Genetikus algoritmus – Permutációs kódolású mintafeladat

Egy futárnak öt helyszínt kell meglátogatnia:

A, B, C, D, E

A cél olyan bejárási sorrend meghatározása, amelynél a futár által megtett **teljes távolság a lehető legkisebb**.

A feladat egyszerűsítése érdekében:

- a futár mindig **A helyszínről indul**;
- minden helyszínt pontosan egyszer látogat meg;
- az utolsó helyszínről **nem kell visszatérnie A-ba**.

A helyszínek közötti távolságok

	A	B	C	D	E
A	–	4	7	6	8
B	4	–	3	5	6
C	7	3	–	4	5
D	6	5	4	–	2
E	8	6	5	2	–

Például az

A – B – C – D – E

útvonal hossza:

$$ 4+3+4+2=13 $$

A genetikus algoritmus célja a teljes távolság **minimalizálása**.

1. Kromoszóma-reprezentáció

Ebben a feladatban egy lehetséges útvonal egy kromoszómának felel meg.

Például:

A – C – B – E – D

A kromoszóma génjei tehát a meglátogatandó helyszínek.

Az egyedeket így **permutációs kódolással** reprezentáljuk.

Például:

$$ [A,B,C,D,E] $$

vagy:

\[[A,D,E,C,B] \]

egy-egy lehetséges kromoszóma.

Mivel minden helyszínt pontosan egyszer kell meglátogatni, egy érvényes kromoszómában minden helyszín pontosan egyszer szerepel.

Érvényes egyed

[A,C,B,E,D]

Minden helyszín pontosan egyszer szerepel.

Érvénytelen egyed

[A,C,B,C,D]

Ez nem érvényes, mert:

- C kétszer szerepel,
- E pedig hiányzik.

2. Kezdő populáció

Legyen a kezdő populáció:

1.	[A,B,C,D,E]
2.	[A,C,B,E,D]
3.	[A,D,E,C,B]
4.	[A,E,D,B,C]

A populáció kiértékelése

Az egyedek jóságának (fitnesszértékének) meghatározásához kiszámítjuk az általuk reprezentált útvonal teljes hosszát.

1. egyed

\[[A,B,C,D,E] \]

\[4+3+4+2=13 \]

2. egyed

$$[A,C,B,E,D]$$

$$7+3+6+2=18$$

3. egyed

$$[A,D,E,C,B]$$

$$6+2+5+3=16$$

4. egyed

$$[A,E,D,B,C]$$

$$8+2+5+3=18$$

Az eredmények:

Egyed	Kromoszóma	Úthossz (Fitnessz)
1.	[A,B,C,D,E]	**13**
2.	[A,C,B,E,D]	18
3.	[A,D,E,C,B]	16
4.	[A,E,D,B,C]	18

Mivel **minimalizálási problémáról** van szó, a kisebb érték jelenti a jobb megoldást.

A jelenlegi populáció legjobb egyede:

$$\boxed{[A,B,C,D,E]}$$

13 egységnyi teljes távolsággal.

3. Szelekció

A szelekció feladata annak meghatározása, hogy mely egyedek vehessenek részt a következő generáció létrehozásában.

Használjunk **tornaszelekciót**.

Két-két egyedet összehasonlítunk, és közülük a jobb megoldást választjuk.

Első torna

Az 1. egyed:

$$ [A,B,C,D,E] \rightarrow 13 $$

A 2. egyed:

$$ [A,C,B,E,D] \rightarrow 18 $$

Mivel:

$$ 13<18 $$

az első kiválasztott szülő:

$$ \boxed{P_1=[A,B,C,D,E]} $$

Második torna

A 3. egyed:

$$ [A,D,E,C,B] \rightarrow 16 $$

A 4. egyed:

$$ [A,E,D,B,C] \rightarrow 18 $$

Mivel:

$$ 16<18 $$

a második kiválasztott szülő:

$$ \boxed{P_2=[A,D,E,C,B]} $$

4. Keresztezés

Permutációs kódolásnál ügyelni kell arra, hogy az utód továbbra is **érvényes permutáció** legyen.

Egypontos

Egyszerű egypontos keresztezést nem feltétlenül használhatunk.

Például:

$$P1: A B C | D E \\ P2: A D E | C B$$

Ha egyszerű egypontos keresztezéssel kombinálnánk a két szülőt:

$$A B C | C B$$

az eredmény:

[A,B,C,C,B]

Ez nem megfelelő kromoszóma, mert:

- B kétszer szerepel,
- C kétszer szerepel,
- D hiányzik,
- E hiányzik.

Ezért permutációs kódolás esetén speciális keresztezési operátort kell alkalmazni.

Sorrendi keresztezés

Használjunk **sorrendi keresztezést**.

A két szülő:

\[P1: A B | C D | E \\ P2: A D | E C | B \]

1. lépés – Egy szakasz megtartása

Az első szülőből megtartjuk a kijelölt szakaszt:

\[_ _ | C D | _\]

Az utód jelenleg:

\[[_,_,C,D,_] \]

2. lépés – A hiányzó elemek meghatározása

A második szülő (A D E C B):

C és D már szerepelnek az utódban, ezért ezeket kihagyjuk.

Marad:

A E B

3. lépés – A hiányzó helyek feltöltése

A hiányzó helyeket a megmaradt elemekkel töltjük fel:

\[A E | C D | B\]

Az új utód:

$$\boxed{U=[A,E,C,D,B]}$$

–

Az utód kiértékelése

Az új útvonal:

$$[A,E,C,D,B]$$

A hozzá tartozó távolságok:

- A → E = 8
- E → C = 5
- C → D = 4
- D → B = 5

Ezért:

$$8+5+4+5=22$$

Az új egyed értéke:

$$\boxed{22}$$

A keresztezéssel kapott utód tehát rosszabb lett, mint mindkét szülő:

$$f(P_1)=13$$

$$f(P_2)=16$$

$$f(U)=22$$

Ez nem jelent hibát.

> **A keresztezés nem garantálja, hogy az utód jobb lesz a szülőknél.**

5. Mutáció

Permutációs reprezentáció esetén egy egyszerű mutációs operátor a **cserés mutáció (swap mutation)**.

Ennek során két gén pozícióját felcseréljük.

A keresztezéssel kapott utód:

$$A\ E\ C\ D\ B$$

Cseréljük fel E és B helyét:

$$ A B C D E $$

A mutáció után kapott egyed:

$$ \boxed{[A,B,C,D,E]} $$

—

A mutálodott egyed kiértékelése

Az útvonal:

$$ A-B-C-D-E $$

Távolsága:

$$ 4+3+4+2=13 $$

Tehát:

$$ \boxed{f=13} $$

A mutáció előtt:

$$ [A,E,C,D,B] \rightarrow 22 $$

A mutáció után:

$$ [A,B,C,D,E] \rightarrow 13 $$

Ebben az esetben a mutáció jelentősen javította a megoldást.

> **Fontos:** a mutáció nem garantálja a javulást. A mutáció javíthatja vagy ronthatja is az egyedet.

6. Az új populáció kialakítása

A keresztezés és mutáció után el kell döntenünk, hogy mely egyedek kerüljenek a **következő generációba**.

Tartsuk meg a populáció eredeti méretét:

$$ N=4 $$

A régi populáció:

Egyed	Kromoszóma	Úthossz
c_1	[A,B,C,D,E]	**13**
c_2	[A,C,B,E,D]	18
c_3	[A,D,E,C,B]	16
c_4	[A,E,D,B,C]	18

Az újonnan létrehozott egyedek:

Egyed	Kromoszóma	Úthossz
\(c_5\)	[A,E,C,D,B]	22
\(c_6\)	[A,B,C,D,E]	**13**

6.1 Elitizmus

Alkalmazzunk **1 egyedes elitizmust**.

Ez azt jelenti, hogy a régi populáció legjobb egyedét automatikusan megtartjuk.

A legjobb régi egyed:

\[[A,B,C,D,E] \]

értéke:

\[13 \]

Ez tehát biztosan bekerül a következő generációba.

Az új generáció többi helyét az utódokkal, illetve további kiválasztott egyedekkel tölthetjük fel.

6.2 Diverzitás

A genetikus algoritmusban nem feltétlenül célszerű mindig kizárólag a legjobb célfüggvényértékkel rendelkező egyedeket továbbvinni.

Ha folyamatosan csak a legjobb egyedeket választjuk, akkor a populáció egyedei egyre hasonlóbbá válhatnak. Ez a **diverzitás csökkenéséhez** vezethet, és növelheti annak veszélyét, hogy az algoritmus idő előtt egy lokális optimum környezetében ragad.

Ezért az új populáció kialakításánál figyelembe vehetjük az egyedek **diverzitását** is.

6.2.1 Hamming távolság

Két kromoszóma Hamming-távolsága azt mutatja meg, hogy hány pozícióban különböznek egymástól.

Permutációs reprezentáció esetén ugyanazt az elvet alkalmazhatjuk.

Vegyük például az alábbi két útvonalat:

\[c_1 = [A,B,C,D,E] \\ c_2 = [A,D,E,C,B] \]

Pozíciónként összehasonlítva:

Pozíció	\(c_1\)	\(c_2\)	Eltérés

1.	A	A	0
2.	B	D	1
3.	C	E	1
4.	D	C	1
5.	E	B	1

Ezért:

\[d_H(c_1,c_2)=4 \]

Általánosan:

\[d_H(c_i,c_j)= \sum_{k=1}^{m} I(c_{ik}\neq c_{jk}) \]

ahol:

- \(m\) a kromoszóma hossza,
- \(I(\cdot)\) értéke 1, ha az adott pozícióban különböznek a gének,
- \(I(\cdot)\) értéke 0, ha megegyeznek.

6.2.2 Az útvonalak diverzitásának meghatározása

Az elitizmussal a következő generációba továbbvitt elem(ek) kivételével megnézzük a diverzitásokat.

Öt egyedünk van, ezért összesen:

\[\binom{5}{2}=10 \]

különböző egyedpár Hamming-távolságát kell meghatároznunk.

\(c_2\) és \(c_3\)

Pozíciónként:

Pozíció	\(c_2\)	\(c_3\)	Megegyezik?	Érték
1.	A	A	igen	0
2.	C	D	nem	1
3.	B	E	nem	1
4.	E	C	nem	1
5.	D	B	nem	1

Ezért:

\[d_H(c_2,c_3) = 0+1+1+1+1 = \boxed{4} \]

\(c_2\) és \(c_4\)

Pozíció	\(c_2\)	\(c_4\)	Megegyezik?	Érték

1.	A	A	igen	0
2.	C	E	nem	1
3.	B	D	nem	1
4.	E	B	nem	1
5.	D	C	nem	1

Ezért:

$$ d_H(c_2,c_4) = 0+1+1+1+1 = \boxed{4} $$

c_2 és c_5

Pozíció	c_2	c_5	Megegyezik?	Érték
1.	A	A	igen	0
2.	C	E	nem	1
3.	B	C	nem	1
4.	E	D	nem	1
5.	D	B	nem	1

Ezért:

$$ d_H(c_2,c_5) = 0+1+1+1+1 = \boxed{4} $$

c_2 és c_6

Pozíció	c_2	c_6	Megegyezik?	Érték
1.	A	A	igen	0
2.	C	B	nem	1
3.	B	C	nem	1
4.	E	D	nem	1
5.	D	E	nem	1

Ezért:

$$ d_H(c_2,c_6) = 0+1+1+1+1 = \boxed{4} $$

c_3 és c_4

Pozíció	c_3	c_4	Megegyezik?	Érték
1.	A	A	igen	0
2.	D	E	nem	1
3.	E	D	nem	1
4.	C	B	nem	1
5.	B	C	nem	1

Ezért:

$$ d_H(c_3,c_4) = 0+1+1+1+1 = \boxed{4} $$

c_3 és c_5

Pozíció	c_3	c_5	Megegyezik?	Érték
1.	A	A	igen	0
2.	D	E	nem	1
3.	E	C	nem	1
4.	C	D	nem	1
5.	B	B	igen	0

Ezért:

$$ d_H(c_3,c_5) = 0+1+1+1+0 = \boxed{3} $$

c_3 és c_6

Pozíció	c_3	c_6	Megegyezik?	Érték
1.	A	A	igen	0
2.	D	B	nem	1
3.	E	C	nem	1
4.	C	D	nem	1
5.	B	E	nem	1

Ezért:

$$ d_H(c_3,c_6) = 0+1+1+1+1 = \boxed{4} $$

c_4 és c_5

Pozíció	c_4	c_5	Megegyezik?	Érték
1.	A	A	igen	0
2.	E	E	igen	0
3.	D	C	nem	1
4.	B	D	nem	1
5.	C	B	nem	1

Ezért:

$$ d_H(c_4,c_5) = 0+0+1+1+1 = \boxed{3} $$

c_4 és c_6

Pozíció	c_4	c_6	Megegyezik?	Érték
1.	A	A	igen	0
2.	E	B	nem	1
3.	D	C	nem	1

4.	B	D	nem	1
5.	C	E	nem	1

Ezért:

$$ d_H(c_4,c_6) = 0+1+1+1+1 = \boxed{4} $$

c_5 és c_6

Pozíció	c_5	c_6	Megegyezik?	Érték
1.	A	A	igen	0
2.	E	B	nem	1
3.	C	C	igen	0
4.	D	D	igen	0
5.	B	E	nem	1

Ezért:

$$ d_H(c_5,c_6) = 0+1+0+0+1 = \boxed{2} $$

Ez a vizsgált párok közül a legkisebb Hamming-távolság, tehát c_5 és c_6 egymáshoz viszonylag hasonló kromoszómák.

Hamming-távolság mátrix

A páronként kiszámított Hamming-távolságokat egyetlen mátrixban is összefoglalhatjuk:

	c_2	c_3	c_4	c_5	c_6
c_2	0	4	4	4	4
c_3	4	0	4	3	4
c_4	4	4	0	3	4
c_5	4	3	3	0	2
c_6	4	4	4	2	0

A Hamming-távolság mátrix **szimmetrikus**, mert:

$$ d_H(c_i,c_j)=d_H(c_j,c_i) $$

Például:

$$ d_H(c_3,c_5)=d_H(c_5,c_3)=3 $$

A főátlóban minden érték 0:

$$ d_H(c_i,c_i)=0 $$

hiszen egy egyed önmagával teljesen megegyezik.

Az egyedek összesített Hamming-távolsága

Az egyes egyedek diverzitását a populáció többi tagjától vett **Hamming-távolságok összegével** határozhatjuk meg.

Az \(i\). egyed diverzitási értéke: \[\boxed{ D_i= \sum_{ \substack{j=1,j \neq i}}^{n} d_H(c_i,c_j) }\] ahol:

D_i a c_i jelölt összesített diverzitása, $d_H(c_i,c_j)$ a két egyed Hamming-távolsága, n a vizsgált jelöltek száma.

A vizsgált halmazban:

\[n=5 \]

ezért minden egyed esetében négy másik egyedtől mért távolság átlagát számítjuk.

A diverzitási értékek összesítése

Egyed	Kromoszóma	Összesített Hamming-távolság
\(c_2\)	[A,C,B,E,D]	16
\(c_3\)	[A,D,E,C,B]	15
\(c_4\)	[A,E,D,B,C]	15
\(c_5\)	[A,E,C,D,B]	12
\(c_6\)	[A,B,C,D,E]	14

Ha a nagyobb Hamming-távolságot tekintjük nagyobb diverzitásnak, akkor a rangsor:

\[\boxed{ c_2 > c_3=c_4 > c_6 > c_5 } \]

vagy a diverzitási értékekkel:

\[\boxed{ 16 > 15=15 > 14 > 12 } \]

A \(c_3\) és \(c_4\) egyedek között holtverseny alakult ki.

A holtverseny feloldható például:

- véletlenszerűen, vagy - a jobb fitnessértékkel rendelkező egyed előresorolásával.

A legnagyobb diverzitással rendelkező jelölt c_2=[A,C,B,E,D], mivel D_2=16. A legkisebb diverzitással rendelkező jelölt c_5=[A,E,C,D,B], mivel D_5=12.

Minél nagyobb egy egyed összesített Hamming-távolsága, annál inkább eltér a vizsgált populáció többi tagjától, vagyis annál nagyobb mértékben járul hozzá a populáció **diverzitásához**.

=== 6.3 Túlélési valószínűség meghatározása A jelölteket az összesített diverzitási értékük alapján csökkenő sorrendbe rendezzük.

Rang	Jelölt	Diverzitás
1.	\(c_2\)	**16**

2.	\(c_3\)	**15**
3.	\(c_4\)	**15**
4.	\(c_6\)	**14**
5.	\(c_5\)	**12**

A \(c_3\) és \(c_4\) jelölt diverzitása megegyezik. A holtverseny feloldható például véletlenszerűen vagy a jobb fitnessérték alapján.

Legyen:

\[P=0,5 \]

A rang alapján a túlélési valószínűségek:

\[P_1=P=0,5 \]

\[P_2=(1-P)P=0,5\cdot0,5=0,25 \]

\[P_3=(1-P)^2P=0,5^2\cdot0,5=0,125 \]

\[P_4=(1-P)^3P=0,5^3\cdot0,5=0,0625 \]

Az utolsó jelölt esetében:

\[P_5=(1-P)^4=0,5^4=0,0625 \]

Így:

Rang	Jelölt	Diverzitás	Túlélési valószínűség
1.	\(c_2\)	16	**0,5000**
2.	\(c_3\)	15	**0,2500**
3.	\(c_4\)	15	**0,1250**
4.	\(c_6\)	14	**0,0625**
5.	\(c_5\)	12	**0,0625**

Ellenőrzés:

\[0,5+0,25+0,125+0,0625+0,0625=1 \]

A túlélési valószínűségek alapján **rulettkerék-szelekcióval** választhatjuk ki a következő generáció egyedeit.

Kumulált valószínűségek meghatározása

Először meghatározzuk a kumulált túlélési valószínűségeket.

\(c_2\)

$$ K_2=0,5 $$

c_3

$$ K_3=0,5+0,25=0,75 $$

c_4

$$ K_4=0,75+0,125=0,875 $$

c_6

$$ K_6=0,875+0,0625=0,9375 $$

c_5

$$ K_5=0,9375+0,0625=1 $$

Így a következő intervallumokat kapjuk:

Jelölt	Túlélési valószínűség	Kumulált valószínűség	Kiválasztási intervallum
c_2	0,5000	0,5000	$0 \leq r < 0,5000$
c_3	0,2500	0,7500	$0,5000 \leq r < 0,7500$
c_4	0,1250	0,8750	$0,7500 \leq r < 0,8750$
c_6	0,0625	0,9375	$0,8750 \leq r < 0,9375$
c_5	0,0625	1,0000	$0,9375 \leq r \leq 1$

Véletlen számok generálása

Mivel az új populáció mérete:

$$ N=4 $$

Ezért a már elitizmussal továbbvitt 1 egyed mellé még három egyedet kell kiválasztanunk.

Generáljunk három véletlen számot a

$$ [0,1] $$

intervallumból.

Például:

$$ r_1=0,32 $$

$$ r_2=0,81 $$

$$ r_3=0,91 $$

A generált véletlen számokat megkeressük a kiválasztási intervallumokban.

Első húzás

$$ r_1=0,32 $$

Mivel:

$$ 0\leq0,32<0,5 $$

ezért:

$$ \boxed{c_2} $$

kerül kiválasztásra.

Második húzás

$$ r_2=0,81 $$

Mivel:

$$ 0,75\leq0,81<0,875 $$

ezért:

$$ \boxed{c_4} $$

kerül kiválasztásra.

Harmadik húzás

$$ r_3=0,91 $$

Mivel:

$$ 0,875\leq0,91<0,9375 $$

ezért:

$$ \boxed{c_6} $$

kerül kiválasztásra.

A három kiválasztott egyed:

$$ c_2,\quad c_4,\quad c_6 $$

Így a következő generáció:

Új populáció	Egyed	Kromoszóma
1.	c_1	[A,B,C,D,E]
2.	c_2	[A,C,B,E,D]
3.	c_4	[A,E,D,B,C]
4.	c_6	[A,B,C,D,E]

Ez lesz a genetikus algoritmus következő generációjának populációja.

> Az új populáció kiválasztására többféle stratégia alkalmazható. Az elitizmus biztosítja, hogy a korábban megtalált legjobb megoldás ne vesszen el.

From:
https://edu.iit.uni-miskolc.hu/ - **Institute of Information Science - University of Miskolc**

Permanent link:
https://edu.iit.uni-miskolc.hu/mesterseges_intelligencia_alapjai_gtk_genetikus_permutacio

Last update: **2026/11/05 18:25**