ELTE Természettudományi Kar
Operációkutatás Tanszék

Genetikus algoritmusok Markov-lánc modellje

Szakdolgozat

Készítette: **Boncz András**

Témavezető: **Kovács Margit**
Docens

1996.

Tartalomjegyzék

B e v e z e t é s

Az utóbbi időben egyre hangsúlyosabb szerepet kapnak a különböző kereső algoritmusok. Alapvetően két különböző típusú keresési feladat van:

a) szűkebb értelemben vett **keresés**: adott tulajdonságú elem keresése, illetve

b) **optimalizálás**: egy adott reláció szerinti legjobb elem keresése.

A két feladat ekvivalens egymással, mivel az is egy tulajdonság, hogy valamely elem egy adott reláció szerint a legjobb és úgy is definiálhatunk egy relációt, hogy minden elemhez, melynek megvan a keresett tulajdonsága, 1-t rendelünk, a többihez pedig 0-t. Ennek ellenére a két különböző megfogalmazás gyökeresen eltérő megoldási módszereket igényelhet. Jelen dolgozatban csak az optimalizálás egy lehetséges módszeréről lesz szó.

A legkézenfekvőbb módszer -amely mindkét megfogalmazás esetén használható- a **leszámlálás**. Ez abból áll, hogy az ember az összes szóbajöhető elemet megvizsgálja, ekkor el lehet dönteni, hogy melyik a legjobb. Ennek természetesen megvan az a hátránya, hogy sok ideig tart, ha nagy a keresési tér.

A folytonos kereső módszerekien (pl. Newton- ill. gradiens-módszer) alapuló **mohó algoritmus** akkor használható, ha a keresési térnek van valamilyen szomszédsági struktúrája. Ekkor kiindulva egy tetszőleges elemből, a szomszédai közül a legjobbat választjuk ki. Addig folytatjuk ezt a műveletet, amíg a szomszédai közt van nála jobb. Ekkor természetesen egy lokálisan legjobb elemhez jutunk. Vagy előzetesen tudjuk, hogy csak egyetlen lokális optimum van, vagy több különböző elemből kiindulva a kapott lokális optimumok közül kiválasztjuk a legjobbat. Ez a módszer akkor használható, ha a reláció kellően ”sima”, azaz nincs túlságosan sok lokális optimum.

A **genetikus algoritmusok** ötlete biológiai példa hatására merült fel. Itt már mi generáljuk a struktúrát, számunkra kényelmes módon kódoljuk a tér elemeit. Nem egyetlen elemet vizsgálunk, hanem egy egész populációt és úgynevezett genetikus operátorok segítségével hozunk létre újabb populációkat. Ezt addig folytatjuk, amíg az új populáció eltérése az előzőtől egy küszöb alatt marad. Ezt a módszert vizsgálom a dolgozatban.

Az úgynevezett "**simulated annealing**" algoritmus -amelyet talán lassú hűtésnek lehetne fordítani- pedig fizikai analógián alapul [1]. Eszerint ha hirtelen lehűtünk egy rendszert, az valamely magasabb energiájú állapotban "befagy". Viszont ha lassan hűtjük le, akkor folyamatosan egyre kisebb energiájú állapotokon keresztül eléri a minimális energiájú állapotot.

A genetikus algoritmusok vizsgálatával többen foglalkoztak már [2,5]. Mindezen művek azonban azonos szemszögből járják körül a problémát. Ebben a dolgozatban felhasználom a Markov-láncok elméletét, amely hatékony vizsgálati eszköz lehet. Az első fejezet a [3] alapján épül fel, ebben a később felhasználandó tételeket írtam le. A második fejezetben az algoritmusban szereplő fogalmakat definiáltam újra, hogy illeszkedjenek a Markov-láncok elméletéhez. Ebben nagy segítségemre volt [4], innen vettem a *metahalmaz* ötletét. A harmadik fejezet tartalmazza az alkalmazást.

I. Markov-láncok

I.1 Markov-láncok definíciói

1. Definíció: Egy $\{X_t, t \in \mathrm{T}\}$ sztochasztikus folyamatot Markov-láncnak nevezünk, ha

a) állapotainak halmaza megszámlálható

b) $\mathbf{P}(X_t = b | X_{t_1} = a_1, X_{t_2} = a_2, \ldots, X_{t_n} = a_n) = \mathbf{P}(X_t = b | X_{t_n} = a_n)$
$(\forall\ t_1 < t_2 < \ldots < t_n < t)$

2. Definíció: Az X_t Markov-láncot diszkrét paraméterűnek nevezzük, ha $\mathrm{T} \subset \mathbb{N}$.

3. Definíció: Az X_t diszkrét paraméterű Markov-láncot homogén átmenetvalószínűségűnek nevezzük, ha $\mathbf{P}(X_{i+j} = b | X_j = a)$ független j-től.
Jelölések: $P_{i,j}^{(n)} = \mathbf{P}(X_{t+n} = i | X_t = j)$, de $n = 1$ esetén csak $P_{i,j}$-t írunk.

I.2 Markov-láncok állapotainak jellemzése

I.2.1. Definíció: Legyen I egy diszkrét paraméterű, homogén átmenetvalószínűségű Markov-lánc állapotainak halmaza. $i \in \mathrm{I}$ állapotot elérhetetlennek nevezzük, ha $\forall j \in \mathrm{I}\ \ P_{j,i} = 0$.

I.2.2. Definíció: Legyen $i, j \in \mathrm{I}$ egy diszkrét paraméterű, homogén átmenetvalószínűségű Markov-lánc két állapota. Azt mondjuk, hogy i-ből elérhető j, ha $\exists n \in \mathbb{N}$, hogy $P_{i,j}^{(n)} > 0$. Jelölése: $i \to j$. (Természetesen $P_{i,j}^{(0)} = \delta_{i,j}$ értelmezés miatt $i \to i$.) Azt mondjuk, hogy i és j kölcsönösen elérhetőek,ha $i \to j$ és $j \to i$. Jelölése: $i \leftrightarrow j$.

I.2.1 Állítás: A $\leftrightarrow$ reláció ekvivalenciareláció.

Bizonyítás: $i \leftrightarrow i$ a definíció alapján teljesül. A szimmetria is fennáll, ugyanis $\leftrightarrow$ definíciójában i és j szerepe szimmetrikus. Végül pedig a tranzitivitás igazolásához elég belátni, hogy $\to$ tranzitív. Ha $i \to j$ és $j \to k$, akkor $\exists n_1, n_2 \in \mathbb{N}$, hogy $P_{i,j}^{(n_1)} > 0$ és $P_{j,k}^{(n_2)} > 0$. Ekkor $P_{i,k}^{(n_1+n_2)} \geq P_{i,j}^{(n_1)} P_{j,k}^{(n_2)} > 0$. Azaz $i \to k$.

I.2.3. Definíció: Legyen I egy diszkrét paraméterű, homogén átmenetvalószínűségű Markov-lánc állapotainak halmaza. $i \in \mathrm{I}$ állapotot lényegesnek nevezzük, ha $\forall j \in \mathrm{I}$, amelyre $i \to j$, teljesül $j \to i$ is. Egy nem lényeges, elérhető állapotot átmenetinek nevezünk.

I.2.2 Állítás: A lényegesség és az átmenetiség osztálytulajdonság, azaz minden ekvivalenciaosztály elemeire egyszerre teljesülnek vagy nem teljesülnek.

Bizonyítás: Legyen i lényeges és legyen $i \leftrightarrow j$. Ekkor $\forall k \in \mathrm{I}$, amelyre $j \to k$, teljesül $i \to k$ is. Mivel i lényeges, $k \to i$, ebből pedig $k \to j$ is következik. Azaz j is lényeges. Legyen most i átmeneti és legyen $i \leftrightarrow j$. Ez azt jelenti, hogy $\exists k \in \mathrm{I}$, amelyre $i \to k$, de $k \not\to i$. Ekkor $j \to k$ is teljesül és ha $k \to j$ lenne, akkor következne $k \to i$ is, ami ellentmondás. Emiatt j is átmeneti.

I.2.4. Definíció: Az i állapot periódusának nevezzük azon n számok legnagyobb közös osztóját, melyekre teljesül $P_{i,i}^{(n)} > 0$. Jelölése: d_i.

I.2.3 Állítás: Legyen i állapot periódusa d_i és legyen $i \leftrightarrow j$. Ekkor $d_j = d_i$, ezenkívül $\exists r \in \mathbb{N}$ $P_{i,j}^{(n)} > 0 \Rightarrow n \equiv r (\mathrm{mod}\ d_i)$.

Bizonyítás: $i \leftrightarrow j$ miatt $\exists n, m :\ P_{i,j}^{(n)}, P_{j,i}^{(m)} > 0$. Ekkor $\forall s$, amelyre $P_{i,i}^{(s)} > 0$, $P_{j,j}^{(m+n+s)} > 0$ is teljesül. $d_j | m+n$ és $d_j | m+n+s$ miatt $d_j | s$. Mivel d_i az ilyen s számok legnagyobb közös osztója, $d_j | d_i$ is teljesül. Mivel i és j szerepe szimmetrikus, $d_j | d_i$. Eszerint tehát $d_i = d_j$. A másik állítás ennek egyszerű következménye.

I.2.5. Definíció: Jelölje $C_r(i) \subseteq C(i)$ i osztályának azon j elemeit, amelyekre $P_{i,j}^{(n)} > 0 \Rightarrow n \equiv r (\mathrm{mod}\ d_i)$.

I.2.6. Definíció: Legyen $f_{i,j}^{(n)} = \mathbf{P}(X_{t+n} = j | X_t = i, \forall 0 < k < j,\ X_{t+k} \neq j)$.
Legyen $f_{i,j}^{*} = \sum\limits_{n \in \mathbb{N}} f_{i,j}^{(n)}$. $i \in \mathrm{I}$ állapotot visszatérőnek nevezzük, ha $f_{i,i}^{*} = 1$.

I.2.7. Definíció: Legyen $i \in \mathrm{I}$ egy visszatérő állapot. Az $n \mapsto f_{i,i}^{(n)}$ valószínűségeloszlású változót első visszatérési időnek nevezzük. Legyen $m_{i,i} \doteq m_i$ az első visszatérési idő várható értéke. Ha $m_i < \infty$, akkor pozitív visszatérő állapotnak nevezzük, különben pedig nullvisszatérőnek.

I.2.4 Állítás:

$$P_{i,j}^{(n)} = \sum_{k=1}^{n} f_{i,j}^{(k)} \cdot P_{j,j}^{(n-k)}$$

Bizonyítás: Az úgynevezett első elérés módszerével bizonyítunk. Legyen

$$\kappa = \inf\{n \geq 1 : \; X_n = j | X_0 = i\}.$$

Szétválasztjuk az eseményeket κ értéke szerint.$P(\kappa = k) = f_{i,j}^{(k)}$. Így az állítás éppen a teljes valószínűség tétele.

I.2.8. Definíció: $g_{i,j} \doteq \mathbf{P}(X_n = j$ végtelen sok $n-$re$|X_0 = i)$
Mértékelméleti megfontolások alapján

$$g_{i,j} = \lim_{m \to \infty} \mathbf{P}(\exists n > m : \; X_n = j | X_0 = i) = \lim_{m \to \infty} \sum_{k \in I} P_{i,k}^{(m)} \cdot f_{kj}^{*}.$$

I.2.5 Állítás: $g_{i,j} = f_{i,j}^{*} \cdot g_{j,j}$

Bizonyítás: Az első elérés módszerével bizonyítható.

I.2.6 Állítás: $g_{i,j} = f_{i,j}^{*}$, ha j visszatérő, különben pedig $g_{i,j} = 0$.

Bizonyítás: $g_{i,j}(m) \doteq \mathbf{P}(\exists n_m > \ldots > n_1 > 0 : \; X_{n_k} = j | X_0 = i)$.
Az első elérés módszerével belátható $g_{i,j}(m+1) = f_{i,j}^{*} \cdot g_{j,j}(m)$, így $g_{i,j}(m+1) = f_{i,j}^{*} \cdot (f_{j,j}^{*})^{m}$.
$g_{i,j} = \lim_{m \to \infty} g_{i,j}(m)$-ből az állítás következik.

I.3 Markov-láncok határeloszlása

I.3.1 Tétel:

a) $\sum\limits_{n\in\mathbb{N}} P_{i,i}^{(n)}$ divergens, ha i visszatérő állapot

b) $\sum\limits_{n\in\mathbb{N}} P_{i,i}^{(n)} = (1 - f_{i,i}^*)^{-1}$, ha i nem visszatérő.

Bizonyítás: Felhasználjuk a következő lemmát bizonyítás nélkül. Legyenek $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ valós számsorozatok a következő tulajdonságokkal:

a)

$$a_n \geq 0, \ \lim_{n\to\infty} \frac{a_n}{\sum\limits_{i=0}^{n} a_i} = 0$$

b) $\exists b \in \mathbb{R} \ \lim\limits_{n\to\infty} b_n = b$

Ekkor

$$\lim_{n\to\infty} \frac{\sum\limits_{i=0}^{n} a_i \cdot b_{n-i}}{\sum\limits_{i=0}^{n} a_i} = b$$

Az I.2.4 állítás alapján

$$\sum_{n=1}^{N} P_{i,j}^{(n)} = \sum_{n=1}^{N} \sum_{k=0}^{n-1} f_{i,j}^{(n-k)} \cdot P_{j,j}^{(k)} = \sum_{k=0}^{N-1} P_{j,j}^{(k)} \cdot \sum_{n=k+1}^{N} f_{i,j}^{(n-k)} = \sum_{k=0}^{N-1} P_{j,j}^{(k)} \cdot \sum_{n=1}^{N-k} f_{i,j}^{(n)}.$$

$a_n = P_{j,j}^{(n)}$ és $b_n = \sum\limits_{i=1}^{n} f_{i,j}^{(i)}$ esetre felhasználva a lemmát

$$\lim_{n\to\infty} \frac{\sum\limits_{n=1}^{N} P_{i,j}^{(n)}}{\sum\limits_{n=0}^{N} P_{j,j}^{(n)}} = f_{i,j}^*$$

egyenlőséget kapjuk, amiből $j = i$ esetben

$$\lim_{n\to\infty} \frac{\sum\limits_{n=1}^{N} P_{i,i}^{(n)}}{1 + \sum\limits_{n=1}^{N} P_{i,i}^{(n)}} = f_{i,i}^*$$

Ebből az állítás következik.

I.3.2 Tétel: Legyen i visszatérő állapot d_i periódussal. Ekkor $\lim\limits_{n\to\infty} P_{i,i}^{(n\cdot d_i)} = \dfrac{d_i}{m_i}$, emellett $P_{i,i}^{(n\cdot d_i+r)} = 0$, ha $r \not\equiv 0 \pmod{d_i}$.

Bizonyítás: Az állítás második fele nyilvánvaló következménye a periódus definíciójának. Az elsőnek a belátásához felhasználjuk a következő lemmát bizonyítás nélkül: d_i azon n számok legnagyobb közös osztója is, amelyekre $f_{i,i}^{(n)} > 0$ teljesül.

Legyen $r_k = \sum\limits_{n=k+1}^{\infty} f_{i,i}^{(n)}$, ekkor $\sum\limits_{k=0}^{\infty} r_k = \sum\limits_{k=1}^{\infty} k \cdot f_{i,i}^{(k)} = m_i$. Az I.2.4 állításba $f_{i,i}^{(k)} = -(r_k - r_{k-1})$-t helyettesítve:

$$P_{i,i}^{(n)} = -\sum_{k=1}^{n} (r_k - r_{k-1}) \cdot P_{i,i}^{(n-k)}, \text{ azaz}$$

$$\sum_{k=0}^{n} r_k \cdot P_{i,i}^{(n-k)} = \sum_{k=0}^{n-1} r_k \cdot P_{i,i}^{(n-1-k)} = \ldots = r_0 \cdot P_{i,i}^{(0)} = 1$$

Legyen $\lambda = \limsup\limits_{n\to\infty} P_{i,i}^{(n\cdot d_i)}$. Ekkor $\exists n_k$ részsorozat, melyre

$$\lambda = \lim_{k\to\infty} P_{i,i}^{(n_k\cdot d_i)}.$$

Legyen s tetszőleges olyan szám, melyre $f_{i,i}^{(s)} > 0$. Ekkor a lemma alapján $d_i | s$. Ekkor

$$\lambda = \liminf_{k\to\infty} P_{i,i}^{(n_k\cdot d_i)} =$$

$$= \liminf_{k\to\infty} \Big(f_{i,i}^{(s)} \cdot P_{i,i}^{(n_k\cdot d_i - s)} + \sum_{r=1,\ r\neq s}^{n_k} f_{i,i}^{(r)} \cdot P_{i,i}^{(n_k\cdot d_i - r)} \Big) \leq$$

$$\leq f_{i,i}^{(s)} \cdot \liminf_{k\to\infty} P_{i,i}^{(n_k\cdot d_i - s)} + \limsup_{k\to\infty} \sum_{r=1,\ r\neq s}^{n_k} f_{i,i}^{(r)} \cdot P_{i,i}^{(n_k\cdot d_i - r)} =$$

$$= f_{i,i}^{(s)} \cdot \liminf_{k\to\infty} P_{i,i}^{(n_k\cdot d_i - s)} + \lambda \cdot (1 - f_{i,i}^{(s)}).$$

Az egyenlőtlenségből $\lambda \leq \liminf\limits_{k\to\infty} P_{i,i}^{(n_k\cdot d_i - s)}$ következik, amely λ definíciójával együtt $\lambda = \lim\limits_{k\to\infty} P_{i,i}^{(n_k\cdot d_i - s)}$ állításra vezet. Egy ismert számelméleti eredmény alapján ebből $\lim\limits_{n\to\infty} P_{i,i}^{(n\cdot d_i)} = \lambda$ is következik. Emiatt

$$1 = \lim_{n\to\infty} \sum_{k=0}^{n} r_k \cdot P_{i,i}^{(n-k)} = \lambda \cdot \sum_{k=0}^{\infty} r_{k\cdot d_i} = \frac{\lambda}{d_i} \cdot \sum_{k=0}^{\infty} r_k = \lambda \cdot \frac{m_i}{d_i}.$$

Ez azt jelenti, hogy

$$\lambda = \frac{d_i}{m_i}.$$

I.3.3 Tétel: Legyen j visszatérő állapot d_j periódussal. Ekkor ha

a) i ugyanabba az osztályba tartozik és $j \in C_r(i)$, akkor

$$\lim_{n\to\infty} P_{i,j}^{(n\cdot d_j+r)} = \frac{d_j}{m_j}, \text{ emellett } \lim_{n\to\infty} P_{i,j}^{(n\cdot d_j+k)} = 0, \text{ ha } k \not\equiv r \pmod{d_i}$$

b) i másik, lényeges osztályba tartozik, akkor $\forall n \in \mathbb{N} : P_{i,j}^{(n)} = 0$

c) i átmeneti, akkor

$$\lim_{n\to\infty} P_{i,j}^{(n\cdot d_j+r)} = f_{i,j}^*(r) \cdot \frac{d_j}{m_j}, \text{ ahol } f_{i,j}^*(r) = \sum_{k\equiv r \pmod{d_j}} f_{i,j}^{(r)}$$

Bizonyítás: Ha i és j különböző lényeges osztályok, akkor $i \not\to j$ és $j \not\to i$. Azaz valóban $\forall n \in \mathbb{N} : P_{i,j}^{(n)} = 0$. Ha $j \in C_r(i)$, akkor $k \not\equiv r \pmod{d_i}$ esetén az I.2.3 állítás alapján $P_{i,j}^{(n\cdot d_j+k)} = 0$. Az a) és a c) esetre vonatkozó állítás az I.3.2 tétel bizonyításának módszerével belátható.

I.3.4 Tétel: Ha j nem visszatérő állapot, akkor $\forall i \in I : \sum_{n=0}^{\infty} P_{i,j}^{(n)} < \infty$, emiatt $\lim_{n\to\infty} P_{i,j}^{(n)} = 0$.

Bizonyítás: Az I.2.4 állítás alapján

$$\sum_{n=1}^{N} P_{i,j}^{(n)} = \sum_{k=0}^{N-1} P_{j,j}^{(k)} \cdot \sum_{n=1}^{N-k} f_{i,j}^{(n)} \le \sum_{k=0}^{N-1} P_{j,j}^{(k)}, \text{ emiatt}$$

$$\sum_{n=0}^{\infty} P_{i,j}^{(n)} \le \sum_{n=0}^{\infty} P_{j,j}^{(n)} < \infty$$

.

II. Általános modell

II.1 Alapvető definíciók

II.1.1. Definíció: Legyen X egy megszámlálható halmaz. Jelölje Ξ az $X \longmapsto X$ függvények halmazát. Jelölje $\mathcal{P}(X)$ az $\Omega \longmapsto \Xi$ függvények, az úgynevezett sztochasztikus operátorok halmazát.

A sztochasztikus operátorokat úgy is fel lehet fogni, hogy $\forall x \in X$ elemhez egy valószínűségi változót rendelnek, $f(x) : \omega \mapsto f(\omega)(x)$.

II.1.2. Definíció: X-beli metahalmaznak nevezünk egy $m : X \longmapsto \mathbb{N}$ függvényt, ha $c(m) \doteq \sum_{x\in X} m(x) < \infty$. A $c(m)$ számot a metahalmaz számosságának nevezzük. A $\mathcal{H}(m) = \{x \in X | m(x) > 0\}$ halmaz m tartóhalmaza. Legyen

$$\mathcal{M}^{(n)}(X) = \{m\ X\text{−beli metahalmaz}|\ c(m) = n\}$$

$$\text{valamint } \mathcal{M}(X) = \bigcup_{n\in\mathbb{N}} \mathcal{M}^{(n)}(X).$$

II.1.3. Definíció: Legyen $f \in \mathcal{P}(X)$ tetszőleges sztochasztikus operátor. Ekkor képezhető a következő, $\mathcal{M}(f) \in \mathcal{P}(\mathcal{M}(X))$ sztochasztikus operátor: $m \in \mathcal{M}(X)$ esetén legyen

$$((\mathcal{M}(f))(m))(y) \doteq \sum_{x\in X} \sum_{i=1}^{m(x)} \chi\{f_{x,i}(x) = y\}$$

valószínűségi változó, ahol $f_{x,i}$ operátorok függetlenek és eloszlásuk megegyezik f eloszlásával.

II.2 Genetikus tér

II.2.1. Definíció: Legyen Γ egy megszámlálható halmaz, amelyet a genotípusok halmazának nevezünk. $\mathcal{M}(\Gamma)$ halmazt az összes populáció halmazának, míg $\Pi \subseteq \mathcal{M}(\Gamma)$ halmazt a lehetséges populációk halmazának nevezzük.

II.2.2. Definíció: Tegyük fel, hogy Γ előáll $\Gamma_1 \times \ldots \times \Gamma_k$ alakban, ahol $k > 1$. Ekkor Γ-t dekomponálhatónak nevezzük. Γ_i elemei az i. gén alléljai.

II.2.3. Definíció: Egy $\Psi \in \mathcal{P}(\Pi)$ sztochasztikus operátort genetikus operátornak nevezünk. A $(\Pi, \Psi_1, \ldots, \Psi_n)$ struktúrát **genetikus térnek** nevezzük, ha

a) Π a lehetséges populációk halmaza

b) Ψ_i $i = 1, \ldots, n$ Π-n értelmezett független genetikus operátorok.

II.2.1 Állítás: Legyen $(\Pi, \Psi_1, \ldots, \Psi_n)$ genetikus tér. Legyen ξ_0 tetszőleges Π-n értelmezett valószínűségi változó.
A $\xi_{k+1} \doteq \Psi_1(\ldots(\Psi_n(\xi_k)\ldots)$ rekurzióval definiált sztochasztikus folyamat homogén átmenetvalószínűségű Markov-lánc.

Bizonyítás: Mivel ξ_{k+1}-t ξ_k függvényeként definiáltuk, Markov-láncot kapunk. $\forall k \in \mathbb{N}$ esetén ugyanazt a függvényt használjuk, így homogén átmenetvalószínűségű is lesz.

II.2.4. Definíció: S genetikus operátort **szelekciónak** nevezzük, ha

$$\forall m \in \Pi: \ \mathbf{P}(\mathcal{H}(S(m)) \subseteq \mathcal{H}(m)) = 1.$$

II.2.5. Definíció: $P \in \mathcal{P}(\Gamma)$ sztochasztikus operátort **mutációnak** nevezzük, ha

a) $\forall \gamma_1, \gamma_2$ esetén $P_{\gamma_1,\gamma_2} \doteq \mathbf{P}(P(\gamma_1) = \gamma_2) > 0$

b) $\sup_{\gamma \in \Gamma} P_{\gamma,\gamma} < 1$ és $\inf_{\gamma \in \Gamma} P_{\gamma,\gamma} > 0$.

Legyen $\epsilon_*(P) \doteq 1 - \sup_{\gamma \in \Gamma} P_{\gamma,\gamma}$ és $\epsilon^*(P) \doteq 1 - \inf_{\gamma \in \Gamma} P_{\gamma,\gamma}$.
$\epsilon^*(P)$-t mutációs rátának nevezzük.

II.2.6. Definíció: Legyen Γ dekomponálható. A $CO: \Gamma \times \Gamma \longmapsto \Gamma^{\Omega}$ operátort **keresztezésnek** nevezzük, ha $\gamma^1 = (\gamma^1_1, \ldots, \gamma^1_k), \gamma^2 = (\gamma^2_1, \ldots, \gamma^2_k)$ esetén $\gamma^* \doteq CO(\gamma^1, \gamma^2)$ olyan valószinűségi változó, melyre $\forall \omega \in \Omega: \; \gamma^*_j(\omega) = \gamma^1_j$ vagy $\gamma^*_j(\omega) = \gamma^2_j$.

A mutáció és a keresztezés nem genetikus operátorok. Viszont belőlük kiindulva képezhetünk genetikus operátorokat:

a) $M = \mathcal{M}(P)|\Pi$ már genetikus operátor, nevezzük ezt is mutációnak

b) $C = \mathcal{M}(Cr)|\Pi$-t pedig nevezhetjük keresztezésnek, ahol $Cr \in P(\Gamma)$ a következő sztochasztikus operátor:
Legyen $m \in \Pi$ rögzített, γ_2 valószínűségi változó, $P(\gamma_2 = \gamma) = \frac{m(\gamma)}{c(m)}$.
Ekkor $Cr: \; \gamma_1 \mapsto CO(\gamma_1, \gamma_2)$. A jelölés nem pontos, mert Cr kiterjesztésénél minden egyes $m \in \Pi$ esetén ezzel az m-mel tekintjük Cr-t, nem pedig egy rögzítettel.

Ahhoz, hogy a kiterjesztett mutációra teljesüljenek a megkívánt tulajdonságok, Π nem lehet akármilyen. Először is:

$$\forall n \in \mathbb{N} \; (\Pi \bigcap \mathcal{M}^{(n)}(\Gamma) = \emptyset) \vee (\mathcal{M}^{(n)}(\Gamma) \subseteq \Pi).$$

Legyen ugyanis $m \in \mathcal{M}^{(n)}(\Gamma)$ tetszőleges. Ekkor

$$\forall m^* \in \mathcal{M}^{(n)}(\Gamma) \;\; \mathbf{P}((\mathcal{M}(P))(m) = m^*) > 0.$$

Azaz $m \in \Pi$ esetén $\mathcal{M}^{(n)}(\Gamma) \subseteq \Pi$-nek is teljesülnie kell.

Emellett a *II.2.5.b* tulajdonság teljesüléséhez szükséges, hogy $\exists K \in \mathbb{N}: \; m \in \Pi \Rightarrow c(m) \leq K$.

Mivel a genetikus algoritmusoknál $\Pi = \mathcal{M}^{(K)}(\Gamma)$ eset teljesül, a továbbiakban ezt kikötjük.

II.3 Szelekció

II.3.1 Állítás: Ha $m_1 \overset{S}{\leftrightarrow} m_2$, akkor $\mathcal{H}(m_1)=\mathcal{H}(m_2)$.

Bizonyítás: A *II.2.4* tulajdonság miatt $m_1 \overset{S}{\rightarrow} m_2$-ből $\mathcal{H}(m_2) \subseteq \mathcal{H}(m_1)$ következik.

II.3.1. Definíció: Legyenek az S-lényeges állapotosztályok $\Sigma=\{\sigma_1, \sigma_2, \ldots, \sigma_n, \ldots\}$. Jelölje az

$$\bigcup_{n\in\mathbb{N}} \{S^{(n)}(m) \in \sigma_i | S^{(n)}(m) \in \bigcup_{\sigma_j\in\Sigma} \sigma_j\}$$

eseményt $\pi_i(m)$. A $\mathbf{P}(\pi_i(m))$ értéket jelöljük $s_i(m)$-mel.

II.3.2. Definíció: Az $\bigcup\{m \in \Pi : s_i(m) > 0\}$ halmazt σ_i vonzáskörzetének nevezzük és $\mathcal{S}(\sigma_i) \doteq \mathcal{S}_i$-vel jelöljük.
Az $\bigcup\{m \in \Pi : s_i(m) > 1-\delta\}$ halmaz neve pedig σ_i δ-vonzáskörzete, jele $\mathcal{S}_i(\delta)$.

II.3.3. Definíció: $m \in \mathcal{S}_i$ esetén legyen

$$\tau_i(m)=\inf\{n \in \mathbb{N} : S^{(n)}(m) \in \sigma_i | \pi_i(m)\}$$

valószínűségi változó. Ezenkívül legyen

$$\tau(m)=\inf\{n \in \mathbb{N} : S^{(n)}(m) \in \bigcup_{\sigma_i\in\Sigma} \sigma_i\}.$$

II.3.1 Tétel: $\mathbf{E}\tau(m)=\sum\limits_{i:m\in\mathcal{S}_i} \mathbf{E}\tau_i(m) \cdot s_i(m)$.

Bizonyítás: A $\pi_i(m)$ eseményen $\tau(m)=\tau_i(m)$, erre alkalmazhatjuk a teljes-várhatóérték tételt.

II.3.2 Tétel:

$$\mathbf{E}\tau_i(m)=1+\sum_{m^*\in\mathcal{S}_i} \mathbf{E}\tau_i(m^*) \cdot \frac{\mathbf{P}(S(x)=m^*|x=m)}{\mathbf{P}(S(m)\in\mathcal{S}_i)}, \quad \text{ha } m \notin \sigma_i.$$

Bizonyítás: Szintén a teljes-várhatóérték tételből következik.

II.3.3 Tétel:

$$s_i(m) = \sum_{m^* \in \mathcal{S}_i} s_i(m^*) \cdot \frac{\mathbf{P}(S(x)=m^*|x=m)}{\mathbf{P}(S(m) \in \mathcal{S}_i)}, \quad \text{ha } m \notin \sigma_i.$$

Bizonyítás: Ez pedig a teljes valószínűség tételből következik.

Ezek az egyenletek az $\mathbf{E}\tau_i(m) = 0$ és $s_i(m) = 1$, ha $m \in \sigma_i$ feltételekkel együtt egy egyértelműen megoldható lineáris egyenletrendszert alkotnak.

II.4 Mutáció

II.4.1 Állítás: Legyen Γ véges. Ekkor a $\overset{M}{\leftrightarrow}$ reláció szerint egyetlen osztály van, amely lényeges, sőt pozitív visszatérő. Ezenfelül a állapotok periódusa 1.

Bizonyítás: $\forall \gamma_1, \gamma_2 : \; P(\gamma_1, \gamma_2) > 0$, emiatt $\forall m_1, m_2 \in \Pi : \; M(m_1, m_2) > 0$, azaz $\forall m_1, m_2 \in \Pi : \; m_1 \overset{M}{\to} m_2$. Mivel Γ véges és $\Pi = \mathcal{M}^{(K)}(\Gamma)$, $\exists \epsilon = \min\limits_{m_1, m_2 \in \Pi} M(m_1, m_2) > 0$. Tehát egyetlen osztály van, ez nyilvánvalóan lényeges. Legyen m egy rögzített állapot. Elég ha erről belátjuk, hogy pozitív visszatérő, mert ez osztálytulajdonság. Legyen tehát $f_n = \mathbf{P}(M^{(k)}(m) \neq m, k=1, \ldots, n)$. $1 - f_n = \sum\limits_{k=1}^{n} f_{m,m}^{(k)}$ miatt m pontosan akkor visszatérő, ha $\lim\limits_{n\to\infty} f_n = 0$. Mivel legalább ϵ valószínűséggel m állapotba kerülünk, akármi is volt az előző állapot, $f_{n+1} \leq f_n \cdot (1 - \epsilon)$. Ebből $f_n \leq (1 - \epsilon)^n$ következik, azaz $\lim\limits_{n\to\infty} f_n = 0$. Tehát m visszatérő, így $\sum\limits_{k=n+1}^{\infty} f_{m,m}^{(k)} = 1 - \sum\limits_{k=1}^{n} f_{m,m}^{(k)} = f_n$, ebből

$$m_m = \sum_{n=1}^{\infty} n \cdot f_{m,m}^{(n)} = \sum_{n=1}^{\infty} \sum_{k=n+1}^{\infty} f_{m,m}^{(k)} = \sum_{n=1}^{\infty} f_n \leq \frac{1}{\epsilon} < \infty.$$

Ez azt jelenti, hogy m pozitív visszatérő is. Már csak azt kell belátni, hogy a periódusa 1, ez pedig $M_{m,m} > 0$ miatt egyértelmű.

Következmény: Legyen $\Psi \in \mathcal{P}(\Pi)$ tetszőleges genetikus operátor. Ekkor a $\overset{M \circ \Psi}{\leftrightarrow}$ reláció szerint is egyetlen, pozitív vissztérő osztály van.

II.4.1. Definíció: Legyen Γ dekomponálható, $\Gamma = \Gamma_1 \times \cdots \times \Gamma_k$. Legyen P_i mutáció Γ_i-n. Ekkor a $P : \gamma \mapsto \gamma^*$ mutációt a $P_1, \ldots, P_k$ mutációk kompozíciójának nevezzük, ha $\gamma_i^* = P_i(\gamma_i)$.
Jelölése: $P = P_1 \times \cdots \times P_k$.

II.4.2 Állítás: Ha P a $P_1, \ldots, P_k$ mutációk kompozíciója, akkor

$$\epsilon^*(P) = 1 - \prod_{i=1}^{k}(1 - \epsilon^*(P_i)) \approx \sum_{i=1}^{k} \epsilon^*(P_i), \text{ ha} \sum_{i=1}^{k} \epsilon^*(P_i) \ll 1.$$

Bizonyítás: A közelítés $e^{-x} \approx 1 - x$-ből adódik, ha $x \ll 1$. Az egyenlőség pedig abból adódik, hogy $P_1, \ldots, P_k$ függetlenek, így

$$\mathbf{P}(P(\gamma_1, \ldots, \gamma_k) = (\gamma_1, \ldots, \gamma_k)) = \prod_{i=1}^{k} \mathbf{P}(P_i(\gamma_i) = \gamma_i).$$

Megjegyzés: Az állítás úgy pontosabb, -bár kevésbé fejezi ki a lényeget- hogy

$$\exists \delta : \sum_{i=1}^{k} \epsilon^*(P_i) < \delta \Rightarrow \epsilon^*(P) < 2 \cdot \sum_{i=1}^{k} \epsilon^*(P_i).$$

II.4.3 Állítás:

$$\epsilon^*(M) \approx K \cdot \epsilon^*(P), \text{ ha } K \cdot \epsilon^*(P) \ll 1.$$

Bizonyítás: A konvenciónk értelmében $\Pi = \mathcal{M}^{(K)}(\Gamma)$. Az előző állításhoz hasonló közelítést alkalmazunk. Egy populációt a mutáció nem változtat meg, ha egyetlen egyed sem változik. Ehhez képest elenyészően kicsi annak a valószínűsége, hogy az egyedek pontosan úgy változzanak, hogy a populáció másik egyede legyen belőlük.

Következmény: Ha $P = P_1 \times \cdots \times P_k$ és $K \cdot \sum_{i=1}^{k} \epsilon^*(P_i) < \delta$, akkor $\epsilon^*(M) < 4K \cdot \sum_{i=1}^{k} \epsilon^*(P_i)$.

II.4.2. Definíció: Legyen Γ dekomponálható. Jelölje d_H az úgynevezett Hamming-távolságot, amit a következőképp értelmezünk:

$$d_H(\gamma_1,\gamma_2)=card\{i\in\mathbb{N}:\ \gamma_1^i\neq\gamma_2^i\}.$$

Legyen d_H^Π Π-beli távolság d_H alapján:

$$d_H^\Pi(m_1,m_2)=\min\{\sum_{k=1}^{K}d_H(\gamma_{1,k},\gamma_{2,k})\ :\ \forall\gamma\in\Gamma$$

$$card\{i\in\mathbb{N}:\ \gamma_{1,i}=\gamma\}=m_1(\gamma),\ card\{i\in\mathbb{N}:\ \gamma_{2,i}=\gamma\}=m_2(\gamma)\}.$$

Ez az alábbi definíció azt az elképzelést fejezi ki, hogy két populáció távolsága úgy áll elő, hogy párosítjuk az egyedeket a két populációból és a párok távolságainak összegét tekintjük. Természetesen ez a távolság függene a párosítástól, ezért vesszük az összes lehetséges párosításra a minimumot.

II.5 Keresztezés

II.5.1. Definíció: Legyen Γ dekomponálható, $\Gamma=\Gamma_1\times\ldots\times\Gamma_k$. A $\Delta=\Gamma_1^*\times\ldots\times\Gamma_k^*$ halmazt **sémának** nevezzük, ha $i=1,\ldots,k:\ \Gamma_i^*\subseteq\Gamma_i$.

II.5.2. Definíció: Δ sémát egyszerűnek nevezzük, ha $\Gamma_i^*=\Gamma_i$ vagy $\Gamma_i^*=\{\gamma^i\}$. Legyen $I=\{i:\ \Gamma_i^*=\{\gamma^i\}\}$ és $r(\Delta)=card(I)$, ekkor a sémát $\Delta^I_{\gamma^{i_1},\ldots,\gamma^{i_r}}$-vel jelöljük.

II.5.3. Definíció: Legyen $H\subseteq\Gamma$ egy tetszőleges nemüres halmaz. A H által generált sémának nevezzük a legszűkebb $\Delta(H)$ sémát, melyre $H\subseteq\Delta(H)$ teljesül. Az m populáció által generált sémán a $\Delta(\mathcal{H}(m))$ sémát értjük.

II.5.4. Definíció: Legyen Δ tetszőleges séma. Legyen

$$\Pi(\Delta)=\{m\in\Pi:\ \Delta(\mathcal{H}(m))\subseteq\Delta\}.$$

$\Pi(\Delta)$-t a Δ sémára illeszkedő populációk halmazának nevezzük.

II.5.1 Állítás: A keresztezés az m populációt egy $\Delta(\mathcal{H}(m))$-ra illeszkedő populációba viszi.

Bizonyítás: Cr_m operátor $\forall\gamma \in \Delta(\mathcal{H}(m))$ genotípust $\Delta(\mathcal{H}(m))$-beli genotípusba visz át. Ha tehát tekintjük C-t, a kiterjesztés tulajdonságai miatt az eredmény egy $\Delta(\mathcal{H}(m))$-ra illeszkedő populáció.

II.5.5. Definíció: m populáció átmérőjén az $r(\Delta(\mathcal{H}(m))$ számot értjük.

II.6 Fitness

II.6.1. Definíció: Az $f : \Gamma \longmapsto \mathbb{R}$ függvényt fitnessnek nevezzük, ha $\forall\gamma \in \Gamma$ esetén $f(\gamma) > 0$.

II.6.2. Definíció: Legyen f fitness, $m \in \Pi$ tetszőleges populáció.

$$f(m) \doteq \sum_{x\in\Gamma} \sum_{i=1}^{m(x)} f(x).$$

Az $f_m(\gamma) \doteq \frac{f(\gamma)}{f(m)}$, $(m(\gamma) > 0)$ értéket γ egyed relatív fitnessének nevezzük.

II.6.3. Definíció: Legyen f fitness, $m \in \Pi$ tetszőleges populáció, Δ tetszőleges séma.

$$f(\Delta) \doteq \sum_{x\in\Delta} \sum_{i=1}^{m(x)} f(x).$$

Az $f_\Delta(\gamma) \doteq \frac{f(\gamma)}{f(\Delta)}$, $(\gamma \in \Delta,\ m(\gamma))$ értéket γ egyed Δ sémára illeszkedő relatív fitnessének nevezzük.

II.6.4. Definíció: Legyen f fitness, $m \in \Pi$ tetszőleges populáció. Legyen $f^* : \Pi \longmapsto \mathcal{M}^{(1)}(\Gamma)$ a következő sztochasztikus operátor:
$f^*(m)(\gamma) = 1,\ m(\gamma) \cdot f_m(\gamma)$ valószínűséggel.

Az

$$S_f : m \mapsto \sum_{i=1}^{c(m)} f_i^*(m)$$

sztochasztikus operátort az f fitness által vezérelt szelekciónak nevezzük.

II.6.1 Állítás: S_f szelekció.

Bizonyítás: f^* definíciója alapján $\mathcal{H}(f^*(m)) \subseteq \mathcal{H}(m)$ teljesül.

$$\mathcal{H}(S_f(m)) = \bigcup_{i=1}^{c(m)} \mathcal{H}(f_i^*(m))$$

miatt az állítás igaz.

II.6.2 Állítás: Az S_f-lényeges állapotok pontosan azok, melyeknek a tartóhalmaza egyelemű. Mindegyik ilyen állapot egyetlen osztályt alkot, melynek a vonzáskörzetében pontosan azok a populációk vannak, amelyek tartóhalmaza tartalmazza azt az egy genotípust.

Bizonyítás: $\mathcal{H}(m^*) \subseteq \mathcal{H}(m)$ esetén $\mathbf{P}(S_f(m) = m^*) > 0$. Azaz $m \to m^* \iff \mathcal{H}(m^*) \subseteq \mathcal{H}(m)$. Ebből következik az állítás.

III. Genetikus algoritmusok

III.1 Definíciók

III.1.1. Definíció: Legyen $\Gamma=\{0,1\}^k$ a genotípusok halmaza. Legyen P_* mutáció $\{0,1\}$-n, $\mathbf{P}(P_*(x)=x) = 1-p,\ x=0,1$. Legyen $P = P_* \times \cdots \times P_*$ mutáció Γ-n. Legyen CO keresztezés Γ-n, $\mathbf{P}(CO_i(\gamma^1,\gamma^2) = \gamma_i^1) = \frac{1}{2}$. Legyen $f\ :\ \Gamma \longmapsto \mathbb{R}_+$ fitness. Legyen $\Pi=\mathcal{M}^{(K)}(\Gamma)$, S_f az f fitness által generált szelekció, M a P által generált mutáció, C pedig a CO által generált keresztezés Π-n. A következő eljárást nevezzük **genetikus algoritmusnak**:
Kiválasztunk egy $m_0 \in \Pi$ populációt, végrehajtjuk a szelekciót, aztán a keresztezést végül a mutációt, így kapjuk $m_1 = M(C(S_f(m_0)))$ populációt. Ugyanígy kapható m_{n+1} populáció m_n-ből. Egy megállítási kritérium alapján eldöntjük, mely n-re hagyjuk abba a műveletet (ez lehet előre rögzített n, de lehet pl. olyan kritérium is, hogy $|f(m_{n+1} - f(m_n))| < \epsilon$).

III.2 Eredmények

A II.4.1 állítás következménye alapján egyetlen, pozitív visszatérő osztály van. Az I.3.3 tétel alapján tehát a Markov-láncnak $\exists!$ stacionárius eloszlása és bármely kezdeti eloszlásból kiindulva ehhez a stacionárius eloszláshoz konvergál. Ebben az eloszlásban minden állapot valószínűsége pozitív.

Mivel a genetikus algoritmussal a Markov-láncnak egy trajektóriáját járjuk végig, minden időpontban egy konkrét állapotban vagyunk. Az állapotok stacionárius eloszlásbeli valószínűsége azt adja meg, aszimptotikusan mennyi időt töltünk abban az állapotban.

Az átmenetvalószínűségmátrix ismeretében a stacionárius eloszlás meghatározható, legalábbis elvileg. A mátrix 1 sajátértékhez tartozó sajátvektorát kell meghatározni. Genetikus algoritmusok esetén pontosan egy ilyen sajátérték van, de az állapotok hatalmas száma miatt az egyenletek felírása is nehézséget okoz, nemhogy a megoldásuk.

Szerencsére nekünk elég az a tény, hogy 1 valószínűséggel lesz olyan állapot, amely tartalmaz maximális fitness-szel rendelkező genotípust. Sőt 1 valószínűséggel végtelen sok ilyen állapot van. A minket érdeklő kérdések inkább azok, hogy

a) mikor jutunk el egy ilyen állapotba?

b) hogyan ismerjük fel, hogy ilyen állapotba jutottunk?

Az első kérdésre adott válaszunk azon alapul, hogy csak a mutáció és a keresztezés hoz létre új genotípusokat. Annak a valószínűsége, hogy m populáció a mutációval olyan populációvá alakul, amely már tartalmaz γ^* genotípust: $P_m = 1 - \prod_{\gamma \in \Gamma} \prod_{i=1}^{m(\gamma)} (1 - \mathbf{P}(P(\gamma) = \gamma^*))$. Legyen $\epsilon = \min_{m \in \Pi} P_m$, ekkor $\frac{1}{\epsilon}$ felső becslése az ilyen állapotba való eljutás várható értékének. Ha figyelembe vesszük a keresztezést is, elég, ha olyan populációba mutálódik, amelyre a generált sémának eleme γ^*. Ez jóval kisebb várható időtartamot ad, ezután már csak meg kell várnunk, hogy a keresztezés "kikeverje" a megfelelő genotípust. Természetesen a szelekció meggyorsíthatja a folyamatot, azaz olyan populáció irányába mozoghat, ahonnan könnyebb elérni a kívánt genotípust.

A második kérdésre adott válasznál kell figyelembe vennünk jelentősebb mértékben a szelekciót. Az S-lényeges állapotok éppen az egyetlen genotípust tartalmazó populációk. Minél magasabb fitness tartozik az adott genotípushoz, annál nagyobb az állapotnak a δ-vonzáskörzete. Azaz a mutáció és a keresztezés kisebb valószínűséggel vezetnek ki az adott vonzáskörzetből. Ez azt jelenti, hogy nagyobb az ebben az állapotban tartózkodás várható ideje. Tehát meghatározhatunk egy olyan időintervallumot, hogy ha az algoritmus legalább ilyen hosszú ideig tartózkodik egy adott állapot környezetében, akkor megtaláltuk a maximális fitnessű genotípust. Természetesen ehhez alacsony mutációs ráta kell, hogy a szelekció elég gyorsan visszatérjen a lényeges állapotba, azaz átmeneti állapotban kevés mutáció legyen.

Köszönetnyilvánítás

Szeretnék köszönetet mondani:

Kovács Margitnak, témavezetőmnek
a türelméért
Fritz Józsefnek
kritikai megjegyzéseiért
Török Jánosnak és Czirók Andrásnak
a TeX-ben nyújtott segítségükért
Richard Dawkinsnak
a könyveiből merített ihletért

H i v a t k o z á s o k

[1] van Laarhoven, Petrus Joseph Maria:
Theoretical and Computational Aspects of Simulated Annealing

[2] Holland, John H.:
Adaptation in Natural and Artificial Systems
Cambridge, Massachusetts - London

[3] Chung, Kai Lai:
Markov chains with stationary transition probabilities
Berlin - Heidelberg - New York
Springer, 1969.

[4] Jones, Terry:
Evolutionary Algorithms, Fitness Landscapes and Search
the University of New Mexico, Albequerque, New Mexico
1995. May

[5] Feistel, R.; Eberling, W.:
Evolution of Complex Systems
Berlin
1989.