



Πρόγραμμα Μεταπτυχιακών Σπουδών
"Προηγμένες Τεχνολογίες Υπολογιστών
και Πληροφορικής"

Υπολογιστική Νοημοσύνη

Μάθημα 6^ο

Δρ. Βασίλειος Γ. Καμπουρλάζος
Δρ. Ανέστης Γ. Χατζημιχαηλίδης

Τμήμα Μηχανικών Πληροφορικής Τ.Ε.
ΤΕΙ Ανατολικής Μακεδονίας και Θράκης

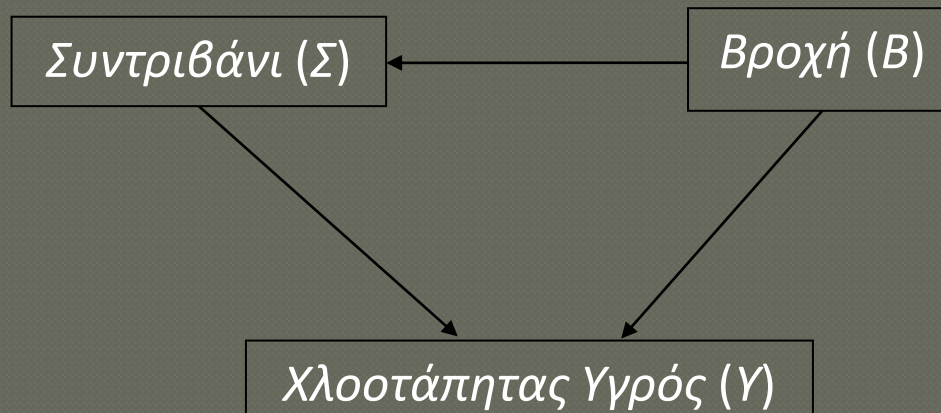
2016-2017

Περιεχόμενα

Διευρυμένη Υπολογιστική Νοημοσύνη (ΥΝ)

- ◉ Επεκτάσεις της Κλασικής ΥΝ.
- ◉ Μεθοδολογίες Στατιστικές, Πιθανοτήτων, Δυνατοτήτων.
- ◉ Μεθοδολογίες με Γράφους.

- Σε κάποιες περιπτώσεις στα δεδομένα του προβλήματος υπάρχει δομή η οποία μπορεί να αναπαρασταθεί με έναν γράφο. Π.χ.



Μεθοδολογίες με Γράφους

- Δίκτυα Bayes.
- Δένδρα Αποφάσεων.
- Γνωσιακοί Χάρτες.

Δίκτυα Bayes (ΔB)

- ❖ Τα Δίκτυα Bayes (ΔB) είναι εναλλακτικά γνωστά με ονόματα ως δίκτυα πεποίθησης (**belief networks**) καθώς και ως πιθανοτικά μοντέλα κατευθυνόμενων ακυκλικών γράφων (ΚΑΓ) (**directed acyclic graphs (DAGs)**).
- ❖ Οι κόμβοι του ΚΑΓ αναπαριστούν τυχαίες μεταβλητές, ενώ οι σύνδεσμοι αναπαριστούν εξαρτήσεις μεταξύ των μεταβλητών.

Συμπερασμός ΔB

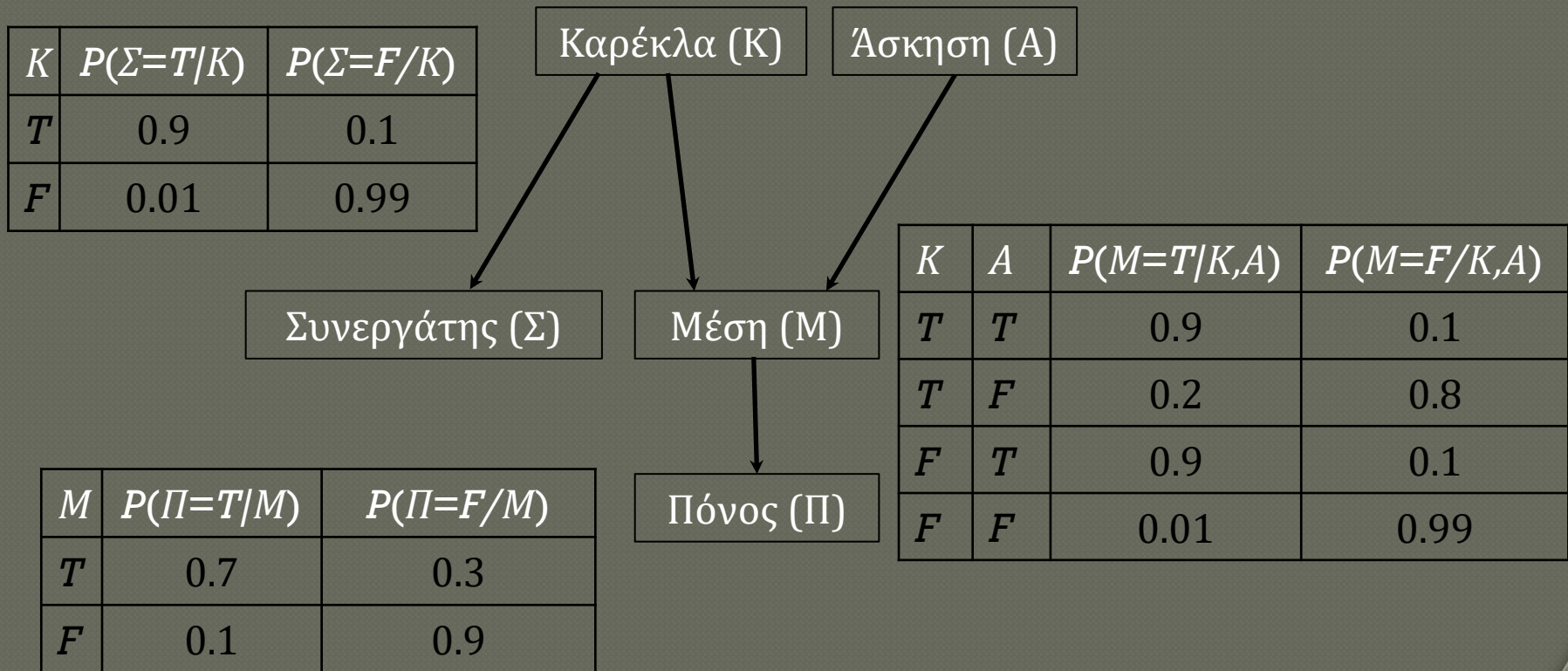
- ◉ Ένα ΔB υποστηρίζει τους ακόλουθους τύπους συμπερασμού:
 - ❖ Συμπερασμός πρόβλεψης για έναν κόμβο X χρησιμοποιώντας γνώση των γονέων του X ή/και άλλων προγόνων του X – Συλλογιστική από πάνω προς τα κάτω.
 - ❖ Συμπερασμός διάγνωσης για έναν κόμβο X χρησιμοποιώντας γνώση των παιδιών του X ή/και άλλων απογόνων του X – Συλλογιστική από κάτω προς τα πάνω.

Εφαρμογή Συμπερασμού Διάγνωσης

- Έστω ότι ένας άνθρωπος υποφέρει από πόνο στη μέση. Χρησιμοποιώντας συλλογιστική από κάτω προς τα πάνω μπορούμε να υπολογίσουμε την πεποίθηση ότι υπάρχει μια μη-αναπνευστική καρτέκλα στο γραφείο :

Παράδειγμα ΔB (διάγνωσης)

$P(K=T)$	$P(K=F)$	$P(A=T)$	$P(A=F)$
0.8	0.2	0.02	0.98



Συμπερασμός ΔB (διάγνωσης)

- Αντί η από κοινού συνάρτηση κατανομής πιθανότητας να υπολογίζεται με τον (γενικό) κανόνα αλυσίδας:

$$P(K, A, \Sigma, M, \Pi) = P(K)P(A | K)P(\Sigma | A, K)P(M | \Sigma, A, K)P(\Pi | M, \Sigma, A, K)$$

Υπολογίζεται ως

$$P(K, A, \Sigma, M, \Pi) = P(K)P(A)P(\Sigma | K)P(M | A, K)P(\Pi | M)$$

Συμπερασμός ΔB (διάγνωσης)

- Δοθέντος ότι ένας άνθρωπος έχει πόνο υπολογίζουμε την *πεποίθηση* ότι υπάρχει μια μη-αναπαυτική καρέκλα στο γραφείο:

$$P(K = T | \Pi = T) = \frac{P(K = T, \Pi = T)}{P(\Pi = T)}$$

όπου

$$P(K = T, \Pi = T) = \sum_{A, \Sigma, M \in \{T, F\}} P(K = T)P(A)P(\Sigma | K = T)P(M | A, K = T)P(\Pi = T | M)$$

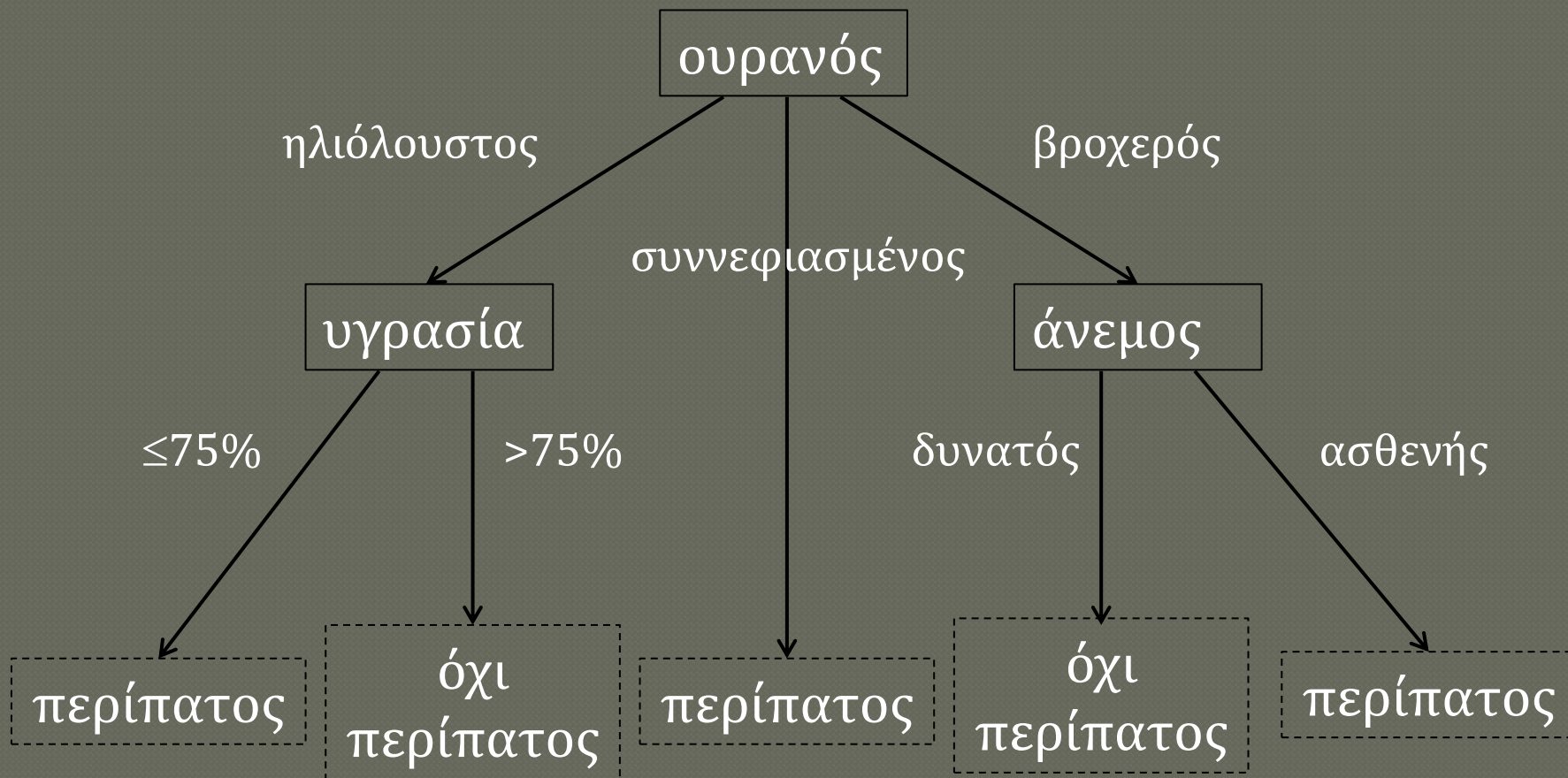
και

$$P(\Pi = T) = \sum_{K, A, \Sigma, M \in \{T, F\}} P(K)P(A)P(\Sigma | K)P(M | A, K)P(\Pi = T | M)$$

Δένδρα Αποφάσεων

❖ Ένα δένδρο αποφάσεων χρησιμοποιείται για λήψη αποφάσεων. Για παράδειγμα, έστω ότι η απόφασή μας να κάνουμε περίπατο εξαρτάται από τις καιρικές συνθήκες σύμφωνα με το ακόλουθο Σχήμα, το οποίο δόθηκε από κάποιον εμπειρογνώμονα.

Δένδρο Αποφάσεων



Σύνολο *T* δεδομένων για επαγωγή Δένδρου Απόφασης

ουρανός	θερμοκρασία (°C)	υγρασία (%)	άνεμος	κατηγορία
ηλιόλουστος	20	70	δυνατός	περίπατος
ηλιόλουστος	25	90	δυνατός	όχι περίπατος
ηλιόλουστος	27	85	ασθενής	όχι περίπατος
ηλιόλουστος	19	95	ασθενής	όχι περίπατος
ηλιόλουστος	18	70	ασθενής	περίπατος
συννεφιασμένος	19	90	δυνατός	περίπατος
συννεφιασμένος	26	78	ασθενής	περίπατος
συννεφιασμένος	17	65	δυνατός	περίπατος
συννεφιασμένος	25	75	ασθενής	περίπατος
βροχερός	19	80	δυνατός	όχι περίπατος
βροχερός	15	70	δυνατός	όχι περίπατος
βροχερός	20	80	ασθενής	περίπατος
βροχερός	18	80	ασθενής	περίπατος
βροχερός	19	96	ασθενής	περίπατος

Υπολογισμός Δένδρου Απόφασης με Επαγωγή

❖ Τυπικά, ένα δένδρο αποφάσεων υπολογίζεται με τεχνική τύπου **διαίρει-και-βασίλευε (divide-and-conquer)**. Συγκεκριμένα, έστω (α) ένα σύνολο T δεδομένων, όπου ένα δεδομένο περιλαμβάνει N γνωρίσματα τα οποία είναι είτε αριθμητικά, είτε μη-αριθμητικά και (β) υπάρχουν συνολικά k κατηγορίες C_1, C_2, \dots, C_k .

Τρεις είναι οι δυνατές περιπτώσεις:

(Υπολογισμός Δένδρου Απόφασης με Επαγωγή)

1. Το σύνολο T είναι κενό.
2. Το σύνολο T περιλαμβάνει ένα ή περισσότερα δεδομένα τα οποία ανήκουν σε μία μόνον κατηγορία, έστω C_j . Σ' αυτήν την περίπτωση το δένδρο απόφασης είναι ένα φύλλο, το οποίο αντιστοιχεί στην κατηγορία C_j .
3. Το T περιλαμβάνει δεδομένα σε περισσότερες από μία κατηγορίες. Στόχος είναι ο **διαμερισμός (partition)** του συνόλου T σε υποσύνολα, των οποίων όλα τα δεδομένα ανήκουν σε μία μόνον κατηγορία. Επιλέγουμε μια **δοκιμασία (test)**, η οποία τυπικά χρησιμοποιεί ένα μόνον γνώρισμα, με αποτέλεσμα στο σύνολο $\{O_1, O_2, \dots, O_n\}$. Έτσι το T διαμερίζεται σε υποσύνολα T_1, T_2, \dots, T_n .

(Υπολογισμός Δένδρου Απόφασης με Επαγωγή)

Συγκεκριμένα,

1. Έστω ότι επιλέγουμε μια δοκιμασία με βάση το (πρώτο) γνώρισμα «ουρανός». Στο αποτέλεσμα «συννεφιασμένος» αντιστοιχεί μόνον η κατηγορία «περίπατος».
2. Το υποσύνολο που αντιστοιχεί στο αποτέλεσμα «ηλιόλουστος» μπορεί να διαμεριστεί αν επιλέξουμε μια δοκιμασία με βάση το (τρίτο) γνώρισμα «υγρασία» με δύο αποτελέσματα: «υγρασία \leq 75%», «υγρασία $>$ 75%».
3. Το υποσύνολο που αντιστοιχεί στο αποτέλεσμα «βροχερός» μπορεί να διαμεριστεί αν επιλέξουμε μια δοκιμασία με βάση το (τέταρτο) γνώρισμα «άνεμος» με δύο δυνατά αποτελέσματα: «δυνατός» και «ασθενής».

Αξιολόγηση Δοκιμασιών

Θέλουμε κάθε φύλλο του δένδρου απόφασης που θα υπολογιστεί να περιλαμβάνει ένα σημαντικό αριθμό δεδομένων ή, ισοδύναμα, θέλουμε να διαμερίσουμε το σύνολο T των δεδομένων στο μικρότερο δυνατό αριθμό υποσυνόλων. Σε κάθε περίπτωση επιζητούμε τον υπολογισμό μικρών δένδρων απόφασης.

Ένας τρόπος εύρεσης του μικρότερου δένδρου απόφασης είναι με **τυφλή αναζήτηση (blind search)**. Όμως, για τα λίγα δεδομένα του προηγούμενου Πίνακα υπάρχουν περισσότερα από 10^6 δένδρα απόφασης.

(Αξιολόγηση Δοκιμασιών)

- Μια δημοφιλής ευρετική συνάρτηση που χρησιμοποιείται για την επιλογή μιας δοκιμασίας είναι η συνάρτηση κέρδους πληροφορίας η οποία βασίζεται στη συνάρτηση εντροπία (Shannon, 1948). Όσο πιο σπάνιο είναι ένα γεγονός, τόσο μεγαλύτερη να είναι η πληροφορία που προσλαμβάνουμε, όταν αυτό συμβεί. Συγκεκριμένα, έστω P_A η πιθανότητα να συμβεί ένα γεγονός A . Τότε η πληροφορία που μεταφέρει το A ορίζεται ως $-\log_2(P)$ και μετριέται σε **bits**. Για παράδειγμα, όταν από οκτώ ισοπίθανα γεγονότα συμβεί το ένα, τότε θεωρούμε ότι προσλαμβάνουμε πληροφορία ίση με $-\log_2(8) = 3$ **bits**.

(Αξιολόγηση Δοκιμασιών)

- Έστω ότι είναι δυνατόν να συμβούν n πιθανά γεγονότα A_1, \dots, A_n με πιθανότητες P_1, \dots, P_n , αντίστοιχα. Τότε ως **εντροπία (entropy)**, συμβολικά $\text{info}(A_1, \dots, A_n)$, ορίζεται ο μέσος όρος της πληροφορίας όλων των γεγονότων, δηλ.
$$\text{info}(A_1, \dots, A_n) = - \sum_{i=1}^n P_i \log_2(P_i).$$
- Τα προηγούμενα χρησιμοποιούνται στη συνέχεια για να οριστεί μια ευρετική συνάρτηση κέρδους.

(Αξιολόγηση Δοκιμασιών)

- Έστω ένα σύνολο T δεδομένων πληθικότητας $|T|$ και έστω $\text{freq}(C_j, T)$ ο συνολικός αριθμός των δεδομένων της κατηγορίας C_j , $j \in \{1, \dots, k\}$. Σύμφωνα με τα προηγούμενα, η κατηγορία C_j μεταφέρει πληροφορία ίση με $\log_2\left(\frac{\text{freq}(C_j, T)}{|T|}\right)$ bits. Άρα, η εντροπία του συνόλου των κατηγοριών στο

σύνολο T είναι:
$$\text{info}(T) = - \sum_{j=1}^k \frac{\text{freq}(C_j, T)}{|T|} \log_2\left(\frac{\text{freq}(C_j, T)}{|T|}\right)$$

- Η μέση πληροφορία στο διαμερισμό του συνόλου T σε n υποσύνολα T_1, \dots, T_n ως αποτέλεσμα εφαρμογής μιας δοκιμασίας \mathbf{X} :
$$\text{info}_{\mathbf{X}}(T) = - \sum_{i=1}^n \frac{|T_i|}{|T|} \text{info}(T_i)$$

(Αξιολόγηση Δοκιμασιών)

- Η ποσότητα $\text{gain}(X) = \text{info}(T) - \text{info}_X(T)$ ονομάζεται **κέρδος πληροφορίας (information gain)** ή εναλλακτικά **αμοιβαία πληροφορία (mutual information)**, όταν το σύνολο T διαμερίζεται με τη δοκιμασία X .
- Η επιδίωξή μας είναι να επιλέξουμε μια δοκιμασία η οποία να μεγιστοποιεί το κέρδος πληροφορίας.

Παράδειγμα

- Θεωρήστε το προηγούμενο σύνολο T δεδομένων.
- Υπάρχουν δύο κατηγορίες: η κατηγορία «περίπατος» με 9 δεδομένα και η κατηγορία «όχι περίπατος» με 5 δεδομένα. Η εντροπία του συνόλου T είναι

$$\text{info}(T) = -\frac{9}{14} \log_2\left(\frac{9}{14}\right) - \frac{5}{14} \log_2\left(\frac{5}{14}\right) = 0.940 \text{ bits.}$$

(Παράδειγμα)

- Χρησιμοποιώντας το γνώρισμα «ουρανός» ως δοκιμασία X διαμερίζουμε το σύνολο T σε τρία υποσύνολα με εντροπία

$$\text{info}_X(T) = \frac{5}{14} \left[-\frac{2}{5} \log_2 \left(\frac{2}{5} \right) - \frac{3}{5} \log_2 \left(\frac{3}{5} \right) \right] + \frac{4}{14} \left[-\frac{4}{4} \log_2 \left(\frac{4}{4} \right) - \frac{0}{4} \log_2 \left(\frac{0}{4} \right) \right] \\ + \frac{5}{14} \left[-\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) \right] = 0.694 \text{ bits.}$$

Άρα, το κέρδος είναι $0.940 - 0.694 = 0.246 \text{ bits}$.

- Χρησιμοποιώντας το γνώρισμα «άνεμος» το σύνολο T διαμερίζεται σε δύο υποσύνολα με εντροπία $\text{info}_X(T) =$

$$\frac{6}{14} \left[-\frac{3}{6} \log_2 \left(\frac{3}{6} \right) - \frac{3}{6} \log_2 \left(\frac{3}{6} \right) \right] + \frac{8}{14} \left[-\frac{6}{8} \log_2 \left(\frac{6}{8} \right) - \frac{2}{8} \log_2 \left(\frac{2}{8} \right) \right] = 0.892 \text{ bits.}$$

Άρα, το κέρδος πληροφορίας είναι $0.940 - 0.892 = 0.048 \text{ bits}$.

- Συνεπώς, σύμφωνα με το κριτήριο κέρδους η δοκιμασία με το γνώρισμα «ουρανός» είναι προτιμότερη.

❖ Το κριτήριο κέρδους είναι μεροληπτικό υπέρ δοκιμασιών με πολλά (έστω $|T|$) αποτελέσματα. Συγκεκριμένα, σ' αυτήν την περίπτωση η μέση πληροφορία είναι

$$\text{info}_X(T) = - \sum_{i=1}^n \frac{1}{|T|} \text{info}(T_i) = 0 \text{ διότι } \text{info}(T_i) = 0.$$

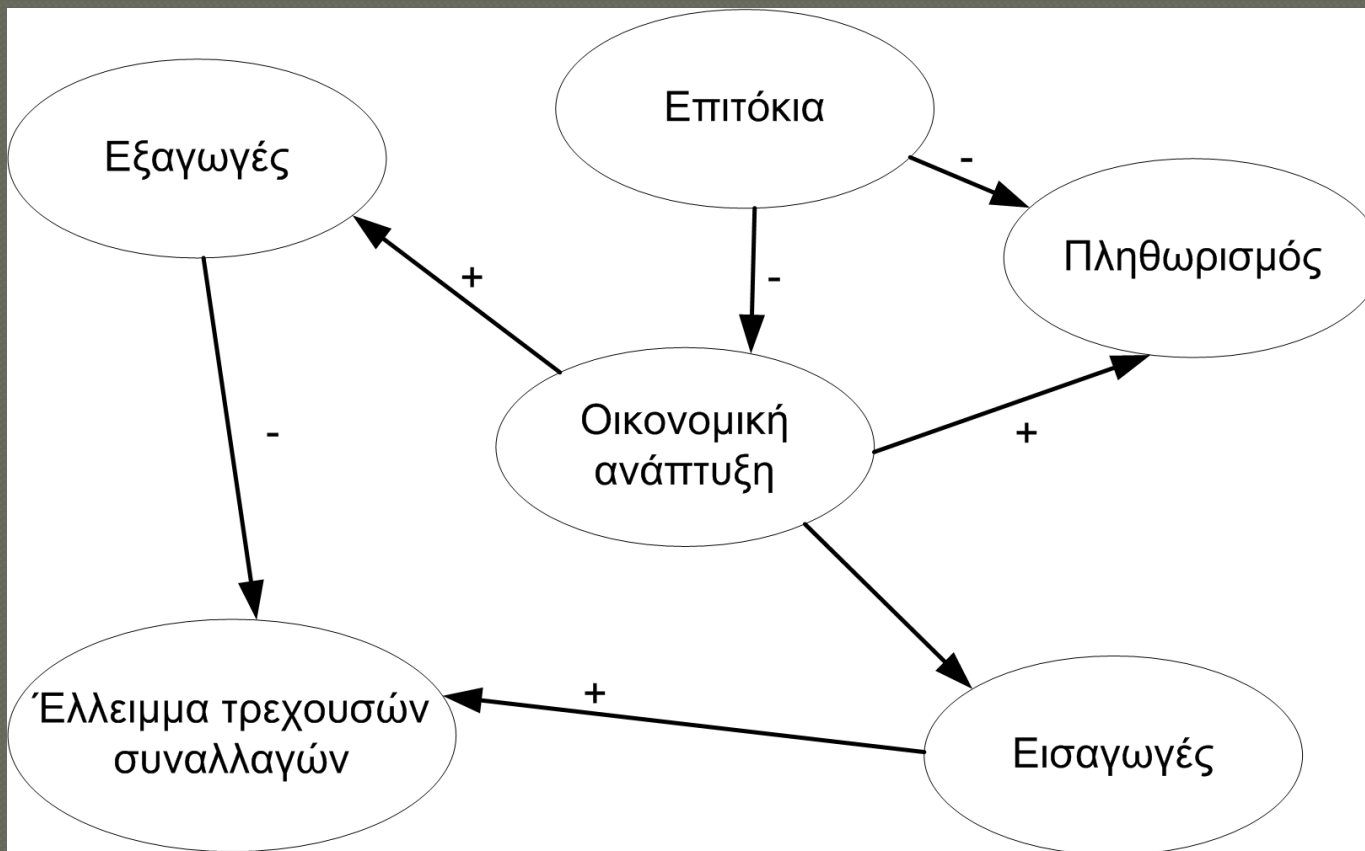
❖ Η προαναφερθείσα μεροληψία του κριτηρίου κέρδους μπορεί να αρθεί με **κανονικοποίηση (normalization)** η οποία επιτυγχάνεται με την ευρετική συνάρτηση **λόγος κέρδους (gain ratio)** $\text{gain ratio}(X) = \text{gain}(X)/\text{splitInfo}(X)$,

$$\text{όπου } \text{splitInfo}(X) = - \sum_{i=1}^n \frac{|T_i|}{|T|} \log_2 \left(\frac{|T_i|}{|T|} \right)$$

Γνωσιακοί Χάρτες

- ❖ Οι γνωσιακοί χάρτες (ΓΧ) είναι προσημασμένοι κατευθυνόμενοι γράφοι για λήψη αποφάσεων.
- ❖ Οι κόμβοι σε αυτούς τους χάρτες αναπαριστούν έννοιες ή μεταβλητές σχετιζόμενοι με κάποιο φυσικό σύστημα.
- ❖ Οι αιτιώδεις σύνδεσμοι μεταξύ των εννοιών του χάρτη ονομάζονται ακμές.
- ❖ Οι ακμές χαρακτηρίζονται από ένα πρόσημο και μία κατεύθυνση που περιγράφουν τον τρόπο με τον οποίο αλληλεπιδρούν οι κόμβοι μεταξύ τους. Συγκεκριμένα, μία ακμή από τον κόμβο Α στον Β μπορεί να είναι είτε θετική είτε αρνητική περιγράφοντας αντίστοιχα την προωθητική ή την ανασταλτική επίδραση του κόμβου Α στον κόμβο Β.

Παράδειγμα Γνωσιακού Χάρτη

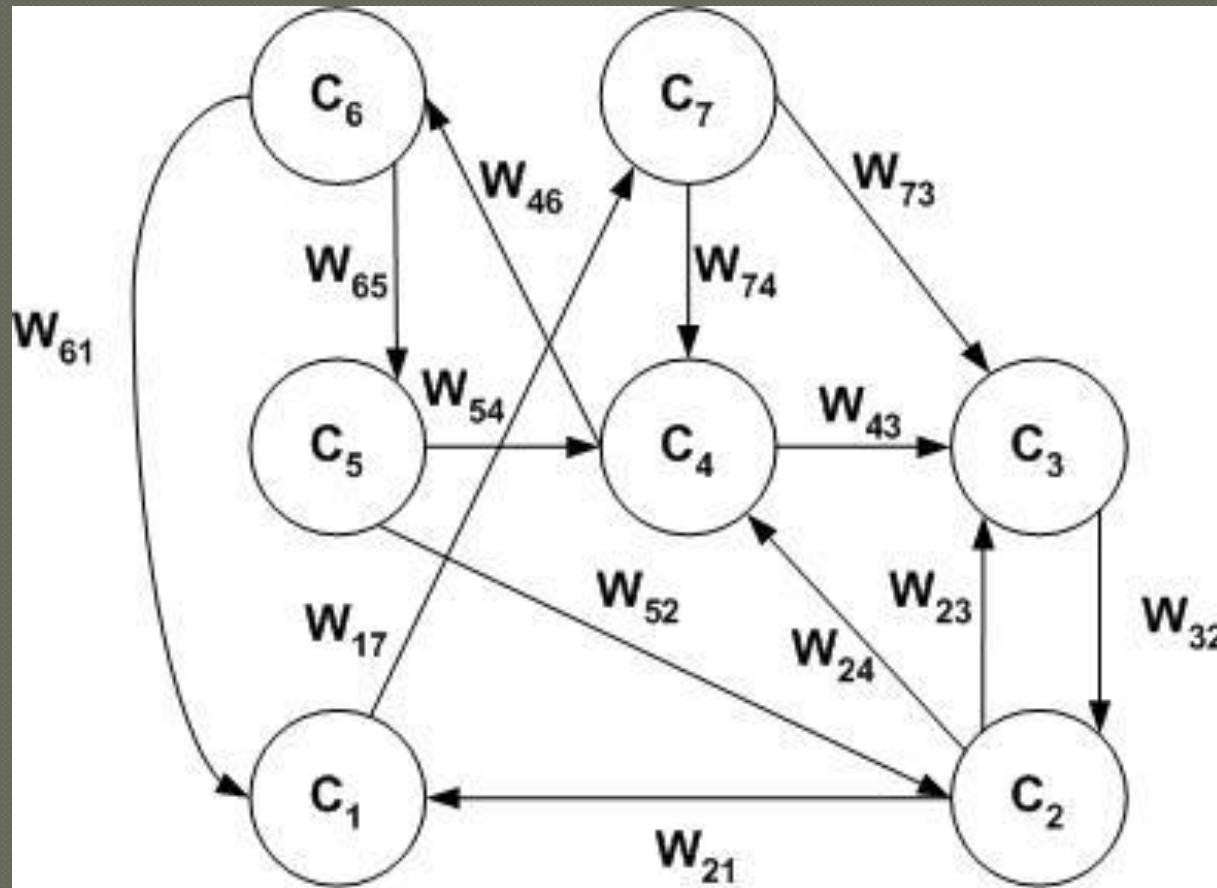


Ασαφείς Γνωσιακοί Χάρτες

Οι ασαφείς γνωσιακοί χάρτες (ΑΓΧ) (**fuzzy cognitive maps (FCM)**) είναι επέκταση των κλασικών γνωσιακών χαρτών με τρεις σημαντικές διαφορές από τους ΓΧ:

1. Οι ακμές μεταξύ των κόμβων μπορούν να λαμβάνουν αριθμητικές τιμές που ονομάζονται *βάρη*.
2. Οι ΑΓΧ μπορούν να μοντελοποιούν δυναμικά συστήματα. Έτσι, οι ΑΓΧ μοιάζουν με τα *ανατροφοδοτούμενα ΤΝΔ*.
3. Όπως τα ΤΝΔ, έτσι και οι ΑΓΧ μπορούν να προσαρμόζουν τα βάρη τους μέσω της διαδικασίας της μάθησης.

Παράδειγμα ΑΓΧ



Ασαφείς Γνωσιακοί Χάρτες

❖ Κάθε χρονική στιγμή η τιμή κάθε έννοιας A_i υπολογίζεται από το άθροισμα των επιρροών όλων των υπολοίπων εννοιών σε αυτή την έννοια και τον περιορισμό της συνολικής επίδρασης με τη χρήση μίας συνάρτησης φραγής f σύμφωνα με τον παρακάτω κανόνα:

$$A_i^{t+1} = f\left(A_i^t + \sum_{i=1, i \neq j} W_{ji} A_j^t\right)$$

Ασαφείς Γνωσιακοί Χάρτες

❖ Σε κάθε βήμα μία νέα κατάσταση των εννοιών προκύπτει μέσω του προαναφερθέντος κανόνα. Μετά από έναν αριθμό επαναλήψεων ο ΑΓΧ μπορεί να καταλήξει σε μία από τις παρακάτω καταστάσεις:

1. Σε ένα συγκεκριμένο σημείο ισορροπίας.
2. Σε έναν περιορισμένο κύκλο.
3. Σε χαοτική συμπεριφορά.

❖ Όταν ο ΑΓΧ καταλήξει σε ένα συγκεκριμένο σημείο ισορροπίας λέμε ότι ο ΑΓΧ έχει συγκλίνει.

Παράρτημα Β

- Στατιστική Αξιολόγηση Μοντέλων ΥΝ.

Απλή Αξιολόγηση

- ❖ Είναι η απλούστερη τεχνική αξιολόγησης, η οποία περιλαμβάνει μόνο έναν υπολογιστικό κύκλο. Συγκεκριμένα, σε κάποιες περιπτώσεις, ένα σύνολο δεδομένων δίδεται διαμερισμένο σε ένα **υποσύνολο εκπαίδευσης** και ένα **υποσύνολο εξέτασης**.
- ❖ Σκοπός είναι η εκπαίδευση ή, ισοδύναμα, μάθηση έτσι ώστε να προκύψει μια αποδεκτή *ικανότητα γενίκευσης*.

Σταυρωτή Αξιολόγηση

❖ Είναι μια τεχνική αξιολόγησης πολλών υπολογιστικών κύκλων, η οποία αποτιμά την ικανότητα γενίκευσης ενός μοντέλου εφαρμόζοντας μια απλή αξιολόγηση ανά υπολογιστικό κύκλο (π.χ. εξαντλητική σταυρωτή αξιολόγηση, άσε-k-έξω σταυρωτή αξιολόγηση).

Στοιχεία Επικοινωνίας

Δρ. Βασίλειος Γ. Καμπουράζος

vgkabs@teiemt.gr

Τηλ. 6945224802

Γραφείο B122 (Κτήριο βιβλιοθήκης)

ΕΥΧΑΡΙΣΤΩ