

16.7.2010.

Ovaj dokument ukratko opisuje sadržaj i način korištenja programskog paketa načinjenog u okviru iProjekta "Korisnički prilagođeni postupci raspoređivanja".

Projekt pokazuje primjenu genetskog programiranja (GP) u strojnom učenju (evoluciji) heuristika raspoređivanja prilagođenih zadanom okruženju i kriteriju.

Više informacija o projektu može se pronaći na <http://gp.zemris.fer.hr/iprojekt/>

Sadržaj:

1. Sadržaj podatkovne strukture.....	1
2. Podržana okruženja i kriteriji raspoređivanja.....	2
2.1 Opis okruženja i kriterija raspoređivanja.....	2
2.2 Zadavanje željenog okruženja i kriterija izrade rasporeda.....	4
2.3 Definiranje dostupnih podataka u procesu učenja	4
2.4 Stvaranje novog skupa ispitnih primjera.....	4
3. Korištenje programskih komponenti.....	4
3.1 Postupak učenja (evolucije) GP heuristike	5
3.2 Postupak ispitivanja GP heuristike.....	5
4. Korištenje izvornih tekstova programa	6
4.1 Učenje uz pomoć paketa <i>OpenBeagle</i>	7
4.2 Učenje uz pomoć paketa <i>ECF</i>	7
4.3 Evaluacija.....	7
4.4 Pomoćni programi.....	7

1. Sadržaj podatkovne strukture

Paket sadrži sljedeće direktorije:

- **dokumentacija:** pismeni opis problema, metodologije rješavanja i pregled rezultata i literature
- **ucenje:** programi za učenje (evoluciju) heuristika raspoređivanja uz pomoć genetskog programiranja (GP)
- **evaluacija:** programi za ocjenu učinkovitosti postojećih i evoluiranih heuristika raspoređivanja
- **okruzenja:** gotove konfiguracijske datoteke za opis pojedinih okruženja i kriterija raspoređivanja te izvršni program za generiranje ispitnih primjera
- **projekti:** programska okruženja za razvijene komponente

Način korištenja programa u direktorijima *ucenje* i *evaluacija* opisan je u poglavlju 3.

2. Podržana okruženja i kriteriji raspoređivanja

U ovom projektu do sada je ostvarena podrška za nekoliko okolina raspoređivanja, navedenih u nastavku.

Okolina jednog stroja (oznaka *single*)

- statička ili dinamička (s vremenima pripravnosti poslova)
- sa ili bez trajanja postavljanja između poslova
- sa ili bez ovisnosti u redoslijedu poslova

Okolina paralelnih jednolikih strojeva (oznaka *uniform*)

- statička ili dinamička
- sa ili bez trajanja postavljanja između poslova

Okolina paralelnih nesrodnih strojeva (oznaka *unrelated*)

- statička ili dinamička

Okolina proizvoljne obrade (oznaka *jobshop*)

- statička ili dinamička

Osnovni kriteriji ocjene rasporeda

- težinsko zaostajanje (Twt)
- težinska zakašnjelost (Uwt)
- ukupna duljina rasporeda (Cmax)
- trajanje protjecanja (Fwt)

U postupku učenja odnosno ocjenjivanja moguće su kombinacije navedenih okruženja i kriterija.

2.1 Opis okruženja i kriterija raspoređivanja

Svako okruženje opisano je skupom datoteka i zadano određenim brojem *ispitnih primjera*, odnosno skupova poslova koje je potrebno rasporediti. Datoteke koje opisuju pojedino okruženje su:

- *fitness.txt* (obavezna): sadrži parametre okruženja i postupka učenja
- *fitness_deadline.txt* (obavezna): sadrži željena vremena završetka svih poslova (engl. *due date*)
- *fitness_duration.txt* (obaveza): sadrži trajanja svih poslova (ili aktivnosti unutar poslova)
- *fitness_SP.txt* (obavezna): sadrži zbroj trajanja svih poslova za sve ispitne primjere

- *fitness_weight.txt* (obavezna): sadrži težine (važnosti) svih poslova
- *fitness_ready.txt* (opcionalna): sadrži vremena pripravnosti svih poslova
- *fitness_constraints.txt* (opcionalna): sadrži ograničenja u redoslijedu poslova

Datoteka *fitness.txt* sadrži sljedeće parametre koji definiraju okruženje raspoređivanja:

- okruženje raspoređivanja, zapisano jednim parametrom: single | uniform | unrelated | jobshop
- broj skupova (ispitnih primjera): sets <broj>
- najveći broj poslova u nekom primjeru: max_jobs <broj>
- najveći broj strojeva u nekom primjeru: max_machines <broj>
- najdulje trajanje nekog posla: max_length <broj>
- brojevi poslova po skupovima: jobs <broj> <broj> ...
- brojevi strojeva po skupovima: machines <broj> <broj>
- postoje li trajanja postavljanja među poslovima: setup 1|0
- postoje li vremena pripravnosti: ready <datoteka>
- postoje li ograničenja u redoslijedu: constraints <datoteka>
- datoteka s trajanjima: duration <datoteka>
- datoteka sa željenim vremenima završetka: deadline <datoteka>
- datoteka s ukupnim trajanjima po skupovima: SP <datoteka>
- datoteka s težinskim faktorima za kriterije: weight_<kriterij> <datoteka>
- kriteriji: protjecanje (F), zakasnjelost (T), preuranjenost (E), prekoracenje (N)

Ista datoteka također definira kriterije postupka strojnog učenja:

- kriterij optimizacije: fitness Twt | Nwt | FTwt | Cmax
- aktivni terminali u GP stablu: definiraju se ovisno o okruženju (nazivi i značenje terminala za pojedino okruženje opisani su u dokumentu "*Pregled rezultata za ostvarena okruženja raspoređivanja*")
- uređivanje rješenja (rekurzivna primjena matematičkih pravila u cilju smanjenja prikaza funkcije prioriteta): editing 1|0
- opcija za gradnju rasporeda u postupku ispitivanja: evaluation 0|1

Primjeri sadržaja ovih datoteka za različita okruženja mogu se naći unutar strukture direktorija.

2.2 Zadavanje željenog okruženja i kriterija izrade rasporeda

Prilikom evolucije algoritma raspoređivanja (postupkom učenja), korisnik mora definirati (ili odabrati iz ponuđenih) skup ispitnih primjera za učenje. To podrazumijeva odabir nekog od postojećih skupova primjera za učenje ili definiranje vlastitog skupa.

Okruženje i skup ispitnih primjera jednoznačno su definirani skupom datoteka koji uključuje osnovnu datoteku *fitness.txt* i ostale potrebne datoteke s podacima, ovisno o okruženju. Kako bi se proces učenja mogao provesti, potrebno je odabrani skup datoteka postaviti u isti direktorij kao i program za GP evoluciju (npr. u direktorij **ucenje**).

2.3 Definiranje dostupnih podataka u procesu učenja

U postupku učenja genetskom su programu na raspolaganju varijable koje opisuju trenutno stanje sustava i poslova koji se raspoređuju. Ove varijable ovisne su o okruženju u kojemu se obavlja raspoređivanje. Popis varijabli koje se koriste u određenim okruženjima već je naveden u *fitness.txt* datoteci za odgovarajuće okruženje, a značenje tih varijabli pojašnjeno je u dokumentu "*Pregled rezultata za ostvarena okruženja raspoređivanja*".

Prilikom postupka učenja, korisnik može koristiti samo neke od ponuđenih varijabli, a korištenjem dostupnih izvornih tekstova programa moguće je i definirati nove varijable za neko okruženje.

2.4 Stvaranje novog skupa ispitnih primjera

Stvaranje novog skupa ispitnih primjera moguće je na sljedeći način:

1. definirati vlastito okruženje i svojstva ispitnih primjera u datoteci *fitness.txt* na temelju svojstava opisanih u odjeljku 2.1
2. postaviti datoteku u isti direktorij s izvršnim programom *generiranje.exe* koji se nalazi u direktoriju **okruženja**
3. pokrenuti program *generiranje.exe*

Nakon završetka rada programa, u istom direktoriju stvaraju se datoteke koje opisuju skup ispitnih primjera. Ove datoteke mogu se potom uporabiti u postupku učenja.

3. Korištenje programskih komponenti

Opisana programska podrška namijenjena je za izvođenje u *Windows 2000/XP* operacijskom sustavu. Za korištenje u *Windows Vista/7* okruženju moguća je potreba ponovnog prevođenja programa.

3.1 Postupak učenja (evolucije) GP heuristike

Postupak učenja koristi skup datoteka koje definiraju zadano okruženje i ispitne primjere te izvršni program koji izvodi evoluciju genetskim programiranjem. Izvršni program *gp.exe* se može naći u direktoriju **ucenje**. Postupak učenja je sljedeći:

1. u isti direktorij staviti datoteke *go.bat*, *gp.exe* i *gp.txt* (navedene datoteke nalaze se u direktoriju **ucenje**)
2. u datoteci *gp.txt* podesiti parametre evolucije (velicina populacije, broj generacija...)
3. dodati skup datoteka koje definiraju okruženje i skup ispitnih primjera (*fitness.txt* i pripadajuće datoteke)
4. u datoteci *fitness.txt* opcija *evaluation* mora biti isključena ("evaluation 0")
5. pokrenuti *go.bat*
 - a. tijekom evolucije stvaraju se datoteke *beagle.log* i *beagle.log_hof.txt*
6. na kraju evolucije, najbolja rješenja u XML obliku po generacijama nalaze se u datoteci *beagle.log_hof.txt*

3.2 Postupak ispitivanja GP heuristike

U postupku ispitivanja, evoluirano rješenje (funkcija prioriteta u obliku stabla) uspoređuje se s postojećim heuristikama raspoređivanja za zadano okruženje. Usporedba se može obaviti na istom skupu primjera koji je korišten za učenje, no uputno je učinkovitost odrediti na posebnom skupu za ocjenu. Postupak ocjene moguće se provesti na sljedeći način:

1. u isti direktorij staviti datoteke *test.bat*, *gp.exe*, *test.conf*, *evaluacija.exe* i *ispitivanje.exe* (navedene datoteke nalaze se u direktoriju **evaluacija**)
2. dodati skup datoteka koje definiraju okruženje (*fitness.txt* i pripadajuće datoteke)
3. u datoteci *fitness.txt* mora biti uključena opcija *evaluation* ("evaluation 1")
4. postaviti datoteku s rješenjima dobivenu evolucijskim postupkom (pretpostavljeno ime je *beagle.log_hof.txt*, vidi odjeljak 3.1)
5. pokrenuti program *ispitivanje.exe* i slijediti upute
 - a. program traži ulaznu datoteku s rješenjima
 - b. korisnik unosi generaciju evolucijskog postupka iz koje se uzima najbolja funkcija prioriteta
 - c. program pokreće generiranje rasporeda za zadani skup ispitnih primjera uz zadanu funkciju prioriteta iz datoteke s rješenjima
 - d. program pokreće generiranje rasporeda za postojeće algoritme raspoređivanja i uspoređuje učinkovitost po zadanim kriterijima

Alternativno, umjesto programa *ispitivanje.exe* može se ručno definirati funkcija prioriteta na sljedeći način:

1. postaviti rješenje (evoluirano za istu okolinu) u XML obliku u datoteku *seed.obm*, npr:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<Beagle>
<Seeds>
<Individual>
<Genotype type="gptree" size="29" depth="9"><SUB><SL/><SUB><DIV><w/>
<DIV><SUB><ADD><pt/><w/></ADD><ADD><DIV><SL/><SL/></DIV><SPr/></AD
D></SUB><ADD><SL/><MUL><SL/><DIV><ADD><SL/><pt/></ADD><w/></DIV></
MUL></ADD></DIV></DIV><ADD><DIV><pt/><w/></DIV><w/></ADD></SUB></SU
B></Genotype>
</Individual>
</Seeds>
</Beagle>
```

2. pokrenuti *test.bat*

- a. stvaraju se datoteke *raspored_GP.txt* i *rezultat_GP.txt* s rasporedom i postignutim kriterijima za GP heuristiku

3. pokrenuti *evaluacija.exe*

- b. ispisuju se rezultati i postotci dominacije; stvaraju se datoteke s rezultatima postojećih heuristika za zadano okruženje

3.3 Uporaba funkcije prioriteta u proizvoljnom okruženju

Funkcija prioriteta izvedena uporabom genetskog programa zapisana je u simboličkom obliku u XML formatu. Primjerice, dobivena funkcija

```
<Genotype type="gptree" size="10"
depth="5"><MUL><DIV><pt/><LVL/></DIV><ADD><POS><DIV><pt/><w/></DIV></
POS><pt/></ADD></MUL></Genotype>
```

ekvivalentna je matematičkom zapisu

$$\frac{pt}{LVL} \cdot \left(pos\left(\frac{pt}{w}\right) + pt \right).$$

Projektna programska podrška koristi simbolički zapis za generiranje rasporeda. Za potrebe automatskog raspoređivanja (tj. uporabom računala) u nekom drugom sustavu, ovaj je zapis potrebno implementirati u raspoređivač.

4. Korištenje izvornih tekstova programa

U direktoriju **projekti** nalaze se izvorni tekstovi programa korištenih u izradi projekta. Uz pomoć ovih datoteka moguće je izraditi vlastite inačice izvršnih programa (npr. za različiti operacijski sustav). Sadržaj poddirektorija opisan je u nastavku poglavlja.

4.1 Učenje uz pomoć paketa *OpenBeagle*

U direktoriju **ucenje_GP_OBeagle** nalazi se Microsoft Visual Studio projekt i izvorni tekstovi programa za evoluciju funkcija prioriteta u zadanom okruženju. Ovaj projekt moguće je prevesti uz prethodnu instalaciju besplatnog paketa *Open Beagle*, dostupnog na <http://beagle.gel.ulaval.ca/>.

U izradi projekta korištena je inačica Beagle paketa 2.2.0. U međuvremenu su za dostupne i novije inačice, no u zadnje vrijeme ovaj se razvojni paket nije često obnavljao.

4.2 Učenje uz pomoć paketa *ECF*

U direktoriju **ucenje_GP_ECF** nalazi se Microsoft Visual Studio projekt i izvorni tekstovi programa za evoluciju funkcija prioriteta koji koriste besplatno programsko okruženje *ECF - Evolutionary Computation Framework*. Ovo okruženje razvijeno je od strane autora projekta i studentskih suradnika na matičnoj ustanovi (FER), a moguće ga je dohvatiti na stranici <http://gp.zemris.fer.hr/ecf/>.

Nakon instalacije okruženja ECF, datoteke u navedenom direktoriju potrebno je postaviti u direktorij *examples* paketa ECF te pokrenuti proces prevođenja u Visual Studio programskom paketu.

Prevođenje je moguće izvesti i na drugim operacijskim sustavima, u kojem slučaju je potrebno dodatno definirati skup datoteka koje čine projekt (npr. Eclipse).

4.3 Evaluacija

U direktoriju **evaluacija** nalaze se izvorni tekstovi programa koji obuhvaćaju:

- simulator zadanih okruženja raspoređivanja
- ostvarenje postojećih postupaka raspoređivanja
- upravljački program koji ocjenjuje postojeće postupke i rješenje dobiveno genetskim programiranjem, na temelju zadanog okruženja
- upravljački program koji generira skup ispitnih primjera na temelju zadanog okruženja i parametara ispitivanja

4.4 Pomoćni programi

U direktoriju **pomocni** nalaze se izvorni tekstovi programa koji obavljaju pomoćne operacije za potrebe ocjene učinkovitosti evoluiranih pravila raspoređivanja:

- generiranje rasporeda i ocjena učinkovitosti za jedno rješenje iz zadane generacije postupka evolucije
- generiranje rasporeda i ocjena učinkovitosti za sva rješenja jednog evolucijskog procesa (iz svih generacija)
- generiranje rasporeda i ocjena učinkovitosti za sva rješenja iz iste generacije iz više procesa evolucije (za potrebe ocjenjivanja nakon višekratnog pokretanja postupka učenja)