

Bevezetés a lágy számítás módszereibe

Genetikus algoritmusok
Evolúció, genetika

Werner Ágnes

Villamosmérnöki és Információs Rendszerek Tanszék

e-mail: werner.agnes@virt.uni-pannon.hu

Evolúció

- darwini evolúció, törzsfejlődés ismerete
- egyedek közötti versengés
 - a táplálékszerzés ügyessége
 - a táplálékká válás elkerülésének képessége
 - fajon belüli harc a nőstényekért
 - tűrőképesség mértéke
 - alkalmazkodás képessége
- minden irányítottság nélkül

Genetika

- azonos fajhoz tartozó élőlények nem egyformák
- tökéletesebb utódok
- "két fekete kecskének tarka utódja"
- gének
- szaporodás
- mutáció

1. Gének

- az öröklött tulajdonságokat a gének határozzák meg
- két fontos jellemző:
 1. funkció
 2. lókuszt (hely)
- allélok
- kromoszóma
- kromoszómaszelvény
- a gén egyik allélja domináns
- fenotípus
- genotípus

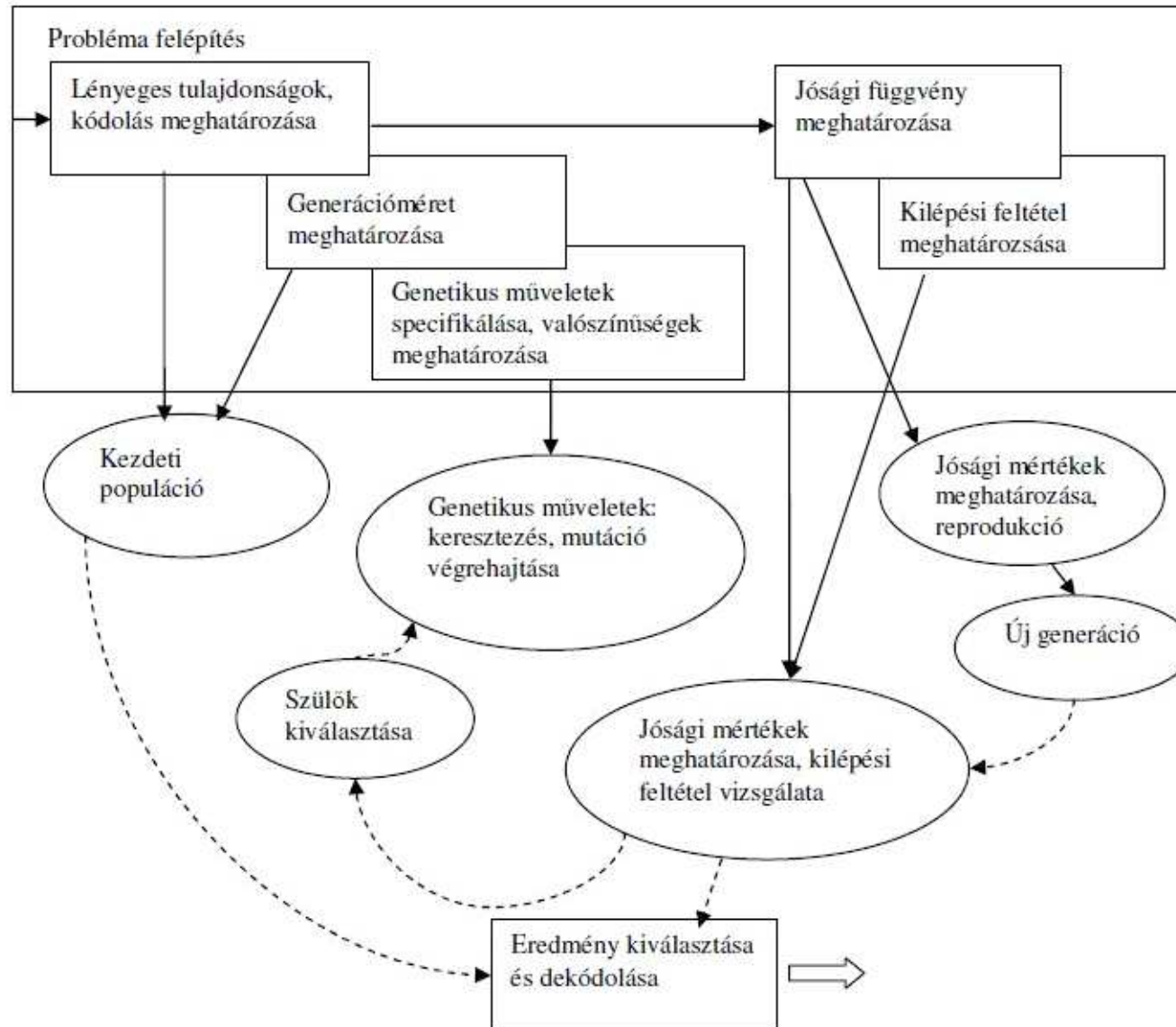
2. Szaporodás

- ivartalan (utódnak egy szülője van, megegyezik a szülővel)
- ivaros (utódnak két szülője van, genetikai anyag keveredik)

3. Mutáció

- megváltozhat egy gén értéke
- kromoszóma részek maradhatnak el
- kromoszóma részek kettőződhetnek meg
- kromoszóma részek fordulhatnak meg
- a mutációt lehet pozitív értelemben is tekinteni → elősegíti a genetikai változatoságot

A genetikus algoritmus működési sémája



Evolúciós algoritmus alapechnikák

1. Hogyan ábrázolható az egyed
2. Milyen rekombinációs és mutációs műveletet alkalmazunk
3. Milyen szelekciós és visszahelyező művelet jöhet szóba
4. Fitness függvény definiálása
5. Egyes feladatoknál a lehetséges megoldásokat feltételekkel korlátozzuk (pl. büntető függvény)
6. Általános paraméterek, az egyes műveletek paramétereinek meghatározása

Az egyedek ábrázolási formája

- **Valós (egész) vektor**

Fenotípus formát jelent

Az egyed tulajdonságait mint valós (egész) értékű változókat adjuk meg és az egyedeket vektorként ábrázoljuk:

$E = (x_1, x_2, \dots, x_n)$, ahol x_i az i -edik tulajdonsághoz tartozó változó.

- **Permutáció**

Adott számú objektumot valamilyen sorrendben el kell helyezni

Permutáció: $(\pi_1, \pi_2, \dots, \pi_n)$ az első n pozitív egész szám permutációja

A sorrend rögzített

Más ábrázolási forma:

$\pi_1, \pi_2, \dots, \pi_n$ sorrend

x_1, x_2, \dots, x_n pozícióhoz rendelt objektum

Fenotípus forma

Az egyedek ábrázolási formája

- **Bináris vektor**

Genotípus formát jelent

Jelölje az (x_1, x_2, \dots, x_n) valós (egész) vektor az egyed tulajdonságait.

Bináris ábrázolásban az egyed egy sztringként jelenik meg:

$x_1, x_2, \dots, x_i, \dots, x_n$

...00|11010111|01... az x_i változó kódolt értékei

Bináris kódoló és dekódoló függvény

- **Standard bináris kódolás**

az x_i változót kettes számrendszerbeli számmá konvertálja és a kapott bitsorozatot egy véges hosszúságú rész sztringre képezi le;

csak egy $[b_i, c_i]$ intervallumba eső, adott pontosságú számot tud ábrázolni;

a másik irányú transzformációhoz jelölje D a dekódoló függvényt,

h_i az i . rész sztring hosszát,

a_{iz} a rész sztring z . pozícióján lévő bit értékét,

az i . rész sztring dekódolt értéke:

$$D(a_{i1}, \dots, a_{in}) = b_i + \frac{c_i - b_i}{2^{b_i} - 1} \left(\sum_{z=1}^{h_i} a_{i(h_i - z + 1)} 2^{z-1} \right) = x_i$$

Bináris kódoló és dekódoló függvény

- **Gray-kódolás**

A standard kódolásnál az egymás melletti számok Hamming távolsága különböző számoknál más és más. Ez a rekombinációnál hibákat okozhat.

A Gray-kódolás kijavítja a Hamming távolságot: egységesen, bármely egymás melletti szám távolsága 1 lesz.

A standard bináris kódból a Gray-kód:

$$g_z = \begin{cases} a_z & \text{ha } z = 1 \\ a_{z-1} \oplus a_z & \text{különben} \end{cases}$$

A Gray-kódból a bináris kód:

$$a_z = \bigoplus_{k=1}^z g_k$$

Szelekció

A populáció átlagos minőségét hivatott javítani. A minőséget a fitness függvénnyel mérjük. A jobb minőségű egyedeket nagyobb valószínűséggel használja a GA az új populáció kialakításához.

A szelekciós műveletek összehasonlítása:

- **szelekciós intenzitás:** a szelekció hatására bekövetkező, a populációk átlagos fitness értékeinek változását mutatja
$$Int = (M^* - M) / \sigma$$
 ahol M^* és M a szelekció előtti és utáni átlagos fitness értékek, σ a fitness értékek szórását jelöli az új populációban
- **változatosság elvesztése:** a populáció azon egyedeinek D aránya, amelyeket nem választott ki a szelekciós művelet
- **szelekciós variancia:** a populációbeli fitness értékek varianciájának változása a szelekció hatására
$$V = (\sigma^*)^2 / \sigma^2$$
 σ és σ^* a fitness értékek szórása a szelekció előtt és után

Rulett szelekció

- Fitnessz arányos szelekció, amely az egyedeket fitnessz értékük abszolút értékének arányában választja ki a szelekciós állományból.
- Visszatevéses művelet
- Egy egyed kiválasztását a szelekciós valószínűség határozza meg:

$$P(E_i) = \frac{f(E_i)}{\sum_{j=1}^n f(E_j)}$$

f a fitnessz függvény, E_i ($i = 1, \dots, n$) az egyedek

Rulett szelekció

- veszünk egy rulettet
- feleltessünk meg minden E_i egyednek valamely kiindulási pontból folyamatosan egy-egy körszeletet
- generálunk egy $[0, 1]$ -beli véletlen számot, a véletlen számot ívhossznak tekintjük
- azt az egyedet választjuk, amelynek körszeletében az ív végződik
- egy μ elemű szelekciós halmazból a választást μ -ször kell megismételni, amíg kialakul a szülők állománya

A kiválasztott egyedek közt $\mu * p(E_i)$ ($i = 1, \dots, n$) várható számú másolata lesz az E_i egyednek.

Sztokasztikus univerzális mintavétel

- Fitnessz arányos szelekció
- Minimalizálja a szelekció során kapott duplikációk számát
- Minden E_i egyedhez a várható másolatok számával azonos hosszúságú körívet rendel:

$$V(E_i) = \mu * p(E_i)$$

Sztochasztikus univerzális mintavétel

Lépések:

1. **Input:** A szelekciós állomány E_i elemei és a hozzá tartozó $V(E_i)$ ($i = 1, \dots, n$) várható másolatok száma.
2. **Output:** A populáció a szelekció után (szülők állománya): E'_i ($i = 1, \dots, n$)
3. $s = 0$; $j = 1$
4. $mutato = Rnd$ (véletlen szám a $[0, 1]$ intervallumból)
5. **for** $i = 1$ **to** μ **do**
6. $s = s + V(E_i)$
7. **while** ($s > mutato$) **do**
8. $E'_j = E_i$; $j = j + 1$; $mutato = mutato + 2r\pi/\mu$
9. **od**
10. **od**

A kiválasztott egyedek közt $\mu * p(E_i)$ ($i = 1, \dots, n$) várható számú másolata lesz az E_i egyednek.

Versengő szelekció

- Az egyedek fitness értékeinek sorrendjét használja fel.
- Nem fog növekedni az egyed duplikációk száma.
- Több egyedből a legjobb fitness értékű egyedet választja ki. (Biológiai szelekciót modellezi.)

Versengő szelekció

Lépések:

1. **Input:** A szelekciós állomány E_i elemei és $f(E_i)$ fitness értékei ($i = 1, \dots, n$),
 $tour$ paraméter
2. **Output:** A populáció a szelekció után (szülők állománya): E'_i ($i = 1, \dots, n$)
3. **for** $i = 1$ **to** μ **do**
4. **for** $k = 1$ **to** $tour$ **do**
5. válasszunk egy $j \in \{1, \dots, n\}$ indexet véletlenszerűen
6. $T_k = E_j$
7. **od**
8. $E'_i = T_j$ ha $f(T_j) = \max(f(T_1), \dots, f(T_{tour}))$
9. **od**

A kiválasztott egyedek közt $\mu * p(E_i)$ ($i = 1, \dots, n$) várható számú másolata lesz az E_i egyednek.

Csonkolásos szelekció

- Csak a legjobb egyedeket választjuk ki.
- A fitness értékek sorrendjét használja fel. (Mesterséges eljárás.)
- Lépések:
 1. **Input:** A szelekciós állomány E_i elemei és $f(E_i)$ fitness értékei ($i = 1, \dots, n$), a $T \in [0, 1]$ korlát
 2. **Output:** A populáció a szelekció után (szülők állománya): E'_i ($i = 1, \dots, n$)
 3. Legyen J a fitness értékek alapján növekvőbe rendezett szelekciós halmaz.
 4. **for** $i = 1$ **to** μ **do**
 5. válasszunk egy $k \in \{[(1 - T) * \mu], \dots, \mu\}$ indexet véletlenszerűen
 6. $E'_i = J_k$
 7. **od**

Lineáris sorrend alapú szelekció

- Minden egyedet a fitness értékeik alapján sorba rendezünk, és sorszámot rendelünk hozzájuk (1 a legrosszabb egyed sorszáma).
- Az egyedek kiválasztását a szelekciós valószínűség határozza meg, amely lineárisan függ az egyed sorszámától.
- $p_1 = \eta^- / \mu$ a legrosszabb egyed szelekciós valószínűsége
 $\eta^- \in [0, 1]$
- μ egyedet kell választani
- $p_\mu = (2 - \eta^-) / \mu$
- η^- / μ és η^+ / μ a legrosszabb és a legjobb egyed kiválasztásának valószínűsége $\eta^+ = 2 - \eta^-$

Lineáris sorrend alapú szelekció

Lépések:

1. **Input:** A szelekciós állomány E_i elemei és $f(E_i)$ fitness értékei ($i = 1, \dots, n$),
 $\eta^- \in [0, 1]$
2. **Output:** A populáció a szelekció után (szülők állománya): E'_i ($i = 1, \dots, n$)
3. Legyen J a fitness értékek alapján növekvőbe rendezett szelekciós halmaz.
4. $S_0 = 0$
5. **for** $i = 1$ **to** μ **do**
6. $S_i = S_{i-1} + p_i$
7. **od**
8. **for** $i = 1$ **to** μ **do**
9. $r = Rnd$; $E'_i = J_z$, ha $S_{z-1} \leq r < S_z$
10. **od**