Θ.Ε. ΠΛΗ22 (2012-13) – ΓΡΑΠΤΗ ΕΡΓΑΣΙΑ #4.

Έκδοση v2 με διόρθωση τυπογραφικού λάθους στο ερώτημα 6.3

Στόχος:

Βασικό στόχο της 4ης εργασίας αποτελεί η εξοικείωση με τα μέτρα ποσότητας πληροφορίας τυχαίων μεταβλητών (Κεφάλαιο 1), τις σχετικές έννοιες και τα μέτρα διακριτών πηγών χωρίς μνήμη και με μνήμη (Κεφάλαιο 2), καθώς και με τις ιδιότητες κωδίκων και την εφαρμογή αλγορίθμων κωδικοποίησης πηγής (Κεφάλαιο 2). Επίσης, στους στόχους συμπεριλαμβάνεται και η εξοικείωση με θέματα σχετικά με τον μέγιστο ρυθμό μετάδοσης (χωρητικότητα) και την κωδικοποίηση πληροφορίας ψηφιακών καναλιών επικοινωνίας (Κεφάλαιο 3).

ΘΕΜΑ 1

Στόχος της άσκησης *είναι η εξοικείωση με τα μέτρα ποσότητας πληροφορίας καθώς και με τις έννοιες της υπό συνθήκη και αμοιβαίας ποσότητας πληροφορίας δύο τυχαίων μεταβλητών.*
Σχετικές ασκήσεις: *Θ1/ΓΕ4/2004-5, Θ1/ΓΕ4/2005-6, Θ1/ΓΕ4/2006-7, Θ1/ΓΕ4/2009-10.*

Στον τελικό των Play-Offs του Βορειοαμερικανικού Πρωταθλήματος Καλαθοσφαίρισης NBA, πρωταθλήτρια θα αναδειχθεί εκείνη η ομάδα από τις δύο που έφθασαν στον τελικό που θα πετύχει 4 νίκες σε μία σειρά το πολύ έως 7 αγώνων. (Κάθε αγώνας ολοκληρώνεται με τη νίκη της μίας ή της άλλης ομάδας, δεν υφίσταται το ισόπαλο αποτέλεσμα.) Ορίζουμε Χ την τυχαία μεταβλητή που αναπαριστά το αποτέλεσμα της σειράς των αγώνων του τελικού των Play-Offs ανάμεσα στην ομάδα Α και την ομάδα Β. Δηλαδή η μεταβλητή Χ μπορεί να πάρει τις τιμές ΑΑΑΑ, ΒΒΒΒ, ΑΒΑΑΑ, ΒΑΒΒΒ, κλπ., όπου το ‘ΑΑΑΑ’ σημαίνει ότι νικήτρια των τεσσάρων αγώνων ήταν η ομάδα Α κ.ο.κ. Επιπλέον, ορίζουμε και την τυχαία μεταβλητή Υ που αναπαριστά το πλήθος των αγώνων που απαιτούνται μέχρι την ανάδειξη της πρωταθλήτριας ομάδας, που κυμαίνεται από 4 έως 7. Υποθέτουμε ότι οι ομάδες Α και Β είναι ισοδύναμες και επομένως με την ίδια πιθανότητα μπορούν να κατακτήσουν τη νίκη σε έναν αγώνα, όπως επίσης ότι το αποτέλεσμα κάθε αγώνα είναι ανεξάρτητο από τα προηγούμενα.

Ζητείται

1. Ποιές είναι οι δυνατές τιμές των τυχαίων μεταβλητών Χ και Υ και ποιές οι πιθανότητές τους;
2. Οι μέσες ποσότητες πληροφορίας των Χ και Υ.
3. Οι υπό συνθήκη ποσότητες πληροφορίας Η(Υ/Χ) και Η(Χ/Υ).
4. Η αμοιβαία πληροφορία *I(X;Y)*.

(Υπόδειξη: προσδιορίστε πρώτα τον αριθμό των δυνατών συνδυασμών για κάθε τιμή της τυχαίας μεταβλητής Υ, συμβολίζοντας όπως ανωτέρω τη μία ομάδα με Α και την άλλη ομάδα με Β. Παρατηρήστε επίσης ότι η ομάδα που τελικά κερδίζει το πρωτάθλημα (είτε η ομάδα Α είτε η ομάδα Β) θα καταλαμβάνει οπωσδήποτε την τελευταία θέση στη σειρά των αγώνων ενώ στις υπόλοιπες θέσεις της σειράς των αγώνων θα καταλαμβάνει οπωσδήποτε 3. Έτσι, *για παράδειγμα*, σε σειρά 7 αγώνων (Υ=7), οι δυνατοί έγκυροι συνδυασμοί είναι 40.)

Ενδεικτική Μεθοδολογία: *Πρώτα, αφού προσδιορίσετε όλα τα ενδεχόμενα (συνδυασμούς) για κάθε τιμή των τυχαίων μεταβλητών Χ και Υ, να υπολογίσετε τις πιθανότητες κάθε γεγονότος και ακολούθως να εφαρμόσετε τους τύπους με τους οποίους υπολογίζονται τα ζητούμενα μέτρα.*

ΘΕΜΑ 2

Στόχος της άσκησης *είναι η εξοικείωση με τις ιδιότητες που πρέπει να πληρούν οι κώδικες πηγής, δηλαδή να είναι μοναδικά αποκωδικοποιήσιμοι και άμεσοι.*
Σχετικές ασκήσεις: *ΓΕ4/2011-2/Θ3.*

Ορίζουμε τον κώδικα-γινόμενο (product code) δύο δυαδικών κωδίκων πηγής C1 και C2, αντίστοιχα, ως τον κώδικα C1xC2 του οποίου οι κωδικές λέξεις είναι όλοι οι διαφορετικοί συνδυασμοί (ακολουθίες) της μορφής $c_i d_j$, όπου $c_i \in C1$ και $d_j \in C2$. Για παράδειγμα αν C1={0,01} και C2={0, 10} , τότε C1 x C2 = {00, 010, 0110}.
Ζητούνται τα ακόλουθα:

1. Είναι οι κώδικες C1 και C2 του παραδείγματος μοναδικά αποκωδικοποιήσιμοι και άμεσοι;
2.
   a. Δείξτε ότι αν στη γενική περίπτωση οι C1 και C2 είναι άμεσοι τότε και ο κώδικας-γινόμενο C1 x C2 είναι άμεσος. *(Υπόδειξη: να κάνετε χρήση του ορισμού, σύμφωνα με τον οποίο αν ένας κώδικας C είναι άμεσος, δηλαδή ελεύθερος προθέματος, τότε δεν μπορεί να ισχύει ότι $c_j=c_k a$, όπου c_j, c_k είναι κωδικές λέξεις του C και το a μπορεί να είναι οποιαδήποτε ακολουθία δυαδικών ψηφίων ακόμα και το $\emptyset$.)*
   b. Αν ένας από τους 2 κώδικες (ή και οι δύο) του προηγούμενου ερωτήματος είναι μη άμεσοι υπάρχει περίπτωση ο C1 x C2 να είναι άμεσος;
3. Πάλι στη γενική περίπτωση, αν οι C1 και C2 είναι μοναδικώς αποκωδικοποιήσιμοι (αλλά όχι απαραίτητα άμεσοι), ισχύει πάντοτε ότι ο C1 x C2 είναι μοναδικώς αποκωδικοποιήσιμος; Αποδείξτε το δίνοντας ένα αντιπαράδειγμα όπου C1 και C2 είναι μοναδικώς αποκωδικοποιήσιμοι αλλά ο C1 x C2 δεν είναι.

Ενδεικτική Μεθοδολογία: *Για την απάντηση του ερωτήματος 1, απλά εξετάστε αν οι δεδομένοι κώδικες πληρούν τις επιθυμητές ιδιότητες. Για την απάντηση του ερωτήματος 2.a, να αξιοποιήσετε την ανωτέρω υπόδειξη, σύμφωνα με την οποία για να είναι άμεσος ο κώδικας – γινόμενο C1xC2 δύο άμεσων κωδίκων C1 και C2, αρκεί να δείξετε ότι και ο κώδικας-γινόμενο C1xC2 είναι ελεύθερος προθέματος. Για την απάντηση των ερωτημάτων 2.b και 3, αξιοποιείστε κατάλληλα παραδείγματα κωδίκων.*

ΘΕΜΑ 3

Στόχος της άσκησης *είναι η εξοικείωση με τις ιδιότητες κωδίκων συμπίεσης γενικά και ειδικότερα με χαρακτηριστικά αλγορίθμων κωδικοποίησης πηγής.*
Σχετικές ασκήσεις: *Θ3/ΓΕ4/2011-12, Θ4/ΓΕ4/2011-12, Θ3/ΓΕ4/2010-11, Θ2/ΓΕ4/2009-10, Θ3/ΓΕ4/2008-09, Θ4/ΓΕ4/2006-7 και Θ3/ΓΕ/2004-5.*

Θεωρούμε πηγή με αλφάβητο τα σύμβολα που απεικονίζονται διαταγμένα στον παρακάτω πίνακα με φθίνουσα πιθανότητα εκπομπής

Σύμβολο	**Πιθανότητα p_i**
Α	½
Β	¼
Γ	x
Δ	1/16
Ε	y

Επίσης, θεωρούμε ότι το μέσο πληροφορικό περιεχόμενο των συμβόλων της ως άνω πηγής είναι ίσο με 1,875 bits. Στη συνέχεια το αλφάβητο αυτό κωδικοποιείται με έναν άριστο δυαδικό κώδικα, με τα εξής μήκη κωδικών λέξεων: $l_i = \{1,2,3,4,4\}$ $i = 1,..5$, αντίστοιχα για τα σύμβολα {Α, Β, Γ, Δ, Ε}.

Ζητούνται:

1. Να σχηματίσετε τον άριστο αυτό κώδικα για το δεδομένο αλφάβητο της πηγής, αφού προηγουμένως προσδιορίσετε τις πιθανότητες εμφάνισης των συμβόλων Γ και Ε, σύμφωνα με τα ανωτέρω δεδομένα.
2. Είναι ο κώδικας αυτός μοναδικά αποκωδικοποιήσιμος και άμεσος?
3. Αν η πηγή εκπέμψει 10^8 σύμβολα, τα οποία κωδικοποιούνται σύμφωνα με τον ως άνω κώδικα που σχηματίσατε, να βρεθεί ο αριθμός των bits που έχει εκπέμψει η πηγή ως πλεονασμό και τα οποία οφείλονται στις δεδομένες πιθανότητες εμφάνισης των συμβόλων της πηγής. Υπάρχει πλεονασμός κωδικοποίησης;

<u>Ενδεικτική Μεθοδολογία:</u> *Αρχικά να βρεθούν οι πιθανότητες των συμβόλων, να χαρακτηρισθεί ο κώδικας και ακολούθως να επιλεγεί ο κατάλληλος αλγόριθμος κωδικοποίησης. Τέλος να εφαρμοσθεί ο ορισμός του πλεονασμού. Υπενθυμίζεται ότι ο πρόσθετος πλεονασμός που εισάγεται από μη άριστο κώδικα ονομάζεται πλεονασμός κωδικοποίησης.*

ΘΕΜΑ 4

Στόχος της άσκησης *είναι η εξάσκηση στην εφαρμογή των αλγορίθμων κωδικοποίησης.*
Σχετικές ασκήσεις: *Θ4/ΓΕ4/2011-12, Θ3/ΓΕ4/2010-11, Θ2/ΓΕ4/2009-10.*

Δίδεται διακριτή πηγή που παράγει 8 διαφορετικά σύμβολα, *Α, Β, Γ, Δ, Ε, Ζ, Η* και Θ, με τις ακόλουθες πιθανότητες εμφάνισης, αντίστοιχα: {0.10, 0.15, 0.20, 0.25, 0.05, 0.10, 0.05, 0.10}.

Ζητείται:

1. Να σχηματιστεί δυαδικός κώδικας σύμφωνα με τον αλγόριθμο Huffman.
2. Να σχηματιστεί δυαδικός κώδικας σύμφωνα με τον αλγόριθμο Fano.
3. Να σχηματισθεί δυαδικός κώδικας σύμφωνα με τον αλγόριθμο Shannon.
4. Να συγκριθούν οι κώδικες που προκύπτουν στα ερωτήματα (1) – (3) ως προς την επίδοσή τους.

Ενδεικτική Μεθοδολογία: *Εφαρμόζονται οι αλγόριθμοι κωδικοποίησης Huffman, Shannon και Fano, καθώς και ο τύπος της επίδοσης του κώδικα.*

ΘΕΜΑ 5

Στόχος της άσκησης *είναι η εξοικείωση με τον υπολογισμό του ρυθμού πληροφορίας πηγής και τη σημασία του δεύτερου θεωρήματος κωδικοποίησης του Shannon σχετικά με τη δυνατότητα ή μη μετάδοσης χωρίς σφάλματα για δεδομένο ρυθμό πληροφορίας πηγής, χωρητικότητα καναλιού και SNR.*
Σχετικές ασκήσεις*: ΓΕ 4/2011-12, ΓΕ4/2005-6/Θ5, ΓΕ4/2006-7/Θ7, ΓΕ4/2008-9/Θ5.*

Το αλφάβητο μιας στατικής πηγής χωρίς μνήμη αποτελείται από 128 σύμβολα. Τα (πρώτα) 32 από αυτά τα σύμβολα είναι ισοπίθανα, εμφανίζονται με πιθανότητα 1/128 και έχουν διάρκεια τ_A=0.1 msec**.** Τα (επόμενα) 32 είναι και αυτά ισοπίθανα, εμφανίζονται με πιθανότητα 1/64 και έχουν διάρκεια τ_B=0.2 msec**.** Τα υπόλοιπα είναι και αυτά ισοπίθανα με χρονική διάρκεια τ_C=0.3 msec. Αν τα σύμβολα εκπέμπονται χωρίς χρονικό κενό μεταξύ τους, ζητείται:

1. Να υπολογιστεί η εντροπία H και ο ρυθμός παραγωγής συμβόλων R.
2. Να δείξετε ότι δεν είναι δυνατή η μετάδοση της πληροφορίας της πηγής χωρίς σφάλματα σε ένα κανάλι με εύρος ζώνης 8KHz και SNR 10 dB.
3. Τι SNR θα απαιτούνταν για μετάδοση της ως άνω πληροφορίας χωρίς σφάλματα σε κανάλι εύρους ζώνης 8KHz;

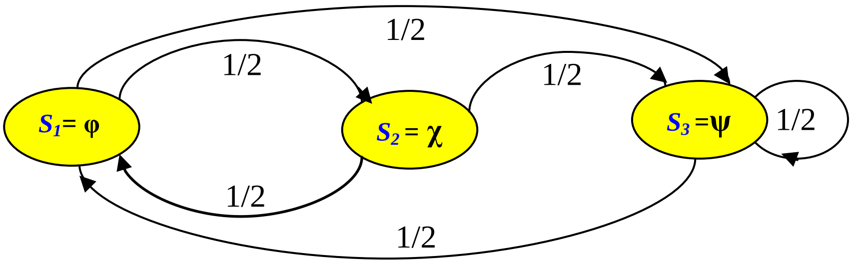
Ενδεικτική Μεθοδολογία: *Ισχύει* $C = W \times \log_2(1+SNR)$, *όπου C η χωρητικότητα του καναλιού, W το εύρος ζώνης και SNR ο λόγος σήματος προς θόρυβο. Επίσης, να βασιστείτε στο δεύτερο θεώρημα κωδικοποίησης του Shannon.*

ΘΕΜΑ 6

Στόχος της άσκησης *είναι η εξοικείωση με τις βασικές αρχές-έννοιες πηγών πληροφορίας με μνήμη που μοντελοποιούνται με αλυσίδες Markoff.*

Σχετικές ασκήσεις: *ΓΕ4/2005-6/Θ5, ΓΕ4/2006-7/Θ7, ΓΕ4/2008-9/Θ5, ΓΕ4/2011-12/Θ6.*

Δίδεται στατική πηγή Markoff 1ης τάξης, με αλφάβητο τα σύμβολα φ, χ, και ψ. Ο πίνακας μετάβασης της πηγής περιγράφεται στο ακόλουθο διάγραμμα:



Ζητούνται τα ακόλουθα:

1. Οι στατικές πιθανότητες εκπομπής των συμβόλων φ, χ και ψ.
2. Η εντροπία της πηγής, *H(S)*.
3. Οι συνδυασμένες πιθανότητες εκπομπής μηνυμάτων αποτελούμενων από δύο σύμβολα: $m_1=(\varphi,\varphi)$, $m_2=(\varphi,\chi)$, $m_3=(\varphi,\psi)$, $m_4=(\chi,\varphi)$, $m_5=(\chi,\chi)$, $m_6=(\chi,\psi)$, $m_7=(\psi,\varphi)$, $m_8=(\psi,\chi)$, $m_9=(\psi,\psi)$.
4. Βέλτιστος δυαδικός κώδικας των συμβόλων της πηγής αυτής Markoff, οποίος να κωδικοποιεί ανά δύο τα σύμβολα, ανεξαρτήτως της κατάστασης της πηγής.

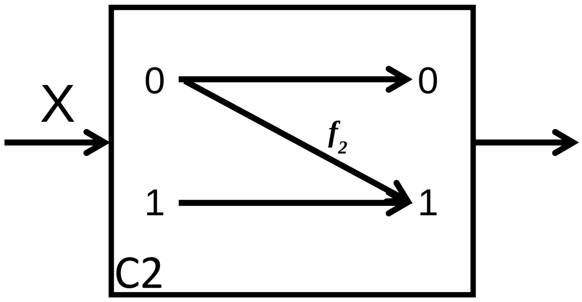
<u>Ενδεικτική Μεθοδολογία:</u> *Για την απάντηση του ερωτήματος 1, να καταστρώσετε και επιλύσετε κατάλληλο σύστημα εξισώσεων ως προς τις ζητούμενες πιθανότητες. Για το ερώτημα 2, εφαρμόζετε τους σχετικούς τύπους του βιβλίου και για την απάντηση του ερωτήματος 4, αφού υπολογίσετε τις πιθανότητες του ερωτήματος 3, εφαρμόζετε κατάλληλο αλγόριθμο κωδικοποίησης πηγής.*

ΘΕΜΑ 7

Στόχος της άσκησης *είναι η εξοικείωση με θέματα διακριτών καναλιών επικοινωνίας και ιδίως με τις έννοιες του πίνακα μετάβασης, της αμοιβαίας πληροφορίας μεταξύ εισόδου και εξόδου και της χωρητικότητας ενθόρυβου καναλιού επικοινωνίας.*

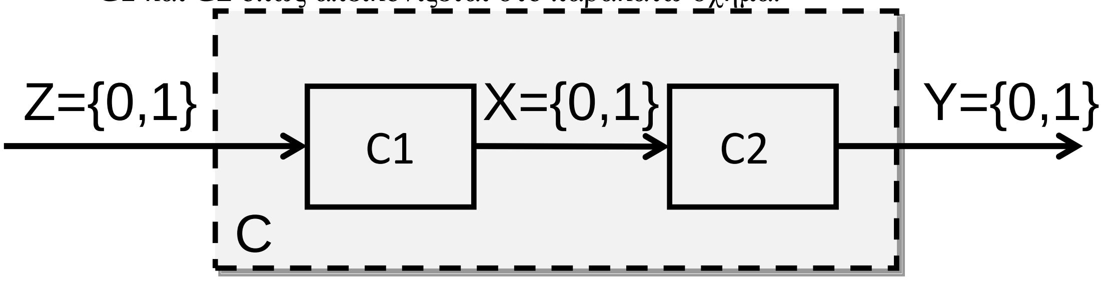
Σχετικές ασκήσεις: *Α.Α. 3.2, ΓΕ4/2005-06/Θ4, ΓΕ4/2010-11/Θ7, ΓΕ4/2011-12/Θ7, ΕΞ2008Α/Θ5.*

1. Δίνεται το κανάλι **C2** το οποίο απεικονίζεται στο παρακάτω σχήμα όπου f_2=1/4.



Ζητείται τι είδους κανάλι είναι το κανάλι **C2** και στη συνέχεια να βρείτε την χωρητικότητά του καθώς και τις πιθανότητες εμφάνισης των συμβόλων στην είσοδο του καναλιού για τις οποίες επιτυγχάνεται η χωρητικότητα αυτή.

2. Ακολούθως, το κανάλι **C2** συνδέεται με τις εξόδους ενός καναλιού **C1** που είναι ένα δυαδικό συμμετρικό κανάλι (BSC) με πιθανότητα ορθής μετάδοσης 3/4 και προκύπτει το κανάλι **C** το οποίο είναι συνδυασμός των δύο καναλιών **C1** και **C2** όπως απεικονίζεται στο παρακάτω σχήμα.



Αν υποθέσουμε ότι οι πιθανότητες εμφάνισης των συμβόλων εξόδου του καναλιού C1 είναι ίδιες με αυτές των πιθανοτήτων εισόδου του καναλιού **C2** που δίνονται στο προηγούμενο ερώτημα, να απαντηθούν τα κάτωθι:

   a. Να σχεδιαστεί το κανάλι **C** συναρτήσει των δύο επιμέρους καναλιών και να βρεθεί ο πίνακας μετάβασής του P;
   b. Να υπολογίσετε την αμοιβαία πληροφορία I(Z;Y) του καναλιού **C**.

<u>Ενδεικτική Μεθοδολογία:</u> *Για τις απαντήσεις των ερωτημάτων, πρώτα προσδιορίζετε τις σχετικές πιθανότητες μετάβασης και στη συνέχεια εφαρμόζετε τους τύπους υπολογισμού της χωρητικότητας και της αμοιβαίας πληροφορίας. Για τη σχεδίαση του καναλιού C, να λάβετε υπόψη τη δεδομένη σχεδίαση για το κανάλι C2 και τη γνωστή απεικόνιση ενός δυαδικού συμμετρικού καναλιού.*

Τρόπος – Ημερομηνία Παράδοσης

1. Η εργασία σας θα πρέπει να έχει αποσταλεί στον Καθηγητή-Σύμβουλό σας μέχρι την 7η Απριλίου 2013, ώρα 23:59.
2. Περιμένουμε όλες οι εργασίες να σταλούν με χρήση της υπηρεσίας ανάρτησης και διαχείρισης ΓΕ του ΕΑΠ, μέσω του συνδέσμου http://moodle.eap.gr και να είναι γραμμένες σε επεξεργαστή κειμένου (π.χ. MS-Word).
3. Τη 12η Απριλίου 2013 θα δημοσιευθεί ενδεικτική απάντηση για την επίλυση της εργασίας στο site της Θ.Ε. στο http://moodle.eap.gr και στην ιστοσελίδα της ΠΛΗ-22 "http://p-comp.di.uoa.gr/eap/index.html".

Κριτήρια αξιολόγησης:

ΘΕΜΑ 1	**15**	
Ερώτημα 1		6
Ερώτημα 2		3
Ερώτημα 3		3
Ερώτημα 4		3
ΘΕΜΑ 2	**15**	
Ερώτημα 1		2
Ερώτημα 2.a		8
Ερώτημα 2.b		2
Ερώτημα 3		3
ΘΕΜΑ 3	**10**	
Ερώτημα 1		6
Ερώτημα 2		2
Ερώτημα 3		2
ΘΕΜΑ 4	**12**	
Ερώτημα 1		3
Ερώτημα 2		3
Ερώτημα 3		3
Ερώτημα 4		3
ΘΕΜΑ 5	**12**	
Ερώτημα 1		6
Ερώτημα 2		3
Ερώτημα 3		3
ΘΕΜΑ 6	**14**	
Ερώτημα 1		4
Ερώτημα 2		4
Ερώτημα 3		3
Ερώτημα 4		3
ΘΕΜΑ 7	**22**	
Ερώτημα 1		10
Ερώτημα 2.a		6
Ερώτημα 2.b		6
ΣΥΝΟΛΟ	**100**	

Ο συνολικός βαθμός θα διαιρεθεί δια 10, ώστε να προκύψει ο τελικός βαθμός της εργασίας.

Καλή Επιτυχία!!!