

Neizrazito, evolucijsko i neuroračunarstvo: zadatak 4

Tema ovog zadatka je izrada evolucijskog algoritma.

Opis problema

Za neki sustav poznata je njegova prijenosna karakteristika:

$$f(x, y) = \sin(\beta_0 + \beta_1 \cdot x) + \beta_2 \cdot \cos(x \cdot (\beta_3 + y)) \cdot \frac{1}{1 + e^{(x - \beta_4)^2}}$$

gdje su x i y ulazi u sustav a β_0 do β_4 parametri sustava koji korisniku unaprijed nisu poznati. Stoga je korisnik jedno vrijeme promatrao rad jednog takvog sustava te je prikupio 250 mjerenja (za neki x i y , zabilježio je što je bilo na izlazu sustava). Ta mjerenja dostupna su u tekstovnoj datoteci `zad4-dataset1.txt` koja je dostupna u repozitoriju na Ferku. U svakom retku te datoteke nalazi zapis jednog mjerenja (vrijednost od x , vrijednost od y te izmjeren izlaz sustava; vrijednosti su razdvojene znakom tab). Ova mjerenja izmjerena su vrlo preciznim instrumentom.

Datoteka `zad4-dataset2.txt` sadrži mjerenja obavljena nad istim sustavom ali malo manje kvalitetnim instrumentom; izmjereni podatci za izlaz sustava sadrže određen šum.

Zadatak

Uporabom genetskog algoritma pronađite vrijednosti parametara β_0 do β_4 . Napišite dvije inačice genetskog algoritma:

- kanonski generacijski genetski algoritam (omogućite kontrolu elitizma preko booleove zastavice – želimo li elitizam ili ne)
- kanonski eliminacijski genetski algoritam (jednostavna troturnirska selekcija: biramo tri jedinke, dvije bolje su roditelji, nastalo dijete zamjenjuje preostalog/trećeg najlošijeg roditelja).

Omogućite jednostavno definiranje potrebnih parametara algoritma poput veličine populacije, vjerojatnosti mutacije te drugih parametara (pokušajte pronaći dobar skup tih parametara).

Pri radu algoritma, svaki puta kada pronađete novo najbolje rješenje, ispišite ga (zajedno s informacijom o broju trenutne generacije te kvaliteti/kazni) kako biste lakše pratili napredak.

Isprobajte rad nad oba skupa podataka.

Da biste provjerili koliko je neki skup parametara sustava dobar, za svaki zapis iz skupa podataka izračunajte koju vrijednost ti parametri predviđaju za izlaz, oduzmite to od stvarno izmjerenog izlaza, kvadrirajte tu razliku i potom sumirajte tako dobiveni broj za sve uzorke. Konačnu sumu podijelite s ukupnim brojem uzoraka. Tako izračunata mjera zove se *srednja kvadratna pogreška*. Nju koristite kao mjeru lošće (odnosno funkciju kazne).

Prilikom rješavanja zadatka pretpostavite da se svi parametri nalaze negdje između -4 i +4.

Kao reprezentaciju rješenja koristite polje decimalnih brojeva te prikladne operatore križanja i mutacije. Sve nespecificirano odredite samostalno.