

***SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA***

***PROJEKT
Algoritmi zasnovani na evolucijskom računanju***

UMJETNI IMUNOLOŠKI SUSTAVI

Krešimir Đuretec
Voditelj: Doc. dr. sc. Marin Golub

Zagreb, Prosinac, 2008.godina

1.	Uvod	1
2.	Algoritam klonske selekcije	2
2.1	Inspiracija	2
2.2	Algoritam	2
2.3	Raspoznavanje uzoraka	3
2.4	Optimizacijski problemi	5
2.4.1	Optimizacija višedimenzionalne funkcije	5
2.4.2	Kombinatorička optimizacija (Problem trgovačkog putnika)	7
3.	Eksperimentalni rezultati	8
3.1	Parametar n	8
3.2	Parametar β	9
3.3	Parametar d	9
3.4	Parametar ϕ	10
4.	Umjetna imunološka mreža (aiNet).....	11
4.1	Inspiracija	11
4.2	Algoritam	12
5.	Primjena klonske selekcije i umjetne imunološke mreže	14
6.	Zaključak	15
7.	Literatura	16

1. Uvod

Čovjek je, kao i svaki drugi živi organizam, svakodnevno izložen opasnim virusima, bakterijama i ostalim štetnim organizmima (dalje antigeni¹) koji bi mogli oštetiti njegovo tijelo. To se ipak ne događa jer posjeduje imunološki sustav. Imunološki sustav je vrlo kompleksan sustav koji se sastoji od mnoštva različitih stanica (B-stanice i T-stanice) koje sprečavaju strana tijela da oštete organizam. Svaki organizam prilikom svog rođenja posjeduje urođeni imunološki sustav. Taj sustav ima veliku ulogu u razvitku cjelovitog imunološkog sustava. Naime, tijekom godina kako je tijelo napadnuto određenim antigenom imunološki sustav nema ulogu samo uništavanja tog antigena, nego već i njegovog pamćenja. Na taj način imunološki sustav efikasno može eliminirati već viđene antigene (ovo je svojstvo da čovjek ne može oboljeti dva puta od iste bolesti npr. vodenih kozica).

Uočavaju se dva vrlo važna svojstva imunoloških sustava : prepoznavanje i pamćenje.

Upravo su ta dva svojstva potaknula znanstvenike da pokušaju primijeniti ideju imunoloških sustava za rješavanje računarskih problema.

Proučavanje umjetnih imunoloških sustava je započelo devedesetih godina 20. stoljeća te je do danas razvijen veliki broj algoritama koji imaju različitu primjenu. Algoritmi se mogu podijeliti, ovisno prema svojstvu imunološkog sustava koje koriste, u četiri cjeline [3]: negativna selekcija, klonska selekcija, imunološke mreže i algoritam dendritskih stanica.(DCA).

U ovom će radu biti opisan algoritam klonske selekcije tj. jedan tip algoritma i umjetna imunološka mreža koja temelji svoj rad na klonskoj selekciji.

¹ iako se cijeli virus ili bakterija ne može zvati antigen

2. Algoritam klonske selekcije

2.1 Inspiracija

Kada je tijelo napadnuto antigenom određeni broj B-stanica odgovara stvarajući antitijela. Svaka B-stanica stvara određeno antitijelo. Stimulirane antigenom B-stanice se dijele i sazrijevaju u plazma stanice. Plazma stanice imaju mogućnost proizvodnje velike količine antitijela. Prilikom stvaranja plazma stanica B-stanica može sazreti i u memorijsku B-stanicu koja može vrlo dugo (preko 20 godina) ostati nepromijenjena u tijelu domaćina. Na taj način se stvara imunološki sustav koji ima mogućnost pamćenja te uz pomoć memorije na ponovnu pojavu antigena reagira vrlo brzo.

Da bi imunološki sustav što brže proizveo antitijela koja će najbolje prepoznavati antigen, koji je i pokrenuo reakciju, B-stanice koje proizvode bolja antitijela su klonirana u većem broju i podložne slabijoj mutaciji, dok B-stanice koje proizvode lošija antitijela su klonirana u manjem broju, ali su podložne snažnijoj mutaciji.

2.2 Algoritam

Prema ovom principu Leonardo N. de Castro i Fernando J. Von Zuben su predložili algoritam koji će moći rješavati probleme raspoznavanja uzoraka i optimizacijske probleme. Algoritam je dobio ime CLONALG [1] (*eng.* CLONal selection ALGORITHM).

Algoritam ima sljedeća svojstva :

1. održavanje određenog memorijskog uzorka
2. selekcija i kloniranje antitijela koja su najviše stimulirana
3. odumiranje nestimuliranih antitijela
4. sazrijevanje prema dobroti
5. selekcija antitijela proporcionalno prema dobroti
6. generiranje i održavanje raznolikosti

Algoritam temelji svoj rad na dvama operatorima: kloniranje i mutacija.

Kloniranje radi na principu umnažanja određene jedinke nekoliko puta. Koliko puta će neka jedinka biti umnožena ovisi o njejoj dobroti.

Mutacija također ovisi o dobroti jedinke. Što je jedinka bolja to je mutacija manja i obrnuto.

Algoritam karakteriziraju dva skupa (populacije). Neka su to populacija A i populacija G. Populacija A predstavlja "antitijela" , dok populacija G predstavlja "antigene". Populacija A se tijekom niza generacija podvrgava kloniranju , mutiranju i selekciji najboljih, a populacija G ostaje nepromijenjena.

Kako je algoritam namijenjen za rješavanje različitih problema (optimizacijskih i problema raspoznavanja uzoraka) tako i populacije A i G imaju različite uloge.

2.3 Raspoznavanje uzoraka

Da bi algoritam bio prilagođen za rješavanje problema raspoznavanja uzoraka skup G predstavlja uzorke koji će se morati moći prepoznati, a skup A predstavlja skup generiranih uzoraka (uzorci koji će moći prepoznati uzorke iz skupa G). Skup A je podijeljen na dva podskupa A_m i A_r . A_m predstavlja memorijski skup tj. skup koji će pamtit uzorke, dok A_r predstavlja ostatak uzoraka iz skup A .

Algoritam u svakoj generaciji uspoređuje uzorke iz skupa A s uzorcima iz skupa G . Najbolje uzorke klonira i mutira. Iz skupa mutiranih uzoraka uzima najbolji uzorak i ako je on bolji od svog "pretka" u skupu A_m tada ga sprema u skup A_m . Na kraju iz skupa A_r zamijeni određen broj najlošijih uzoraka s novim slučajno generiranim uzorcima. Ova zamjena omogućuje stvaranje korisnog materijala za daljnje kloniranje i mutaciju.

ulaz : skup G koje treba moći prepoznavati,

d broj najlošijih jedinki koje će biti zamijenjene ,

n broj uzoraka koji se uzima za kloniranje

izlaz: skup A koji ima mogućnost prepoznavanja skupa G

početak

stvari slučajni početni skup uzoraka A

ponavljaj dok (nije dosegnut određen broj generacija)

za svaki uzorak u skupu G

odredi sličnost (dobrotu) sa svakim uzorkom iz skupa A

proporcionalno sličnosti kloniraj n najboljih uzorka iz skupa A

proporcionalno sličnosti mutiraj generirane klonove

odredi najbolji klon za trenutni uzorak iz G

ako je bolji od postojećeg u skupu A_m

spremi ga u skup A_m

kraj

zamijeni d najlošijih uzoraka iz skupa A_r s slučajno generiranim uzorcima

kraj

kraj

kraj

Slika 2.1 Pseudo kod algoritma za raspoznavanje uzoraka

Ako pretpostavimo da je na prvom mjestu jedinka koja ima najveću sličnost tada prema formuli (2.1) možemo lako generirati broj klonova proporcionalan sličnosti pojedinog uzorka.

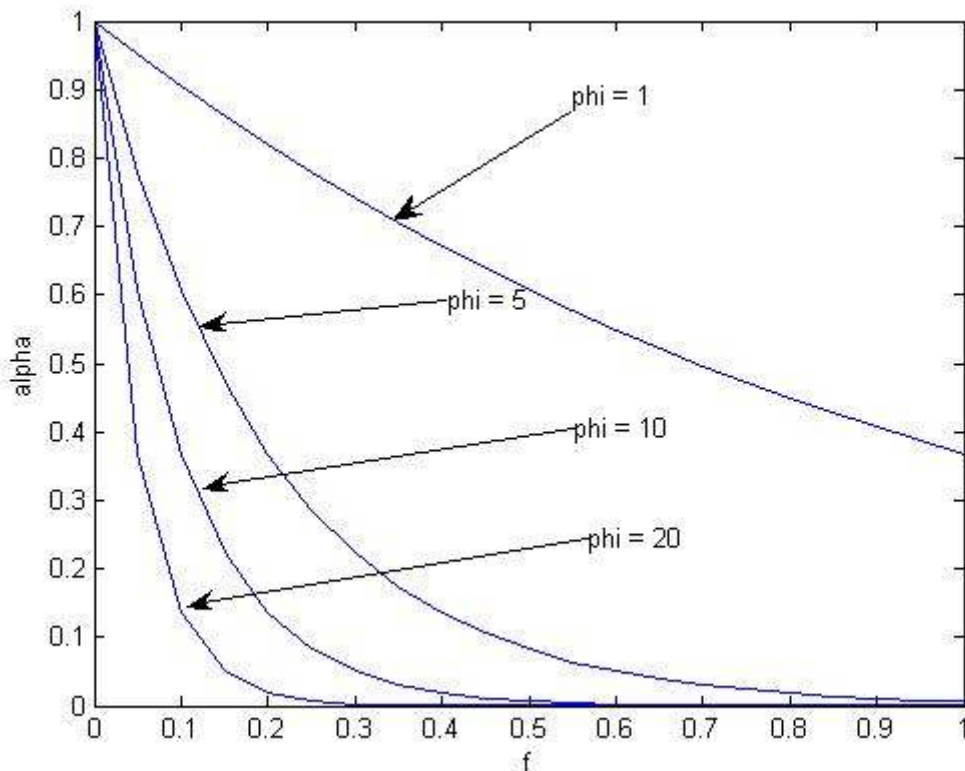
$$N_c = \sum_{i=1}^n \text{round}\left(\frac{\beta \cdot N}{i}\right) \quad (2.1)$$

N_c je ukupan broj generiranih klonova, β konstantni faktor, N ukupan broj uzoraka u skupu A .

Kako je i vjerojatnost mutacije određena dobrotom uzorka predložena je formula (2.2). f predstavlja sličnost uzorka predstavljenu realnim brojem.

$$\alpha = e^{(-\varphi \cdot f)} \quad (2.2)$$

Vrijednost f je skalirana na interval $[0,1]$.



Slika 2.2 Ovisnost α o parametru φ

2.4 Optimizacijski problemi

Postoji nekoliko vrsta optimizacijskih problema. U ovom radu biti će obrađena dva problema: optimizacija funkcije i optimizacija kombinatoričkog problema.

Problem optimizacije funkcije zahtijeva pronalaženje optimuma (maksimuma ili minimuma) funkcije. Kod kombinatoričkog problema također je potrebno pronaći optimum, ali problem nije zadan kao funkcija² nego već kao svojstvo npr. najkraći put u problemu trgovačkog putnika.

Algoritam iz prethodnog poglavlja vrlo je lako prilagoditi za optimizacijske probleme. Više nije potreban skup G jer ne postoji skup koji treba biti prepoznat. Također skup A više ne mora biti podijeljen na dva dijela. Sada svaki uzorak iz skupa A predstavlja moguće rješenje pa nije potrebno kloniranje prema dobroti ,nego je dovoljno prema formuli (2.3) . Vjerojatnost mutacije ostaje prema formuli (2.2) s napomenom da je f i dalje normaliziran na interval $[0,1]$.

$$N_c = \sum_{i=1}^n \text{round}(\beta \cdot N) \quad (2.3)$$

Ovisno o selekciji n najboljih jedinki (koje će biti spremljene u skup A) postoje dvije vrste algoritma : CLONALG1 i CLONALG2.

CLONALG1 : u generaciji t svaka jedinka će biti zamijenjena s najboljom jedinkom iz skupa mutiranih klonova $\beta \cdot N$.

CLONALG2: u populaciju u generaciji $t+1$ će biti ubačeno n najboljih klonova iz mutirane populacije u generaciji t .

CLONALG1 i CLONALG2 mogu biti primijenjeni na oba problema.

2.4.1 Optimizacija višedimenzionalne funkcije

Za optimizaciju višedimenzionalne funkcije korišten je CLONALG1. Naime, kod optimizacije cilj je pronaći uz globalni optimum što više lokalnih optimuma, a za to je korisno svojstvo algoritma CLONALG1.

Također u formuli (2.3) može se umjesto n uvrstiti N te će tako sve jedinke iz skupa A biti klonirane i mutirane. Za parametar d (broj najlošijih jedinki za zamjenu) može se uvrstiti $d=0$. Ako se želi da u generaciji jedinka ne može biti zamijenjena lošijom jedinkom tada se prilikom mutiranja jednu jedinku ostavi nemutiranu. Na taj način se ostavlja mogućnost da se jedinka – roditelj ponovno vrati u skup A , ako kloniranje i mutacija nisu postigli poboljšanje.

² u pozadini je funkcija , ali nju je vrlo teško pronaći

ulaz : funkcija f kojoj je potrebno naći optimume

d broj najlošijih jedinki koje će biti zamijenjene ,

n broj uzoraka koji se uzima za kloniranje

izlaz: skup A optimumi funkcije

početak

stvari slučajni početni skup uzoraka A

ponavlja dok (nije dosegnut određen broj generacija)

odredi vrijednost (dobrotu) funkcije za svaki uzorak $u \in A$

kloniraj n najboljih uzorka iz skupa A

proporcionalno dobroti mutiraj generirane klonove

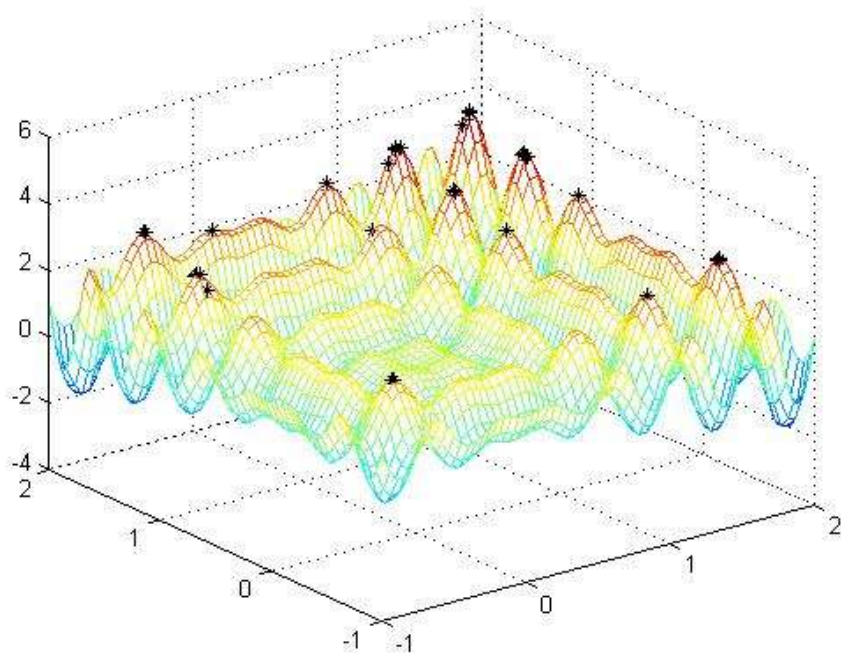
u skupu A određenu jedinku zamijeni s pripadajućim najboljim mutiranim klonom

zamijeni d najlošijih uzoraka iz skupa A sa slučajno generiranim uzorcima

kraj

kraj

Slika 2.3 Pseudo kod optimizacijskog algoritma CLONALG1



Slika 2.4 Optimizacija funkcije $f = x \cdot \sin(4\pi x) - y \sin(4\pi y + \pi) + 1$

2.4.2 Kombinatorička optimizacija (Problem trgovačkog putnika)

Problem trgovačkog putnika spada u skupinu problema koji imaju NP vrijeme izvođenja te su zbog toga ne rješivi u realnom vremenu za obične algoritme.

Za rješavanje problema trgovačkog putnika primijenjen je CLONALG2, jer sada nije cilj pronaći više lokalnih optimuma, nego jedan globalni. Upravo nam to omogućuje svojstvo algoritma CLONALG2 jer iz generacije u generaciju ubacuje u populaciju najbolje jedinke koje je stvorio kloniranjem i mutiranjem.

ulaz : svojstvo koje je potrebno optimirati (najkraća ruta)

d broj najlošijih jedinki koje će biti zamijenjene ,

n broj uzoraka koji se uzima za kloniranje

izlaz: skup A optimumi funkcije

početak

stvari slučajni početni skup uzoraka A

ponavljaj dok (nije dosegnut određen broj generacija)

odredi vrijednost (dobrotu) funkcije za svaki uzorak u A

kloniraj n najboljih uzorka iz skupa A

proporcionalno dobroti mutiraj generirane klonove

ubaci u skup A n najboljih klonova

zamijeni d najlošijih uzoraka iz skupa A s slučajno generiranim uzorcima

kraj

kraj

2.5 Pseudokod optimizacijskog algoritma CLONALG2

3. Eksperimentalni rezultati

Rad algoritma ovisi o parametrima algoritma.

Parametri algoritma :

1. N - veličina populacije
2. n - broj najboljih jedinki za kloniranje
3. β - faktor kloniranja
4. d - broj najlošijih jedinki za zamjenu
5. φ - faktor mutacije
6. generacija - broj generacija

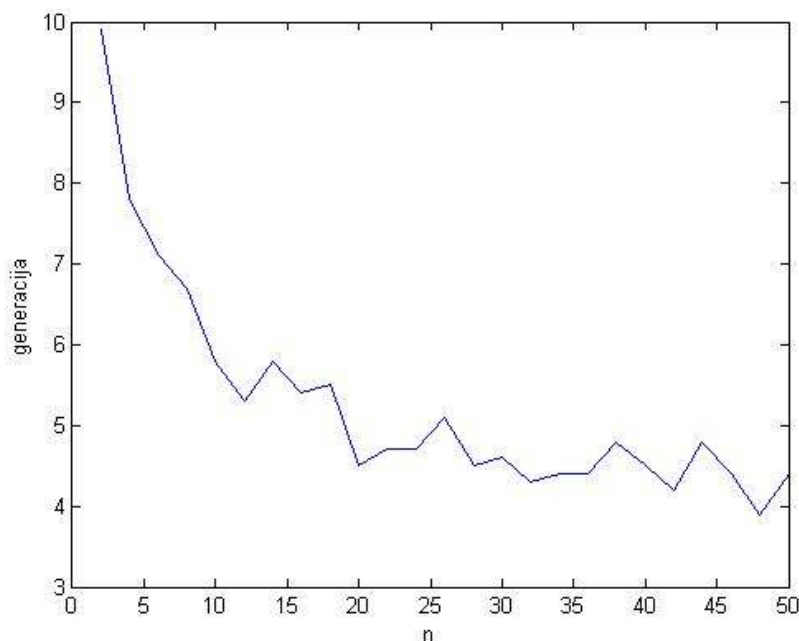
Razumljivo je da povećanje faktora N izravno povećava mogućnost algoritma da brže pronađe optimum, ali također i povećava procesorsko vrijeme računanja. Parametri n , β , d , φ su važni parametri jer utječu na algoritam da pronađe više optimuma, ili brže jedan globalni optimum. Ti parametri (neki od njih) također utječu na procesorsko vrijeme pa ih je potrebno vrlo precizno postaviti da bi se dobio željeni rezultat u zadovoljavajućem vremenu.

Testirani su parametri n , β , d , φ na optimizacijskom problemu pronalaženja optimuma (maksimuma) funkcije (3.1) (Slika 2.4) .

$$f = x \cdot \sin(4\pi x) - y \sin(4\pi y + \pi) + 1 \quad (3.1)$$

3.1 Parametar n

Parametar n označava broj najboljih jedinki koje se selektiraju za kloniranje i mutiranje. Budući da parametar n utječe na procesorsko vrijeme (što je veći n algoritam ima više posla za odraditi u jednoj generaciji) , potrebno je ustanoviti kako parametar utječe na brzinu otkrivanja globalnog optimuma.

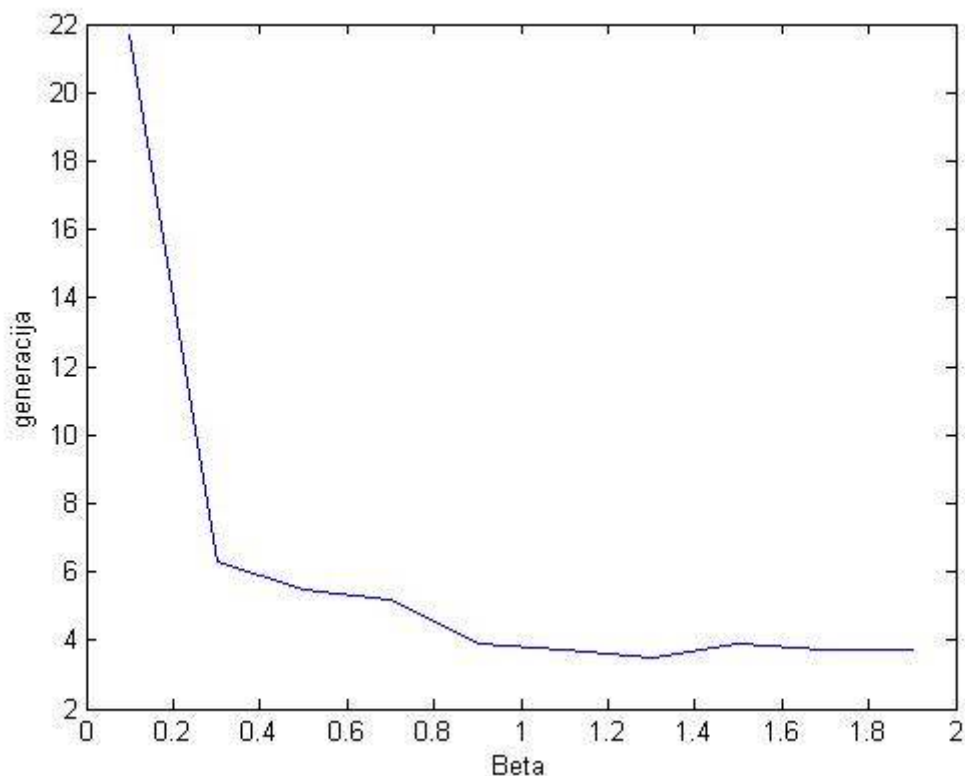


Slika 3.1 Utjecaj n na brzinu konvergencije algoritma

Kada je $n \ll N$ tada algoritam ima problema s pronalaženjem globalnog optimuma. Također znatnije povećanje iznosa n ne pridonosi brzini konvergencije algoritma, ali zato usporava rad algoritam.

3.2 Parametar β

Parametar β utječe na broj klonova pojedine jedinke. Na taj način utječe i na procesorsko vrijeme izvođenja algoritma.



Slika 3.2 Utjecaj parametra β na brzinu konvergencije algoritma

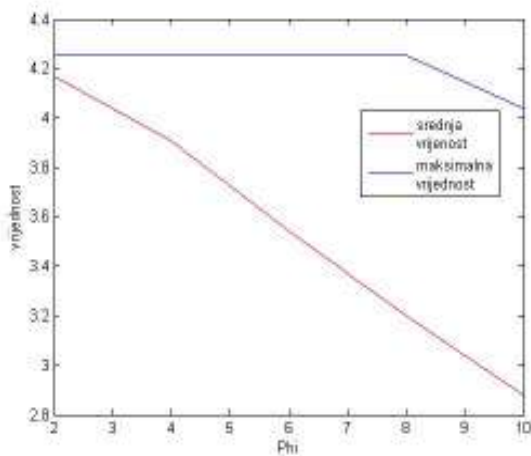
Iz slike 3.2 se uočava da pri vrlo malim iznosima parametra β algoritam ne pronalazi globalni optimum, dok povećanje parametra iznad 1 ne pridonosi poboljšanju konvergencije.

3.3 Parametar d

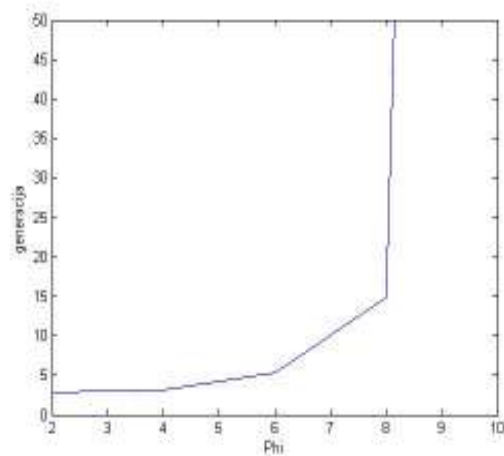
Parametar d omogućuje unošenje novih slučajno generiranih jedinki u algoritam. Za vrijednost parametra se uzima $[0, 20]$ % cijele populacije (N). Jasno je da preveliki iznosi parametra d pretvaraju algoritam u algoritam slučajnog pretraživanja prostora rješenja, a to nije cilj postići.

3.4 Parametar φ

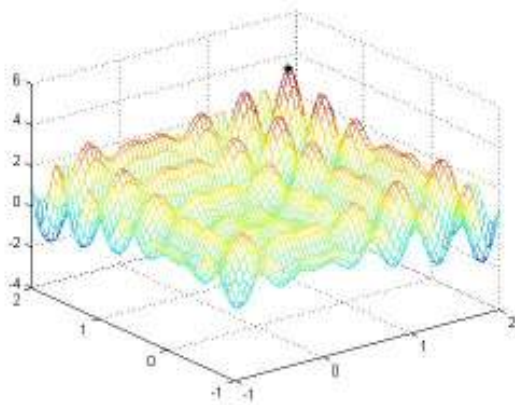
Parametar φ je vrlo važan parametar jer utječe na mutaciju. Kako se vidi i prema slici 2.2 što je veći parametar to se smanjuje vjerojatnost mutacije za bolje jedinke. To svojstvo nam omogućuje da algoritam prilagodimo po želji za brzo otkrivanje globalnog optimuma ili za nešto sporije otkrivanje više lokalnih optimuma. Također je jasno da preveliki iznosi parametra φ onemogućuju algoritam u pronalazanju globalnog optimuma(kao što se vidi i sa slike 3.4).



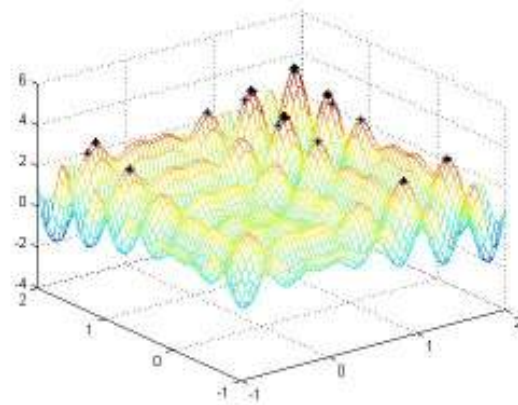
Slika 3.3 Ponašanje populacije



Slika 3.4 Brzina konvergencije



Slika 3.5 Parametar $\varphi = 2$



Slika 3.6 Parametar $\varphi = 7$

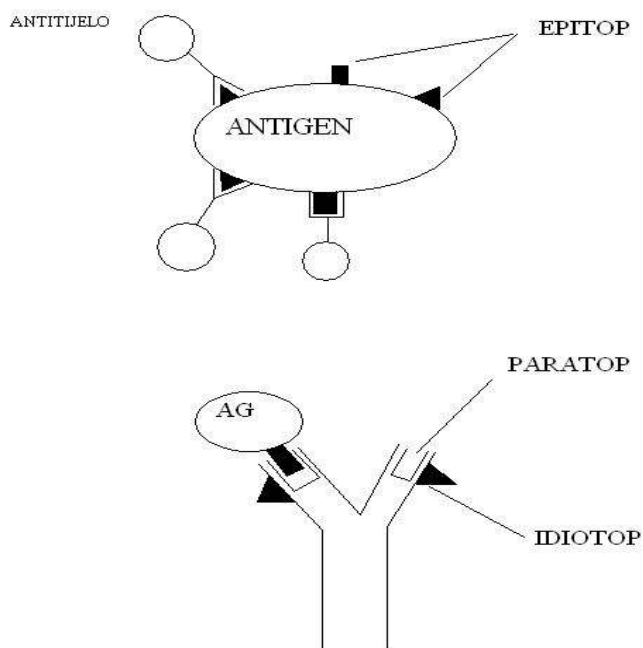
Sa slika 3.5 i 3.6 se uočava razlika u broju otkrivenih optimuma za različite vrijednosti parametra φ . Također se na slici 3.3 uočava ovisnost srednje vrijednosti cijele populacije o parametru φ .

4. Umjetna imunološka mreža (aiNet)

4.1 Inspiracija

Prema teoriji imunološke mreže (Niels K. Jerne 1974. godine) imunološki sustav se sastoji od niza stanica (B-stanice) i molekula (antitijela). Antitijela ne prepoznaju samo antigene, nego već mogu prepoznati i neka druga antitijela u tijelu.

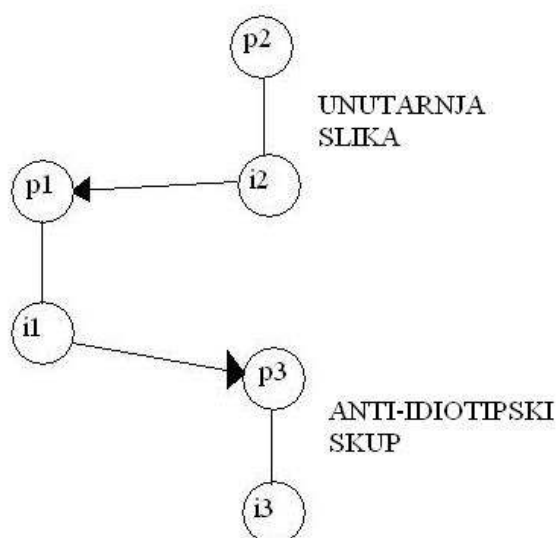
Antitijelo prepoznaje dio antigena koji se naziva epitop (*eng. epitope*) (Slika 4.1). Također na tijelu antitijela se nalazi nekoliko (jedinostvenih) epitopa koje mogu prepoznati druga antitijela. Ti epitopi se nazivaju idiotopi (*eng. idiotope*) (Slika 4.1). Dio antitijela koji je zadužen za prepoznavanje antigena se naziva paratop (*eng. paratope*) (Slika 4.1 **Error! Reference source not found.**).



Slika 4.1 Građa antitijela

Slika 4.2 prikazuje reakciju imunološke mreže na antigen. Kada se u tijelu pojavi antigen njegovi epitopi su prepoznati s nizom različitih paratopa (p_1). Također ti paratopi mogu prepoznati skup idiotopa (i_2) koji se naziva unutarnja slika antigena. Skup idiotopa i_1 prepoznaje skup paratopa p_3 . Taj skup se naziva antiidotopski skup. Takva interakcija se nastavlja dalje.

Kada ne postoji antigen koji bi stimulirao mrežu, mreža se nalazi u ravnotežnom stanju jer svaki skup paratopa kontrolira skupa idiotopa koje prepoznaje.



Slika 4.2 Reakcija imunološke mreže

Prilikom pojave antigena određeni skup paratopa koji prepoznaju antigene je smanjen. To automatski aktivira dvije akcije. Unutarnja slika antigena više nije blokirana i ima mogućnost povećanja broja antitijela. Antiidotopski skup je manje stimuliran (jer je manji broj paratopa koje može prepoznati) pa se njegova koncentracija povećava te je zbog toga uklonjen (ne potpuno) od nekog drugog skupa antitijela.

Prema tome imunološki sustav može reagirati ili pozitivno (prepoznat je antigen) ili negativno (nema antigena). Pozitivna reakcija dovodi do rasta broja određenih stanica i izlučivanja antitijela, dok negativna reakcija dovodi do suspenzije i dovođenja mreže u stabilno stanje.

4.2 Algoritam

Umjetna imunološka mreža (aiNet[4]) je težinski graf koji se sastoji od čvorova (antitijela) i rubova s težinama. Nakupine čvorova (*eng.* clusters) će predstavljati unutarnju sliku postojećih nakupina podataka. Algoritam će koristiti određenu metriku da stvori graf kojemu će težine predstavljati sličnost (udaljenost) između pojedinih čvorova. Broj čvorova unutar nakupine će biti puno manji od broja podataka unutar te nakupine.

ulaz : skup G koji treba moći prepoznavati,

n broj uzoraka koji se uzima za kloniranje,

ξ postotak uzimanja odraslih klonova nakon klonske selekcije

σ_d granica smrti uzorka

σ_s granica suspresije

izlaz: skup M koji ima mogućnost prepoznavanja skupa G

skup S skup sličnosti između pojedinih uzoraka u skupu M

početak

stvari slučajni početni skup uzoraka A

za svaki uzorak u skupu G

odredi sličnost (dobrotu) sa svakim uzorkom iz skupa A

proporcionalno sličnosti kloniraj n najboljih uzorka iz skupa A

proporcionalno sličnosti mutiraj generirane klonove

odaberi ξ % najboljih uzoraka i spremi ih u skup M_j

odredi sličnosti (skup S) između uzoraka u skupu M_j

eliminiraj one uzorke (iz M_j) čija sličnost s nekim uzorkom prelazi granicu σ_s (1)

spoji skup M_j s M

eliminiraj one uzorke (iz M) čija sličnost s nekim uzorkom prelazi granicu σ_s (2)

zamijeni d najlošijih uzoraka iz skupa A s slučajno generiranim uzorcima

kraj

kraj

U algoritmu postoje dvije suspresije klonska (1) i mrežna (2) Klonska suspresija uklanja previše identične uzorke u pojedinoj nakupini podataka, dok mrežna suspresija traži sličnosti između pojedinih nakupina. Na taj način se iz grafa uklanjaju redundantni podaci.

5. Primjena klonske selekcije i umjetne imunološke mreže

Algoritam klonske selekcije je prvenstveno namijenjen problemu raspoznavanja uzoraka, no kako posjeduje određena korisna svojstva vrlo lako ga se može prilagoditi i za optimizacijske probleme.

Umjetna imunološka mreža koristi grupiranje podataka u grozdove te ju to čini pogodnom za rješavanje problema raspoznavanja uzoraka, donošenje odluka i strojno učenje, dubinska analiza podataka, segmentacija slike, oporavljanje dokumenata. Također je primjenjiva i na optimizacijske probleme.

Tablica 5.1 Primjene klonske selekcije i umjetne imunološke mreže

KLONSKA SELEKCIJA	<ul style="list-style-type: none">- raspoznavanje uzoraka- optimizacija funkcije- kombinatorička optimizacija
UMJETNA IMUNOLOŠKA MREŽA	<ul style="list-style-type: none">- grupiranje podataka u grozdove- optimizacijski problemi- pronalaženje znanja- dubinska analiza podataka- uklanjanje redundancije

6. Zaključak

Iako su umjetni imunološki sustavi relativno novo područje istraživanja unutar područja evolucijskog računanja dosada su postignuti značajni rezultati. Umjetni imunološki sustavi ne predstavljaju samo jedan algoritam, nego čitavu skupinu algoritama koji temelje svoj rad na prirodnom imunološkom sustavu.

U radu su predstavljena dva takva algoritma. Prvi algoritam (CLONALG) spada u skupinu algoritama koji svoj rad temelje na klonskoj selekciji, dok drugi algoritam temelji svoj rad na imunološkoj mreži.

Analizom CLONALG-a pokazano je da ga se iako je namijenjen za probleme raspoznavanja uzoraka može vrlo jednostavno prilagoditi za rješavanje optimizacijskih problema. Na tom području (posebice optimizaciji funkcija) se pokazao sposobnim pronaći optimum u zadovoljavajućem vremenu. Također je pokazan utjecaj parametara algoritma na njegovo izvođenje. Iz dobivenih rezultata se može zaključiti da loše podešeni parametri algoritam značajno degradiraju rad samog algoritma.

Na kraju ovog rada se može zaključiti da su umjetni imunološki sustavi s razlogom posebno poglavlje unutar evolucijskog računanja i da uz daljnja istraživanja obećavaju dobre rezultate.

7. Literatura

- [1] Leandro N. de Castro, Fernando J. Von Zuben, Learning and Optimization Using the Clonal Selection Principle , IEEE Transactions on Evolutionary Computation, Special Issue on Artificial Immune Systems, vol. 6, n. 3, pp. 239-251, 2002. god.
- [2] Leandro N. de Castro, Fernando J. Von Zuben, Artificial immune systems Part I Basic theory and applications, 1999. god.
- [3] AISWeb, <http://www.artificial-immune-systems.org/>
- [4] Leandro N. de Castro, Fernando J. Von Zuben, aiNet: An Artificial Immune Network for Dana Analysis
- [5] Artificial Imunne Systems , <http://ais.cs.memphis.edu/>
- [6] Vincenzo Cutello, Giuseppe Narzisi, Giuseppe Nicosia, Mario Pavone, Clonal Selection Algorithms:A comparative Case Study Using Effective Mutation Potentials , 4th Int. Conference on Artificial Immune Systems, ICARIS 2005, August 14-17, 2005, Banff, Canada. Springer, LNCS 3627:13-28, 2005.