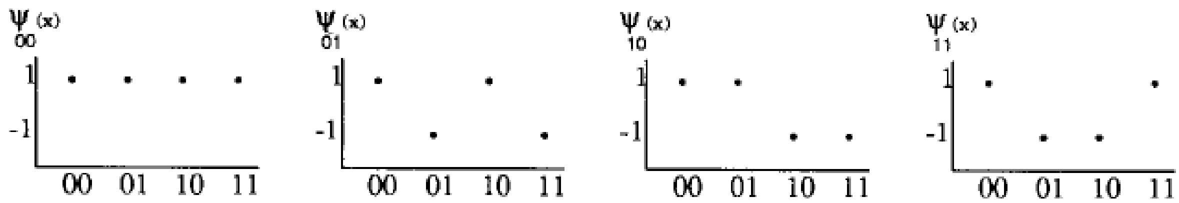
4.4.2.2. Walshove funkcije, Walshova dekompozicija i Walshovi polinomi

Walshove funkcije su kompletan ortogonalan skup baznih funkcija koje uvode transformacije slične Fourierovim transformacijama. Međutim, Walshove funkcije se razlikuju od ostalih baza (npr. trigonometrijske funkcije) po tome što imaju samo dvije vrijednosti, +1 i -1. Bethke je pokazao kako upotrijebiti te bazne funkcije kako bi se konstruiralo funkcije sa različitim stupnjem težine za GA. Kako bi to napravio, Bethke je upotrijebio diskretnu verziju Walshove originalno kontinuirane funkcije. Ove funkcije tvore ortogonalnu bazu za realne vrijednosti funkcija definiranih na $\{0, 1\}^1$.

Diskretne Walshove funkcije preslikavaju bit nizove na {1, -1}. Svaka Walshova funkcija je povezana sa pojedinim dijelom prostora pretrage. Walshova funkcija odgovarajuća j dijelu (gdje indeks j označava bit niz) je definirana na slijedeći način:

$$\Psi_j(x) = \begin{cases} 1 & \text{ako } x \wedge j \text{ imaju paran paritet} \\ -1 & \text{inace} \end{cases}$$

Znak $\wedge$ predstavlja logičku operaciju AND. Treba primjetiti da $\psi_j(x)$ ima svojstvo da su samo bitovi u x koji doprinose njegovoj vrijednosti oni koji odgovaraju vrijednostima 1 u j.

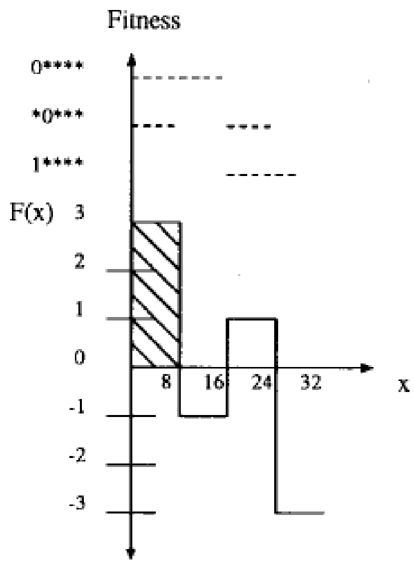


Slika 3. Prikaz 4 Walshove funkcije definirane nad 2 bita.

Pošto Walshove funkcije tvore bazični skup, bilo koja funkcija F(x) definirana na $\{0, 1\}^l$ može biti zapisana kao linearna kombinacija Walshovih funkcija.

$$F(x) = \sum_{j=0}^{2^l-1} \omega_j \psi_j(x)$$

gdje je x bitovni niz, l njegova dužina, i svaki ω_j je realan koeficijent zvan Walshov koeficijent. Na primjer, funkcija prikazana slikom



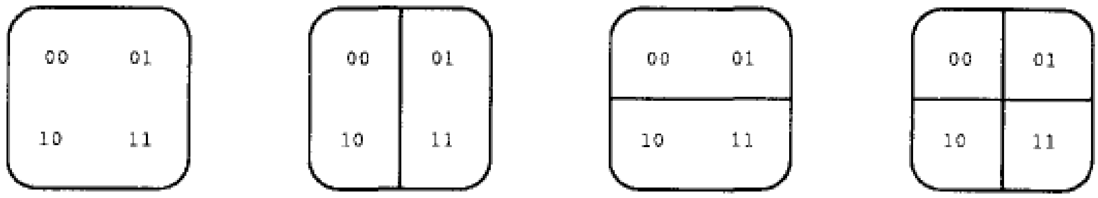
Slika 4. Prikaz funkcije dobrote

može biti zapisana kao

$$F(x) = 2\psi_{01000}(x) + \psi_{10000}(x)$$

Walshovi koeficijenti ω_j dane funkcije F mogu biti dobiveni pomoću Walshove transformacije, koja je slična Fourierovoj transformaciji. Prikaz funkcije F u svojstvima skupa Walshovih koeficijenata ω_j je zvan F Walshova dekompozicija ili Walshovi polinomi.

Kao jednostavan primjer Walshove transformacije, uzmimo funkciju F(x) = X^2, gdje je x niz od 2 bita. Prostor niza od 2 bita može biti razdvojen u skupove shema na 4 različita načina kao što je prikazano na slijedećoj slici:



A. Partition = **, j = 00 B. Partition = *d, j = 01 C. Partition = d*, j = 10 D. Partition = dd, j = 11

Slika 5. Skupovi shema za prostor od 2 bita

Walshova transformacija radi na način da transformira F(x) u sumu Walshovih izraza dobivenih formulom $F(x) = \sum_{j=0}^{2^l-1} \omega_j \psi_j(x)$, u kojima rastuće duže sume osiguravaju postepeno bolje procjene vrijednosti F(x). Izrazi u sumi su dobiveni od prosječne vrijednosti F u progresivno manjim particijama elemenata.

U ovom primjeru koristimo Walshovu analizu kako bismo dobili sve bolje i bolje procjene za F(11) (= 9).

Razmotrimo prvo prosječnu vrijednost F u cijelom prostoru, koja je jednaka kao prosječna dobrota f(**) sheme ** u particiji j = 00 (dio A slike 5).

$$f(**) = \overline{F} = (F\ 00) + F\ 01) + F\ 10) + F\ 11)\ )/4 = 14/4$$

Neka je ω_{00} = f(**) = $\overline{F}$. To se može smatrati procjenom nultog reda za F(11).
Sada, kako bi dobili bolju procjenu za F(11), neke ispravke se moraju napraviti nad procjenom nultog reda. Jedan način da se to napravi je pogledati prosječnu vrijednost F na manjim particijama elemenata koji sadrži npr. F(11) - *1 (desna strana elementa B slike 5).
Prosječna vrijednost sheme *1 je

$$f(*1) = \omega_{00} - \text{ispravka}_{*1}$$

tj. jednaka je prosjeku cijelog prostora minus devijacija f(*1) od globalnog prosjeka. Slično, prosječna vrijednost komplementarne sheme *0 je

$$f(*0) = \omega_{00} + \text{ispravka}_{*1}$$

pošto je

$$f(*1) + f(*0) = 2f(**) = 2\ \omega_{00}$$

Dodjeljivanje + ili – znaka ispravci$_{*1}$ je proizvoljno. Magnituda ispravka je ista za obje sheme (*1 i *0) u particiji *d. Nazovimo tu magnitudu ω_{01}.
Bolja procjena za F(11) je tada $\omega_{00} - \omega_{01}$.

Ista stvar se može napraviti za drugu shemu reda-1 koja sadrži 11: 1*. Neka je ω_{10} devijacija prosječne vrijednosti u d* od globalnog prosjeka. Tada,

$$f(1*) = \omega_{00} - \omega_{10}.$$

Još bolja procjena za F(11) je $\omega_{00} - \omega_{01} - \omega_{10}$. To je procjena prvog reda (temeljena na 1-bitnoj shemi). Dva izraza ispravka su neovisni jedan o drugom, s obzirom da oni ispravljaju razlike u prosječnoj vrijednosti shema definiranih nad različitim bitovima. Ako je funkcija linearna, ovaj izraz bi dao točnu vrijednost za F(11). Pošto je funkcija $F(x) = x^2$ nelinearna, još jedan dodatni ispravak treba biti napravljena kako bi se nadoknadila razlika između ove procjene i prosjeka sheme reda-2, niza 11 samog:

$$F(11) = \omega_{00} - \omega_{01} - \omega_{10} + ispravak_{11}.$$

Magnituda ispravka reda-2 je ista za svaki F(x). To može biti pokazano na slijedeći način. Znamo da vrijedi:

$$F(11) = \omega_{00} - \omega_{01} - \omega_{10} + ispravak_{11}$$

i sličnom analizom

$$F(10) = \omega_{00} + \omega_{01} - \omega_{10} + ispravak_{10}\ .$$

Dodavši oba dvije strane ovih 2 jednadžbi dobivamo

$$F(11) + F(10) = 2\omega_{00} - 2\omega_{10} + ispravak_{11} + ispravak_{10}$$

ali vrijedi

$$F(11) + F(10) = 2f(1*)$$

pa imamo

$$F(11) + F(10) = 2f(1*) = 2\omega_{00} - 2\omega_{10}\ ,$$

kako je $f(1*) = \omega_{00} - \omega_{10}$, tako je $ispravak_{11} = - ispravak_{10.}$
Isto tako,

$$F(01) = \omega_{00} - \omega_{01} + \omega_{10} + ispravak_{01},$$

pa slijedi

$$F(11) + F(01) = 2\omega_{00} - 2\omega_{01} + ispravak_{11} + ispravak_{01},$$

i kako je

$$F(11) + F(01) = 2f(*1) = 2\omega_{00} - 2\omega_{01},$$

tada je ispravak $ispravak_{11} = - ispravak_{01}$.
Konačno,

$$F(00) = \omega_{00} + \omega_{01} + \omega_{10} + ispravak_{00}$$

pa slijedi

$$F(00) + F(01) = 2\omega_{00} + 2\omega_{10} + ispravak_{11} + ispravak_{01}$$

i kako je

$$F(00) + F(01) = 2f(0*) = 2\omega_{00} + 2\omega_{10},$$

imamo da je $ispravak_{00} = - ispravak_{01}$.
Tako su magnitude ispravki reda-2 jednake. Tu zajedničku magnitudu nazovimo ω_{11}.

Pokazali smo da, za ovu jednostavnu funkciju, svaka particija j' ima jednu $\omega_{j'}$ povezanu sa njom, koja predstavlja devijaciju stvarne prosječne dobrote svake sheme u particiji j' od procjena dobivenih kombinacijama $\omega_{j'}$ nižeg reda. Magnituda ovih devijacija je ista za sve sheme u particiji j'. To je bilo jednostavno za pokazati za particije reda-1 i reda-2. Generalno govoreći, za bilo koju particiju j, prosječna dobrota shema je obostrano ograničena na načine slične onima prikazanim gore, i jedinstvenost ω_j može biti pokazana na sličan način za j bilo kojeg reda.
Sada, kada smo pokazali kako vrijednosti funkcije mogu biti izračunane kao Walshovi koeficijenti, koji predstavljaju uznapredovalo točniji ispravak uvjeta prema procjenama nižeg reda u uvjetima prosjeka sheme, na slijedeći način računamo ω_j:

$\Omega_{00} = u(**) = (0+1+4+9)/4 = 14/4$
$\Omega_{01} = \omega_{00} – u(*1) = (0+1+4+9)/4 – 1\ (1+9)/2 = (0-1+4-9)/4 = -6/4$
$\Omega_{10} = \omega_{00} – u(1*) = (0+1+4+9)/4 – (4+9)/2 = (0+1-4-9)/4 = -12/4$
$\omega_{011} = F(11)$ – procjena reda-1 $= F(11) – (\omega_{00} - \omega_{01} – \omega_{10}) = 9 – (14/4 +6/4 +12/4) = 4/4$

i kako bi provjerili:

$$F(11) = \omega_{00} - \omega_{01} – \omega_{10} + \omega_{11} = 14/4 + 6/4 + 12/4 + 4/4 = 9.$$

Generalno vrijedi,

$$\omega_j = \frac{1}{2l} \sum_{x=0}^{2^l-1} F(x)\psi_j(x)$$

što je Walshova transformacija. Jednom kada su ω_j nađeni, F se može izračunati kao:

$$F(x) = \sum_{j=0}^{2^l-1} \omega_j \psi_j(x)$$

Ovaj izraz je zvan Walshov polinom.

Kako se odabire dodaje li se ili oduzima devijacija ω_j u ovom izrazu? Odgovor je temeljen na konvenciji: npr. neka f(*1) iznosi $\omega_{00} - \omega_{01}$ ili $\omega_{00} + \omega_{01}$. Kada se odabere konvencija, ona nameće ograničenja hoće li Walshovi koeficijenti višeg reda biti dodani ili oduzeti u izrazu za F(x). Ako je x član sheme s čiji prosjek ima devijaciju u pozitivnom smjeru od procjene nižeg reda, tada pozitivna vrijednost ω_j koja odgovara particiji s ulazi u sumu. Sve što je potrebno je konzistentan način pridruživanja predznaka, ovisno o particiji j i koji element od j je dani bit niz unutra. Svrha Walshovih funkcija $\psi_j(x)$ je da osigura konzistentan način pridruživanja predznaka izrazu ω_j, pomoću logičkog I pariteta. To nije jedina moguća metoda; ponešto drugačiju metodu je razvio Holland za svoju transformaciju hiperravnine.

4.4.2.3. Walsh – shema transformacija

Postoji bliska veza između Walshove transformacije i shema. Walsh – shema transformacija formalizira tu vezu. Bethke je predložio metodu za računanje prosječne dobrote f(s) za shemu s (npr. *1). Nazvao je tu metodu Walsh – shema transformacija. Ova transformacija daje uvid u način na koji se smatra da se obrada sheme dešava u GA. Ona je također dozvolila Bethkeu da izrazi neke uvjete pod kojima će funkcija biti lagana za GA da ju optimizira, i dozvolila mu je da konstruira funkcije koje su teške za GA jer sheme nižeg reda vode pretragu u krivom smjeru.

Koristeći isti primjer kao prije, prosječna dobrota sheme *1 je f(*1) = $\omega_{00} - \omega_{01}$; to dolazi od definicije ω_{01}. Vrijednost f(*1) ne ovisi o npr. ω_{10}; ona ovisi samo o Walshovim koeficijentima na dijelu koji sadrži ili *1 ili nadskup od *1 (**). Generalno, dio j uključuje shemu s ako sadrži shemu s' takvu da je s' strogi podskup od s. Na primjer, shema od 3 bita 10* je uključena sa 4 particije: dd*, d**, *d* i ***, koje odgovaraju j vrijednostima 110, 100, 010 i 000, respektivno. Treba primjetiti da j uključuje s onda i samo onda kada svaki definirani bit u j (svaki 1) odgovara definiranom bitu u s (0 ili 1, ne *).

Walsh – shema transformacija izražava dobrotu sheme s kao sumu progresivnih Walshovih koeficijenata višeg reda ω_j, analogno izrazu F(x) kao sumi progresivih ω_j višeg reda. Kao što je svaki ω_j u izrazu za F(x) ispravak za prosječnu dobrotu neke sheme u dijelu j koji sadrži x, svaki ω_j u izrazu za f(s) je izraz ispravka, koji ispravlja procjenu dobivenu nekom shemom nižeg reda koja sadrži s. Razlika je da za F(x) svi 2^1 koeficijenti particije moraju biti sumirani (mada neki od njih mogu biti 0). No, kako bismo izračunali f(s), samo koeficijenti uključenih particija moraju biti sumirani.

Gore naveden primjer je prejednostavan kako bi dobro predstavio ove ideje, ali ako ga proširimo na 3 bita može poslužiti.

Neka je F(x) = X^2. Prosječna dobrota sheme *01 je suma koeficijenata dijelova koji sadrže sheme ***, **1, *0* i *01. Jednostavan način kako bi se odredio predznak uključenih koeficijenata ω_j je uzeti bilo koju instancu od s i izračunati $\psi_j(x)$. Ova vrijednost će biti ista za sve x ε s, dok god je j uključena particija, s obzirom da su sve vrijednosti 1 u j u paru s istim bitovima u bilo kojoj instanci od s. Na primjer, particija **d (j = 001) uključuje shemu *11, i $\psi_j(x) = -1$ za bilo koji x ε *11. Koristeći sličnu metodu kako bismo dobili predznake za druge koeficijente dobivamo:

$$F(*11) = \omega_{000} - \omega_{001} - \omega_{100} + \omega_{011}$$

Generalno vrijedi

$$f(s) = \sum_{j:j\ ukljucuje\ s} \omega_j \psi_j(s)$$

gdje je $\Psi_j(x)$ vrijednost od $\psi_j(x)$ (+1 ili -1) za bilo koji x ε s.
Suma

$$f(*11) = \omega_{000} - \omega_{001} - \omega_{100} + \omega_{011}$$

daje dojam kako GA ustvari procjenjuje f(*11). Populacija nizova u GA može biti smatrana kao određen broj uzoraka različitih shema, i GA radi koristeći dobrotu nizova u populaciji kako bi procijenio dobrotu shema. GA iskorištava dobre sheme pomoću reprodukcije na način da pridružuje više uzoraka njima, i istražuje nove sheme pomoću križanja kombinirajući dobre sheme niskog reda kako bi stvorio uzorke shema višeg reda koje će također biti uspješne. Generalno postoji mnogo više instanci shema niskog reda u nekoj populaciji nego shema visokog reda (npr. u slučajno stvorenoj populaciji, oko pola nizova će biti instance od 1**...***, ali vrlo mali broj, ako će ih biti uopće, će biti instance od 111...1). Tako točne procjene dobrote će biti dobivene puno ranije za sheme niskog reda nego za sheme visokog reda. GA procjena neke sheme s može biti smatrana kao proces postupnog poboljšanja, gdje algoritam inicijalno temelji svoju procjenu na informacijama o shemama niskog reda koje sadrže s, i postepeno poboljšava svoju procjenu na temelju informacija o shemama višeg reda koje sadrže s. Slično tome, izrazi u sumi napisanoj gore predstavljaju povećano poboljšanje procjene kako dobra je shema *11. Izraz ω_{000} daje prosjek populacije (odgovarajući na prosječnu dobrotu sheme ***) i sve viši i viši redovi od ω_j u sumi predstavljaju poboljšanja višeg reda procjene dobrote za *11, gdje su poboljšanja postignuta sumiranjem ω_j odgovarajućih particijama višeg reda koje sadrže *11.

Tako, jedan način za opisivanje operacija u GA na funkciji dobrote f je taj da GA radi sve bolje i bolje procjene kakvi su Walshovi koeficijenti za funkciju f, i vodi pretragu prema dijelovima j koji imaju visoku magnitudu ω_j, i ka elementima particija (shemama) za koje su izrazi ispravki pozitivni.

4.4.2.4. Walsh – shema teorem i GA – decepcijske funkcije

Bethke (1980) je upotrijebio Walshovu analizu kako bi djelomično okarakterizirao funkcije koje će biti lagane za GA da ih optimizira. Ta karakterizacija dolazi od dvije činjenice o prosječnoj dobroti sheme . Prvo, pošto f(s) ovisi samo o ω_j za koje j uključuje s, tada ako red od j (broj jedinica u j) prelazi red od s (broj definiranih bitova u s) tada ω_j ne utječe na f(s). Na primjer, ω_{111} ne utječe na f(*11): *11 ima dvije instance 011 i 111 koje dobivaju suprotne doprinose predznaka od ω_{111}. Drugo, ako je definirajuća dužina od j (udaljenost između prve i zadnje vrijednosti 1 u j) veća nego definirajuća dužina od s (udaljenost između prvog i posljednjeg definiranog bita u s), tada f(s) ne ovisi o ω_j. Na primjer, ω_{101} ne utječe na f(*11) – opet dvije instance od f(*11) dobivaju različite doprinose predznaka od ω_{101}.

Bethke je predložio da ako se Walshovi koeficijenti funkcije brzo smanjuju sa povećanjem reda i definirajuće dužine od j – tj. da su najvažniji koeficijenti povezani sa kratkim particijama niskog reda – tada će funkcija biti lagana za GA da ju optimizira. U takvim slučajevima, lokacija globalnog optimuma može biti određena pomoću procjenjenog prosjeka dobrote shema s kratkom dužinom i niskog reda. Kao što je opisano iznad, takve sheme dobivaju mnogo više uzoraka nego sheme veće dužine i višeg reda: „niski red" znači da one definiraju veće podskupove prostora pretrage a „kratka dužina" znači da imaju sklonost ne biti promijenjene križanjem. Pritom GA može procijeniti njihovu prosječnu dobrotu brže nego onu od shema s većom dužinom i višim redom.

Tako bi, ako je sve ostalo jednako, funkcija čija Walshova dekompozicija uključuje j visokog reda sa značajnim koeficijentima trebala biti teža za GA da ju optimizira nego funkcija koja ima samo j nižeg reda, pošto će GA trebati više truda kako bi konstruirao dobre procjene shema višeg reda koje pripadaju particijama j višeg reda.

Bethkeova analiza nije bila namijenjena kao praktičan alat za odlučivanje je li dani problem lagan ili težak za GA. Kao što je spomenuto ranije, dobrota funkcija korištenih u mnogim GA aplikacijama nije oblika koji je lagan za izraziti pomoću Walshovog polinoma. Čak ako funkcija F može biti tako izražena, Walshova transformacija funkcije F zahtijeva procjenu F-a u svakoj točki u prostoru argumenata (isto vrijedi za Brzu Walshovu transformaciju), te je tako nerazumna operacija za većinu funkcija dobrote koje su nam zanimljive. Puno je efikasnije upotrijebiti GA na početnoj funkciji i mjeriti njegov učinak direktno nego rastavljati funkciju u Walshove koeficijente i tada određivati iz tih koeficijenta hoće li GA uspjeti.

Međutim, Walshova analiza može biti upotrijebljena kao teoretski alat za razumijevanje vrsta svojstava koja mogu učiniti problem teškim za GA. Na primjer, Bethke je koristio Walsh – shema transformaciju za konstrukciju funkcija koje navode GA na krivi trag, na način da je direktno dodjeljivao vrijednosti Walshovih koeficijenta tako da su prosječne vrijednosti shema niskog reda davale zavaravajuće podatke o prosječnim vrijednostima poboljšanja viših redova tih shema (shema viših redova sadržanih u shemama nižih redova). Posebno, Bethke je odabirao koeficijente tako da su kratke sheme niskog reda imale relativno nisku prosječnu dobrotu, i tada je birao druge koeficijente takve da su te sheme niske dobrote ustvari sadržavale globalni optimum. Takve funkcije su kasnije prozvane decepcijskim.

Walshova analiza može biti upotrijebljena za konstrukciju problema sa različitim stupnjem decepcije, te učinak GA može biti proučavan empirijski na takvim funkcijama.

4.4.3. Konstrukcija potpuno decepcijskih funkcija na temelju rada Whitleya

Whitley je došao na ideju konstrukcije proizvoljno duge decepcijske funkcije reda većeg od 2 na slijedeći način. Sortiraju se binarni nizovi prema njihovoj relativnoj udaljenosti u Hammingovom prostoru. Kada su nizovi sortirani na taj način (u odnosu na njihovu Hammingovu udaljenost od optimalnog rješenja ili njegovog komplementa) tada je distribucija nizova u podskupinama koje sadržavaju specifičan broj 0 ili 1 binomno raspodijeljena. Na primjer, u problemu reda-3 takav sortirani redoslijed bi bio 111, 011, 101, 110, 001, 010, 100, 000. Postoji više nego jedan pravilan raspored jer podgrupe koje imaju isti broj nula ili jedinica mogu biti posložene u bilo kojem redoslijedu. Kada je raspored odabran, nizovima se pridijele brojevi od 1 do N. Niz 1 je globalni optimum a niz N decepcijski atraktor, komplement niza 1. Nizu 2 se pridjeli neka pozitivna vrijednost B (u najjednostavnijem slučaju može biti i 0). Za bilo koji niz X, gdje je x > 2, decepcijska funkcija F_d pridjeljuje vrijednosti na slijedeći način:

$$F_d(X)=F_d(X-1)+C,$$

gdje je C neka pozitivna konstantna vrijednost.
Konačno, optimumu se pridjeljuje vrijednost $F_d(1)=F_d(N)+C$.

Slijedeća potpuno decepcijska funkcija reda-4 je konstruirana pomoću navedenog algoritma. U ovom slučaju, vrijednost B je 0, i vrijednost C je 2. Problem je takav da sve informacije o hiperravninama nižeg reda vode pretragu od 1111 prema 0000.

Decepcijska funkcija 2:

f(1111) = 30	f(0100) = 22	f(0110) = 14	f(1110) = 6
f(0000) = 28	f(1000) = 20	f(1001) = 12	f(1101) = 4
f(0001) = 26	f(0011) = 18	f(1010) = 10	f(1011) = 2
f(0010) = 24	f(0101) = 16	f(1100) = 8	f(0111) = 0

Treba primijetiti da svaka relevantna shema reda-3 ili manje koja je sastavljena isključivo od 0 i * je nadmoćna prema svim primarnim natjecateljima hiperravnine.

Decepcijska funkcija prikazana gore, kao i sve umjetno stvorene decepcijske funkcije, ima decepcijski atraktor koji je lokalni optimum u Hammingovom prostoru s obzirom da je ta točka superiorna svakoj susjednoj točki u Hammingovom prostoru. Razlog zašto je funkcija iznad (i decepcijska funkcija 1) potpuno decepcijska funkcija je jasan: u Hammingovom prostoru decepcijski optimum ima bazu privlačenja koja pokriva cijeli prostor osim jednog šiljka koji predstavlja pravi optimum. Međutim, decepcijski atraktori potpuno decepcijskih funkcija ne moraju uvijek biti lokalni optimumi u Hammingovom prostoru.

4.5. Teorem decepcijskog atraktora

Slijedeći teorem pokazuje da za svako preraspoređivanje funkcije skupa nizova u skup vrijednosti, rezultirajuće preraspoređivanje neće biti potpuno decepcijsko u očekivanom određenom redu osim ako postoji globalni atraktor koji je komplement globalnog optimuma. Međutim, to ne znači da komplement mora imati visoku vrijednost dobrote koja je usporedna sa globalnim optimumom. Ovaj teorem vrijedi kako za decepcijske funkcije tako i za decepcijske građevne blokove.

Pojam atraktor predlaže da je dinamični sustav okarakteriziran stabilnim atraktorom. Pojam decepcijski atraktor je upotrijebljen neprecizno ali dokazi nalažu da:

1) ako je problem potpuno decepcijski
2) decepcijski atraktor je komplement globalnog optimuma
3) rekombinacija uvijek razara informacije reda-N koje identificiraju globalni optimum,

tada će decepcijski atraktor biti stabilna točka za više populacija koje dostatno prikazuju relevantnu decepcijsku hiperravninu reda manjeg od N.

Preciznije rečeno, ako definiramo neki idealizirani GA gdje je pretraga dominirana od strane decepcijske hiperravnine reda N-1 ili manjeg, kada pretražujemo potpuno decepcijski problem reda N tada decepcijski atraktor mora biti stabilna točka za bilo koju potpuno decepcijsku funkciju: sva natjecanja hiperravnina reda N-1 vode prema atraktoru.

Unatoč tome, teorem koji slijedi ne počiva na opažanjima o dinamičnim sustavima nego na promatranjima o vezi binarnih nizova i hiperravnina u N-dimenzionalnoj kocki. Slijedeće definicije su stoga temeljene na statičnim odnosima. Ako je GA modeliran kao dinamičan sustav ti teoremi ne moraju više vrijediti (ali će bar biti polazna točka).

Teorem 2:

Kako bi funkcija ili građevni blok reda N bili konzistentno decepcijski na svim relevantnim hiperravninama nižeg reda, decepcijski atraktor mora biti komplement niza koji predstavlja globalni optimum u decepcijskoj funkciji, ili u slučaju decepcijskog građevnog bloka, decepcijski atraktor mora biti komplement sheme koja predstavlja „globalnog pobjednika“ relevantnog primarnog natjecanja hiperravnina reda N koji je nadmoćan svim ostalim natjecateljima.

Dokaz kontradikcijom:

Pretpostavimo decepcijski atraktor reda N takav da dijeli barem jednu bit vrijednost sa „globalnim pobjednikom“ relevantnog natjecanja hiperravnina reda N. Odabere se jedan bit koji ima vrijednost koju dijeli globalni pobjednik i decepcijski atraktor. Neka je G bit koji odgovara globalnom pobjedniku, a D bit koji odgovara decepcijskom atraktoru. Oni su izabrani tako da ti bitovi imaju istu vrijednost (0 ili 1). Problem nije konzistentno decepcijski za neki određeni redoslijed osim ako je f(*...*G*...*) < f(*...*C*...*), no to nije moguće, s obzirom da su hiperravnine predstavljene shemama koje imaju iste bitove i tako te hiperravnine imaju iste vrijednosti dobrote.

Neka je C komplement od bitova G i D (koji su u stvari isti bitovi). Kako bi postojala decepcija, potrebno je da f(*...*G*...*) < f(*...*C*...*), što također implicira da f(*...*D*...*) < f(*...*C*...*). Tako, ako je problem decepcijski na svim hiperravninama, natjecanje hiperravnina mora biti dobiveno od strane sheme reda-1 koja sadrži bit C i ne bit D. Iz toga slijedi da predloženi decepcijski problem sa decepcijskim atraktorom koji ima bit D nije konzistentno decepcijski na svim hiperravninama naznačenog reda.

QED

Rečeno jednostavnije, relevantno natjecanje hiperravnine reda-1 može se slagati samo sa globalnim pobjednikom reda N ili njegovim komplementom. Tako, kako bi bila potpuno ili konzistentno decepcijska, jedina hiperravnina koja može služiti kao decepcijski atraktor je komplement globalnog pobjednika relevantnog natjecanja hiperravnine reda N. Naravno, ovaj zaključak pretpostavlja statičan pogled na odnose hiperravnina. Kada se gleda dinamička situacija, problem postaje puno kompleksniji, ali neka se opažanja mogu dati. Prvo, teorem ima snažnije implikacije za potpuno decepcijske funkcije, s obzirom da sve hiperravnine nižeg reda vode ka istoj točki; to omogućuje stvaranje snažnijih tvrdnji o statusu decepcijskog atraktora kao pravog atraktora dinamičkog sustava. Također, treba primjetiti da postoje i drugi atraktori ako problem nije potpuno ili konzistentno decepcijski. Biti će drugih atraktora ako hiperravnine nižeg reda ne dominiraju pretragom (zbog predrasude uzorkovanja shema višeg reda); u stvari globalno rješenje je jedan takav natjecateljski atraktor. Unatoč tome, u potpuno ili konzistentno decepcijskom problemu hiperravnine reda-1 moraju voditi pretragu prema decepcijskom atraktoru što će imati utjecaj na druga relevantna natjecnja hiperravnina. Sve to vodi do slijedećeg korolara, definiranog sa statičkog gledišta.

Korolar 1:

Ako decepcijski atraktor reda N nije komplement globalnog pobjednika relevantnog natjecanja hiperravnina reda N, tada problem nije niti potpuno niti konzistentno decepcijski. To slijedi iz dokaza teorema decepcijskog atraktora jer taj teorem razmatra hiperravnine *...*G*...*, *...*D*...* i *...*C*...*. S obzirom da to predstavlja neki skup hiperravnina s definiranim jednim bitom, ako *...*D*...* nije decepcijska hiperravnina, tada problem nije niti potpuno niti konzistentno decepcijski.

4.6. Decepcijski atraktor i lokalni optimum

Postavke iznesene u prethodnom poglavlju govore da decepcijski atraktor mora biti komplement globalnog pobjednika relevantnog natjecanja hiperravnina reda N. Međutim, ako ograničimo naše razmatranje na decepcijske funkcije za trenutak, ne slijedi da komplement globalnog pobjednika natjecanja hiperravnine reda N mora imati visoku vrijednost dobrote. Hiperravnine u potpuno decepcijskoj funkciji mogu voditi ka decepcijskom atraktoru koji nije lokalni optimum u Hammingovom prostoru. Ova postavka također vrijedi za decepcijske građevne blokove, s obzirom da se takav problem može konstruirati konkatenacijom nekoliko decepcijskih funkcija gdje decepcijski atraktor nije lokalni optimum.

Razmotrimo slijedeći primjer:
decepcijska funkcija 3:

f(1111) = 30	f(0100) = 27	f(0110) = 5	f(1110) = 0
f(0000) = 10	f(1000) = 28	f(1001) = 5	f(1101) = 0
f(0001) = 25	f(0011) = 5	f(1010) = 5	f(1011) = 0
f(0010) = 26	f(0101) = 5	f(1100) = 5	f(0111) = 0

Funkcija prikazana gore je konstruirana tako da je jednostavna za razumijevanje. To je potpuno decepcijska funkcija, ali decepcijski atraktor ima vrijednost dobrote 1/3 globalnog optimuma. Decepcijski atraktor očito nije lokalni optimum u Hammingovom prostoru jer je slabiji nego bilo koji susjedni niz. Decepcijski atraktor reda N mora biti skriven od strane jakih nizova lociranih susjedno u odnosu na njega u Hammingovom prostoru ako se želi da problem bude potpuno decepcijski. Slijedi da taj faktor nije ključan za konzistentno decepcijske funkcije. U svakom slučaju, primjer dan iznad pokazuje da sam decepcijski atraktor ne treba biti jaki natjecateljski niz, nego mora biti zamaskiran jakim nizovima koji su mu susjedni u Hammingovom prostoru kako bi zamaskirali njegovu slabost. Može se doći i do zaključka o dobroti decepcijskog atraktora u odnosu na druge nizove (ili relevantne sheme) koje su mu blizu u Hammingovom prostoru.

Teorem 3:

Decepcijski atraktor reda N za binarno kodiran problem ne može zadržati potpunu decepciju reda N ako je slabiji nego bilo koji niz ili shema koja se razlikuje od decepcijskog atraktora za točno 2 bita.

Dokaz:

Razmotrimo potpuno decepcijsku funkciju reda N. Neka je Z binarni niz (ili shema) sa bit vrijednostima 0 koje predstavljaju decepcijski atraktor i f(Z) vrijednošću povezanom sa decepcijskim atraktorom. Neka je X niz (ili shema) koja se razlikuje od Z za točno 2 bita (tj. ima 2 bita sa vrijednošću 1) i koji ima veću vrijednost nego bilo koji niz koji se razlikuje od decepcijskog atraktora za 2 bita. Ako je pokazano da f(Z) mora biti veća od f(X), tada f(Z) mora biti veća nego vrijednost bilo kojeg niza koji se razlikuje za točno 2 bita od decepcijskog atraktora.

Još su dva niza (ili sheme) važni. S1 i S2 su dva jedinstvena niza (ili sheme) koji se razlikuju u jednom bitu od decepcijskog atraktora i od niza (sheme) X. Drugim riječima, oba imaju vrijednost 1 za jedan bit. Sada se pridruže imena obim nizovima tako da niz S1 predstavlja niz (ili shemu) koji ima vrijednost dobrote veću ili jednaku S2. To također implicira natjecanje hiperravnina između ta 4 niza (ili shema) uključujući shemu reda N-1. Pretpostavimo da S1 i X dijele bit nazvan B1; također S2 i X dijele drugi bit nazvan B2. Niz S1 i X se natječu protiv S2 i Z kada se gleda shema reda N-1 koja zamjenjuje bit B2 sa „bilo što“ operatorom *. Teorem decepcijskog atraktora i definicija potpune decepcije impliciraju:

$$f(Z) + f(S2) > f(S1) + f(X)$$

pa slijedi

$$f(Z) > f(S1) - f(S2) + f(X)$$

S obzirom da je f(S1) >= f(S2), neka je f(S1) – f(S2) = K, gdje je K pozitivan cijeli broj.
Iz toga slijedi:

$$f(Z) > f(X) + K, \text{ tako da je } f(Z) > f(X)$$

QED

Očita poteškoća kod održavanja potpune decepcije reda N se pojavljuje kada su hiperravnine reda N-1 umiješane. Sada odaberimo neku decepcijsku funkciju reda-4. Neka je 0000 Z, 10001 X, 1000 i 00001 su dva jedinstveno definirana niza koji se razlikuju za jedan bit od decepcijskog atraktora i niza 10001. Odabere se bilo koji niz koji će imati višu dobrotu; npr. f(1000) > f(0001).

To također implicira relevantno natjecanje hiperravnina između shema reda N-1. Potpuna decepcija implicira f(000*) > f(100*), što opet implicira

$$f(0000) > + f(0001) > f(1000) + f(1001)$$

Korolar 2:

U potpuno decepcijskoj funkciji, ako decepcijski atraktor nije lokalni optimum u Hammingovom prostoru, lokalni optimum mora postojati za neki niz koji je susjedan decepcijskom atraktoru u Hammingovom prostoru. Drugim riječima, ili decepcijski atraktor ili neki niz koji se razlikuje od decepcijskog atraktora za 1 bit mora biti lokalni optimum u Hammingovom prostoru.

To slijedi iz teorema 2. Ako decepcijski atraktor mora imati veću vrijednost dobrote nego bilo koji niz koji se razlikuje od decepcijskog atraktora za točno 2 bita, tada je decepcijski atraktor (i niz koji se razlikuje od decepcijskog atraktora za jedan bit) okružen točkama u Hammingovom prostoru koje su slabije nego decepcijski atraktor. To implicira da je ili decepcijski atraktor lokalni optimum, ili je jedan od nizova koji se razlikuju od decepcijskog atraktora za jedan bit lokalni optimum u Hammingovom prostoru.

Sada promotrimo konzistentno decepcijski problem koji nije potpuno decepcijski. Ako ublažimo uvjete za decepciju reda N možemo sastaviti problem koji je konzistentno decepcijski ali nije potpuno decepcijski; od posebnog su interesa problemi koji su konzistentno decepcijski i gdje mnoge, ali ne sve hiperravnine nižeg reda vode ka decepcijskom atraktoru, ali gdje niti jedna ne vodi ka globalnom pobjedniku natjecanja hiperravnina reda N. Koristeći skup vrijednosti iz prethodnog primjera, ako se vrijednost niza 0000 promijeni iz 10 u 7, problem više nije potpuno decepcijski jer f(000*) > f(1000*), (f000*) = f(010*) i f(00*0) = f(10*0), ali sve informacije hiperravnina i dalje vode ka 0000 – decepcijskom atraktoru. Međutim, u takvim slučajevima i dalje ima smisla govoriti o decepciji reda-4 koja je manja nego potpuno decepcijska, jer se decepcijski problem reda-4 ne može svesti na decepcijski problem reda-3. U stvari, nekoliko hiperravnina reda-3 i dalje nastavlja biti decepcijsko u odnosu na ispravan građevni blok reda-4. Hiperravnine *000 i 0*00 i dalje imaju više vrijednosti dobrote nego bilo koji od ostalih natjecatelja i tako nastavljaju podržavati decepciju na nivou reda-4. Iako funkcija nije potpuno decepcijska s obzirom da sve hiperravnine nižeg reda ne vode prema decepcijskom atraktoru, problem je i dalje decepcijski na svim hiperravninama. Niti na jednom natjecanju hiperravnina ne pobjeđuje ona hiperravnina koja vodi ka ispravnom građevnom bloku.

Smanjili smo stupanj decepcije, ali i dalje ostaju siva područja između potpuno decepcijskog problema reda-4 i potpuno decepcijskog problema reda-3. Takav tip situacije pokazuje razliku između potpunog decepcijskog problema i konzistentno decepcijskog problema. U ovom trenutku ne postoje empirijski dokazi o teškoći konzistentnih decepcijskih problema.

4.7. Strategije preraspodjeljivanja decepcijskog atraktora

Koje praktično značenje sve ovo ima? Spoznaja da decepcijski atraktor mora biti komplement globalnog pobjednika je korisna ako se decepcija pojavljuje na poznatoj ili predvidivoj lokaciji (npr. unutar postavki parametara); rekodiranje problema koje će značajno promijeniti Hammingovu udaljenost između globalnog pobjednika i decepcijskog atraktora natjecanja hiperravnina će smanjiti stupanj decepcije koja postoji u tom trenutku. Male promjene možda neće smanjiti nivo decepcije; na primjer, u decepcijskoj funkciji kreiranoj u prethodnom odlomku okretanje rasporeda između 1000 i 0000 ne smanjuje stupanj decepcije – u stvari to pojačava decepciju. Ali preraspoređivanje velikog broja nizova tako da globalni pobjednik bude preseljen bliže decepcijskom atraktoru i onim nizovima koji pomažu održanju decepcije u Hammingovom prostoru će u mnogim slučajevima smanjiti nivo decepcije.

Također, čini se mogućim da preraspoređivanje može stvoriti decepciju u drugim dijelovima prostora pretrage, mada mogućnost da se to dogodi ne implicira ništa o vjerojatnosti takvog događaja. Problem kod rekodiranja ili preraspoređivanja strategija je da znanje o točnoj lokaciji decepcije nije generalno praktično, jer to uključuje nalaženje binarne maske koja pokazuje lokacije gdje se decepcija događa; nalaženje takve binarne maske uključuje prostor pretrage velik koliko i prostor optimizacije funkcije. Također nije jasno da li jedna maska za decepciju zadovoljava: nužno je znati koliko različitih slučajeva decepcije postoji i gdje se nalaze. Problem nalaženja bitova koji tvore decepcijski građevni blok također uključuje takva jednostavna rješenja za decepciju kao što su inverzija konačnog rješenja ili inverzija nizova tokom pretrage i ocjenjivanje njihovih komplemenata. To nisu dobre vijesti za strategiju preraspoređivanja osim ako se ne može preraspodijeliti slučajno odabrane dijelove koda i nadati se da su tim zahvaćeni i dijelovi decepcije.

4.8. Decepcija i problemi povezanosti

Decepcija je samo dio problema. Drugi faktor koji čini problem GA-teškim je povezanost između bitova u decepcijskom građevnom bloku. Decepcija je puno ozbiljniji problem ako se pojavljuje u nekim slučajnim (i time, nepoznatim) kombinacijama bitova u prikazu. Imati slabu povezanost implicira da su decepcijski bitovi vrlo razdvojeni u prikazu. Hiperravnine prikazane bitovima koji su distibuirani po dugom nizu su loše uzorkovane genetskim algoritmom zbog veće stope rascjepa tijekom rekombinacije. Jednostavnije rečeno, ako su decepcijski bitovi blizu jedan drugome, tada je niz koji čini neku pogodnu kombinaciju bitova vjerojatniji za prijenos odgovarajuće sheme netaknute svojim potomcima (jer križanje rascjepljuje kritične sheme rjeđe) i GA neće biti naveden na krivi put. Uzmimo primjer kada se decepcija pojavljuje u decepcijskom građevnom bloku dužine 3 bita. Bilo bi očekivano da se točna kombinacija (111) pojavljuje u 12.5% slučajeva u inicijalnoj slučajnoj populaciji; ako kritični bitovi imaju sklonost prolaska kao jedan blok, tada će se proširiti u populaciji i decepcija koja se pojavljuje na hiperravninama nižeg reda neće imati veliku važnost za pretragu. Međutim, ako se ta tri bita razdvoje, tada je manje vjerojatno da će biti naslijeđeni u očuvanom obliku (zbog rascjepljujućeg utjecaja križanja) i tako informacije hiperravnina nižeg reda postaju kritične. Kada se to dogodi, GA će vjerojatno skrenuti ka netočnom rješenju.

Pogledajmo decepcijsku funkciju 1. Također, pretpostavimo da su 3 bita koja čine decepcijsku funkciju razdvojeni u binarnom prikazu sastavljenom od nekoliko decepcijskih funkcija. Kako se natjecanje između 000 (sa vrijednošću 28) i 111 (sa vrijednošću 30) pojačava, nema dovoljno selektivnog pritiska kako bi se prevladao rascjepljujući utjecaj rekombinacije. Time natjecanje pada na nivo relevantnih hiperravnina nižeg reda, i 000 pobjeđuje zbog decepcije ugrađene u problem.

Kako bi se izgradila pogodna test funkcija koja je vjerojatna da će biti teška za GA za riješiti, Goldberg, Korb i Deb (1989.) su definirali 30-bitnu funkciju sastavljenu od 10 kopija potpuno decepcijske funkcije od 3 bita. Problem je težak (ili „ružan“) kada su bitovi u takvom rasporedu da je svaki od 3 bita podfunkcije uniformno i maksimalno raspodijeljen unutar prikaza. Tako svaka podfunkcija od 3 bita i, ima bitove na lokaciji i, i + 10, i + 20. Goldberg i suradnici su pokazali da standardni GA konzistentno skreće ka netočnom rješenju problema, sa svakom podfunkcijom koja konvergira ka 000 umjesto prema 111. Opet, to je u skladu s opažanjima o uzorkovanju hiperravnina i predrasudama poznatim da postoje u GA protiv shema velikih dužina.

5. Načini rješavanja problema decepcije

5.1. Uvod

Danas postoji puno različitih načina rješavanja decepcijskih problema gdje svaka metoda ima svoje prednosti i mane, no niti jedna ne nudi definitivno rješenje problema decepcije. U nastavku je napravljen presjek nekoliko metoda rješavanja decepcijskih problema.

5.2. Poboljšani GA za rješavanje decepcijskih problema

Prema teoremu Whitleya, za potpuno decepcijske probleme možemo dobiti globalni optimum tako da komplementiramo prikaz lokalnog optimuma tijekom izvođenja GA. Nadalje, za konzistentno decepcijske probleme, kada provodimo komplementarnu operaciju nad optimalnim jedinkama u trenutnoj populaciji GA, dobiti ćemo globalni optimum direktno ako je decepcijski atraktor lokalni optimum, ili ćemo dobiti binarni niz čija je Hammingova udaljenost od optimalnog niza 1 ako decepcijski atraktor nije lokalni optimum. Međutim, u praksi se češće pojavljuju nekonzistentno decepcijski ili parcijalno decepcijski problemi. Kako bi se poboljšao GA treba napraviti slijedeće:

Neka je $x = (x_1, x_2, ..., x_L)^\tau$ binarno kodirana jedinka, gdje je $x_i \varepsilon \{0, 1\}$, $(i = 1,2,...,L)$ i-ti gen jedinke x. Funkcija dobrote f(x) je preraspoređivanje od X ka R^+, i optimalni cilj f(x) je naći $x^* \varepsilon$ X, koji zadovoljava uvjet da je $f(x^*)$ maksimum od f(x). Nadalje, neka je $P_i \subset X$ koji predstavlja i-tu generaciju populacije tijekom evolucijskog procesa GA.
Prvo postavimo neki cijeli broj γ. Ako se optimalna jedinka u grupi tijekom γ uzastopnih generacija ne promijeni, pretpostavimo da je GA naišao na lokalni optimum ili decepcijski atraktor. Tada, postavimo vrijednosti jedinkama u trenutnoj populaciji od najveće do najmanje prema njihovoj dobroti. Posljedično, mislimo na m elitističkih jedinki, označenih kao $(x_{1k}, x_{2k}, ...,x_{lk})$, k = 1, 2, ...,m i definiramo mjeru raznolikosti i-tog gena od tih m elitističkih jedinki kao

$$z(i) = 1 - \frac{1}{m}(\max\left\{\sum_{j=1}^{m}(x_{ij}), \sum_{j=1}^{m}(1 - x_{ij})\right\} - \min\left\{\sum_{j=1}^{m}(x_{ij}), \sum_{j=1}^{m}(1 - x_{ij})\right\})$$

gdje je z (i) ε [0, 1](i= 1, 2,..., l). Ako je z (i) = 0 tada je raznolikost i-tog gena potpuno nestala, a ako je z(i) = 1, tada je raznolikost gena dosegla maksimum. Tako, kada raznolikost nekog gena nestane u potpunosti, možemo zaključiti da je taj gen potencijalni kandidat za uvođenje decepcije sheme. Zato možemo naći sve lokacije gena koji mogu uzrokovati decepciju u trenutnoj populaciji. Tada provodimo komplementarnu operaciju za te lokacije gena od gornjih m elitističkih jedinki u trenutnoj populaciji a druge gene zadržavamo sa starom vrijednošću. Ovom metodom, dobivamo m jedinki. Također, provodimo komplementarnu operaciju jedinke koja predstavlja optimalno rješenje u trenutnoj populaciji kako bi dobili (m+1) jedinki. Tih m+1 jedinki tada zamjenjuje najgorih m+1 jedinki u trenutnoj populaciji P_i kako bi se uništila decepcijska shema u GA. Tada se provode operacije selekcije, križanja i mutacije za trenutnu populaciju.

5.3. Strukturirani GA

Centralna karakteristika strukturiranog GA (sGA) je upotreba hijerarhije i izraza gena u genotipu. Geni na bilo kojem nivou mogu biti ili aktivni ili pasivni. Geni na visokom stupnju aktiviraju ili deaktiviraju skupove gena na niskom stupnju. Tako su aktivnosti gena na bilo kojem nivou, bez obzira jesu li su oni izraženi fenotipski ili ne (u genotip – fenotip preraspoređivanju), odlučene njihovim genima na visokom stupnju. U sGA genomi su uključeni u kromosom i predstavljeni su kao skupovi binarnih podnizova. Ovaj model također koristi konvencionalne genetske operatore i kao kriterij opstanka najsposobnijih kako bi dobio što bolje potomke. Kao efekt, ovaj model osigurava mehanizam za genetsku evoluciju u kojoj je raznolikost održana zaštitom dodatnog genetskog materijala i kontrolom njihovih izraza tijekom dekodiranja. To povećava vjerojatnost pretrage potencijalnih područja fenotipskog prostore s velikom učinkovitošću. Kako komadići informacija uključenih u akciju u narednim generacijama mogu imati precizno kompatibilne komponenete s velikom varijabilnošću, može se izbjegnuti da ostane uhvaćena u zamku u lokalnom optimumu što uzrokuje preuranjena konvergencija.

Tako sGA osigurava mehanizam za rješavanje problema održavanjem kvazi – stabilnog stanja genetske raznolikosti tijekom cijelog evolucijskog procesa. Takav prikaz omogućuje znatnu fleksibilnost kada se primijeni na decepcijske probleme.

U implementaciji sGA definirajuća dužina fizičke sheme i izražene sheme je različita. Generalno, definirajuća dužina fizičke sheme varira ovisno o uzorku aktiviranja gena visokog stupnja i njihovih definiranih bita niskog stupnja koji se preslikavaju. Taj je stupanj uvijek veći ili jednak onome izražene sheme gdje je definicija izražene sheme slična onoj u shemi jednostavnog GA i dekodira prostor rješenja. Mada postoji dugačka definirajuća dužina u sGA prikazu, uništavanje križanjem nije tako vjerojatno da će zavarati potragu kao u slučaju jednostavnog GA, obzirom da je nešto od efekta absorbirano redundantnim segmentima koji su uobičajeno rašireni duž kromosoma. Drugim riječima, križanje u dvije točke i mutacija mogu dovesti do bilo koje točke na genotipu niskog stupnja i uobičajeno ostavlja za posljedicu manje efekte razaranja na izraženom dijelu kromosoma. Konvergencija s GA indicira homogenost u izraženim genima, ne u cijelom genotipu, tako da omogućuje da se viši stupanj raznolikosti održava u kromosomu duž njegove neutralne forme.

5.4. Prošireni GA

Ovaj način predpostavlja da SBBH vrijedi za GA te se njime uz male promjene na osnovnom GA mogu riješiti decepcijski problemi.

Puno istraživanja o decepciji uključuje funkcije za koje je komplement globalnog optimuma decepcijski atraktor (Liepins i Vose 1991., Whitley 1991.). U stvari, postoje argumenti da sve potpuno decepcijske funkcije imaju to svojstvo. Za potrebe ovog algoritma rješavanja problema, pretpostavimo da taj argument vrijedi i pretpostavimo da će se GA ponašati u skladu sa SBBH na potpuno decepcijskim problemima. Drugim riječima, pretpostavimo da GA stvarno skreće ka komplementu globalnog optimuma.Tada je algoritam za nalaženje globalnog optimuma u svim takvim problemima jednostavan:

1. vrti GA dok ne konvergira ka nekom x
2. rješenje je ili x ili komplement od x, što god je bolje.

Pod cijenu jedne dodatne procjene, ovaj algoritam proširuje klasu problema koji mogu biti optimizirani pomoću GA kako bi uključio sve potpuno decepcijske probleme.

Međutim, češći su slučajevi u kojima su problemi samo djelomično decepcijski, tj. problemi koji imaju decepcijsku komponentu i ne-decepcijsku komponentu. Takvi problemi mogu biti rješeni pomoću proširenog GA (aGA) prikazanog pseudokodom ispod. Linije prikazane znakom * predstavljaju promjene na standardnom generacijskom GA. U ovoj verziji, postoje 3 odvojene populacije veličine N, nazvane P, Q i R. Tijekom svake generacije, ažuriramo P u skladu s originalnim GA, postavimo članove populacije Q na komplemente odgovarajućih elemenata u populaciji P, i kreiramo R križanjem slučajno izabranih roditelja iz P i Q.

Razmotrimo neki problem za koji originalni GA (onaj koji nije označen linijom *) nalazi prihvatljivo rješenje u vremenu t, koristeći populaciju veličine N. Tada prošireni GA nalazi rješenje koje je barem jednako dobro kao ono originalnog GA, trebajući maksimalno 3t vremena (pretpostavljajući da vrijeme procjenjivanja ima dominantnu vremensku konstantu u algoritmu). Kao i jednostavnija verzija, prošireni algoritam rješava bilo koji potpuno decepcijski problem koji ima svojstvo da je globalni optimum binarni komplement decepcijskog atraktora, jer čim populacija P stvori kopiju decepcijskog atraktora, populacija Q stvori kopiju globalnog optimuma.

Dozvolivši rekombinaciju između slučajno izabranih članova populacije P i Q, SBBH predviđa da možemo također dobiti optimalno rješenje za parcijalnu decepciju. Za bilo koju komponentu koja nije decepcijska, elementi u P će konvergirati ka komplementu optimalne vrijednosti komponente. Tako će komplementi decepcijskih komponenti – optimalne vrijednosti komponenata – biti pohranjeni u populaciji Q. Izvodeći križanje u više točaka nad populacijama P i Q trebalo bi proizvesti optimalne komponenete duž cijelog kromosoma u populaciji R. Tako prošireni GA može doći do konačnog odgovora koji nikad nije lošiji nego onaj dobiven originalnim GA, te rješava brige oko decepcijskih problema, naravno pod cijenu utrostručenja vremena izračunavanja.

Može se zaključiti da decepcija, kako je trenutačno definirana, nije ozbiljan problem za GA, s obzirom da se može riješiti malim promjenama na originalnom algoritmu. Nažalost, ova rasprava je teoretska, s obzirom da predloženo rješenje, kao i decepcija sama, počiva na SBBH i time nije vjerojatno da daje korisne rezultate za realne GA.

Algoritam proširenog GA:

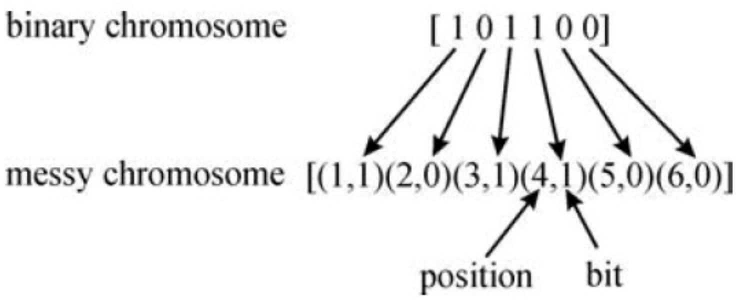
```
procedure  prosireni GA
begin
      t=0;
      inicijalizacija P(t);
      procjena struktura u P(t);
      while uvjet završenja nije ispunjen
      begin
            t=t+1;
            odabir P(t) iz P(t-1);
            rekombinacija struktura u P(t);
            (*) Q(t)=komplement od P(t);
            (*)stvaranje R(t) rekombinacijom P(t) i Q(t);
            procjena struktura u P(t);
            (*)procjena struktura u R(t);
            (*)procjena struktura u R(t);
            prikaz najbolje strukture u P(t)∪R(t) ∪Q(t);
      end
end
```

5.5. Messy GA

Messy GA je izmišljen kako bi se ubrzala konvergencija genetskog algoritma (Goldberg i suradnici 1989.). Ovaj tip GA ima varijabilnu dužinu kromosoma i gena koji ne ovise o poziciji. *Messy* GA dodjeljuje broj pozicije svakoj binarnoj znamenci. Uređeni par je predstavljen kao:

(pozicija gena, vrijednost bita)

Tako geni nisu pozicijski zavisni u kromosomu. Kromosom također nema fiksiranu vrijednost.



Slika 6. Pozicija je dodjeljena svakom bitu u mGA

Slijedeća slika prikazuje dva *messy* kromosoma, koji imaju dužinu manju od 6 bita. Prvi kromosom je prespecificiran, jer ima dvije vrijednosti za poziciju gena 3. Drugi kromosom je nedovoljno specificiran jer ima vrijednosti za gene 1, 4 i 6, ali nema za gene 2, 3 i 5.

	A	**B**
prespecificirani	(3,1) (2, 0) (1, 0) (5, 0)	(3, 0) (6, 1) (4, 0)
	C	**D**
nedovoljno specificirani	(4, 1) (6, 0)	(1, 0)

Slika 7. Prikaz križanja u mGA

Prva faza u mGA se naziva primordialna faza. Ona počinje sa inicijalnom populacijom koja ima sve moguće građevne blokove u populaciji određene dužine. Slijedeće, GA se vrti nekoliko generacija koristeći samo križanje. Pola kromosoma je odbačeno svakih nekoliko generacija. Tada *juxtapositional*-na faza poziva druge genetske operatore dok se ne nađe zadovoljavajuće rješenje.

Dva specijalna mGA operatora koji se koriste u *juxtapositional* fazi su rezanje i spajanje. Slučajni rez se radi na kromosomima (odrezani dijelovi su nazvani A, B, C i D). Dva dijela su slučajno odabrana (ne mogu biti isti) i spojeni su zajedno kako bi tvorili nove kromosome.

(3,1) (2, 0) (1, 0) (5, 0) (4, 1) (6, 0) **A – C** točno specificirani
(3,1) (2, 0) (1, 0) (5, 0) (1, 0) **A – D** previše i premalo specificirani
(1, 0) (3, 0) (6, 1) (4, 0) **D – B** premalo specificirani
(4,1) (6, 0) (3, 0) (6, 1) (4, 0) **C – B** previše i premalo specificirani

Slika 8. Rezultirajući potomci gdje su neke pozicije prespecificirane, neke nedovoljno i neke točno specificirane.

Slika iznad prikazuje 4 od 12 mogućih rezultirajućih kromosoma. Oni imaju različite dužine. Goldberg je izvjestio da su rezultati na većem broju primjera decepcijskih funkcija ohrabrujući.

6. Zaključak

Genetski algoritmi predstavljaju moćnu metodu optimizacije. No, u radu genetskog algoritma se mogu pojaviti neki problemi. Jedan od tih problema je decepcija, gdje decepcija predstavlja konvergiranje rješenja prema nekom rješenju različitom od globalnog optimuma.

Koliki problem decepcija uopće predstavlja? De Jong je rekao da je decepcija problem samo ako tražimo globalni optimum – najbolje rješenje. Često je dovoljno naći samo rješenje koje zadovoljava uvjete i u tim slučajevima decepcija nije toliko važna. Jednostavno, omjer rizik / isplativost je previsok da bi se tražilo najbolje rješenje.

Decepcija može biti različitog stupnja – od potpune gdje sva natjecanja hiperravnina vode ka globalnom ataktoru pa do parcijalne koja uopće ne mora zavarati genetski algoritam. Danas postoji više metoda za rješavanje različitih stupnjeva decepcije. Svaka metoda bi trebala davati odgovor na dva pitanja: hoće li ovaj problem biti decepcijski i ako hoće, kako riješiti tu decepciju?

Neke metode za rješavanje decepcije sam spomenuo, a postoje i mnoge druge, no te su metode uglavnom testirane ne na realnim problemima već na testnim funkcijama za koje se unaprijed zna koji stupanj decepcije imaju.

Na pitanje hoće li problem biti decepcijski nema odgovora osim onih dobivenih empirijskim putem. Pronalazak uvjeta pojave decepcije se čini mogućim odgovorom na to pitanje. Što je do sada napravljeno po tom pitanju?

Belgijski matematičari Naudts i Vershoren su pokazali da se decepcija ne pojavljuje u slučajevima gdje je niska epistaza (koliko doprinos jednog gena u dobroti ovisi o vrijednostima drugih gena) i da funkcije s visokom epistazom nisu nužno decepcijske. Lopez je pokazao da je za dugačke funkcije zamke (veće od 10^4 bitova), kompleksnost (Kolmogorov 1965., Chaitin 1974., Lempel i Ziv 1976., kompleksnost je mjera količine informacija) slučajno odabrane funkcije zamke obrnuto proporcionalna vjerojatnosti da je funkcija potpuno decepcijska.

Ovih nekoliko primjera pokazuje da je decepcija u sprezi sa drugim parametrima sustava. No ovi primjeri daju samo dovoljne uvjete decepcije. Izazov je naći nužan uvjet (ili skup uvjeta) za pojavu decepcija.

Neka rješenja na polju određivanja stupnja decepcije već postoje: Li Yun-qiang je osmislio Brzi algoritam za izračunavanje stupnja decepcije u funkcijama cilja. No taj algoritam vrijedi samo uz uvođenje nove definicije stupnja decepcije i za sada nije istraženo može li ovaj algoritam odigrati bitnu ulogu u nalaženju općenitog algoritma za određivanje stupnja decepcije.

7. Literatura:

[1] Marin Golub, Genetski algoritam prvi dio, Zagreb 2004.
[2] Carlos A. Coello Coello, Gary B. Lamont, David A. Van Veldhuizen, Evolutionary Algorithms for Solving Multi-Objective Problems, Springer 2007
[3] Paul B.Parker, Genetic Algorithms and Their Use in Geophysical Problems, doktorska disertacija, 1999
[4] Thomas Weise, Global Optimization Algorithms – Theory and Application, dostupno na internet adresi: http://www.it-weise.de/
[5] Randy L. Haupt, Sue Ellen Haupt, Practical Genetic Algorithms, John Wiley & Sons, 2004
[6] Michael D. Vose, Alden H. Wright, The Simple Genetic Algorithm and the Walsh Transform: part I, The Theory, Computer Science Dept. The University of Tennessee, Computer Science Dept.The University of Montana
[7] Michael D. Vose, Alden H. Wright, The Simple Genetic Algorithm and the Walsh Transform: part II, The Inverse, Computer Science Dept. The University of Tennessee, Computer Science Dept.The University of Montana
[8] John J. Grefenstette, Deception Considered Harmful, Navy Center for Applied Research in Artificial Intelligence
[9] L. Darrell Whitley, Fundamental Principles of Deception in Genetic Search, Dapartment of Computer Science, Colorado State University
[10] Jeffrey Horn, David E. Goldberg, Genetic Algorithm Difficulty and the Modality of Fitness Landscapes, Illinois Genetic Algorithms Laboratory, University of Illinois u Urbana-Champaign, 1994
[11] Dipankar Dasgupta, Handling Deceptive Problems Using a different Genetic Search
[12] David E. Goldberg, Kalyanmoy Deb, Jeffrey Horn, Massive Multimodality, Deception, and Genetic Algorithms, Illinois Genetic Algorithms Laboratory, University of Illinois u Urbana-Champaign, 1992
[13] Mitchell Melanie, An Introduction to Genetic Algorithms, The MIT Press, 1999
[14] Hillol Kargupta, Ka.lyanmoy Deb, David E. Goldberg, Ordering genetic algorithms and deception, Illinois Genetic Algorithms Laboratory, University of Illinois u Urbana-Champaign
[15] Stephanie Forrest, Melanie Mitchell, What Makes a Problem Hard for a Genetic Algorithm? Some Anomalous Results and Their Explanation, Department of Computer Science, University of New Mexico, Artificial Intelligence Laboratory, University of Michigan, 1993
[16] Gunar E. Liepins, Michael D. Vose, Deceptivness and Genetic Algorithm Dynamics, Computer Science Dept. The University of Tennessee, Oak Ridge National Laboratory
[17] B. Naudts, A. Verschoren,Epistasis and Deceptivity
[18] Luis R. Lopez, Are Deception and Complexity Conjugate Variables in Genetic Learning?, U.S. Amy Space and Strategic Defense Command, Computer Resources Engineering office
[18] Jianwu Li, Minqiang Li,An Improved Genetic Algorithm for Solving Deceptive Problems, Department of Computer Science and Engineering Beijing Institute of Technology, School of Management Tianjin University