

## GA ძირითადი თეორემა

GA ეფუძნება ყველაზე უფრო შეგუებადი ინდივიდების (ქრომოსომების) ტრანსფორმაციის პრინციპს.  $P(0)$ -ინდივიდების საწყისი პოპულაცია,  $P(k)$ - ალგორითმის  $K$ -ური იტერაცია. თითოეული  $P(k)=0,1,2,\dots$  პოპულაციიდან სელექციის მეთოდით ამოირჩევა მაღალი შეგუებადობის მატჩენბლის მქონე ქრომოსომები, რომლებიც ჩაერთვებიან მშობლების პულში (matching pool)  $M(k)$ -ში. შემდეგ  $M(k)$  – პოპულაციიდან შეირჩევა ე.წ მშობელთა წყვილები და  $P_c$ -შეუღლების ოპერატორის ალბათობით და  $P_m$  მუტაციის ოპერატორის ალბათობით მოხდება ახალი  $P(k+1)$  პოპულაციის ფორმირება, რომელშიც შევლენ  $M(k)$  პოპულაციის შთამომავლები.

მაშასადამე ყოველი სქემისათვის რომელსაც გააჩნია კარგი ამონახსნი, სასურველი იქნებოდა, რომ ყოველი  $k$ -ი ტერაციის შემდეგ გაზრდილიყო ამ სქემის შესაბამისი ქრომოსომების რიცხვი.

გენეტიკურ ალგორითმებში სქემის სიცოცხლისუნარიანობაზე გავლენას ახდენს 3 ფაქტორი: ქრომოსომების სელექცია, შეუღლება და მუტაცია. გავანალიზოთ თითოეული ფაქტორის ზეგავლენა სქემის სიცოცხლისუნარიანობაზე.

$S$ -ით აღვნიშნოთ გასახილველი სქემა,  $C(S,k)$ -აღვნიშნოთ  $S$ -სქემის შესაბამისი ქრომოსომების რაოდენობა  $P(k)$  პოპულაციიდან, ამრიგად  $C(S,k)$  – არის  $P(k)$ -ის  $S$  სიმრავლის ელემენტების რაოდენობა (სიმრავლის სიმძლავრე).

გამოვარკვეოთ სელექციის გავლენა.

სელექციის შემდეგ ქრომოსომების კოპირება ხდება  $P(k)$  დან  $M(k)$ -ში.

$$P_s(ch_i) = \frac{F(ch_i)}{\sum_{i=1}^N F(ch_i)} \text{ ალბათობით}$$

ვთავქთ  $F(S_1 k)$  აღნიშნავს  $P(k)$  პოპულაციაში შემავალი ქრომოსომების შეგუებადობის ფუნქციის საშუალოს  $S$ -სქემისთვის

$$\text{თუ } P(k) S = \{ch_{i_1}, \dots, ch_{i_{C(S,k)}}\} \text{ მაშინ}$$

$$F(S,k) = \frac{\sum_{i=1}^{C(S,k)} F(ch_i)}{C(S,k)} \quad (1)$$

$F(s,k)$  – ასევე უწოდებენ  $S$  სქემის შეგუებადობას  $k$ - ურ იტერაციაზე.

ვთქვათ  $L(k)$  – აღნიშნავს  $N$  სიმძლავრის  $P(k)$  – პოპულაციის ქრომოსომების შეგუებადობის უნქციების მნიშვნელობის ჯამს, მაშინ

$$L = \sum_{i=1}^N \bar{F}(ch_i^{(k)}) \quad (2)$$

აღვ.  $\bar{F}(k)$  – ამ პოპულაციაში შემავალი ქრომოსომების შეგუებადობის ფუნქციის მნიშვნელობების საშუალო.

$$\bar{F}(k) = \frac{1}{N} L(k) \quad (3)$$

ვთქვათ  $ch_r^{(k)}$  აღნიშნავს  $M(k)$  მშობელთა პულის ელემენტს.  $ch_r^{(k)} \in M(k)$   $i=1, \dots, c(S,k)$  ალბათობა იმისა, რომ  $ch_r^{(k)} = ch_i$  განისაზღვრება შეფარდებით  $\frac{F(ch_i)}{F(k)}$ , ამიტომ  $M(k)$  პოპულაციაში შემავალი იმ ქრომოსომების მოსალოდნელი რაოდენობა, რომელსაც  $ch_i$  - ის ტოლია

$$N \frac{F(ch_i)}{L(k)} = \frac{F(ch_i)}{F(k)} \quad (4)$$

ამრიგად  $P(k)$   $S$  სიმრავლიდან აღებული ქრომოსომების მოსალოდნელი რაოდენობა, რომლებიც უნდა ჩავრთოთ  $M(k)$ -ში გამოითვლება შემდეგნაირად

$$\sum_{i=1}^{C(s,k)} \frac{F(ch_i)}{F(k)} = \frac{F(ch_s)}{F(k)} \quad (5)$$

და ეს მიიღება (1) დან .

ვინაიდან  $M(k)$ -ს ყოველი ქრომოსომა იმავდროულად ეკუთვნის  $P(k)$ -ს ამიტომ ის ქრომოსომები რომლებიც შედიან  $M(k)$   $\cap$   $S$  - ეს იგივე ინდივიდებია (ქრომოსომებია) რომლებიც აღებულია  $P(k)$   $\cap$   $S$  სიმრავლიდან -  $M(k)$  პოპულაციაში ჩასართველად. მაშინ თუ  $b(s,k)$  - აღვ.  $S$  სქემის შესაბამის  $M(k)$  პოპულაციაში შემავალი ქრომოსომების რაოდენობას, მაშინ

**შედეგი 1:**

$$E[b(s,k)] = C(s,k) \frac{F(s,k)}{F(k)} \quad (6)$$

ე.ი თუ  $s$  სქემა შეიცავს ისეთ ქრომოსომებს რომელთა შეგუებადობის ფუნქციის მნიშვნელობა აჭარბებს საშუალო მაჩვენებელს ე.ი  $s$  სქემის შეგუებადობის მნიშვნელობა  $k$ -ურ იტერაციაზე აღმოჩნდება მეტი ვიდრე  $P(k)$  პოპულაციის ქრომოსომების შეგუებადობის ფუნქციის საშუალო მნიშვნელობა ამიტომაც  $\frac{F(s,k)}{F(k)}$ , მაშინ  $s$  სქემის  $M(k)$  ქრომოსომების მოსალოდნელი რაოდენობა მეტი იქნება  $P(k)$  პოპულაციის ქრომოსომების რაოდენობაზე. მაშასადამე შეიძლება დავასკვნათ რომ სელექციის შედეგად გავრცელდება „საშუალოზე მაღალი“ შეგუებადობის სქემები და გაქრება „უარესი“ შეგუებადობის მაჩვენებლიანი სქემები.

შეუღლებს ოპერატორის გავლენა სქემის შეგუებადობის მაჩვენებელზე

განვიხილოთ  $M(k)$  ო  $S$  სიმრავლიდან რომელიმე ქრომოსომა, ანუ  $S$  სქემის ქრომოსომები  $M(k)$  - დან. ალბათობა იმისა რომ ეს ქრომოსომა მიიღებს მონაწილეობას შეუღლებაში  $P_e$  - ს ტოლია. ეი შეუღლების წერტილი უნდა მდებარეობდეს მოცემული სქემის პირველ და უკანასკნელ მუდმივ წევრებს შორის. ამის ალბათობა ტოლია  $\frac{d(s)}{L-1}$ .

**შედეგი 2:**  $M(k)$  ო  $S$  რომელიმე ქრომოსომისათვის ალბათობა იმისა, რომ ის იქნება ამორჩეული შეუღლებისათვის და მისი არცერთი შთამომავალი არ იქნება  $S$  სქემაში წარმოდგენილი შემოსაზღვრულია ზემოდან სიდიდით

$$P_e \frac{d(s)}{L-1}$$

ამ სიდიდეს ვუწოდებთ  $S$  სქემის განადგურების, გაქრობის ალბათობის სიდიდეს.

**შედეგი 3:**  $M(k)$  ო  $S$  რომელიმე ქრომოსომისათვის ალბათობა იმისა, რომ ის იქნება ამორჩეული შეუღლებისათვის და შემდეგ მისი ერთი მაინც შთამომავალი არ იქნება  $S$  - ის წარმოდგენილი შემოსაზღვრულია ქვევიდან და უდრის

$$1 - P_e \frac{d(s)}{L-1}$$

ამ სიდიდეს ვუწოდებთ  $S$  სქემის გადარჩენის ალბათობას.

თუ მოცემული ქრომოსომა ეკუთვნის  $S$  სქემას და ამორჩეულია შეუღლებისათვის და თუ მეორე ქრომოსომაც ეკუთვნის  $S$  სქემას, მაშინ მათი შთამომავალი ქრომოსომაც ეკუთვნის  $S$  სქემას. მიღებული შედეგების გათვალისწინებით შეიძლება დავადგინოთ რამდენად მნიშვნელოვანია  $d(s)$  სქემის გაქრობის მაჩვენებელი სქემის გაქრობა-გადარჩენის საკითხებში.

განვიხილოთ  $M(k)$ -ში მუტაციის გავლენა, მუტაციის ოპერატორი  $P_m$  ალბათობით ცვლის ქრომოსომაში კონკრეტულ პოზიციაზე 0-ს 1-ანით და/ან პირიქით. ცხადია სქემა გადაურჩება მუტაციას მხოლოდ იმ შემთხვევაში თუ მისი ყველა მუდმივი პოზიციები დარჩებიან უცვლელი მუტაციის ოპერაციის შემდეგ.

**შედეგი 4:**  $S$  სქემის ქრომოსომა  $M(k)$ -დან დარჩება ამ სქემაში მაშინ და მხოლოდ მაშინ, როცა ამ ქრომოსომის მუდმივ პოზიციაზე მყოფი არცერთი სიმბოლო არ შეიცვლება მუტაციის შემდეგ. ამის ალბათობა ტოლია  $(1 - P_m)^{O(s)}$ . შედეგი ალბათობა იმისა რომ რომელიმე ქრომოსომა  $M(k)$  ო  $S$  დან ეკუთვნის  $S$  სქემის მუტაციის შემდეგ ტოლია  $1 - P_m^{O(s)}$  - ეს სიდიდე გვიჩვენებს სქემის გარჩენის ალბათობას მუტაციის შემდეგ.

თუ  $P_m$  ალბათობა მცირეა ( $P_m \ll 1$ ) მაშინ მიახლოებით შეიძლება ვთქვათ რომ  $S$  სქემის გადარჩენის ალბათობა მუტაციის შემდეგ გამოისახება

$$1 - P_m^{O(s)}.$$

შევეცადოთ განვსაზღვროთ სელექციის შეუღლების და მუტაციის ერთობლივი მოქმედების შედეგები, თუ რომელიმე ქრომოსომა  $S$  სქემიდან  $M(k)$  ო  $S$  დარჩება და

მოხვდება  $P(k+1) \cap S$  სქემაში, მაშინ

$$E[C(s,k+1)] \geq C(s,k) \frac{F(s,k)}{F(k)} (1 - P_c \frac{d(s)}{L-1}) (1 - P_m)^{o(s)} \quad (7)$$

ეს გამოსახულება გვიჩვენებს, როგორ იცვლება მოცემული სქემის შესაბამისი ქრომოსომების რაოდენობა პოპულაციიდან პოპულაციამდე. ეს ცვლილებები განპირობებულია შემდეგი 3 ფაქტორით; კერძოდ:  $\frac{F(s,k)}{F(k)}$  გვიჩვენებს შეგუებადობის ფუნქციის საშუალო მნიშვნელობის გავლენას, ხოლო  $1 - P_c \frac{d(s)}{L-1}$  შეუღლების ოპერატორის გავლენას და  $(1 - P_m)^{o(s)}$  მუტაციის ოპერატორის გავლენას. რაც უფრო დიდია თითოეული ფაქტორის მნიშვნელობა,  $S$  სქემისთვის ქრომოსომების მით მეტი მოსალოდნელი რაოდენობა აღმოჩნდება შემდგომ პოპულაციაში. (7) – შეიძლება შემდეგნაირად გადავწეროთ.

$$E[C(s,k+1)] \geq c(s,k) \frac{F(s,k)}{F(k)} (1 - P_c \frac{d(s)}{L-1} - P_m^{o(s)}) \quad (8)$$

დიდი პოპულაციისთვის შეიძლება მოვახდინოთ (8)–ის აპროქსიმაცია

$$C(s,k+1) \geq c(s,k) \frac{F(s,k)}{F(k)} (1 - P_c \frac{d(s)}{L-1} - P_m^{o(s)}) \quad (9)$$

შეიძლება დავასკვნათ რომ  $S$  სქემის შესაბამისი ქრომოსომების მოსალოდნელი რაოდენობა შემდგომ პოპულაციაში დამოკიდებულია  $S$  სქემის შემადგენლობასთან და ასევე სქემის რიგსა და სიგრძეზე. შევნიშნოთ რომ, საშუალოზე მაღალი შეგუებადობის უნქციის მაჩვენებლიანი, მცირე რიგისა და სიგრძის სქემები ხასიათდებიან იმით რომ შემდგომ პოპულაციებში მათი წარმომადგენლების რაოდენობა ყოველთვის იმატებს. ეს ზრდადობა ჩანს (6) – ში. დიდი პოპულაციების შემთხვევაში (6) შეიძლება შევცვალოთ რეკურენტული დამოკიდებულებით.

$$C(s,k+1) = C(s,k) \frac{F(s,k)}{F(k)} \quad (10)$$

თუ დავუშვებთ რომ  $S$  სქემას გააჩნია შეგუებადობა საშუალოზე  $\varepsilon\%$  ით მეტი ე.ი

$$F(s,k) = \overline{F(k)} + \varepsilon \overline{F(k)} \quad (11)$$

თუ ჩავსვავთ (9) გამოსახულებას (8) – ში, და იმის გათვალისწინებით, რომ  $\varepsilon$  იცვლება დროში, მივიღებთ

$$C(s,k) = C(s,0)(1 + \varepsilon)^k$$

$$\varepsilon = \frac{F(s,k) - \overline{F(k)}}{\overline{F(k)}} \quad (12)$$

მაშასადამე საშუალოზე მაღალი შეგუებადობის ფუნქციის  $S$  სქემებისთვის  $\varepsilon > 0$ , ხოლო წინააღმდეგ შემთხვევაში  $\varepsilon < 0$ .

(12) – აღწერს გეომეტრიულ პროგრესიას. შეიძლება ვთქვათ რომ საშუალოზე უკეთესი სქემების რაოდენობა GA ალგორითმის იტერაციის შემდგომ ეტაპზე იზრდება ყოველთვის,

მაშინ როცა საშუალოზე უარესი სქემების რიცხვი შესაბამისად იკლებს. ყველა ეს დამოკიდებულება სამართლიანია იმ შემთხვევებისათვის როცა შეგუებადობის F ფუნქცია ყოველთვის დადებით მნიშვნელობებს ღებულობს, ხოლო იმ შემთხვევაში როცა საოპტიმიზაციო ფუნქცია ღებულობს უარყოფით მნიშვნელობებს, მაშინ საჭიროა დამატებითი კორექტივების შეტანა.

### გენეტიკური ალგორითმების სქემების თეორიები

დაბალი რიგის, მცირე სიგრძის და საშუალოზე მაღალი შეგუებადობის ფუნქციის მქონე სქემების რაოდენობა ყოველთვის იმატებს გენეტიკური ალგორითმის შემდგომ შთამომავლობაში.

ამ თეორიებიდან გამომდინარე დიდი მნიშვნელობა ენიჭება კოდირების საკითხებს, რაზეცაა დამოკიდებული დაბალი რიგის, მცირე სიგრძის და საშუალოზე მაღალი შეგუებადობის ფუნქციების მქონე სქემების აგებაზე.

### GA და ევოლუციის პროცესი

#### GA მეთოდი და გამოყენება

#### GA ძირითადი ცნებები და განმარტებები

GA - გენეტიკური ალგორითმების იდეა გამოთქვა ჯ. ჰოლანდმა XX საუკუნის სამოცდაათიან წლებში. ჯ. ჰოლანდი დაინტერესებული იყო ცოცხალი და ხელოვნური სისტემების ბუნებრივ გარემოსთან ადაპტაციის საკითხებით [1]. ევოლუციის პროცესში ქრომოსომები განიცდიან ევოლუციას და არა თვითონ ცოცხალი ორგანიზმები.

GA გამოიყენება ხელოვნური ინტელექტის (AI) სისტემებში, ოპტიმიზაციის ამოცანებში, ხელოვნურ ნეიტრონულ ქსელებში (NN), არამკაფიო სისტემებში (Fuzzy) და ცოდნის სხვა დარგებში პროგრამული უზრუნველყოფის დასამუშავებლად.

GA-ში გამოყენებულია ყველაზე უფრო კარგად შეგუებული ინდივიდების გადარჩენის ევოლუციური პრინციპი. ტრადიციული ოპტიმიზაციის მეთოდებისაგან განსხვავებით GA იყენებს: 1) კოდირებულ პარამეტრებს და 2) ამონახსნის საძიებლად არა ერთ საწყის წერტილს, არამედ საწყის პოპულაციას; 3) მიზნის ფუნქციას მხოლოდ და არა მის წარმოებულს ან რაიმე დამატებით ინფორმაციას; 4) ამორჩევის ალბათურ წესს და არა დეტერმინირებულს. ჩამოთვლილი ოთხი თვისება განაპირობებს GA-ს ეფექტურობას სხვა დანარჩენ ტექნოლოგიებთან.

GA ძირითადი ცნებები. GA-ში გამოყენებულია გენეტიკიდან გადმოღებული ცნებები და განმარტებები: გენი, ქრომოსომა, პოპულაცია და სხვა.

*პოპულაცია* – ინდივიდების სასრული სიმრავლე;

*ინდივიდები* – კოდირებული ქრომოსომები ანუ ამოცანის პარამეტრები – ამონახსნთა სივრცის წერტილები;

*ქრომოსომა* – კოდური მიმდევრობები ანუ ჯაჭვები, გენების მოწესრიგებული მიმდევრობა;

*გენი* – ქრომოსომის ატომარული ელემენტი ანუ გენოტიპის ელემენტი;

*გენოტიპი* – მოცემული ინდივიდის ქრომოსომების ერთობლიობა;

*ფენოტიპი* – დეკოდირებული სტრუქტურა ანუ ამოცანის პარამეტრების სიმრავლე;

*აღელი* – კონკრეტული გენის მნიშვნელობა;

*ლოკუსი* – ქრომოსომაში კონკრეტული გენის პოზიცია;

*რეპროდუქცია* – აღნიშნავს ახალი ქრომოსომების შექმნას მშობლების გენების რეკომბინაციის შედეგად.

*რეკომბინაცია* – ამ პროცესის შედეგად მიიღება გენების ახალი კომბინაციები, რისთვისაც გამოიყენება ორი ძირითადი ოპერატორი (ოპერატორი) – შეუღლება და მუტაცია.

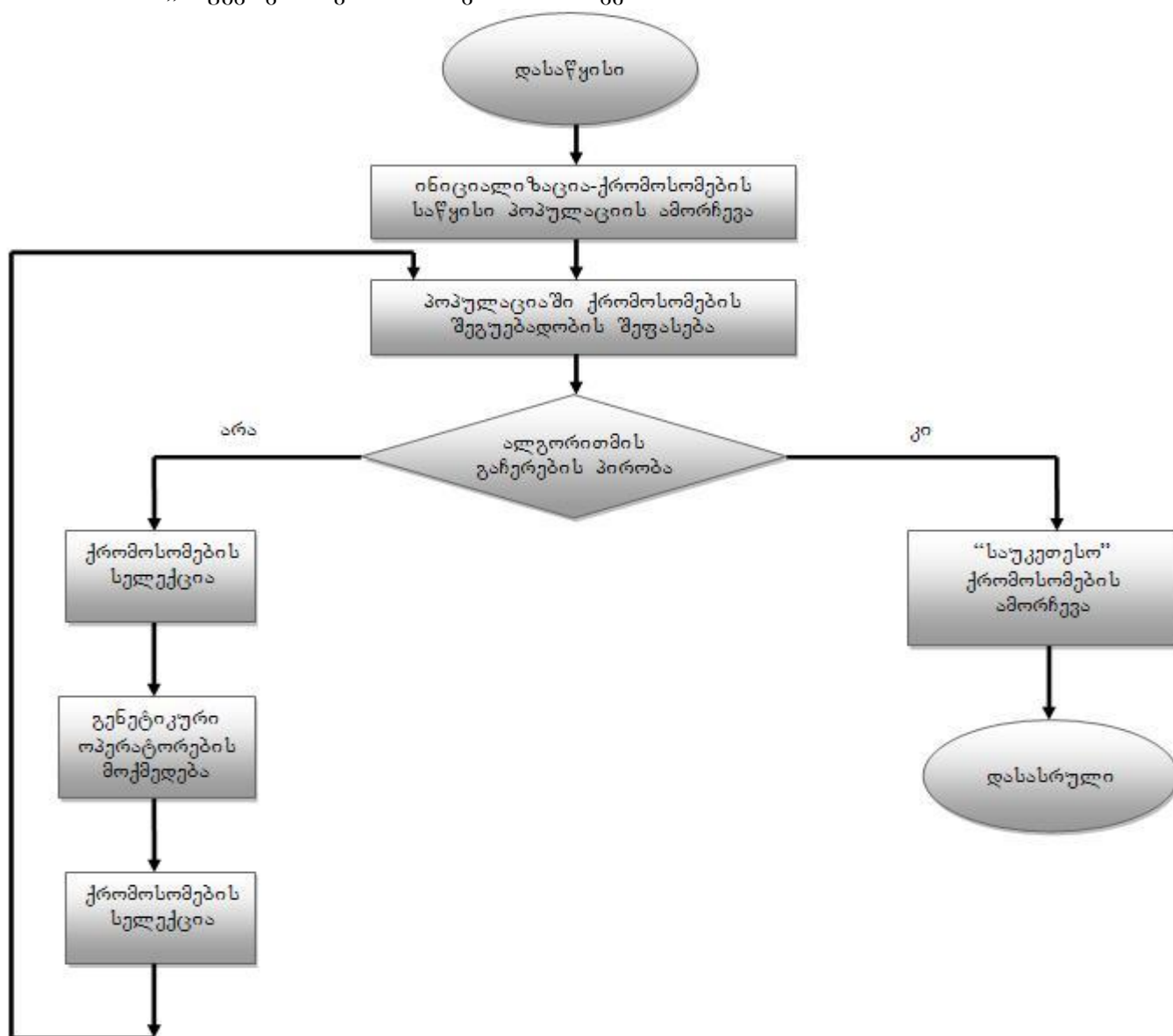
GA-ს ერთ-ერთი ძირითადი ცნებაა შეგუებადობის ფუნქციის ცნება (fitness function), ანუ შეფასების ფუნქცია, რომელიც გვიჩვენებს პოპულაციაში მოცემული ინდივიდის შეგუებადობის ზომას. ამ ფუნქციის საშუალებით ამოირჩევენ პოპულაციიდან უფრო მეტად შეგუებულ ინდივიდებს, რაც ევოლუციის პრინციპიდან გამომდინარეობს. გადარჩება ის ინდივიდი, რომელსაც მაღალი შეგუებადობის ფუნქციის მაჩვენებელი გააჩნია, ანუ „ძლიერის“ გადარჩენის პრინციპი. ეს ფუნქცია უმნიშვნელოვანესია GA ფუნქციონირების პროცესში. ოპტიმიზაციის ამოცანებში ამ ფუნქციას მიზნის ფუნქციას უწოდებენ და ხდება მისი ოპტიმიზაცია (უფრო ზუსტად, მაქსიმიზაცია), მართვის თეორიაში შეიძლება იყოს ცდომილების ფუნქცია, ხოლო თამაშთა თეორიაში – ფასის ფუნქცია. GA-ს ფუნქციონირებისას ყოველ იტერაციაზე თითოეული ინდივიდი ფასდება შეგუებადობის ფუნქციის საშუალებით და მის მიხედვით მიიღება ახალი პოპულაცია შემდგომი თაობისათვის.

## კლასიკური GA

ძირითადად, კლასიკური GA ან ზოგჯერ მას უწოდებენ ელემენტარულს ან

მარტივ GA-ს, შედგება შემდეგი ბიჯებისაგან:

1. ინიციალიზაცია ანუ ქრომოსომების საწყისი პოპულაციის ამორჩევა;
2. პოპულაციაში ქრომოსომების შეგუებადობის შეფასება;
3. ალგორითმის გაჩერების პირობა;
4. ქრომოსომების სელექცია;
5. გენეტიკური ოპერატორების მოქმედება;
6. ახალი პოპულაციის ფორმირება;
7. „საუკეთესო“ ქრომოსომების ამორჩევა.



ქრომოსომების შეგუებადობის შეფასება. პოპულაციაში თითოეული ქრომოსომის შეგუებადობის ფუნქციას გამოითვლიან, რაც მაღალია ამ ფუნქციის მნიშვნელობა, მით მაღალია ქრომოსომის „ხარისხი“. შეგუებადობის ფუნქციის ფორმა დამოკიდებულია ამოსახსნელი ამოცანის ტიპზე. კლასიკურ GA-ში იგულისხმება, რომ შეგუებადობის ფუნქციის მნიშვნელობა ყოველთვის არაუარყოფითია.

ალგორითმის გაჩერების პირობის შემოწმება. GA-ს გაჩერების პირობა დამოკიდებულია კონკრეტულ პირობებზე. ოპტიმიზაციის ამოცანებში თუ ცნობილია შეგუებადობის ფუნქციის მაქსიმუმის (ან მინიმუმის) მნიშვნელობა, მაშინ ალგორითმი

გაჩერდება თუ მიიღწევა მოსალოდნელი მნიშვნელობა, შესაძლებელია – მოცემული სიზუსტით. ალგორითმის გაჩერება შეიძლება მაშინაც, როცა მიღებული შედეგები აღარ უმჯობესდება. ალგორითმი შეიძლება გაჩერდეს განსაზღვრული დროის შემდეგ ან განსაზღვრული რაოდენობის ოპერაციების შესრულების შემდეგ. თუ გაჩერების პირობა სრულდება, მაშინ ალგორითმი გადადის „საუკეთესო“ ქრომოსომის ამორჩევის ეტაპზე. წინააღმდეგ შემთხვევაში შემდეგ ეტაპზე სრულდება სელექცია.

ქრომოსომების სელექცია – ამ ეტაპზე იმ ქრომოსომების ამორჩევა ხდება, რომელთა შეგუებადობის ფუნქციაც ამის საშუალებას იძლევა, ე.ი. ისინი მონაწილეობას მიიღებენ შემდეგ ეტაპზე ახალი თაობის ფორმირებაში. ბუნებრივი გადარჩევის პრინციპის თანახმად მაღალი შეგუებადობის ფუნქციის ქრომოსომები შემდეგ ეტაპზე იძლევიან შთამომავლობას.

სელექციის სხვადასხვა მეთოდებიდან ერთ-ერთი ყველაზე პოპულარულია რულეტკის მეთოდი, სახელწოდება ცნობილი რულეტკიდან მოდის (Roulette wheel selection). თითოეული ქრომოსომის შეგუებადობის ფუნქციის მნიშვნელობა პროპორციულად შეესაბამება რულეტკის ბარაბანზე შესაბამისი სექტორის ზომას, ანუ რაც მეტია ფუნქციის მნიშვნელობა, მით მეტია სექტორის ზომა. პოპულაციის ყველა ქრომოსომის შეგუებადობის ფუნქციების ჯამი შეესაბამება რულეტკის ბორბლის მთელ ზედაპირს. თუ თითოეული ქრომოსომას აღნიშნავთ  $\chi_i$ , სადაც  $i=1,2,\dots,N$ .  $N$  აღნიშნავს პოპულაციაში ქრომოსომების რაოდენობას, მაშინ  $\chi_i$  ქრომოსომის შესაბამის რულეტკის სექტორს  $v(\chi_i)$ -ით გამოვსახავთ პროცენტებში, გვეჩვენება

$$v(\chi_i) = P_s(\chi_i) \cdot 100\%, \quad \text{სადაც} \quad P_s(\chi_i) = \frac{F(\chi_i)}{\sum_{i=1}^N F(\chi_i)}.$$

$F(\chi_i)$  – არის  $\chi_i$  ქრომოსომის შეგუებადობის ფუნქციის მნიშვნელობა.

$P_s(\chi_i)$  – არის  $\chi_i$  ქრომოსომის სელექციის ალბათობა (ამორჩევის ალბათობა).

ქრომოსომის სელექცია ნიშნავს (არის) რულეტკის ბარაბანის იმ სექტორის ამორჩევა („მოგება“), რომელიც ამ ქრომოსომას შეესაბამება. სელექციის პროცესის შედეგად იქმნება მშობელთა პოპულაცია, ანუ mating pool,  $N$ -რაოდენობის ქრომოსომებით.

გენეტიკური ოპერატორები. სელექციის შედეგად ამორჩეულ ქრომოსომების პოპულაციაში გამოიყენება GA-ს ორი ძირითადი გენეტიკური ოპერატორი: შეუღლების crossover ოპერატორი და მუტაციის mutation ოპერატორი. მუტაციის ოპერატორი შეუღლების ოპერატორთან შედარებით მეორეხარისხოვან როლს ასრულებს. კლასიკურ GA-ში შეუღლება ყოველთვის სრულდება  $0.5 \leq P_c \leq 1$ , ხოლო მუტაცია კი ძალიან იშვიათად  $0 \leq P_m \leq 0.1$ . მუტაცია შეიძლება მოხდეს ან შეუღლებამდე ან შეუღლების შემდეგ. განვიხილოთ თითოეული მათგანი.

შეუღლების ოპერატორი – პირველ ეტაპზე მშობლების პოპულაციიდან ამოირჩევა ორი ქრომოსომა. ამ ეტაპზე ხდება ამ ქრომოსომების შეწყვილება შემთხვევით  $P_c$



ალბათობით. შემდეგ ისევ შემთხვევით აიღებენ შეუღლების წერტილს, თუ ქრომოსომა შედგება L-გენისაგან (ე.ი. ქრომოსომის სიგრძეა 4), მაშინ  $i_k$ -შეუღლების წერტილი არის ნატურალური რიცხვი ნაკლები L-ზე, ანუ  $i_k$  წერტილი შემთხვევით აირჩევა [1, L-1] ინტერვალდან. შეუღლების შედეგად მიიღება შთამომავლების წყვილი.

- 1) შთამომავლის ქრომოსომები 1-დან  $i_k$ -მდე შედგება პირველი მშობლის გენებისაგან, ხოლო  $i_{k+1}$  პოზიციიდან L-მდე – მეორე მშობლის გენებისაგან.
- 2) შთამომავლის ქრომოსომები 1-დან  $i_k$ -მდე შედგება მეორე მშობლის გენებისაგან, ხოლო  $i_{k+1}$  პოზიციიდან L-მდე – პირველი მშობლის გენებისაგან.

მუტაციის ოპერატორი  $P_m$  ალბათობით იცვლება ქრომოსომაში გენის მნიშვნელობა მისი საპირისპიროთი. ანუ 0 იცვლება 1-ით და პირიქით.

შეუღლების ოპერატორი crossover

1 მშობელი 01.....10.....1 შეუღლება 01.....11.....0

2 მშობელი 11.....01.....0 11.....00.....1

ლოკუსი 12,.....  $i_k$ .....L

მუტაცია 10.....1.....0 მუტაცია 10.....0.....0

ლოკუსი 12,.....  $i_k$ ..... L

### ახალი პოპულაციის ფორმირება

გენეტიკური ოპერატორების მოქმედების შედეგად მიღებული ქრომოსომები გაერთიანდებიან ახალ პოპულაციაში, მას უკვე ვუწოდებთ მიმდინარე პოპულაციას GA მოცემული იტერაციისათვის. ყოველ შემდგომ იტერაციაზე გამოითვლება ქრომოსომების შეგუებადობის ფუნქცია და მოწმდება ალგორითმის გაჩერების პირობა ან გადავდივართ ალგორითმის შესრულების შემდეგ ეტაპზე – სელექციის ეტაპზე. კლასიკურ GA-ში ყოველთვის ამორჩეული ქრომოსომების პოპულაცია მეორე ეტაპზე იცვლება ახალი პოპულაციით, ხოლო ქრომოსომების რიცხვი ყველა პოპულაციაში თანაბარია.

### საუკეთესო ქრომოსომის ამორჩევა

თუ (ამოცანის) ალგორითმის გაჩერების პირობა შესრულდება, მაშინ უნდა გამოვიტანოთ მუშაობის შედეგი, ე.ი. ამონახსნი. საუკეთესო ამონახსნად ითვლება ის

ქრომოსომა, რომლის შეგუებადობის ფუნქციის მნიშვნელობა უდიდესია. კლასიკური GA ლიტერატურაში ცნობილია სამი ძირითადი ინსტრუმენტის ერთობლიობით. ესენია: სელექცია, შეუღლება და მუტაცია.

ამოცანის პარამეტრების კოდირების მეთოდი. კლასიკურ GA-ში საწყის პოპულაციაში წარმოდგენილი ქრომოსომებისათვის გამოყენებულია ორობითი კოდირება ანუ გენები ტოლია 0-ის ან 1-ის. ქრომოსომების სიგრძე დამოკიდებულია ამოცანის პირობაზე, უფრო ზუსტად, ამონახსნთა სივრცეში წერტილების რაოდენობაზე.

## GA სქემა

გენეტიკური ალგორითმის ფუნქციონირების უკეთ გასაგებად გამოიყენება სქემის ცნება. იმ ქრომოსომებისათვის, რომლებსაც რაღაც ერთნაირი თვისებები გააჩნიათ ანუ ერთმანეთის მსგავსნი არიან, შემოიღეს სქემა, ანუ სქემა არის იმ ქრომოსომების სიმრავლე, რომლებსაც წინასწარ განსაზღვრულ ერთ ან რამდენიმე პოზიციაზე ნულები და ერთიანები აქვთ. სქემის განხილვისას გამოიყენება გაფართოებული ალფავიტი  $\{0,1,*\}$ . \*-ით აღნიშნულ პოზიციაზე იგულისხმება გენის ორივე დასაშვები მნიშვნელობა - 0 ან 1.

მაგალითად

$$10*1=\{1001, 1011\}, \quad *01*10=\{001010, 001110, 101010, 101110\}.$$

ამბობენ, რომ ქრომოსომა ეკუთვნის  $S$  მოცემულ სქემას, თუ ყოველი  $j$ -ური პოზიციისათვის  $j=1,2,\dots,L$ , სადაც  $L$  ქრომოსომების (რიცხვი) სიგრძეა, სიმბოლოა, რომელსაც  $S$  სქემის  $j$ -ური პოზიცია უკავია. \* - სიმბოლოს შეესაბამება როგორც 0, ისე 1. ერთი და იგივე გამოთქმები ქრომოსომა წარმოადგენს სქემას და ქრომოსომა შეესაბამება სქემას. თუ სქემაში არის  $m$ -რაოდენობის \*, მაშინ ეს სქემა შეიცავს  $2^m$  ქრომოსომას; ამასთანავე, ყოველი  $L$  სიგრძის ქრომოსომა ეკუთვნის  $2^L$  სქემას.

განსაზღვრება.  $S$  სქემის რიგი (order) აღნიშნეთ  $O(S)$  – ეს არის მუდმივი პოზიციების რიცხვი სქემაში.

მაგალითი.  $O(10*1)=3$ ;  $O(*01*10)=4$ ;  $O(*0*1)=2$ ;  $O(*101*)=3$ ;  $O(****)=0$ .

$O(L)$  – ყოველთვის მთელი რიცხვია  $[0,L]$ -ინტერვალთან.

სქემის რიგი ტოლია  $L$  სიგრძეს გამოკლებული  $*$ -ის რაოდენობა.

სქემის სიგრძე (განსაზღვრება).  $S$  სქემის სიგრძის (დეფინიგ ლენგჰტ)  $d(S)$  – ეს არის მანძილი პირველ და უკანასკნელ მუდმივ სიმბოლოს შორის, ანუ სხვაობა მარჯვენა და მარცხენა უკიდურეს პოზიციებს შორის, რომლებიც მუდმივ სიმბოლოებს შეიცავენ.

მაგალითი.

$$d(10^*1)=4-1=3; d(*01^*10)=6-2=4; d(**0^*1^*)=5-3=2; d(**1^*)=0$$

$d(S)$ -ით ხასიათდება სქემაში გამოხატული ინფორმაციის შინაარსობრიობა.

$d(S)$  – მოელი რიცხვია  $[0, L-1]$  ინტერვალდან.

### მოდიფიცირებული GA

კლასიკურ GA-ში გამოყენებულია (პარამეტრების) ქრომოსომების ორობითი კოდირება, რულეტკის მეთოდით ქრომოსომების სელექცია და წერტილოვანი შეუღლების ოპერატორი (ერთწერტილოვანი შეუღლება). GA-ს ეფექტურობის გასაზრდელად გამოყენებულია GA-ის მოდიფიცირებული ალგორითმები. ერთწერტილოვანი შეუღლების ოპერატორის ნაცვლად იყენებენ ორწერტილოვან, მრავალწერტილოვან და სხვა შეუღლების ოპერატორებს, ქრომოსომების სელექციის სხვადასხვა მეთოდებს. GA-ს ზოგიერთი ვერსიით შესაძლებელია არამარტო გლობალური ოპტიმუმების, არამედ ლოკალური ოპტიმუმების მოძებნაც. ეს ალგორითმები იყენებენ ე.წ. ნიშებს, ბუნებრივი ეკოლოგიური ნიშების მსგავსად. GA-ს სხვადასხვა ვერსიებით შეიძლება მრავალკრიტერიუმიანი ოპტიმიზაციის ამოცანების ამოხსნა და სხვ. გენეტიკური მიკროალგორითმები წარმატებით მუშაობენ მცირე ზომის პოპულაციებთან და ასევე მცირე სიგრძის ქრომოსომების შემთხვევაში.

### სელექციის მეთოდები

რულეტკის მეთოდი სელექციის ძირითადი მეთოდია კლასიკურ GA-ში, მაგრამ ამ მეთოდს იყენებენ მაშინ, როცა  $F(\chi)$  შეგუებადობის ფუნქციის მნიშვნელობები დადებითია. ამრიგად, ეს მეთოდი წარმატებით გამოიყენება ფუნქციის მაქსიმიზაციის ამოცანებში. სელექციის სხვა მეთოდებიდან განვიხილოთ ტურნირის მეთოდი (tournament selection). ამ მეთოდის გამოყენების შემთხვევაში პოპულაციის ყველა ინდივიდი იყოფა ორ ჯგუფად, თითოეული ჯგუფიდან ვირჩევთ ერთ „საუკეთესო“ ინდივიდს, რომელთა შეგუებადობის ფუნქცია უკეთესია. განიხილავენ ამორჩევის ორ წესს: დეტერმინირებული და შემთხვევითი ამორჩევის წესს. დეტერმინირებული

ამორჩევის დროს ამორჩევა ხორციელდება 1-ის ტოლი ალბათობით. შემთხვევითი ამორჩევის დროს ალბათობა 1-ზე ნაკლებია. პოპულაციის ქვეჯგუფები შეიძლება სხვადასხვა ზომის იყოს, ხშირად მთელი პოპულაცია იყოფა ისეთ ქვეჯგუფებად, რომლებიც 2-3 ინდივიდისაგან შედგება. ტურნირის მეთოდი გამოიყენება, როგორც ფუნქციის მაქსიმიზაციის, ისე მინიმიზაციის შემთხვევაში, ასევე მრავალკრიტერიუმიანი ოპტიმიზაციის ამოცანებში. შეიძლება ქვეჯგუფებში პოპულაციის ზომების შეცვლა.

რანგული სელექციის მეთოდის შემთხვევაში პოპულაციის რანჟირება ხდება ინდივიდების შეგუებადობის ფუნქციის მნიშვნელობების მიხედვით. შედეგად თითოეულ ინდივიდს სიაში (ცხრილში) მიეწერება ის რიცხვი, რომლითაც განისაზღვრება მისი ადგილი სიაში და ეს არის მისი რანგი.

### რეპროდუქციის განსაკუთრებული პროცედურები

ელიტარული სტრატეგია (elitist strategy) გულისხმობს საუკეთესო ქრომოსომების აუცილებელ გადასვლას იტერაციის შემდგომ ეტაპზე. კლასიკურ GA-ში ზოგჯერ ყველაზე შეგუებული ქრომოსომები ყოველთვის არ ხვდებიან შემდგომ თაობებში. ეს ნიშნავს, რომ  $P(k+1)$  პოპულაციაში არ მოხვდა ის ქრომოსომა, რომელსაც ყველაზე მაღალი მნიშვნელობების შეგუებადობის ფუნქცია გააჩნია  $P(k)$ -ში. ელიტარული სტრატეგია გამოიყენება ასეთი შემთხვევების დროს და ეს ინდივიდი არ დაიკარგება და აუცილებლად ჩაერთვება ახალ პოპულაციაში.

GA პოპულაციის ნაწილობრივი შეცვლით ან GA დაფიქსირებული მდგომარეობით (steady state). ასეთი სტრატეგიით პოპულაციის ნაწილი ყოველგვარი ცვლილების გარეშე (შეუღლებისა და მუტაციის ოპერატორების გამოყენების გარეშე) გადადის შემდგომ თაობაში. ასეთი სტრატეგია ხშირად გამოიყენება ისეთ ამოცანებში, როდესაც მომხმარებელს სურს პოპულაციის რაღაც ნაწილის შენარჩუნება.

გენეტიკური ოპერატორები. კლასიკური GA-ს განსხვავებით მოდიფიცირებულ GA-ში გამოიყენება ორწერტილოვანი, მრავალწერტილოვანი და თანაბარი შეუღლების ოპერატორი.

ორწერტილოვანი შეუღლება (two-point crossover). ამ შემთხვევაში გვაქვს შეუღლების ორი წერტილი და შთამომავლები ღებულობენ მშობლების გენებს შემთხვევით ამორჩეულ ორი წერტილით მიღებული მშობლების ქრომოსომების ფრაგმენტებს. *ერთწერტილოვანი შეუღლების მაგალითი:*

პირველი მშობელი: [0011 0011 1010]	} $\xrightarrow{\text{შეუღლება}}$	[0011 1111 1010] შთამომავალი 1
მეორე მშობელი : [1010 1101 1011]		[1010 0001 1011] შთამომავალი 2

მრავალწერტილიანი შეუღლება (multiple-point crossover). იგი წარმოადგენს წინა ოპერატორის განზოგადოებულ შემთხვევას და ხასიათდება შეუღლების მრავალი წერტილით. სამწერტილოვანი შეუღლების მაგალითი:

$$\left. \begin{array}{l} \text{პირველი მშობელი: } [0011 \ 0011 \ 1010] \\ \text{მეორე მშობელი : } [1010 \ 1101 \ 1011] \end{array} \right\} \xrightarrow{\text{შეუღლება}} \begin{array}{l} [0011 \ 1111 \ 1010] \text{ შთამომავალი 1} \\ [1010 \ 0001 \ 1011] \text{ შთამომავალი 2} \end{array}$$

$l_k : 4 \quad 6 \quad 9$

თანაბარი შეუღლება (uniform crossover). სხვანაირად მას უწოდებენ მონოლითურ შეუღლებას. სრულდება შემთხვევით ამორჩეული ეტალონის შესაბამისად, რომელიც გვიჩვენებს, თუ რომელი გენები გადაეცემა მემკვიდრეობით მშობლებისაგან შთამომავლობას. დანარჩენი გენები აიღება მეორე მშობლისაგან. ვთქვათ, არჩეულია ეტალონი, ეტალონი 010110111011, რომელშიც 1 აღნიშნავს, რომ ამ პოზიციაზე აიღება პირველი მშობლის შესაბამისი გენი, ხოლო 0 გვიჩვენებს მეორე მშობლისაგან აღებულ გენს. ასე ხდება პირველი შთამომავლის ფორმირება და ანალოგიურად მიიღება მეორე შთამომავალიც.

თანაბარი შეუღლების მაგალითი:

$$\left. \begin{array}{l} \text{პირველი მშობელი: } [0011 \ 0011 \ 1010] \\ \text{მეორე მშობელი : } [1010 \ 1101 \ 1011] \end{array} \right\} \xrightarrow{\text{შეუღლება}} \begin{array}{l} [1011 \ 0111 \ 1010] \text{ შთამომავალი 1} \\ [0010 \ 1001 \ 1011] \text{ შთამომავალი 2} \end{array}$$

ლოკუსი: 1 2 3 4 5 6 7 8 9 10 11 12

ეტალონი: 0 1 0 1 1 0 1 1 1 0 1 1

ინვერსიის ოპერატორი. ქრომოსომაში შემთხვევით ამორჩეულ ორ პოზიციას შორის მოთავსებულ გენებში იცვლება ამ გენების თანმიმდევრობა.

მაგალითი: მოძებნეთ მაქსიმუმი  $f(x,y) = \frac{1}{1+x^2+y^2}$

I

	x	y	F
$C_0$	-1	2	0.167
$C_1$	-2	3	0.007
$C_2$	1.5	0	0.31
$C_3$	0.5	-1	0.44

II

	x	y	F
$C_0(x_{c_8}y_{c_2})$	0.5	0	0.8
$C_1(x_{c_2}y_{c_8})$	1.5	-1	0.24
$C_2(x_{c_8}y_{c_0})$	0.5	2	0.19
$C_3(x_{c_0}y_{c_8})$	0.5	-1	0.44

საშუალო I F=0.231

საშუალო II F=0.39



