

νοήματα επιλογής

Χ. Δίνεται η πρόταση: "Όταν ο αλγόριθμος A^* επιλέγει για επεξεργασία έναν κόμβο, ο κόμβος αυτός βρίσκεται στο ελάχιστο μονοπάτι". Για την πρόταση αυτή ισχύει ότι:

Α. Ισχύει για κάθε ευρετική.
 Β. Ισχύει για κάθε αποδεκτή ευρετική.
 Γ. Ισχύει για κάθε συνεπή ευρετική.
 Δ. Η πρόταση αυτή δεν ισχύει.

Συνεπής η $h(k)=0$
 Άρα όχι

C

Αποδεκτές (admissible) ευρετικές: Έστω $h^*(k)$ το κόστος του βέλτιστου μονοπατιού από την k στο στόχο G .

Μία ευρετική $h(k)$ λέγεται αποδεκτή όταν: $0 \leq h(k) \leq h^*(k)$

Θεώρημα: Ο αλγόριθμος A^* είναι πλήρης και βρίσκει τη βέλτιστη λύση για κάθε αποδεκτή ευρετική

Συνεπείς (consistent) ή μονότονες (monotone) ευρετικές: Έστω $h(k)$ μία ευρετική. Έστω k ένας κόμβος και k' ένας επόμενός του. Η $h(k)$ λέγεται συνεπής ή μονότονη όταν:

$h(k) \leq c(k, k') + h(k')$, όπου $c(k, k')$ το κόστος μετάβασης από την k στην k' .

• Ισχύει ότι οι συνεπείς ευρετικές είναι αποδεκτές

Θεώρημα: Ο αλγόριθμος A^* που χρησιμοποιεί συνεπείς ευρετικές, όταν επεκτείνει ένα μονοπάτι, αυτό είναι το βέλτιστο μονοπάτι προς το τελευταίο στοιχείο του

Χ. Δίνονται δύο ευρετικές h_1 και h_2 , για τις οποίες ισχύει ότι $h_1(k) \leq h_2(k)$, για κάθε κόμβο k . Ισχύει ότι:

Α. Η h_1 είναι πιο ακριβής από την h_2 .
 Β. Η h_2 είναι πιο ακριβής από την h_1 .
 Γ. Αν οι h_1, h_2 είναι συνεπείς, τότε η h_1 είναι πιο ακριβής από την h_2 .
 Δ. Αν οι h_1, h_2 είναι συνεπείς, τότε η h_2 είναι πιο ακριβής από την h_1 .

Ακρ. Συν
 όχι συν →

d

Ακρίβεια (accuracy) ευρετικών: Έστω h_1, h_2 δύο συνεπείς ευρετικές τέτοιες ώστε για κάθε k : $h_1(k) \leq h_2(k)$. Τότε η $h_2(k)$ θα λέγεται πιο ακριβής (accurate) από την $h_1(k)$.

Θεώρημα: Έστω h_1, h_2 δύο συνεπείς ευρετικές με h_2 πιο ακριβή από την h_1 . Έστω A^*1 ο αλγόριθμος A^* που χρησιμοποιεί την h_1 και A^*2 ο αλγόριθμος A^* που χρησιμοποιεί την h_2 . Αν υπάρχει λύση, όλοι οι κόμβοι που επεκτείνονται από τον A^*2 επεκτείνονται και από τον A^*1 .

(εκτός ίσως από ορισμένους για τους οποίους $f_1(k) = f_2(k) = c^*$)

3. Ποια από τις παρακάτω προτάσεις σχετικά με τους αλγορίθμους αναζήτησης σε παίγνια δυο αντιπάλων είναι εσφαλμένη;

Α. Ο alpha-beta μπορεί να μειώσει το χρόνο υπολογισμού σε σχέση με τον min-max κλαδεύοντας τμήματα του δέντρου αναζήτησης.
 Β. Ο alpha-beta επιστρέφει πάντοτε την ίδια τιμή με τον min-max για τη ρίζα του δέντρου.
 Γ. Ο alpha-beta επιστρέφει πάντοτε την ίδια τιμή με τον min-max για όλους τους κόμβους του δέντρου.
 Δ. Ο alpha-beta είναι βέλτιστος.

C

Διοτι ο AB δεν επιστρέφει τιμή για όλους τους κόμβους

4. Δίνεται μία γνώση K και έστω p μία πρόταση της Γνωρίζουμε ότι η p συνεπάγεται λογικά κάθε άλλη πρόταση της γνώσης K . Με βάση αυτό μπορούμε να συμπεράνουμε ότι:
- A. Η p είναι ταυτολογία.
 B. Η p είναι αντίφαση.
 Γ. Η p είναι συνεπής, αλλά δεν γνωρίζουμε αν είναι ταυτολογία.
 Δ. Δεν ισχύει καμία από τις άλλες προτάσεις.

$$p = \text{arc} \rightarrow \text{True}$$

$$p = a, K = \{a \vee b, a \vee c\}$$

Ⓐ

Μια πρόταση p **συνεπάγεται** (entails) μια πρόταση q ανν όλες οι ερμηνείες της γνώσης που περιέχει τις δύο προτάσεις και ικανοποιούν την p ικανοποιούν και την q .
 Μία πρόταση p ονομάζεται **ταυτολογία** (tautology) ανν την ικανοποιεί κάθε ερμηνεία της.
 Μία πρόταση p ονομάζεται **συνεπής** (consistent) ανν υπάρχει τουλάχιστον μία ερμηνεία που την ικανοποιεί, αλλιώς ονομάζεται **αντίφαση** (clash).

5. Δίνεται η γνώση $\{[a, b], [-a, c]\}$ και η ανάθεση $\{a = 0, b = 1\}$. Η ανάθεση αυτή είναι:

- A. Ικανοποιητική.
 B. Ψευδοποιητική.
 Γ. Ανεπλήρη.
 Δ. Δεν γνωρίζουμε αν είναι ικανοποιητική, ψευδοποιητική ή ανεπλήρη, αφού δεν είναι πλήρης.

$$0 \vee 1 \rightarrow 1, 1 \vee \dots = 1$$

Ⓐ

Backtracking: $\{[p2, p3], [\neg p1, \neg p4], [\neg p2, p4]\}$

Αναθέσεις (assignments): $\{p1 = \{0 \text{ ή } 1\}\}$ $\{p1 = 0, p2 = 1, p3 = 0, p4 = 0\}$

Μερική ανάθεση: $\{p1 = 0, p3 = 0\}$

Πλήρης ανάθεση: $\{p1 = 0, p2 = 1, p3 = 0, p4 = 0\}$

Ψευδοποιητική ανάθεση: $\{p1 = 1, p4 = 1\}$

Ικανοποιητική ανάθεση: $\{p1 = 0, p2 = 1, p4 = 1\}$

Ανεπλήρη ανάθεση: $\{p1 = 0, p2 = 0, p4 = 1\}$

$$K = \{[a, b], [-a, c]\}$$

$$I = \{a = 0, b = 1\}$$

$$\text{val}(K) = \{[0, 1], [1, 1]\} = 1$$

6. Έστω ένα σύνολο από n διαφορετικές προτάσεις Horn, κάθε μία μεγέθους το πολύ 4. Έστω ότι το σύνολο περιέχει δύο θετικές προτάσεις, μεγέθους το πολύ 2. Το μέγιστο πλήθος διαφορετικών επιλογών εφαρμογής του κανόνα της ανάλυσης είναι:
- A. $2n - 4$.
 B. $2n - 2$.
 Γ. $8n - 16$.
 Δ. $8n - 14$.

$$(\neg a \vee \neg b) \vee (\neg b \vee \neg c) \leftarrow (n-2)$$

$$(\neg a \vee \neg c) \vee (\neg b \vee \neg d) \leftarrow (n-2)$$

Ⓑ

$$\begin{aligned} & \{[a, b, c, d], \\ & [a, b, e, z]\} \quad n-2 \text{ αρνητικές} \\ & \{[a, b], [b, a]\} \quad +2 \text{ θετικές} \end{aligned}$$

συνολικές αναλύσεις:

$$(n-2) + (n-2) + 2 = 2n-2$$

7. Πόσες πιθανότητες πρέπει να γνωρίζουμε κατ' ελάχιστο ώστε το δίκτυο πίστης (belief network) της εικόνας να είναι πλήρως προσδιορισμένο;
- A. 9
 B. 4
 Γ. 14
 Δ. 2

εκτός 9πώς?
Ⓐ

$$p_i: \# \text{ πιθανοτήτων} = 2^{\# \text{ κόμβων}}$$

$$2^0 + 2^1 + 2^1 + 2^1 = 9$$

$$\begin{matrix} p_1 & p_2 & p_3 & p_4 \\ \downarrow & \downarrow & \downarrow & \downarrow \end{matrix}$$

8. Η επαγωγική μεροληψία ενός ταξινομητή σχετίζεται:
- Α. Μόνο με το σύνολο δεδομένων εκπαίδευσης.
 - Β. Μόνο με το σύνολο υποθέσεων.
 - Γ. Με το σύνολο υποθέσεων και τον ιδανικό ταξινομητή (όχι με το σύνολο δεδομένων εκπαίδευσης).
 - Δ. Με το σύνολο υποθέσεων, τον ιδανικό ταξινομητή και το σύνολο δεδομένων εκπαίδευσης.

(b)

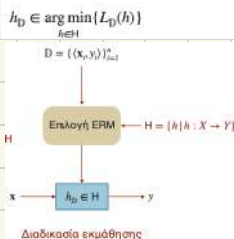
50%

Η επιλογή της μορφής του h είναι προκαθορισμένη, καθορίζεται από το σύνολο H (σύνολο υποθέσεων). Συνεπώς, κάθε περιορισμός του συνόλου H λειτουργεί μεροληπτικά για την μορφή που τελικά θα έχει η h .

Το φαινόμενο αυτό αναφέρεται ως **επαγωγική μεροληψία (inductive bias)** και γενικά αυξάνει την πιθανότητα σφάλματος

Για να μην επηρεάζεται σημαντικά το σφάλμα από την επαγωγική μεροληψία, πρέπει η επιλογή της μορφής των υποθέσεων να βασίζεται σε πρότερη γνώση για το πρόβλημα.

$H = \{h : h : X \rightarrow Y\}$: σύνολο ταξινομητών $X \rightarrow Y$



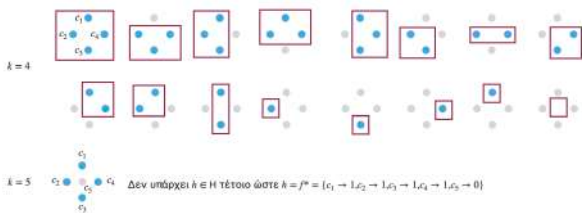
9. Για έναν ταξινομητή σε χώρο δύο διαστάσεων που χρησιμοποιεί το σύνολο υποθέσεων όλων των παράλληλων στους άξονες, παραλληλόγραμμων ισχύει ότι:
- Α. Είναι PAC learnable αφού τα παράλληλα στους άξονες παραλληλόγραμμα έχουν διάσταση VC $k = 4$.
 - Β. Είναι PAC learnable αφού τα παράλληλα στους άξονες παραλληλόγραμμα έχουν διάσταση VC $k = 5$.
 - Γ. Είναι PAC learnable αφού τα παράλληλα στους άξονες παραλληλόγραμμα έχουν άπειρη διάσταση VC.
 - Δ. Δεν ισχύει καμία από τις άλλες προτάσεις.

(a)

Παράδειγμα (παράλληλα στους άξονες παραλληλόγραμμα)

$H = \{h(a_1, a_2, b_1, b_2) : a_1 \leq a_2 \text{ και } b_1 \leq b_2\}$, όπου $h(a_1, a_2, b_1, b_2) = 1$ αν και μόνο αν $a_1 \leq x_1 \leq a_2$ και $b_1 \leq x_2 \leq b_2$

Αν θεωρήσουμε $C = \{c_1, c_2, \dots, c_5\}$, ποιο το μέγιστο k για το οποίο για κάθε $f : C \rightarrow \{0,1\}$ υπάρχει ένα $h \in H$ τέτοιο ώστε $h = f$?



10. Στην ενισχυτική μάθηση, η εκτιμώμενη ανταμοιβή μιας ενέργειας α στον αλγόριθμο Upper Confidence Bound (UCB) προσοφάνεται κατά μια τιμή, η οποία:

- Α. Είναι ανάλογη του πλήθους των φορών τις οποίες έχουμε εκτελέσει τη συγκεκριμένη ενέργεια στο παρελθόν.
- Β. Είναι μεγαλύτερη για ενέργειες με μικρότερη εκτιμώμενη ανταμοιβή.
- Γ. Αυξάνεται με την αβεβαιότητα ως προς την εκτιμώμενη τιμή της ανταμοιβής.
- Δ. Αποβαρύνει την εξερεύνηση.

(c)

ΕΚΤΟΣ!

Ιδέα: Τη στιγμή ! επέλεξε την ενέργεια " με βάση το συνδυασμό της έως τώρα εκτιμώμενης ανταμοιβής $\mu_t(a)$ και της αβεβαιότητας της εκτίμησής μας η οποία προκύπτει από το πόσες φορές έχουμε εκτελέσει αυτή την ενέργεια στο παρελθόν.

Αν η ενέργεια " δεν έχει δοκιμαστεί πολλές φορές μέχρι τώρα, τότε δεν είμαστε πολύ βέβαιοι για την εκτίμηση $\mu_t(a)$. Στην περίπτωση αυτή δίνουμε μεγάλο "bonus" στην ενέργεια αυτή αυξάνοντας την πιθανότητα να την επιλέξουμε. Έτσι, διευκολύνουμε την εξερεύνηση.

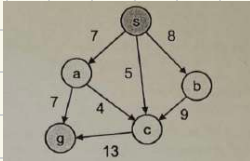
Το bonus είναι αντιστρόφως ανάλογο του μετρητή $\mu_t(a)$: όσες πιο πολλές φορές έχουμε εκτελέσει την ενέργεια " τόσο λιγότερο αβέβαιοι είμαστε για την εκτίμησή μας.

Αν μια ενέργεια έχει δοκιμαστεί πολλές φορές και έχει μικρή εκτιμώμενη ανταμοιβή $\mu_t(a)$ τότε δεν θα επιλεγεί.

Το bonus ενθαρρύνει την εξερεύνηση αλλά δεν πρέπει να είναι τόσο μεγάλο ώστε να επιλέγουμε ενέργειες που δεν έχουν καμία ελπίδα να είναι βέλτιστες.

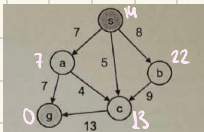
Θεμα 1

Ζητούμενο 1

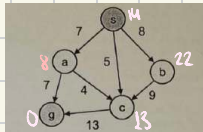


1. Δίνεται το παρακάτω γράφημα που αναπαριστά ένα πρόβλημα με αφετηρία την κατάσταση s και τερματισμό την κατάσταση g . Στο γράφημα αυτό εμφανίζεται το πραγματικό κόστος μετάβασης από μία κατάσταση σε μία άλλη.

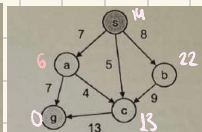
(α') Να βρείτε τις εξής μη τετρωμένες (να μην ταυτίζονται με τη βέλτιστη, ούτε να είναι μηδενική) ευρετικές, ακεραίων τιμών, την h_1 που δεν είναι αποδεκτή, την h_2 που είναι αποδεκτή αλλά όχι συνεπής, την h_3 που είναι αποδεκτή και συνεπής και την h_4 που είναι πιο ακριβής από την h_3 . Για κάθε μία από τις ευρετικές, να γράψετε τις συνθήκες που διασφαλίζουν ή παραβιάζουν τις αντίστοιχες ιδιότητες των ευρετικών.



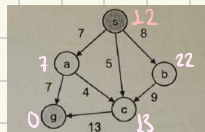
h_1 : βέλτιστη / πραγματικό



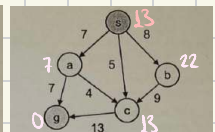
h_2 : μη αποδεκτή



h_2 : αποδεκτή αλλά όχι συνεπής



h_3 : συνεπής



h_4 : πιο ακριβής από h_3

h_1 : $h_1(a) = 7 < h^*(a) = 7 \rightarrow$ μη αποδεκτή

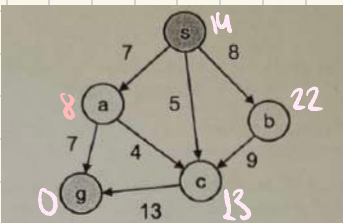
h_2 : $h_2(s) = 14 \leq h^*(s) = 14 \rightarrow$ αποδεκτή

$h_2(s) = 14 > c(s,a) + h_2(a) = 7 + 6 = 13 \rightarrow$ μη συνεπής

h_3 : συνεπής

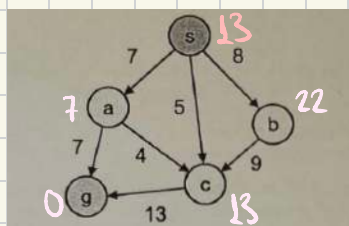
h_4 : $h_4(s) = 12 \leq h_3(s) = 13 \rightarrow$ πιο ακριβής

(β') Να εκτελέσετε τον αλγόριθμο A^* , χρησιμοποιώντας εναλλακτικά τις ευρετικές h_1 και h_4 , δίνοντας σε όλα τα βήματα, το περιεχόμενο του μετώπου αναζήτησης και του κλειστού συνόλου. Τι παρατηρείτε για τις λύσεις σε κάθε περίπτωση;



→ επιτυχία:
sag (κόστος=14)

frontier	closed	current	child nodes
1. $(s, 0, 14)^s$	{}	s	$(a, 7, 15)$ $(b, 8, 30)$ $(c, 5, 18)$
2. $(a, 7, 15)^{sa}$ $(c, 5, 18)^{sc}$ $(b, 8, 30)^{sb}$	{(s, 0)}	a	$(c, 11, 24)$ $(g, 14, 14)$
3. $(g, 14, 14)^{sag}$ $(c, 5, 18)^{sc}$ $(b, 8, 30)^{sb}$	{(s, 0), (a, 7)}	g	-



→ επιτυχία:
sag (κόστος = 14)

	frontier	closed	current	child nodes
1.	$(s, 0, 13)^s$	{ }	s	$(a, 7, 14)$ $(b, 8, 30)$ $(c, 5, 18)$
2.	$(a, 7, 14)^{sa}$ $(c, 5, 18)^{sc}$ $(b, 8, 30)^{sb}$	{(s, 0)}	a	$(c, 11, 24)$ $(g, 14, 14)$
3.	$(g, 14, 14)^{sag}$ $(c, 5, 18)^{sc}$ $(b, 8, 30)^{sb}$	{(s, 0), (a, 7)}	g	-

Ζητούμενο 2

2. Δίνεται ο ακόλουθος πληθυσμός κωδικοποιημένων λύσεων ενός προβλήματος στη γενιά-0 (G_0):

(1) 11011, (2) 01011, (3) 11001, (4) 10111

Οι υποψηφίες λύσεις αξιολογούνται με χρήση της ακόλουθης συνάρτησης απόδοσης (fitness function): $f = \sum_{i=0}^4 (1 - \frac{1}{2^{b_i+1}})$ όπου b_i είναι το i -στό bit της συμβολοσειράς (με $i = 0$ αναπαριστούμε το λιγότερο σημαντικό bit και με $i = 4$ το περισσότερο σημαντικό bit). Εφαρμόζεται γενετικός αλγόριθμος με ρυθμό μετάλλαξης ίσο με 0 και διασταύρωση μονού σημείου.

Δίνεται επίσης γεννήτρια τυχαίων αριθμών (ρουλέτα) με εξόδους: 0.25 0.73 0.15 0.52 0.81 0.65. Οι εξοδοί αυτές χρησιμοποιούνται κατά σειρά σε κάθε διασταύρωση για την επιλογή: αρχικά των γονέων και στη συνέχεια του εκάστοτε σημείου διασταύρωσης. Συγκεκριμένα, το καθένα από τα 4 πιθανά σημεία διασταύρωσης επιλέγεται, από δεξιά προς τα αριστερά, με ίση πιθανότητα χρησιμοποιώντας την εκάστοτε έξοδο της παραπάνω γεννήτριας. Για παράδειγμα, ο τυχαίος αριθμός 0.22 συνεπάγεται σημείο διασταύρωσης μεταξύ του bit-0 και του bit-1, ο τυχαίος αριθμός 0.45 συνεπάγεται σημείο διασταύρωσης μεταξύ του bit-1 και του bit-2, κ.ο.κ.

Η νέα γενιά προκύπτει με πλήρη αντικατάσταση της προηγούμενης, δηλαδή τα δύο παιδιά που προκύπτουν από μία διασταύρωση αντικαθιστούν τους γονείς στη νέα γενιά, επομένως χρειάζονται ακριβώς δύο διασταυρώσεις για να δημιουργηθεί η γενιά-1 (G_1).

$V_1 = 11011$

$V_2 = 01011$

$V_3 = 11001$

$V_4 = 10111$

fitness function:

$$f = \sum_{i=0}^4 \frac{1 - (-1)^{b_i+1}}{2}$$

(α') Να υπολογίσετε την απόδοση (fitness) κάθε ατόμου της G_0 , καθώς και τη μέση συνολική απόδοση της G_0 .

$eval(V_1) = f(V_1 = 11011) = 1 + 0 + 0 + 0 + 1 = 2$ - $p_1 = 0,2$ - $q_1 = 0,2$

$eval(V_2) = f(V_2 = 01011) = 1 + 0 + 0 + 0 + 0 = 1$ - $p_2 = 0,1$ - $q_2 = 0,3$

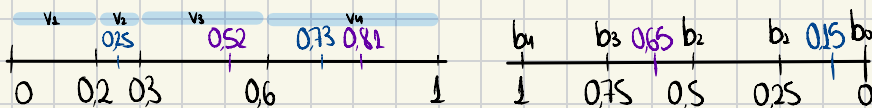
$eval(V_3) = f(V_3 = 11001) = 1 + 1 + 0 + 0 + 1 = 3$ - $p_3 = 0,3$ - $q_3 = 0,6$

$eval(V_4) = f(V_4 = 10111) = 1 + 0 + 1 + 1 + 1 = 4$ - $p_4 = 0,4$ - $q_4 = 1$

συνολική απόδοση πληθυσμού: $f = \sum_{i=1}^4 eval(V_i) = 2 + 1 + 3 + 4 = 10$ | πιθανότητα επιλογής: $p_i = \frac{eval(p_i)}{f}$ | αθροιστική πιθανότητα: $q_i = \sum_{j=1}^i p_j$

(β') Να προσδιορίσετε τα άτομα της G_1 , εφαρμόζοντας τη μέθοδο της ρουλέτας.

γονείς	σημείο	γονείς	σημείο
διασταύρωση 1: 0,25 0,73	0,15	$V_2 \ V_4$	$b_1 - b_2$
διασταύρωση 2: 0,52 0,81	0,65	$V_3 \ V_4$	$b_2 - b_3$



διασαύρωση 1:

$$\begin{array}{l} V_2: 0101 \mid 1 \\ V_4: 1011 \mid 1 \end{array} \rightarrow \begin{array}{l} V_1': 01011 \\ V_2': 10111 \end{array}$$

διασαύρωση 2:

$$\begin{array}{l} V_3: 11001 \\ V_4: 10111 \end{array} \rightarrow \begin{array}{l} V_3': 11111 \\ V_4': 10001 \end{array}$$

$$\Rightarrow G_1: \{01011, 10111, 11111, 10001\}$$

(γ') Να υπολογίσετε την απόδοση (fitness) κάθε ατόμου της G_1 , καθώς και τη μέση συνολική απόδοση της G_1 .

$$\text{eval}(V_1') = f(V_1' = 01011) = 1$$

$$\text{eval}(V_2') = f(V_2' = 10111) = 4$$

$$f = \sum_{i=1}^4 \text{eval}(V_i') = 1 + 4 + 3 + 4 = 12$$

$$\text{eval}(V_3') = f(V_3' = 11111) = 3$$

$$\text{eval}(V_4') = f(V_4' = 10001) = 4$$

(δ') Να εξηγήσετε ποιες συμβολοσειρές ευνοεί (επιβραβεύει) η παραπάνω συνάρτηση απόδοσης f .

$f = \sum_{i=0}^4 \frac{1 - (-1)^{b_{i+1}}}{2}$ λαμβάνει μέγιστη τιμή για $f(10101) = 5$ συνεπώς ευνοεί συμβολοσειρές που πλησιάζουν την 10101

Θέμα 2

Δίνεται μια γνώση $\mathcal{K}$ που αποτελείται από τις παρακάτω προτάσεις:

- $\forall x. (C(x) \Rightarrow \exists y. (r(x, y) \wedge D(y)))$
- $\forall x. (E(x) \Rightarrow (C(x) \wedge A(x)))$
- $\forall x. \forall y. ((C(x) \wedge A(x) \wedge r(x, y) \wedge D(y)) \Rightarrow \exists z. (s(y, z) \wedge B(z)))$
- $\forall x. \forall y. \forall z. ((r(x, z) \wedge s(z, y) \wedge B(y)) \Rightarrow G(x))$
- $E(a)$

1. Να βρείτε τρία μοντέλα για την γνώση $\mathcal{K}$, τα οποία έχουν πεδίο ερμηνείας ένα σύνολο που περιέχει 1, 2 και 3 στοιχεία, αντίστοιχα, το καθένα από αυτά. Να τροποποιήσετε όποια από τα αξιώματα επιθυμείτε, ώστε η γνώση να μην έχει μοντέλα με πεδίο ερμηνείας ένα σύνολο που περιέχει λιγότερα από 3 στοιχεία.

$$\underline{\text{CNF}}: \{ [\neg C(x), r(x, f(x))]_1, [\neg C(x), D(f(x))]_2, [\neg E(x), C(x)]_3, [\neg E(x), A(x)]_4,$$

$$[\neg C(x), \neg A(x), \neg r(x, y), \neg D(y), s(y, f(x, y))]_5, [\neg C(x), \neg A(x), \neg r(x, y), \neg D(y), B(f(x, y))]_6,$$

$$[\neg r(x, z), \neg s(z, y), \neg B(y), G(x)]_7, [E(a)]_8 \}$$

$$\underline{\text{CNF}}: \{ [\neg C(x), r(x, f(x))]_1, [\neg C(x), D(f(x))]_2, [\neg E(x), C(x)]_3, [\neg E(x), A(x)]_4,$$

$$[\neg C(x), \neg A(x), \neg r(x, y), \neg D(y), s(y, f(x, y))]_5, [\neg C(x), \neg A(x), \neg r(x, y), \neg D(y), B(f(x, y))]_6,$$

$$[\neg r(x, z), \neg s(z, y), \neg B(y), G(x)]_7, [E(a)]_8 \}$$

$$\sim D = \{a\}:$$

$$\textcircled{1} E(a) \rightarrow \neg$$

$$\textcircled{3} C(a) \rightarrow \neg$$

$$\textcircled{4} A(a) \rightarrow \neg$$

$$\textcircled{1} r(a, f(a)) \rightarrow \neg$$

$$\textcircled{2} D(f(a)) \rightarrow \neg$$

$$\textcircled{5} s(a, f(a, a)) \rightarrow \neg$$

$$\textcircled{6} B(f(a, a)) \rightarrow \neg$$

$$\textcircled{7} G(a) \rightarrow \neg$$

$$\sim D = \{a, b\}$$

$$\textcircled{1} E(a) \rightarrow \neg$$

$$\textcircled{3} C(b) \rightarrow \neg$$

$$\textcircled{4} A(a) \rightarrow \neg$$

$$\textcircled{1} r(b, f(a, b)) \rightarrow \neg$$

$$\textcircled{2} D(f(a)) \rightarrow \neg$$

$$\textcircled{5} s(a, f(a, b)) \rightarrow \neg$$

$$\textcircled{6} B(f(a, b)) \rightarrow \neg$$

$$\textcircled{7} G(a) \rightarrow \neg$$

$$D = \{a, b, c\}$$

$$\textcircled{1} E(a) \rightarrow \neg$$

$$\textcircled{3} C(b) \rightarrow \neg$$

$$\textcircled{4} A(c) \rightarrow \neg$$

$$\textcircled{1} r(a, f(a)) \rightarrow \neg$$

$$\textcircled{2} D(f(c)) \rightarrow \neg$$

$$\textcircled{5} s(a, f(a, c)) \rightarrow \neg$$

$$\textcircled{6} B(f(a, c)) \rightarrow \neg$$

$$\textcircled{7} G(a) \rightarrow \neg$$

2. Να αποδείξετε, με χρήση του αλγόριθμου της ανάλυσης, ότι $\mathcal{K} \models G(a)$.

$$\text{Έστω } K' = K \cup \{\neg G(a)\} \Rightarrow K' = \{K, [\neg G(a)]_9\}$$

$$\textcircled{1} + \textcircled{4} \xrightarrow{x=a} [\neg r(a, z), \neg s(z, y), \neg B(y)]_6 \xrightarrow{y=f(a)} [\neg s(f(a), y), \neg B(y), \neg C(a)]_8 \xrightarrow{x=a} [\neg s(f(a), y), \neg B(y), \neg E(a)]_2 \xrightarrow{y=f(a)} [\neg s(f(a), y), \neg B(y)]_6$$

$$[\neg s(f(a), y), \neg B(y)]_6 \xrightarrow{y=f(a)} [\neg B(f(a)), \neg C(a), \neg A(x), \neg r(x, f(a)), \neg D(f(a))]_4 \xrightarrow{f(x,y)=f(a)} [\neg C(a), \neg A(x), \neg r(x, f(a)), \neg D(f(a))]_5$$

$$\xrightarrow{x=a} [\neg C(a), \neg A(a), \neg r(a, f(a))]_6 \xrightarrow{y=f(a)} [\neg C(a), \neg E(a), \neg r(a, f(a))]_4 \xrightarrow{x=a} [\neg E(a), \neg r(a, f(a))]_2 \xrightarrow{y=f(a)} [\neg E(a), \neg C(a)]_4 \xrightarrow{x=a} [\neg E(a)]_8$$

$$[\neg E(a)]_8 \xrightarrow{y=f(a)} [\neg E(a)]_8 \Rightarrow \text{όχι } K \models G(a)$$

2. Να αποδείξετε, με χρήση του αλγόριθμου της ανάλυσης, ότι $K \models G(a)$. $\rightarrow K' = KU \{ \neg G(a) \}$

$$K' = \{ [\neg C(x), r(x, f(x))]_1, [\neg C(x), D(f(x))]_2, [\neg E(x), C(x)]_3, [\neg E(x), A(x)]_4,$$

$$[\neg C(x), \neg A(x), \neg r(x, y), \neg D(y), s(y, f(x, y))]_5, [\neg C(x), \neg A(x), \neg r(x, y), \neg D(y), B(f(x, y))]_6,$$

$$[\neg r(x, z), \neg s(z, y), \neg B(y), G(x)]_7, [E(a)]_8, [\neg G(a)]_9 \}$$

$$x=a: \textcircled{7}: [\neg r(a, z), \neg s(z, y), \neg B(y), G(a)] \rightarrow [\neg r(a, z), \neg s(z, y), \neg B(y)]_{10}$$

$$\textcircled{9}: [\neg G(a)]$$

$$x=a: \textcircled{1}: [\neg C(a), r(a, f(a))] \rightarrow [\neg C(a), \neg s(f(a), y), \neg B(y)]_{11}$$

$$z=f(a): \textcircled{10}: [\neg r(a, f(a)), \neg s(f(a), y), \neg B(y)]$$

$$x=a: \textcircled{3}: [\neg E(a), C(a)] \rightarrow [\neg E(a), \neg s(f(a), y), \neg B(y)]_{12}$$

$$\textcircled{11}: [\neg C(a), \neg s(f(a), y), \neg B(y)]$$

$$x=a: \textcircled{4}: [\neg E(a), A(a)] \rightarrow [A(a)]_{13} \quad \textcircled{3}: [\neg E(a), C(a)] \rightarrow [C(a)]_{14}$$

$$\textcircled{8}: [E(a)]$$

$$\textcircled{8}: [E(a)]$$

$$x=a: \textcircled{1}: [\neg C(a), r(a, f(a))] \rightarrow [r(a, f(a))]_{15} \quad \textcircled{2}: [\neg C(a), D(f(a))] \rightarrow [D(f(a))]_{16}$$

$$\textcircled{19}: [C(a)]$$

$$\textcircled{19}: [C(a)]$$

$$x=(a): \textcircled{5}: [\neg C(a), \neg A(a), \neg r(a, f(a)), \neg D(f(a)), s(f(a), f(a, f(a)))] \rightarrow [s(f(a), f(a, f(a)))]_{17}$$

$$y=f(a): \textcircled{19}: [C(a)] \quad \textcircled{15}: [r(a, f(a))]$$

$$\textcircled{13}: [A(a)]$$

$$\textcircled{16}: [D(f(a))]$$

$$x=a: \textcircled{6}: [\neg C(a), \neg A(a), \neg r(a, f(a)), \neg D(f(a)), B(f(a, f(a)))] \rightarrow [B(f(a, f(a)))]_{18}$$

$$y=f(a): \textcircled{19}: [C(a)] \quad \textcircled{15}: [r(a, f(a))]$$

$$\textcircled{13}: [A(a)]$$

$$\textcircled{16}: [D(f(a))]$$

$$x=a: \textcircled{7}: [\neg r(a, f(a)), \neg s(f(a), f(f(a))), \neg B(f(f(a))), G(a)] \rightarrow [G(a)]_{19}$$

$$y=f(f(a)): \textcircled{15}: [r(a, f(a))] \quad \textcircled{17}: s(f(a), f(f(a))) \quad \textcircled{18}: [B(f(f(a)))]$$

$$\textcircled{9}: [G(a)] \rightarrow \text{αντίφαση} \Rightarrow K \models G(a)$$

$$\textcircled{19}: [\neg G(a)]$$

Θέμα 3

Δίνεται το σύνολο δεδομένων D με στοιχεία που ανήκουν σε μία από δύο κλάσεις A και B :
 $D = \{[(0, 1, 2), A], [(4, 0, -1), B], [(1, 2, 0), A], [(3, -1, 0), B], [(-2, 1, -2), A]\}$.

1. Να βρείτε έναν ταξινομητή perceptron με μηδενικό εμπειρικό σφάλμα για το συγκεκριμένο πρόβλημα. Σε ποια κλάση θα ταξινομήσει το συγκεκριμένο perceptron το στοιχείο $(3, 1, 2)$;

$\vec{x}_1 = (0, 1, 2) \rightarrow y_1 = A$ Perceptron με 3 εισόδους.

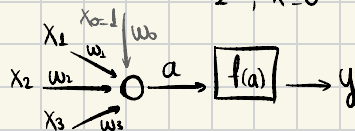
$\vec{x}_2 = (4, 0, -1) \rightarrow y_2 = B$ συνάρτηση ενεργητικότητας:

$\vec{x}_3 = (1, 2, 0) \rightarrow y_3 = A$

$\vec{x}_4 = (3, -1, 0) \rightarrow y_4 = B$

$\vec{x}_5 = (-2, 1, -2) \rightarrow y_5 = A$

$$f(x) = \text{sign}(x) = \begin{cases} 1, & x > 0 \rightarrow \text{κλάση A} \\ -1, & x \leq 0 \rightarrow \text{κλάση B} \end{cases}$$



$$a = \sum_{i=1}^3 x_i \cdot w_i \text{ και } y = f(a) = f(\sum_{i=1}^3 x_i w_i)$$

$$\langle \vec{x}_1, y_1 \rangle \Rightarrow f(a(0, 1, 2)) \geq 0 : w_0 + 0w_1 + 1w_2 + 2w_3 \geq 0 \Rightarrow w_2 \geq -2w_3 - w_0$$

$$\langle \vec{x}_2, y_2 \rangle \Rightarrow f(a(4, 0, -1)) < 0 : w_0 + 4w_1 + 0w_2 - 1w_3 < 0 \Rightarrow w_3 > 4w_1 + w_0$$

$$\langle \vec{x}_3, y_3 \rangle \Rightarrow f(a(1, 2, 0)) \geq 0 : w_0 + 1w_1 + 2w_2 + 0w_3 \geq 0 \Rightarrow w_1 \geq -2w_2 - w_0$$

$$\langle \vec{x}_4, y_4 \rangle \Rightarrow f(a(3, -1, 0)) < 0 : w_0 + 3w_1 - w_2 + 0w_3 < 0 \Rightarrow w_2 > 3w_1 - w_0$$

$$\langle \vec{x}_5, y_5 \rangle \Rightarrow f(a(-2, 1, -2)) \geq 0 : w_0 - 2w_1 + w_2 - 2w_3 \geq 0 \Rightarrow -2w_1 + w_2 - 2w_3 \geq 0$$

$$\Rightarrow w_0 = 0, w_1 = -1, w_2 = 2, w_3 = 1 \Rightarrow a(\vec{x}) = x_1 + 2x_2 + x_3$$

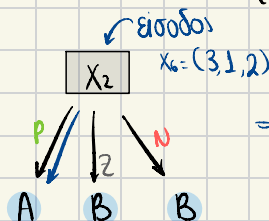
στοιχείο $(3, 1, 2)$: $a = -1 + 2 + 1 = 2$ και $f(a) = f(2) = 1 \rightarrow$ κλάση A

2. Να σχεδιάσετε, με βάση τον αλγόριθμο CART, ένα δέντρο αποφάσεων, θεωρώντας τρεις κατηγορίες τιμών για τα χαρακτηριστικά: για θετικές τιμές την κατηγορία P, για αρνητικές τιμές την κατηγορία N και για την τιμή μηδέν την κατηγορία Z. Σε ποια κλάση θα ταξινομήσει το συγκεκριμένο δέντρο αποφάσεων το στοιχείο $(3, 1, 2)$;

	x_1	x_2	x_3	gt
$\vec{x}_1$	0	1	3	A
$\vec{x}_2$	4	0	-1	B
$\vec{x}_3$	1	2	0	A
$\vec{x}_4$	3	-1	0	B
$\vec{x}_5$	-2	1	-2	A

	x_1	x_2	x_3	gt
$\vec{x}_1$	Z	P	P	A
$\vec{x}_2$	P	Z	N	B
$\vec{x}_3$	P	P	Z	A
$\vec{x}_4$	P	N	Z	B
$\vec{x}_5$	N	P	N	A

παρατηρούμε ότι μπορούμε να οδηγηθούμε σε απόφαση με w του x_2 καθώς

$$\begin{cases} x_2 = P \rightarrow \vec{x}_i \in A \\ x_2 = N \vee Z \rightarrow \vec{x}_i \in B \end{cases}$$


$\Rightarrow$ το ταξινομήσει στην κλάση A