eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

Lengoaia eta Sistema Informatikoak Saila

Bilboko Industria Ingeniaritza Teknikoko Unibertsitate Eskola

Lengoaiak, Konputazioa eta Sistema Adimendunak

Kudeaketaren eta Informazio Sistemen Informatikaren Ingeniaritzako Gradua

2. Maila

2015-16 ikasturtea

2. Gaia: Lengoaiak

JOSE GAINTZARAIN IBARMIA

Azken eguneraketa: 2015-9-2

GAIEN AURKIBIDEA

2. LENGOAIAK

Egoera-makinek karaktere-kateak prozesatu beharko dituzte. Gai honetan karaktere-kateekin erlazionatutako kontzeptuak azalduko dira: sinboloak, alfabetoak, hitzak, hitzen gaineko eragiketak, alfabeto baten gainean era daitezkeen hitz denen multzoa, hitz-multzoak, lengoaiak, alfabeto baten gainean defini daitezkeen lengoaia guztiez osatutako multzoa eta, bukatzeko, lengoaien gaineko eragiketak.

2.1 Alfabetoak eta Hitzak

Azpi-atal honetan, sinbolo kontzeptutik abiatu eta, hasteko, alfabetoa, hitza eta alfabeto baten gainean defini daitezkeen hitz denen multzoa definituko dira. Bukatzeko, hitzen gaineko eragiketak azalduko dira.

2.1.1 Sinboloak

Sinboloak ordenagailuan erabil daitezkeen letrak, digituak, puntuazio-markak eta beste elementuak (geziak, karaktere bereziak eta abar) izango dira.

2.1.2 Alfabetoak

Alfabeto bat sinboloz osatutako multzo **finitua** eta **ez hutsa** da. Jarraian adibide batzuk ikus daitezke:

$$A_1 = \{a, b, c, \ldots, z\} \qquad A_2 = \{a, b\} \qquad A_3 = \{x, y, z\}$$

$$A_4 = \{0, 1\} \qquad A_5 = \{0, 1, 2, 3, \ldots, 9\} \qquad A_6 = \{1, \#, \sqcup\}$$

$$A_7 = \{a, b, c, A, B, \#, \sqcup, @, \&\} \qquad A_8 = \{0, 1, \leftarrow, \uparrow, \rightarrow, \downarrow\} \qquad A_9 = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$$

2.1.3 Hitzak

A alfabetoa emanda, A alfabetoaren gaineko **hitza** A alfabetoko sinboloz osatutako sekuentzia **finitua** da. Sinboloak errepikatuta ager daitezke hitzean. Adibidez, aurreko ataleko A_2 alfabetoa hartzen badugu, honako sinbolo-sekuentzia hauek A_2 alfabetoaren gaineko hitzak dira:

$$aaaaa \qquad ba \qquad a \qquad bbbbabbbbaab$$

Sekuentzia hutsa izan daiteke eta sekuentzia hutsa edo **hitz hutsa** ε sinbolo bereziaren[1] bidez adierazten da.

A alfabetoaren gaineko hitzaren kontzeptua honela defini daiteke errekurtsibitatearen bidez era formalagoan:

- ε sekuentzia hutsa A alfabetoaren gaineko hitza da.
- A alfabetoko α sinbolo bat eta A alfabetoaren gaineko w hitz bat hartuz, αw hitza A alfabetoaren gaineko hitza da.

[1] ε: epsilon

- Aurreko bi puntu horietara egokitzen diren elementuak bakarrik dira A alfabetoaren gaineko hitzak.

Edozein hitz ε hitzetik abiatu eta ezkerretik sinboloak erantsiz osa daiteke. Esate baterako $A = \{a, b, c\}$ alfabetoaren gaineko $abaac$ hitza urratsez urrats honela osa daiteke:

$$\varepsilon \Rightarrow c\varepsilon \Rightarrow ac\varepsilon \Rightarrow aac\varepsilon \Rightarrow baac\varepsilon \Rightarrow abaac\varepsilon$$

ε bakarrik ez dagoenean ez da idazten, beraz, $abaac\varepsilon$ idatzi beharrean $abaac$ idatziko da, nahiz eta kontzeptualki berdinak izan.

A alfabetoa emanda, A alfabetoaren gainean definitutako **hitz finitu denez osatutako multzoa** A^* bezala adieraziko da. Adibidez, $A = \{0, 1\}$ baldin bada, orduan

$$A^* = \{\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111, 0000, \ldots\}$$

A^* multzoa **infinitua** izango da edozein A alfabetorentzat.

2.1.4 Hitzen gaineko eragiketak

Jarraian hitzen gaineko sei eragiketa definituko dira.

Hitzen luzera

w A alfabetoaren gaineko hitza baldin bada, w-ren **luzera** w hitza osatzen duten sinboloen kopurua da eta $|w|$ bezala adieraziko da. Luzera kalkulatzen duen funtzioaren mota honako hau da:

$$A^* \to I\!N$$

Motaren bidez, A^* motako elementu bat (hau da, hitz bat) hartuta zenbaki arrunt[2] bat lortuko dela adierazten da.

Adibidez, $A = \{a, b\}$ alfabetoaren gainean definitutako $aaaaa$, ba, a eta $bbbbabbbbaab$ hitzak hartzen baditugu, $|aaaaa| = 5$, $|ba| = 2$, $|a| = 1$ eta $|bbbbabbbbaab| = 12$ beteko da. Hitz hutsaren luzera 0 izango da, hau da, $|\varepsilon| = 0$.

Sinbolo baten agerpen-kopurua *hitz batean*

A alfabetoaren gainean definitutako w hitz bat eta A alfabetoko α sinbolo bat hartuz, $|w|_\alpha$ espresioaren bidez α sinboloa w hitzean zenbat aldiz agertzen den adieraziko dugu. Funtzio horren mota honako hau da:

$$A^* \times A \to I\!N$$

Motaren bidez, A^* motako elementu bat (hau da, hitz bat) eta A alfabetoko sinbolo bat hartuta zenbaki arrunt bat itzuliko dela adierazten da.

Adibidez, $A = \{a, b\}$ alfabetoaren gaineko ba, a, $bbbbabbbbaab$ eta ε hitzak hartuz, $|ba|_b = 1$, $|a|_b = 0$, $|bbbbabbbbaab|_a = 3$, $|bbbbabbbbaab|_b = 9$ eta $|\varepsilon|_b = 0$ betetzen da.

[2] Zenbaki arrunten multzoari $I\!N$ deituko diogu eta bertan osoak diren zenbaki ez-negatiboak daude: $I\!N = \{0, 1, 2, 3, \ldots\}$

Posizio bateko sinboloa

A alfabetoaren gainean definitutako w hitz bat eta 1 eta $|w|$-ren arteko k zenbaki oso bat (hau da, $1 \leq k \leq |w|$) hartzen baditugu, $w(k)$-ren bidez ezkerretik hasita k posizioan dagoen sinboloa adieraziko dugu. Funtzio horren mota honako hau da:

$$A^* \times \mathbb{N} \rightarrow A$$

Motaren bidez, A^* motako elementu bat (hau da, hitz bat) eta zenbaki arrunt bat hartuta, A alfabetoko sinbolo bat itzuliko dela adierazten da.

Adibidez, $A = \{a, b\}$ alfabetoaren gainean definitutako $bbbbaaab$ hitza hartzen badugu, honako hauek beteko dira:

$$bbbbaaab(2) = b \qquad\qquad bbbbaaab(5) = a$$

Bi hitz kateatu

A alfabetoaren gainean definitutako w eta v bi hitz emanda, bi hitz horiek kateatuz lortzen de z hitza wv hitza da. Funtzio horren mota honako hau da:

$$A^* \times A^* \rightarrow A^*$$

Motaren bidez, A^* motako bi elementu (hau da, bi hitz) hartuta, A^* motako elementu bat (beste hitz bat) itzuliko dela adierazten da.

w eta v hitzak kateatuz lortzen den z hitzak honako propietate hauek izango ditu:

- $|w| = m$ eta $|v| = n$ baldin badira, orduan $|z| = m + n$
- $|w| = m$ eta $1 \leq j \leq m$ baldin badira, orduan $z(j) = w(j)$
- $|w| = m$, $|v| = n$ eta $m + 1 \leq j \leq m + n$ betetzen badira, orduan $z(j) = v(j - m)$

Adibidez, $A = \{a, b\}$ alfabetoaren gainean definitutako $w = aa$ eta $v = ba$ hitzak hartzen baditugu, wv hitza $aaba$ izango da, eta vw hitza $baaa$ izango da.

Kateatze-eragiketa **elkarkorra** da, hau da, w, v eta u alfabeto beraren gainean definitutako hiru hitz baldin badira, $w(vu) = (wv)u$ beteko da. Beraz, parentesiak erabili beharrik ez dago. Adibidez, u, v, w eta z hitzak kateatzeko ez genuke parentesirik erabiliko, zuzenean $uvwz$ idatziko genuke.

Kateatze-eragiketarentzat **elementu neutroa** ε hitz hutsa da. Beraz, $\varepsilon w = w = w\varepsilon$ beteko da w edozein hitz izanda ere.

Berreketa

A alfabetoaren gainean definitutako w hitza eta j zenbaki arrunta emanda, w ber j kalkulatzerakoan lortzen den z hitza w^j bezala adieraziko dugu eta w hitza j aldiz kateatuz lortzen da: $\underbrace{w \cdots w}_{j \text{ aldiz}}$.

Berreketaren mota honako hau da:

$$A^* \times \mathbb{N} \rightarrow A^*$$

Motaren bidez, A^* motako elementu bat (hau da, hitz bat) eta zenbaki arrunt bat hartuta, A^* motako elementu bat (beste hitz bat) itzuliko dela adierazten da.

w^j kalkulatuz lortutako z hitzak honako propietate hau beteko du: $|w| = n$ betetzen bada, orduan $|z| = n \times j$ (hemen $\times$ sinboloak zenbakien arteko biderketa adierazten du).

Gainera $w^j w^k = w^{j+k}$, $w^1 = w$ eta $w^0 = \varepsilon$ beteko dira.

Adibidez, $A = \{0, 1\}$ alfabetoaren gainean definitutako $w = 100$ hitza eta $j = 3$ balioa hartzen baditugu, w^j hitza 100100100 izango da. Bertan w hitza 3 aldiz agertzen da: $\underbrace{100}_{w}\underbrace{100}_{w}\underbrace{100}_{w}$.

Alderantzizko hitza

A alfabetoaren gainean definitutako w hitza emanda, w-ren alderantzizkoa bere sinboloak alderantzizko ordenean ipiniz lortzen da eta w^R bezala adieraziko dugu.[3]
Alderantzizkoa kalkulatzen duen funtzioaren mota honako hau da:

$$A^* \to A^*$$

Motaren bidez, A^* motako elementu bat (hau da, hitz bat) hartuta, A^* motako elementu bat (beste hitz bat) itzuliko dela adierazten da.

w-ren alderantzizkoa kalkulatzerakoan lortutako w^R hitzak honako propietateak izango ditu:

- $|w^R| = |w|$
- $\forall k(1 \leq k \leq |w| \to w^R(k) = w(|w| + 1 - k)$

Adibidez, $A = \{0, 1\}$ alfabetoaren gainean definitutako $w = 100$ hitza hartzen badugu, w^R hitza 001 izango da.

Kasu berezi bezala, α sinboloa A alfabetoko elementu bat baldin bada, $\alpha^R = \alpha$ beteko da. Adibidez, $A = \{a, b\}$ alfabetoa hartzen badugu, $a^R = a$ eta $b^R = b$ beteko da. Hitz hutsaren kasuan $\varepsilon^R = \varepsilon$ beteko da.

2.2 Lengoaiak: Definizioa eta eragiketak

2.2.1 Definizioa

2.1.3 atalean esan den bezala, A^* multzoa A alfabetoaren gainean defini daitezkeen hitz guztiez osatuta dago eta multzo infinitua da beti.

A^*-ren edozein azpimultzo (finitua edo infinitua, hutsa edo ez hutsa) A **alfabetoaren gainean definitutako lengoaia bat da**. A^* bera ere A alfabetoaren gainean definitutako lengoaia da. A^* lengoaia A**-ren gainean definitutako lengoaia unibertsala** bezala ezagutzen da. A-ren gainean definitutako beste lengoaia denak A^*-ren azpimultzoak dira.

Adibidez, $A = \{0, 1\}$ alfabetoa hartzen badugu, jarraian aurkezten diren L_1, L_2, L_3, L_4 eta L_5 multzoak A-ren gainean definitutako lengoaiak dira:

$$L_1 = \{00, 101, 111, 0000\} \quad L_2 = \{\varepsilon, 0, 00, 000, 0000, \ldots\} \quad L_3 = \varnothing \quad L_4 = \{\varepsilon\} \quad L_5 = \{0, 1\}$$

L_1 lengoaia finitua da, bakarrik lau hitz ditu. L_2 lengoaia infinitua da eta batekorik ez duten eta $A = \{0, 1\}$ alfabetoaren gainean definitutakoak diren hitz denez osatuta dago. Lengoaia bat era formalean definitzerakoan eten-puntuak ezin dira erabili. L_2 era formalean honela definitu beharko genuke:

$$L_2 = \{w \mid w \in A^* \wedge \forall k(1 \leq k \leq |w| \to w(k) = 0)\}$$

Beraz A^* multzoko w hitza L_2 lengoaiakoa izango da 1 eta w-ren luzeraren artean dagoen edozein posiziorentzat, posizio horretako elementua 0 baldin bada. ε hitzarentzat ere baldintza hori bete egiten da, izan ere, $|\varepsilon| = 0$ denez, eremu hutseko formula unibertsala izango baikenuke kasu horretan.

L_2 beste era honetan ere defini dezakegu:

$$L_2 = \{w \mid w \in A^* \wedge \underbrace{|w|_1 = 0}_{\text{zero bateko}}\}$$

[3] R hori ingeleseko "Reverse" hitzetik dator

L_2 definitzeko hirugarren era bat honako hau da:

$$L_2 = \{w \mid w \in A^* \wedge \underbrace{|w| = |w|_0}_{\text{luzera = zero-kopurua}}\}$$

L_2 definitzeko laugarren era bat honako hau da:

$$L_2 = \{w \mid w \in A^* \wedge \exists k (k \geq 0 \wedge w = 0^k)\}$$

L_2 definitzeko bosgarren aukera bat beste hau da:

$$L_2 = \{w \mid w \in A^* \wedge \neg\exists v, u (v \in A^* \wedge u \in A^* \wedge w = v1u)\}$$

L_2 definitzeko bosgarren aukera hau ulertzeko, honako hau hartu beharko da kontuan: Hitz batek 1 sinboloaren agerpenen bat badu, orduan hitz hori hiru azpihitzetan zati daiteke erdikoa 1 sinboloaz osatutako hitza izanda. Adibidez 00101001 hitza honela zati daiteke: $\underbrace{00}_{v}1\underbrace{01001}_{u}$. Beraz hitz hori ez da L_2 lengoaiakoa. Hitz hori bera beste era honetan ere zati daiteke: $\underbrace{0010}_{v}1\underbrace{001}_{u}$. Eta baita beste era honetan ere: $\underbrace{0010100}_{v}1\underbrace{\varepsilon}_{u}$. Baina 0000 hitza hartzen badugu, hitz hori ezin da $v1u$ erako azpihitzetan banandu, eta horregatik 0000 hitza L_2 lengoaiakoa da. Beraz, L_2-ren bosgarren definizio honetan L_2 lengoaia zatiketa hori onartzen ez duten hitzez osatuta dagoela esaten da.

L_3 lengoaia hutsa da, ez du hitzik. Multzo hutsa $\varnothing$ sinbolo bereziaren bidez adieraziko dugu. Bestalde, L_4 lengoaiak hitz bakarra du, hitz hutsa. L_3 eta L_4 desberdinak direla garbi eduki behar da. L_5 lengoaia A alfabetoaren berdina da.

A **alfabetoaren gainean defini daitezkeen lengoaia denen multzoa** $\mathcal{P}(A^*)$ bezala edo 2^{A^*} bezala adierazten da. Teknikoki 2^{A^*} espresioak $A^* \to \{0,1\}$ erako funtzio denez osatutako multzoa adierazten du. Demagun $A = \{a, b, c\}$ alfabetoa eta jarraian erakusten den bezala definitutako f eta g funtzioak ditugula:

$$f : A^* \to \{0,1\}$$

$$f(w) = \begin{cases} 1 & |w|_b = 0 \text{ eta } |w|_c = 0 \text{ baldin badira} \\ 0 & \text{kontrako kasuan} \end{cases}$$

$$g : A^* \to \{0,1\}$$

$$g(w) = \begin{cases} 1 & |w| \leq 2 \text{ baldin bada} \\ 0 & \text{kontrako kasuan} \end{cases}$$

f funtzioak b-rik eta c-rik ez duten hitzez osatutako lengoaia infinitua definitzen du:

$$L_f = \{\varepsilon, a, aa, aaa, \ldots\}$$

g funtzioak 2 edo txikiagoa den luzera duten hitzez osatutako lengoaia finitua definitzen du:

$$L_g = \{\varepsilon, a, b, c, aa, ab, ac, ba, bb, bc, ca, cb, cc\}$$

Kasu bietan, funtzioak (f-k edo g-k) hitz bati 1 balioa ematen badio, orduan hitza lengoaian dago eta bestela ez.

Era berean L lengoaia bat multzo eran definituta ematen badigute, dagokion f_L funtzioa kalkula dezakegu. Adibidez, L lengoaia zortzi a dituzten hitzez osatuta baldin badago:

$$L = \{w \mid w \in A^* \wedge \underbrace{|w|_a = 8}_{\text{zortzi } a}\}$$

dagokion f_L funtzioa honako hau izango da:

$$f_L : A^* \to \{0,1\}$$

$$f_L(w) = \begin{cases} 1 & |w|_a = 8 \\ 0 & \text{kontrako kasuan} \end{cases}$$

2^{A^*} multzoa A^*-ko elementuekin osa daitezkeen multzo denez osatutako multzoa da:

$$\{\varnothing, \{\varepsilon\}, \{a\}, \{b\}, \{c\}, \{\varepsilon, a, aa, b\}, \{a, b, cba\}, \{aa, aaa, aaaa, \ldots\}, \ldots\}$$

Beraz, 2^{A^*} multzoa A alfabetoaren gainean **defini daitezkeen lengoaia denez osatutako multzoa** da.

Edozein A alfabeto hartuta ere, 2^{A^*} **beti infinitua** izango da.

2.2.2 Lengoaien gaineko eragiketak

Definizioa kontuan hartuz, lengoaiak sinboloz eratutako kateez osatutako multzoak dira. Beraz, lengoaien gainean definitzen diren eragiketa batzuk multzoen gaineko eragiketak dira eta beste eragiketa batzuk hitzen gaineko eragiketetan oinarrituta daude. Gogoratu multzo batean ez dagoela errepikatutako elementurik, hau da, elementu bakoitza behin bakarrik agertzen da. Adibidez, C multzoa hitzez osatutako multzo bat baldin bada (hau da, lengoaia bat baldin bada) eta 0110 hitza C multzokoa baldin bada, 0110 hitza behin bakarrik agertuko da C multzoan. w hitza L lengoaiakoa dela adierazteko $w \in L$ idatziko dugu eta w hitza L lengoaiakoa ez dela adierazteko $w \notin L$ idatziko dugu.

Bilkura

L_1 eta L_2 A alfabetoaren gainean definitutako bi lengoaia baldin badira, L_1 eta L_2 bilduz lortzen den lengoaia L_1-ekoak edo L_2-koak (gutxienez bietako batekoak) diren hitz denez osatutako lengoaia izango da.

$$L_1 \cup L_2 = \{w \mid w \in A^* \wedge (w \in L_1 \vee w \in L_2)\}$$

Bilkuraren **mota** honako hau da:

$$2^{A^*} \times 2^{A^*} \to 2^{A^*}$$

Motaren bidez A^*-ren bi azpimultzo emanda, A^*-ko beste azpimultzo bat lortuko dela adierazten da. Beste era batera esanda, 2^{A^*} multzoko bi elementu emanda, 2^{A^*} multzoko elementu bat lortuko da, hau da, bi lengoaia emanda, beste lengoaia bat lortuko da.

Bilkuraren **propietateak**:

- Bilkura **elkarkorra** da. L_1, L_2 eta L_3 edozein hiru lengoaia izanda ere, honako hau beteko da

  $$L_1 \cup (L_2 \cup L_3) = (L_1 \cup L_2) \cup L_3$$

  Beraz parentesiak ez dira beharrezkoak: $L_1 \cup L_2 \cup L_3$

- Bilkura **trukakorra** da. L_1 eta L_2 edozein bi lengoaia izanda ere, honako hau beteko da

  $$L_1 \cup L_2 = L_2 \cup L_1$$

- $\cup$ eragilearentzat (hau da, bilkurarentzat) **elementu neutroa** lengoaia hutsa da. Beraz, L edozein lengoaia izanda ere, honako hau beteko da

$$\varnothing \cup L = L \cup \varnothing = L$$

- L edozein lengoaia izanda ere, honako hauek beteko dira: $L \cup L = L$ eta $L \cup A^* = A^*$.

Gogoratu $\varnothing \neq \{\varepsilon\}$ betetzen dela. Horregatik L lengoaia bat hartzen badugu, orokorrean $L \cup \{\varepsilon\} \neq L$ beteko da. $L \cup \{\varepsilon\} = L$ betetzeko, ε hitz hutsak L lengoaiakoa izan beharko du.

Ebaketa

L_1 eta L_2 A alfabetoaren gainean definitutako bi lengoaia baldin badira, L_1 eta L_2 ebakiz lortzen den lengoaia aldi berean L_1-ekoak eta L_2-koak diren hitz denez osatutako lengoaia izango da:

$$L_1 \cap L_2 = \{w \mid w \in A^* \land w \in L_1 \land w \in L_2\}$$

Ebaketaren mota honako hau da:

$$2^{A^*} \times 2^{A^*} \to 2^{A^*}$$

Motaren bidez A^*-ren bi azpimultzo emanda, A^*-ko beste azpimultzo bat lortuko dela adierazten da. Beste era batera esanda, 2^{A^*} multzoko bi elementu emanda, 2^{A^*} multzoko elementu bat lortuko da, hau da, bi lengoaia emanda, beste lengoaia bat lortuko da.

Ebaketaren **propietateak**:

- Ebaketa **elkarkorra** da. L_1, L_2 eta L_3 edozein hiru lengoaia izanda, honako hau beteko da

$$L_1 \cap (L_2 \cap L_3) = (L_1 \cap L_2) \cap L_3$$

  Beraz parentesiak ez dira beharrezkoak: $L_1 \cap L_2 \cap L_3$

- Ebaketa **trukakorra** da. L_1 eta L_2 edozein bi lengoaia izanda, honako hau beteko da

$$L_1 \cap L_2 = L_2 \cap L_1$$

- $\cap$ eragiketarentzat (hau da, ebaketarentzat) **elementu neutroa** A^* multzoa da. Beraz L edozein lengoaia izanda, honako hau beteko da

$$A^* \cap L = L \cap A^* = L$$

- L edozein lengoaia izanda ere, honako hauek beteko dira: $L \cap \varnothing = \varnothing$ eta $L \cap L = L$.

$\{\varepsilon\}$ lengoaiari dagokioenez, L lengoaia bat hartzen badugu eta $\varepsilon \in L$ betetzen bada, orduan $L \cap \{\varepsilon\} = \{\varepsilon\}$ beteko da. Baina $\varepsilon \notin L$ betetzen bada, $L \cap \{\varepsilon\} = \varnothing$ beteko da.

Kenketa

L_1 eta L_2 A alfabetoaren gainean definitutako bi lengoaia baldin badira, L_1 lengoaiari L_2 lengoaia kenduz lortzen den lengoaia L_2-koak ez diren L_1-eko hitzez osatutako lengoaia izango da.

$$L_1 \setminus L_2 = \{w \mid w \in A^* \land w \in L_1 \land w \notin L_2\}$$

Kenketaren mota honako hau da:

$$2^{A^*} \times 2^{A^*} \to 2^{A^*}$$

Motaren bidez A^*-ren bi azpimultzo emanda, A^*-ko beste azpimultzo bat lortuko dela adierazten da. Beste era batera esanda, 2^{A^*} multzoko bi elementu emanda, 2^{A^*} multzoko elementu bat lortuko da, hau da, bi lengoaia emanda, beste lengoaia bat lortuko da.

Kenketaren **propietateak**:

- Kenketa **ez da elkarkorra**. Hona hemen adibidea. $L_1 = \{00, 01, 10, 11\}$, $L_2 = \{00, 000\}$ eta $L_3 = \{00, 101, 0\}$ lengoaiak hartzen baditurgu:

  $$L_1 \setminus (L_2 \setminus L_3) \neq (L_1 \setminus L_2) \setminus L_3$$

  Izan ere $L_1 \setminus (L_2 \setminus L_3) = L_1$ baina $(L_1 \setminus L_2) \setminus L_3 = \{01, 10, 11\}$:

- Kenketa **ez da trukakorra**. Adibidez, $L_1 = \{00, 01, 10, 11\}$ eta $L_2 = \{00, 000\}$ lengoaiak hartuko ditugu.

  $$L_1 \setminus L_2 \neq L_2 \setminus L_1$$

  Izan ere $L_1 \setminus L_2 = \{01, 10, 11\}$ eta $L_2 \setminus L_1 = \{000\}$.

- L edozein lengoaia izanda ere $L \setminus \varnothing = L$, $L \setminus L = \varnothing$ eta $L \setminus A^* = \varnothing$ beteko dira.

Osagarria

L A alfabetoaren gainean definitutako lengoaia baldin bada, L-ren osagarria $\overline{L}$ bezala adierazten da eta L-koak ez diren A^* multzoko hitzez osatutako lengoaia da.

$$\overline{L} = \{w \mid w \in A^* \wedge w \notin L\}$$

Osagarria kenketa erabiliz ere defini daiteke: $\overline{L} = A^* \setminus L$.

Adibidez, $A = \{a, b\}$ alfabetoaren gainean definitutako $L = \{w \mid w \in A^* \wedge |w| \geq 3\}$ lengoaia hartzen badugu, L lengoaia infinitua hiru sinbolo edo sinbolo gehiagoz osatutako A alfabetoaren gaineko hitzez osatuta dago. L-ren osagarria $\overline{L} = \{\varepsilon, a, b, aa, ab, ba, bb\}$ izango litzateke, hau da, L-n ez dauden A^*-ko hitzez osatutako lengoaia. Adibide honetan $\overline{L}$ lengoaia finitua da, baina lengoaia infinitu baten osagarria infinitua izatea ere gerta daiteke. Har dezagun $H = \{w \mid w \in A^* \wedge |w| = |w|_a\}$ lengoaia infinitua. H lengoaian b sinboloaren agerpenik ez duten hitzak daude. Bere osagarria, hau da, $\overline{H}$ lengoaia, infinitua da.

Osagarria kalkulatzen duen funtzioaren mota honako hau da:

$$2^{A^*} \to 2^{A^*}$$

Beraz A^*-ren azpimultzo bat emanda, A^*-ko beste azpimultzo bat lortuko dela adierazten da. Beste era batera esanda, 2^{A^*} multzoko elementu bat emanda, 2^{A^*} multzoko elementu bat lortuko da, hau da, lengoaia bat emanda, beste lengoaia bat lortuko da.

Osagarriaren **propietateak**:

- L edozein lengoaia izanda ere $\overline{\overline{L}} = L$ beteko da: L-ren osagarriaren osagarria L da.
- L_1 eta L_2 edozein bi lengoaia izanda ere $\overline{L_1 \cup L_2} = \overline{L_1} \cap \overline{L_2}$ beteko da.
- L_1 eta L_2 edozein bi lengoaia izanda ere $\overline{L_1 \cap L_2} = \overline{L_1} \cup \overline{L_2}$ beteko da.
- $\overline{\varnothing} = A^*$ eta $\overline{A^*} = \varnothing$.

Kateaketa

L_1 eta L_2 A alfabetoaren gainean definitutako bi lengoaia baldin badira, L_1 eta L_2 lengoaiak kateatuz L_1-eko hitz bakoitza L_2-ko hitz bakoitzarekin kateatuz osatzen diren hitz guztiez osatutako lengoaia lortzen da.

$$L_1L_2 = \{w \mid w \in A^* \wedge \exists u, v(w = uv \wedge u \in L_1 \wedge v \in L_2)\}$$

Adibidez $L_1 = \{aa, cc\}$ eta $L_2 = \{bcb, c\}$ lengoaiak hartzen baditugu, L_1L_2 lengoaia $\{\underline{aa}bcb, \underline{aa}c, \underline{cc}bcb, \underline{cc}c\}$ izango da. Bertan L_1-eko hitzak azpimarratuta ageri dira irakurgarritasuna hobetzeko.

Kateaketaren mota honako hau da:

$$2^{A^*} \times 2^{A^*} \rightarrow 2^{A^*}$$

Motaren bidez A^*-ren bi azpimultzo emanda, A^*-ko beste azpimultzo bat lortuko dela adierazten da. Beste era batera esanda, 2^{A^*} multzoko bi elementu emanda, 2^{A^*} multzoko elementu bat lortuko da, hau da, bi lengoaia emanda, beste lengoaia bat lortuko da.

Kateaketaren **propietateak**:

- Kateaketa **elkarkorra** da. L_1, L_2 eta L_3 edozein hiru lengoaia izanda ere, honako hau beteko da

  $$L_1(L_2L_3) = (L_1L_2)L_3$$

  Beraz parentesiak ez dira beharrezkoak: $L_1L_2L_3$

- Kateaketa **ez da trukakorra**. Esate baterako $L_1 = \{aa, cc\}$ eta $L_2 = \{bcb, b\}$ lengoaiak hartzen baditugu:

  $$L_1L_2 = \{\underline{aa}bcb, \underline{aa}c, \underline{cc}bcb, \underline{cc}c\} \qquad L_2L_1 = \{bcb\underline{aa}, bcb\underline{cc}, b\underline{aa}, b\underline{cc}\}$$

  L_1-eko hitzak azpimarratuta ageri dira irakurgarritasuna hobetzeko.

- Kateaketarentzat **elementu neutroa** hitz hutsaren lengoaia da, hau da, $\{\varepsilon\}$ lengoaia. Beraz, L edozein lengoaia izanda ere, honako hau beteko da:

  $$\{\varepsilon\}L = L\{\varepsilon\} = L$$

- Lengoaia hutsari dagokionez, L edozein lengoaia izanda ere, honako hau beteko da:

  $$L\varnothing = \varnothing L = \varnothing$$

Berreketa

L A alfabetoaren gainean definitutako lengoaia baldin bada eta k osoa eta ez negatiboa ($k \geq 0$) den zenbaki bat baldin bada, L^k lengoaia honela definitzen da:

- $\{\varepsilon\}$, $k = 0$ baldin bada
- LL^{k-1} , $k \geq 1$ baldin bada.

Beraz L^k lengoaia L lengoaia k aldiz kateatuz lortzen da: $\underbrace{L\ldots L}_{k \text{ aldiz}}$. Ondorioz, w hitza L^k lengoaiako edozein hitz izanda ere $w_1, w_2, \ldots, w_k$ k azpihitzetan deskonposa daiteke $w = w_1w_2\ldots w_k$ eta $w_j \in L$ betez, $j \in \{1, \ldots, k\}$ balio denentzat, hau da, $w_1, w_2, \ldots, w_k$ hitz denak L-koak izanda.

k balioa 1 edo handiagoa denean L^k beste era honetara ere defini daiteke:

$$L^k = \{w \mid w \in A^* \wedge \exists u, v(u \in L \wedge v \in L^{k-1} \wedge w = uv)\}$$

Beraz $k \geq 1$ denean, w hitza L^k lengoaiako edozein hitz izanda ere, L-koa den u hitz bat eta L^{k-1}-ekoa den v beste hitz baten kateaketa bezala adierazi ahal izango da: $w = uv$.

Adibidez, $L = \{ab, cb\}$ lengoaia hartzen badugu, L^3 lengoaia LLL lengoaia da, hau da,

$$\{ababab, ababcb, abcbab, abcbcb, cbabab, cbabcb, cbcbab, cbcbcb\}$$

L^3 lengoaia ab eta cb hitzak hirukoteak osatuz konbinatzeko dauden zortzi aukerez osatuta dago.

Berreketaren mota honako hau da:

$$2^{A^*} \times \mathbb{N} \rightarrow 2^{A^*}$$

Hau da, A^*-ren azpimultzo bat eta zenbaki arrunt bat emanda (0 baino handiagoa edo berdina den zenbaki oso bat), A^*-ren beste azpimultzo bat lortuko da. Beste era batera esanda, 2^{A^*} multzoko elementu bat eta zenbaki arrunt bat emanda, hau da, lengoaia bat eta zenbaki arrunt bat emanda, beste lengoaia bat lortuko da.

Berreketaren **propietateak**:

- L edozein lengoaia izanda eta j eta k edozein bi zenbaki arrunt izanda: $L^kL^j = L^{k+j}$.
- L edozein lengoaia izanda eta j eta k edozein bi zenbaki arrunt izanda: $(L^k)^j = L^{k\times j}$.
- $k \geq 0$ izanda, $\{\varepsilon\}^k = \{\varepsilon\}$.
- $k \geq 1$ izanda, $\varnothing^0 = \{\varepsilon\}$ eta $\varnothing^k = \varnothing$.
- L edozein lengoaia izanda, $L^1 = L$ eta $L^0 = \{\varepsilon\}$.

Itxidura

L lengoaia A alfabetoaren gainean definitutako lengoaia baldin bada, L-ren **itxidura** L^* bezala adierazten da eta honela definitzen da:

$$L^* = \bigcup_{k\geq 0} L^k$$

Era ez formalean, definizioa honako hau da:

$$L^* = L^0 \cup L^1 \cup L^2 \cup L^3 \ldots$$

Beraz L^* lengoaia infinitu lengoaia bilduz lortzen da: L-ren berredura posible denak bilduz hain zuzen ere.

L^* definitzeko beste era formal bat honako hau da:

$$L^* = \{w \mid w \in A^* \wedge \exists k(k \geq 0 \wedge w \in L^k)\}$$

L^* lengoaia A alfabetoaren gainean definitutako lengoaia bat da eta ondorioz, A^*-ren azpimultzoa da: $L^* \subseteq A^*$.

Adibidez, $A = \{a, b\}$ alfabetoaren gainean definitutako $L = \{aa\}$ lengoaia hartzen badugu, L^* lengoaia a-kopuru bikoitia eta b-rik ez duten hitzez osatutako lengoaia izango da:

$$L^* = \{\varepsilon, aa, aaaa, aaaaaa, \ldots\} = \{w \mid w \in A^* \wedge \underbrace{|w| = |w|_a}_{a \text{ bakarrik}} \wedge \underbrace{|w| \bmod 2 = 0}_{a\text{-kopurua bikoitia}}\}$$

Adibidez aba hitza ez da L^*-koa baina aba hitza A^*-koa da. L^* beti A^*-ren azpimultzoa izango da.

Itxiduraren mota honako hau da:

$$2^{A^*} \to 2^{A^*}$$

Beraz A^*-ren azpimultzo bat emanda, A^*-ren beste azpimultzo bat lortuko da. Beste era batera esanda, 2^{A^*} multzoko elementu bat emanda, 2^{A^*} multzoko beste elementu bat lortuko da, hau da, lengoaia bat emanda beste lengoaia bat lortuko da.

Itxiduraren **propietateak**:

- $(L^*)^* = L^*$
- $\{\varepsilon\}^* = \{\varepsilon\}$
- $\varnothing^* = \{\varepsilon\}$

Itxidura positiboa

L A alfabetoaren gainean definitutako lengoaia bat baldin bada, L-ren **itxidura positiboa** L^+ bezala adierazten da eta honela definitzen da:

$$L^+ = \bigcup_{k \geq 1} L^k$$

Era informalean definizioa honako hau da:

$$L^+ = L^1 \cup L^2 \cup L^3 \ldots$$

Itxidura positiboa eta itxidura mota bereko eragiketak dira:

$$2^{A^*} \to 2^{A^*}$$

Lengoaia bat emanda beste lengoaia bat lortuko da.

L^+ definitzerakoan L^0 ez da kontuan hartzen, hau da, $\{\varepsilon\}$ multzoa ez da biltzen. Hala ere ε hitz hutsa L^+ lengoaian ager daiteke. Hori horrela gertatuko da ε hitz hutsa L lengoaiakoa baldin bada. Ondorioz, $L^* = L^+ \cup \{\varepsilon\}$ beti beteko da baina kasu batzuetan $L^* \setminus \{\varepsilon\}$ eta L^+ ez dira berdinak izango, izan ere ε hitza L^+-koa baldin bada, orduan $L^* \setminus \{\varepsilon\} \neq L^+$ beteko baita.

Edozein L lengoaia hartuta, $L^+ \subseteq L^*$ beti beteko da eta ondorioz L^+ ere A alfabetoaren gainean definitutako beste lengoaia bat izango da eta L^+ multzoa A^*-ren azpimultzoa izango da.

Adibidez, $A = \{a, b\}$ alfabetoaren gainean definitutako $L = \{aa\}$ lengoaia hartzen badugu, L^+ lengoaia b-rik ez eta a-kopuru bikotia duten hitz ez hutsez osatuta egongo da:

$$L^+ = \{aa, aaaa, aaaaaa, \ldots\} = \{w \mid w \in A^* \wedge \underbrace{|w|_a \geq 1}_{\text{gutxienez } a \text{ bat}} \wedge \underbrace{|w| = |w|_a}_{a\text{-k bakarrik}} \wedge \underbrace{|w| \bmod 2 = 0}_{a\text{-kopurua bikoitia}}\}$$

Adibide horretan ε hitza ez da L^+ lengoaiakoa eta ondorioz $L^+ \neq L^*$.

Lengoaia hutsaren kasuan $\varnothing^+ = \varnothing$ betetzen da.

Alderantzizkoa

L A alfabetoaren gainean definitutako lengoaia baldin bada, L-ren alderantzizkoa L^R bezala adierazi-ko da eta L-ko hitzen alderantzizkoez osatuta egongo da. L^R ere A^*-ren azpimultzoa izango da.

$$L^R = \{w \mid w \in A^* \wedge w^R \in L\}$$

Adibidez, $A = \{a, b, c\}$ alfabetoaren gainean definitutako $L = \{a, ab, aab, ac\}$ lengoaia hartzen badugu, $L^R = \{a, ba, baa, ca\}$ izango da.

Alderantzizkoa kalkulatzen duen funtzioaren mota honako hau da:

$$2^{A^*} \to 2^{A^*}$$

Beraz A^*-ren azpimultzo bat emanda, A^*-ren beste azpimultzo bat lortuko da. Beste era batera esanda, 2^{A^*} multzoko elementu bat emanda, 2^{A^*} multzoko beste elementu bat lortuko da, hau da, lengoaia bat emanda beste lengoaia bat lortuko da.

Alderanzketaren **propietateak**:

- L edozein lengoaia izanda, $(L^R)^R = L$ beteko da.
- L_1 eta L_2 edozein bi lengoaia izanda, $(L_1 \cup L_2)^R = L_1^R \cup L_2^R$ beteko da.
- L_1 eta L_2 edozein bi lengoaia izanda, $(L_1 \cap L_2)^R = L_1^R \cap L_2^R$ beteko da.
- L_1 eta L_2 edozein bi lengoaia izanda, $(L_1 L_2)^R = L_2^R L_1^R$ beteko da.
- $\varnothing^R = \varnothing$ eta $\{\varepsilon\}^R = \{\varepsilon\}$.

2.3 Multzo zenbagarriak eta zenbaezinak: A^* eta 2^{A^*}-ren kasua

2.3.1 Funtzio bijektiboak

Hasteko funtzio bat **bijektiboa** izateak zer esan nahi duen gogoratuko dugu. Demagun Q eta S bi multzo direla. $Q \to S$ erako f funtzio bat bijektiboa dela esaten da honako baldintza hauek betetzen baditu:

(a) k elementua Q multzoko balio bat baldin bada, $f(k) = i$ betetzen duen i elementuren bat existitzen da S multzoan.

(b) i elementua S multzoko balio bat baldin bada, $f(k) = i$ betetzen duen k elementuren bat existitzen da Q multzoan.

(c) Q multzoko k_1 eta k_2 elementuak desberdinak badira, $f(k_1)$ eta $f(k_2)$ desberdinak izango dira.

Hiru baldintza horien ondorio bezala, S multzoko i_1 eta i_2 elementuak desberdinak badira, $f(k_1) = i_1$ eta $f(k_2) = i_2$ betetzen duten k_1 eta k_2 bi balio desberdin existituko dira Q multzoan.

Era laburrean esanda, $Q \to S$ erako f funtzio bijektiboren bat existitzen bada, orduan Q eta S multzoetan elementu-kopuru bera daukagu eta Q eta S multzoetako elementuen artean bat-bat erako erlazioa daukagu.

2.3.2 Multzo bat zenbagarria izateko bete beharreko baldintza

Multzo bateko elementu denak zerrenda batean ipini eta zerrendako ezkerreko ertzetik abiatuz, multzokoa den edozein elementutara urrats-kopuru finituan iristea baldin badaukagu, orduan multzoa **zenbagarria** da. Irizpide honek multzo finituentzat eta infinituentzat balio du. Hala ere multzo infinituen kasuan irizpide hori era formalagoan eman daiteke: infinitua den C multzo bat zenbagarria dela esaten da bijektiboa den $\mathbb{N} \rightarrow C$ erako funtzio bat defini baldin badaiteke, hau da, $\mathbb{N}$-ko elementu desberdin bakoitzari C-ko elementu desberdin bat eta C-ko elementu desberdin bakoitzari $\mathbb{N}$-ko elementu desberdin bat egokitzen dion funtzioa defini baldin badaiteke.

Multzo finituak zenbagarriak dira. Izan ere, multzo finitu bateko elementuak zerrenda batean ipini ondoren eta ezkerreko ertzetik abiatuz, elementu denetara urrats-kopuru finituan irits gaitezke. Adibidez, $A = \{a, b, c\}$ alfabetoa eta 2 luzera duten A^*-ko hitzez osatutako D multzo finitua hartzen baditugu, esate baterako honako zerrenda hau osa dezakegu:

$$[\, aa, ab, ac, ba, bb, bc, ca, cb, cc \,]$$

eta ezkerreko ertzetik abiatuz, zerrenda horretako edozein elementutara urrats-kopuru finituan irits gaitezke. Demagun aa hitzera iristeko zero urrats behar direla (ezkerreko ertzean dagoelako). Ondorioz, bc hitzera iristeko bost urrats beharko genituzke, ac hitzera iristeko bi urrats beharko genituzke eta abar.

D multzoko elementuak zerrendan ipintzeko beste aukera batzuk ere badaude. Horretarako elementuen ordena aldatzearekin nahikoa da:

$$[\, bb, bc, aa, ba, cc, ab, ca, cb, ac \,]$$

edo

$$[\, ca, bc, bb, ac, cb, ba, ab, cc, aa \,]$$

eta abar.

Multzo infinituen kasuan, batzuk zenbagarriak dira eta beste batzuk ez.

- Zenbaki arrunten $\mathbb{N}$ multzoa (negatiboak ez diren zenbaki osoez eratutako multzoa) zenbagarria da. Bere elementuak zerrendan ipintzeko aukera bat honako hau da:

  $$[\, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, \ldots \,]$$

  Ezkerretik abiatuta 0 elementura iristeko zero urrats behar direla kontsideratuz, 8 elementura iristeko zortzi urrats beharko ditugu, 15 elementura iristeko hamabost urrats beharko ditugu eta 46 zenbakira iristeko berrogeita sei urrats beharko ditugu. Guztira, edozein n zenbaki emanda, n urratsetan irits gaitezke zenbaki horretara.

  Zerrenda horri dagokion funtzio bijektiboa honako hau da:

  $$\begin{array}{l} f : \mathbb{N} \rightarrow \mathbb{N} \\ f(n) = n \end{array}$$

  Funtzio bijektibo horrek honako hau adierazten du:

  $$\begin{array}{l} f(0) = 0 \\ f(1) = 1 \\ f(2) = 2 \\ f(3) = 3 \\ f(4) = 4 \\ \vdots \end{array}$$

$\mathbb{N}$ multzoko elementuak zerrendan ipintzeko beste aukera bat honako hau da:

$$[\,5, 0, 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, \ldots\,]$$

Ezkerretik abiatuta 5 elementura iristeko zero urrats behar direla kontsideratuz, 3 elementura iristeko lau urrats beharko ditugu, 10 elementura iristeko hamar urrats beharko ditugu eta 46 zenbakira iristeko berrogeita sei urrats beharko ditugu.

Zerrenda horri dagokion funtzio bijektiboa honako hau da:

$$f : \mathbb{N} \to \mathbb{N}$$
$$f(n) = \begin{cases} 5 & n = 0 \text{ baldin bada} \\ n-1 & 1 \leq n \leq 5 \text{ baldin bada} \\ n & n \geq 6 \text{ baldin bada} \end{cases}$$

Funtzio bijektibo horrek honako hau adierazten du:

$$\begin{array}{l} f(0) = 5 \\ f(1) = 0 \\ f(2) = 1 \\ f(3) = 2 \\ f(4) = 3 \\ f(5) = 4 \\ f(6) = 6 \\ f(7) = 7 \\ \vdots \end{array}$$

Hala ere, zerrenda osatzerakoan, $\mathbb{N}$ multzoko elementuak edozein ordenetan ematerik ez dago. Har dezagun, esate baterako, hasteko zenbaki bikoiti denak eta gero zenbaki bakoiti denak dituen zerrenda:

$$[\,0, 2, 4, 6, 8, 10, 12, 14, \ldots, 1, 3, 5, 7, 9, 11, 13, 15, \ldots\,]$$

Ezkerretik abiatuta 0 elementura iristeko zero urrats behar direla kontsideratuz, adibidez, 6 elementura iristeko hiru urrats beharko ditugu eta 12 elementura iristeko sei urrats beharko ditugu. Baina zer gertatzen da zenbaki bakoitiekin? Zenbat urrats behar dira 1 zenbakira iristeko? 1 zenbakira iristerik ba al dago? Zenbaki bakoitietara iristerik ba al dago zerrenda horretan? Erantzuna ezezkoa da. 0 zenbakitik abiatu eta zerrenda eskuinerantz zeharkatuz joaten bagara, inoiz ez gara iritsiko zenbaki bakoitietara, beraien aurretik infinitu zenbaki bikoiti baitaude.

Zenbaki arrunten adibide honetan garbi ikusten da multzo infinitu bat zenbagarria dela frogatzeko, elementuen ordena garrantzitsua dela zerrenda osatzerakoan. Edozein ordena ez da egokia, orden egoki bat aurkitu behar da.

- $\mathcal{Z}$ zenbaki osoen multzoa zenbagarria da. Bere elementuak zerrendan ipintzeko aukera bat honako hau da:

$$[\,0, 1, -1, 2, -2, 3, -3, 4, -4, 5, -5, 6, -6, 7, -7, 8, -8, 9, -9, 10, -10, 11, -11, 12, -12, \ldots\,]$$

Ezkerretik abiatuta 0 elementura iristeko zero urrats behar direla kontsideratuz, 3 elementura iristeko bost urrats beharko ditugu, -3 elementura iristeko sei urrats beharko ditugu eta -7 zenbakira iristeko hamalau urrats beharko ditugu. Guztira, edozein n zenbaki emanda, zenbaki horretara urrats-kopuru finituan irits gaitezke.

Zerrenda horri dagokion funtzio bijektiboa honako hau da:

$$f : I\!N \to \mathcal{Z}$$
$$f(n) = \begin{cases} -(n \;\; div \;\; 2) & n \text{ bikoitia baldin bada} \\ (n \;\; div \;\; 2) + 1 & n \text{ bakoitia baldin bada} \end{cases}$$

Funtzio bijektibo horrek honako hau adierazten du:

$$\begin{aligned} &f(0) = 0 \\ &f(1) = 1 \\ &f(2) = -1 \\ &f(3) = 2 \\ &f(4) = -2 \\ &\vdots \end{aligned}$$

Zerrenda osatzerakoan, zenbaki osoen multzoko elementuak edozein ordenetan ematerik ez dago. Har dezagun, esate baterako, hasteko zenbaki ez negatibo denak (zero eta positiboak) eta gero zenbaki negatibo denak dituen zerrenda:

$$[\,0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots, -1, -2, -3, -4, -5, -6, -7, -8, -9, -10, -11, \ldots\,]$$

Ezkerretik abiatuta 0 elementura iristeko zero urrats behar direla kontsideratuz, adibidez, 3 elementura iristeko hiru urrats beharko ditugu eta 6 elementura iristeko sei urrats beharko ditugu. Baina zer gertatzen da zenbaki negatiboekin? Zenbat urrats behar dira -1 zenbakira iristeko? -1 zenbakira iristerik ba al dago? Zenbaki negatiboetara iristerik ba al dago zerrenda horretan? Erantzuna ezezkoa da. 0 zenbakitik abiatu eta zerrenda eskuinerantz zeharkatuz joaten bagara, inoiz ez gara iritsiko zenbaki negatiboetara, beraien aurretik infinitu zenbaki positibo baitaude.

Zenbaki osoen adibide honetan ere garbi ikusten da multzo infinitu bat zenbagarria dela frogatzeko, elementuen ordena garrantzitsua dela zerrenda osatzerakoan. Edozein ordena ez da egokia, orden egoki bat aurkitu behar da.

Zenbaki arrunten multzoa eta zenbaki osoen multzoa zenbagarriak izateak, multzo biek neurri bera dutela (elementu kopuru bera dutela) adierazten du, nahiz eta ez eman horrela izan daitekeenik.

- Zenbaki arruntez eratutako bikoteez osatutako $I\!N \times I\!N$ multzoa zenbagarria da. Bere elementuak zerrendan ipintzeko aukera bat honako hau da:

$$[\;\underbrace{(0,0)}_{\text{batura = 0}}\;,\underbrace{(0,1),(1,0)}_{\text{batura = 1}},\underbrace{(0,2),(1,1),(2,0)}_{\text{batura = 2}},\underbrace{(0,3),(1,2),(2,1),(3,0)}_{\text{batura = 3}},\ldots\,]$$

Ezkerretik abiatuta $(0, 0)$ elementura iristeko zero urrats behar direla kontsideratuz, $(2, 0)$ elementura iristeko bost urrats beharko ditugu, $(0, 3)$ elementura iristeko sei urrats beharko ditugu eta $(3, 0)$ zenbakira iristeko bederatzi urrats beharko ditugu. Guztira, edozein (m, n) bikote emanda, bikote horretara urrats-kopuru finituan irits gaitezke.

Zerrenda horri dagokion $f : I\!N \to I\!N \times I\!N$ erako funtzio bijektiboa ez dugu emango, baina funtzio bijektibo horrek honako hau adieraziko luke:

$$
\begin{aligned}
f(0) &= (0,0)\\
f(1) &= (0,1)\\
f(2) &= (1,0)\\
f(3) &= (0,2)\\
f(4) &= (1,1)\\
&\vdots
\end{aligned}
$$

Zerrenda osatzerakoan, $\mathbb{N} \times \mathbb{N}$ multzoko elementuak edozein ordenetan ematerik ez dago. Har dezagun, esate baterako, hasteko lehenengo osagai bezala 0 balioa duten bikote denak, gero lehenengo osagai bezala 1 balioa duten bikote denak eta abar dituen zerrenda:

$$[\,(0,0),(0,1),(0,2),(0,3),\ldots,(1,0),(1,1),(1,2),(1,3),\ldots,(2,0),(2,1),(2,2),(2,3),\ldots]$$

Ezkerretik abiatuta $(0,0)$ bikotera iristeko zero urrats behar direla kontsideratuz, adibidez, $(0,3)$ elementura iristeko hiru urrats beharko ditugu eta $(0,20)$ elementura iristeko hogei urrats beharko ditugu. Baina zer gertatzen da lehenengo osagai bezala 0 zenbakia ez duten bikoteekin? Zenbat urrats behar dira $(1,0)$ bikotera iristeko? $(1,0)$ bikotera iristerik ba al dago? Erantzuna ezezkoa da. $(0,0)$ bikotetik abiatu eta zerrenda eskuinerantz zeharkatuz joaten bagara, inoiz ez gara iritsiko $(1,0)$ bikotera, bere aurretik lehenengo osagai bezala 0 balioa duten infinitu bikote baitaude.

$\mathbb{N} \times \mathbb{N}$ multzoaren adibide honetan garbi ikusten da, berriz ere, multzo infinitu bat zenbagarria dela frogatzeko, elementuen ordena garrantzitsua dela zerrenda osatzerakoan. Edozein ordena ez da egokia, orden egoki bat aurkitu behar da.

Zenbaki arrunten multzoa, zenbaki osoen multzoa eta $\mathbb{N} \times \mathbb{N}$ multzoa zenbagarriak izateak, hiru multzo horiek neurri bera dutela (elementu kopuru bera dutela) adierazten du, nahiz eta ez eman horrela izan daitekeenik.

- Azkenik, zenbaki errealez osatutako $\mathbb{R}$ multzoa hartzen badugu, multzoa zenbaezina dela (zenbagarria ez dela) frogatuta dago matematikoki (2.3.5 atalean emango da froga matematiko hori). $\mathbb{R}$-ko elementuak zerrenda batean ipini eta gero, $\mathbb{R}$-ko elementu guztietara urrats-kopuru finituan iristeko erarik ez dago. Ondorioz, $\mathbb{N}$ eta $\mathbb{R}$ multzoak infinituak izan arren, $\mathbb{R}$ multzoan $\mathbb{N}$ multzoan baino elementu gehiago dago. Beraz, infinitutasun desberdinak daude.

2.3.3 A^* zenbagarria da

Adibidez, $A = \{0,1\}$ alfabetoa hartzen badugu, A^* multzo infinitua zenbagarria da. A^* multzoko elementuak zerrendan ipintzeko era egoki bat honako hau da:

$$[\,\varepsilon,0,1,00,01,10,11,000,001,010,011,100,101,110,111,0000,\ldots]$$

Zerrenda horretan elementuak, hasteko, luzeraren arabera ordenatuta daude eta luzera bera dutenak orden alfabetikoa jarraituz ordenatuta daude. $\mathbb{N}$-ko zenbaki bakoitzari A^*-ko elementu bat egokitzen dion $f : \mathbb{N} \to A^*$ funtzio bijektiboa honela defini dezakegu:

$$
\begin{aligned}
f(0) &= \varepsilon\\
f(1) &= 0\\
f(2) &= 1\\
f(3) &= 00\\
f(4) &= 01\\
&\vdots
\end{aligned}
$$

f funtzioak n zenbaki bat emanda, n urratsetan zein hitzetara iritsiko garen adierazten digu. Kontua elementuen artean ordena ezartzean datza. Adibide honetan, w hitz batek 0-rik ez badu eta bere luzera n baldin bada ($|w| = n$), bere hurrengo hitza zeroz bakarrik osatutako $n + 1$ luzerako hitza izango da (adibidez 11-en hurrengoa 000 izango da). w hitz batek 0-ren bat baldin badu eta bere luzera n baldin bada ($|w| = n$), bere hurrengo hitzaren luzera ere n izango da eta orden alfabetikoa jarraituz aukeratuko da (adibidez 100-en hurrengoa 101 izango da). Ideia hau edozein alfabetotara egoki daiteke. Beraz, A edozein alfabeto izanda ere, A^* multzo infinitua **zenbagarria** da.

Har dezagun orain A^*-ko elementuak zerrendan ipintzeko beste aukera hau:

$$[\,\varepsilon, 0, 00, 000, 0000, 00000, \ldots\,]$$

Zerrenda horretan, hasteko, batekorik ez duten hitz denak agertzen dira. Zerrenda honek ez du balio A^* zenbagarria dela frogatzeko. Izan ere, batekoren bat duten hitzetara iristerik ez dago, hitz horiek iristezinak dira. Adibide honekin, A^* zenbagarria dela frogatzeko, elementuak zerrendan ipintzerakoan edozein ordenak ez duela balio ikus dezakegu. Orden egokia aurkitu behar da.

Zenbagarriak diren multzo infinituek $I\!N$ multzoak adina elementu dituzte.

2.3.4 2^{A^*} zenbaezina da

2^{A^*} multzo infinitua **ez da zenbagarria**, $I\!N \to 2^{A^*}$ erako funtzio bijektiborik ezin baita definitu.

Era horretako funtziorik ezin dela definitu frogatzeko **kontraesanaren teknika** erabiliko dugu. Teknika hori erabiltzeko, 2^{A^*} zenbagarria dela suposatuko dugu eta suposizio horren ondorio bezala kontraesan bat sortuko dugu. Horrela, 2^{A^*} zenbagarria dela suposatzeak ezinezkoa den egoera batera eramango gaituenez, 2^{A^*} zenbagarria ez dela ondorioztatuko dugu.

Demagun 2^{A^*} zenbagarria dela. 2^{A^*} zenbagarria baldin bada, $I\!N \to 2^{A^*}$ erako g funtzio bijektibo bat existituko da. Beraz, g funtzioaren bidez $I\!N$-ko zenbaki desberdin bakoitzari 2^{A^*} multzoko lengoaia desberdin bat dagokio eta 2^{A^*} multzoko lengoaia desberdin bakoitzari $I\!N$-ko zenbaki desberdin bat dagokio. Zenbaki arrunt eta lengoaien artean bat-bat erako erlazioa edukiko dugu g-ren bidez. Horrela, $g(0)$ lengoaia bat da, $g(1)$ beste lengoaia bat da, $g(2)$ beste lengoaia bat da, $g(3)$ beste lengoaia bat da, eta abar. Badakigu 2^{A^*} multzoko lengoaiak A^*-ren azpimultzoak direla, beraz $g(0)$ A^*-ren azpimultzoa da, $g(1)$ A^*-ren azpimultzoa da, $g(2)$ A^*-ren azpimultzoa da, eta abar. Horren arabera, 2^{A^*}-ko lengoaiak zerrenda batean ipini daitezke:

$$[\,g(0), g(1), g(2), g(3), \ldots, g(j), \ldots\,]$$

Bestalde, badakigu A^* multzoa zenbagarria dela eta $I\!N \to A^*$ erako f funtzio bijektibo bat existitzen dela. Ondorioz, f funtzioak $I\!N$-ko zenbaki desberdin bakoitzari A^* multzoko hitz desberdin bat egokitzen dio eta A^* multzoko hitz desberdin bakoitzari $I\!N$-ko zenbaki desberdin bat egokitzen dio. Beraz, $f(0)$ A^*-ko hitz bat da, $f(1)$ A^*-ko beste hitz bat da, $f(2)$ A^*-ko beste hitz bat da, $f(3)$ A^*-ko beste hitz bat da eta abar. Ondorioz A^*-ko hitzak zerrenda batean ipini daitezke:

$$[\,f(0), f(1), f(2), f(3), \ldots, f(j), \ldots\,]$$

Laburtuz, alde batetik $g(0), g(1), g(2), g(3), \ldots$ lengoaiak dira (2^{A^*} multzoko elementuak) eta beste aldetik $f(0), f(1), f(2), f(3), \ldots$ hitzak dira (A^* multzoko elementuak).

Orain g eta f funtzioak erabiliz "ezinezkoa" den C lengoaia bat (2^{A^*} multzoko elementu bat) eraikiko dugu.

C lengoaia honela eraikiko dugu: $I\!N$ multzoko k zenbaki bakoitzeko, $f(k)$ hitza $g(k)$ lengoaiakoa ez bada, orduan $f(k)$ hitza C multzoan sartuko dugu. Aldiz, $f(k)$ hitza $g(k)$ lengoaiakoa baldin bada, orduan $f(k)$ ez dugu C multzoan sartuko. Beraz, $f(0)$ hitza $g(0)$ lengoaian ez badago, orduan $f(0)$ hitza C lengoaian sartuko dugu eta $f(0)$ hitza $g(0)$ lengoaian baldin badago, orduan $f(0)$ hitza ez dugu C lengoaian sartuko. Era berean, $f(1)$ hitza $g(1)$ lengoaian ez badago, orduan $f(1)$ hitza C lengoaian sartuko dugu eta $f(1)$ hitza $g(1)$ lengoaian baldin badago, orduan $f(1)$ hitza ez dugu C lengoaian sartuko. Era berean, $f(2)$ hitza $g(2)$ lengoaian ez badago, orduan $f(2)$ hitza C lengoaian sartuko dugu eta $f(2)$ hitza $g(2)$ lengoaian baldin badago, orduan $f(2)$ hitza ez dugu C lengoaian sartuko. Eta horrela zenbaki arrunt denekin. Era horretara eraikitako C multzoa hitzez osatutako multzoa da eta ondorioz lengoaia bat da, hau da, 2^{A^*} multzoko lengoaia edo elementu bat da. Hori horrela izanda, g funtzioak 2^{A^*}-ko lengoaia desberdin bakoitzari $I\!N$-ko zenbaki bat egokitzen dionez, C lengoaiari ere $I\!N$-ko zenbaki bat egokituko dio. Demagun C-ri dagokion zenbakia j dela. Horrek esan nahi du $g(j) = C$ beteko dela. Arazoa $f(j)$ hitza hartu eta $f(j)$ hitza $g(j)$ lengoaiakoa al den erabakitzerakoan sortzen da, hau da, arazoa $f(j)$ hitza C lengoaiakoa al den erabakitzerakoan sortzen da. C-ren definiziora itzultzen bagara, $f(j)$ hitza $g(j)$ lengoaiakoa baldin bada, orduan $f(j)$ hitza ez da C multzokoa izango eta $f(j)$ hitza $g(j)$ lengoaiakoa ez bada, orduan $f(j)$ hitza C lengoaiakoa izan go da. C lengoaia $g(j)$ lengoaia denez, $f(j)$ hitza C lengoaiakoa bada, orduan $f(j)$ hitza ez da C lengoaiakoa eta $f(j)$ hitza C lengoaiakoa ez bada, orduan $f(j)$ hitza C lengoaiakoa da. Baina hori kontraesana da, hitz batek ezin baitu izan multzo batekoa eta aldi berean multzo horretakoa ez izan.

Beraz g funtzio bijektiboa existitzen dela suposatzeak kontraesanera eraman gaituenez, g funtzio hori ez dela existitzen ondoriozta dezakegu eta horrek 2^{A^*} zenbaezina dela esan nahi du.

2.3.5 $I\!R$ zenbaezina da

2^{A^*} multzo infinitua zenbaezina dela frogatzeko erabili den teknika hobeto ulertzeko, $I\!R$ zenbaezina dela frogatuko da teknika bera jarraituz.

$I\!R$ zenbaezina dela frogatzeko, $I\!R$-ren azpimultzoa den $[0..1)$ multzo infinitua[4] zenbaezina dela frogatuko dugu, $[0..1)$ zenbaezina bada, $I\!R$ ere zenbaezina izango baita. Horretarako, $I\!N \to [0..1)$ erako funtzio bijektiborik ezin dela definitu erakutsiko dugu.

Kontraesanaren teknika erabiliz, era horretako funtziorik ezin dela definitu frogatzeko, $[0..1)$ zenbagarria dela suposatuko dugu eta suposizio horren ondorio bezala kontraesan bat sortuko dugu. Horrela, $[0..1)$ zenbagarria dela suposatzeak ezinezkoa den egoera batera eramango gaituenez, $[0..1)$ zenbaezina dela ondorioztatuko dugu.

Demagun $[0..1)$ zenbagarria dela. $[0..1)$ zenbagarria baldin bada, $I\!N \to [0..1)$ erako h funtzio bijektibo bat existituko da. Beraz, h funtzioaren bidez $I\!N$-ko zenbaki desberdin bakoitzari $[0..1)$ multzoko zenbaki desberdin bat dagokio eta $[0..1)$ multzoko zenbaki desberdin bakoitzari $I\!N$-ko zenbaki desberdin bat dagokio. Zenbaki arrunt eta $[0..1)$ multzoko zenbakien artean bat-bat erako erlazioa edukiko dugu h-ren bidez. Horrela, $h(0)$ elementua $[0..1)$ multzoko zenbaki bat da, $h(1)$ elementua $[0..1)$ multzoko beste zenbaki bat da, $h(2)$ elementua $[0..1)$ multzoko beste zenbaki bat da, $h(3)$ balioa $[0..1)$ multzoko beste zenbaki bat da, eta abar. Horren arabera, $[0..1)$ multzoko zenbakiak zerrenda batean ipini daitezke:

[4] $[0..1)$ multzoan zero bera eta "zero koma zerbait" erako zenbaki erreal denak daude, baina 1 zenbakia ez dago.

$$[\,h(0), h(1), h(2), h(3), \ldots, h(j), \ldots\,]$$

Bestalde, badakigu $[0..1)$ multzoko zenbaki bakoitza "zero koma eta infinitu digitu" erakoa dela. Esate baterako zero zenbakia $0,00000\ldots$ da, infinitu digiturekin (infinitu 0rekin). Beraz h funtzio bijektiboa erabiliz eraiki den $[\,h(0), h(1), h(2), h(3), \ldots, h(j), \ldots\,]$ zerrendako elementuek honako egitura hau izango dute:

$$h(0) = 0, d_0^0\, d_0^1\, d_0^2\, \ldots\, d_0^j\, \ldots$$

$$h(1) = 0, d_1^0\, d_1^1\, d_1^2\, \ldots\, d_1^j\, \ldots$$

$$h(2) = 0, d_2^0\, d_2^1\, d_2^2\, \ldots\, d_2^j\, \ldots$$

$$\vdots$$

$$h(j) = 0, d_j^0\, d_j^1\, d_j^2\, \ldots\, d_j^j\, \ldots$$

$$\vdots$$

Hor d_m^k erako elementu bakoitza $\{0, 1, 2, 3, \ldots, 9\}$ multzoko balio bat izango da.

Orain h funtzioa erabiliz "ezinezkoa" den $[0..1)$ multzoko q zenbaki bat eraikiko dugu. Eraikiko dugun q zenbaki hori $[0..1)$ multzokoa denez, honako egitura hau izango du:

$$q = 0, q_0\, q_1\, q_2\, \ldots\, q_j\, \ldots$$

hor q_i osagai bakoitza $\{0, 1, 2, 3, \ldots, 9\}$ multzokoa izango da.

q zenbakia honela eraikiko dugu: $I\!N$ multzoko ℓ zenbaki bakoitzeko, $h(\ell)$ zenbakiko d_ℓ^ℓ digitua zero bada, orduan q zenbakiko ℓ-garren digitua, hau da, q_ℓ digitua, 1 izango da eta $h(\ell)$ zenbakiko d_ℓ^ℓ digitua zero ez bada, orduan q zenbakiko ℓ-garren digitua, hau da, q_ℓ digitua, 0 izango da. Beraz, $h(0)$ zenbakiko d_0^0 digitua zero bada, orduan q_0 digitua 1 izango da eta $h(0)$ zenbakiko d_0^0 digitua zero ez bada, orduan q_0 digitua 0 izango da. Era berean, $h(1)$ zenbakiko d_1^1 digitua zero bada, orduan q_1 digitua 1 izango da eta $h(1)$ zenbakiko d_1^1 digitua zero ez bada, orduan q_1 digitua 0 izango da. Irizpide bera jarraituz, $h(2)$ zenbakiko d_2^2 digitua zero bada, orduan q_2 digitua 1 izango da eta $h(2)$ zenbakiko d_2^2 digitua zero ez bada, orduan q_2 digitua 0 izango da. Eta horrela zenbaki arrunt denekin. Era horretara eraikitako q zenbakia $[0..1)$ multzokoa izango da. Hori horrela izanda, h funtzioak $[0..1)$ multzoko zenbaki bakoitzari $I\!N$-ko zenbaki bat egokitzen dionez, q zenbakiari ere $I\!N$-ko zenbaki bat egokituko dio. Demagun q-ri dagokion zenbakia j dela. Horrek esan nahi du $h(j) = q$ beteko dela. Definitzeko eragatik, q zenbakiko digitu denak 0 edo 1 dira. Arazoa q zenbakia hartu eta q_j digitua 0 ala 1 al den erabakitzerakoan sortzen da. q-ren definiziora itzultzen bagara, $h(j)$ zenbakiko d_j^j digitua zero bada, orduan q_j digitua 1 izango da eta $h(j)$ zenbakiko d_j^j digitua zero ez bada, orduan q_j digitua 0 izango da. Baina $h(j) = q$ denez, hasteko $d_j^0 = q_0, d_j^1 = q_1, d_j^2 = q_2, \ldots, d_j^j = q_j, \ldots$ beteko da. Eta gainera, $h(j)$ zenbakiko d_j^j digitua zero bada, orduan q zenbakiko q_j digitua (hau da, d_j^j digitua) 1 izango da eta $h(j)$ zenbakiko d_j^j digitua zero ez bada, orduan q zenbakiko q_j digitua (hau da, d_j^j digitua) 0 izango da. Baina hori kontraesana da, digitu batek, eta zehazki d_j^j digituak, ezin baitu izan aldi berean zero eta 1 eta ezin baitu izan aldi berean zeroren desberdina eta zero.

Beraz h funtzio bijektiboa existitzen dela suposatzeak kontraesanera eraman gaituenez, h funtzio hori ez dela existitzen ondoriozta dezakegu eta horrek $[0..1)$ zenbaezina dela esan nahi du. Bukatzeko, $[0..1)$ zenbaezina bada, $\mathbb{R}$ ere zenbaezina ezina da.

2.3.6 Zenbagarritasunaren eta zenbaezintasunaren esanahia konputagarritasunari dagokionez

A^* eta 2^{A^*} infinituak dira baina lehenengoa zenbagarria den bitartean, bestea zenbaezina da. Beraz A^*-ren infinitutasuna eta 2^{A^*}-ren infinitutasuna desberdinak dira. Konputazioaren ikuspuntutik multzo zenbagarriak zenbaezinak baino maneiagarriagoak dira.

Multzo bat zenbagarria baldin bada, bertako elementu denez osatutako zerrenda bat eratuz joateko gai den algoritmo bat existituko da, multzo horretako edozein elementu aukeratuta ere, algoritmoak elementu hori urrats kopuru finitu batean sortuko duelarik.

Multzo bat zenbaezina baldin bada, bertako elementu denez osatutako zerrenda bat eratuz joateko gai den algoritmorik ez da existituko. Multzo horretako elementuez osatutako zerrenda bat eratzen saiatzen den edozein algoritmo hartuta ere, beti existituko da algoritmoak inoiz sortu ezingo duen elementuren bat.

2.4 Ariketak: lengoaien definizio formalen ulermena

$A = \{a, b, c\}$ alfabetoaren gainean definitutako honako lengoaia hauetakoak diren hitz batzuk eta lengoaia horietakoak ez diren hitz batzuk eman:

1. $H_1 = \{w \mid w \in A^* \wedge \exists x(x \in AA \wedge w = xx^R x)\}$
2. $H_2 = \{w \mid w \in A^* \wedge ww = www\}$
3. $H_3 = \{w \mid w \in A^* \wedge \exists u, v(u \in A^* \wedge v \in A^* \wedge uvw = wvu)\}$
4. $H_4 = \{w \mid w \in A^* \wedge \exists u(u \in A^* \wedge www = uu)\}$

2.5 Ariketak: Lengoaien definizio formala

$A = \{a, b, c\}$ alfabetoaren gainean definitutako honako lengoaia hauen deskribapen formala eman, beharren arabera multzo-notazioa, hitzen gaineko eragiketak eta lengoaien gaineko eragiketak erabiliz. Ariketa hauetako asko aurreko urteetako azterketetakoak dira eta puntuazioa zehazten da beraien arteko zailtasun desberdintasuna hobeto azaltzeko:

1. L_1 – aa, bb eta ac hitzez osatutako lengoaia.
2. L_2 – ε, bbc eta acc hitzez osatutako lengoaia.
3. L_3 – Lau sinbolo dituzten (4 luzera duten) hitzez osatutako lengoaia. Adibidez, $aaaa$, $bcab$ eta $cbbb$ L_3 lengoaiakoak dira baina ε, a, bc eta $bcbcba$ ez.
4. L_4 – a sinboloaren agerpen bakarra eta guztira lau sinbolo dituzten hitzez osatutako lengoaia. Adibidez, $abcb$, $ccac$ eta $cbca$ L_4 lengoaiakoak dira baina ε, abc, bc, $aabc$ eta $cccbb$ ez.
5. L_5 – Errepikatutako sinbolorik ez duten hitzez osatutako lengoaia. Adibidez, ε, a, ac eta acb L_5 lengoaiakoak dira baina aa, $bcac$ eta $accaaa$ ez.
6. L_6 (0,075 puntu) Gutxienez bi sinbolo desberdin dituzten hitzez osatutako L_6 lengoaiaren definizio formala eman. Adibidez, aab, $accccabab$ eta $cccbc$ hitzak L_6 lengoaiakoak dira baina aaa, b eta ε ez.
7. L_7 (0,100 puntu) Desberdinak diren bi sinbolo edo gehiago ez dituzten, hau da, sinbolo bakar baten zero edo errepikapen gehiagoz eratutako hitzez osatutako L_7 lengoaiaren definizio formala eman. Adibidez, ε, bbb, aa eta $cccc$ L_7 lengoaiakoak dira baina ac, $baaa$ eta $aaccb$ ez.
8. L_8 (0,025 puntu) Luzera bikoitia duten hitzez osatutako L_8 lengoaiaren definizio formala eman. Adibidez, ε, ab, $aaaa$ eta $cabb$ L_8 lengoaiakoak dira baina a, bab eta $accaa$ ez.
9. L_9 (0,025 puntu) Luzera bakoitia duten hitzez osatutako L_9 lengoaiaren definizio formala eman. Adibidez, a, bab eta $accaa$ L_9 lengoaiakoak dira baina ε, ab, $aaaa$ eta $cabb$ ez.
10. L_{10} (0,100 puntu) Desberdinak diren bi sinbolo edo gehiago ez dituzten eta luzera bikotia duten hitzez osatutako L_{10} lengoaiaren definizio formala eman. Adibidez, ε, $bbbb$, aa eta $cccc$ L_{10} lengoaiakoak dira baina $baaa$, aaa eta $aaccb$ ez.
11. L_{11} (0,025 puntu) a-z hasten diren hitzez osatutako L_{11} lengoaiaren definizio formala eman. Adibidez, a, aa , $abcc$, $abaa$ eta acb L_{11} lengoaiakoak dira baina ε, bc eta $cbab$ ez.

12. L_{12} (0,025 puntu) a-z hasten ez diren hitzez osatutako L_{12} lengoaiaren definizio formala eman. Adibidez, ε, bc eta $cbab$ L_{12} lengoaiakoak dira baina a, aa , $abcc$, $abaa$ eta acb ez.

13. L_{13} (0,025 puntu) a-z bukatzen diren hitzez osatutako L_{13} lengoaiaren definizio formala eman. Adibidez, a, $ccca$, aaa eta $abaa$ L_{13} lengoaiakoak dira baina ε, aab, b eta ccc ez.

14. L_{14} (0,025 puntu) a-z bukatzen ez diren hitzez osatutako L_{14} lengoaiaren definizio formala eman. Adibidez, ε, c, ccb, aac eta $abac$ L_{14} lengoaiakoak dira baina a, aa, baa eta $acbaaa$ ez.

15. L_{15} (0,050 puntu) a-z hasi eta a-z bukatzen diren hitzez osatutako L_{15} lengoaiaren definizio formala eman. Adibidez, a, aa, $abba$ eta $acaaabba$ L_{15} lengoaiakoak dira baina ε, c, ab, bbc eta $ccaa$ ez.

16. L_{16} (0,050 puntu) ε hitz hutsaz gain, a-ren desberdina den sinbolo batez hasi eta a-ren desberdina den sinbolo batez bukatzen diren hitzez osatutako L_{16} lengoaiaren definizio formala eman. Adibidez, ε, b, $baac$, ccc eta $ccaaabac$ L_{16} lengoaiakoak dira baina a, abb, $abba$ eta caa ez.

17. L_{17} – Luzera bakoitia edukitzeaz gain erdiko posizioan a sinboloa duten hitzez osatutako L_{17} lengoaiaren definizio formala eman. Adibidez, a, aaa, $ababc$ eta $ccaabba$ L_{17} lengoaiakoak dira baina ε, b, aa, abc eta $abcc$ ez.

18. L_{18} – b sinboloaz bukatzen diren hitzez osatutako L_{18} lengoaiaren definizio formala eman. Adibidez, b, aab, bbb eta $bacb$ L_{18} lengoaiakoak dira baina ε, c, ba, ccc eta $abbbc$ ez.

19. L_{19} – a sinboloaz hasi eta b sinboloaz bukatzen diren hitzez osatutako L_{19} lengoaiaren definizio formala eman. Adibidez, ab, $aaacb$ eta $abcab$ L_{19} lengoaiakoak dira baina ε, a, ca, bca eta $bbbcb$ ez.

20. L_{20} – a sinboloaz hasi edo b sinboloaz bukatzen diren hitzez osatutako L_{20} lengoaiaren definizio formala eman. Adibidez, a, b, ab, ac, cb, aaa, $aacb$ eta ccb L_{20} lengoaiakoak dira baina ε, c, cca eta $baac$ ez.

21. L_{21} – a sinboloaz hasi baina b sinboloaz bukatzen ez diren hitzez osatutako L_{21} lengoaiaren definizio formala eman. Adibidez a, ac, aaa, $abbc$ eta $abbba$ L_{21} lengoaiakoak dira baina ε, b, ab, ccb eta $cacb$ ez.

22. L_{22} (0,025 puntu) a sinboloa b sinboloa baino gehiagotan duten hitzez osatutako L_{22} lengoaiaren definizio formala eman. Adibidez, a, acc, $baac$ eta aaa L_{22} lengoaiakoak dira baina ε, ab, $bbac$, bbb eta $cccc$ ez.

23. L_{23} (0,025 puntu) a kopuru bikoitia duten hitzez osatutako L_{23} lengoaiaren definizio formala eman. Adibidez, ε, b, aa, $baab$, $caba$, $aaaa$ eta ccc L_{23} lengoaiakoak dira baina a, bac, aaa eta $ccab$ ez.

24. L_{24} (0,025 puntu) a sinboloa b sinboloa baino gehiagotan ez duten hitzez osatutako L_{24} lengoaiaren definizio formala eman. Adibidez, ε, ab, ccc, bc eta $bacc$ L_{24} lengoaiakoak dira baina a, aba, ca eta aaa ez.

25. L_{25} (0,075 puntu) a kopuru bikoitia eta a sinboloa b sinboloa baino gehiagotan duten hitzez osatutako L_{25} lengoaiaren definizio formala eman. Adibidez, aa, $caba$ eta $aaaa$ L_{25} lengoaiakoak dira baina ε, b, $aaab$, ccb eta acc ez.

26. L_{26} (0,025 puntu) b-rik eta c-rik ez duten hitzez osatutako L_{26} lengoaiaren definizio formala eman. Adibidez, ε, a, aa eta aaa L_{26} lengoaiakoak dira baina b, $abca$, ccc eta abb ez.

27. L_{27} (0,025 puntu) a-rik eta c-rik ez duten hitzez osatutako L_{27} lengoaiaren definizio formala eman. Adibidez, ε, b, bb eta $bbbb$ L_{27} lengoaiakoak dira baina c, aaa, ac, bac eta bcc ez.

28. L_{28} – c sinboloaren agerpenik ez edukitzeaz gain a-ren agerpen denak (a-rik baldin badago) ezkerreko aldean eta b-ren agerpen denak (b-rik baldin badago) eskuineko aldean dituzten hitzez osatutako L_{28} lengoaiaren definizio formala eman. Adibidez, ε, aab, $aaabbbb$, aaa eta bb hitzak L_{28} lengoaiakoak dira baina caa, $abcb$, $bbaaa$ eta ccc ez dira L_{28} lengoaiakoak.

29. L_{29} (0,200 puntu) c-rik ez duten eta, a-rik baldin badago, a-ren agerpen denak alde batean (ezkerreko aldean edo eskuineko aldean) jarraian eta, b-rik baldin badago, b-ren agerpen denak beste aldean jarraian dituzten hitzez osatutako L_{29} lengoaiaren definizio formala eman. Adibidez, ε, $aabbb$, $baaa$, bbb eta $aaaa$ hitzak L_{29} lengoaiakoak dira baina $aabaa$, $aaaccbb$ eta $abaaa$ ez dira L_{29} lengoaiakoak.

30. L_{30} (0,100 puntu) c-rik ez, a eta b sinboloak kopuru berean eta a denak (a-rik baldin badago) ezkerreko aldean elkarren jarraian eta b denak (b-rik baldin badago) eskuineko aldean elkarren jarraian dituzten hitzez osatutako L_{30} lengoaiaren definizio formala eman. Adibidez, ε, ab, $aabb$ eta $aaabbb$ hitzak L_{30} lengoaiakoak dira baina $aabbb$, $aaacbb$, aaa eta $bbaa$ ez dira L_{30} lengoaiakoak.

31. L_{31} (0,125 puntu) b kopurua a kopurua baino handiagoa eta c kopurua b kopurua baino handiagoa izateaz gain, a-rik baldin badago, a-ren agerpen denak ezkerreko aldean elkarren jarraian, b-ren agerpen denak erdiko aldean elkarren jarraian eta c-ren agerpen denak eskuineko aldean elkarren jarraian dituzten hitzez osatutako L_{31} lengoaiaren definizio formala eman. Adibidez, $bccc$, $abbccc$, $aabbbcccccc$ eta $bbcccc$ hitzak L_{31} lengoaiakoak dira baina ε, $aabbb$, $aaacbb$, aaa, ccc eta $bbaa$ ez dira L_{31} lengoaiakoak.

32. L_{32} (0,025 puntu) b eta c kopuruen baturaren berdina den a kopurua duten hitzez osatutako L_{32} lengoaiaren definizio formala eman. Adibidez, ε, $aabc$, $acccaa$ eta $cabaca$ hitzak L_{32} lengoaiakoak dira baina aaa, b eta $accb$ ez.

33. L_{33} (0,100 puntu) a-ren agerpen bakoitzaren jarraian gutxienez bi b dituzten hitzez osatutako L_{33} lengoaiaren definizio formala eman. Adibidez, ε, $bcbbcabb$, $abbbabbabbc$ eta $cccc$ L_{33} lengoaiakoak dira baina $baaa$, ab eta $aaccb$ ez.

34. L_{34} (0,025 puntu) b-rik eta c-rik ez duten eta a-kopuru bikoitia duten hitzez osatutako L_{34} lengoaiaren definizio formala eman. Adibidez, ε, $aaaa$ eta aa L_{34} lengoaiakoak dira baina $baaa$, bb, $cbbb$, c, aaa eta $aaccb$ ez.

35. L_{35} Jarraian aipatzen diren baldintzetakoren bat (gutxienez bat) betetzen duten hitzez osatutako lengoaia:

    - b eta c sinbolorik ez edukitzea
    - a sinboloaren agerpen-kopurua bikoitia izatea.

    Adibidez, ε, aaa, $aaaa$, $abca$, bb eta aa L_{35} lengoaiakoak dira baina $baaa$, bab, $cbbbaaa$, ca, $aaca$ eta $aaccba$ ez.

36. L_{36} (0,025 puntu) Gutxienez a bat eta gutxienez c bat duten hitzez osatutako L_{36} lengoaiaren definizio formala eman. Adibidez, ca, $aabbbbaabc$ eta $ccccaa$ L_{36} lengoaiakoak dira baina ε, $baaa$, bb, $cbbb$, c eta aaa ez.

37. L_{37} (0,050 puntu) ac katea edo ca katea (gutxienez bietako bat) gutxienez behin duten hitzez osatutako L_{37} lengoaiaren definizio formala eman. Adibidez, ca, $acabbbbccaac$ eta $acaccbaac$ L_{37} lengoaiakoak dira baina ε, $cbaaa$, bba, $cbbab$, bbb, c eta aaa ez.

38. L_{38} (0,100 puntu) a eta c elkarren jarraian (ez ac bezala eta ez ca bezala) ez dituzten hitzez osatutako L_{38} lengoaiaren definizio formala eman. Adibidez, ε, $cbaaa$, $bcba$, $cbbb$, c eta aaa L_{38} lengoaiakoak dira baina ca, $aabbbbaac$ eta $ccccaa$ ez.

39. L_{39} (0,025 puntu) a eta b sinboloak kopuru berean dituzten hitzez osatutako L_{39} lengoaiaren definizio formala eman. Adibidez, $aabacbcb$, ccc, ε, $aaabbb$, $abab$ eta $bccaccc$ hitzak L_{39} lengoaiakoak dira baina b, ca, $aabbbbca$ eta $ccccaa$ ez.

40. L_{40} (0,025 puntu) a-z hasi, b-z bukatu eta a eta b sinboloak kopuru berean dituzten hitzez osatutako L_{40} lengoaiaren definizio formala eman. Adibidez, $aabacbcb$, acb, $aababb$ eta $accbbcaab$ L_{40} lengoaiakoak dira baina $abba$ ez da L_{40} lengoaiakoa, a eta b kopuru berean agertu arren, hitza ez delako b-z bukatzen.

41. L_{41} (0,025 puntu) aa azpikatea duten hitzez osatutako L_{41} lengoaiaren definizio formala eman. Adibidez, $aaaaa$, $aabacbcb$, $acaaab$, $cbaabaab$ eta $accbaaaab$ L_{41} lengoaiakoak dira baina ε, b, ca, $abbbca$ eta $cccc$ ez.

42. L_{42} (0,075 puntu) aa eta cc azpikate biak dituzten hitzez osatutako L_{42} lengoaiaren definizio formala eman. cc azpikatea aa baino lehenago ager daiteke edo ez. Lengoaia honetako hitz bakoitzak azpikate biak izan behar ditu gutxienez behin. Adibidez, $ccaaaaa$, $aabacbcccb$, $accaaab$, $ccbaabaab$ eta $accbaaaabcc$ L_{42} lengoaiakoak dira baina $bacbcc$ ez da L_{42} lengoaiakoa aa azpikatea ez duelako.

43. L_{43} (0,050 puntu) aa azpikatea ez duten hitzez osatutako L_{43} lengoaiaren definizio formala eman. Adibidez, $cabbccaba$, $cccabbbb$, $cccc$, ε eta $accbbbabab$ hitzak L_{43} lengoaiakoak dira.

44. L_{44} (0,100 puntu) b-rik agertzen bada, b guztiak batera (jarraian) dituzten hitzez osatutako L_{44} lengoaiaren definizio formala eman. Adibidez, $ccaaaaa$, $aabbbccca$, ccc, $bbaccaaa$, ε, $bbbb$ eta $ccbbb$ hitzak L_{44} lengoaiakoak dira. Bestalde, $bacbcc$ hitza ez da L_{44} lengoaiakoa b denak ez daudelako jarraian.

45. L_{45} (0,050 puntu) Luzera gutxienez 2 eta hasieran eta bukaeran sinbolo bera duten hitzez osatutako L_{45} lengoaiaren definizio formala eman. Adibidez, $aabacbca$, bcb, $babb$ eta $ccccc$ hitzak L_{45} lengoaiakoak dira baina $cbbb$ ez da L_{45} lengoaiakoa hasieran eta bukaeran ez duelako sinbolo bera. Beste aldetik, c hitza ere ez da L_{45} lengoaiakoa bere luzera 2 baino txikiagoa baita.

46. L_{46} (0,050 puntu) ab hitza nahi adina aldiz errepikatuz eratutako hitzez osatutako L_{46} lengoaiaren definizio formala eman. Adibidez, $ababab$, ab eta ε hitzak L_{46} lengoaiakoak dira baina aba, $bababa$ eta $cabc$ hitzak ez dira L_{46} lengoaiakoak.

47. L_{47} (0,050 puntu) aa azpikatea edo cc azpikatea duten hitzez osatutako L_{47} lengoaiaren definizio formala eman. Hitz bakoitzak gutxienez azpikate horietako bat gutxienez behin eduki behar du. Adibidez, $ccaaaaa$, $bacbcccb$, $acaaab$, $cccc$, $ccba$ eta $aabccccb$ hitzak L_{47} lengoaiakoak dira baina $bacbca$ hitza ez da L_{47} lengoaiakoa aa eta cc azpikateak ez baitira agertzen hitz horretan.

48. L_{48} (0,050 puntu) a sinboloaz hasi, b sinboloaz bukatu eta gutxienez c bat duten hitzez osatutako L_{48} lengoaiaren definizio formala eman. Adibidez, $accaaaab$, $aabbcbccbb$, acb eta $aaccbaccb$ hitzak L_{48} lengoaiakoak dira baina ε, $bacbcc$ eta $bbbb$ hitzak ez dira L_{48} lengoaiakoak.

49. L_{49} (0,025 puntu) Hiru baino handiagoa den luzera eta gainera hirugarren posizioan a sinboloa duten hitzez osatutako L_{49} lengoaiaren definizio formala eman. Adibidez, $aaaa$, $ccab$, $cbabbbaac$, $ccabcbaaaa$ eta $bcaccc$ hitzak L_{49} lengoaiakoak dira. Baina ε, aa, aaa, $aabbca$, ba eta bba ez dira L_{49} lengoaiakoak.

50. L_{50} (0,075 puntu) a-z hasi, b-z bukatu, c bakarra, hasierako a eta c bakarraren artean nahi adina b (zero edo gehiago) baina a-rik ez eta c bakarraren eta bukaerako b-aren artean nahi adina a (zero edo gehiago) baina b-rik ez duten hitzez osatutako L_{50} lengoaiaren definizio formala eman. Adibidez, $abbbcaab$, acb, $acaaab$ eta $abbbbcb$ L_{50} lengoaiakoak dira baina $abba$, ε, $abbcaba$, $abbcac$, $acbbb$, aaa eta ab ez dira L_{50} lengoaiakoak.

51. L_{51} (0,050 puntu) Hasieran a sinboloaren agerpen batzuk (zero edo gehiago) gero b sinboloaren agerpen batzuk (bat edo gehiago) eta bukatzeko, c sinboloaren agerpen batzuk, (justu a sinboloaren agerpen-kopuru bera) dituzten hitzez osatutako L_{51} lengoaiaren definizio formala eman. Adibidez, $aabcc$, $bbbb$, b, $abbc$ eta $aabbbcc$ L_{51} lengoaiakoak dira. Baina bc, ac, ε, $aaccbbb$ eta $aaabbb$ ez dira L_{51} lengoaiakoak.

52. L_{52} (0,075 puntu) abc azpikatea hasieran edo bukaeran (edo bietan) duten hitzez osatutako L_{52} lengoaiaren definizio formala eman. abc azpikatea leku gehiagotan ere ager daiteke hitzaren erdian. Adibidez, $abcaaaa$, abc, $accaaabc$, $abcbbbabc$ eta $abccabcaaa$ L_{52} lengoaiakoak dira baina ε, a eta $bacbcc$ ez dira L_{52} lengoaiakoak.

53. L_{53} (0,025 puntu) L_{52} lengoaiakoak ez diren hitzez osatutako L_{53} lengoaiaren definizio formala eman.

54. L_{54} (0,075 puntu) b-rik agertzen bada, c-rik ez duten hitzez osatutako L_{54} lengoaiaren definizio formala eman. Adibidez, $ccaaaaa$, $aabbba$, ccc, $aaaa$, ε, $bbbb$ eta $acaac$ hitzak L_{54} lengoaiakoak dira. Bestalde, $bacbcc$ hitza ez da L_{54} lengoaiakoa.

55. L_{55} (0,075 puntu) a-z hasi eta gero c-rik ez baina gutxienez bi b edo a-z hasi eta gero dena c duten hitzez osatutako L_{55} lengoaiaren definizio formala eman. Adibidez, abb, $aababa$, $aabaaab$, $aababab$ eta $acccc$ hitzak L_{55} lengoaiakoak dira baina ε, $aabbcb$, $caacbb$, $cccc$ eta bbc ez dira L_{55} lengoaiakoak.

56. L_{56} – Gutxienez sinbolo bat edukitzeaz gain posizio bikoitietan a sinboloa eta posizio bakoitietan b sinboloa duten hitzez osatutako lengoaia. Adibidez, $babab$, b eta $bababa$ hitzak L_{56} lengoaiakoak dira baina ε, $aabbcb$, $caacbb$, $cccc$ eta bbc ez dira L_{56} lengoaiakoak.

57. L_{57} – Luzera bikoitia edukitzeaz gain posizio bikoitietan a sinboloa eta posizio bakoitietan b sinboloa duten hitzez osatutako lengoaia. Adibidez, ε, $baba$, ba eta $bababa$ hitzak L_{57} lengoaiakoak dira baina $aabbcb$, $caacbb$, $cccc$, $babab$ eta bbc ez dira L_{57} lengoaiakoak.