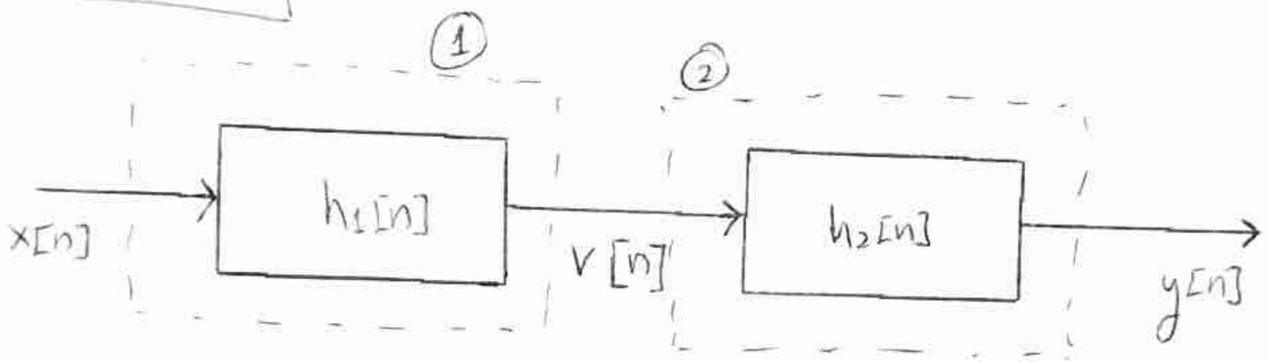
Επεξεργασία Φωνής και Φυσικής Γλώσσας

1η Σειρά Αναλυτικών Ασκήσεων

Άσκηση 1



1. Θέλω να βρω την κρουστική απόκριση $h[n]$ του παραπάνω συστήματος για την οποία ισχύει:

$$y[n] = h[n] * x[n]$$

Έχω για το 1ο υποσύστημα: $v[n] = h_1[n] * x[n]$

Έχω για το 2ο υποσύστημα: $y[n] = h_2[n] * v[n]$ $\Big\} \Rightarrow$

άρα

$$\Rightarrow y[n] = h_2[n] * \left(h_1[n] * x[n]\right) \Rightarrow y[n] = h_2[n] * h_1[n] * x[n] \Rightarrow$$

$$\xrightarrow[\text{of convolution}]{\text{commutativity}} y[n] = h_1[n] * h_2[n] * x[n], \text{ άρα: } \boxed{h[n] = h_1[n] * h_2[n]}$$

(απόδειξη στη συνέχεια)

2. Θα δείξουμε ότι ισχύει η ιδιότητα commutativity στη συνέλιξη, δηλαδή ότι:

$$h_1[n] * h_2[n] = h_2[n] * h_1[n].$$

έχω ότι $h_1[n] * h_2[n] = \sum_{m=-\infty}^{+\infty} h_1[m] \cdot h_2[n-m]$

με αλλαγή μεταβλητής θέτοντας $u = n - m \Rightarrow m = n - u$ έχω:

- για $m = -\infty \Rightarrow u = +\infty$
- για $m = +\infty \Rightarrow u = -\infty$

Άρα: $h_1[n] * h_2[n] = \sum_{u=-\infty}^{u=+\infty} h_1[n-u] \cdot h_2[u] = h_2[n] * h_1[n]$

Άρα, ισχύει $\boxed{h_1[n] * h_2[n] = h_2[n] * h_1[n].}$

Επομένως, η ολική κρουστική απόκριση δεν εξαρτάται με την σειρά με την οποία εμφανίζονται τα συστήματα.

3. $$H(z) = \left(\sum_{r=0}^{M} b_r z^{-r} \right) \cdot \left(\frac{1}{1 - \sum_{k=1}^{N} \alpha_k z^{-k}} \right) = H_1(z) \cdot H_2(z)$$

και είναι: $Y(z) = H(z) \cdot X(z) \Rightarrow$

$$Y(z) = \left(\sum_{r=0}^{M} b_r z^{-r} \right) \left(\frac{1}{1 - \sum_{k=1}^{N} \alpha_k z^{-k}} \right) X(z) \Rightarrow$$

$$Y(z) - Y(z) \cdot \left(\sum_{k=1}^{N} \alpha_k z^{-k} \right) = X(z) \cdot \left(\sum_{r=0}^{M} b_r z^{-r} \right) \overset{Z^{-1}\{\cdot\}}{\Rightarrow}$$

$$\boxed{y[n] - \sum_{k=1}^{N} \alpha_k \cdot y[n-k] = \sum_{r=0}^{M} b_r \cdot x[n-r]}$$ *

4. Αν έχω τα δύο συστήματα με ανάποδη σειρά δεν αλλάζει η ολική κρουστική απόκριση, ούτε αλλάζουν οι εξισώσεις διαφορών σε σχέση με πριν.*

$$\boxed{\text{Άσκηση 2}}$$

1. $\tilde{X}(k) = \sum_{m=-\infty}^{\infty} x(m)\, e^{-j\frac{2\pi}{N}km}$ που προκύπτει από δειγματοληψία του $X(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m)\, e^{-j\omega m}$ στις συχνότητες $\omega_k = 2\pi k/N$, $k = 0, 1, \dots, N-1$

άρα τα δείγματα είναι (DFT): $\tilde{x}(n) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} \tilde{X}(k) e^{j\frac{2\pi}{N}kn}$

Θέλουμε να δείξουμε ότι $\tilde{x}(n) = \sum_{r=-\infty}^{\infty} x(n+rN)$

Έχουμε:

$$\tilde{x}(n) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} \tilde{X}(k) e^{j\frac{2\pi}{N}kn} =$$

$$= \frac{1}{N} \cdot \sum_{k=0}^{N-1} \sum_{m=-\infty}^{\infty} x(m)\, e^{-j\frac{2\pi}{N}\cdot km} \cdot e^{j\frac{2\pi}{N}kn} =$$

$$= \sum_{m=-\infty}^{\infty} x(m) \cdot \frac{1}{N} \cdot \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}kn} \cdot e^{-j\frac{2\pi}{N}km}.$$

Η ανωτέρω ταυτότητα εκφράζει την ορθογωνικότητα των μιγαδικών εκθετικών:

$$\frac{1}{N}\cdot\sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}kn}\cdot e^{-j\frac{2\pi}{N}km} =$$

$$= \frac{1}{N}\cdot\sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}k(n-m)} = \begin{cases} 1, & n-m=r\cdot N,\ r\in\mathbb{Z} \\ 0, & \text{αλλού} \end{cases}$$

άρα ο υπολογισμός έχει νόημα για $n-m=r\cdot N \Rightarrow \boxed{m=n-rN}$ $r\in\mathbb{Z}$

Άρα ισοδύναμα παίρνουμε με αλλαγή μεταβλητής

$$n-m=r\cdot N \Rightarrow r=\frac{n-m}{N}$$

$$m\to+\infty \Rightarrow r\to-\infty$$

$$m\to-\infty \Rightarrow r\to+\infty$$

Άρα: $$\tilde{x}(n)=\sum_{r=-\infty}^{+\infty} x(n-rN)$$

και με αλλαγή μεταβλητής όπου r το $-r$:

$$\boxed{\tilde{x}(n)=\sum_{-\infty}^{+\infty} x(n+rN)}$$

2. Θέλουμε να μην έχουμε παραμόρφωση στο πεδίο του χρόνου όταν το $X(e^{j\omega})$ δειγματοληπτείται.

Το ω είναι συνεχής μεταβλητή, άρα υπάρχει άπειρος αριθμός πιθανών τιμών από το 0 μέχρι το 2π ή έστω από το $-\pi$ μέχρι το π. Έτσι, το $X(e^{j\omega})$ μπορεί να υπολογιστεί μόνο σε ένα πεπερασμένο σύνολο συχνοτήτων. Χρησιμοποιούμε, λοιπόν, N δείγματα γύρω από τον μοναδιαίο κύκλο:

$\omega_k = \frac{2\pi k}{N}$, $k = 0, 1, \dots, N-1$ και να πάρω

N δείγματα του $X(e^{j\omega})$ για το $X[k]$

Ωστόσο δε μπορούμε να υπολογίσουμε πάντα το $X[k]$ αφού το $x(n)$ μπορεί να είναι άπειρο και να αθροίζουμε για όλα τα n. Ακόμα και αν το $x(n)$ είναι πεπερασμένο δε μπορούμε να ανακτήσουμε πάντα τέλεια το $x(n)$ από το $X[k]$ επειδή το $X[k]$ ουσιαστικά είναι τα δείγματα του DTFT. Επομένως, υπάρχουν κάποιες απαραίτητες προϋποθέσεις για τις οποίες θα μπορεί να ανακτηθεί το $x(n)$ από το $X[k]$.

Η δειγματοληψία στο χρόνο είναι κατ' αντιστοιχία όμοια με αυτή στο πεδίο των συχνοτήτων. Προκειμένου να γίνει πλήρης ανακατασκευή θα πρέπει να αποφευχθεί το aliasing. Επομένως, το σήμα θα πρέπει να περιορίζεται σε μια συχνότητα Ω_0 και το διάστημα δειγματοληψίας θα πρέπει να είναι αρκετά σύντομο έτσι ώστε $\Delta T < \frac{2\pi}{\Omega_0}$.

Μπορούμε να εφαρμόσουμε τη δυϊκότητα χρόνου - συχνότητας για να απεικονίσουμε τη διαδικασία δειγματοληψίας στο πεδίο της συχνότητας. Αν το διάστημα δειγματοληψίας στη συχνότητα δεν είναι αρκετά σύντομο έχουμε aliasing στο χρόνο. Εάν το διάστημα δειγματοληψίας στο πεδίο των συχνοτήτων είναι $\Delta\Omega$ τότε αντιστοιχεί σε σήματα στο πεδίο του χρόνου σε κάθε $\frac{2\pi}{\Delta\Omega}$. Προκειμένου να ανακτηθεί το $x(t)$ από το $\tilde{x}(t)$ με time - windowing, το $x(t)$ θα πρέπει να

Περιορίζεται χρονικά σε T_0 και το διάστημα δειγματοληψίας πρέπει να είναι αρκετά μικρό ώστε $\Delta\Omega < \frac{2\pi}{T_0}$.

Έχουμε ουσιαστικά το ίδιο αποτέλεσμα στον διακριτό χρόνο.

Το διάστημα δειγματοληψίας πρέπει να πληροί την ακόλουθη προϋπόθεση:

$2\pi\,\Delta\Omega < M$ (όταν το $X(\omega)$ είναι M σημείων)

Αν υποδηλώσουμε ως N τον αριθμό των δειγμάτων συχνότητας από 0 έως 2π πρέπει να ισχύει $2\pi N \geq M \Delta\Omega$. Υπό αυτήν την προϋπόθεση το $x[n]$ μπορεί να ανακτηθεί τέλεια από τα δείγματα του DTFT.

3. $y(n) = x(nM)$

$v(n) = x(n) \cdot p(n)$

$p(n) = \sum_{r=-\infty}^{\infty} \delta(n+rM)$

$y(n) = v(nM) = x(nM)$

ν.δ.ο. $Y(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} X(e^{j\frac{(\omega-2\pi k)}{M}})$

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} y(n) e^{-j\omega n} = \sum_{n=-\infty}^{+\infty} x(n) \cdot p(n) e^{-j\omega n/M}$$

με $p(n) = \sum_{r=-\infty}^{+\infty} \delta(n+rM) = \begin{cases} 1, & \text{αν } n = kM \\ 0, & \text{αλλού} \end{cases} \Rightarrow$

και είναι $\frac{1}{M} \sum_{k=0}^{M-1} e^{j\frac{2\pi}{M}kn} = \sum_{r=-\infty}^{+\infty} \delta[n-kM]$

άρα $p(n) = \frac{1}{M} \sum_{k=0}^{M-1} e^{j\frac{2\pi}{M}kn}$

Επομένως: $Y(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} \sum_{n=-\infty}^{+\infty} x(n) \cdot e^{-j\frac{(\omega-2\pi k)n}{M}} \Rightarrow$

$$\boxed{Y(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} X(e^{j(\omega-2\pi k)/M})}$$

4. Στο ερώτημα 2 απαντήθηκε επιπρόσθετα το τι πρέπει να ισχύει για το $X(e^{j\omega})$ ώστε να μην υπάρχει παραμόρφωση στο πεδίο της συχνότητας όταν το $x(n)$ δειγματοληπτείται, καθώς έγινε πολύ εκτενής περιγραφή του ζητήματος της δειγματοληψίας

Άσκηση 3

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m) \cdot w(n-m) \cdot x(m+k) \cdot w(n-k-m)$$

1. $$R_n(-k) = \sum_{m=-\infty}^{\infty} x(m) \cdot w(n-m) \cdot x(m-k) \cdot w(n+k-m)$$

Θέτω $m = u+k \Rightarrow u = m-k$

για $m \to +\infty \Rightarrow u \to +\infty$

$m \to -\infty \Rightarrow u \to -\infty$

Άρα $$R_n(-k) = \sum_{u=-\infty}^{+\infty} x(u+k) \cdot w(n-k-u) \cdot x(u) \cdot w(n-u) = R_n(k)$$

άρα άρτια!

2. $$R_n(k) = R_n(-k) = \sum_{m=-\infty}^{+\infty} x(m) \cdot x(m-k) \cdot w(n-m) \cdot w(n+k-m)$$

έστω $h_k(n) = w(n) \cdot w(n+k)$ τότε $h_k(n-m) = w(n-m) w(n-m+k)$

Άρα $$\boxed{R_n(k) = \sum_{m=-\infty}^{+\infty} x(m) \cdot x(m-k) \cdot h_k(n-m)}$$ με $h_k(n) = w(n) \cdot w(n+k)$

3.

$$w(n) = \begin{cases} \alpha^n &, n \geq 0 \\ 0 &, n < 0 \end{cases}$$

έχω $h_k(n) = w(n) \cdot w(n+k)$

$$w(n+k) = \begin{cases} \alpha^{n+k} &, n \geq -k \\ 0 &, n < -k \end{cases}$$

Επομένως $\boxed{h_k(n) = \begin{cases} \alpha^{2n+k} &, n \geq \max\{0, -k\} \\ 0 &, \text{αλλού} \end{cases}}$

δηλαδή αλλιώς: $\boxed{h_k(n) = \begin{cases} \alpha^{2n+k} \cdot u(n) &, k \geq 0 \\ \alpha^{2n+k} \cdot u(n+k) &, k < 0 \end{cases}}$

4. • $\alpha^{2n+k} u[n] = \alpha^k \cdot \alpha^{2n} u[n] = \alpha^k \cdot \alpha^n \cdot \alpha^n \cdot u[n]$

① $\alpha^n u[n] \overset{z}{\longleftrightarrow} \dfrac{1}{1 - az^{-1}}$, $|z| > |a|$ $\Rightarrow$ $\alpha^n \alpha^n u[n] \overset{z}{\longleftrightarrow} \dfrac{1}{1 - a(a^{-1}z)^{-1}} = \dfrac{1}{1 - a^2 z^{-1}}$

άρα για $k \geq 0$: $\underline{H_k(z) = \dfrac{\alpha^k}{1 - a^2 z^{-1}}}$

• $\alpha^{2n+k} \cdot u[n+k] = \alpha^n \cdot \alpha^{n+k} \cdot u[n+k]$ με $\alpha^{n+k} \cdot u[n+k] \overset{z}{\underset{①}{\longleftrightarrow}} \dfrac{z^k}{1 - az^{-1}}$ $\Rightarrow$

$\Rightarrow a^n \cdot a^{n+k} \cdot u[n+k] \overset{z}{\longleftrightarrow} \dfrac{a^{-k} \cdot z^k}{1 - a^2 z^{-1}}$, άρα για $k<0$: $\underline{H_k(z) = \dfrac{\alpha^{-k} z^k}{1 - a^2 z^{-1}}}$

Άρα: $\boxed{H_k(z) = \begin{cases} \dfrac{\alpha^k}{1 - a^2 z^{-1}} &, k \geq 0 \\ \dfrac{\alpha^{-k} \cdot z^k}{1 - a^2 z^{-1}} &, k < 0 \end{cases}}$

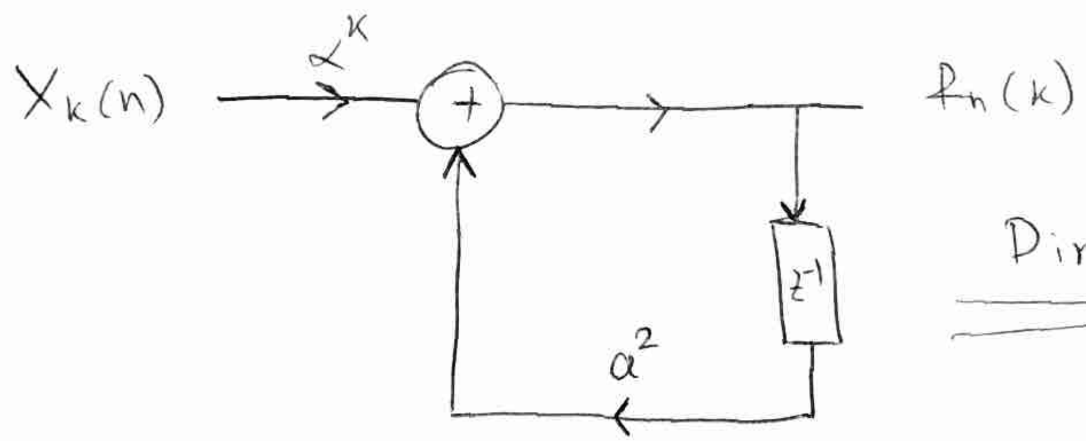
έστω $x_k(m) = x(m) \cdot x(m-k)$

Άρα $R_n(k) = \sum_{m=-\infty}^{+\infty} x_k(m) \cdot h_k(n-m) \Rightarrow$

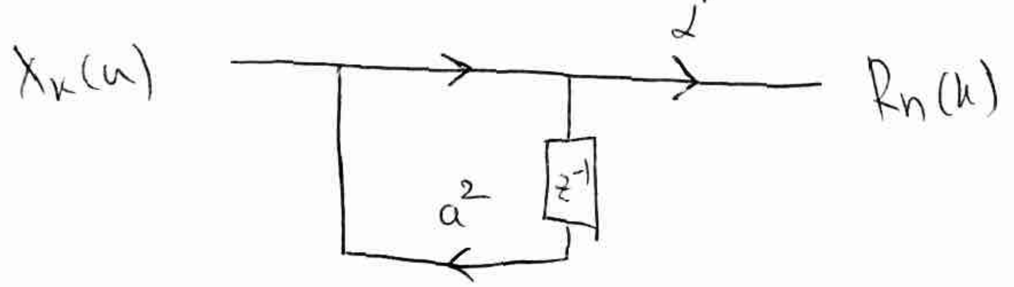
$$\boxed{R_n(k) = x_k(n) * h_k(n)}$$

άρα $\xrightarrow{z}$ $R_n(z) = X_k(z) \cdot H_k(z)$

για $k \geq 0$: $H_k(z) = \dfrac{\alpha^k}{1 - \alpha^2 z^{-1}}$ άρα:



Direct Form I



Direct Form II

άρα $R_n(k) = a^k \cdot X_k(n) + a^2 R_{n-1}(k) \Rightarrow$

$$\boxed{R_n(k) = a^k \cdot x(n) \cdot x(n-k) + a^2 \cdot R_{n-1}(k)} \quad k \geq 0$$

5. $w'(n) = \begin{cases} n \cdot a^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$

$\Rightarrow w'(n) = n \cdot a^n u(n)$

$h'_k(n) = w'(n) \cdot w'(n+k) = n \cdot a^n u(n) \cdot (n+k) \cdot a^{n+k} \cdot u(n+k) =$

$= n \cdot (n+k) \cdot \alpha^{2n+k} \cdot u(n) \cdot u(n+k) = \begin{cases} n(n+k) a^{2n+k} \cdot u(n), & k \geq 0 \\ n(n+k) a^{2n+k} \cdot u(n+k), & k < 0 \end{cases}$

Ομοίως, στο σύνολο παρατηρούμε:

$w'(n) = n \cdot w(n)$ και $w'(n+k) = (n+k) \cdot w(n+k)$

άρα $h'_k(n) = n(n+k) \cdot h_k(n) = n^2 \cdot h_k(n) + k \cdot n \cdot h_k(n)$

$$n^2 \cdot h_k(n) \xleftrightarrow{z} -z \cdot \frac{d}{dz}\left(-z \frac{d H_k(z)}{dz}\right) = -z \frac{d}{dz}\left(\frac{z \cdot a^{k+2}}{(z-a^2)^2}\right) =$$

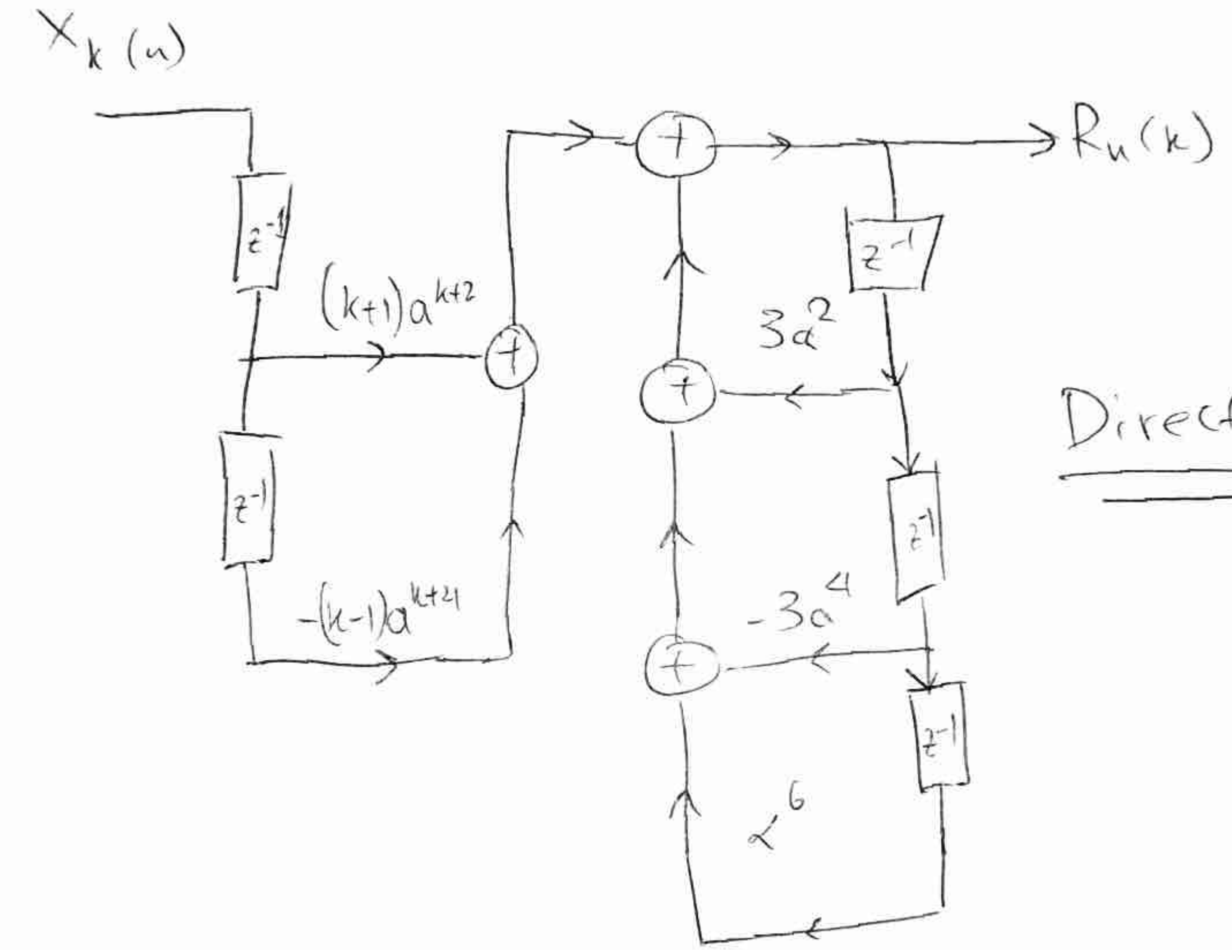
$$= \frac{z \cdot a^{k+2}(z+a^2)}{(z-a^2)^3}$$

$$k n h_k(n) \xleftrightarrow{z} \frac{k \cdot z \cdot a^{k+2}}{(z-a^2)^2}$$

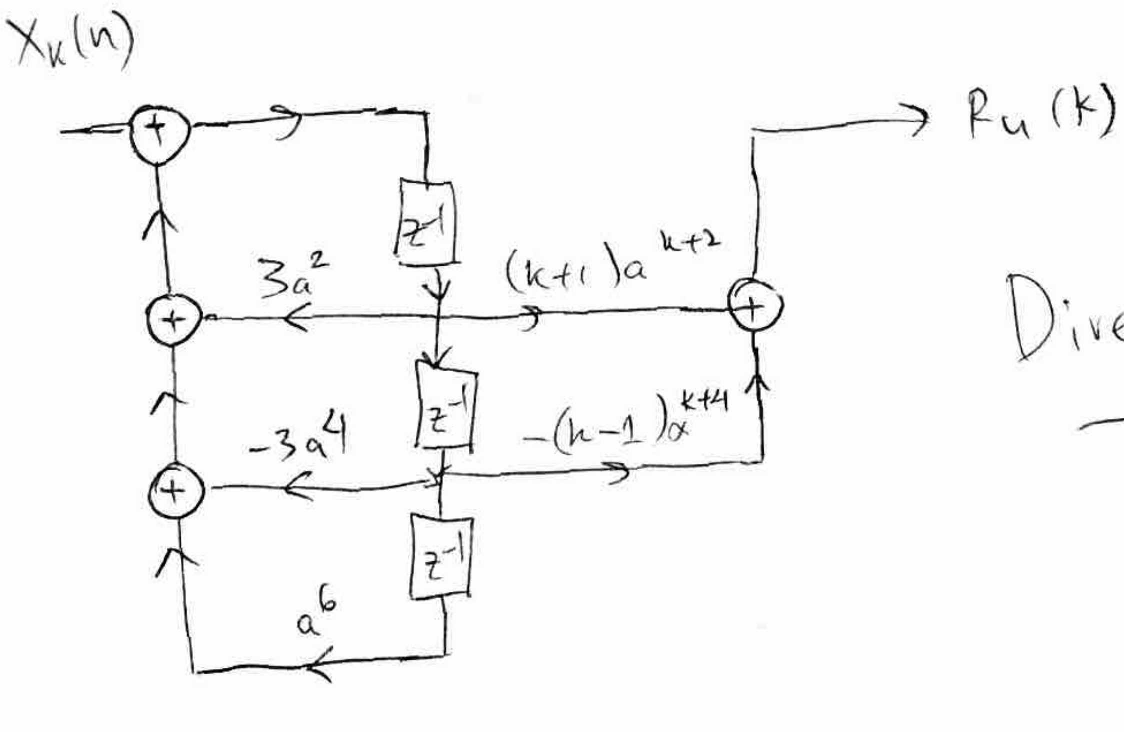
άρα για $k \geq 0$: $\boxed{H'_k(z) = \frac{z \cdot a^{k+2}(z+a^2)}{(z-a^2)^3} + \frac{k \cdot z \cdot a^{k+2}}{(z-a^2)^2}}$

με απλοποιήσεις στη μορφή που επιθυμούμε:

$$\Rightarrow H_{k}'(z) = \frac{(k+1)\,a^{k+2}\cdot z^{-1} - (k-1)\,a^{k+4}\cdot z^{-2}}{1 - 3a^2 z^{-1} + 3a^4 z^{-2} - a^6 \cdot z^{-3}}$$



Direct Form I



Direct form II

Άρα $R_n(k) = (k+1)\, a^{k+2}\, x_n(n-1) - (k-1)\, a^{k+4}\, x_n(n-2) +$

$+ 3a^2 R_{n-1}(k) - 3a^4 R_{n-2}(k) + a^6 R_{n-3}(k) \Rightarrow$

$$\Rightarrow R_n(k) = (k+1)\, a^{k+2} \cdot x(n-1)\, x(n-k-1)$$

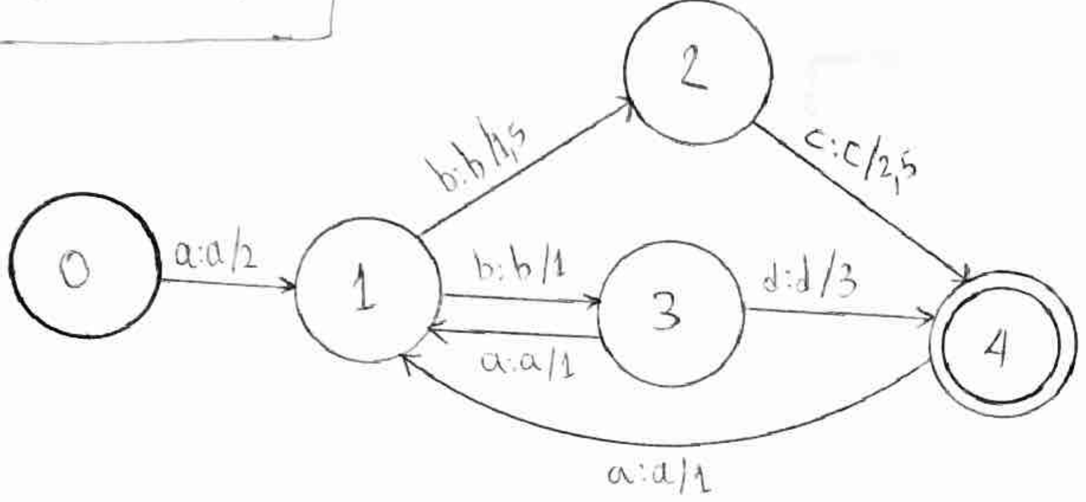
$$- (k-1)\, a^{k+4}\, x(n-2)\, x(n-k-2)$$

$$+ 3a^2 R_{n-1}(k)$$

$$- 3a^4 R_{n-2}(k)$$

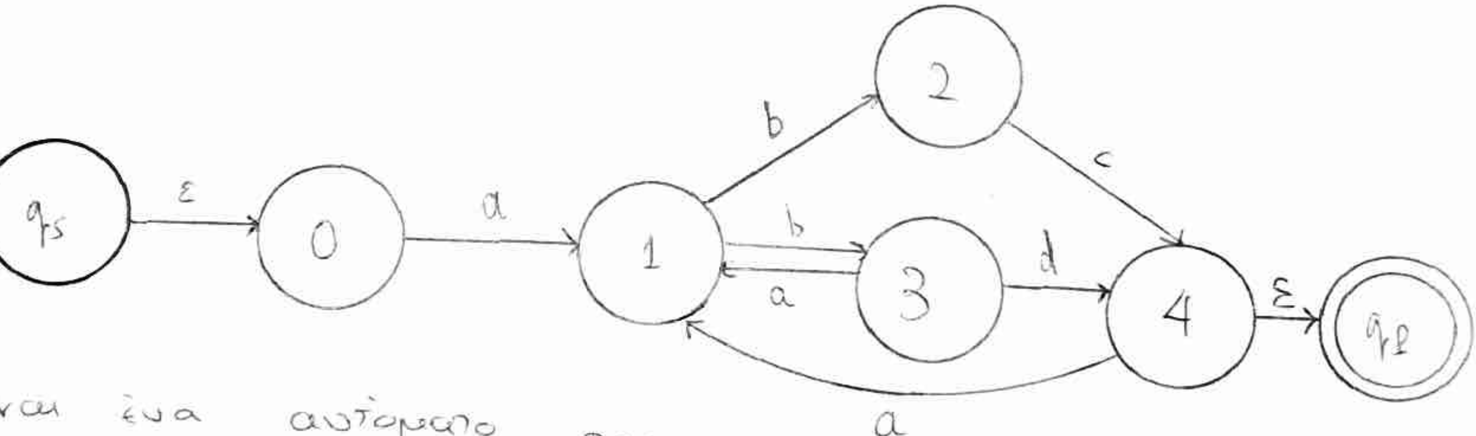
$$+ a^6 R_{n-3}(k)$$

Άσκηση 4



1. Κανονική Έκφραση

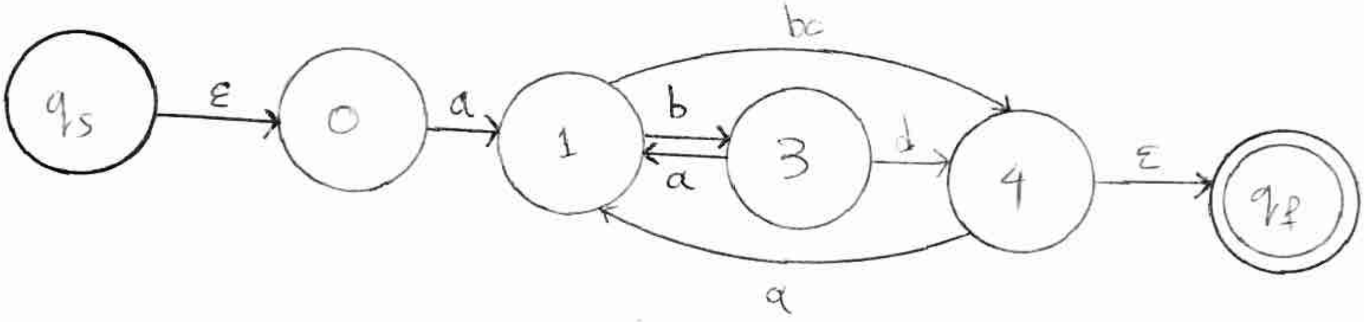
Για την εύρεση της κανονικής έκφρασης της παραπάνω μηχανής θα φτιάξω το παρακάτω γράφημα:



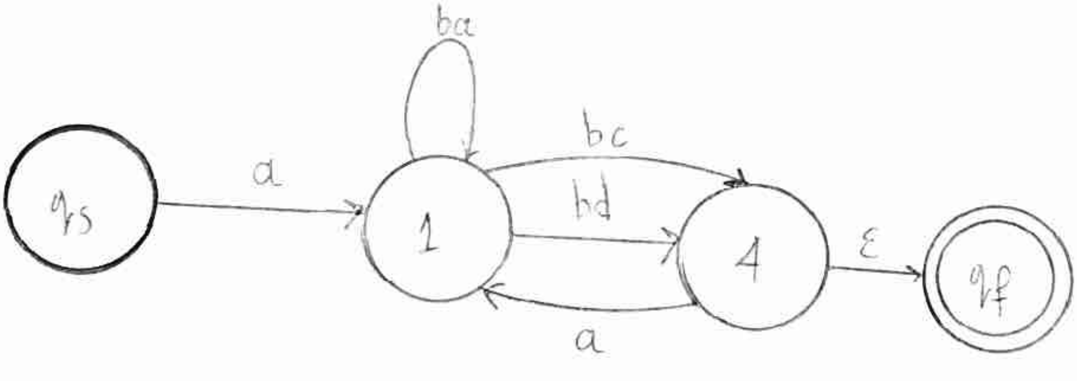
Είναι ένα αυτόματο που αντιστοιχεί στο δοθέν έχοντας προσθέσει δύο νέες καταστάσεις με ε-μετάβαση όπως φαίνεται. Κατά σύμβαση, θα σημειώνω πάνω από κάθε μετάβαση μόνο το σύμβολο/συμβολοσειρά εισόδου καθώς στο δοθέν αυτόματο είναι ίδιο με της εξόδου σε κάθε μετάβαση. Ακόμη, για την εύρεση κανονικής έκφρασης αδιαφορούμε για τα βάρη των μεταβάσεων.

Για την εύρεση της κανονικής έκφρασης αρχικά πρέπει να αφαιρέσουμε τα junk states. Ωστόσο, δε φαίνεται να υπάρχουν junk states. Αφαιρώ τώρα μία μία τις καταστάσεις μέχρι να μείνουν οι q_s, q_f.

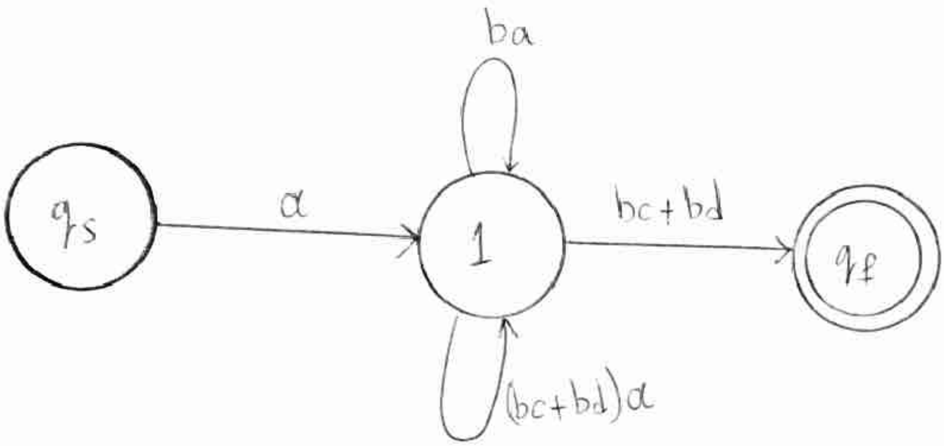
• Αφαιρώ την κατάσταση 2 και παίρνω το εξής ισοδύναμο:



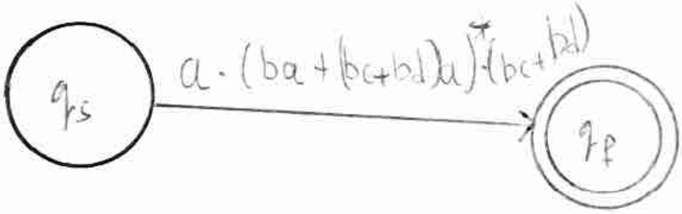
• Αφαιρώ την κατάσταση 0 και την 3 και παίρνω το εξής ισοδύναμο:



• Αφαιρώ την κατάσταση 4 και παίρνω το εξής ισοδύναμο:



• Αφαιρώ και την κατάσταση 1 και παίρνω το τελικό ισοδύναμο:



Άρα η κανονική έκφραση είναι:

$$RE = a\cdot\left(ba+(bc+bd)a\right)^*\cdot(bc+bd) \Rightarrow$$

$$\boxed{RE = a\left(b(a+(c+d)a)\right)^* b\,(c+d)}$$

2. Έστω αυτόματο $A = (Q, \Sigma, \delta, q_0, F)$

$Q = \{0, 1, 2, 3, 4\}$: καταστάσεις

$\Sigma = \{\alpha, b, c, d\}$: αλφάβητο

δ : συνάρτηση που δέχεται είσοδο μια κατάσταση και ένα σύμβολο μετάβασης και επιστρέφει την κατάσταση στην οποία μεταβαίνει και το αντίστοιχο βάρος :

$$\delta(\underbrace{0}_{\text{κατάσταση}}, \underbrace{a}_{\text{σύμβολο}}) = (\underbrace{1}_{\text{κατάσταση}}, \overbrace{2}^{\text{βάρος}})$$

$$\delta(1, b) = (2, 1.5)$$

$$\delta(1, b) = (3, 1)$$

$$\delta(2, c) = (4, 2.5)$$

$$\delta(3, a) = (1, 1)$$

$$\delta(3, d) = (4, 3)$$

$$\delta(4, a) = (1, 1)$$

$q_0 = 0$

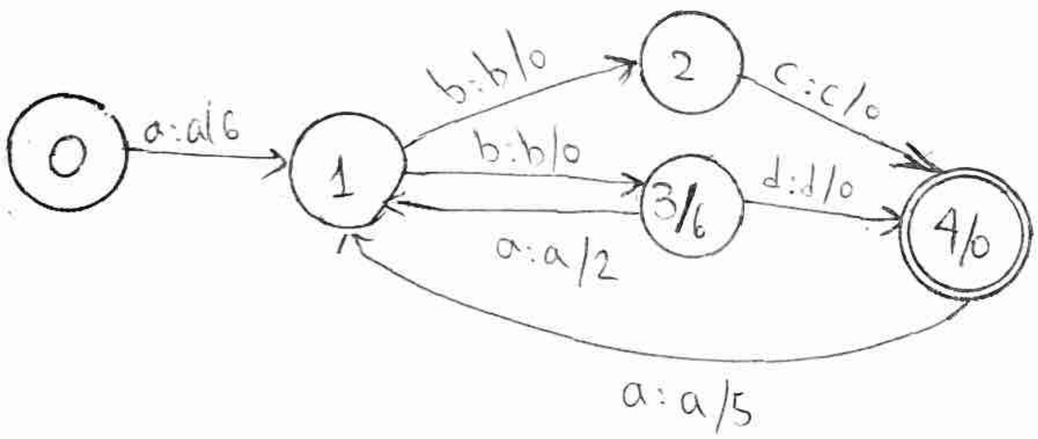
$F = \{4\}$

Έχω $cost(3) = 6$ και $cost(4) = 0$

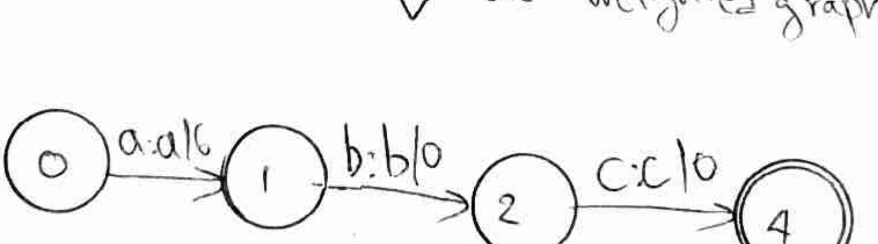
Λαμβάνοντας υπόψη το tropical-semiring $(\mathbb{R}_+, \min, +, \infty, 0)$

Θα αναζητήσω το πιο πιθανό μονοπάτι στο wFA με αρχική κατάσταση 0 και τελική 4, με τις κινήσεις extend και collect.

Με weighted pushing στο wFA παίρνω για διευκόλυνση το αντίστοιχο ισοδύναμο.



το διαχειρίζομαι σαν single source shortest path σε weighted graph



Άρα, πιθανότερη γραμματοσειρά εφόσον χρησιμοποιούμε τον τροπικό ημιδακτύλιο φαίνεται να είναι η abc με κόστος 6.

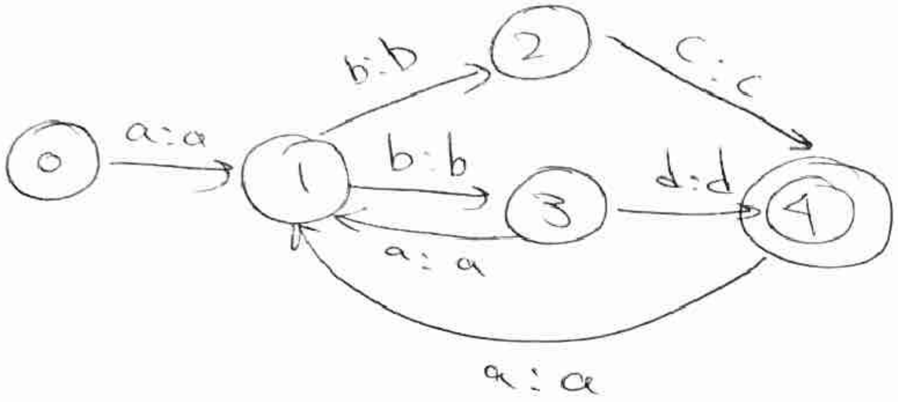
3. $cost(abdababc) = 2+1+3+1+1+1+1{,}5+2{,}5 = 13$

το κόστος καθαρά ακολουθώντας τις τιμές που δίνουν τα συμβολοσειρά.

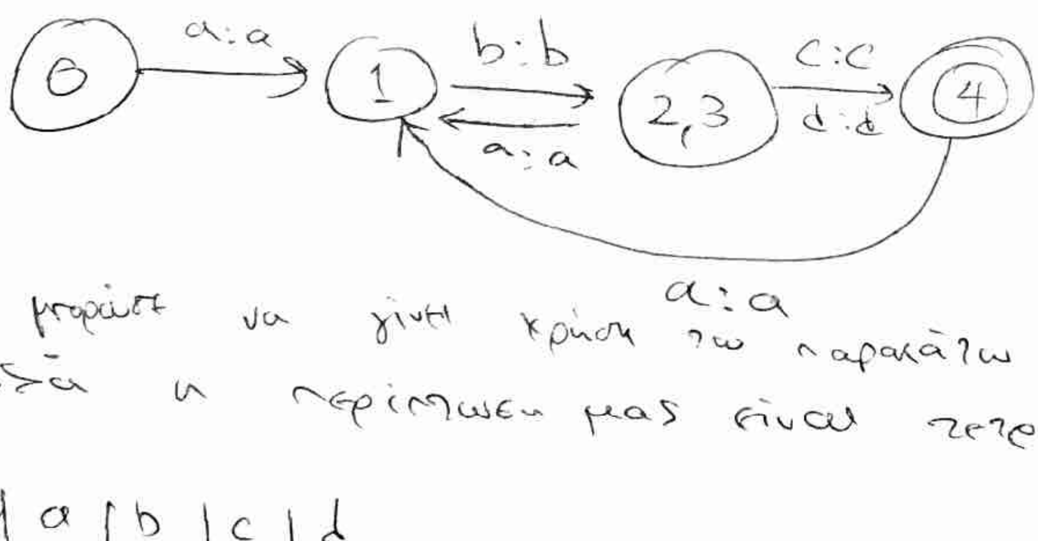
~~Συνολικό κόστος για έλεγχο~~:

- $cost(a) = 2$
- $cost(ab) = cost(a) + 1.5 + 1 = 4.5$
- $cost(abd) = cost(ab) + 3 = 7.5$
- $cost(abda) = cost(abd) + 1 = 8.5$
- $cost(abdab) = cost(abda) + 1 + 1.5 = 11$
- $cost(abdaba) = cost(abdab) + 1 = 12$
- $cost(abdabab) = cost(abdaba) + 1 + 1.5 = 14.5$
- $cost(abdababc) = cost(abdabab) + 2.5 = 17$

4. Αφαιρούμε όλα τα βάρη από το γράφημα:



Για να κατασκευαστεί το αντίστοιχο DFA παρατηρούμε πολύ εύκολα ότι αρκεί να συγχωνεύσουμε τις καταστάσεις 2 και 3 σε μια την 2,3 ως εξής:



Θα μπορούσε να γίνει χρήση του παρακάτω πίνακα αλλά η περίπτωση μας είναι τετριμμένη.

	a	b	c	d
0	1	–	–	–
1	–	2,3	–	–
2	–	–	4	–
3	1	–	–	4
4	1	–	–	–

5. Μπορούμε να φτιάξουμε την εξής ντετερμινιστική γραμματική για το αυτόματο:

$S \to abAC$

$A \to Bd / c$

$B \to \alpha b B / \varepsilon$

$C \to abD$

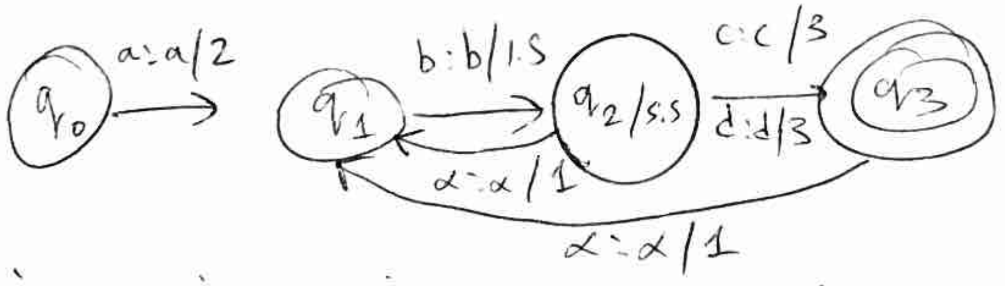
$D \to ED / cD / \varepsilon$

$E \to \alpha b E / \varepsilon$

Αφού η γραμματική μας δεν είναι διφορούμενη, ούτε το αυτόματο μας είναι.

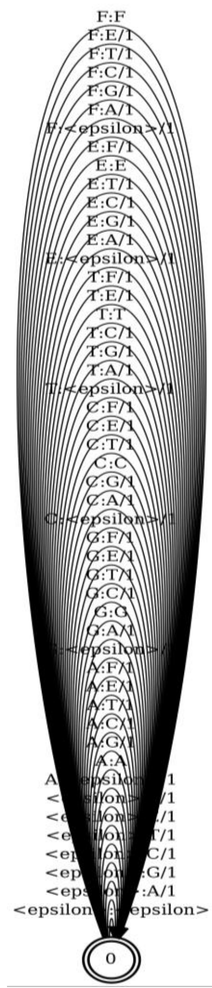
Επίσης το αυτόματο είναι προσβάσιμο σε όλες τις καταστάσεις ξεκινώντας από την αρχική κατάσταση 0.

Το ζητούμενο ντετερμινιστικό αυτόματο με βάση είναι:

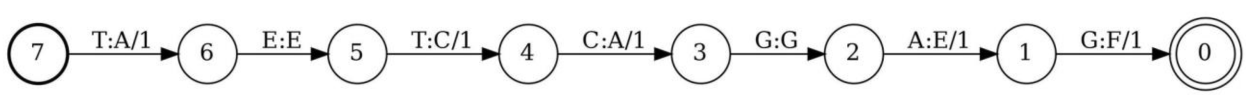


Άσκηση 5

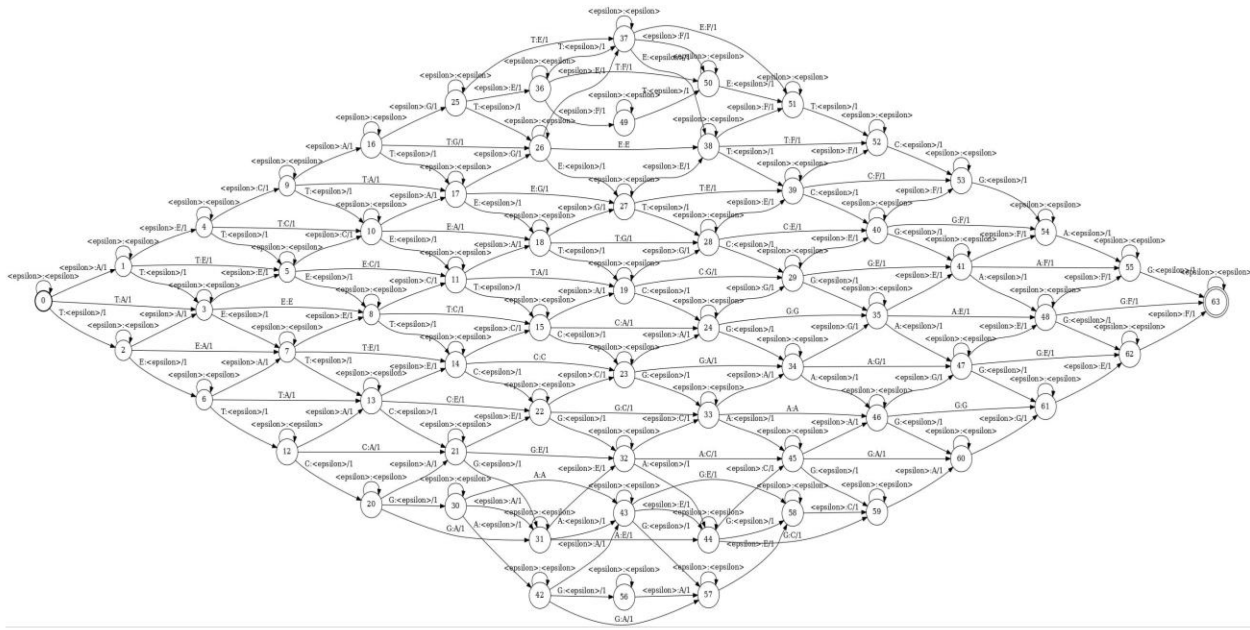
1. Παρακάτω φαίνεται ο μετατροπέας με τις ζητούμενες απαιτήσεις:



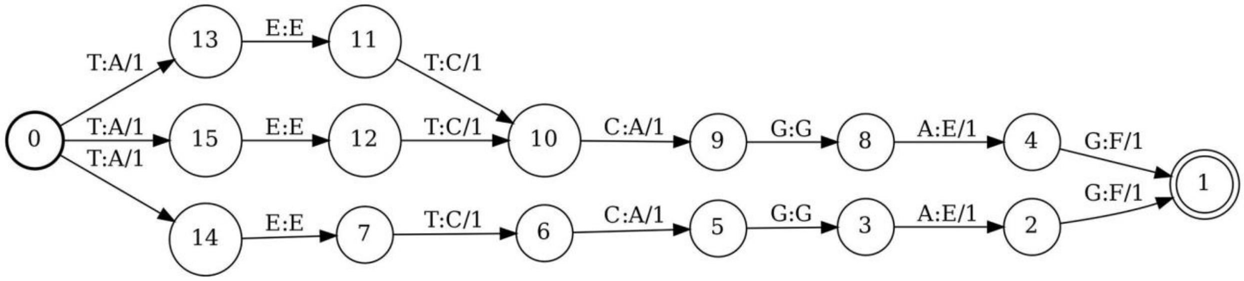
2. Η καλύτερη αντιστοίχιση ανάμεσα στις γραμματοσειρές AECAGEF και TETCGAG φαίνεται στο παρακάτω γράφημα:



Η παραπάνω αντιστοίχιση δίνει Levenshtein distance = 5 , όπως παρατηρεί κανείς.
Η λογική είναι η εξής:
Καλούμαστε να υπολογίσουμε edit distance μεταξύ δυο συμβολοσειρών, με όλα τα edits μοναδιαίου κόστους. Επομένως δεν έχει σημασία από τις δύο θα επιχειρήσουμε να αλλάξουμε ως προς την άλλη καθώς πρόκειται για το ίδιο ακριβώς πρόβλημα. Θέτουμε την μία ως accepted (έστω ένα FSA με μόνο αυτή τη λέξη) και δίνουμε την άλλη ως input. Φτιάχνουμε επίσης έναν μετατροπέα για τα edits και κάνουμε compose τα δύο αυτόματα και στη συνέχεια παίρνω το shortest path και έτσι προκύπτει το παραπάνω. Το FST που προκύπτει από το compose φαίνεται παρακάτω:



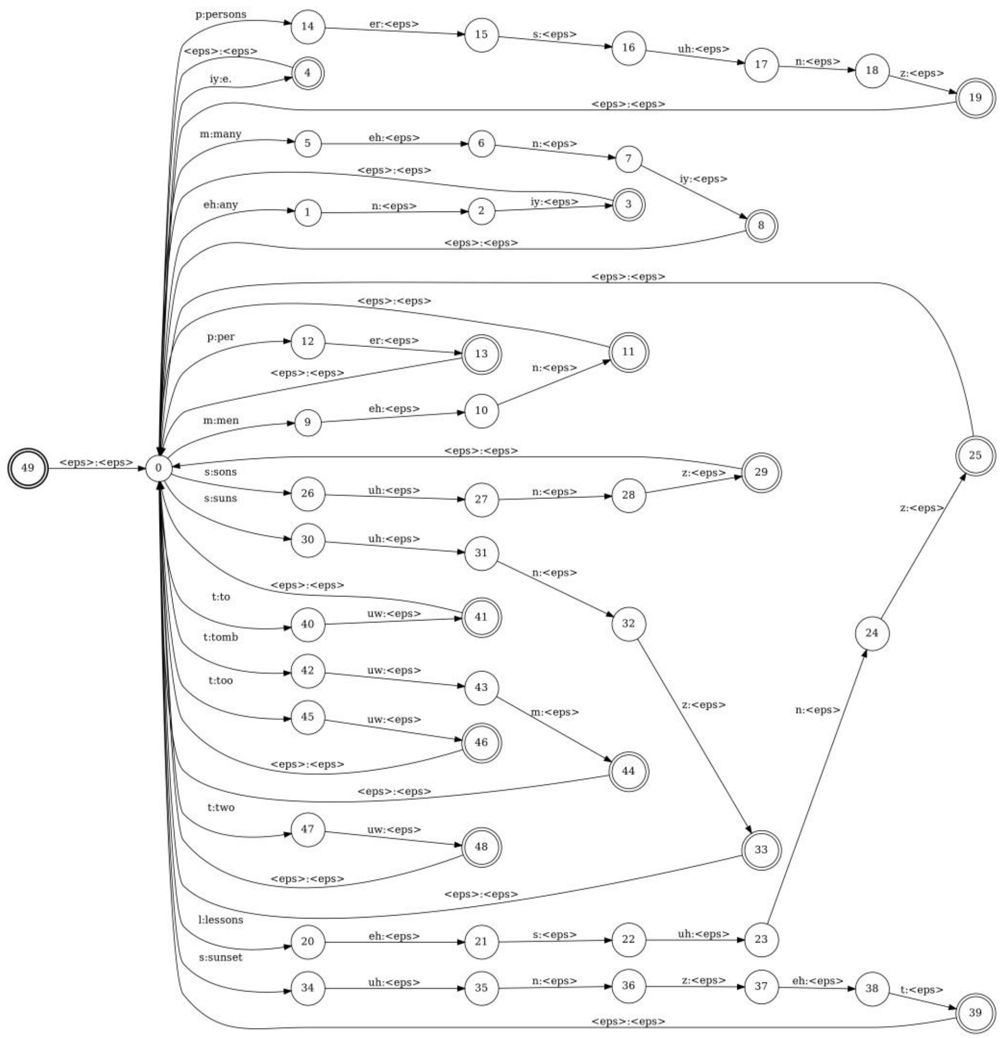
3. Η δεύτερη καλύτερη αντιστοιχία προκύπτει όπως φαίνεται παρακάτω:



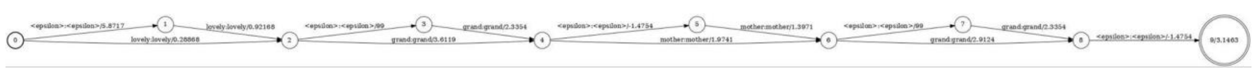
Είδαμε πριν ότι το min distance ήταν 5 όποτε δοκιμάσαμε με shortest path με -nshortest=6, και προέκυψαν οι παραπάνω αντιστοιχίσεις μεταξύ των δύο συμβολοσειρών.

<u>Άσκηση 6</u>

1. Με χρήση της fstclosure παίρνουμε το παρακάτω αποτέλεσμα για την αντιστοίχιση σειρών φωνημάτων σε λέξεις:



2.

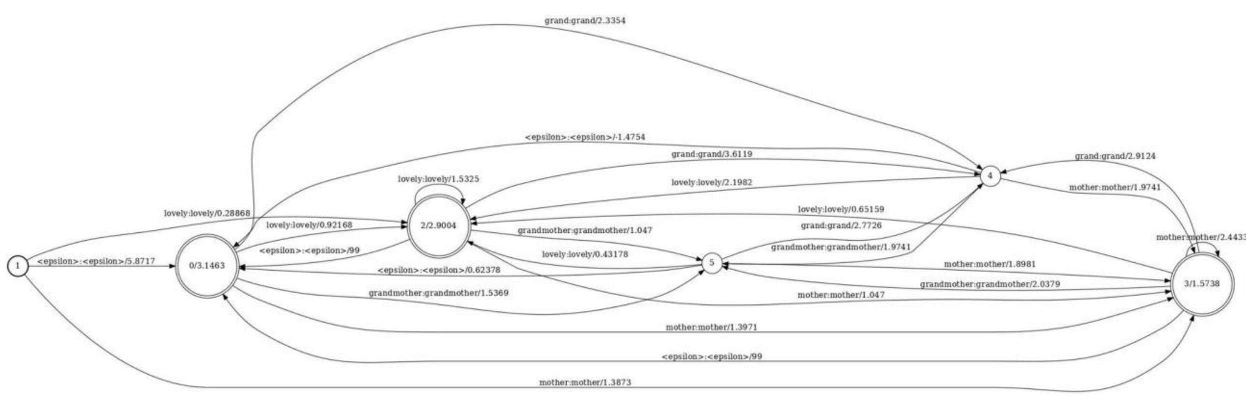


<u>Άσκηση 7</u>

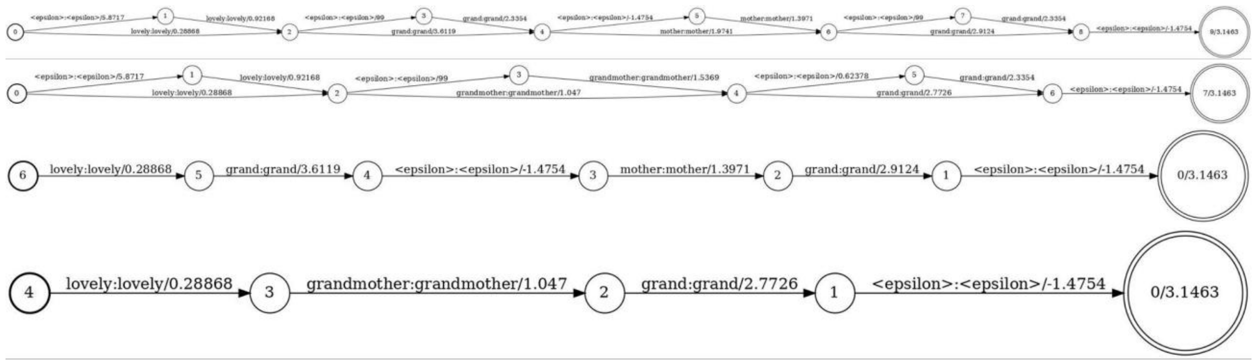
1. Παρακάτω παρουσιάζονται τα βάρη του αντίστοιχου bigram γλωσσικού μοντέλου:

```
\2-grams:
-0.1253732      <s> lovely
-0.6024945      <s> mother
-1.259638       lovely </s>
-0.6655462      lovely lovely
-0.4546929      lovely mother
-1.568636       lovely grand
-0.4546929      lovely grandmother
-0.6834827      mother </s>
-0.2829811      mother lovely
-1.061132       mother mother
-1.264818       mother grand
-0.885041       mother grandmother
-0.954677       grand lovely
-0.8573325      grand mother
-0.8573325      grand grandmother
-0.1875212      grandmother lovely
-0.8243433      grandmother mother
-1.20412        grandmother grand
```

Καθώς και η αναπαράσταση του σε FST:



2.



<epsilon>:<epsilon>/5.8717
lovely:lovely/0.92168
lovely:lovely/0.28868
<epsilon>:<epsilon>/99
grand:grand/2.3354
grand:grand/3.6119
<epsilon>:<epsilon>/-1.4754
mother:mother/1.3971
mother:mother/1.9741
<epsilon>:<epsilon>/99
grand:grand/2.3354
grand:grand/2.9124
<epsilon>:<epsilon>/-1.4754
9/3.1463
<epsilon>:<epsilon>/5.8717
lovely:lovely/0.92168
lovely:lovely/0.28868
<epsilon>:<epsilon>/99
grandmother:grandmother/1.5369
grandmother:grandmother/1.047
<epsilon>:<epsilon>/0.62378
grand:grand/2.3354
grand:grand/2.7726
<epsilon>:<epsilon>/-1.4754
7/3.1463
lovely:lovely/0.28868
grand:grand/3.6119
<epsilon>:<epsilon>/-1.4754
mother:mother/1.3971
grand:grand/2.9124
<epsilon>:<epsilon>/-1.4754
0/3.1463
lovely:lovely/0.28868
grandmother:grandmother/1.047
grand:grand/2.7726
<epsilon>:<epsilon>/-1.4754
0/3.1463