

## Exemplu de utilizare a unui algoritm genetic pentru rezolvarea unui puzzle Criptaritmetica

SEND	MATE
9567 +	1487 +
MORE	INFO
1085	2509
-----	-----
MONEY	COOL
10652	3996

Criptaritmetica este un puzzle criptaritmetic cu litere inlocuite de cifre, unde numerele formate insumate verifica suma data de numarul obtinut prin schimbarea literelor cu cifre in cuvantul final.

Exemple de alte puzzle-uri Criptaritmetica care au solutii:

BASIC + LOGIC = PASCAL  
APPLE + LEMON = BANANA  
HACK + HACKER = REBOOT  
EARTH + URANUS = SATURN  
COMET + SATURN = URANUS  
BROWN + YELLOW = PURPLE

### Exemplul CRIPTARITMETICA tratat in continuare:

doi + trei = cinci

730 + 9480 = 10210 (o potentiala solutie)

Determinam toate literele care sunt diferite in cadrul problemei, adica in cadrul cuvintelor *doi*, *trei* si *cinci*.

Avem 8 variabile: d, o, i, t, r, e, c si n.

Alegem probabilitatea de incrucisare  $p_c = 0.4$  si probabilitatea de mutatie  $p_m = 0.1$ . Vom aplica un algoritm genetic pentru o singura generatie (iteratie) si vom urmari fiecare pas al algoritmului.

### Cum se genereaza un individ (sau cromozom, sau potentiala solutie)?

Lungimea fiecarui cromozom este 10 si codifica cifrele 0, 1, 2, 3, 4, 5, 6, 7, 8 si 9. Initializarea se face in aceeaasi maniera in care se genereau valorile pentru vectorul de la problema aranjarii damelor, cu diferenta ca aici avem 10 valori in loc de 8. A se vedea algoritmul de la slide-ul cu initializarea unei solutii de la cursul *Caracteristici si constrangeri* (<http://inf.ucv.ro/~cstoean/courses/ia/index.html>).

### Cu ce populam fiecare cromozom? Care sunt toate simbolurile care ne apar?

Fiecare cromozom va contine valori numerice pentru toate cele 8 variabile, plus 2 variabile goale pe care le putem nota cu x si y (x si y sunt alese astfel incat sa nu apara printre celelalte variabile).

Deci avem:  $M = \{d, o, i, t, r, e, c, n, x, y\}$

lungime(M) = 10

Consideram ca prima pozitie dintr-un vector este 1, nu 0 ca in limbajul de programare C. Vom construi primul cromozom (notat in continuare cromozom<sub>1</sub>) pornind de la multimea M aplicand algoritmul cu initializarea unei solutii mentionat mai sus.

Generam un numar intre 1 si 10  $\rightarrow r_1 = 6$   
 Luam al 6-lea element din M, adica e  $\rightarrow$  cromozom<sub>1</sub> = [e].  
 Scoatem acel element din M  $\rightarrow \{d,o,i,t,e,c,n,x,y\}$ , lung(M) = 9  
 Generam  $r_2 \in \{1,\dots,9\} \rightarrow 3$  adica litera i  $\rightarrow [e,i]$   
 $M = \{d,o,t,r,c,n,x,y\}$ , lung(M) = 8  
 $r_3 = \{1,\dots,8\} \rightarrow 8 \rightarrow y \rightarrow [e,i,y]$   
 $M = \{d,o,t,r,c,n,x\}$ , lung(M) = 7  
 $r_4 = 1 \rightarrow d \rightarrow [e,i,y,d]$   
 $M = \{o,t,r,c,n,x\}$ , lung(M) = 6  
 $r_5 = 4 \rightarrow c \rightarrow [e,i,y,d,c]$   
 $M = \{o,t,r,n,x\}$ , lung(M) = 5  
 $r_6 = 2 \rightarrow t \rightarrow [e,i,y,d,c,t]$   
 $M = \{o,r,n,x\}$ , lung(M) = 4  
 $r_7 = 1 \rightarrow o \rightarrow [e,i,y,d,c,t,o]$   
 $M = \{r,n,x\}$ , lung(M) = 3  
 $r_8 = 2 \rightarrow n \rightarrow [e,i,y,d,c,t,o,n]$   
 $M = \{r,x\}$ , lung(M) = 2  
 $r_9 = 1 \rightarrow r \rightarrow [e,i,y,d,c,t,o,n,r]$   
 $M = \{x\}$ , lung(M) = 1  
 Am obtinut cromozomul  $c_1 = [e,i,y,d,c,t,o,n,r,x]$

Analog, se obtin  $c_2 = [i,c,x,d,n,r,y,o,t,e]$   
 $c_3 = [x,d,o,e,i,c,t,y,n,r]$   
 $c_4 = [r,d,i,c,t,o,y,e,x,n]$

### Cum masuram performanta acestor indivizi?

	0	1	2	3	4	5	6	7	8	9
$c_1$	e	i	y	d	c	t	o	n	r	x
$c_2$	i	c	x	d	n	r	y	o	t	e
$c_3$	x	d	o	e	i	c	t	y	n	r
$c_4$	r	d	i	c	t	o	y	e	x	n

Ne amintim problema de rezolvat: doi + trei = cinci. Functia de evaluare va masura cat de aproape este suma obtinuta prin adunarea celor doua numere corespunzatoare cuvintelor *doi* si *trei* de numarul corespunzator lui *cinci*. Pentru primul cromozom, *cinci* este reprezentat de 41741, iar *doi* si *trei* vor fi 361, respectiv 5801. Astfel, vom avea evaluarile de mai jos pentru cei 4 cromozomi:

$$f(c_1) = |41741 - (361 + 5801)| = 35579$$

$$f(c_2) = |10410 - (370 + 8590)| = 1450$$

$$f(c_3) = |54854 - (124 + 6934)| = 45894$$

$$f(c_4) = |32932 - (152 + 4072)| = 28708$$

Daca insumam evaluarile tuturor celor 4 cromozomi, obtinem evaluarea intregii populatii de 111631.  
Vom urmari dupa o singura generatie cum se va modifica aceasta.

Aplicam selectia turnir cu  $k = 2$  pentru a selecta 4 parinti:

Generam 2 numere diferite in mod aleatoriu intre 1 si 4.  
Fie ele 1 si 4  $\rightarrow$  al patrulea este selectat pentru ca are o evaluare mai mica, deci mai buna.  
Alte 2 numere intre 1 si 4: 2 si 4  $\rightarrow$  al doilea este selectat.  
Alte 2 numere intre 1 si 4: 3 si 2  $\rightarrow$  al doilea este selectat.  
Alte 2 numere intre 1 si 4: 3 si 1  $\rightarrow$  primul este selectat.

Dupa selectie, populatia arata astfel:

	0	1	2	3	4	5	6	7	8	9	Evaluare
$c_1$	r	d	i	c	t	o	y	e	x	n	28708
$c_2$	i	c	x	d	n	r	y	o	t	e	1450
$c_3$	i	c	x	d	n	r	y	o	t	e	1450
$c_4$	e	i	y	d	c	t	o	n	r	x	35579

Avem  $p_c = 0.4$

pentru  $i = 1, 4$  executa //se iau cromozomii unul cate unul  
generam  $r \in [0, 1]$   
daca  $r < p_c$  atunci  $c_i$  este selectat  
Presupunem ca s-a generat: 0.2, 0.1, 0.5, 0.8  $\rightarrow c_1, c_2$

Daca am fi avut numar impar de parinti, se mai alegea in mod aleatoriu unul.  
Recombinarea intre  $c_1$  si  $c_2$ :

Se genereaza punctul de taietura aleatoriu intre 1 si 9. Fie acesta 3.  
 $c_1 = r d i | c t o y e x n$   
 $c_2 = i c x | d n r y o t e$   
 Folosind recombinarea de la problema aranjarii damelor descrisa in cadrul cursului *Elemente de calcul evolutiv* (<http://inf.ucv.ro/~cstoean/courses/ia/index.html>), se obtin descendenti  $c'_1$  si  $c'_2$ :  
 $c'_1 = r d i n y o t e c x$   
 $c'_2 = i c x t o y e n r d$   
 $f(c'_1) = |82382 - (152 + 6072)| = 76158$   
 $f(c'_2) = |10710 - (940 + 3860)| = 5910$

Lupta pentru supravietuire se da intre  $c_1, c_2, c'_1, c'_2$ . Doar 2 pot merge mai departe; se iau cei mai buni 2, adica  $c_2$  si  $c'_2$ .

Dupa recombinare, populatia obtinuta este urmatoarea:

	0	1	2	3	4	5	6	7	8	9	Evaluare
$c_1$	i	c	x	t	o	y	e	n	r	d	5910
$c_2$	i	c	x	d	n	r	y	o	t	e	1450
$c_3$	i	c	x	d	n	r	y	o	t	e	1450
$c_4$	e	i	y	d	c	t	o	n	r	x	35579

Mutatia cu probabilitatea  $p_m = 0.1$ , deci aproximativ una din 10 gene (pozitii) se modifica la fiecare  $c_i$ . De exemplu, la  $c_3$ , se alege gena a 9-a, cea referitoare la t. Pentru t se genereaza o alta valoare intre 1 si 10 - fie ea 10  $\rightarrow$  din  $c_3$  se obtine  $c'_3 = i c x d n r y o e t$  cu cele doua pozitii schimbate.

$$f(c'_3) = |10410 - (370 + 9580)| = 460 < 1450 \rightarrow c'_3 \text{ il inlocuieste pe } c_3.$$

Mutatia aplicata lui  $c_4$ :

Se alege gena 6 (t) si se alege gena de schimb: 9 (r) si se obtine  $c'_4 = e i y d c r o n t x$

$$f(c'_4) = |41741 - (361 + 8501)| = 32879 < 35579 \rightarrow c'_4 \text{ inlocuieste } c_4$$

Dupa o generatie populatia obtinuta este urmatoarea.

	0	1	2	3	4	5	6	7	8	9	Evaluare
$c_1$	i	c	x	t	o	y	e	n	r	d	5910
$c^2$	i	c	x	d	n	r	y	o	t	e	1450
$c_3$	i	c	x	d	n	r	y	o	e	t	460
$c_4$	e	i	y	d	c	r	o	n	t	x	32879

Performanta totala a populatiei dupa o generatie este de 40699, deci se observa o evolutie mare de la de 111631.

## Problema rucsacului

Avem un rucsac care are o capacitate de 15 kg si 5 obiecte de diferite marimi: 10, 2, 1, 4 si 3. Alegem  $p_c = 0.4$  si  $p_r = 0.1$ .

Cromozomul are 5 gene, cate una pentru fiecare obiect. Valorile posibile pentru fiecare gena: 0 sau 1 (absenta obiectului sau prezenta sa in rucsac). Pentru fiecare gena, se genereaza un numar din multimea  $\{0, 1\}$ .

Pentru  $c_1$ , se genereaza  $[0 \ 1 \ 0 \ 0 \ 1]$ .

Evaluarea va masura cat de aproape ne gasim ocuparea completa a rucsacului:

$$f(c_1) = |15 - (0*10 + 1*2 + 0*1 + 4*0 + 3*1)| = 10$$

Analog, generam:

	1	2	3	4	5	Evaluare
<b>c<sub>1</sub></b>	0	1	0	0	1	10
<b>c<sub>2</sub></b>	1	0	1	0	1	1
<b>c<sub>3</sub></b>	0	1	1	1	0	8
<b>c<sub>4</sub></b>	1	0	0	1	0	1

Evaluarea totala a populatiei: 20.

Selectia – analog cu criptaritmetica:

1, 4 → 4

2, 4 → 2

3, 2 → 2

3, 1 → 3

	1	2	3	4	5	Evaluare
<b>c<sub>1</sub></b>	1	0	0	1	0	1
<b>c<sub>2</sub></b>	1	0	1	0	1	1
<b>c<sub>3</sub></b>	1	0	1	0	1	1
<b>c<sub>4</sub></b>	0	1	1	1	0	8

Se aplica recombinarea binara discreta ( $p_c = 0.4$ ):

Numere generate: 0.5 0.2 0.6 0.1 → c<sub>2</sub> si c<sub>4</sub> →

c<sub>2</sub> [1 0 1 0 1] → 1

c<sub>4</sub> [0 1 1 1 0] → 8

Se genereaza pentru fiecare pozitie un numar intre 0 si 1. Daca este mai mica de 0.5 se alege pentru c'<sub>2</sub> valoarea de la c<sub>2</sub> si pentru c'<sub>4</sub> valoarea de la c<sub>4</sub>, iar daca e mai mare de 0.5 pentru c'<sub>2</sub> se ia de la c<sub>4</sub>, iar pentru c'<sub>4</sub> de la c<sub>2</sub>.

Numere generate: 0.3, 0.7, 0.4, 0.6, 0.8 →

c'<sub>2</sub> = [1 1 1 1 0] → 2

c'<sub>4</sub> = [0 0 1 0 1] → 11

Se aleg c<sub>2</sub> si c'<sub>2</sub>.

	1	2	3	4	5	Evaluare
<b>c<sub>1</sub></b>	1	0	0	1	0	1
<b>c<sub>2</sub></b>	1	0	1	0	1	1
<b>c<sub>3</sub></b>	1	0	1	0	1	1

<b>c<sub>4</sub></b>	1	1	1	1	0	2
----------------------	---	---	---	---	---	---

Aplicam mutatia asupra lui  $c_4$ . Cum? Se selecteaza o gena si isi schimba valoarea din cea existenta in cealalta (din 0 in 1 sau din 1 in 0).

Pentru gena 2  $\rightarrow c'_4 = [1\ 0\ 1\ 1\ 0]$   $\rightarrow$  solutia, deci evaluarea cromozomului este 0  $\rightarrow$  il inlocuieste pe  $c_4$ .

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Evaluare</b>
<b>c<sub>1</sub></b>	1	0	0	1	0	1
<b>c<sub>2</sub></b>	1	0	1	0	1	1
<b>c<sub>3</sub></b>	1	0	1	0	1	1
<b>c<sub>4</sub></b>	1	0	1	1	0	0

Evaluarea intregii populatii este acum 3 (spre deosebire de 20 cu numai o iteratie in urma).