Εφαρμογές Κινητών, Edge Υπολογιστική και Μελλοντικά Δίκτυα

Εφαρμογές Κινητών, Edge Υπολογιστική και Μελλοντικά Δίκτυα

- Η τεχνολογία Mobile Edge Υπολογιστική (MEC) καθορίζει μία καινοτόμα αρχιτεκτονική δικτύου,
  - όπου οι υπηρεσίες υπολογιστικού νέφους παρέχονται από το άκρο του δικτύου, δηλαδή από έξυπνους σταθμούς βάσης και σημεία πρόσβασης στο δίκτυο.
- Το άκρο του δικτύου είναι το τμήμα που είναι πιο κοντά στον τελικό χρήστη.
- Μεταφέροντας τις υπηρεσίες σε αυτό, περιορίζονται οι καθυστερήσεις (delays),
  - καθώς μειώνεται η απόσταση μεταξύ του χρήστη και του τόπου από όπου παρέχονται οι υπηρεσίες

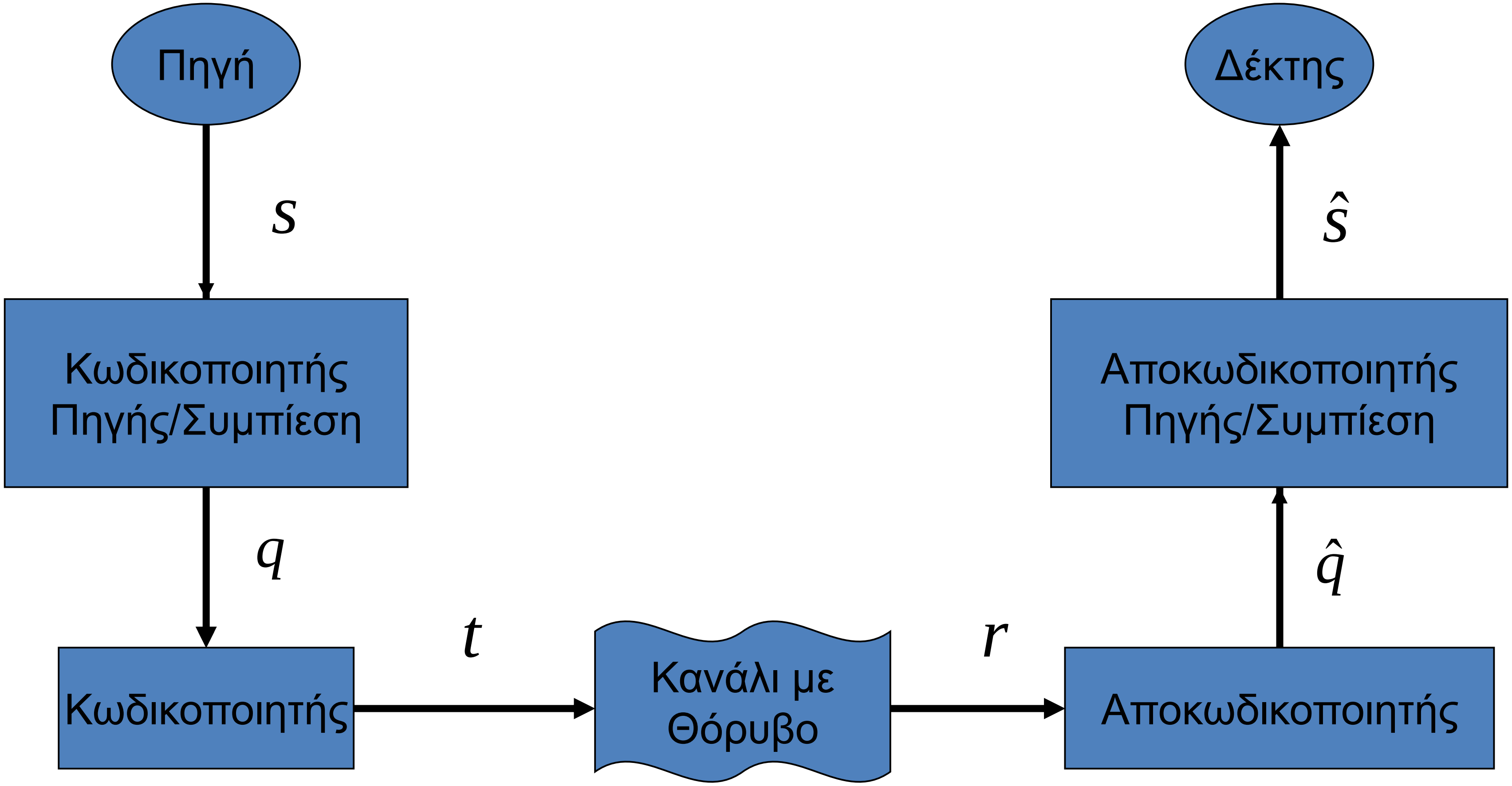
Εφαρμογές Κινητών, Edge Υπολογιστική και Μελλοντικά Δίκτυα

- Η τεχνολογία Mobile Edge Υπολογιστική (MEC) καθορίζει μία καινοτόμα αρχιτεκτονική δικτύου,
  - όπου οι υπηρεσίες υπολογιστικού νέφους παρέχονται από το άκρο του δικτύου, δηλαδή από έξυπνους σταθμούς βάσης και σημεία πρόσβασης στο δίκτυο.
- Το άκρο του δικτύου είναι το τμήμα που είναι πιο κοντά στον τελικό χρήστη.
- Μεταφέροντας τις υπηρεσίες σε αυτό, περιορίζονται οι καθυστερήσεις (delays),
  - καθώς μειώνεται η απόσταση μεταξύ του χρήστη και του τόπου από όπου παρέχονται οι υπηρεσίες

Εφαρμογές Κινητών, Edge Υπολογιστική και Μελλοντικά Δίκτυα

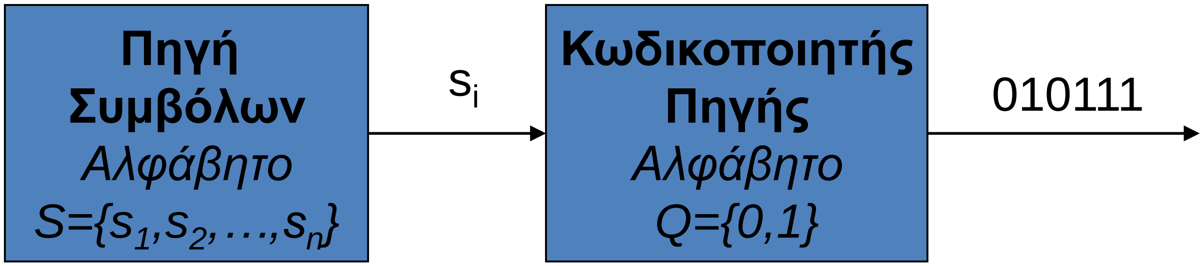
- Επίσης, αναλύονται ζητήματα
  - σχεδιασμού και ανάπτυξης Ασύρματων Δικτύων Τέταρτης και Πέμπτης Γενιάς (4G-5G wireless networks),
  - Πυκνών Δικτύων (Ultra Dense Networks - UDN),
  - Ασύρματων Δικτύων Αισθητήρων (Wireless Sensor Networks - WSN),
  - Δικτύων Καθοριζόμενων από το Λογισμικό (Software Defined Networks - SDN) και Οχηματικών Δικτύων (Vehicular Networks) για περιβάλλοντα έξυπνων πόλεων.

Κωδικοποίηση Πηγής (1)



Κωδικοποίηση Πηγής (2)

- Κωδικοποίηση/συμπίεση της πηγής
  - Είναι η διαδικασία αντιστοίχισης του αλφάβητου των συμβόλων σε ένα άλλο αλφάβητο.
  - Το καινούριο αυτό αλφάβητο ονομάζεται κωδικό αλφάβητο και τα μέλη ονομάζονται κωδικά σύμβολα.
  - ΟΙ ακολουθίες των κωδικών συμβόλων που αντιστοιχούν σε σύμβολα της πηγής λέγονται κωδικές λέξεις



Κωδικοποίηση Πηγής (3)

- Ερωτήματα που προκύπτουν από τις σχέσεις που διέπουν τα σύμβολα και τους κώδικες και τις κωδικές λέξεις
  - Πως μπορούμε να κατασκευάσουμε άμεσους κώδικες με το ελάχιστο προσδοκώμενο μήκος για την συμπιεσμένη αναπαράσταση των συμβόλων μιας πηγής;
  - Υπάρχει σχέση μεταξύ συμβόλου και μήκους κωδικής λέξης;
  - Το πληροφοριακό περιεχόμενο της πηγής των συμβόλων πως επηρεάζει την παραγωγή των κωδικών λέξεων και με ποιό τρόπο;

Κωδικοποίηση Πηγής (4)

- Απαιτήσεις για χρησιμότητα κωδικών
  - Κάθε ακολουθία κωδικών λέξεων πρέπει να μπορεί να αποκωδικοποιηθεί με μοναδικό τρόπο
  - Η αποκωδικοποίηση πρέπει να γίνεται εύκολα και άμεσα
  - Ο κώδικας πρέπει να πετυχαίνει τη βέλτιστη δυνατή συμπίεση

Κωδικοποίηση Πηγής (5)

- **Ορισμοί**
  - **Μη ιδιάζων κώδικας**
    - Όταν όλες οι κωδικές λέξεις είναι διαφορετικές
  - **Μοναδικά αποκωδικοποιήσιμος**
    - Όταν και οι ακολουθίες των κωδικών λέξεων είναι διαφορετικές
  - **Άμεσος ή Προθεματικός κώδικας**
    - Κάθε μοναδικά αποκωδικοποίησιμος κώδικας που επιτρέπει την άμεση αποκωδικοποίηση της κωδικής λέξης χωρίς να χρειάζεται να λάβει υπόψη του τις επόμενες κωδικές λέξεις.
    - Ο άμεσος κώδικας αποτελείται από κωδικές λέξεις οι οποίες δεν αποτελούν μέρος (προθέματα άλλων)

Κωδικοποίηση Πηγής (6)

- **Παράδειγμα**
  - Μη ιδιάζων, I,II,III,IV
  - Μοναδικά αποκωδικοποιήσιμος, II,III,IV. Ο I δεν είναι αφού ΦΦΦΦ, ΦΦΨ, ΨΨ όλα έχουν κωδική λέξη την ίδια, 0000
  - Άμεσοι κώδικες, II και III
  - Ο κώδικας IV δεν είναι άμεσος αφού χρειάζεται να γνωρίζουμε ψηφία που ανήκουν στην επόμενη κωδική λέξη, π.χ. 01101110 0?

	I	**II**	**III**	**IV**
Φ	0	00	0	0
Χ	11	01	10	01
Ψ	00	10	110	011
Ω	01	11	1110	0111

Κωδικοποίηση Πηγής (11)

- **<u>Θεώρημα 12:</u>** Ανισότητα του Kraft
  - Για κάθε άμεσο κώδικα με πλήθος κωδικών συμβόλων q του κωδικού αλφαβήτου Q και μήκη των κωδικών λέξεων l_i, όπου i=1,2,…,n και n το πλήθος των συμβόλων της πηγής ισχύει,
    - $$\boxed{\sum_{i=1}^{n} q^{-l_i} \leq 1}$$
  - Αντίστροφα, αν για ένα σύνολο μηκών κωδικών λέξεων ισχύει η ανισότητα Kraft τότε υπάρχει ένας άμεσος κώδικας του οποίου οι κωδικές λέξεις έχουν αυτά τα μήκη.

Κωδικοποίηση Πηγής (16)

- Με τις πιθανότητες εμφάνισης των συμβόλων τι γίνεται;
- Μπορούμε να βρούμε ένα κώδικα (άμεσο και αποκωδικοποιήσιμο) του οποίου οι κωδικές λέξεις να έχουν το βέλτιστο δυνατό μήκος; Να έχουν δηλαδή τη κατά μέσο όρο τη μικρότερη τιμή μήκους κωδικής λέξης;
  - $\min \sum_i p_i l_i$
- Υπάρχει σχέση μήκους λέξης και πιθανότητας εμφάνισης συμβόλου πηγής;
- Αν αντιστοιχίσουμε κωδικές λέξεις μικρού μήκους σε σύμβολα με μεγάλη πιθανότητα εμφάνισης θα μειωθεί το μέσο μήκος της κωδικής λέξης;
- Ξέρουμε όμως ότι μικρού μήκους κωδικές λέξεις έχουν μεγάλο κόστος αφού χρειάζεται να εισάγουμε μεγάλου μήκους κωδικές λέξεις για την πλήρη αντιστοίχιση των συμβόλων πηγής.
- **Ποια είναι η βέλτιστη συμπίεση που είναι δυνατόν να επιτευχθεί;**

Κωδικοποίηση Πηγής (17)

- **Θεώρημα 13:** *Κωδικοποίησης Πηγής*
  - Έστω μια πηγή παράγει $S=\{s_1,s_2,\ldots,s_n\}$ σύμβολα με πιθανότητα εμφάνισης κάθε συμβόλου $\{p_1,p_2,\ldots,p_n\}$. Τα σύμβολα αυτά κωδικοποιούνται από ένα κωδικό αλφάβητο q συμβόλων και αντιστοιχίζονται σε άμεσο και αποκωδικοποιήσιμο κώδικα n κωδικών λέξεων μήκους l_i η κάθε μία, $i=1,2,\ldots,n$. Αν H(C) είναι το μέσο πληροφοριακό περιεχόμενο των συμβόλων της πηγής τότε ισχύει,
    - $$\boxed{\frac{H(C)}{\log q} \le \sum_{i=1}^{n} p_i l_i}$$
- Για q=2, το βέλτιστο (ελάχιστο) **μέσο** μήκος κωδικής λέξης είναι ίσο με το μέσο πληροφοριακό περιεχόμενο της πηγής των συμβόλων και δεν μπορεί να είναι μικρότερο από αυτό.
- Για ποιές τιμές των l_i ισχύει η ισότητα του θεωρήματος της κωδικοποίησης της πηγής;
- **Ελάχιστο του μήκους των κωδικών λέξεων**
  - Η ισότητα ισχύει όταν

$$\boxed{l_i = \log_q\left(\frac{1}{p_i}\right)}$$

Κωδικοποίηση Πηγής (18)

- Βέλτιστες κωδικές λέξεις για q=2
  - Η μέση τιμή του μήκους των λέξεων ενός κώδικα ελαχιστοποιείται και είναι ίση με την εντροπία της πηγής όταν τα μήκη της κάθε κωδικής λέξης είναι ίσα με το πληροφοριακό περιεχόμενο του κάθε συμβόλου
  - $\boxed{l_i = \log_2\left(\frac{1}{p_i}\right)}$
- Αντίστροφα, οποιαδήποτε επιλογή κώδικα και κωδικών λέξεων έμμεσα ορίζει και μία κατανομή πιθανότητας $\{q_i\}$,
  - $\boxed{q_i = 2^{-l_i}/z}$,z=1 όταν ο κώδικας είναι πλήρης

Κωδικοποίηση Πηγής (19)

- Δεν μπορούμε λοιπόν να συμπιέσουμε λιγότερο από την εντροπία της πηγής.
- Πρακτικά πόσο κοντά σε αυτή την τιμή μπορούμε να φτάσουμε;
- **Θεώρημα 14:**
  - Για κάθε τ.μ X υπάρχει ένας άμεσος και μοναδικά αποκωδικοποιήσιμος κώδικας C του οποίου η μέση τιμή μήκους, L(C,X), ικανοποιεί τη σχέση
    - $H(X) \leq L(C,X) < H(X)+1$
  - Ο κώδικας αυτός έχει κωδικές λέξεις μήκους
    - $$l_i^* = \lceil \log_2 (1/p_i) \rceil$$ όπου $\lceil x \rceil$ είναι ο μικρότερος ακέραιος που είναι μεγαλύτερος του χ.

Συμπίεση Πληροφορίας ή Κωδικοποίηση Πηγής ...

- Αλγόριθμοι κωδικοποίησης
  - FANO
  - SHANNON
  - HUFFMAN
- JPEG και MPEG χρησιμοποιούν μεταξύ άλλων και τον αλγόριθμο Huffman

Αλγόριθμος Κωδικοποίησης FANO …

- **<u>Βήμα 1ο</u>**: Τα σύμβολα (ή τα μηνύματα) ταξινομούνται έτσι ώστε οι πιθανότητές τους είναι σε φθίνουσα ακολουθία.
- **<u>Βήμα 2ο</u>**: Στη συνέχεια τα σύμβολα χωρίζονται σε ομάδες ο αριθμός των οποίων είναι ίσος με τον αριθμό των κωδικών συμβόλων (στην περίπτωση δυαδικού κώδικα οι ομάδες χωρισμού συμβόλων είναι δύο).
  - Το κριτήριο σχηματισμού της κάθε ομάδας είναι τέτοιο ώστε αφενός να διατηρείται η σειρά των συμβόλων όπως αυτή έχει καθοριστεί από το βήμα 1
  - Αφετέρου δε να ελαχιστοποιείται η σχέση
  - $$\left|\sum_{i=1}^{k} p_i - \sum_{i=k+1}^{n} p_i\right|$$
- **<u>Βήμα 3ο</u>**: Για κάθε μία ομάδα συμβόλων που δημιουργήσαμε αντιστοιχίζουμε ένα από τα κωδικά σύμβολα ως το πρώτο τον κωδικών λεξεων που θα προκύψουν
- **<u>Βήμα 4ο</u>**: Επαναλαμβάνουμε τα βήματα 2 & 3 για κάθε μία από τις ομάδες προσθέτοντας κάθε φορά και από ένα κωδικό σύμβολο στην κωδική λέξη μέχρι να δημιουργήσουμε ομάδες με ένα μόνο σύμβολο
- Ο αλγόριθμος FANO δημιουργεί κώδικες όπου όλες οι κωδικές λέξεις είναι του ίδιου μήκους εάν η διαίρεση σε υποομάδες γίνεται πάντα με την ίδια πιθανότητα.

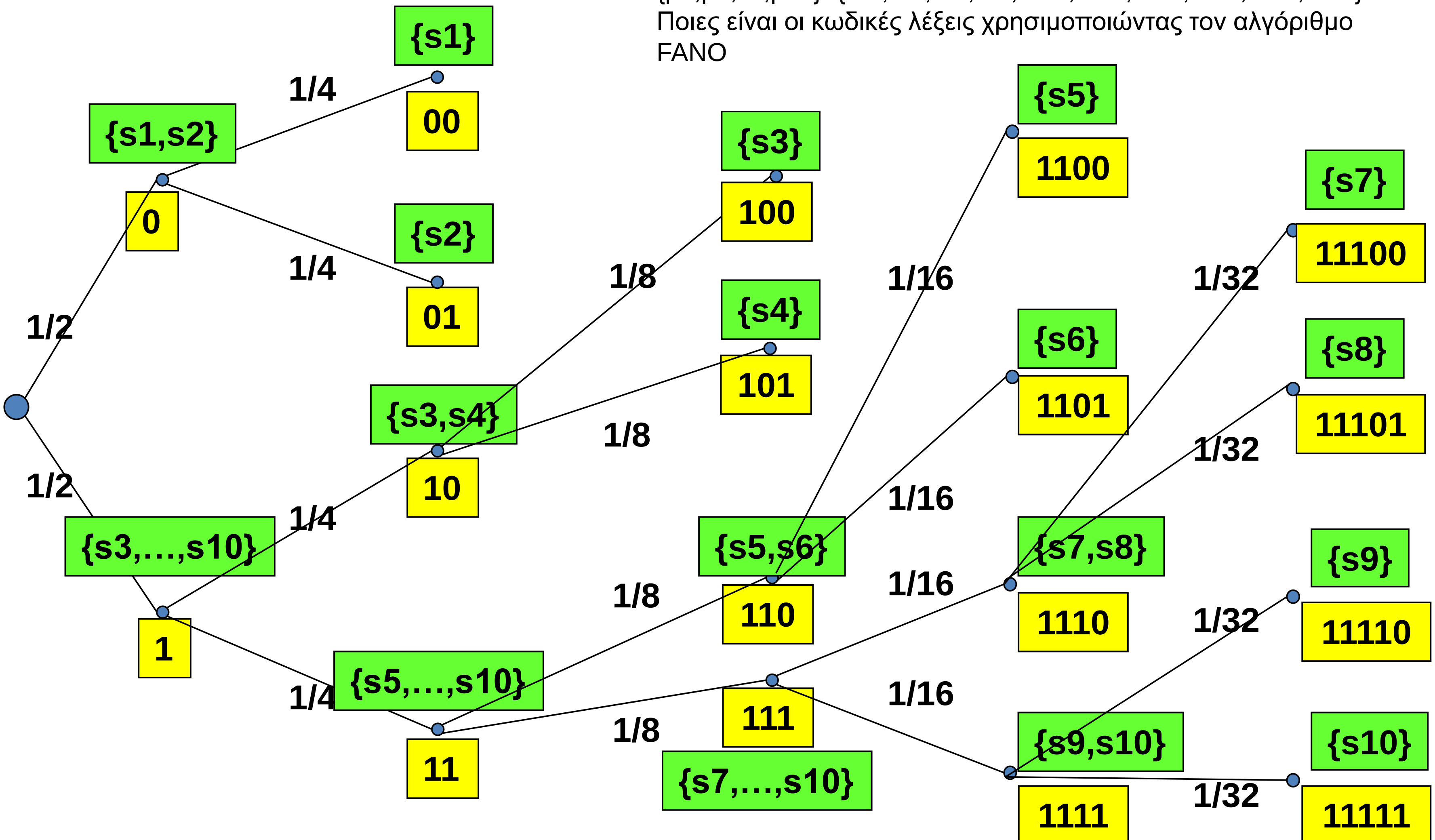
Αλγόριθμος Κωδικοποίησης FANO …

Παράδειγμα

S={s1,s2,…,s10}

{p1,p2,…,p10}={1/4,1/4,1/8,1/8,1/16,1/16,1/32,1/32,1/32,1/32}

Ποιες είναι οι κωδικές λέξεις χρησιμοποιώντας τον αλγόριθμο FANO



Αλγόριθμος κωδικοποίησης SHANNON …

- **<u>Βήμα 1ο</u>**: Τα σύμβολα (ή τα μηνύματα) ταξινομούνται έτσι ώστε οι πιθανότητές τους είναι σε φθίνουσα ακολουθία, όπως ακριβώς και του FANO.
- **<u>Βήμα 2ο</u>**: Για κάθε σύμβολο s_i του οποίου η πιθανότητα εμφάνισης είναι $p(s_i)$ υπολογίζεται η αθροιστική πιθανότητα P_j ως εξής:
  - $$P_i = \sum_{j=1}^{i-1} p(s_j), \quad P_1 = 0, \quad i = 2,\ldots,n$$
- **<u>Βήμα 3ο</u>**: Το μήκος της κωδικής λέξης που αντιστοιχεί στο σύμβολο s_i ισούται με τον ακέραιο αριθμό l_i, που πληροί τη σχέση
  - $$l_i = \lceil \log_2 \left(1/p\left(s_i\right)\right) \rceil$$
- **<u>Βήμα 4ο</u>**: Η κωδική λέξη c_i που αντιστοιχεί στο σύμβολο πηγής s_i είναι το δυαδικό ανάπτυγμα του κλάσματος P_i (μόνο τα πρώτα l_i bits αναπτύγματος)

Αλγόριθμος κωδικοποίησης SHANNON …

Παράδειγμα

Σύμβολα Πηγής	Πιθανότητες Συμβόλων	P_i	Μήκος l_i	Ανάπτυγμα του P_i	Κωδικές Λέξεις
S_1	1/4	$P_1 = 0$	$l_1 = 2$	.00000	00
S_2	1/4	$P_2 = 1/4$	$l_2 = 2$	.01000	01
S_3	1/8	$P_3 = 1/2$	$l_3 = 3$	.10000	100
S_4	1/8	$P_4 = 5/8$	$l_4 = 3$	.10100	101
S_5	1/16	$P_5 = 3/4$	$l_5 = 4$	.11000	1100
S_6	1/16	$P_6 = 13/16$	$l_6 = 4$	.11010	1101
S_7	1/32	$P_7 = 7/8$	$l_7 = 5$	.11100	11100
S_8	1/32	$P_8 = 29/32$	$l_8 = 5$	.11101	11101
S_9	1/32	$P_9 = 15/16$	$l_9 = 5$	.11110	11110
S_{10}	1/32	$P_{10} = 31/32$	$l_{10} = 5$	.11111	11111

Κωδικοποίηση HUFFMAN (1)

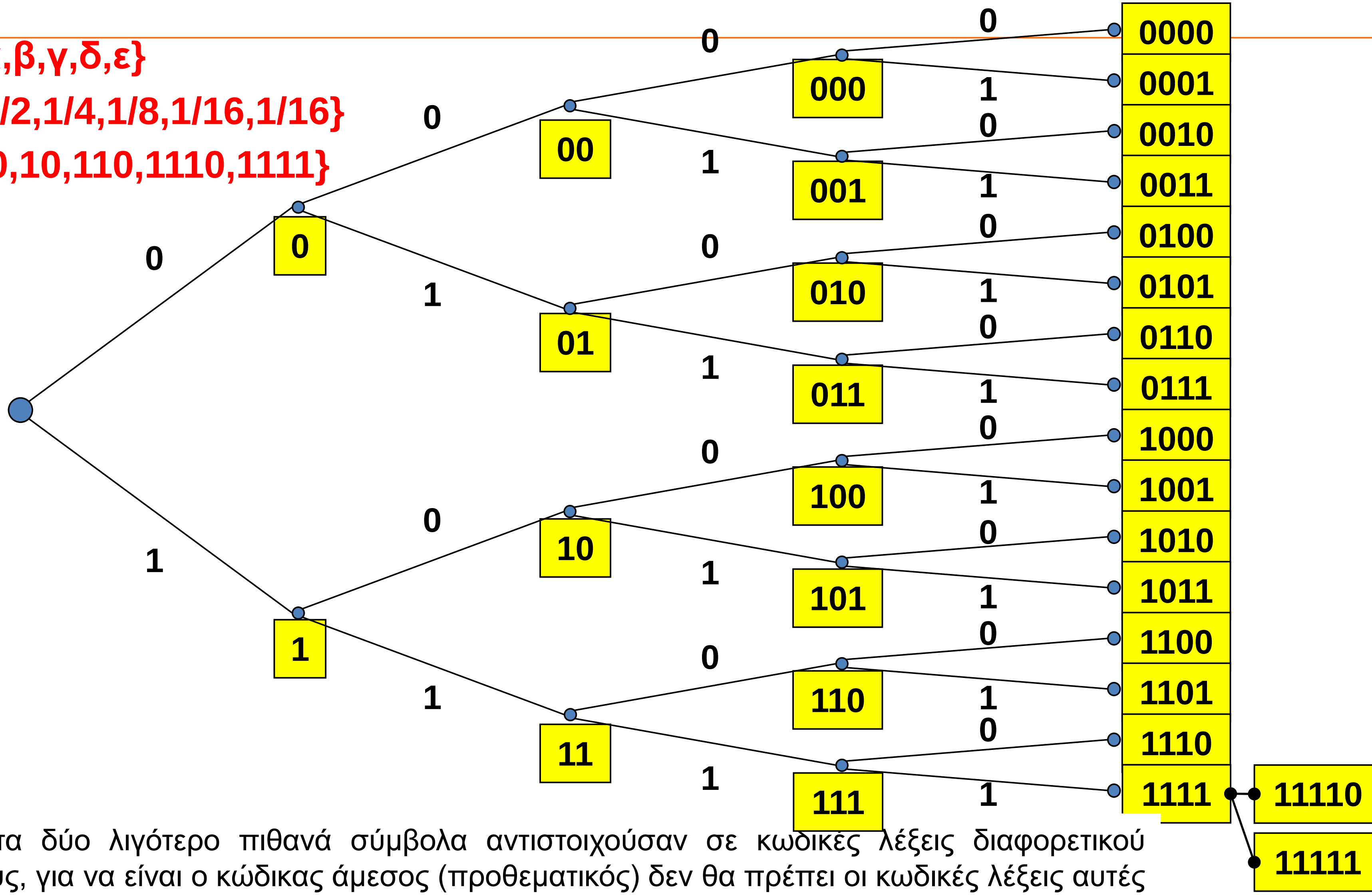
- Δοθείσας μιας πηγής συμβόλων με πιθανότητες εμφάνισης p_i, πως μπορούμε να σχεδιάσουμε ένα βέλτιστο άμεσο κώδικα;
- Με τον όρο βέλτιστο εννοούμε ελαχιστοποίηση του μέσου μήκους των κωδικών λέξεων ακεραίου μήκους.
- Η διαδικασία Huffman βασίζεται σε δύο παρατηρήσεις που έχουν σχέση με βέλτιστους κώδικες:
  - Σε έναν βέλτιστο κώδικα τα σύμβολα που εμφανίζονται συχνότερα (μεγαλύτερη πιθανότητα εμφάνισης) θα πρέπει να αντιστοιχούν σε μικρότερες κωδικές λέξεις απ’ ότι σύμβολα με μικρότερη συχνότητα εμφάνισης
  - Σε ένα βέλτιστο κώδικα τα δύο σύμβολα με τη μικρότερη πιθανότητα εμφάνισης αντιστοιχούν σε κωδικές λέξεις ίδιου μήκους

Κωδικοποίηση HUFFMAN (2)

S={α,β,γ,δ,ε}

P={1/2,1/4,1/8,1/16,1/16}

C={0,10,110,1110,1111}



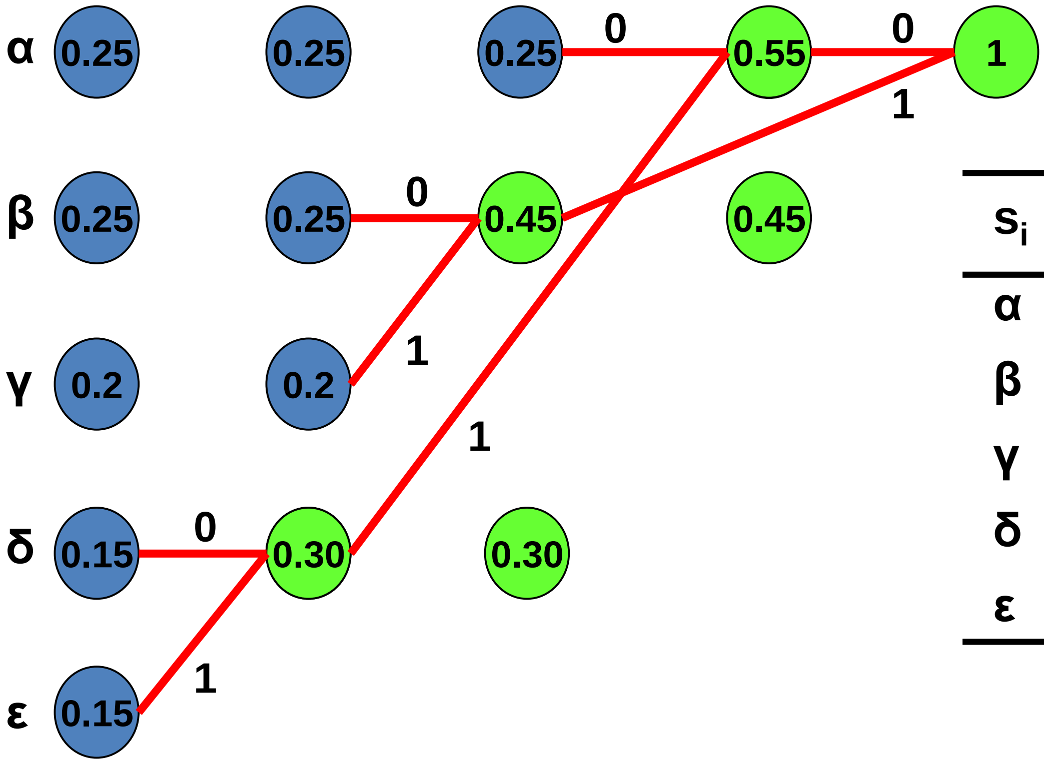
Εάν τα δύο λιγότερο πιθανά σύμβολα αντιστοιχούσαν σε κωδικές λέξεις διαφορετικού μήκους, για να είναι ο κώδικας άμεσος (προθεματικός) δεν θα πρέπει οι κωδικές λέξεις αυτές να αποτελούν προθέματα άλλης κωδικής λέξης ούτε και μεταξύ τους. Άρα αν η μία είναι μεγαλύτερη από την άλλη κατά κ bits καταργώντας τα η λέξη που παίρνουμε είναι επίσης κωδική λέξη και μάλιστα ίδιου μήκους. Διαφέρουν δε μόνο ως προς το τελευταίο bit (0 ή 1)

Κωδικοποίηση HUFFMAN (3)

- **Αλγόριθμος κωδικοποίησης HUFFMAN**
  - Ο αλγόριθμος Huffman κατασκευάζει το δυαδικό δέντρο αρχίζοντας από τα φύλλα του και προχωράει προς τη ρίζα του.
  - **Βήμα 1ο**: Τα σύμβολα (ή τα μηνύματα) ταξινομούνται έτσι ώστε οι πιθανότητές τους είναι σε φθίνουσα ακολουθία.
  - **Βήμα 2ο**: Στη συνέχεια παίρνουμε τα δύο σύμβολα με τις μικρότερες πιθανότητες. Γι’ αυτά μέσα από την διαδικασία του αλγορίθμου θα αναθέσουμε, τις μακρύτερες δυνατές κωδικές λέξεις έτσι ώστε αυτές να έχουν το ίδιο μήκος και να διαφέρουν στο τελευταίο τους ψηφίο.
    - Το βήμα αυτό θα δημιουργήσει το τελευταίο από τα ψηφία της κωδικής λέξης
  - **Βήμα 3ο**: Συνδυάζοντας τα δύο σύμβολα που επιλέξαμε στο βήμα 2 σε ένα και αναθέτοντας στο συνδυασμένο σύμβολο το άθροισμα των πιθανοτήτων των επιμέρους συμβόλων επαναλαμβάνουμε τη διαδικασία από το βήμα 1 μεταξύ των συμβόλων που απομένουν και του συμβόλου που δημιουργήσαμε μέχρις ότου καταλήξουμε σε ένα σύμβολο με πιθανότητα 1.
  - **Βήμα 4ο**: Οι κωδικές λέξεις που αντιστοιχούν στο κάθε σύμβολο αποτελούνται από τις ακολουθίες 0 και 1 που δημιουργούνται αν διατρέξουμε το δένδρο που δημιουργήθηκε από τον κόμβο με το μοναδικό σύμβολο προς τα σύμβολα από τα οποία ξεκινήσαμε

Κωδικοποίηση HUFFMAN (4)

- **Παράδειγμα**

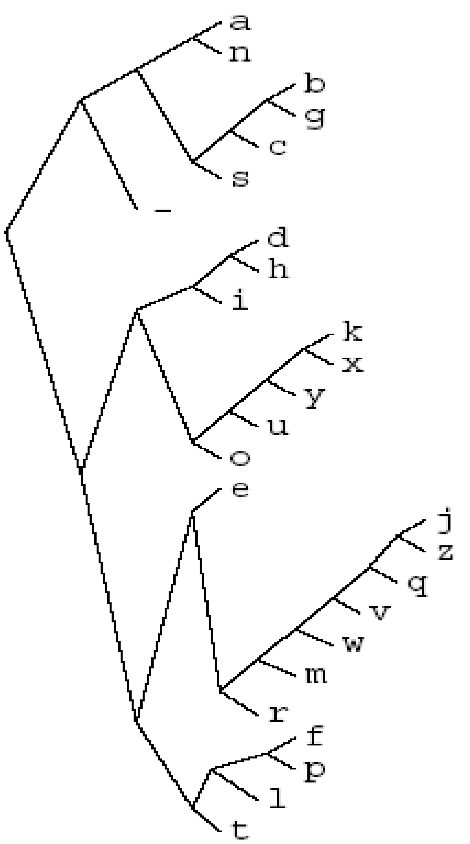


s_i	p_i	H	l_i	$C(s_i)$
α	0,25	2.0	2	00
β	0.25	2.0	2	10
γ	0.2	2.3	2	11
δ	0.15	2.7	3	010
ε	0.15	2.7	3	011
		2.2855	2.30	

Κωδικοποίηση HUFFMAN (5)

- Παράδειγμα
  - Παρατηρείστε τις διαφορές που υπάρχουν μεταξύ βέλτιστου μήκους και ελάχιστου μήκους
  - H(X)=4.11 bits
  - L(C,X)=4.15 bits

a_i	p_i	$\log_2 \frac{1}{p_i}$	l_i	$c(a_i)$
a	0.0575	4.1	4	0000
b	0.0128	6.3	6	001000
c	0.0263	5.2	5	00101
d	0.0285	5.1	5	10000
e	0.0913	3.5	4	1100
f	0.0173	5.9	6	111000
g	0.0133	6.2	6	001001
h	0.0313	5.0	5	10001
i	0.0599	4.1	4	1001
j	0.0006	10.7	10	1101000000
k	0.0084	6.9	7	1010000
l	0.0335	4.9	5	11101
m	0.0235	5.4	6	110101
n	0.0596	4.1	4	0001
o	0.0689	3.9	4	1011
p	0.0192	5.7	6	111001
q	0.0008	10.3	9	110100001
r	0.0508	4.3	5	11011
s	0.0567	4.1	4	0011
t	0.0706	3.8	4	1111
u	0.0334	4.9	5	10101
v	0.0069	7.2	8	11010001
w	0.0119	6.4	7	1101001
x	0.0073	7.1	7	1010001
y	0.0164	5.9	6	101001
z	0.0007	10.4	10	1101000001
–	0.1928	2.4	2	01



Κωδικοποίηση HUFFMAN (6)

- Παρατηρήσεις σχετικά με τον αλγόριθμο Huffman
  - Αποδεικνύεται ότι κανένας άλλος αλγόριθμος δεν μπορεί να οδηγήσει στην κατασκευή κώδικα με μικρότερο μέσο μήκος κωδικών λέξεων για ένα δεδομένο αλφάβητο πηγής.
  - Η κατασκευή του δένδρου γίνεται από τα φύλλα προς τη ρίζα του δένδρου,
  - Σε αντίθεση με τον αλγόριθμο FANO όπου δημιουργεί το δένδρο από τη ρίζα του προς τα φύλλα διαιρώντας κάθε φορά το σύνολο συμβόλων σε δύο σύνολα. Ένας τέτοιος αλγόριθμος είναι πάντα λιγότερο βέλτιστος.

Κωδικοποίηση HUFFMAN (7)

- **<u>ΑΚΟΛΟΥΘΙΕΣ ΣΥΜΒΟΛΩΝ</u>**
- Μειονεκτήματα αλγορίθμου Huffman
  - Υπόθεση: Τα σύμβολα της πηγής παράγονται ανεξάρτητα το ένα από το άλλο. Τι γίνεται όμως αν τα σύμβολα αυτά εξαρτώνται από το ποια σύμβολα έχουν παραχθεί στο άμεσο παρελθόν;
  - Γνωρίζουμε ότι ο αλγόριθμος Huffman παράγει βέλτιστο κώδικα και άρα βάσει του θεωρήματος ισχύει ότι
    - $H(X) \leq L(C,X) < H(X)+1$
    - Άρα κατά μέσο όρο και ανά σύμβολο έχουμε πλεονάζοντα bits μεταξύ 0 και 1.
    - Αν η εντροπία H(X) της πηγής είναι μεγάλη τότε το πλεονάζον αυτό bit, L(C,X)-H(X), θα ήταν αμελητέο στην παραγωγή μηνυμάτων. Αν όμως η εντροπία είναι μικρότερη από 1 bit τότε το πλεονάζον bit θα ήταν καθοριστικό στην παραγωγή μηνυμάτων.
  - Χρειάζεται να ξέρουμε τις πιθανότητες εμφάνισης εκ των προτέρων. Αν όχι τότε θα πρέπει να συλλέξουμε πρώτα τα στατιστικά μιας πηγής και μετά να εφαρμόσουμε τον αλγόριθμο
- **Έτσι λοιπόν παρόλο που οι κώδικες Huffman θεωρούνται βέλτιστοι αυτό αφορά στην παραγωγή συμβόλων και όχι στην παραγωγή μηνυμάτων που είναι και αυτό που χρειαζόμαστε στην πράξη.**

Κωδικοποίηση HUFFMAN (8)

- **<u>Παράδειγμα</u>**
  - Έστω μία πηγή με αλφάβητο $A=\{\alpha_1,\alpha_2,\alpha_3\}$ και $P=\{0.8,0.02,0.18\}$.
  - H(A)=0,816 bits/symbol
  - Εφαρμόζοντας τον αλγόριθμο του Huffman παίρνουμε τις εξής κωδικές λέξεις και μήκη
  - Παρατηρούμε ότι το μέσο μήκος κωδικής λέξης είναι 1.2 bits/symbol το οποίο απέχει κατά 47% από την εντροπία δηλαδή υπάρχει πλεονασμός κατά 0.384 bits/symbol.
  - Σε επίπεδο μηνυμάτων (ακολουθίες συμβόλων) αυτός ο πλεονασμός παίζει καθοριστικό ρόλο
    - Π.χ. Για ακολουθίες μηνυμάτων που αποτελούνται από N=1000 σύμβολα τότε σύμφωνα με την κωδικοποίηση κατά Huffman θα παράγαμε κατά μέσο όρο 384 bits περισσότερα από τα 816 που είναι τα αναγκαία
  - Τι πρέπει να γίνει;

α_i	p_i	H	l_i	$C(\alpha_i)$
α_1	0.8	0.322	1	0
α_2	0. 02	5.644	2	11
α_3	0.18	2.474	2	10
		0,816	1,2	
		H(A)	L(C,A)	

Κωδικοποίηση HUFFMAN (9)

- **Παράδειγμα (συνέχεια)**
  - Να εφαρμόσουμε τον αλγόριθμο όχι σε επίπεδο συμβόλων αλλά σε επίπεδο μηνυμάτων.
  - Έτσι για μηνύματα δύο συμβόλων έχουμε
  - Το μέσο μήκος κάθε κωδικής λέξης που αντιστοιχεί σε μήνυμα 2 συμβόλων είναι 1.7228 bits και άρα κάθε σύμβολο το μέσο μήκος κωδικής λέξης ανά σύμβολο είναι 1.7228/2=0.8614 bits το οποίο συγκρινόμενο με την εντροπία H(A)=0.816 είναι μόλις κατά 5.5% προσαυξημένο
  - Το πρόβλημα που παρουσιάζει αυτή η μέθοδος στην πράξη είναι ότι χρειάζεται να υπολογίσουμε όλες τις πιθανότητες των πιθανών μηνυμάτων. Για ένα αλφάβητο με n σύμβολα και μηνύματα μήκους m τότε το σύνολο όλων των μηνυμάτων είναι m^n, δηλαδή για ένα αλφάβητο 5 συμβόλων και μηνύματα μήκους 10 θα χρειαστεί να υπολογίσουμε περίπου 10 εκ. πιθανότητες πρώτα!!

α_i	p_i	H	l_i	$C(\alpha_i)$
$\alpha_1\alpha_1$	0.64	0.644	1	0
$\alpha_1\alpha_2$	0. 016	5.966	5	10101
$\alpha_1\alpha_3$	0.144	2.796	2	11
$\alpha_2\alpha_1$	0.016	5.966	6	101000
$\alpha_2\alpha_2$	0. 0004	11.288	8	10100101
$\alpha_2\alpha_3$	0.0036	8.118	7	1010011
$\alpha_3\alpha_1$	0. 144	2.796	3	100
$\alpha_3\alpha_2$	0.0036	8.118	8	10100100
$\alpha_3\alpha_3$	0.0324	4.948	4	1011
		1,632	1,7228	
		$H(A^2)$	$L(C,A^2)$	
		2*H(A)	$L(C,A^2)/2$=0,861	