ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

3η Σειρά Ασκήσεων

Θεμελιώδη Θέματα Επιστήμης Υπολογιστών

3ο ΕΞΑΜΗΝΟ

ΔΟΝΤΗ ΕΙΡΗΝΗ
Α.Μ.: 03119839

ΑΘΗΝΑ 2020-2021

<u>Άσκηση 1</u>

1.

Αλγόριθμος Αναρρίχησης Λόφων ή Hill Climbing:

Βήμα 1:

Όρισε τον τρέχοντα κόμβο ως τη ρίζα του δέντρου.

Βήμα 2:

Μέχρι που ο τρέχων κόμβος δεν είναι κόμβος στόχος, εκτέλεσε:

Βήμα $2^{α}$:

Βρες τα παιδιά του τρέχοντος κόμβου, και στη συνέχεια βρες αυτό με την ελάχιστη υπολογιζόμενη υπόλοιπη απόσταση από το στόχο.

Βήμα $2^{β}$:

Εάν ο τρέχων κόμβος δεν έχει παιδιά ή το παιδί που βρέθηκε στο Βήμα $2^{α}$ έχει μεγαλύτερη τιμή ευριστικής από αυτόν πήγαινε στο Βήμα 3.

Βήμα $2^{γ}$:

Όρισε τον κόμβο που βρέθηκε στο Βήμα $2^{α}$ ως τρέχων κόμβο.

Βήμα 3:

Εάν βρήκαμε ένα κόμβο στόχο, τότε ανακοινώνουμε επιτυχία αλλιώς ανακοινώνουμε αποτυχία.

Με αρχικό κόμβο τον κόμβο s, **εκτελούμε** τον παραπάνω αλγόριθμο:

Βήμα 1:

Θεωρούμε τρέχων κόμβο τον s.

Βήμα 2:

Τα παιδιά του κόμβου s είναι οι κόμβοι b(6), c(4) και d(5). Οι τιμές στις παρενθέσεις είναι η τιμή της ευριστικής συνάρτησης για τον εκάστοτε κόμβο.

Ο κόμβος c είναι ο κόμβος - παιδί με τη μικρότερη τιμή ευριστικής που είναι μικρότερη από την τιμή της ευριστικής στο s. Οπότε, ορίζουμε τον κόμβο c ως τρέχων κόμβο.

Έχοντας ως τρέχων κόμβο τον c, παρατηρούμε τα εξής:

Τα παιδιά του κόμβου c είναι οι κόμβοι h(5) και k(2). Ο κόμβος k είναι ο κόμβος - παιδί με τη μικρότερη τιμή ευριστικής που είναι μικρότερη από την τιμή της ευριστικής στο s. Οπότε, ορίζουμε τον κόμβο k ως τρέχων κόμβο.

Έχοντας ως τρέχων κόμβο τον k, παρατηρούμε τα εξής:

Τα παιδιά του κόμβου k είναι οι κόμβοι g(0) και h(5). Ο κόμβος g είναι ο κόμβος - παιδί με τη μικρότερη τιμή ευριστικής που είναι μικρότερη από την τιμή της ευριστικής στο s. Οπότε, ορίζουμε τον κόμβο g ως τρέχων κόμβο.

Βήμα 3:

Ο κόμβος g δεν έχει παιδιά, οπότε πρέπει να αποφανθούμε την επιτυχία ή την αποτυχία.

Παρατηρούμε ότι ο κόμβος g είναι ένας κόμβος στόχος, όπότε ο αλγόριθμος ανακοινώνει **επιτυχία**.

Συνοπτικά μπορούμε, τα παραπάνω βήματα, να τα συμπυκνώσουμε σε έναν πίνακα:

Βήμα	**Μέτωπο Αναζήτησης**	**Κλειστό Σύνολο**
1	[s]	[]
2	[c]	[s]
3	[k]	[s,c]
4	Τέλος Αλγορίθμου	[s,c,k,g]

Αλγόριθμος Α*:

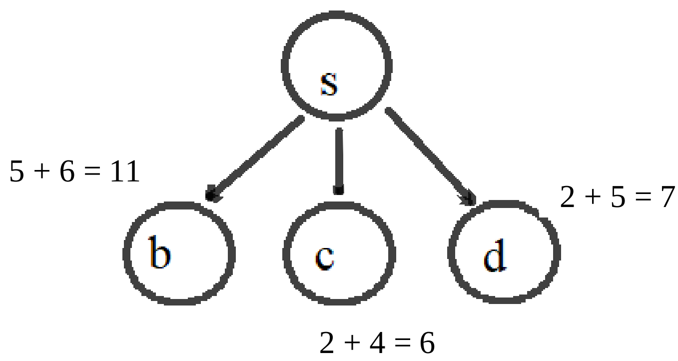
Ο **αλγόριθμος Α*** ακολουθεί την λογική των αλγορίθμων Branch and Bound και Best First.

Θεωρούμε τη σύνθετη ευριστική συνάρτηση F(k) = g(k) + h(k), όπου:

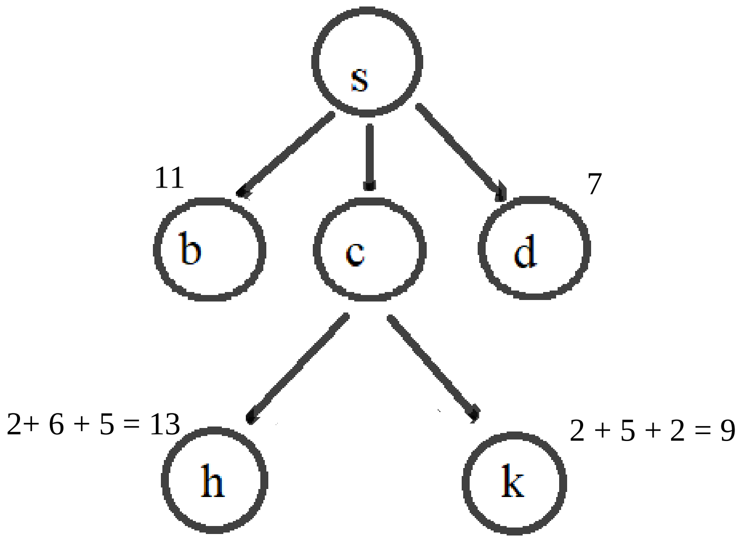
g(k) : η απόσταση της k από την αρχική κατάσταση, η οποία είναι πραγματική και γνωστή.

h(k) : μία εκτίμηση της απόστασης της k από το στόχο (μέσω μίας ευριστικής συνάρτησης).

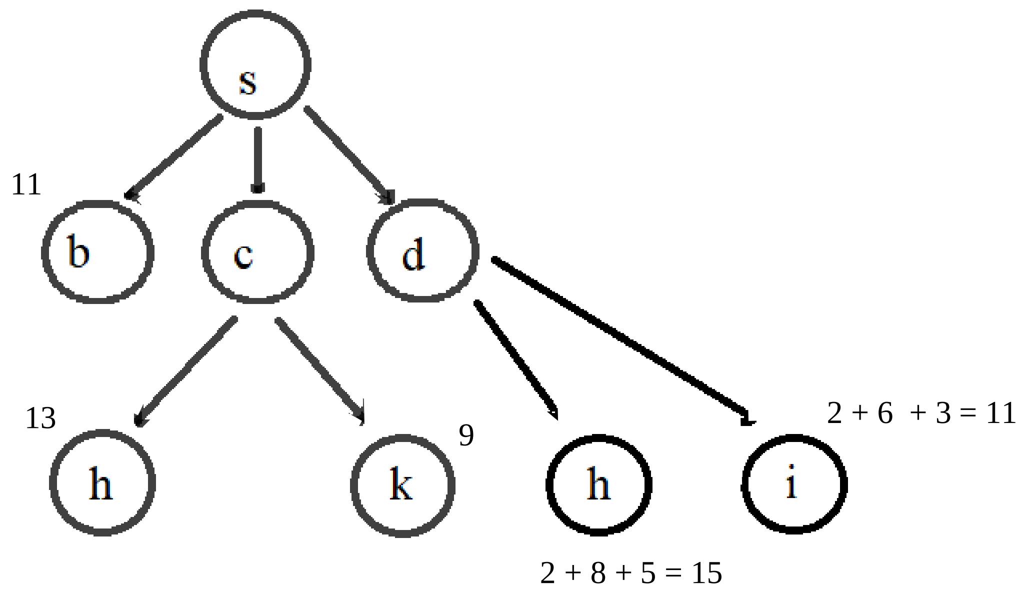
Εκτελούμε τον αλγόριθμο με αρχική κατάσταση την s. Παρατηρούμε τα παιδιά του s και ταξινομούμε σε αύξουσα σειρά τους κόμβους με τις ευριστικές.



Οπότε, έχουμε με την σειρά c, d, b. Μικρότερη τιμή ευριστικής έχει ο κόμβος c, άρα επεκτείνουμε εκεί.

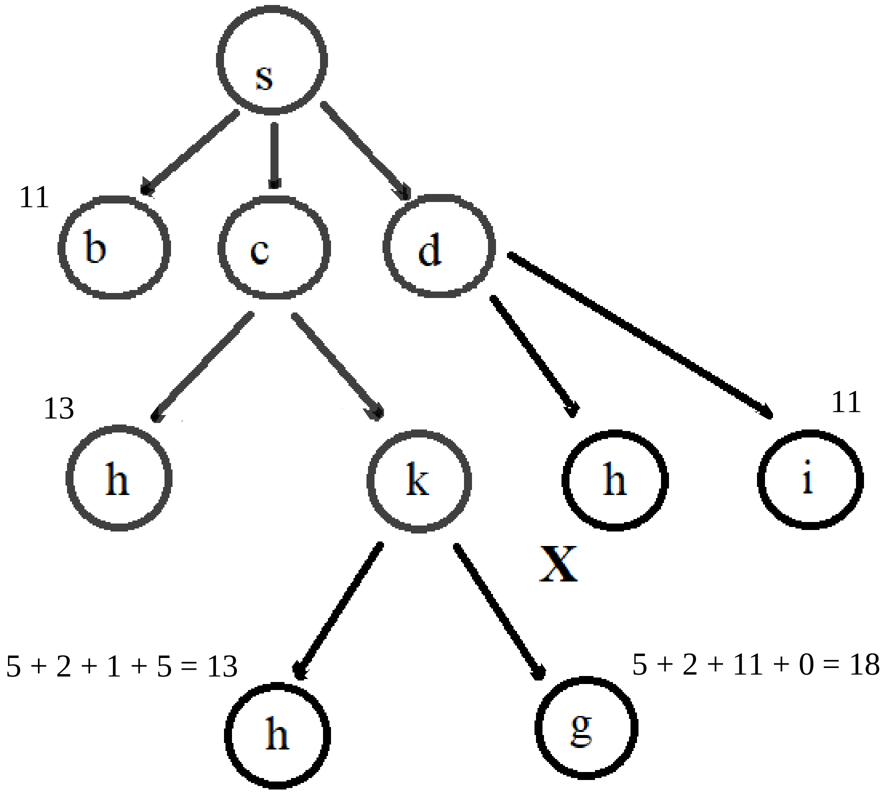


Συνεχίζουμε, να επεκτείνουμε από την μεριά του κόμβου d, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης d, k, b, h. Ο κόμβος d θα έχει δύο παιδιά, όπως παρατηρούμε παρακάτω.



Κλαδεύουμε τον κόμβο h(15), γιατί μπορούμε να φτάσουμε στον κόμβο h με μικρότερη τιμή ευριστικής 13 < 15.

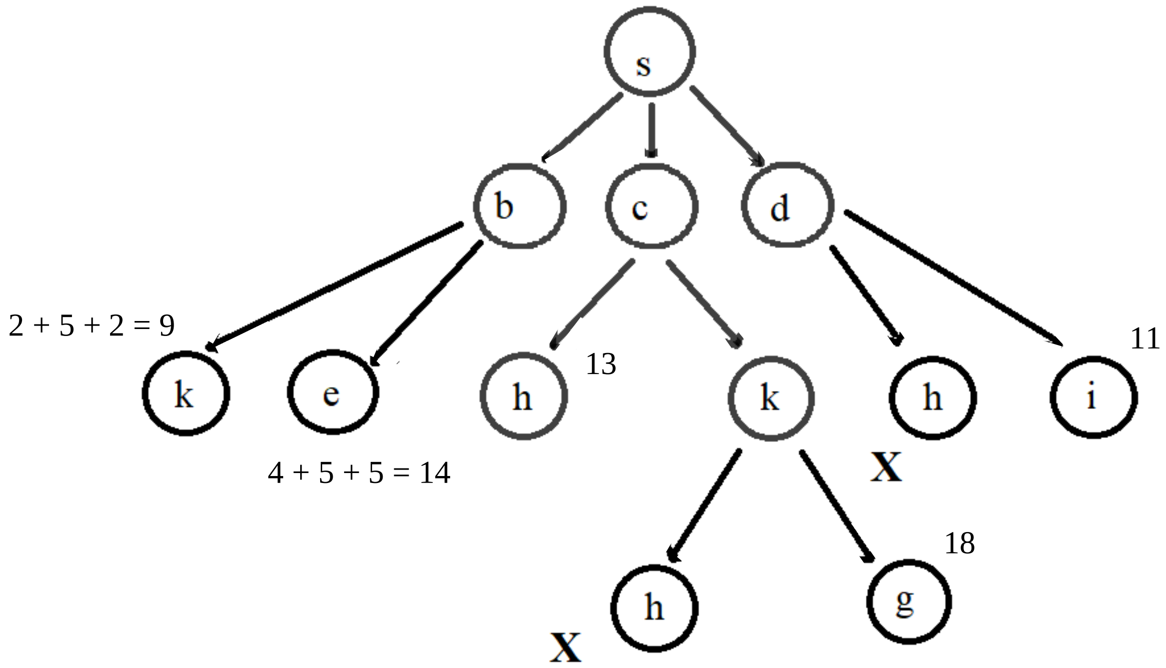
Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου k, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης k, b, i,, h.



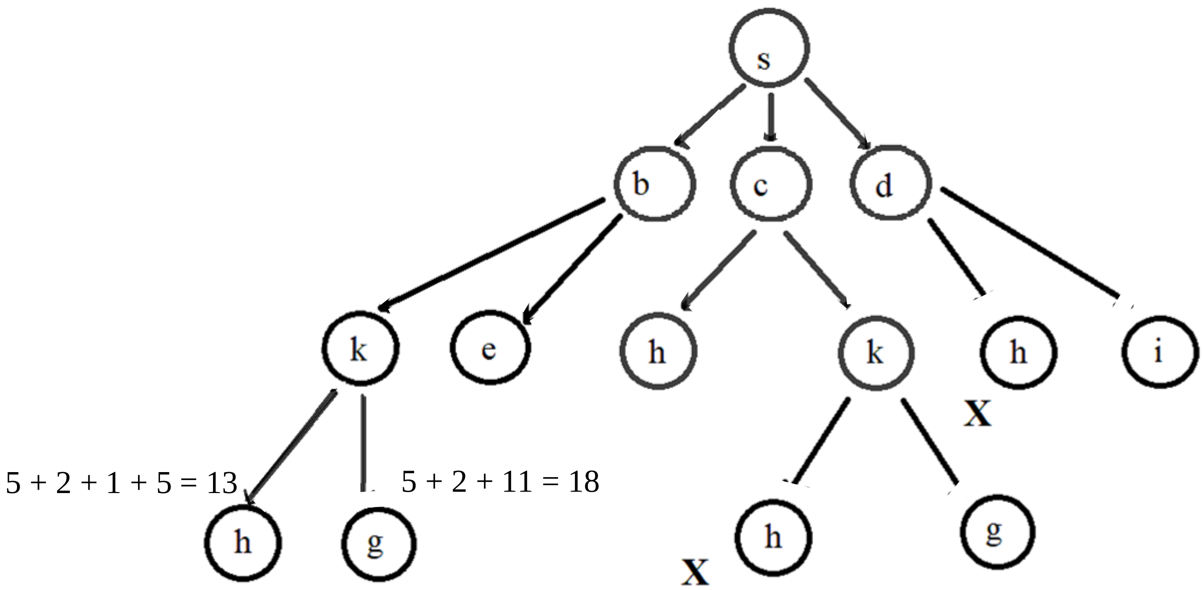
Κλαδεύουμε τον κόμβο – παιδί του κόμβου k, h(13), γιατί μπορούμε να φτάσουμε στον κόμβο h με μικρότερη διαδρομή.

Βρήκαμε πιθανή λύση, με κόστος 18 στο μονοπάτι s → c → k → g. Στην περίπτωση που βρούμε μονοπάτια μήκους μεγαλύτερο από 18, θα τα κλαδεύουμε.

Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου b, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης b, i, h. Κατά σύμβαση, θεωρούμε τον κόμβο b από τον κόμβο i, αφού οι κόμβοι με ίση τιμή ταξινομούνται με αλφαβητική σειρά.



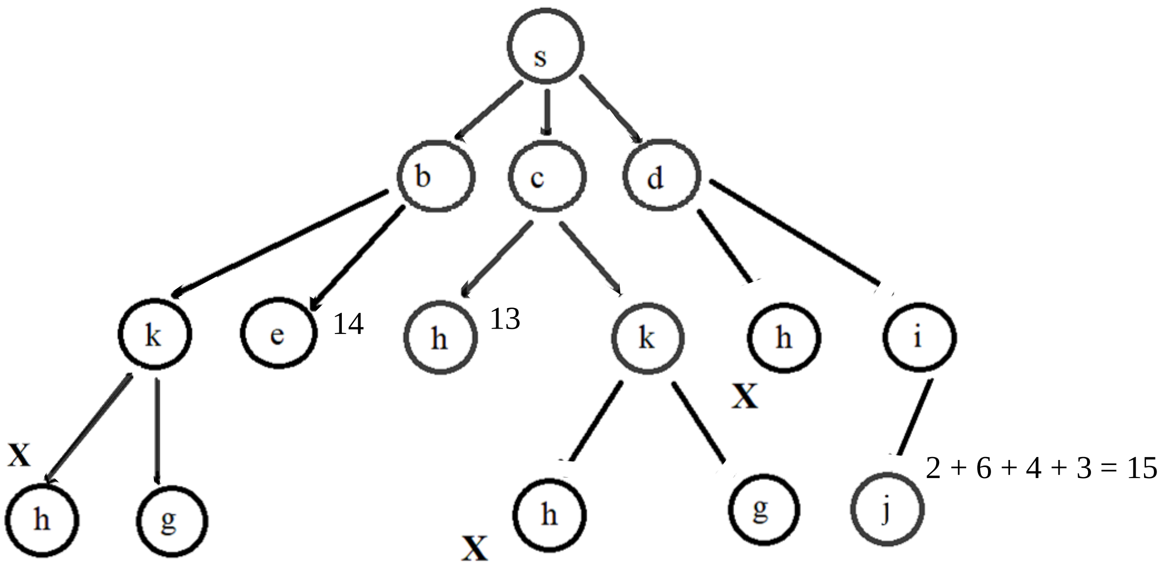
Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου k, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης.



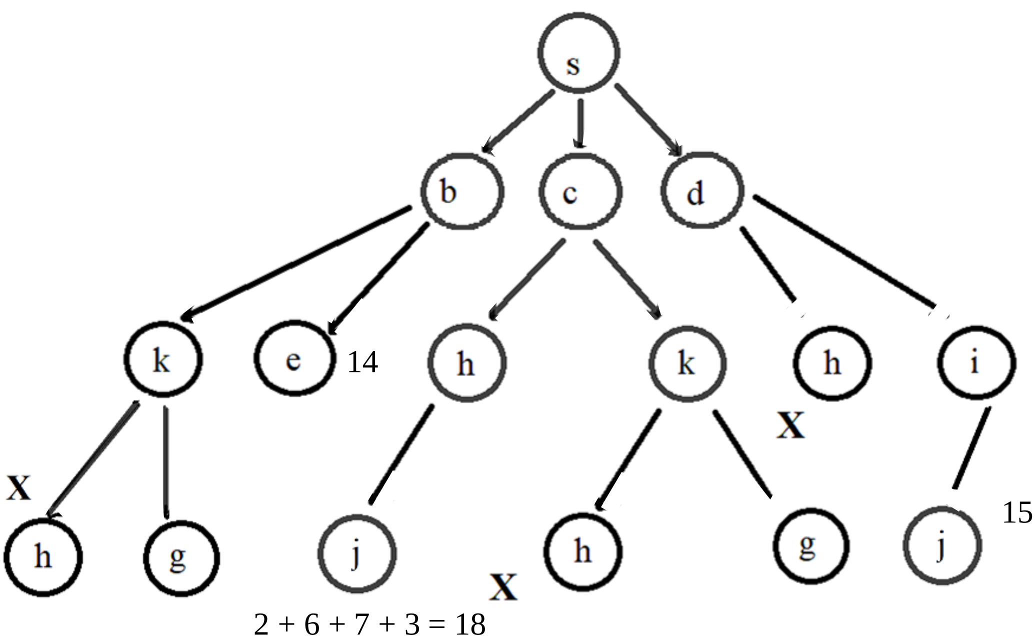
Βρήκαμε πιθανή λύση, με κόστος 18 στο μονοπάτι s → b → k → g.

Κλαδεύουμε τον κόμβο – παιδί του κόμβου k, h(13), γιατί μπορούμε να φτάσουμε στον κόμβο h με μικρότερη διαδρομή.

Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου i, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης.

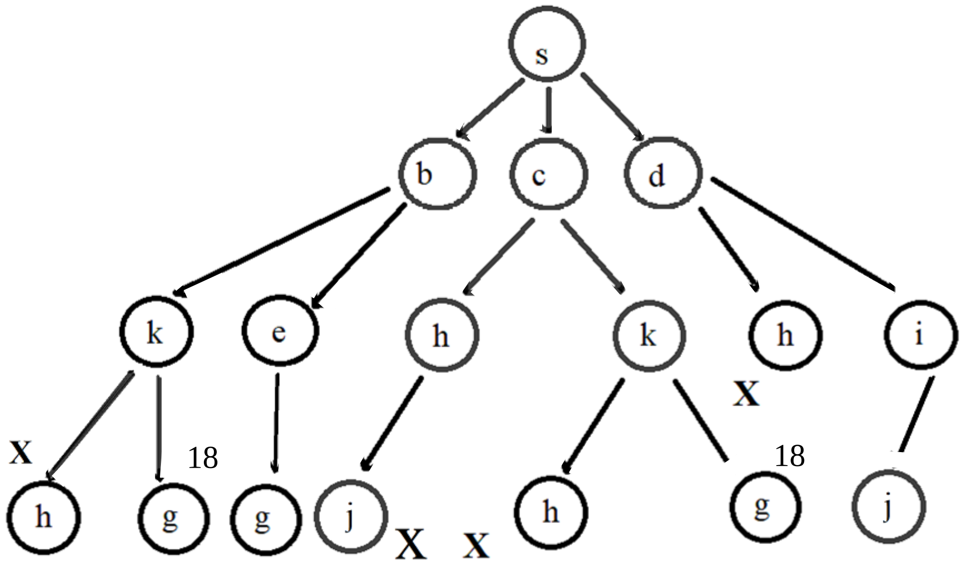


Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου h, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης.



Κλαδεύουμε τον κόμβο – παιδί του κόμβου h, j(18), γιατί μπορούμε να φτάσουμε στον κόμβο j με μικρότερη τιμή ευριστικής 15 < 18.

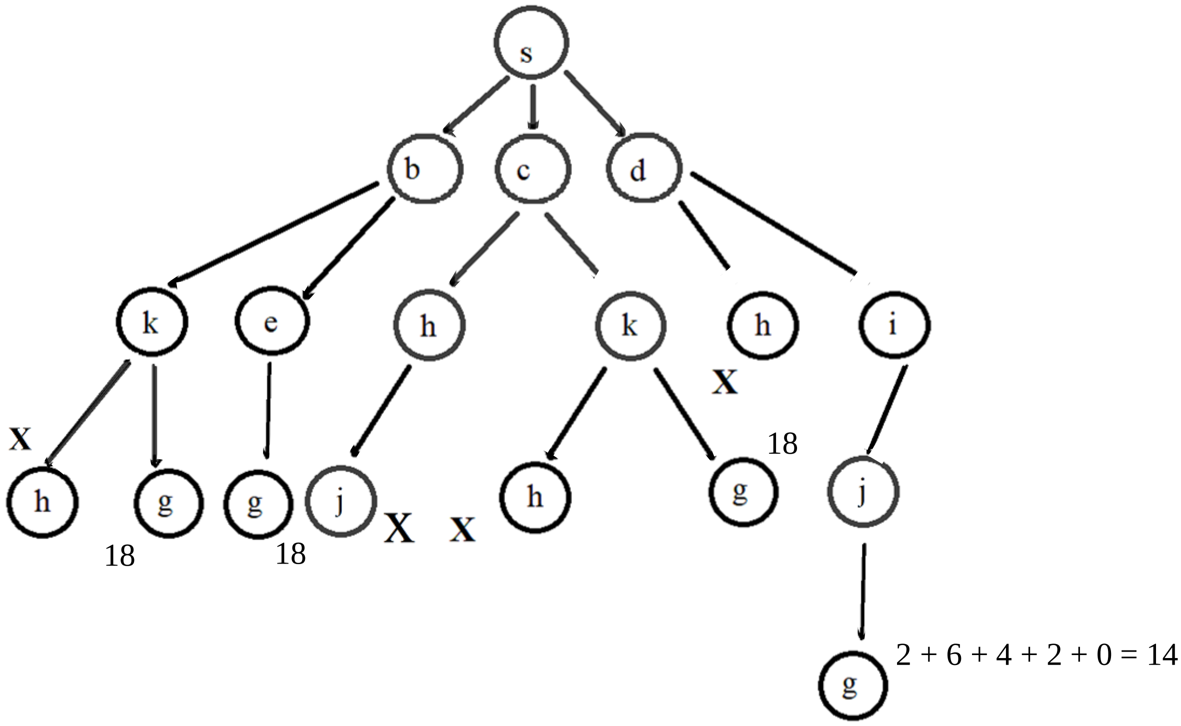
Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου e, γιατί έχει τη μικρότερη ευριστική τιμή από το σύνολο του ταξινομημένου μετώπου αναζήτησης.



5 + 4 + 9 = 18

Βρήκαμε πιθανή λύση, με κόστος 18 στο μονοπάτι s → b → e → g.

Στη συνέχεια, συνεχίζουμε να επεκτείνουμε από την μεριά του κόμβου j.



Βρήκαμε τη **βέλτιστη** λύση, με κόστος 14 στο μονοπάτι s → d → i → j → g(14).

Το μέτωπο αναζήτησης είναι άδειο, οπότε ο αλγόριθμος A* ολοκληρώθηκε.

Συνοπτικά μπορούμε, τα παραπάνω βήματα, να τα συμπυκνώσουμε σε έναν πίνακα:

Βήμα	**Λίστα Μονοπατιών**
1	[]
2	[s→c:6, s→d:7, s→b:11]
3	[s→d:7, s→c→k:9, s→b:11, s→c→h:13]
4	[s→c→k:9, s→b:11, s→d→i:11, s→c→h:13]
5	[s→b:11, s→d→i:11, s→c→h:13, **s→c→k→g:18**]
6	[s→b→k:9, s→d→i:11, s→c→h:13, s→b→e:14, **s→c→k→g:18**]
7	[s→d→i:11, s→c→h:13, s→b→e:14, **s→c→k→g:18, s→b→k→g:18**]
8	[s→c→h:13, s→b→e:14, s→d→i→j:15, **s→c→k→g:18, s→b→k→g:18**]
9	[s→b→e:14, s→d→i→j:15, **s→c→k→g:18, s→b→k→g:18**]
10	[s→d→i→j:15, **s→c→k→g:18, s→b→k→g:18, s→b→e→g:18**]
11	[**s→d→i→j→g:14, s→c→k→g:18, s→b→k→g:18, s→b→e→g:18**]

2.

Το πρόβλημα έχει 4 λύσεις.

Λύση με κόστος 18 στο μονοπάτι s → c → k → g

Λύση με κόστος 18 στο μονοπάτι s → b → k → g.

Λύση με κόστος 18 στο μονοπάτι s → b → e → g

Λύση με κόστος 14 στο μονοπάτι s → d → i → j → g.

Βέλτιστη είναι η τελευταία λύση, αφού έχει κόστος 14 < 18.

Ο αλγόριθμος Hill Climbing απλώς αποφαίνεται αν ο αλγόριθμος καταλήγει σε έναν κόμβο στόχο, ενώ δεν βρίσκει την βέλτιστη λύση. Στην προκειμένη περίπτωση, ο αλγόριθμος επέστρεψε επιτυχία του προγράμματος, αφού κατέληξε στο κόμβο στόχο g.

Αντίθετα, ο αλγόριθμος A*, μπορεί (όχι πάντα) να βρίσκει λύσεις από τις οποίες αποφαίνεται για το ποια είναι η βέλτιστη. Αν για κάθε κατάσταση, η τιμή h(k) είναι μικρότερη ή ίση με την πραγματική απόσταση της k από την τελική κατάσταση, τότε ο αλγόριθμος A* βρίσκει *πάντα* τη βέλτιστη λύση. Με άλλα λόγια, πρέπει να έχουμε αποδεκτό (admissible) ευριστικό μηχανισμό ώστε να υπολογίζεται η βέλτιστη λύση, αλλιώς *δεν είναι σίγουρο* ότι ο αλγόριθμος μπορεί θα υπολογίσει τη βέλτιστη λύση.

Άσκηση 2

1.

Το **περιβάλλον** του προβλήματος περιλαμβάνει:

Ο χώρος του αυτοκινητοδρόμου, τα αυτοκίνητα, οι πεζοί και οι επιβάτες.

Οι **αισθητήρες** του προβλήματος περιλαμβάνουν:

Κάμερες, αισθητήρες υπερήχων, κοντέρ, GPS, οδόμετρο, αισθητήρες λειτουργίας αυτοκινήτου και εγκέφαλος αυτοκινήτου.

Οι **δράσεις** του προβλήματος περιλαμβάνουν:

Γκάζι, φρένο, τιμόνι, φώτα και κόρνα.

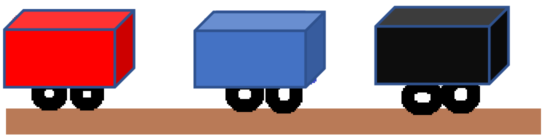
Οι **δείκτες επίδοσης** του προβλήματος περιλαμβάνουν:

Ασφαλής, γρήγορη, άνετη, χωρίς μεταβάσεις μεταφορά στον προορισμό.

2.

Ο **κόσμος** του προβλήματος αποτελείται από τα εξής αντικείμενα:

Αυτοκίνητα και ο δρόμος:



Αντικείμενα	Ιδιότητες	Σχέσεις
Κόκκινο Αυτοκίνητο	Κόκκινο αυτοκίνητο ελεύθερο	Κόκκινο Αυτοκίνητο πίσω από το μπλε
Μπλε Αυτοκίνητο	Μπλε αυτοκίνητο ελεύθερο	Μπλε Αυτοκίνητο πίσω από το μαύρο
Μαύρο Αυτοκίνητο	Μαύρο αυτοκίνητο ελεύθερο	Μαύρο Αυτοκίνητο προπορεύεται όλων
Δρόμος	Ο δρόμος έχει χώρο	Κινούνται πάνω σε αυτό τα αυτοκίνητα

Ένα παράδειγμα κατάστασης του κόσμου είναι η εξής αναπαράσταση του κόσμου σε μία δεδομένη χρονική στιγμή:

Κόκκινο αυτοκίνητο πίσω από το μπλε
Μπλε αυτοκίνητο πίσω από το μαύρο
Μαύρο αυτοκίνητο πάνω στον δρόμο
Μπλε αυτοκίνητο προπορεύεται όλων
Όλα τα αυτοκίνητα κινούνται πάνω στο δρόμο

Κάποιοι **τελεστές μετάβασης** , από τη μία κατάσταση σε μία άλλη, είναι τα παρακάτω παραδείγματα:

- Το κόκκινο αυτοκίνητο προσπερνά το μπλε αυτοκίνητο.
- Το κόκκινο αυτοκίνητο προσπερνά το μαύρο αυτοκίνητο.

3.

Μία ευριστική συνάρτηση που εκτιμά τόσο το κόστος μετάβασης από μία κατάσταση σε μία άλλη, όσο και το υπολοιπόμενο κόστος μέχρι την τελική κατάσταση, είναι το **ντεπόζιτο βενζίνης** του αυτοκινήτου. Συγκεκριμένα, η **κατανάλωση βενζίνης** του αυτοκινήτου είναι εκείνη που θα αποδείξει το κόστος από την αρχική θέση έως τη τελική θέση που θα ακινητοποιηθεί το αυτοκίνητο.

Άσκηση 3

Τα δοσμένα χαρακτηριστικά είναι Boolean, γιατί παίρνουν δύο τιμές. Οπότε, μπορούμε να τα θεωρήσουμε σαν ενδεχόμενα που είναι αληθή ή ψευδή, ώστε να ακολουθήσουμε πιθανοτική προσέγγιση. Επίσης, θεωρούμε ότι τα μηνύματα που περιέχουν τη λέξη «ευκαιρία» με εκείνα που περιέχουν τη λέξη «μοναδικός» είναι ανεξάρτητα.

Συνολικά, έχουμε 1800 μηνύματα τα οποία χωρίζονται σε δύο **κατηγορίες**:

Είναι διαφημιστικά: add

Δεν είναι διαφημιστικά: no_add

Έχουμε δύο **χαρακτηριστικά**:

Περιέχουν τη λέξη «ευκαιρία»: op

Περιέχουν τη λέξη «μοναδικός»: un

Και τα *συμπληρωματικά* τους:

Δεν περιέχουν τη λέξη «ευκαιρία»: no_op

Δεν περιέχουν τη λέξη «μοναδικός»: no_un

Υπολογίζουμε τις παρακάτω πιθανότητες βάσει του δοσμένου πίνακα:

p(add) = $\frac{300}{1800}$, p(no_add) = $\frac{1500}{1800}$

p(op | add) = $\frac{285}{300}$, p(no_op | add) = $\frac{15}{300}$

p(un | add) = $\frac{240}{300}$, p(no_un | add) = $\frac{60}{300}$

p(op | no_add) = $\frac{225}{1500}$, p(no_op | no_add) = $\frac{1275}{1500}$

p(un | no_add) = $\frac{600}{1500}$, p(no_un | no_add) = $\frac{900}{1500}$

Από τον τύπο του Bayes, έχουμε:

$$\boxed{p(i\,|\,x) = \frac{p(i)\cdot\prod_{k=1}^{p} p(x^{(k)}\,|\,i)}{\sum_{j=1}^{c} p(j)\cdot\prod_{k=1}^{p} p(x^{(k)}\,|\,j)}}$$

Χρησιμοποιούμε αυτόν τον τύπο, για να υπολογίσουμε την πιθανότητα να μας έρθει ένα μήνυμα με κλάση i = «είναι διαφημιστικό μήνυμα» ή «add» είτε «δεν είναι» ή «no_add», δεδομένου ότι έχει χαρακτηριστικό x = «μήνυμα που περιέχει τις λέξεις «ευκαιρία» και «μοναδικός»» ή «op, un».

Η πιθανότητα να μας έχει σταλεί ένα **διαφημιστικό** μήνυμα που περιέχει τις λέξεις «ευκαιρία» και «μοναδικός», είναι:

$$p(add|\,op,un) \;=\; \frac{p(add)p(op|add)p(un|add)}{p(add)p(op|add)p(un|add) + p(no\_add)p(op|no\_add)p(un|no\_add)} =$$

$$\frac{\frac{300}{1800}\,\frac{285}{300}\,\frac{240}{300}}{\frac{300}{1800}\,\frac{285}{300}\,\frac{240}{300} + \frac{1500}{1800}\,\frac{225}{1500}\,\frac{600}{1500}} = 0{,}716$$

Η πιθανότητα να μας έχει σταλεί ένα **μη διαφημιστικό** μήνυμα που περιέχει τις λέξεις «ευκαιρία» και «μοναδικός», είναι:

$$p(no\_add|\ op, un)\ =\ \frac{p(no\_add)p(op|no\_add)p(un|no\_add)}{p(add)p(op|add)p(un|add) + p(no\_add)p(op|no\_add)p(un|no\_add)} =$$

$$\frac{\frac{1500}{1800}\frac{225}{1500}\frac{600}{1500}}{\frac{300}{1800}\frac{285}{300}\frac{240}{300}+\frac{1500}{1800}\frac{225}{1500}\frac{600}{1500}} = 0{,}284$$

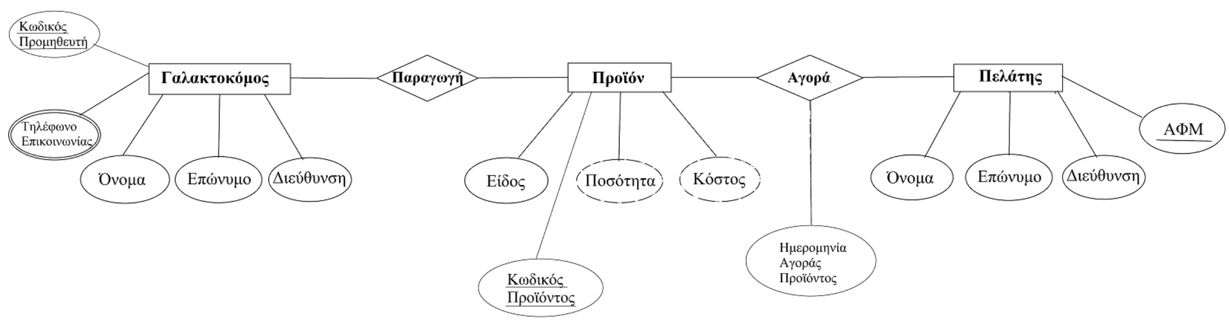
Από τις παραπάνω μετρημένες πιθανότητες παρατηρούμε ότι:

$$p(no\_add|\ op, un)\ +\ p(add|\ op, un)\ =\ 1$$

Οπότε, ένα μήνυμα το οποίο περιέχει τις λέξεις «ευκαιρία» και «μοναδικός», έχει περισσότερες πιθανότητες να είναι διαφημιστικό.

Άσκηση 4

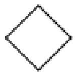
Το ζητούμενο ER διάγραμμα φαίνεται παρακάτω:



Τα γραφικά σύμβολα που χρησιμοποιήθηκαν για το παραπάνω ER διάγραμμα είναι τα εξής:

 Περιγραφή οντότητας:

Στην προκειμένη περίπτωση, οι οντότητες που μας ενδιαφέρουν είναι: ο γαλακτοκόμος, ο πελάτης και το προϊόν.

 Απόδοση σχέσης:

Στην προκειμένη περίπτωση, ο γαλακτοκόμος *παράγει* ένα προϊόν το οποίο *αγοράζουν* οι πελάτες. Εχουμε συνδέσει μόνο με γραμμές το σχήμα αυτό, γιατί θέλουμε να δείξουμε ότι υπάρχουν «πολλές» σχέσεις που συνδέουν τα περιεχόμενα αυτών των σχημάτων με εκείνα των οντοτήτων.

Τέλος, εφόσον μας ενδιαφέρει και η ημερομηνία αγοράς προϊόντων, προσθέτουμε στην σχέση μία απόδοση περιγραφής που υποδεικνύει το ενδιαφέρον για την ημερομηνία αγοράς.

 Απόδοση περιγραφής οντότητας:

Στην προκειμένη περίπτωση, ο γαλακτοκόμος και ο πελάτης περιγράφονται από το *όνοματεπώνυμό* τους και την διεύθυνση τους. Το προϊον περιγράφεται από το *είδος* του.

 Προερχόμενη απόδοση οντότητας:

Στην προκειμένη περίπτωση, η ποσότητα και το *κόστος* του προϊόντος εξαρτάται από το ίδιο το προϊόν.

 Πολλαπλή απόδοση:

Στην προκειμένη περίπτωση, ο γαλακτοκόμος μπορεί να έχει στη διάθεση του παραπάνω από ένα τηλέφωνα επικοινωνίας.

 Πρωταρχικό κλειδί:

Ο γαλακτοκόμος αναγνωρίζεται μοναδικά από τον *κωδικό προμηθευτή*. Επίσης, το προϊόν κατηγοριοποιείται μοναδικά από τον *κωδικό προϊόντος*, ενώ ο πελάτης πιστοποιείται από το μοναδικό αναγνωριστικό του, τον *ΑΦΜ*.