ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

# ΣΗΜΕΙΩΣΕΙΣ
# ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ

ΠΑΝΑΓΙΩΤΗΣ ΣΤΑΜΑΤΟΠΟΥΛΟΣ

ΑΘΗΝΑ 2004

Οι σημειώσεις αυτές μεταφέρθηκαν σε PowerPoint® από τον Παναγιώτη Κόκκαλη, φοιτητή του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Αθηνών, το καλοκαίρι και το φθινόπωρο του 2002. Τον ευχαριστώ πολύ και εκ μέρους των συναδέλφων του, επισημαίνοντας ότι η είσοδος για τη μεταφορά ήταν μία δυσανάγνωστη χειρόγραφη μορφή, συνεπώς για ό,τι λάθη υπάρχουν δεν ευθύνεται εκείνος.

Παναγιώτης Σταματόπουλος

**ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ**
**(ARTIFICIAL INTELLIGENCE)**

- Τι είναι η τεχνητή νοημοσύνη;
- Τι πραγματεύεται η τεχνητή νοημοσύνη;
- Πώς είναι η τεχνητή νοημοσύνη;

—— . ——

Τι είναι η τεχνητή νοημοσύνη;

- Τεχνητή νοημοσύνη είναι η μελέτη του πώς να βάλουμε τους υπολογιστές να κάνουν πράγματα που σήμερα οι άνθρωποι τα κάνουν καλύτερα (Rich, Knight).

- Τεχνητή νοημοσύνη είναι η μελέτη των διαδικασιών που κάνουν δυνατή την αντίληψη, το συλλογισμό και τη δράση (Winston).

- Τεχνητή νοημοσύνη είναι ο κλάδος της επιστήμης των υπολογιστών που ασχολείται με την αυτοματοποίηση της νοήμονος συμπεριφοράς (Luger).

- Τεχνητή νοημοσύνη είναι η σχεδίαση και μελέτη προγραμμάτων που έχουν νοήμονα συμπεριφορά (Dean, Allen, Aloimonos).

Άλλος ένας ορισμός της τεχνητής νοημοσύνης:

Τεχνητή νοημοσύνη είναι η δημιουργία νοημόνων κατασκευασμάτων (Ginsberg).

- Αναδρομικός ορισμός

- Τι είναι νοημοσύνη;

  Νοήμονες λειτουργίες είναι αυτές που οι άνθρωποι κάνουν καλά.

  - Όραση (Vision)
  - Κατανόηση λόγου (Speech understanding)
  - Προφανής συλλογιστική (Commonsense reasoning)
  - Σκάκι (Chess)
  - Απόδειξη θεωρημάτων (Theorem proving)
  - Ιατρική διάγνωση (Medical diagnosis)
  - Σχεδίαση παραγωγής (Production planning)
  - .......................................

- Πείραμα του Turing (Turing test)

  Τροποποίηση του ορισμού:

  Τεχνητή νοημοσύνη είναι η δημιουργία κατασκευασμάτων που, με αξιόπιστο τρόπο, επιτυγχάνουν στο πείραμα του Turing.

- Τι είναι κατασκεύασμα;

  Κατασκεύασμα είναι ένα σύστημα φυσικών συμβόλων (physical symbol system).

- Τι είναι σύστημα φυσικών συμβόλων;
  - Newell, Simon (1976)
  - Οντότητες / Σύμβολα
  - Εκφράσεις / Δομές συμβόλων
  - Διαδικασίες δημιουργίας, τροποποίησης και καταστροφής δομών συμβόλων
  - Σύστημα φυσικών συμβόλων είναι μία “μηχανή” διαχείρισης ενός συνόλου δομών συμβόλων.
  - Δηλωτισμός (Declarativism): Τα σύμβολα ενός συστήματος φυσικών συμβόλων αντιστοιχούν σε αντικείμενα από τον πραγματικό κόσμο.

Νέος ορισμός της τεχνητής νοημοσύνης:

Τεχνητή νοημοσύνη είναι η δημιουργία συστημάτων φυσικών συμβόλων που, με αξιόπιστο τρόπο, επιτυγχάνουν στο πείραμα του Turing.

- Η τεχνητή νοημοσύνη είναι επιστήμη (science) ή τεχνολογία (engineering);
  - Η λέξη “δημιουργία” στον ορισμό παραπέμπει στην τεχνολογία.
  - Χρειάζεται όμως πολλή επιστήμη στα θεμέλια.
  - Έχουν λυθεί όλα τα επιστημονικά προβλήματα;
  - Η κατασκευή πυρηνικών αντιδραστήρων είναι “τεχνολογία” ή “επιστήμη”;

Τι πραγματεύεται η τεχνητή νοημοσύνη;

- Μέθοδοι αναζήτησης (Search methods)
- Αναπαράσταση γνώσης (Knowledge representation)
- Εφαρμογές των παραπάνω

  - Σχεδίαση (Planning)
  - Μάθηση (Learning)
  - Όραση (Vision)
  - Φυσική γλώσσα (Natural language)
  - Έμπειρα συστήματα (Expert systems)
  - ..................................

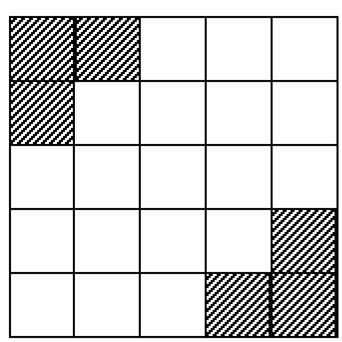
Μέθοδοι αναζήτησης

- Προβλήματα αναζήτησης

  - Κατασκευή σταυρολέξων
  - Σκάκι
  - Προγραμματισμός διακοπών

- Χώρος αναζήτησης

  - $\sim 10^{160}$ διαφορετικά παιγνίδια στο σκάκι
  - $24^4 \approx 3.3 \times 10^5$ πιθανά σταυρόλεξα 2x2
  - $24^{19} \approx 1.7 \times 10^{26}$ πιθανά σταυρόλεξα της μορφής:



  - $24^{200} \approx 1.1 \times 10^{276}$ πιθανά σταυρόλεξα με 200 άσπρα τετραγωνίδια

Πως είναι δυνατόν να λυθούν δύσκολα προβλήματα αναζήτησης (με πολύ μεγάλους χώρους αναζήτησης) σχετικά εύκολα;

**ΑΠΑΝΤΗΣΗ:** Με τη χρήση ευριστικών (heuristic) μεθόδων

Αναπαράσταση γνώσης

- Κατηγορηματική λογική πρώτης τάξης (First order predicate logic)
- Συστήματα διατήρησης της αλήθειας (Truth maintenance systems)
- Μη μονότονη συλλογιστική (Nonmonotonic reasoning)
- Συλλογιστική βασισμένη σε πιθανότητες (Probabilistic reasoning)

Πως είναι η τεχνητή νοημοσύνη;

- Είναι εύκολο να καταλάβει κάποιος τα προβλήματα που αντιμετωπίζει η τεχνητή νοημοσύνη.
- Το δύσκολο είναι να γίνει ουσιαστική πρόοδος στην επίλυση των προβλημάτων της τεχνητής νοημοσύνης.
- Πού είναι το ενδιαφέρον; Στο ότι τόσο απλά πράγματα αποδεικνύονται υπερβολικά πολύπλοκα.

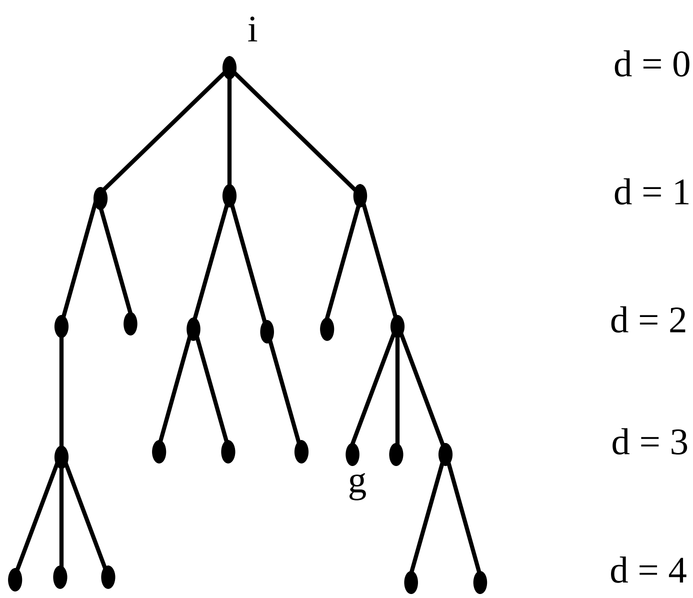
## ΒΙΒΛΙΟΓΡΑΦΙΑ


1. "σύγγραμμα" του μαθήματος:
   Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Ι. Ρεφανίδης, Φ. Κόκκορας, Η. Σακελλαρίου. Τεχνητή Νοημοσύνη. Εκδόσεις Γαρταγάνη, Θεσσαλονίκη, 2002.
2. M. Ginsberg. Essentials of Artificial Intelligence. Morgan Kaufmann, 1993.
3. E. Rich, K. Knight. Artificial Intelligence. McGraw Hill, 2nd edition, 1991.
4. D. H. Winston. Artificial Intelligence. Addison Wesley, 3rd edition, 1992.
5. G. Luger. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Addison Wesley, 4th edition, 2001.
6. T. Dean, J. Allen, Y. Aloimonos. Artificial Intelligence: Theory and Practice. Benjamin/Cummings, 1995.
7. N. Nilsson. Artificial Intelligence: A New Synthesis. Morgan Kaufmann, 1998.
8. E. Charniak, D. McDermott. Introduction to Artificial Intelligence. Addison Wesley, 1987.
9. N. Rowe. Artificial Intelligence through Prolog. Prentice Hall, 1988.
10. A. Bundy (ed.). Catalogue of Artificial Intelligence Techniques. Springer Verlag, 3rd edition, 1990.
11. S. Russel, P. Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall, 1994.

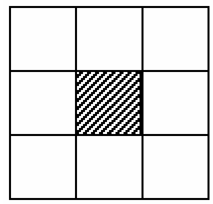
**ΑΝΑΖΗΤΗΣΗ**

**(SEARCH)**

- Χώρος αναζήτησης
  - Δέντρο
  - Γράφος (με ή χωρίς κύκλους)
- Αρχική κατάσταση / κόμβος (≥ 1)
- Τελική κατάσταση / κόμβος (≥ 1)
- Γέννηση διαδόχων καταστάσεων κόμβων
- Το πρόβλημα
  - Εύρεση μονοπατιού από αρχική κατάσταση σε τελική
  - Κατασκευή τελικής κατάστασης από αρχική



- Παράγοντας διακλάδωσης (branching factor)
- Παράδειγμα: Συμπλήρωση άδειου σταυρολέξου με υπαρκτές λέξεις



- Γενική μέθοδος αναζήτησης

Βήμα 1: Έστω L η λίστα των αρχικών κόμβων.
Κάθε στιγμή το L είναι η λίστα των κόμβων που δεν έχουν εξετασθεί.

Βήμα 2: Αν η L είναι κενή, απέτυχες. Αλλιώς, διάλεξε ένα κόμβο n από την L.

Βήμα 3: Αν ο n είναι τελικός κόμβος, να τον επιστρέψεις, καθώς και το μονοπάτι από τον αρχικό κόμβο προς τον n, και να σταματήσεις.

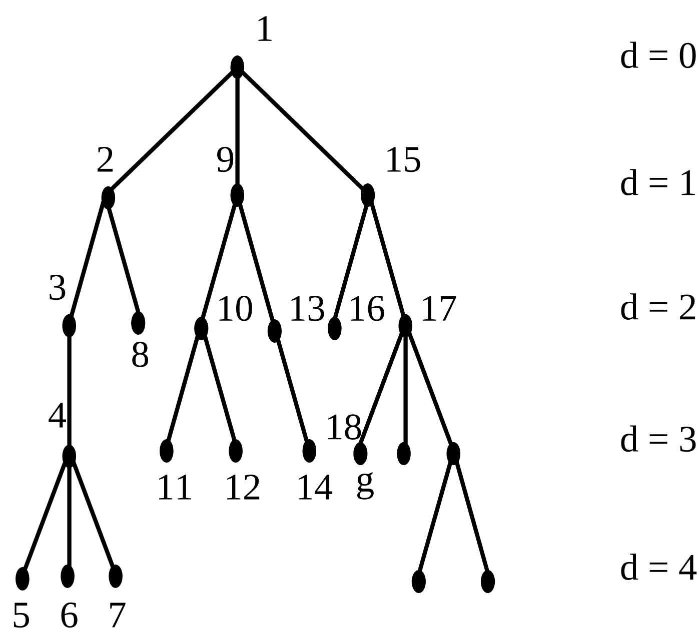
Βήμα 4: Αλλιώς, να διαγράψεις τον n από την L και να προσθέσεις στην L τα παιδιά του n, προσαρτώντας στο καθένα το μονοπάτι από τον αρχικό κόμβο προς αυτό. Πήγαινε στο βήμα 2.

- Σημεία που δεν θίγονται στη γενική μέθοδο αναζήτησης
  - Πώς επιλέγεται ο κόμβος n στο βήμα 2;
  - Σε ποιο σημείο της L τοποθετούνται τα παιδιά του n στο βήμα 4;
  - Τι γίνεται με τους κύκλους;

**Τυφλή αναζήτηση (Blind search)**

- Πρώτα–κατά–βάθος (depth-first) αναζήτηση

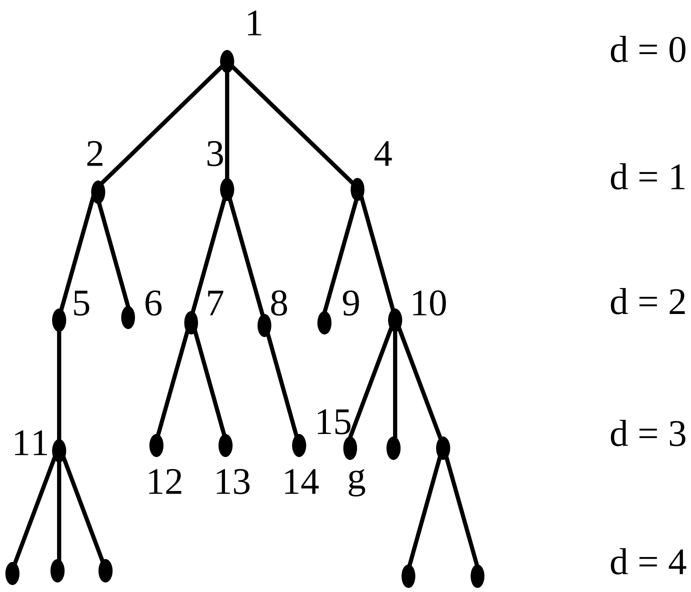
  - Στο βήμα 2 της γενικής μεθόδου αναζήτησης επιλέγεται ο πρώτος κόμβος από την L.
  - Στο βήμα 4 τα παιδιά του n τοποθετούνται στην αρχή της L (stack).




  - Η πρώτα–κατά–βάθος αναζήτηση δεν γίνεται απαραίτητα από–αριστερά–προς–τα–δεξιά (left–to–right).

- Πρώτα–κατά–πλάτος (breadth–first) αναζήτηση

    - Στο βήμα 2 της γενικής μεθόδου αναζήτησης επιλέγεται ο πρώτος κόμβος από την L.
    - Στο βήμα 4 τα παιδιά του n τοποθετούνται στο τέλος της L (queue).

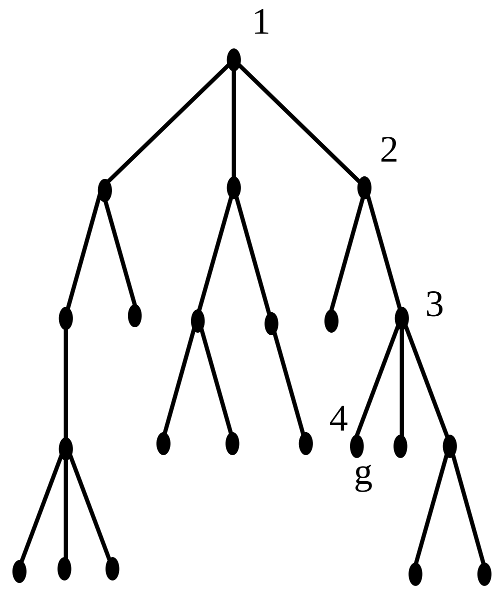



- Η πρώτα–κατά–βάθος μέθοδος αναζήτησης απαιτεί λιγότερη μνήμη από την πρώτα–κατά–πλάτος.

    - Γιατί;
    - Ποια μειονεκτήματα έχει η πρώτα–κατά–βάθος σε σχέση με την πρώτα–κατά–πλάτος μέθοδο αναζήτησης;
    - Στο πρόβλημα του σταυρολέξου ποια συμφέρει και γιατί;

**Ευριστική αναζήτηση (Heuristic search)**

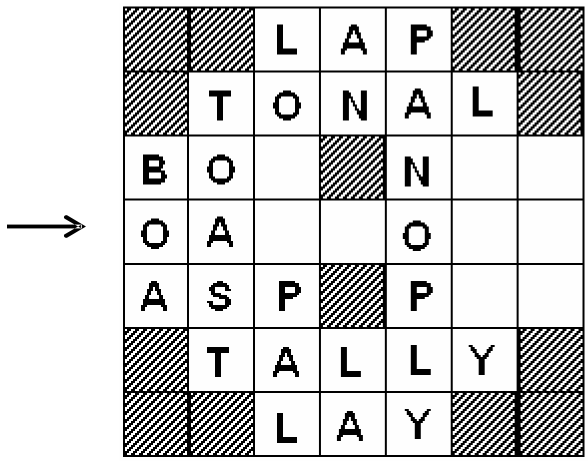
- Σε μία ιδανική μέθοδο αναζήτησης θα έπρεπε η ακολουθία των θεωρούμενων κόμβων να οδηγεί κατ' ευθείαν στον κόμβο–στόχο, π.χ.




- Μία ευριστική μέθοδος αναζήτησης προσπαθεί να προσεγγίσει την ιδανική με την επιλογή, στο βήμα 2 της γενικής μεθόδου αναζήτησης, εκείνου του κόμβου που "φαίνεται" να είναι πιο κοντά στο στόχο.
- Συμβιβασμός μεταξύ της δραστηριότητας βασικού επιπέδου (base–level) και μετα–επιπέδου (meta–level).
- Στην τυφλή αναζήτηση δεν υπάρχει δραστηριότητα μετα–επιπέδου.

**Οπισθοδρόμηση (Backtracking)**

- Χρονολογική (Chronological)
- Βασισμένη στην εξάρτηση (Dependency–directed)




.......... – TONAL – ....... – TOAST – ....... – LAY – LAP –

**Κατεύθυνση αναζήτησης**

- Εμπρόσθια αναζήτηση (Forward search)
- Οπίσθια αναζήτηση (Backward search)
- Βασικό κριτήριο επιλογής: Παράγοντας διακλάδωσης
- Οπίσθια αναζήτηση είναι δυνατή μόνο όταν οι στόχοι είναι δεδομένοι με ρητό τρόπο.

**Προβλήματα αναζήτησης**

- Σκάκι
  - Άνθρωπος: Μετα–επίπεδο
  - Μηχανή: Βασικό επίπεδο

- Τάβλι
  - Μεγάλος παράγοντας διακλάδωσης
  - Ευριστική εκτίμηση τρέχουσας κίνησης

- Ιεραπόστολοι και κανίβαλοι
  - Μικρός χώρος αναζήτησης
  - Μετα–επίπεδο: Ονόματα;

- Γρίφος–8, γρίφος–15, γρίφος–24
  - Αριθμός τετραγωνιδίων εκτός θέσης
  - Συνολική απόσταση Manhattan
  - Παραδεκτές (admissible) ευριστικές συναρτήσεις

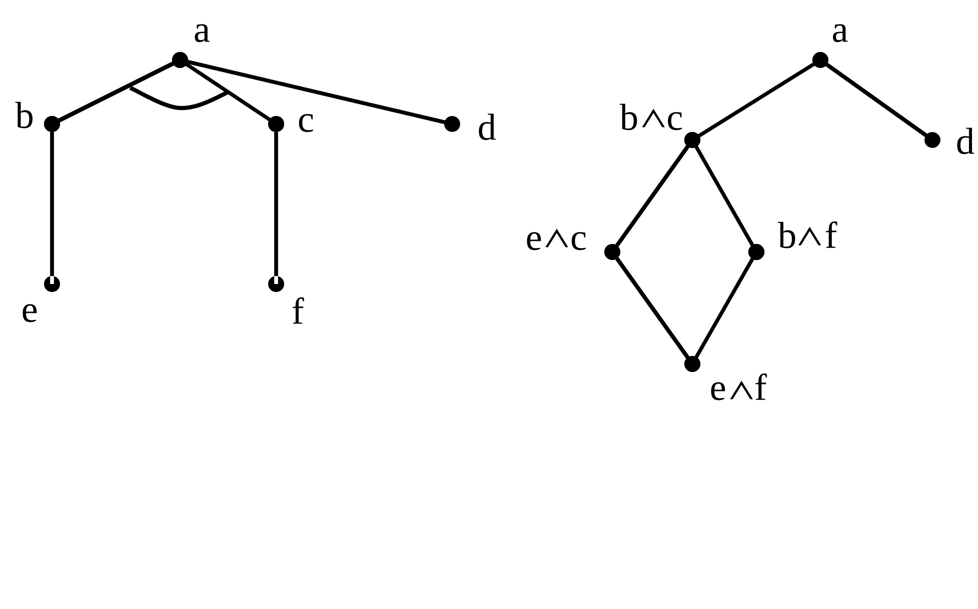
- Πύργοι του Ανόϊ
- Κύβος του Rubik
- Κόσμος των κύβων
- Κρυπτάριθμοι
- Ν βασίλισσες
- Ακρωτηριασμένη σκακιέρα και 31 ντόμινο

**Αναπαράσταση γνώσης και αναζήτηση**

- Απλά προβλήματα αναζήτησης (παιγνίδια, γρίφοι κλπ.) δεν έχουν ιδιαίτερες απαιτήσεις για αναπαράσταση γνώσης.
- Σύνθετα προβλήματα από την τεχνητή νοημοσύνη (κατανόηση φυσικής γλώσσας, έμπειρα συστήματα, όραση κλπ.) είναι επίσης προβλήματα αναζήτησης αλλά η επίλυσή τους απαιτεί κάποια συστηματική μέθοδο για την αναπαράσταση της γνώσης που καλύπτουν.
- Παράδειγμα: Απόδειξη θεωρημάτων (Theorem proving)
  - Προτασιακή λογική, κατηγορηματική λογική κλπ. για αναπαράσταση γνώσης
  - Modus ponens για διάδοχες καταστάσεις

$$\frac{\begin{array}{c} p \Rightarrow q \\ p \end{array}}{q}$$

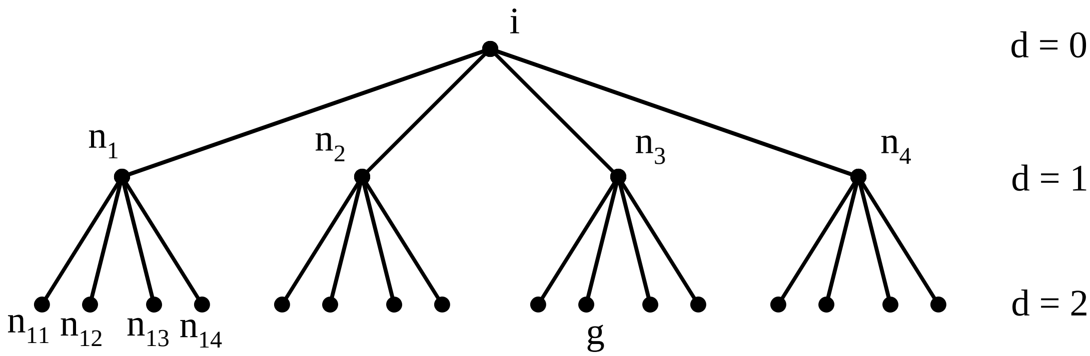
  - Εμπρόσθια ή οπίσθια αναζήτηση;
  - ΚΑΙ–Ή γράφοι (AND–OR graphs): Χρειάζονται;



**ΤΥΦΛΗ ΑΝΑΖΗΤΗΣΗ**

**(BLIND SEARCH)**

**Πρώτα–κατά–πλάτος αναζήτηση**




Διαδοχικές τιμές της L

{i}

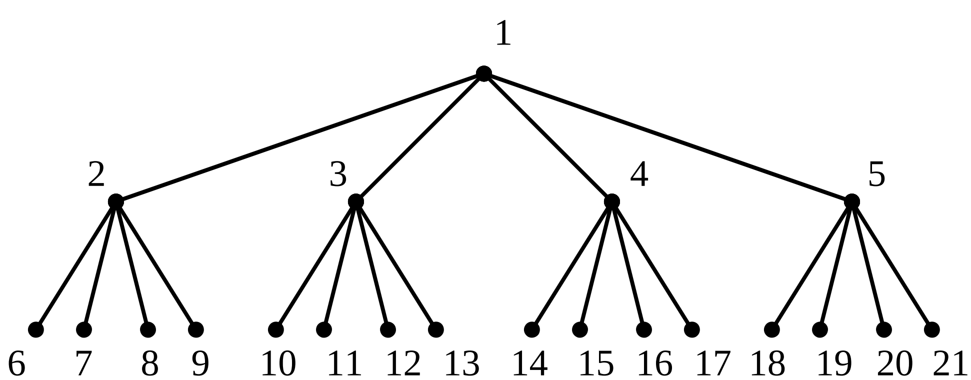
$\{n_1, n_2, n_3, n_4\}$

$\{n_2, n_3, n_4, n_{11}, n_{12}, n_{13}, n_{14}\}$

$\{n_3, n_4, n_{11}, n_{12}, n_{13}, n_{14}, n_{21}, n_{22}, n_{23}, n_{24}\}$

$\{n_4, n_{11}, n_{12}, n_{13}, n_{14}, n_{21}, n_{22}, n_{23}, n_{24}, n_{31}, n_{32}, n_{33}, n_{34}\}$

.......................................



- Απαιτήσεις σε μνήμη της πρώτα–κατά–πλάτος αναζήτησης

  b: παράγοντας διακλάδωσης

  d: βάθος του δέντρου

  Η απαιτούμενη μνήμη είναι ανάλογη του αριθμού των κόμβων–φύλλων, δηλαδή $b^d$.

- Απαιτήσεις σε χρόνο της πρώτα–κατά–πλάτος αναζήτησης

  Ο απαιτούμενος χρόνος είναι ανάλογος του αριθμού των επισκεπτόμενων κόμβων (= όλοι οι κόμβοι στα βάθη από 0 έως d–1 + ο μέσος αριθμός κόμβων στο βάθος d).

  Δηλαδή:

$$(1 + b + b^2 + \ldots\ldots + b^{d-1}) + \frac{b^d + 1}{2} =$$

$$\frac{b^d - 1}{b - 1} + \frac{b^d + 1}{2} = \frac{2\cdot b^d - 2 + b^{d+1} + b - b^d - 1}{2\cdot(b - 1)} =$$

$$\frac{b^{d+1} + b^d + b - 3}{2\cdot(b - 1)}$$

- d >> 1, b >> 1

$$\frac{b^{d+1} + b^d + b - 3}{2\cdot(b - 1)} \overset{d>>1}{\approx} \frac{b^d\cdot(b + 1)}{2\cdot(b - 1)} \overset{b>>1}{\approx} \frac{b^d}{2}$$

- b = 2

$$\frac{b^{d+1} + b^d + b - 3}{2\cdot(b - 1)} = \frac{3\cdot 2^d - 1}{2} \approx 3/2 \cdot 2^d$$

**Πρώτα–κατά–βάθος αναζήτηση**

Διαδοχικές τιμές της L

{i}

$\{n_1, n_2, n_3, n_4\}$

$\{n_{11}, n_{12}, n_{13}, n_{14}, n_2, n_3, n_4\}$

$\{n_{12}, n_{13}, n_{14}, n_2, n_3, n_4\}$

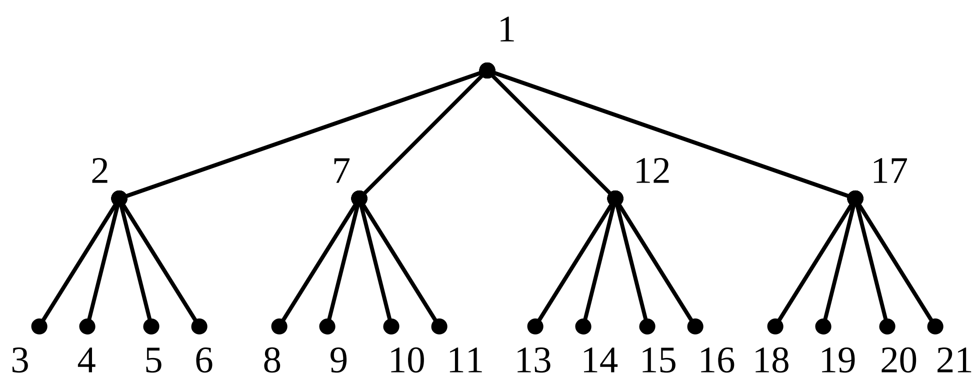
$\{n_{13}, n_{14}, n_2, n_3, n_4\}$

$\{n_{14}, n_2, n_3, n_4\}$

$\{n_2, n_3, n_4\}$

$\{n_{21}, n_{22}, n_{23}, n_{24}, n_3, n_4\}$

.......................................




- Απαιτήσεις σε μνήμη της πρώτα–κατά–βάθος αναζήτησης

  Η απαιτούμενη μνήμη είναι ανάλογη του αριθμού των κόμβων που έχουν φυλαχτεί για μελλοντική εξέταση όταν η αναζήτηση φτάσει για πρώτη φορά σε κόμβο–φύλλο, δηλαδή

  $(d - 1)\cdot(b - 1) + b = d\cdot b - d - b + 1 + b = d\cdot(b - 1) + 1$

Ø Απαιτήσεις σε χρόνο της πρώτα–κατά–βάθος αναζήτησης

Ο απαιτούμενος χρόνος είναι ανάλογος του αριθμού των επισκεπτόμενων κόμβων. Ο ελάχιστος αριθμός είναι d+1 (αν ο κόμβος–στόχος είναι ο πιο αριστερά κόμβος–φύλλο) και ο μέγιστος αριθμός είναι ίσος με το συνολικό αριθμό των κόμβων του δέντρου, δηλαδή $1 + b + .... + b^d = \frac{b^{d+1}-1}{b-1}$ (αν ο κόμβος–στόχος είναι ο πιο δεξιά κόμβος–φύλλο).
Η μέση τιμή του ελάχιστου και του μέγιστου δίνει (γιατί;) τον αναμενόμενο αριθμό των κόμβων που θα εξετασθούν, δηλαδή:

$$\frac{\frac{b^{d+1}-1}{b-1} + d + 1}{2} = \frac{b^{d+1} - 1 + b\cdot d - d + b - 1}{2\cdot(b-1)} = \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)}$$

- d >> 1, b >> 1

$$\frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} \overset{d>>1}{\approx} \frac{b^{d+1}}{2\cdot(b-1)} \overset{b>>1}{\approx} \frac{b^d}{2}$$

- b = 2

$$\frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} = \frac{2^{d+1} + d}{2} \approx 2^d$$

**Σύγκριση μεταξύ πρώτα–κατά–πλάτος και πρώτα–κατά–βάθος αναζήτησης**

- Απαιτήσεις σε μνήμη:
  - Πρώτα–κατά–πλάτος: εκθετική
  - Πρώτα–κατά–βάθος: γραμμική

- Απαιτήσεις σε χρόνο:

$$\frac{\text{πρώτα–κατά–πλάτος}}{\text{πρώτα–κατά–βάθος}} = 1 + \frac{1}{b} = \frac{b+1}{b}$$

| $b$ | $\frac{b+1}{b}$ |
|---|---|
| 2 | 1.5 |
| 3 | 1.3 |
| 5 | 1.2 |
| 10 | 1.1 |
| 25 | 1.04 |
| 100 | 1.01 |

- Η πρώτα–κατά–βάθος αναζήτηση είναι πιο αποδοτική, αλλά .......
  - δεν εγγυάται ότι θα βρει την πιο κοντινή λύση στον αρχικό κόμβο.
  - μπορεί να παγιδευτεί σε ατέρμονα αναζήτηση.

- Το βασικό πρόβλημα της πρώτα–κατά–πλάτος αναζήτησης είναι οι απαγορευτικές απαιτήσεις σε μνήμη.

- **ΕΡΩΤΗΣΗ:**

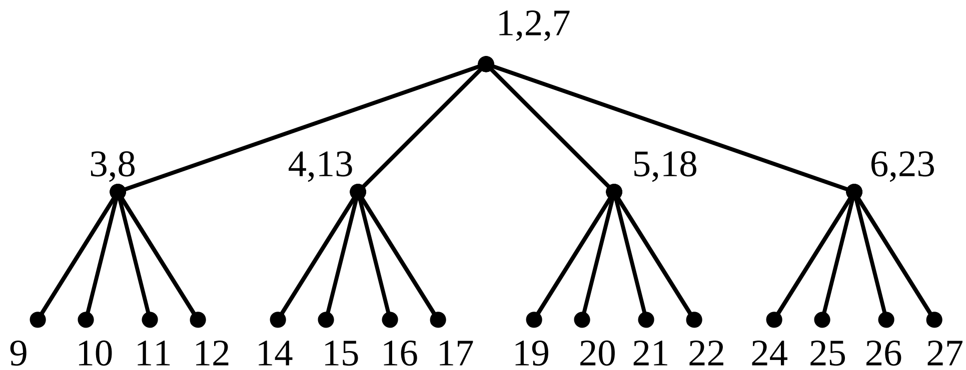
  Μπορούν να συνδυαστούν τα πλεονεκτήματα των δύο μεθόδων σε μία;

  **ΑΠΑΝΤΗΣΗ:**

  **ΝΑΙ**. Είναι η μέθοδος της επαναληπτικής εμβάθυνσης (iterative deepening).

**Επαναληπτική εμβάθυνση**

- Επαναλαμβανόμενες πρώτα–κατά–βάθος αναζητήσεις με συνεχώς αυξανόμενα επάνω φράγματα στο βάθος.




- Η απαιτούμενη μνήμη είναι όση και στην πρώτα–κατά–βάθος αναζήτηση.

- Ο απαιτούμενος χρόνος είναι περισσότερος από την πρώτα–κατά–βάθος αναζήτηση, αλλά πόσο;

Χρειάζονται για κάθε βάθος j (= 0, 1, ...., d–1) εξέταση του αντίστοιχου πλήρους υποδέντρου καθώς και μία τελευταία επιτυχημένη πρώτα–κατά–βάθος αναζήτηση όλου του δέντρου. Συνεπώς, ο αναμενόμενος αριθμός των κόμβων που θα εξετασθούν είναι:

$$\sum_{j=0}^{d-1} (1+b+b^2+.......+b^j) \quad + \quad \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)}$$

$$\sum_{j=0}^{d-1} (1+b+b^2+\ldots\ldots+b^j) + \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} =$$

$$\sum_{j=0}^{d-1} \left( \frac{b^{j+1} - 1}{b-1} \right) + \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} =$$

$$\frac{1}{b-1} \cdot \left[ b \cdot \left( \sum_{j=0}^{d-1} b^j \right) - \sum_{j=0}^{d-1} 1 \right] + \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} =$$

$$\frac{1}{b-1} \cdot \left[ b \cdot \left( \frac{b^d - 1}{b-1} \right) - d \right] + \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} =$$

$$\frac{b^{d+1} - b - b\cdot d + d}{(b-1)^2} + \frac{b^{d+1} + b\cdot d + b - d - 2}{2\cdot(b-1)} =$$

$$\frac{2\cdot b^{d+1} - 2\cdot b - 2\cdot b\cdot d + 2\cdot d + b^{d+2} + b^2\cdot d + b^2 - b\cdot d - 2\cdot b - b^{d+1} - b\cdot d - b + d + 2}{2\cdot(b-1)^2} =$$

$$\frac{b^{d+2} + b^{d+1} + b^2\cdot d + b^2 - 4\cdot b\cdot d - 5\cdot b + 3\cdot d + 2}{2\cdot(b-1)^2}$$

- $d >> 1$

$$\frac{b^{d+2} + b^{d+1} + b^2\cdot d + b^2 - 4\cdot b\cdot d - 5\cdot b + 3\cdot d + 2}{2\cdot(b-1)^2} \approx \frac{(b+1)\cdot b^{d+1}}{2\cdot(b-1)^2}$$

$$\frac{\text{επαναληπτική εμβάθυνση}}{\text{πρώτα–κατά–βάθος}} = \frac{\dfrac{(b+1)\cdot b^{d+1}}{2\cdot(b-1)^2}}{\dfrac{b^{d+1}}{2\cdot(b-1)}} = \frac{b+1}{b-1}$$

| $b$ | $\frac{b+1}{b-1}$ |
|---|---|
| 2 | 3 |
| 3 | 2 |
| 5 | 1.5 |
| 10 | 1.2 |
| 25 | 1.08 |
| 100 | 1.02 |

- $b = 2$

$$\frac{b^{d+2} + b^{d+1} + b^2 \cdot d + b^2 - 4 \cdot b \cdot d - 5 \cdot b + 3 \cdot d + 2}{2 \cdot (b-1)^2} = \frac{2^{d+2} + 2^{d+1} - d - 4}{2} \approx$$

$3 \cdot 2^d$

πρώτα–κατά–βάθος **:** πρώτα–κατά–πλάτος **:** επαναληπτική εμβάθυνση =

$2^d$ **:** $3/2 \cdot 2^d$ **:** $3 \cdot 2^d$

- **ΕΡΩΤΗΣΗ:**

  Είναι δυνατόν μια μέθοδος τυφλής αναζήτησης να απαιτεί λιγότερο χρόνο από εκθετικό ως προς το βάθος;

  **ΑΠΑΝΤΗΣΗ**:

  **ΟΧΙ.** Γιατί το πλήθος των κόμβων–φύλλων, που είναι υποψήφιοι στόχοι, είναι εκθετικό ως προς το βάθος.

- **ΕΡΩΤΗΣΗ:**

  Είναι δυνατόν μια μέθοδος τυφλής αναζήτησης να απαιτεί χώρο λιγότερο από γραμμικό ως προς το βάθος;

  **ΑΠΑΝΤΗΣΗ:**

  **ΟΧΙ.** Γιατί κάθε εκθετικός ως προς το χρόνο αλγόριθμος είναι τουλάχιστον γραμμικός ως προς το χώρο.

  (Hopcroft & Ullman)

- **ΕΡΩΤΗΣΗ:**

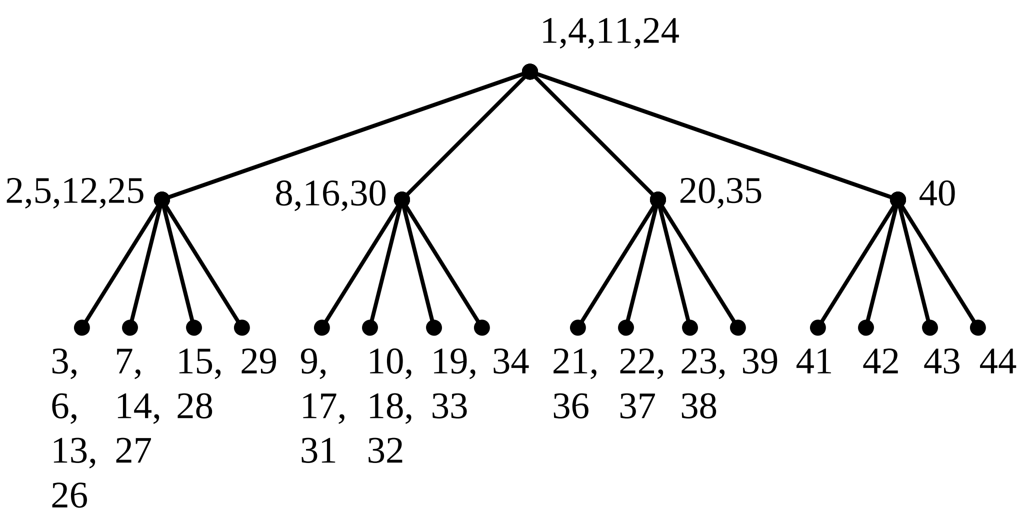
  Αφού λοιπόν μία μέθοδος τυφλής αναζήτησης δεν μπορεί να έχει πολυπλοκότητα ως προς χρόνο καλύτερη από εκθετική και ως προς μνήμη καλύτερη από γραμμική, είναι δυνατόν να ελπίζουμε σε τίποτα καλύτερο από τη μέθοδο της επαναληπτικής εμβάθυνσης;

  **ΑΠΑΝΤΗΣΗ:**

  **ΝΑΙ**. Σε ορισμένες περιπτώσεις (αν υπάρχουν πολλοί κόμβοι–στόχοι) η μέθοδος της επαναληπτικής διεύρυνσης (iterative broadening) αποδεικνύεται καλύτερη στην πράξη, χωρίς να έχει καλύτερη πολυπλοκότητα.

**Επαναληπτική διεύρυνση**

- Επαναλαμβανόμενες πρώτα–κατά–βάθος αναζητήσεις με συνεχώς αυξανόμενα επάνω φράγματα στο πλήθος των εξεταζόμενων κόμβων–παιδιών.



- Στη μέθοδο της επαναληπτικής διεύρυνσης, η απαιτούμενη μνήμη δεν είναι μεγαλύτερη αυτής της πρώτα–κατά–βάθος αναζήτησης.

- Στη μέθοδο της επαναληπτικής διεύρυνσης, θα γίνουν, στη χειρότερη περίπτωση, b πρώτα–κατά–πλάτος αναζητήσεις με παράγοντες διακλάδωσης 1, 2, 3, ........, b, οπότε ο συνολικός απαιτούμενος χρόνος είναι:

  $$\frac{1}{2} \cdot (1 + 2^d + 3^d + ....... + b^d) \approx \frac{1}{2} \cdot \frac{b^{d+1}}{d+1}$$

  $\left(\frac{b}{d+1}\right.$ φορές χειρότερα από την πρώτα–κατά–βάθος αναζήτηση$\left.\right)$

- Η μέθοδος της επαναληπτικής διεύρυνσης υπήρξε η αφορμή για την εισαγωγή της μη–συστηματικής (nonsystematic) αναζήτησης, που έχει δώσει σημαντικά αποτελέσματα σε προβλήματα με εξαιρετικά μεγάλους χώρους αναζήτησης.

**Αναζήτηση σε γράφους**

- Ανοικτές και κλειστές λίστες
- Δυναμική οπισθοδρόμηση (dynamic backtracking)

## ΕΥΡΙΣΤΙΚΗ ΑΝΑΖΗΤΗΣΗ
## (HEURISTIC SEARCH)

- Η εκθετική πολυπλοκότητα χρόνου των μεθόδων τυφλής αναζήτησης είναι ασύμφορη για πραγματικά προβλήματα.

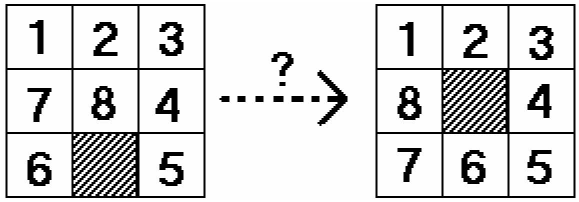
- Η μόνη λύση είναι η αντικατάσταση της τυφλής επιλογής κόμβου για εξέταση με την επιλογή κόμβου σύμφωνα με κάποια γνώση που υπάρχει για το συγκεκριμένο πρόβλημα.

- Η βασική ιδέα είναι να επιλέγεται κάθε φορά για εξέταση ο κόμβος που φαίνεται να είναι ο περισσότερα υποσχόμενος, όπως αυτό αποφασίζεται από μία ευριστική συνάρτηση (heuristic function).

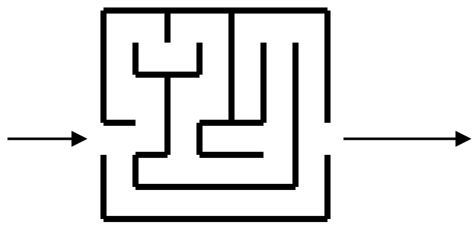
- Πιθανά προβλήματα της ευριστικής συνάρτησης
    – Πόσο ακριβής είναι;
    – Πόση προσπάθεια μετα–επιπέδου απαιτείται;
    – Θα βρεθεί η καλύτερη λύση;

- Σειρά εξέτασης κόμβων
    – Μόνο σε ένα μονοπάτι;
    – Οπουδήποτε στο χώρο αναζήτησης;

- Παραδείγματα
  - – Γρίφος–8 και πλήθος τετραγωνιδίων που είναι λάθος τοποθετημένα



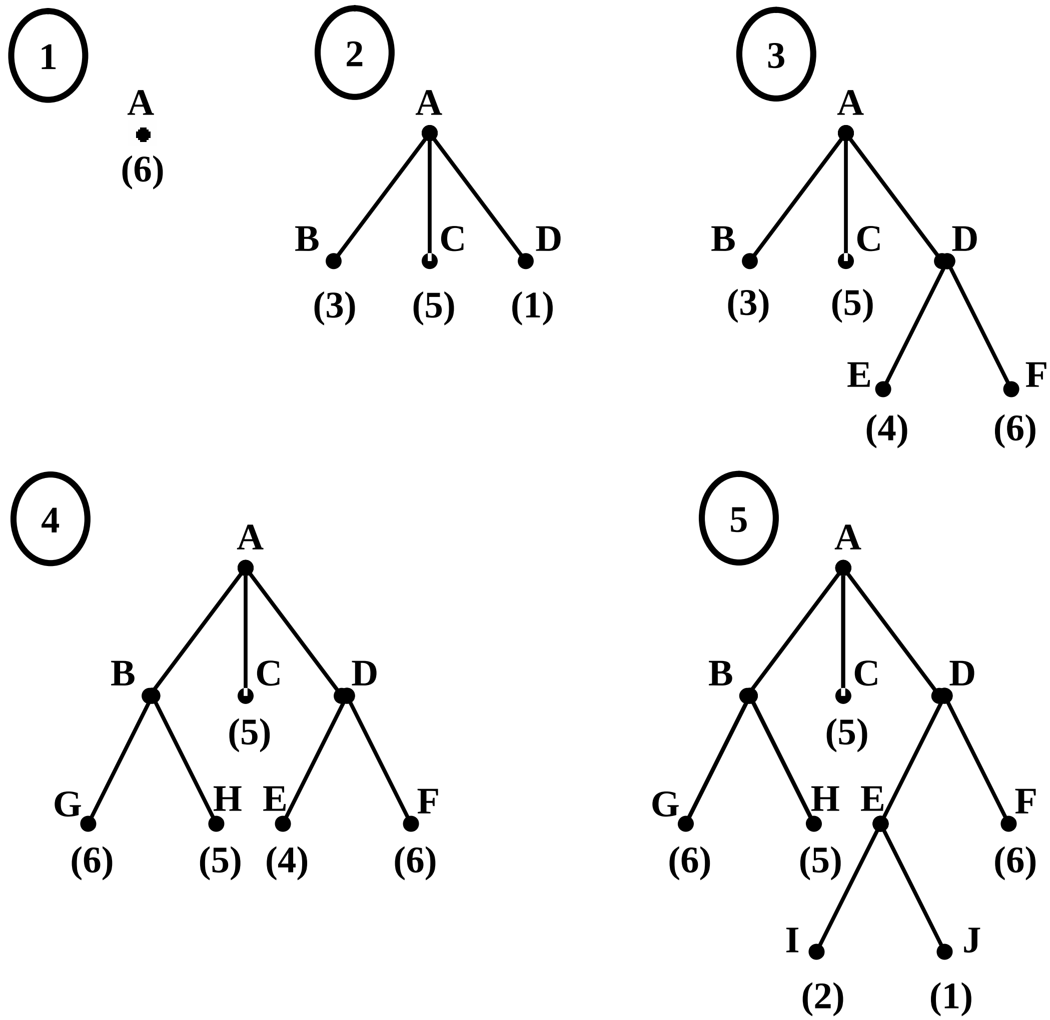

  - – Λαβύρινθος και απόσταση Manhattan για την έξοδο



- Ευριστικές μέθοδοι αναζήτησης που θα εξετασθούν
  - – Πρώτα–ο–καλύτερος (Best–first)
  - – Αναρρίχηση λόφου (Hill climbing)
  - – Προσομοιωμένη ανόπτηση (Simulated annealing)
  - – $A^*$
  - – $IDA^*$ (ID = Iterative Deepening)

**Πρώτα–ο–καλύτερος αναζήτηση**

- Στο βήμα 2 της γενικής μεθόδου αναζήτησης επιλέγεται από την L ο κόμβος που βρίσκεται πλησιέστερα σε κόμβο–στόχο, όπως αυτό αποφασίζεται από μία κατάλληλα επιλεγμένη ευριστική συνάρτηση.




L = {A/6}

L = {B/3, C/5, D/1}

L = {B/3, C/5, E/4, F/6}

L = {G/6, H/5, C/5, E/4, F/6}

L = {G/6, H/5, C/5, I/2, J/1, F/6}

..........................................................

**Αναρρίχηση λόφου**

- Η μέθοδος της αναρρίχησης λόφου αποτελεί μία εξειδίκευση της "πρώτα–ο–καλύτερος" μεθόδου με τα εξής χαρακτηριστικά
  – Η αναζήτηση γίνεται κατά μήκος ενός μόνο μονοπατιού στο χώρο αναζήτησης.
  – Η αναζήτηση προχωρά μόνο από διαδοχικά καλύτερους κόμβους.

Βήμα 1: Τρέχων κόμβος είναι ένας από τους αρχικούς κόμβους.

Βήμα 2: Αν ο τρέχων κόμβος είναι κόμβος–στόχος, τότε τελείωσες με επιτυχία.

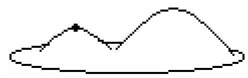
Βήμα 3: Αλλιώς, αν δεν υπάρχει απόγονος του τρέχοντος κόμβου που είναι "καλύτερος" από αυτόν, τότε απέτυχες.

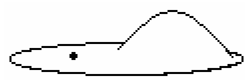
Βήμα 4: Τρέχων κόμβος είναι ένας από τους "καλύτερους" απογόνους του τρέχοντος κόμβου.

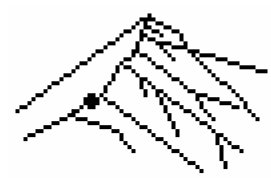
Βήμα 5: Πήγαινε στο βήμα 2.

- Αν στο βήμα 4 επιλέγεται ο "καλύτερος" από τους απογόνους, τότε η μέθοδος ονομάζεται αναρρίχηση λόφου από την πιο απότομη πλαγιά (steepest-ascent hill climbing).
- Η μέθοδος της αναρρίχησης λόφου είναι πολύ αποδοτική σε ορισμένα προβλήματα και εφ' όσον επιλεγεί κατάλληλη ευριστική συνάρτηση.

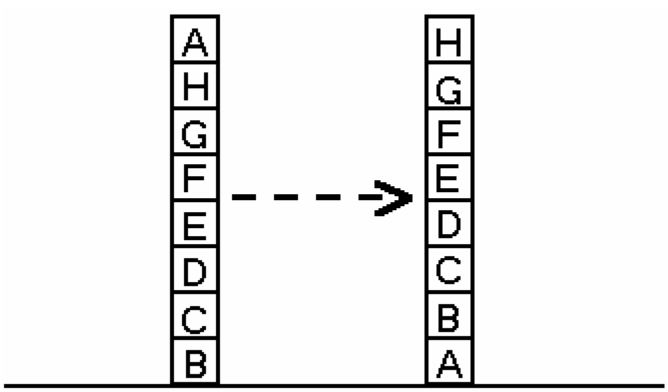
- Προβλήματα της αναρρίχησης λόφου

– τοπικά μέγιστα  (local maximum)

– οροπέδια  (plateau)

– κορυφογραμμές  (ridge)

- Εφαρμογή της αναρρίχησης λόφου στο πρόβλημα του κόσμου των κύβων




αρχική κατάσταση τελική κατάσταση

$f_1 = 4$ $f_1 = 8$

$f_2 = -28$ $f_2 = 28$

Πιθανές ευριστικές συναρτήσεις

– Για κάθε κύβο που βρίσκεται επάνω εκεί που πρέπει να βρίσκεται, πρόσθεσε μία μονάδα. Για κάθε έναν από τους άλλους κύβους, αφαίρεσε μία μονάδα.

– Για κάθε κύβο που έχει τη σωστή δομή υποστήριξης, πρόσθεσε μία μονάδα για κάθε κύβο στη δομή υποστήριξης. Για κάθε έναν από τους άλλους κύβους, αφαίρεσε μία μονάδα για κάθε κύβο στη δομή υποστήριξης.

**Προσομοιωμένη ανόπτηση**

- Η μέθοδος της προσομοιωμένης ανόπτησης είναι μία παραλλαγή της αναρρίχησης λόφου ως προς το ότι στα αρχικά στάδια της αναζήτησης είναι δυνατή η μετάβαση από ένα κόμβο σε κάποιο “χειρότερο” απόγονό του. Η πιθανότητα να συμβεί κάτι τέτοιο μειώνεται με την πάροδο του χρόνου.

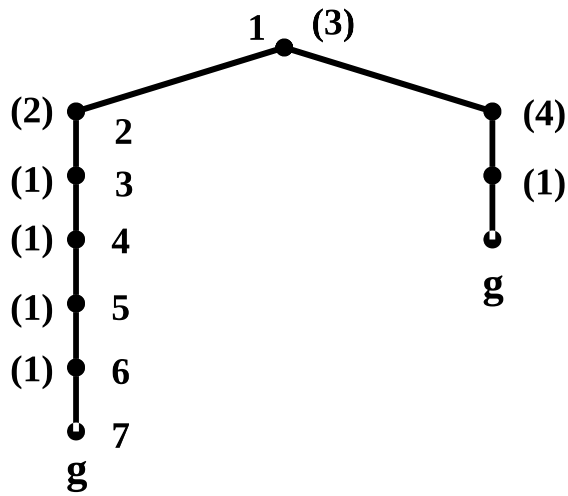
- Η ιδέα προέρχεται από τη σταδιακή ανόπτηση (annealing) ενός τηγμένου μετάλλου μέσω κατάλληλα σχεδιασμένου χρονοδιαγράμματος μείωσης της θερμοκρασίας. Το μέταλλο μεταβαίνει σε καταστάσεις χαμηλότερης ενέργειας, αλλά σε θερμοκρασία Τ είναι δυνατόν να συμβεί αύξηση ενέργειας ΔΕ με πιθανότητα $p = e^{-\Delta E / kT}$ (k είναι η σταθερά Boltzmann).

- Στη μέθοδο της προσομοιωμένης ανόπτησης είναι πάντα δυνατή η μετάβαση σε “καλύτερους” απογόνους, όπως στην αναρρίχηση λόφου, αλλά επιτρέπεται επίσης και η μετάβαση σε “χειρότερους” απογόνους με πιθανότητα $p' = e^{-\Delta f / T'}$, όπου $\Delta f$ είναι η μεταβολή της τιμής της ευριστικής συνάρτησης μεταξύ του τρέχοντος κόμβου και του απογόνου του και $T'$ είναι η “θερμοκρασία” της αναζήτησης.

- Με κατάλληλη εμπειρική επιλογή του χρονοδιαγράμματος μείωσης της “θερμοκρασίας”, η προσομοιωμένη ανόπτηση μπορεί να δώσει λύση εκεί που η αναρρίχηση λόφου αποτυγχάνει.

**Η ευριστική μέθοδος αναζήτησης $A^*$**

- Η μέθοδος αναζήτησης “πρώτα–ο–καλύτερος” δεν εγγυάται ότι η λύση που θα βρεθεί είναι η καλύτερη δυνατή, δηλαδή η πιο κοντινή στην αρχική κατάσταση. Επίσης η επιλογή του κόμβου που θα εξετασθεί γίνεται μόνο με βάση το πόσο κοντά φαίνεται να είναι ο κόμβος αυτός σε ένα κόμβο–στόχο και χωρίς να λαμβάνεται υπόψη το κόστος από την αρχική κατάσταση σ’ αυτόν τον κόμβο.



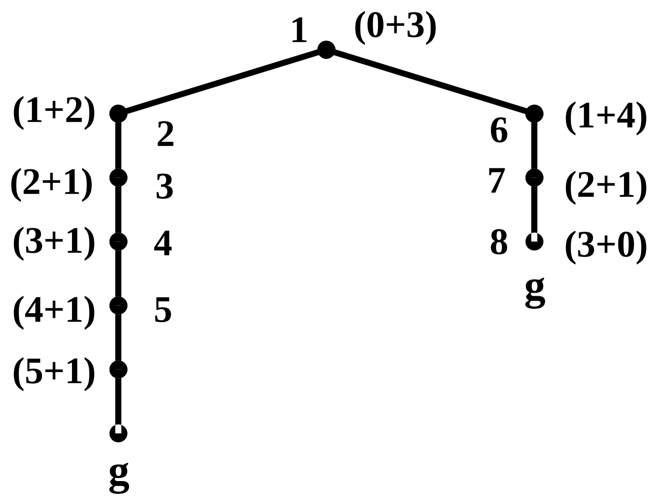

- Για να αντιμετωπισθούν τα προβλήματα της μεθόδου “πρώτα–ο–καλύτερος”, εισάγονται οι εξής συναρτήσεις:

$g(n)$: Το κόστος να μεταβεί κάποιος από τον αρχικό κόμβο στον κόμβο n. Συνήθως, είναι το βάθος του κόμβου n.

$h'(n)$: Μία ευριστική εκτίμηση του κόστους να μεταβεί κάποιος από τον κόμβο σε κόμβο–στόχο.

$f(n) = g(n) + h'(n)$: Μία ευριστική εκτίμηση του κόστους να μεταβεί κάποιος από τον αρχικό κόμβο σε κόμβο–στόχο μέσω του κόμβου n.

- Όλες οι συναρτήσεις g(n), h΄(n) και f(n), σαν συναρτήσεις κόστους, θεωρούνται πάντα ≥ 0.
- Εφαρμογή της f(n) = g(n) + h΄(n)



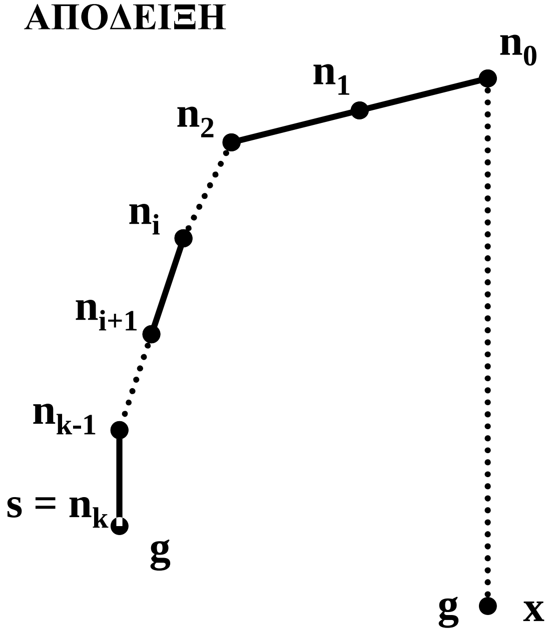

- Η ευριστική μέθοδος αναζήτησης $A^*$ προκύπτει από τη γενική μέθοδο αναζήτησης, όταν στο βήμα 2 της τελευταίας επιλέγεται ο κόμβος n που έχει τη μικρότερη τιμή f(n) = g(n) + h΄(n).
  – Γιατί επιλέχθηκε να εξετασθεί πρώτα ο κόμβος 5 και μετά ο 6;
  – Τι θα συνέβαινε αν ο κόμβος 5 ήταν κόμβος–στόχος;
  – Πότε η μέθοδος $A^*$ εγγυάται ότι θα βρει την καλύτερη λύση;
  – Χρειάζεται να τροποποιηθεί η μέθοδος $A^*$ για να δουλέψει και σε γράφους;

- Έστω h(n) το πραγματικό κόστος μετάβασης από τον κόμβο n σε κόμβο–στόχο και έστω h΄(n) η εκτίμηση αυτού του κόστους από τη θεωρούμενη ευριστική συνάρτηση. Η συνάρτηση h΄(n) λέγεται ότι είναι παραδεκτή (admissible) όταν για κάθε κόμβο n δεν υπερεκτιμά το πραγματικό κόστος, δηλαδή ισχύει $h΄(n) \le h(n)$.

**ΘΕΩΡΗΜΑ**

Η μέθοδος $A^*$ βρίσκει την καλύτερη λύση όταν χρησιμοποιεί παραδεκτή ευριστική συνάρτηση h΄(n).

**ΑΠΟΔΕΙΞΗ**




Έστω ότι παρά το γεγονός ότι η συνάρτηση h΄(n) είναι παραδεκτή, η μέθοδος $A^*$ επιστρέφει σαν λύση τον κόμβο x, ενώ υπάρχει και καλύτερη λύση, ο κόμβος s.

Ισχύουν τα εξής:

$f(n_{i+1}) = g(n_{i+1}) + h΄(n_{i+1}) \le$

$\le g(n_{i+1}) + h(n_{i+1}) = g(s) <$

$< g(x) = g(x) + h΄(x) = f(x)$

Δηλαδή, $f(n_{i+1}) < f(x)$. Συνεπώς ο κόμβος $n_{i+1}$ θα εξετασθεί πριν από τον κόμβο x, εφ’ όσον έχει εξετασθεί προηγουμένως ο κόμβος $n_i$. Τελικά, συμπεραίνεται επαγωγικά ότι αφού οπωσδήποτε θα εξετασθεί ο κόμβος $n_0$, τότε και ο κόμβος $n_k = s$ θα εξετασθεί πριν από τον κόμβο x, όπερ άτοπον.

- Με την ευριστική συνάρτηση $h'(n) = 0$, η μέθοδος $A^*$ συμπίπτει με την πρώτα–κατά–πλάτος.

- Με την ευριστική συνάρτηση $h'(n) = h(n)$, η μέθοδος $A^*$ γίνεται η ιδανική μέθοδος αναζήτησης, δηλαδή κινείται συνεχώς προς το στόχο.

- Μεταξύ δύο ευριστικών συναρτήσεων $h_1'(n)$ και $h_2'(n)$ για τις οποίες ισχύει $h_1'(n) \leq h_2'(n) \leq h(n)$ προτιμότερη είναι η $h_2'(n)$. Παράδειγμα από το πρόβλημα του γρίφου–8:
  – $h_1'(n)$ = αριθμός τετραγωνιδίων που είναι λάθος τοποθετημένα
  – $h_2'(n)$ = συνολική απόσταση Manhattan των λάθος τοποθετημένων τετραγωνιδίων

- Για να εφαρμοσθεί η μέθοδος $A^*$ σε γράφους, αρκεί οι απόγονοι του εξεταζόμενου κόμβου να μην εισάγονται στη λίστα L, εφ’ όσον ήδη βρίσκονται σ’ αυτήν, αλλά να αναθεωρείται το μονοπάτι από τον αρχικό κόμβο, καθώς και το αντίστοιχο κόστος, ανάλογα με τη μέχρι στιγμής καλύτερη εκδοχή.

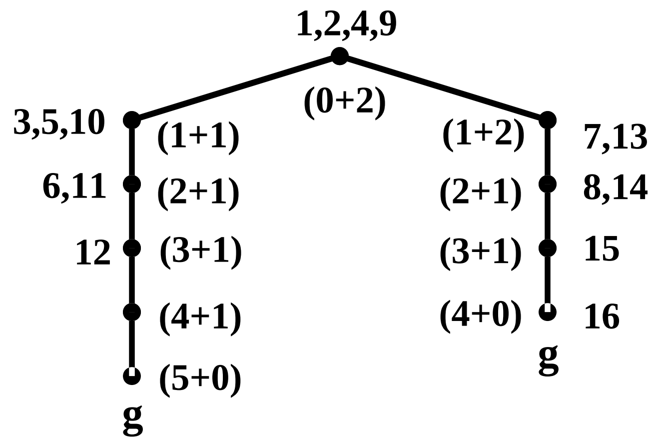
- Η μέθοδος $A^*$ μπορεί να είναι πολύ απαιτητική σε μνήμη, ανάλογα και με τη χρησιμοποιούμενη ευριστική συνάρτηση $h'(n)$.

  **ΕΡΩΤΗΣΗ:** Μπορεί να βελτιωθεί;

  **ΑΠΑΝΤΗΣΗ:** **ΝΑΙ**. Με τη μέθοδο IDA$^*$.

**Η ευριστική μέθοδος αναζήτησης IDA***

- Η μέθοδος IDA* συνίσταται σε διαδοχικές πρώτα–κατά–βάθος αναζητήσεις με συνεχώς αυξανόμενο επάνω φράγμα, όχι στο βάθος του κόμβου n, αλλά στην τιμή της συνάρτησης f(n).
- Αρχική τιμή του επάνω φράγματος είναι 1 και σε κάθε πρώτα–κατά–βάθος αναζήτηση το επάνω φράγμα ισούται με την ελάχιστη υπέρβαση της προηγούμενης αναζήτησης.
- Η μέθοδος IDA* απαιτεί γραμμική ποσότητα μνήμης ως προς το βάθος του κόμβου–στόχου. Η ευριστική συνάρτηση f(n) χρησιμοποιείται για τη διακοπή της τρέχουσας πρώτα–κατά–βάθος αναζήτησης και όχι για τον καθορισμό της σειράς εξέτασης των κόμβων.
- Η μέθοδος IDA* εξετάζει μόνο τους κόμβους που εξετάζει και η A* και αποδεικνύεται ότι σε περίπτωση παραδεκτής ευριστικής συνάρτησης h΄(n) βρίσκει την καλύτερη λύση.



## ΑΝΑΖΗΤΗΣΗ ΣΕ ΠΑΙΓΝΙΔΙΑ
## (ADVERSARY SEARCH)

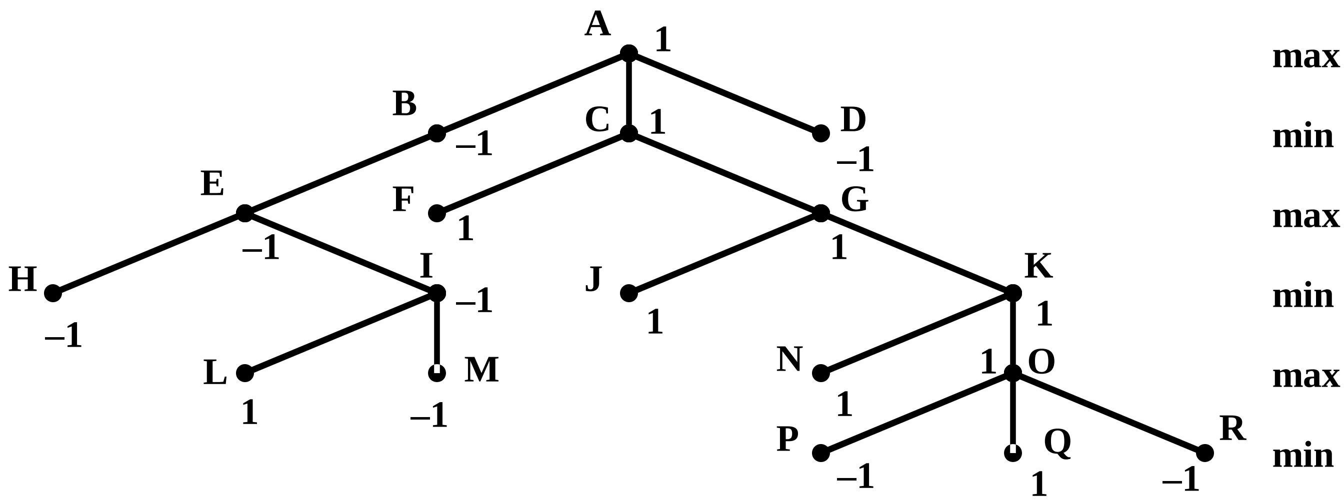
- Παιγνίδια δύο παικτών με εναλλασσόμενες κινήσεις

- Παιγνίδια τέλειας πληροφορίας (perfect information)
  - Σκάκι
  - Ντάμα
  - Τρίλιζα
  - Τάβλι
  - Othello (Reversi)

- Παιγνίδια ατελούς πληροφορίας είναι εκείνα στα οποία οι δύο παίκτες δεν έχουν τα ίδια διαθέσιμα δεδομένα (π.χ. μπριτζ, πόκερ, ναυμαχία, στρατέγκο κλπ.).

- Στην αναζήτηση σε παιγνίδια, για να αποφασίσει κάποιος παίκτης ποια κίνηση θα επιλέξει, πρέπει να εξετάσει με ποιο τρόπο, μετά από κάποιες κινήσεις, θα καταφέρει να κερδίσει τον αντίπαλο του ή, τουλάχιστον, να βρεθεί σε πλεονεκτικότερη θέση απ’ αυτόν.

- Κλασική μέθοδος για αναζήτηση σε παιγνίδια είναι η μέθοδος minimax, στην οποία υπάρχουν ο max παίκτης που προσπαθεί να μεγιστοποιήσει το όφελος του και ο min παίκτης που προσπαθεί να ελαχιστοποιήσει το όφελος του max.




- Τερματικοί κόμβοι με τιμή 1 είναι οι κόμβοι νίκης για τον max, ενώ με τιμή –1 είναι οι κόμβοι νίκης για τον min παίκτη.

- Στη μέθοδο minimax διαδίδονται προς τα επάνω οι τιμές των τερματικών κόμβων με τον κανόνα “ένας max κόμβος έχει τιμή τη μέγιστη τιμή των παιδιών του και ένας min κόμβος έχει τιμή την ελάχιστη τιμή των παιδιών του”.

- Ο στόχος στη μέθοδο minimax είναι να αποφασισθεί η κίνηση που πρέπει να κάνει ο παίκτης που έχει σειρά να παίξει στον κόμβο–ρίζα του χώρου αναζήτησης. Στο παράδειγμα, ο max παίκτης στον κόμβο A πρέπει να κάνει την κίνηση προς τον κόμβο C, γιατί αυτή είναι που θα τον οδηγήσει σε νίκη.

- Κάθε επίπεδο στο χώρο αναζήτησης αναφέρεται και σαν στρώμα (ply).

**Minimax**: Υπολογισμός της τιμής ενός κόμβου n

Βήμα 1 : Γέννησε όλους τους κόμβους κάτω από τον n.

Βήμα 2 : Δώσε την κατάλληλη τιμή στους τερματικούς κόμβους, ανάλογα αν είναι κόμβοι νίκης για τον max ή τον min παίκτη.

Βήμα 3 : Διάλεξε έναν κόμβο χωρίς τιμή του οποίου όλα τα παιδιά έχουν τιμή. Αν δεν υπάρχει τέτοιος κόμβος να επιστρέψεις την τιμή του κόμβου n.

Βήμα 4 : Αν ο κόμβος που διάλεξες είναι max κόμβος δώσε του τιμή τη μέγιστη των παιδιών του. Αν είναι min κόμβος δώσε του τιμή την ελάχιστη τιμή των παιδιών του. Πήγαινε στο βήμα 3.

- Η μέθοδος καλύπτει και την περίπτωση της ισοπαλίας. Απλώς ο αντίστοιχος τερματικός κόμβος πρέπει να πάρει την τιμή 0.

- Όπως περιγράφηκε η μέθοδος minimax οδηγεί σε μία πρώτα–κατά–πλάτος αναζήτηση με τις γνωστές απαιτήσεις σε μνήμη.

**ΕΡΩΤΗΣΗ:** Μπορεί η μέθοδος minimax να μην είναι τόσο “χωροβόρα”;

**ΑΠΑΝΤΗΣΗ:** **ΝΑΙ**. Αν ακολουθήσει μία πρώτα–κατά–βάθος αναζήτηση.

**Minimax**: Υπολογισμός της τιμής ενός κόμβου n με πρώτα–κατά–βάθος αναζήτηση

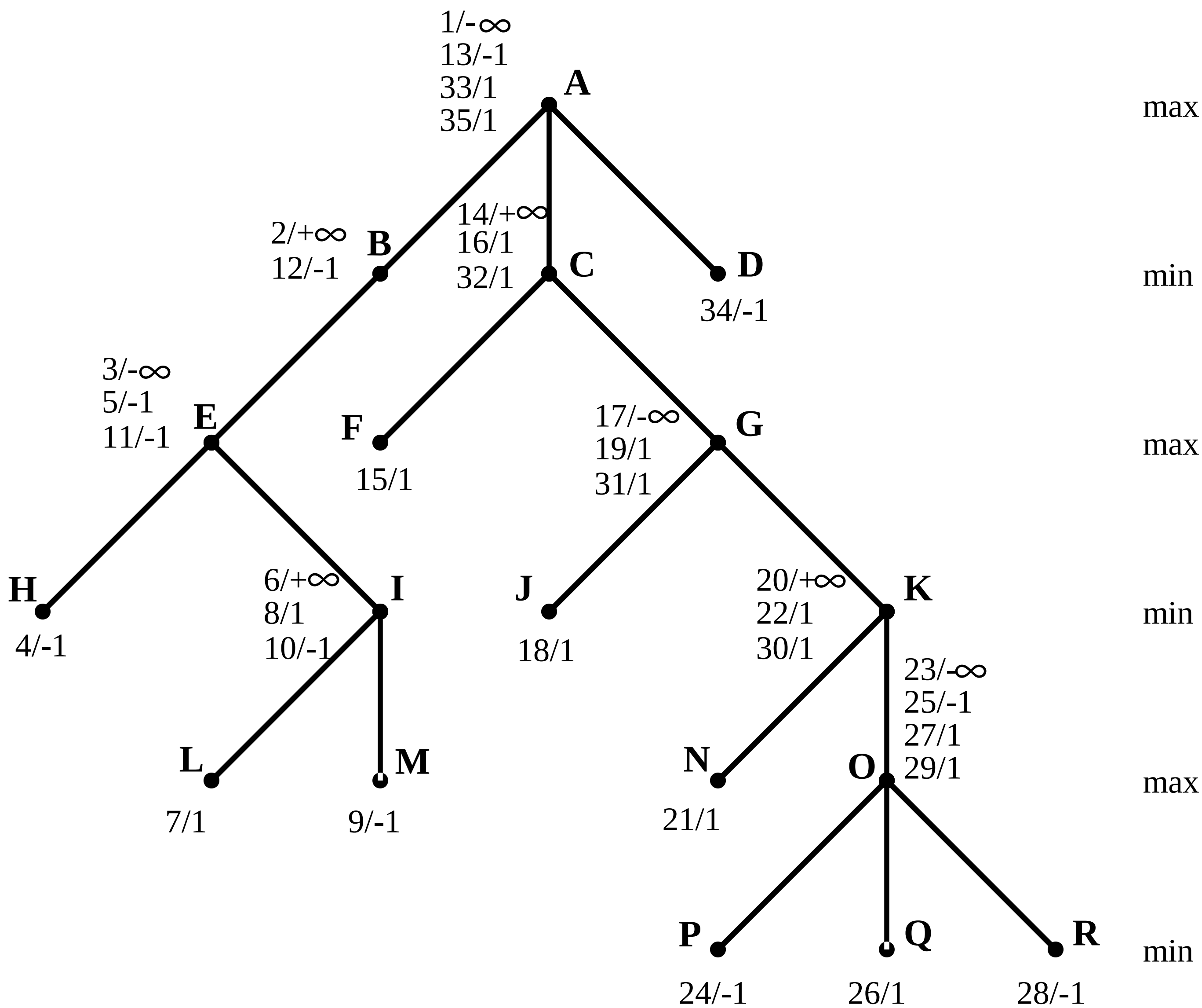
Βήμα 1 : Θέσε L = {n}.

Βήμα 2 : Έστω x ο πρώτος κόμβος της L. Αν x = n και υπάρχει τιμή γι' αυτόν, να επιστρέψεις αυτήν την τιμή.

Βήμα 3 : Αν ο x έχει τιμή $v_x$, έστω p ο πατέρας του x και $v_p$ η τιμή του. Αν ο p είναι min κόμβος, θέσε $v_p = \min(v_p, v_x)$. Αν ο p είναι max κόμβος, θέσε $v_p = \max(v_p, v_x)$. Αφαίρεσε τον x από την L και πήγαινε στο βήμα 2.

Βήμα 4 : Αν ο x δεν έχει τιμή και είναι τερματικός κόμβος, δώσε του την τιμή 1, –1 ή 0, ανάλογα αν είναι κόμβος νίκης για τον max παίκτη, κόμβος νίκης για τον min παίκτη ή κόμβος ισοπαλίας. Πήγαινε στο βήμα 2.

Βήμα 5 : Αν ο x δεν έχει τιμή και δεν είναι τερματικός κόμβος, δώσε του τιμή $-\infty$ αν είναι max κόμβος ή τιμή $+\infty$ αν είναι min κόμβος. Βάλε τα παιδιά του x στην αρχή της L. Πήγαινε στο βήμα 2.




**ΕΡΩΤΗΣΗ:** Έχει πρακτική χρησιμότητα η πρώτα–κατά–βάθος minimax σε πραγματικά παιγνίδια;

**ΑΠΑΝΤΗΣΗ:** **ΟΧΙ**. Οι απαιτήσεις χρόνου είναι απαγορευτικές (εκθετικές ως προς το βάθος).

- Η πλήρης ανάπτυξη του χώρου αναζήτησης είναι πρακτικά αδύνατη σε πραγματικά παιγνίδια.
- Είναι δυνατόν όμως να γίνει ανάπτυξη μέχρις ένα συγκεκριμένο βάθος και εκεί να χρησιμοποιηθεί μία ευριστική συνάρτηση e(n) που να εκτιμά την ποιότητα του κόμβου n ως προς τις δυνατότητες νίκης του max ή του min παίκτη.
- Η διάδοση προς τα επάνω των τιμών της ευριστικής συνάρτησης γίνεται όπως και στην περίπτωση της πλήρους ανάπτυξης του χώρου αναζήτησης.
- Συνήθως, για την ευριστική συνάρτηση ισχύει

$$-1 \leq e(n) \leq 1$$

- Αν η ευριστική συνάρτηση είναι αρκετά επιτυχής, η μέθοδος minimax με πρώτα–κατά–βάθος αναζήτηση και μερική ανάπτυξη έχει εξαιρετικά αποτελέσματα.
- Παράδειγμα ευριστικής συνάρτησης στο σκάκι:

$$e(n) = \frac{w(n) - b(n)}{w(n) + b(n)}$$

w(n) και b(n) είναι οι τιμές “πλεονεκτήματος” του παίκτη με τα άσπρα ή τα μαύρα αντίστοιχα στην κατάσταση n και κάθε μία από αυτές τις τιμές προκύπτει αθροίζοντας κατάλληλα βάρη για κάθε κομμάτι (1 για τα πιόνια, 3 για τους αξιωματικούς, 5 για τους πύργους κλπ.). Η e(n) μπορεί να εμπλουτισθεί και με άλλα κριτήρια (π.χ. προστασία του βασιλιά, έλεγχος του κέντρου κλπ.).

**Minimax**: Υπολογισμός της τιμής ενός κόμβου n με πρώτα–κατά–βάθος αναζήτηση **και μερική ανάπτυξη**

Βήμα 1 : Θέσε L = {n}.

Βήμα 2 : Έστω x ο πρώτος κόμβος της L. Αν x = n και υπάρχει τιμή γι’ αυτόν, να επιστρέψεις αυτήν την τιμή.

Βήμα 3 : Αν ο x έχει τιμή $v_x$, έστω p ο πατέρας του x και $v_p$ η τιμή του. Αν ο p είναι min κόμβος, θέσε $v_p = \min(v_p, v_x)$. Αν ο p είναι max κόμβος, θέσε $v_p = \max(v_p, v_x)$. Αφαίρεσε τον x από την L και πήγαινε στο βήμα 2.

~~Βήμα 4 : Αν ο x δεν έχει τιμή και είναι τερματικός κόμβος, δώσε του την τιμή 1, –1 ή 0, ανάλογα αν είναι κόμβος νίκης για τον max παίκτη, κόμβος νίκης για τον min παίκτη ή κόμβος ισοπαλίας. Πήγαινε στο βήμα 2.~~

Βήμα 5 : ~~Αν ο x δεν έχει τιμή και δεν είναι τερματικός κόμβος~~, δώσε του τιμή $-\infty$ αν είναι max κόμβος ή τιμή $+\infty$ αν είναι min κόμβος. Βάλε τα παιδιά του x στην αρχή της L. Πήγαινε στο βήμα 2.

Αλλιώς

Βήμα 4 : Αν ο x δεν έχει τιμή και είτε είναι τερματικός κόμβος είτε έχουμε αποφασίσει να μην κάνουμε περαιτέρω ανάπτυξη του χώρου αναζήτησης, δώσε του τιμή σύμφωνα με την ευριστική συνάρτηση που έχει επιλεγεί. Πήγαινε στο βήμα 2.

- Προβλήματα της μεθόδου minimax
  - Ασταθής συμπεριφορά της ευριστικής συνάρτησης σε τμήματα του χώρου αναζήτησης (π.χ. φάση ανταλλαγής κομματιών στο σκάκι).
  - Αδυναμία έγκαιρης αντιμετώπισης ενός προβλήματος λόγω της μετάθεσης του πέρα από το βάθος ανάπτυξης που έχει προεπιλεγεί εξ αιτίας κάποιας άσκοπης ενδιάμεσης κίνησης.
  - Τα προηγούμενα προβλήματα, που αναφέρονται σαν έλλειψη σταθερότητας (quiescence) και σύνδρομο του ορίζοντα (horizon effect), είναι πολύ δύσκολο να αντιμετωπισθούν.
  - Μία στοιχειωδώς αποδεκτή μέθοδος αντιμετώπισης είναι αυτή της δευτερεύουσας αναζήτησης (secondary search) σε βάθος μεγαλύτερο από αυτό που έχει αποφασισθεί και στα τμήματα εκείνα του χώρου αναζήτησης που είναι πιο πιθανό να ακολουθηθούν.

**ΕΡΩΤΗΣΗ:** Μήπως η μέθοδος minimax κάνει και άσκοπες αναζητήσεις (π.χ. δέντρο κάτω από τον κόμβο K στο παράδειγμα που έχει παρουσιασθεί); Είναι δυνατόν να βελτιωθεί;

**ΑΠΑΝΤΗΣΗ:** **ΝΑΙ**. Η μέθοδος αναζήτησης α–β είναι η βελτιωμένη έκδοση της minimax.

## Η μέθοδος αναζήτησης α–β

- Στο βήμα 3 της μεθόδου minimax με πρώτα–κατά–βάθος αναζήτηση και μερική ανάπτυξη υπάρχει ενδεχόμενο ο κόμβος p και όλοι οι απόγονοί του να είναι δυνατόν με ασφάλεια να διαγραφούν από την L. Αυτό μπορεί να γίνει στις εξής περιπτώσεις
  - Αν ο p είναι min κόμβος και η τιμή $v_x$ του τρέχοντος κόμβου x είναι μικρότερη ή ίση από τη μέγιστη τιμή α όλων των max προγόνων του p (α–cut).
  - Αν ο p είναι max κόμβος και η τιμή $v_x$ του τρέχοντος κόμβου x είναι μεγαλύτερη ή ίση από την ελάχιστη τιμή β όλων των min προγόνων του p (β–cut).

- Και στις δύο περιπτώσεις (α–cut και β–cut) δεν εξερευνώνται τα υποδέντρα με ρίζες τους απομένοντες αδελφούς κόμβους του x.

- Ο κλάδος του χώρου αναζήτησης που γίνεται η αποκοπή (pruning) δεν μπορεί να είναι ο πρώτος κλάδος ενός κόμβου p και ο κόμβος p, ή κάποιος πρόγονός του άρτιο πλήθος επιπέδων πιο επάνω, δεν είναι ο πρώτος εξεταζόμενος κόμβος μεταξύ των αδελφών του.

**α–β αναζήτηση** : Υπολογισμός της τιμής ενός κόμβου n

Βήμα 1 : Θέσε L = {n}.

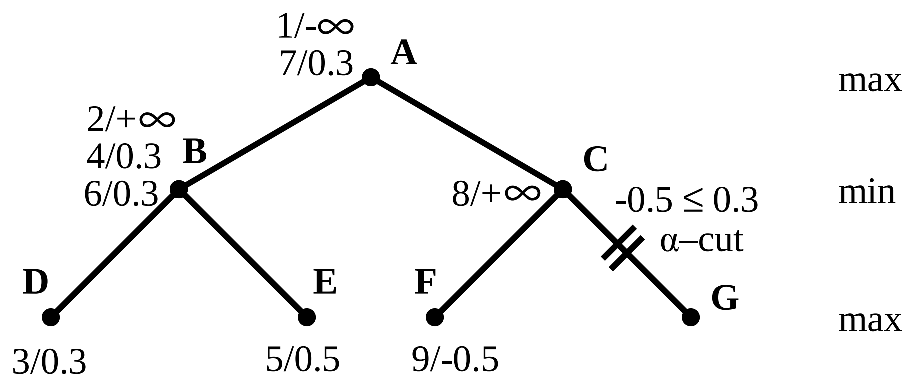
Βήμα 2 : Έστω x ο πρώτος κόμβος της L. Αν x = n και υπάρχει τιμή γι' αυτόν, να επιστρέψεις αυτήν την τιμή.

Βήμα 3 : Αν ο x έχει τιμή $v_x$, έστω p ο πατέρας του x. Αν ο p είναι min κόμβος, θέσε στο α τη μέγιστη τιμή όλων των max προγόνων του p (αν p = n, θέσε $\alpha = -\infty$) και έλεγξε αν $v_x \leq \alpha$. Αν ο p είναι max κόμβος, θέσει στο β την ελάχιστη τιμή όλων των min προγόνων του p (αν p = n, θέσε $\beta = +\infty$) και έλεγξε αν $v_x \geq \beta$. Αν ο έλεγχος ($v_x \leq \alpha$ ή $v_x \geq \beta$) είναι επιτυχής, αφαίρεσε τον κόμβο p καθώς και όλους τους απογόνους του από την L (α–cut ή β–cut αντίστοιχα) και πήγαινε στο βήμα 2. Αν ο έλεγχος δεν είναι επιτυχής, έστω $v_p$ η τιμή του κόμβου p. Αν ο p είναι min κόμβος, θέσε $v_p = \min(v_p, v_x)$. Αν ο p είναι max κόμβος, θέσε $v_p = \max(v_p, v_x)$. Αφαίρεσε τον x από την L και πήγαινε στο βήμα 2.

προσθήκη σε σχέση με την minimax

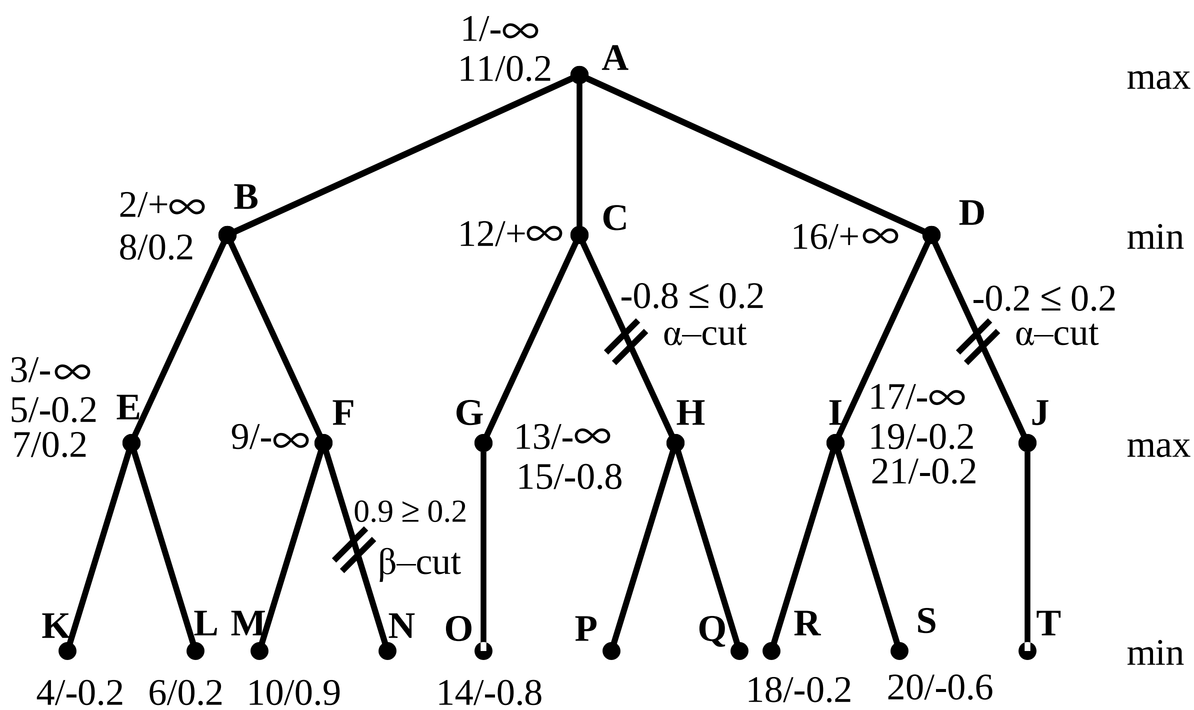
Βήμα 4 : Αν ο x δεν έχει τιμή και είτε είναι τερματικός κόμβος είτε έχουμε αποφασίσει να μην κάνουμε περαιτέρω ανάπτυξη του χώρου αναζήτησης, δώσε του τιμή σύμφωνα με την ευριστική συνάρτηση που έχει επιλεγεί. Πήγαινε στο βήμα 2.

Βήμα 5 : Αλλιώς, δώσε του τιμή $-\infty$ αν είναι max κόμβος ή τιμή $+\infty$ αν είναι min κόμβος. Βάλε τα παιδιά του x στην αρχή της L. Πήγαινε στο βήμα 2.




9 αναθέσεις τιμών (αντί για 13 στη minimax)

3 ευριστικοί υπολογισμοί (αντί για 4 στη minimax)



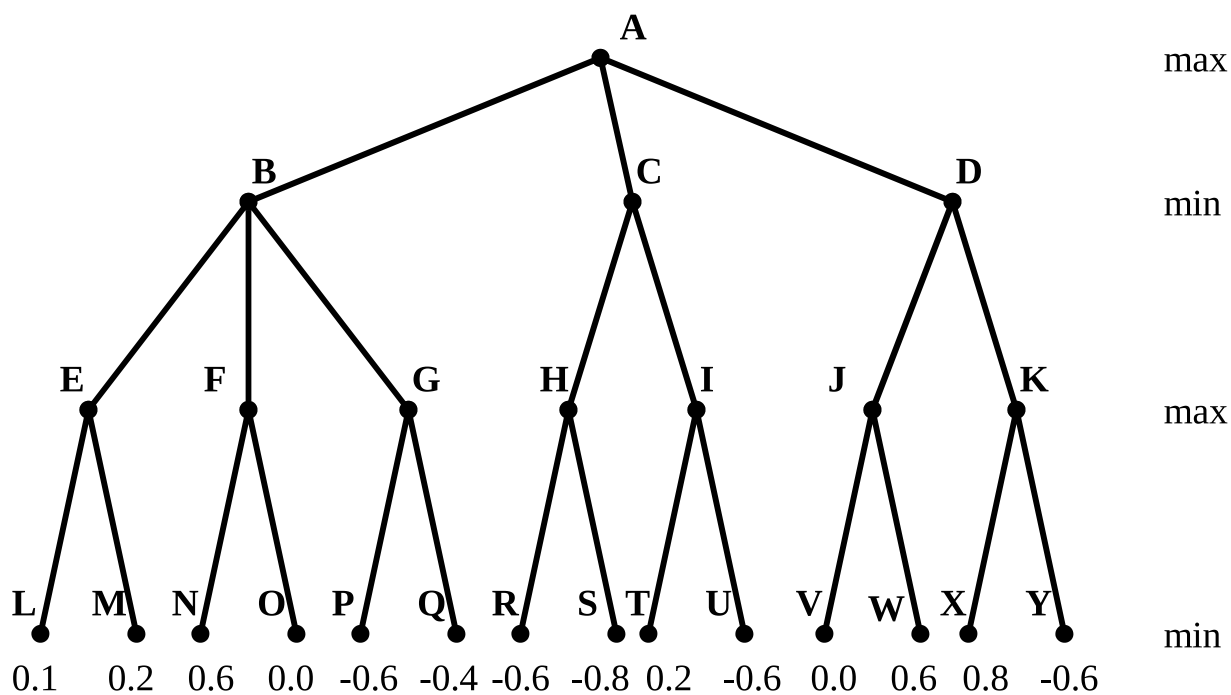

21 αναθέσεις τιμών (αντί για 39 στη minimax)

6 ευριστικοί υπολογισμοί (αντί για 10 στη minimax)

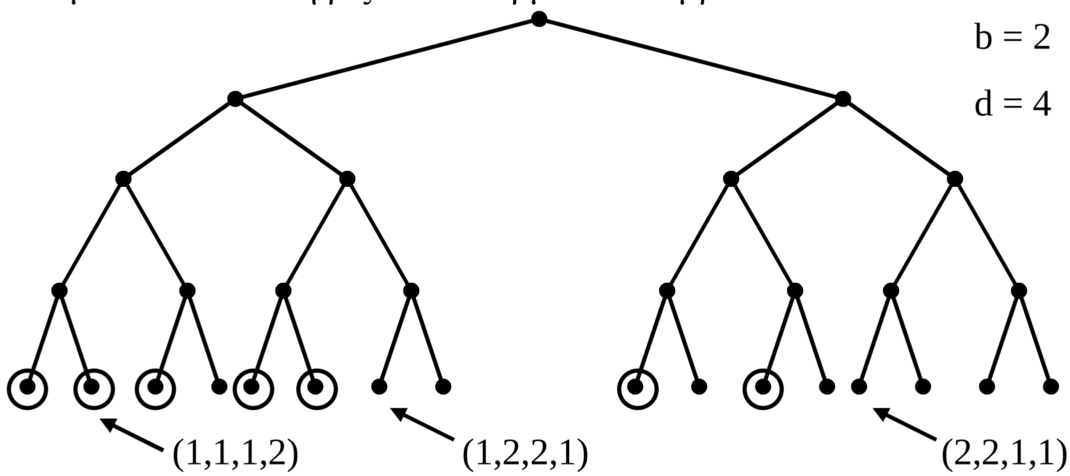
Πότε υπάρχει σημαντικό όφελος στην α–β αναζήτηση σε σχέση με τη μέθοδο minimax;

Τι όφελος υπάρχει στον παρακάτω χώρο αναζήτησης;

Θα μπορούσε να ήταν μεγαλύτερο;



- Σε ένα δέντρο αναζήτησης βάθους d και ομοιόμορφου παράγοντα διακλάδωσης b, κάθε τερματικός κόμβος αντιστοιχεί σε μία ακολουθία της μορφής $(e_1, e_2, \ldots\ldots, e_d)$, όπου για κάθε $e_i$ $(1 \leq i \leq d)$ ισχύει $1 \leq e_i \leq b$. Στην πραγματικότητα, η ακολουθία αυτή δείχνει το μονοπάτι από τη ρίζα στον τερματικό κόμβο.




- Στη χειρότερη περίπτωση, στην α–β αναζήτηση θα πρέπει να εξετασθούν τόσοι τερματικού κόμβοι όσοι και στη μέθοδο minimax, δηλαδή $b^d$.
- Στην καλύτερη περίπτωση, μπορεί να αποφευχθεί η εξέταση εκείνων των τερματικών κόμβων που στην αντίστοιχη ακολουθία τους βρίσκεται κάποιο i τέτοιο ώστε να ισχύουν $e_i \neq 1$ και $e_{i+2n+1} \neq 1$ (γιατί;).
  Οι τερματικοί κόμβοι που πρέπει οπωσδήποτε να εξετασθούν είναι αυτοί που έχουν 1 σε όλες τις περιττές ή σε όλες τις άρτιες θέσεις της αντίστοιχης ακολουθίας τους. Οι κόμβοι αυτοί είναι, για d άρτιο, $2 \cdot b^{d/2} - 1$ (γιατί;) $\approx 2 \cdot b^{d/2}$.
- Στην πιο αισιόδοξη κατάσταση, μία α–β αναζήτηση με παράγοντα διακλάδωσης b ισοδυναμεί με μία minimax αναζήτηση με παράγοντα διακλάδωσης b΄, όπου
  $b'^d = 2 \cdot b^{d/2} \Rightarrow b' \approx \sqrt{b}$
  b΄ = ενεργός παράγοντας διακλάδωσης (effective branching factor)

## ΛΟΓΙΚΗ
## (LOGIC)

- Αναλογίες διαδικασίας επίλυσης προβλημάτων υπολογισμού και προβλημάτων νοημοσύνης
  - Πρόβλημα υπολογισμού
    1. Επινόηση του αλγορίθμου
    2. Επιλογή γλώσσας προγραμματισμού
    3. Κωδικοποίηση του αλγορίθμου σε πρόγραμμα
    4. Εκτέλεση του προγράμματος
  - Πρόβλημα νοημοσύνης
    1. Προσδιορισμός απαιτούμενης γνώσης
    2. Επιλογή γλώσσας αναπαράστασης γνώσης
    3. Κωδικοποίηση της γνώσης του προβλήματος
    4. Δημιουργία νέας γνώσης

- Η λογική είναι μία γλώσσα αναπαράστασης γνώσης (knowledge representation).

- Ύπαρξη διαφόρων “λογικών”
  - Προτασιακή λογική (Propositional logic)
  - Κατηγορηματική λογική (Predicate logic)
  - Λογική πρώτης τάξης (First order logic)

- Σύνταξη (syntax) και σημασία (semantics) στη λογική.

- Το βήμα 4 της διαδικασίας επίλυσης προβλημάτων νοημοσύνης είναι θέμα αναζήτησης.

**Προτασιακή λογική**

Π1: Αν κάνει ζέστη και έχει υγρασία, τότε θα βρέξει

Π2: Αν έχει υγρασία, τότε κάνει ζέστη

Π3: Έχει υγρασία

Π4: Θα βρέξει

P: Κάνει ζέστη

Q: Έχει υγρασία

R: Θα βρέξει

$$\left.\begin{array}{ll} \Pi 1: & P \wedge Q \rightarrow R \\ \Pi 2: & Q \rightarrow P \\ \Pi 3: & Q \\ \Pi 4: & R \end{array}\right\}$$

Πώς είναι δυνατόν η Π4 να παραχθεί αυτόματα από τις Π1, Π2 και Π3;

**Λογική πρώτης τάξης**

Π1: Ο Σωκράτης είναι άνθρωπος

Π2: Κάθε άνθρωπος είναι θνητός

Π3: Ο Σωκράτης είναι θνητός

P(x): Ο x είναι άνθρωπος

Q(x): Ο x είναι θνητός

$$\left.\begin{array}{ll} \Pi 1: & P(\text{Σωκράτης}) \\ \Pi 2: & (\forall x)(P(x) \rightarrow Q(x)) \\ \Pi 3: & Q(\text{Σωκράτης}) \end{array}\right\}$$

Πώς είναι δυνατόν η Π3 να παραχθεί αυτόματα από τις Π1 και Π2;

- Στην κλασική λογική υπάρχουν δύο τιμές αλήθειας (truth values), το αληθές (true, T) και το ψευδές (false, F).

- Στην προτασιακή λογική χρησιμοποιούνται σύμβολα (π.χ. P, Q, ....) που ονομάζονται άτομα (atoms) και τα οποία παριστάνουν προτάσεις (propositions) που μπορεί να είναι αληθείς ή ψευδείς σε δεδομένη κατάσταση του περιβάλλοντος κόσμου.

- Οι καλοσχηματισμένοι τύποι (well–formed formulas) στην προτασιακή λογική είναι είτε απλά άτομα είτε σύνθετοι τύποι που προκύπτουν από το συνδυασμό απλών ατόμων μέσω των λογικών συνδέσμων (logical connections) $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$.

- Η τιμή αλήθειας ενός καλοσχηματισμένου τύπου προκύπτει από τις επιμέρους τιμές αλήθειας των ατόμων που συμμετέχουν σ' αυτόν σύμφωνα με τον παρακάτω πίνακα αλήθειας (truth table).

| P | Q | $\neg$P | P$\wedge$Q | P$\vee$Q | P$\rightarrow$Q | P$\leftrightarrow$Q |
|---|---|---|---|---|---|---|
| T | T | F | T | T | T | T |
| T | F | F | F | T | F | F |
| F | T | T | F | T | T | F |
| F | F | T | F | F | T | T |

- Η κατηγορηματική λογική είναι μία συντακτική επέκταση της προτασιακής λογικής προς την κατεύθυνση της λογικής πρώτης τάξης, χωρίς να προσφέρει κάτι περισσότερο από την προτασιακή λογική στη δυνατότητα παραγωγής νέας γνώσης από υπάρχουσα (τουλάχιστον αυτή είναι η θεώρηση της κατηγορηματικής λογικής από τον Ginsberg που, αν και σε καμία περίπτωση δεν είναι καθολικής αποδοχής, θα τη σεβαστούμε γιατί εξυπηρετεί εκπαιδευτικούς σκοπούς).
- Για την αναπαράσταση γνώσης στην κατηγορηματική λογική χρησιμοποιούνται:
  - Σύμβολα που ονομάζονται σταθερές (constants) και τα οποία παριστάνουν αντικείμενα (objects) του περιβάλλοντος κόσμου. Π.χ. <u>Socrates</u>, <u>20</u>, <u>November</u>, <u>1995</u>, <u>this_slide</u>.
  - Τα λεγόμενα συναρτησιακά σύμβολα (function symbols) τα οποία χρησιμοποιούνται για την κατασκευή δομών που επίσης παριστάνουν αντικείμενα του περιβάλλοντος κόσμου. Π.χ. <u>father_of(Socrates)</u>, <u>date(20, November, 1995)</u>, <u>average(20, 30)</u>, <u>next_to(this_slide)</u>, <u>next_to(next_to(this_slide))</u>.

    Τόσο οι δομές αυτές, όσο και οι σταθερές, ονομάζονται όροι (terms). Κάθε συναρτησιακό σύμβολο έχει ένα βαθμό (arity) που είναι ο αριθμός των όρων επάνω στους οποίους εφαρμόζεται. π.χ. οι βαθμοί των father_of, date, average, next_to είναι 1, 3, 2, 1 αντίστοιχα.

– Σύμβολα που ονομάζονται κατηγορήματα (predicates) τα οποία χρησιμοποιούνται για την κατασκευή ατόμων που, όπως και στην προτασιακή λογική, παριστάνουν προτάσεις που μπορεί να είναι αληθείς ή ψευδείς σε δεδομένη κατάσταση του περιβάλλοντος κόσμου.
  Π.χ. man(Socrates), man(father_of(Socrates)), born(John, date(12, April, 1980)), empty(this_slide) .
  Τα κατηγορήματα, όπως και τα συναρτησιακά σύμβολα, έχουν ένα βαθμό που είναι ο αριθμός των όρων επάνω στους οποίους εφαρμόζονται. π.χ. οι βαθμοί των man, born, empty είναι 1, 2, 1 αντίστοιχα.
– Καλοσχηματισμένοι τύποι που προκύπτουν από το συνδυασμό ατόμων μέσω των γνωστών λογικών συνδέσμων, όπως και στην προτασιακή λογική.
  Π.χ. man(Socrates) $\rightarrow$ mortal(Socrates)

- Η λογική πρώτης τάξης αποτελεί επέκταση της (κατά Ginsberg) κατηγορηματικής λογικής ως προς τα εξής:
  – Χρησιμοποιείται και μία άλλη κατηγορία συμβόλων, οι μεταβλητές (variables), που παριστάνουν τυχαία αντικείμενα του περιβάλλοντος κόσμου. Π.χ. x, y, z. Οι μεταβλητές είναι επίσης όροι.
  – Σε ένα καλοσχηματισμένο τύπο (συνεπώς και σε ένα απλό άτομο) που περιέχει μεταβλητές δεν είναι δυνατόν να ανατεθεί κάποια τιμή αλήθειας, εκτός αν οι μεταβλητές αυτές ποσοτικοποιηθούν με κάποιον από τους ποσοδείκτες $\forall$ (καθολικός) ή $\exists$ (υπαρξιακός). Π.χ.
  ($\exists$x)(loves(John, x)), ($\forall$x)(man(x) $\rightarrow$ mortal(x))

– Αν ένας καλοσχηματισμένος τύπος F περιέχει μία μη ποσοτικοποιημένη μεταβλητή x, τότε τα $(\forall x)(F)$ και $(\exists x)(F)$ είναι επίσης καλοσχηματισμένοι τύποι.

– Ο τύπος $(\forall x)(F)$ είναι αληθής αν για κάθε δυνατή τιμή του x ο τύπος F είναι αληθής, αλλιώς ο $(\forall x)(F)$ είναι ψευδής. Ο τύπος $(\exists x)(F)$ είναι αληθής αν υπάρχει κάποια τιμή του x τέτοια ώστε ο τύπος F να είναι αληθής, αλλιώς ο $(\exists x)(F)$ είναι ψευδής.

- Αν, για κάποιον περιβάλλοντα κόσμο, A είναι το σύνολο όλων των πιθανών ατόμων (για τη λογική πρώτης τάξης, των ατόμων χωρίς μεταβλητές), ένα υποσύνολο I του A ($I \subseteq A$) ονομάζεται ερμηνεία (interpretation). Μία ερμηνεία αντιπροσωπεύει μία πιθανή κατάσταση του περιβάλλοντος κόσμου με την έννοια ότι περιέχει ακριβώς εκείνα τα άτομα που είναι αληθή στην κατάσταση αυτή. Μία τέτοια ερμηνεία ονομάζεται μοντέλο (model) του περιβάλλοντος κόσμου. Ένας καλοσχηματισμένος τύπος ισχύει (holds) σε μία ερμηνεία αν είναι αληθής για τις δεδομένες τιμές αλήθειας, σε σχέση με τη συγκεκριμένη ερμηνεία, των ατόμων που τον συνθέτουν. Δύο καλοσχηματισμένοι τύποι λέγονται ισοδύναμοι (equivalent) όταν ισχύουν ακριβώς στις ίδιες ερμηνείες.

P: Κάνει ζέστη

Q: Έχει υγρασία

R: Θα βρέξει

| P | Q | R | $P\wedge Q$ | $P\wedge Q\rightarrow R$ | $Q\rightarrow P$ | $(P\wedge Q\rightarrow R)\wedge(Q\rightarrow P)\wedge Q$ |
|---|---|---|---|---|---|---|
| T | T | T | T | T | T | T |
| T | T | F | T | F | T | F |
| T | F | T | F | T | T | F |
| T | F | F | F | T | T | F |
| F | T | T | F | T | F | F |
| F | T | F | F | T | F | F |
| F | F | T | F | T | T | F |
| F | F | F | F | T | T | F |

- Κάθε γραμμή σε ένα πίνακα αλήθειας αντιπροσωπεύει μία πιθανή ερμηνεία.

- Σε περιβάλλοντα κόσμο που μπορεί να μοντελοποιηθεί στην προτασιακή λογική με N προτάσεις, οι πιθανές ερμηνείες είναι $2^N$.

- Αν ένας καλοσχηματισμένος τύπος ισχύει για όλες τις πιθανές ερμηνείες, τότε λέγεται ταυτολογία (tautology) ή έγκυρος (valid) τύπος. Π.χ. $((P\rightarrow Q)\wedge P)\rightarrow Q$

- Αν ένας καλοσχηματισμένος τύπος δεν ισχύει για καμία από τις πιθανές ερμηνείες, τότε λέγεται αντίφαση (contradiction) ή ασυνεπής (inconsistent) τύπος. Π.χ. $(P\rightarrow Q)\wedge(P\wedge\neg Q)$

- Για οποιουσδήποτε καλοσχηματισμένους τύπους P και Q, οι τύποι $P \rightarrow Q$ και $\neg P \vee Q$ είναι ισοδύναμοι.

- Για οποιουσδήποτε καλοσχηματισμένους τύπους P και Q, οι τύποι $P \rightarrow Q$ και $\neg Q \rightarrow \neg P$ είναι ισοδύναμοι.

- Αν για δύο καλοσχηματισμένους τύπους P και Q αυτό που συμβαίνει είναι ο Q να ισχύει σε κάθε ερμηνεία που ισχύει και ο P, τότε λέγεται ότι ο Q είναι λογική συνέπεια (logical consequence) του P ή ότι ο P συμπεραίνει (entails) τον Q και γράφεται $P \models Q$.
  Για δύο ισοδύναμους τύπους ισχύει ότι ο ένας συμπεραίνει τον άλλον. Αν ένας τύπος P συμπεραίνει τον Q ($P \models Q$), τότε ο τύπος $P \rightarrow Q$ είναι έγκυρος και αντίστροφα.

- Η σχέση του “συμπεραίνειν” (entailment) $\models$ είναι πολύ βασική για την τεχνητή νοημοσύνη. Συνήθως αυτό που συμβαίνει είναι ότι υπάρχει δεδομένο ένα σύνολο γνώσης που αναπαρίσταται από τους καλοσχηματισμένους τύπους της λογικής πρώτης τάξης $P_1$, $P_2$, ......., $P_n$ και το ζητούμενο είναι αν από αυτή τη γνώση μπορεί κάποιος να συμπεράνει κάτι που αναπαρίσταται από τον καλοσχηματισμένο τύπο Q. Δηλαδή πρέπει να διερευνηθεί αν ισχύει $P_1 \wedge P_2 \wedge \ldots\ldots \wedge P_n \models Q$. Αυτό είναι το βήμα 4 της διαδικασίας επίλυσης προβλημάτων νοημοσύνης που είναι ουσιαστικά ένα πρόβλημα αναζήτησης.

- Το να αποδειχθεί αν ισχύει ή όχι το $P \models Q$ με τη βοήθεια του ορισμού του είναι εφικτό μόνο στην προτασιακή λογική, επειδή εκεί υπάρχει πεπερασμένο πλήθος από ερμηνείες.

- Στη λογική πρώτης τάξης εισάγεται η έννοια των κανόνων συμπερασμού (inference rules).

  Π.χ. modus ponens

$$\frac{\begin{array}{c} x \rightarrow y \\ x \end{array}}{y}$$

  Δηλαδή, αν το $x \rightarrow y$ και το x βρίσκονται μέσα στη βάση γνώσης P, πρόσθεσε και το y στην P.

  Συνέχισε αυτή τη διαδικασία μέχρι να βάλεις το Q μέσα στη βάση. Αυτή είναι μία διαδικασία συμπερασμού (inference procedure) του Q από το P και γράφεται σαν $P \vdash Q$.

- Μία διαδικασία συμπερασμού $\vdash$ λέγεται ορθή (sound) όταν οποτεδήποτε συμβαίνει το $P \vdash Q$, ισχύει επίσης και το $P \models Q$.
- Μία διαδικασία συμπερασμού $\vdash$ λέγεται πλήρης (complete) όταν οποτεδήποτε ισχύει το $P \models Q$, συμβαίνει επίσης και το $P \vdash Q$.
- Οι ιδανικές διαδικασίες συμπερασμού είναι εκείνες που είναι και ορθές αλλά και πλήρεις, αλλά τι κόστος έχουν;
- Το πρόβλημα του “συμπερασμού” ($P \models Q$) στη λογική πρώτης τάξης είναι ημι–αποφασίσιμο (semidecidable).

## ΚΑΤΗΓΟΡΗΜΑΤΙΚΗ ΛΟΓΙΚΗ
## (PREDICATE LOGIC)

Υπενθυμίζεται ο (κατά Ginsberg) ορισμός της κατηγορηματικής λογικής: = λογική πρώτης τάξης χωρίς μεταβλητές και, συνεπώς, χωρίς ποσοδείκτες (δηλαδή εκφραστικά ισοδύναμη με την προτασιακή λογική).

Μία βάση γνώσης

Ο Γιάννης είναι δικηγόρος.

Οι δικηγόροι είναι πλούσιοι.

Οι πλούσιοι άνθρωποι έχουν μεγάλα σπίτια.

Τα μεγάλα σπίτια χρειάζονται πολλή δουλειά για να συντηρηθούν από τους ιδιοκτήτες τους.

Κωδικοποίησή τους σε λογική πρώτης τάξης

| | | |
|---|---|---|
| Σταθερές: | John | (Γιάννης) |
| Συναρτησιακά σύμβολα: | house-of(p) | (το σπίτι του p) |
| Κατηγορήματα: | lawyer(p) | (ο p είναι δικηγόρος) |
| | rich(p) | (ο p είναι πλούσιος) |
| | house(h, p) | (το h είναι σπίτι του p) |
| | big(h) | (το h είναι μεγάλο) |
| | work(h) | (το h χρειάζεται πολλή δουλειά για να συντηρηθεί από τον ιδιοκτήτη του) |

- Καλοσχηματισμένοι τύποι που αναπαριστούν τη γνώση μας

lawyer(John)

$(\forall p)(lawyer(p) \rightarrow rich(p))$

$(\forall p)(house(house\text{-}of(p), p))$ (τι χρειάζεται;)

$(\forall p)(\forall h)(house(h, p) \wedge rich(p) \rightarrow big(h))$

$(\forall h)(big(h) \wedge (\exists p)(house(h, p)) \rightarrow work(h))$

**ΕΡΩΤΗΣΗ:** Πώς είναι δυνατόν από την γνώση αυτή να βγει το συμπέρασμα ότι "το σπίτι του Γιάννη χρειάζεται πολλή δουλειά για να συντηρηθεί απ' αυτόν", ή αλλιώς, work(house-of(John));

**ΑΠΑΝΤΗΣΗ:** Υπάρχει τρόπος να γίνει αυτό στη λογική πρώτης τάξης, αλλά αρχικά θα θεωρήσουμε ένα στιγμιότυπο της βάσης γνώσης μας για να δουλέψουμε στην κατηγορηματική λογική. Το στιγμιότυπο αυτό συνίσταται στην εξειδίκευση των προηγούμενων καλοσχηματισμένων τύπων για p = John και h = house-of(John). Λόγω της ύπαρξης του συναρτησιακού συμβόλου house-of οι πιθανές εξειδικεύσεις είναι άπειρες στο πλήθος.

- Καλοσχηματισμένοι τύποι που αναπαριστούν τη γνώση μας στην κατηγορηματική λογική (εξειδίκευση αυτών στη λογική πρώτης τάξης).

lawyer(John)

lawyer(John) $\rightarrow$ rich(John)

house(house-of(John), John)

house(house-of(John), John) $\wedge$ rich(John) $\rightarrow$ big(house-of(John))

big(house-of(John)) $\wedge$ house(house-of(John), John) $\rightarrow$ work(house-of(John))

Να γιατί δεν έχουμε μεγαλύτερη εκφραστική δύναμη από την προτασιακή λογική:

L $\equiv$ lawyer(John)

R $\equiv$ rich(John)

H $\equiv$ house(house-of(John), John)

B $\equiv$ big(house-of(John))

W $\equiv$ house(house-of(John))

$$\left.\begin{array}{l} L \\ L \rightarrow R \\ H \\ H \wedge R \rightarrow B \\ B \wedge H \rightarrow W \end{array}\right\} \text{Πως μπορεί να αποδειχθεί το W;}$$

- Από τα L και L$\rightarrow$R μπορεί να αποδειχθεί το R μέσω του κανόνα “modus ponens”.

$$\begin{array}{c} a \rightarrow b \\ a \\ \hline b \end{array}$$

- Από τα H, R και H $\wedge$ R $\rightarrow$B μπορεί να αποδειχθεί το B μέσω της εξής επέκτασης του κανόνα “modus ponens”.

$$\begin{array}{c} a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow b \\ a_1 \\ \vdots \\ a_m \\ \hline b \end{array}$$

- Στην πραγματικότητα αρκεί απλώς η εξής επέκταση του κανόνα “modus ponens”.

$$\begin{array}{c} a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow b \\ a_i \\ \hline \underbrace{a_1 \wedge \ldots\ldots\ldots \wedge \hat{a}_i \wedge \ldots\ldots\ldots \wedge a_m}_{\text{σύζευξη όλων των } a_j \ (1 \leq j \leq m) \text{ πλην του } a_i} \rightarrow b \end{array}$$

Η προηγούμενη έκδοση του “modus ponens” για m=1 γίνεται

$$\begin{array}{c} a_1 \rightarrow b \\ a_1 \\ \hline T \rightarrow b \quad = \ b \end{array}$$

- Ένα βήμα πριν την τελική μας έκδοση του κανόνα "modus ponens" είναι να θεωρήσουμε ένα $d = a_i$: και να λάβουμε υπόψη μας ότι $T \rightarrow d = d$, οπότε:

  Αν $d = a_i$:

$$\frac{\begin{array}{r} a_1 \wedge \ldots\ldots \wedge a_m \rightarrow b \\ T \rightarrow d \end{array}}{a_1 \wedge \ldots\ldots \wedge \hat{a}_i \wedge \ldots\ldots \wedge a_m \rightarrow b}$$

- Τέλος, αν το d στηρίζεται σε κάποιες υποθέσεις έχουμε την τελική μορφή του "modus ponens"

  modus ponens:

  Αν $d = a_i$:

$$\frac{\begin{array}{r} a_1 \wedge \ldots\ldots \wedge a_m \rightarrow b \\ c_1 \wedge \ldots\ldots \wedge c_n \rightarrow d \end{array}}{a_1 \wedge \ldots\ldots \wedge \hat{a}_i \wedge \ldots\ldots \wedge a_m \wedge c_1 \wedge \ldots\ldots \wedge c_n \rightarrow b}$$

- Εύκολα αποδεικνύεται ότι:

  "Ο κανόνας συμπερασμού modus ponens είναι ορθός (sound)"

με την έννοια ότι αν η βάση γνώσης μας (το σύνολο των καλοσχηματισμένων τύπων που την αναπαριστούν) είναι P και με διαδοχικές εφαρμογές του "modus ponens" αποδειχθεί το Q, τότε $P \models Q$ (με βάση τον ορισμό του "συμπεραίνειν" μέσω ερμηνειών)

- Εφαρμογή του “modus ponens” για να αποδειχθεί το W από τα:

$$\begin{array}{ll} (1) & L \\ (2) & L \rightarrow R \\ (3) & H \\ (4) & H \wedge R \rightarrow B \\ (5) & B \wedge H \rightarrow W \end{array}$$

---

$$\begin{array}{lll} & (4) & H \wedge R \rightarrow B \\ (3)\ \ H \dashrightarrow & (3)' & T \rightarrow H \\ \hline & (6) & R \rightarrow B \end{array}$$

---

$$\begin{array}{ll} (5) & B \wedge H \rightarrow W \\ (6) & R \rightarrow B \\ \hline (7) & R \wedge H \rightarrow W \end{array}$$

---

$$\begin{array}{lll} & (7) & R \wedge H \rightarrow W \\ (3)\ \ H \dashrightarrow & (3)' & T \rightarrow H \\ \hline & (8) & R \rightarrow W \end{array}$$

---

$$\begin{array}{lll} & (2) & L \rightarrow R \\ (1)\ \ L \dashrightarrow & (1)' & T \rightarrow L \\ \hline & (9) & R \end{array}$$

---

$$\begin{array}{lll} & (8) & R \rightarrow W \\ (9)\ \ R \dashrightarrow & (9)' & T \rightarrow R \\ \hline & & W \end{array}$$

---

- Μία βάση γνώσης που αποτελείται από καλοσχηματισμένους τύπους της μορφής:

$$a_1 \wedge \ldots\ldots\ \wedge a_m \rightarrow b$$

ή αλλιώς

$$\neg a_1 \vee \ldots\ldots \vee \neg a_m \vee b$$

λέγεται βάση Horn (Horn database) όταν τα $a_i$ και b είναι άτομα.

Π.χ. lawyer(John) $\rightarrow$ rich(John) =

$\neg$ lawyer(John) $\vee$ rich(John)

Αλλά το

lawyer(John) $\vee$ waiter(John) =

T $\rightarrow$ (lawyer(John) $\vee$ waiter(John)) =

$\neg$lawyer(John) $\rightarrow$ waiter(John)

δεν μπορεί να συμμετάσχει σε μία βάση Horn.

- Αποδεικνύεται ότι:

“Ο κανόνας συμπερασμού modus ponens είναι πλήρης (complete) για βάσεις Horn”

με την έννοια ότι αν η βάση γνώσης μας είναι P και για κάποιο Q ισχύει $P \models Q$ (με βάση τον ορισμό του “συμπεραίνειν” μέσω ερμηνειών), τότε το Q μπορεί να αποδειχθεί με διαδοχικές εφαρμογές του “modus ponens”.

- Μία διαδικασία διαδοχικής εφαρμογής του “modus ponens”, που διατηρεί την ορθότητα και πληρότητά του, είναι η εξής:
    1. Για κάθε τύπο $T \rightarrow c$ στη βάση:
    2. Αν υπάρχει ένας άλλος τύπος $p_1 \wedge ....... \wedge p_n \rightarrow c'$ στη βάση τέτοιος ώστε $p_1 = c$, πρόσθεσε στη βάση και τον τύπο $p_2 \wedge ....... \wedge p_n \rightarrow c'$ αν δεν υπάρχει ήδη.

- Η προηγούμενη διαδικασία όταν εφαρμοσθεί σε μία βάση Horn από την προτασιακή λογική ή την κατηγορηματική (κατά Ginsberg) λογική δεν χρειάζεται χρόνο περισσότερο από $n \cdot d$ για να αποδειχθεί ότι ένα συγκεκριμένο άτομο προκύπτει από τη βάση, όπου n είναι ο μέγιστος αριθμός υποθέσεων των τύπων της βάσης και d είναι το μέγεθος της βάσης (πλήθος τύπων).

- **ΕΡΩΤΗΣΗ:**

  Τι εμποδίζει τον “modus ponens” να είναι πλήρης και για βάσεις που δεν είναι Horn;

  **ΑΠΑΝΤΗΣΗ:**

  Η πιθανή ύπαρξη διάζευξης στο συμπέρασμα ενός τύπου. Για παράδειγμα, από τα $a \vee b$ , $a \rightarrow q$ , $b \rightarrow q$ λογικά προκύπτει το q, αλλά δεν μπορεί να αποδειχθεί με τον “modus ponens”. Συνεπώς, τι κάνουμε;

  **ΑΠΑΝΤΗΣΗ Β΄:**

  Χρησιμοποιούμε τον κανόνα της ανάλυσης (resolution rule).

- Μία γενική κατηγορία τύπων από αυτούς που δεν μπορούν να συμμετάσχουν σε μία βάση Horn είναι η

$$a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow b_1 \vee \ldots\ldots\ldots \vee b_k$$

ή αλλιώς

$$\neg a_1 \vee \ldots\ldots\ldots \vee \neg a_m \vee b_1 \vee \ldots\ldots\ldots \vee b_k$$

όπου τα $a_i$ και $b_j$ είναι άτομα. Μία βάση γνώσης που αποτελείται από τύπους αυτής της μορφής λέγεται ότι είναι σε κανονική μορφή (normal form).

– Για $m = 0$ έχουμε:

$$T \rightarrow b_1 \vee \ldots\ldots\ldots \vee b_k =$$
$$b_1 \vee \ldots\ldots\ldots \vee b_k$$

– Για $k = 1$ έχουμε τη μορφή Horn

– Για $k = 0$ έχουμε:

$$a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow F =$$
$$\neg( a_1 \wedge \ldots\ldots\ldots \wedge a_m ) =$$
$$\neg a_1 \vee \ldots\ldots\ldots \vee \neg a_m$$

– Για $m = 0$ και $k = 0$ έχουμε:

$$T \rightarrow F =$$
$$F$$

που είναι ο αντιφατικός τύπος

κανόνας της ανάλυσης:

Αν $d_j = a_i$

$$a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow b_1 \vee \ldots\ldots\ldots \vee b_k$$
$$c_1 \wedge \ldots\ldots\ldots \wedge c_n \rightarrow d_1 \vee \ldots\ldots\ldots \vee d_q$$



$$a_1 \wedge \ldots \wedge \hat{a}_i \wedge \ldots \wedge a_m \wedge c_1 \wedge \ldots \wedge c_n \rightarrow b_1 \vee \ldots \vee b_k \vee d_1 \vee \ldots \vee \hat{d}_j \vee \ldots \vee d_q$$

- Ο κανόνας της ανάλυσης είναι μία γενίκευση του κανόνα “modus ponens”.

- Αποδεικνύεται ότι:

  “Ο κανόνας της ανάλυσης είναι ορθός”

- Εφαρμογή του κανόνα της ανάλυσης για να αποδειχθεί το q από τα:

(1) $a \lor b$

(2) $a \to q$

(3) $b \to q$

---

$$(1)\quad a \lor b \dashrightarrow \quad \begin{array}{ll} (2) & a \to q \\ (1)' & T \to a \lor b \\ \hline (4) & b \lor q \end{array}$$

---

$$(4)\quad b \lor q \dashrightarrow \quad \begin{array}{ll} (3) & b \to q \\ (4)' & T \to b \lor q \\ \hline (5) & q \lor q \;=\; q \end{array}$$

---

- Εκτός από τον προφανή εμπρόσθιο (forward) τρόπο απόδειξης ενός q από μία βάση γνώσης D μέσω του κανόνα της ανάλυσης, ο τελευταίος μπορεί να υποστηρίξει και την προς τα πίσω (backward) απόδειξη του q ως εξής:

  Προστίθεται στη βάση η άρνηση $\neg q$ του αποδεικτέου και με διαδοχικές εφαρμογές του κανόνα της ανάλυσης γίνεται προσπάθεια να αποδειχθεί η αντίφαση $T \to F$.

Εφαρμογή του κανόνα της ανάλυσης με οπίσθιο τρόπο

για να αποδειχθεί το W από τα:

(1) $L$

(2) $L \rightarrow R$

(3) $H$

(4) $H \wedge R \rightarrow B$

(5) $B \wedge H \rightarrow W$

| | Αποδεικνυόμενοι τύποι | Αιτιολόγηση |
|---|---|---|
| a | $W \rightarrow F$ | άρνηση αποδεικτέου |
| b | $B \wedge H \rightarrow F$ | από (a), (5) |
| c | $B \rightarrow F$ | από (b), (3) |
| d | $H \wedge R \rightarrow F$ | από (c), (4) |
| e | $R \rightarrow F$ | από (d), (3) |
| f | $L \rightarrow F$ | από (e), (2) |
| g | $T \rightarrow F$ | από (f), (1) |

- Αποδεικνύεται ότι:

  "Ο κανόνας της ανάλυσης είναι πλήρης "

- **ΕΡΩΤΗΣΗ:**

  Δεν χρειάζεται η βάση να είναι σε κανονική μορφή

  για να είναι ο κανόνας της ανάλυσης πλήρης;

**ΑΠΑΝΤΗΣΗ:**

Είναι δυνατόν να αποδειχθεί ότι:

"Αν S είναι ένα σύνολο τύπων στην προτασιακή ή την κατηγορηματική λογική, υπάρχει ένα ισοδύναμο με το S σύνολο τύπων S΄ που είναι σε κανονική μορφή."

- Διαδικασία μετασχηματισμού ενός αυθαίρετου τύπου στην προτασιακή ή την κατηγορηματική (κατά Ginsberg) λογική σε ένα σύνολο τύπων σε κανονική μορφή:
  1. Απαλοιφή του $\rightarrow$ ( $p \rightarrow q = \neg p \vee q$ ) και, εφ' όσον υπάρχει, του $\leftrightarrow$
  2. Εφαρμογή των κανόνων του de Morgan, δηλαδή $\neg(p \wedge q) = \neg p \vee \neg q$ και $\neg(p \vee q) = \neg p \wedge \neg q$, και του $\neg(\neg p) = p$
  3. Εφαρμογή της επιμεριστικότητας των $\wedge$ και $\vee$ για το μετασχηματισμό του τύπου σε μία σύζευξη διαζεύξεων
  4. Διαχωρισμός της σύζευξης σε διακριτούς τύπους
  5. Συνδυασμός των ατόμων με άρνηση (αντίστροφη εφαρμογή de Morgan)
  6. Επαναεισαγωγή του $\rightarrow$ σε κάθε τύπο

Παράδειγμα:
Αν το σπίτι του Γιάννη είναι μεγάλο, τότε χρειάζεται πολλή δουλειά για να συντηρηθεί εκτός αν έχει προσληφθεί κάποιος για να το καθαρίζει και δεν υπάρχει κήπος στο σπίτι.

$$big(H) \wedge house(H, J) \rightarrow work(H) \vee (cleans(C, H) \wedge \neg garden(G, H))$$

big(H) $\equiv$ b

house(H, J) $\equiv$ h

work(H) $\equiv$ w

cleans(C, H) $\equiv$ c

garden(G, H) $\equiv$ g

Μετασχηματισμός του $b \wedge h \rightarrow w \vee (c \wedge \neg g)$ σε κανονική μορφή:

$b \wedge h \rightarrow w \vee (c \wedge \neg g)$

1. $\neg(b \wedge h) \vee w \vee (c \wedge \neg g)$
2. $\neg b \vee \neg h \vee w \vee (c \wedge \neg g)$
3. $(\neg b \vee \neg h \vee w \vee c) \wedge$
   $(\neg b \vee \neg h \vee w \vee \neg g)$
4. – $\neg b \vee \neg h \vee w \vee c$
   – $\neg b \vee \neg h \vee w \vee \neg g$
5. – $\neg(b \wedge h) \vee w \vee c$
   – $\neg(b \wedge h \wedge g) \vee w$
6. – $b \wedge h \rightarrow w \vee c$
   – $b \wedge h \wedge g \rightarrow w$

- Ο κανόνας της ανάλυσης εφαρμόζεται ορθώς και πλήρως και σε βάσεις που δεν είναι σε κανονική μορφή, αφού πρώτα αυτές μετασχηματισθούν (δηλαδή κάθε τύπος που περιλαμβάνουν) σε κανονική μορφή.

## ΛΟΓΙΚΗ ΠΡΩΤΗΣ ΤΑΞΗΣ
## (FIRST–ORDER LOGIC)

- Υπενθυμίζεται ότι στη λογική πρώτης τάξης τα εκφραστικά μέσα είναι αυτά που υπάρχουν και στην κατηγορηματική (κατά Ginsberg) λογική επαυξημένα με τις μεταβλητές και τους ποσοδείκτες.

- Έστω πάλι η γνωστή βάση γνώσης μας

(1)΄ lawyer(John)

(2)΄ $(\forall p)(\text{lawyer}(p) \rightarrow \text{rich}(p))$

(3)΄ $(\forall p)(\text{house}(\text{house-of}(p), p))$

(4)΄ $(\forall p)(\forall h)(\text{house}(h, p) \wedge \text{rich}(p) \rightarrow \text{big}(h))$

(5)΄ $(\forall h)(\text{big}(h) \wedge (\exists p)(\text{house}(h, p)) \rightarrow \text{work}(h))$

- Ο τύπος (3)΄ εκφράζει ουσιαστικά την ίδια γνώση με τον τύπο $(\forall p)(\exists h)(\text{house}(h, p))$. Από αυτόν τον τελευταίο τύπο μπορεί να γίνει με ασφάλεια η απαλοιφή του υπαρξιακού ποσοδείκτη μέσω ενός μετασχηματισμού που ονομάζεται Σκολεμοποίηση (Skolemization).

Η Σκολεμοποίηση του τύπου $(\forall v_1) \ldots\ldots (\forall v_n) (\exists x)(q(v_i, x))$ δίνει τον τύπο $(\forall v_1) \ldots\ldots (\forall v_n)(q(v_i, S_{k-n}(v_1, \ldots\ldots, v_n)))$ όπου το $S_{k-n}$ είναι ένα συναρτησιακό σύμβολο βαθμού n που δεν εμφανίζεται ξανά στη βάση μας.

Π.χ.

$$(\exists h)(\text{house}(h, \text{John})) \xrightarrow{\text{Σκολεμοποίηση}} \text{house}(S_{k-1}, \text{John})$$

$$(\forall p)(\exists h)(\text{house}(h, p)) \longrightarrow (\forall p)(\text{house}(S_{k-2}(p), p))$$

$$(\exists h)(\forall p)(\text{house}(h, p)) \longrightarrow (\forall p)(\text{house}(S_{k-3}, p))$$

- Αποδεικνύεται ότι:

  Αν D είναι μία βάση γνώσης και $S_k(D)$ η βάση που προκύπτει από τη Σκολεμοποίηση της D, τότε οι D και $S_k(D)$ είναι ισοδύναμες με την έννοια ότι για κάθε τύπο q που δεν περιέχει Skolem σύμβολα ισχύει:

  $D \models q$ αν και μόνο αν $S_k(D) \models q$

- Στον τύπο (5)΄ της βάσης μας, μπορεί να γίνει απαλοιφή του υπαρξιακού ποσοδείκτη (όχι όμως με Σκολεμοποίηση – γιατί άραγε;) ως εξής:

  $(\forall h)(big(h) \wedge (\exists p)(house(h, p)) \rightarrow work(h))$ =

  $(\forall h)(\neg (big(h) \wedge (\exists p)(house(h, p))) \vee work(h))$ =

  $(\forall h)(\neg big(h) \vee \neg(\exists p)(house(h, p)) \vee work(h))$ =

  $(\forall h)(\neg big(h) \vee (\forall p)(\neg house(h, p)) \vee work(h))$ =

  $(\forall h)(\forall p)(\neg big(h) \vee \neg house(h, p) \vee work(h))$ =

  $(\forall h)(\forall p)(\neg(big(h) \wedge house(h, p)) \vee work(h))$ =

  $(\forall h)(\forall p)(big(h) \wedge house(h, p) \rightarrow work(h))$

- Τελικά, αφού σε κάθε τύπο της βάσης μας, οι μεταβλητές του ποσοτικοποιούνται στην αρχή, και μόνο εκεί, με καθολικό ποσοδείκτη, μπορούμε να μη γράψουμε τα $(\forall ....)$.

  Τέλος, αν μετονομάσουμε και τις μεταβλητές ώστε να μην υπάρχουν κοινά ονόματα μεταξύ των τύπων, έχουμε:

  (1) lawyer(John)

  (2) $lawyer(p_2) \rightarrow rich(p_2)$

  (3) $house(house\text{-}of(p_3), p_3)$

  (4) $house(h_4, p_4) \wedge rich(p_4) \rightarrow big(h_4)$

  (5) $big(h_5) \wedge house(h_5, p_5) \rightarrow work(h_5)$

- Ένας όρος ή τύπος στη λογική πρώτης τάξης που περιέχει μεταβλητές μπορεί να εξειδικευθεί περισσότερο αν σε μία ή περισσότερες από τις μεταβλητές του αποδοθεί συγκεκριμένη τιμή. Για παράδειγμα, ο τύπος

$$\text{big}(h_5) \wedge \text{house}(h_5, p_5) \longrightarrow \text{work}(h_5)$$

για $h_5$ = house-of(John) γίνεται

$$\text{big(house-of(John))} \wedge \text{house(house-of(John)}, p_5) \longrightarrow \text{work(house-of(John))}$$

- Μία λίστα δεσμεύσεων (binding list) σ είναι ένα σύνολο από στοιχεία της μορφής v = e, όπου v είναι μεταβλητή και e είναι όρος.

- Η εφαρμογή μίας λίστας δεσμεύσεων σ σε μία έκφραση (όρο ή τύπο) p συμβολίζεται με $p|_\sigma$.

Π.χ. $\text{work}(h_5)|_{\{h5 = \text{house-of(John)}\}} = \text{work(house-of(John))}$

$\text{lawyer}(p)|_{\{p = \text{John}\}} = \text{lawyer(John)}$

- Μία λίστα δεσμεύσεων σ που εφαρμοζόμενη σε δυο εκφράσεις p και q τις κάνει πανομοιότυπες, δηλαδή $p|_\sigma = q|_\sigma$, ονομάζεται ενοποιητής (unifier) των p και q και η διαδικασία εύρεσής της ονομάζεται ενοποίηση (unification).

Π.χ. $\text{likes(John, x)}|_{\{x = \text{Mary}, y = \text{John}\}} =$

$\text{likes(y, Mary)}|_{\{x = \text{Mary}, y = \text{John}\}} =$

likes(John, Mary)

- Αν $\sigma_1$ και $\sigma_2$ είναι δύο λίστες δεσμεύσεων, η $\sigma_1$ είναι γενικότερη (more general) από τη $\sigma_2$, αν για κάθε έκφραση p, η $p|_{\sigma_2}$ αποτελεί μία εξειδίκευση της $p|_{\sigma_1}$, ή, αλλιώς, αν υπάρχει μία λίστα δεσμεύσεων $\sigma_3$, τέτοια ώστε για κάθε έκφραση p να ισχύει $p|_{\sigma_2} = (p|_{\sigma_1})|_{\sigma_3}$

  Π.χ. Η $\sigma_1$ = {y = John} είναι γενικότερη
  από την $\sigma_2$ = {x = Mary, y = John} και
  η $\sigma_1$ = {h = house-of(p)} είναι γενικότερη
  από την $\sigma_2$ = {h = house-of(John)}

- Αυτό που ενδιαφέρει κατά τη διαδικασία της ενοποίησης είναι η εύρεση του γενικότερου ενοποιητή (most general unifier) δύο εκφράσεων p και q.

  Η διαδικασία εύρεσής του υπόκειται στους εξής κανόνες:
  - Δύο σταθερές ενοποιούνται εφ' όσον είναι ίδιες, χωρίς να συνεισφέρουν κάποια δέσμευση στο γενικότερο ενοποιητή.
  - Μία μεταβλητή v ενοποιείται με οποιοδήποτε όρο e, προσθέτοντας τη δέσμευση v = e στο γενικότερο ενοποιητή.
  - Δύο άτομα $p(s_1, ...., s_n)$ και $q(u_1, ...., u_m)$ ή δύο όροι $f(s_1, ...., s_n)$ και $g(u_1, ...., u_m)$ ενοποιούνται όταν p = q ή f = g αντίστοιχα και επίσης οι αντίστοιχοι όροι–ορίσματα είναι ίσου πλήθους και ενοποιούνται αντίστοιχα, δηλαδή n = m και $s_i$ ενοποιείται με κάθε $u_i$ για κάθε i.

- Παραδείγματα γενικότερων ενοποιητών mgu(p, q) των εκφράσεων p και q
  - p = x , q = John
    mgu(p, q) = {x = John}
  - p = x, q = house-of(John)
    mgu(p, q) = {x = house-of(John)}
  - p = lawyer(x), q = lawyer(John)
    mgu(p, q) = {x = John}
  - p = likes(John, x), q = likes(y, Mary)
    mgu(p, q) = {x = Mary, y = John}
  - p = likes(John, x), q = likes(x, Mary)
    $\nexists$ mgu(p, q)
  - p = f(x, y), q = f(y, a)
    mgu(p, q) = {x = a, y = a}
  - p = likes(x, house-of(x)), q = likes(y, y)
    mgu(p, q) = {x = y, y = house-of(x)}

- Το τελευταίο παράδειγμα εισάγει το πρόβλημα του ελέγχου εμφάνισης (occurs check). Κανονικά η απόπειρα ενοποίησης των συγκεκριμένων p και q θα έπρεπε να αποτυγχάνει επειδή οδηγεί στη δέσμευση x = house-of(house-of(....... , αλλά η αναγνώριση αυτής της κατάστασης έχει μεγάλο κόστος και στα πρακτικά συστήματα που απαιτούν ενοποίηση (π.χ. Prolog) δεν γίνεται.

- Η τελική μορφή της βάσης γνώσης μας προέκυψε από τη Σκολεμοποίηση των τύπων που περιλάμβανε (για την απαλοιφή των υπαρξιακών ποσοδεικτών) και τη διαγραφή των καθολικών ποσοδεικτών (αφού προτάσσονται σε κάθε τύπο).

- Ένας τύπος προς απόδειξη μπορεί επίσης να περιέχει ποσοδείκτες. Στην προς τα πίσω απόπειρα απόδειξης του τύπου πρέπει να προστεθεί η άρνησή του στη βάση και να αποδειχθεί η αντίφαση. Στην άρνηση του τύπου υπάρχει ενδεχόμενο για Σκολεμοποίηση ή/και διαγραφή καθολικών ποσοδεικτών.

  Παραδείγματα

  – $q = (\forall x)(\text{lawyer}(x))$

    Ισχύει ότι κάθε x είναι δικηγόρος;

    $q = \neg(\forall x)(\text{lawyer}(x)) =$

    $(\exists x)(\neg \text{lawyer}(x)) =$ (Σκολεμοποίηση)

    $\neg \text{lawyer}(S_{k-4}) =$

    $\text{lawyer}(S_{k-4}) \rightarrow F$

  – $q = (\exists x)(\text{work}(x))$

    Ισχύει για κάποιο x ότι χρειάζεται πολλή δουλειά να συντηρηθεί από τον ιδιοκτήτη του;

    $q = \neg(\exists x)(\text{work}(x)) =$

    $(\forall x)(\neg \text{work}(x)) =$ (διαγραφή $\forall$ )

    $\neg \text{work}(x) =$

    $\text{work}(x) \rightarrow F$

- Στην πράξη, η μορφή του δεύτερου παραδείγματος είναι η συχνότερα εμφανιζόμενη.

- Γενίκευση του “modus ponens” στη λογική πρώτης τάξης

modus ponens:

Αν $mgu(d, a_i) = \sigma$

$$\frac{\begin{array}{r} a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow b \\ c_1 \wedge \ldots\ldots\ldots \wedge c_n \rightarrow d \end{array}}{(a_1 \wedge \ldots\ldots\ldots \wedge \hat{a}_i \wedge \ldots\ldots\ldots \wedge a_m \wedge c_1 \wedge \ldots\ldots\ldots \wedge c_n \rightarrow b)\Big|_{\sigma}}$$

- Και στη λογική πρώτης τάξης αποδεικνύεται ότι:

  “Ο κανόνας modus ponens είναι ορθός και,

  για βάσεις Horn, πλήρης”

- Σε μία βάση Horn, στη λογική πρώτης τάξης, χωρίς συναρτησιακά σύμβολα, η διαδικασία απόδειξης αν ένα συγκεκριμένο άτομο προκύπτει από τη βάση δεν χρειάζεται χρόνο περισσότερο από $n \cdot o^v \cdot d$, όπου n είναι ο μέγιστος αριθμός υποθέσεων των τύπων της βάσης, o είναι το πλήθος των σταθερών στη βάση, v είναι το μέγιστο πλήθος των μεταβλητών σε κάθε τύπο της βάσης και d είναι το μέγεθος (πλήθος τύπων) της βάσης.

- Παρά τη “σκληρή” υπόθεση που έγινε, ότι δηλαδή δεν υπάρχουν συναρτησιακά σύμβολα, ο παράγοντας $o^v$, που δεν υπάρχει στο αντίστοιχο φράγμα για την προτασιακή ή την κατηγορηματική (κατά Ginsberg) λογική, μπορεί να έχει δραματικές επιπτώσεις στην πράξη.

- Γενίκευση του “κανόνα της ανάλυσης” στη λογική πρώτης τάξης.

  κανόνας της ανάλυσης:

  Αν $mgu(d_j, a_i) = \sigma$

$$\frac{\begin{array}{l} a_1 \wedge \ldots\ldots\ldots \wedge a_m \rightarrow b_1 \vee \ldots\ldots\ldots \vee b_k \\ c_1 \wedge \ldots\ldots\ldots \wedge c_n \rightarrow d_1 \vee \ldots\ldots\ldots \vee d_q \end{array}}{(a_1 \wedge \ldots \wedge \hat{a}_i \wedge \ldots \wedge a_m \wedge c_1 \wedge \ldots \wedge c_n \rightarrow b_1 \vee \ldots \vee b_k \vee d_1 \vee \ldots \vee \hat{d}_j \vee \ldots \vee d_q) \Big|\, \sigma}$$

- Εφαρμογή του κανόνα της ανάλυσης για την εύρεση “κάποιου x που χρειάζεται πολλή δουλειά για να συντηρηθεί από τον ιδιοκτήτη του” με δεδομένη τη βάση γνώσης:

  (1) lawyer(John)

  (2) $lawyer(p_2) \rightarrow rich(p_2)$

  (3) $house(house\text{-}of(p_3), p_3)$

  (4) $house(h_4, p_4) \wedge rich(p_4) \rightarrow big(h_4)$

  (5) $big(h_5) \wedge house(h_5, p_5) \rightarrow work(h_5)$

  Η ερώτηση είναι:

  $(\exists x)(work(x))$

  Για την εφαρμογή του κανόνα της ανάλυσης με οπίσθιο τρόπο, πρέπει να προστεθεί στη βάση η άρνηση της ερώτησης, δηλαδή η

  $\neg(\exists x)(work(x)) = (\forall x)(\neg work(x))$

  άρα με απαλοιφή του $(\forall x)$ η $\neg work(x)$ και να γίνει στη συνέχεια προσπάθεια απόδειξης της αντίφασης $T \rightarrow F$.

| | Αποδεικνυόμενοι τύποι | Αιτιολόγηση | Δεσμεύσεις |
|---|---|---|---|
| a | $work(x) \rightarrow F$ | άρνηση αποδεικτέου | |
| b | $big(h_5) \wedge house(h_5, p_5) \rightarrow F$ | από (a), (5) | $x = h_5$ |
| c | $big(house\text{-}of(p_3)) \rightarrow F$ | από (b), (3) | $h_5=house\text{-}of(p_3)$, $p_5=p_3$ |
| d | $house(house\text{-}of(p_3), p_4) \wedge rich(p_4) \rightarrow F$ | από (c), (4) | $h_4 = house\text{-}of(p_3)$ |
| e | $rich(p_3) \rightarrow F$ | από (d), (3) | $p_4 = p_3$ |
| f | $lawyer(p_3) \rightarrow F$ | από (e), (2) | $p_2 = p_3$ |
| g | $T \rightarrow F$ | από (f), (1) | $p_3 = John$ |

Άρα το ζητούμενο αποδείχθηκε για

$x = h_5 = house\text{-}of(p_3) = house\text{-}of(John)$

- Ο κανόνας της ανάλυσης εξακολουθεί να είναι ορθός και πλήρης και στη λογική πρώτης τάξης. Για να εφαρμοστεί όμως σε κάποια βάση γνώσης πρέπει αυτή να βρίσκεται σε κανονική μορφή, δηλαδή οι τύποι της να είναι της μορφής:

$$a_1 \wedge \ldots\ldots \wedge a_m \rightarrow b_1 \vee \ldots\ldots \vee b_k$$

όπου όλα τα $a_i$ και $b_j$ είναι άτομα.

- Και στη λογική πρώτης τάξης υπάρχει μεθοδολογία, ουσιαστικά επέκταση αυτής που δόθηκε για την κατηγορηματική λογική, με την οποία ένα σύνολο τύπων S μετασχηματίζονται σε ένα άλλο S΄ που βρίσκεται σε κανονική μορφή. Η επέκταση συνίσταται στο χειρισμό των ποσοδεικτών.

• Διαδικασία μετασχηματισμού ενός αυθαίρετου τύπου της λογικής πρώτης τάξης σε ένα σύνολο τύπων σε κανονική μορφή:

1. Απαλοιφή του $\rightarrow$ ($p \rightarrow q = \neg p \vee q$) και, εφ’ όσον υπάρχει, του $\leftrightarrow$
2. Εφαρμογή των κανόνων του de Morgan, δηλαδή $\neg(p \wedge q) = \neg p \vee \neg q$ και $\neg(p \vee q) = \neg p \wedge \neg q$, και του $\neg(\neg p) = p$ για τη μεταφορά της άρνησης σε απλά άτομα

* 3. Διαφοροποίηση ονομάτων μεταβλητών που ποσοτικοποιούνται από διαφορετικούς ποσοδείκτες

* 4. Απαλοιφή των υπαρξιακών ποσοδεικτών $\exists$ με Σκολεμοποίηση

* 5. Μετακίνηση των καθολικών ποσοδεικτών $\forall$ στην αρχή του τύπου

* 6. Διαγραφή της καθολικής ποσοτικοποίησης

7. Εφαρμογή της επιμεριστικότητας των $\vee$ και $\wedge$ για το μετασχηματισμό του τύπου σε μία σύζευξη διαζεύξεων
8. Διαχωρισμός της σύζευξης σε διακριτούς τύπους
9. Συνδυασμός των ατόμων με άρνηση (αντίστροφη εφαρμογή de Morgan)
10. Επαναεισαγωγή του $\rightarrow$ σε κάθε τύπο

* 11. Διαφοροποίηση των ονομάτων μεταβλητών μεταξύ των διαφόρων τύπων

Στην προηγούμενη διαδικασία, τα βήματα που είναι σημειωμένα με * αποτελούν την επέκταση σε σχέση με την αντίστοιχη διαδικασία για την προτασιακή ή την κατηγορηματική (κατά Ginsberg) λογική.

- Μετασχηματισμός σε κανονική μορφή του:

$(\forall h)(big(h) \wedge house(h) \rightarrow work(h) \vee$
$((\exists x)(cleans(x, h)) \wedge \neg(\exists x)(garden(x, h))))$

1. $(\forall h)(\neg(big(h) \wedge house(h)) \vee work(h) \vee$
   $((\exists x)(cleans(x, h)) \wedge \neg(\exists x)(garden(x, h))))$

2. $(\forall h)(\neg big(h) \vee \neg house(h) \vee work(h) \vee$
   $((\exists x)(cleans(x, h)) \wedge (\forall x)(\neg garden(x, h))))$

3. $(\forall h)(\neg big(h) \vee \neg house(h) \vee work(h) \vee$
   $((\exists m)(cleans(m, h)) \wedge (\forall g)(\neg garden(g, h))))$

4. $(\forall h)(\neg big(h) \vee \neg house(h) \vee work(h) \vee$
   $(cleans(S_k\text{-}cleans(h), h) \wedge (\forall g)(\neg garden(g, h))))$

5. $(\forall h)(\forall g)(\neg big(h) \vee \neg house(h) \vee work(h) \vee$
   $(cleans(S_k\text{-}cleans(h), h) \wedge \neg garden(g, h)))$

6. $\neg big(h) \vee \neg house(h) \vee work(h) \vee$
   $(cleans(S_k\text{-}cleans(h), h) \wedge \neg garden(g, h))$

7. $(\neg big(h) \vee \neg house(h) \vee work(h) \vee cleans(S_k\text{-}cleans(h), h)) \wedge$
   $(\neg big(h) \vee \neg house(h) \vee work(h) \vee \neg garden(g, h))$

8. – $\neg big(h) \vee \neg house(h) \vee work(h) \vee cleans(S_k\text{-}cleans(h), h)$
   – $\neg big(h) \vee \neg house(h) \vee work(h) \vee \neg garden(g, h)$

9. – $\neg(big(h) \wedge house(h)) \vee work(h) \vee cleans(S_k\text{-}cleans(h), h)$
   – $\neg(big(h) \wedge house(h) \wedge garden(g, h)) \vee work(h)$

10. – $big(h) \land house(h) \rightarrow work(h) \lor cleans(S_k\text{-}cleans(h), h)$

– $big(h) \land house(h) \land garden(g, h) \rightarrow work(h)$

11. – $big(h_1) \land house(h_1) \rightarrow work(h_1) \lor cleans(S_k\text{-}cleans(h_1), h_1)$

– $big(h_2) \land house(h_2) \land garden(g_2, h_2) \rightarrow work(h_2)$

## Μερικές λυμένες ασκήσεις

1. Τη γνώση “οι δικηγόροι είναι πλούσιοι” την αναπαραστήσαμε στη λογική πρώτης τάξης με τον τύπο $(\forall x)(lawyer(x) \rightarrow rich(x))$. Συνεπώς, τη γνώση “κάποιος δικηγόρος είναι πλούσιος” πρέπει να την αναπαραστήσουμε με τον τύπο $(\exists x)(lawyer(x) \rightarrow rich(x))$; Αν όχι, ποια είναι η σωστή απάντηση;

### Λύση

Η σωστή αναπαράσταση του “κάποιος δικηγόρος είναι πλούσιος” είναι $(\exists x)(lawyer(x) \land rich(x))$. Ο τύπος $(\exists x)(lawyer(x) \rightarrow rich(x)) = (\exists x)(\neg lawyer(x) \lor rich(x))$ δεν εκφράζει αυτό που μας ενδιαφέρει γιατί μπορεί να ισχύει επειδή είτε υπάρχει κάποιος που δεν είναι δικηγόρος είτε κάποιος που είναι πλούσιος (κάθε ένα πολύ πιθανό) αλλά όχι κάποιος που είναι ταυτόχρονα και δικηγόρος και πλούσιος. Αυτή η απαίτηση του “ταυτοχρόνου” εισάγει τη σύζευξη. Αντίθετα, το “οι δικηγόροι είναι πλούσιοι” είναι το ίδιο με το “αν κάποιος είναι δικηγόρος, τότε είναι πλούσιος”, οπότε η γνώση που μας ενδιαφέρει πρέπει να διατυπωθεί στη λογική πρώτης τάξης με συνεπαγωγή.

2. Έστω ότι μας είναι γνωστό ότι

“ένα μήλο την ημέρα το γιατρό τον κάνει πέρα”.

Κωδικοποιήστε το σε λογική πρώτης τάξης και μετασχηματίστε το σε κανονική μορφή. Αφού κωδικοποιήσετε σε κανονική μορφή και την ερώτηση

“ισχύει ότι ένα μήλο την ημέρα το γιατρό τον κάνει πέρα;”

εφαρμόστε διαδοχικά τον κανόνα της ανάλυσης για να πάρετε απάντηση στην ερώτηση αυτή.

Λύση

Έστω ότι έχουμε τα εξής κατηγορήματα:

day(x) : το x είναι ημέρα

apple(y) : το y είναι μήλο

keeps_doctor_away(y, x) : το μήλο y το γιατρό τον κάνει πέρα την ημέρα x

Τότε η γνώση

“ένα μήλο την ημέρα το γιατρό τον κάνει πέρα”

κωδικοποιείται σαν

$(\forall x)(day(x) \rightarrow (\exists y)(apple(y) \wedge keeps\_doctor\_away(y, x)))$

Ο τύπος αυτός μετασχηματίζεται σε κανονική μορφή ως εξής:

$(\forall x)(day(x) \rightarrow (\exists y)(apple(y) \wedge keeps\_doctor\_away(y, x))) =$

$(\forall x)(\neg day(x) \vee (\exists y)(apple(y) \wedge keeps\_doctor\_away(y, x))) =$

$(\forall x)(\neg day(x) \vee (apple(S_k\text{-}a(x)) \wedge keeps\_doctor\_away(S_k\text{-}a(x), x))) =$

$\neg day(x) \vee (apple(S_k\text{-}a(x)) \wedge keeps\_doctor\_away(S_k\text{-}a(x), x)) =$

$(\neg day(x) \vee apple(S_k\text{-}a(x)))\ \wedge$

$(\neg day(x) \vee keeps\_doctor\_away(S_k\text{-}a(x), x))\ =$

$$\left\{\begin{array}{l} \neg day(x) \vee apple(S_k\text{-}a(x)) \\ \neg day(x) \vee keeps\_doctor\_away(S_k\text{-}a(x), x) \end{array}\right\} =$$

$$\left\{\begin{array}{l} day(x) \rightarrow apple(S_k\text{-}a(x)) \\ day(x) \rightarrow keeps\_doctor\_away(S_k\text{-}a(x), x) \end{array}\right\} =$$

$$\left\{\begin{array}{ll} day(x_1) \rightarrow apple(S_k\text{-}a(x_1)) & (1) \\ day(x_2) \rightarrow keeps\_doctor\_away(S_k\text{-}a(x_2), x_2) & (2) \end{array}\right\}$$

Για να βρεθεί η απάντηση στην ερώτηση "ισχύει ότι ένα μήλο την ημέρα τον γιατρό τον κάνει πέρα;" πρέπει να κωδικοποιηθεί σε λογική πρώτης τάξης η άρνηση του αποδεικτέου και, αφού μετασχηματισθεί σε κανονική μορφή, να εφαρμοσθεί με οπίσθιο τρόπο ο κανόνας της ανάλυσης, με βάση φυσικά τη γνώση (1) και (2), για να αποδειχθεί η αντίφαση $T \rightarrow F$. Έχουμε λοιπόν:

$\neg(\forall x)(day(x) \rightarrow (\exists y)(apple(y) \wedge keeps\_doctor\_away(y, x)))\ =$

$\neg(\forall x)(\neg day(x) \vee (\exists y)(apple(y) \wedge keeps\_doctor\_away(y, x)))\ =$

$(\exists x)(\neg(\neg day(x)) \wedge \neg(\exists y)(apple(y) \wedge keeps\_doctor\_away(y, x)))\ =$

$(\exists x)(day(x) \wedge (\forall y)(\neg apple(y) \vee \neg keeps\_doctor\_away(y, x)))\ =$

$day(S_k\text{-}d) \wedge (\forall y)(\neg apple(y) \vee \neg keeps\_doctor\_away(y, S_k\text{-}d))\ =$

$(\forall y)(day(S_k\text{-}d) \wedge (\neg apple(y) \vee \neg keeps\_doctor\_away(y, S_k\text{-}d)))\ =$

$day(S_k\text{-}d) \wedge (\neg apple(y) \vee \neg keeps\_doctor\_away(y, S_k\text{-}d))\ =$

$$\left.\begin{array}{l} day(S_k\text{-}d) \\ \neg apple(y) \vee \neg keeps\_doctor\_away(y, S_k\text{-}d) \end{array}\right\} =$$

$$\left.\begin{array}{l} day(S_k\text{-}d) \\ \neg(apple(y) \wedge keeps\_doctor\_away(y, S_k\text{-}d)) \end{array}\right\} =$$

$$\left.\begin{array}{ll} T \rightarrow day(S_k\text{-}d) & (3) \\ apple(y) \;\wedge\; keeps\_doctor\_away(y, S_k\text{-}d) \rightarrow F & (4) \end{array}\right\}$$

Εφαρμόζοντας τώρα διαδοχικά τον κανόνα της ανάλυσης στους τύπους (1), (2), (3) και (4) έχουμε:

(1), (3) : $T \rightarrow apple(S_k\text{-}a(S_k\text{-}d))$ (5)

(2), (3) : $T \rightarrow keeps\_doctor\_away(S_k\text{-}a(S_k\text{-}d), S_k\text{-}d)$ (6)

(4), (5) : $keeps\_doctor\_away(S_k\text{-}a(S_k\text{-}d), S_k\text{-}d) \rightarrow F$ (7)

(7), (6) : $T \rightarrow F$ (ο.ε.δ.)

Άρα η απάντηση στην ερώτησή μας είναι καταφατική (όπως το περιμέναμε άλλωστε).

3. Έστω το ακόλουθο μαθηματικό πρόβλημα: "Υπάρχουν άρρητοι αριθμοί x και y τέτοιοι ώστε το $x^y$ να είναι ρητός ;". Ιδού μία απόδειξη ότι υπάρχουν τέτοιοι αριθμοί: "Έστω $z = \sqrt{2}^{\sqrt{2}}$. Αν ο z είναι ρητός, τότε αποδείχτηκε το ζητούμενο (αφού το $\sqrt{2}$ είναι γνωστό ότι είναι άρρητος). Αν ο z είναι άρρητος, ας θεωρήσουμε το $z^{\sqrt{2}}$. Έχουμε τότε $z^{\sqrt{2}} = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{(\sqrt{2}\cdot\sqrt{2})} = \sqrt{2}^{2} = 2$. Συνεπώς η ύψωση του άρρητου $z = \sqrt{2}^{\sqrt{2}}$ στη δύναμη $\sqrt{2}$ (άρρητος και αυτός) δίνει το ρητό 2. Συνεπώς πάλι αποδείχθηκε το ζητούμενο". Διατυπώστε αυτήν την απόδειξη με διαδοχικές εφαρμογές του κανόνα της ανάλυσης, αφού πρώτα κωδικοποιήσετε την απαιτούμενη γνώση σε λογική πρώτης τάξης καθώς επίσης και την απαραίτητη ερώτηση που πρέπει να απαντηθεί (ή θεώρημα προς απόδειξη).

Λύση (.... τμήμα της)

Θα χρησιμοποιήσουμε τα κατηγορήματα :

times(x, y, z) : $x \cdot y = z$

expt(x, y, z) : $x^y = z$

rational(x) : ο x είναι ρητός

Η απαιτούμενη γνώση για να διατυπωθεί η απόδειξη των ζητούμενων μέσω του κανόνα της ανάλυσης είναι:

α) Ισχύει ότι $(a^x)^y = a^{x \cdot y}$ ή αλλιώς:

Αν $z = x \cdot y$, $a^x = b$ και $a^z = c$, τότε $b^y = c$

$(\forall x)(\forall y)(\forall z)(\forall a)(\forall b)(\forall c)(\text{times}(x, y, z) \land \text{expt}(a, x, b) \land$
$\text{expt}(a, z, c) \rightarrow \text{expt}(b, y, c))$

β) Ισχύει ότι $\sqrt{2} \cdot \sqrt{2} = 2$

$\text{times}(\sqrt{2}, \sqrt{2}, 2)$

γ) Ισχύει ότι $\sqrt{2}^{\,2} = 2$

$\text{expt}(\sqrt{2}, 2, 2)$

δ) Ο αριθμός $\sqrt{2}$ είναι άρρητος

$\neg\text{rational}(\sqrt{2})$

ε) Ο αριθμός 2 είναι ρητός

rational(2)

στ) Ισχύει ότι η ύψωση σε δύναμη είναι καλώς ορισμένη

$(\forall x)(\forall y)(\exists z)(\text{expt}(x, y, z)) =$

$(\forall x)(\forall y)(\text{expt}(x, y, \text{pow}(x, y)))$

Το ζητούμενο προς απόδειξη είναι:

$(\exists x)(\exists y)(\exists z)(\text{expt}(x, y, z) \land \neg\text{rational}(x) \land$
$\neg\,\text{rational}(y) \land \text{rational}(z))$

Για να χρησιμοποιηθεί ο κανόνας της ανάλυσης με οπίσθιο τρόπο, πρέπει να προσθέσουμε στη βάση γνώσης μας και την άρνηση του αποδεικτέου, δηλαδή:

$$\neg(\exists x)(\exists y)(\exists z)(expt(x, y, z) \land \neg rational(x) \land \neg rational(y) \land rational(z) =$$
$$(\forall x)(\forall y)(\forall z)(\neg expt(x, y, z) \lor rational(x) \lor rational(y) \lor \neg rational(z))$$

Συνεπώς, μετά από τις απαραίτητες διαγραφές καθολικών ποσοδεικτών, μετονομασίες μεταβλητών και διατυπώσεις μέσω συνεπαγωγών, πρέπει να αποδειχθεί η αντίφαση $T \rightarrow F$ από τα εξής:

$$times(x_1, y_1, z_1) \land expt(a, x_1, b) \land expt(a, z_1, c) \rightarrow expt(b, y_1, c) \quad (1)$$
$$T \rightarrow times(\sqrt{2}, \sqrt{2}, 2) \quad (2)$$
$$T \rightarrow expt(\sqrt{2}, 2, 2) \quad (3)$$
$$rational(\sqrt{2}) \rightarrow F \quad (4)$$
$$T \rightarrow rational(2) \quad (5)$$
$$T \rightarrow expt(x_2, y_2, pow(x_2, y_2)) \quad (6)$$
$$expt(x_3, y_3, z_3) \land rational(z_3) \rightarrow rational(x_3) \lor rational(y_3) \quad (7)$$

Η συνέχεια αφήνεται για άσκηση.

4. “Ο Μότσαρτ επισκέφθηκε τη Βιέννη τρεις φορές και πέθανε εκεί. Σε ποια από τις τρεις επισκέψεις πέθανε;”. Διατυπώστε το πρόβλημα σε λογική πρώτης τάξης και χρησιμοποιήστε τον κανόνα της ανάλυσης για να βρείτε την απάντηση. Ποια επιπλέον γνώση χρειάζεται για να μπορεί να προκύψει η απάντηση από τα δεδομένα;

Λύση

$T \rightarrow$ visit(Mozart, Vienna, first) (1)

$T \rightarrow$ visit(Mozart, Vienna, second) (2)

$T \rightarrow$ visit(Mozart, Vienna, third) (3)

$T \rightarrow$ place_died(Mozart, Vienna) (4)

$(\forall x)(\forall y)(\forall z)$(place_died(x, y) $\wedge$ last_visit(x, y, z) $\rightarrow$ visit_died(x, z)) = (μετασχηματισμός σε κανονική μορφή)

place_died(x, y) $\wedge$ last_visit(x, y, z) $\rightarrow$ visit_died(x, z) (5)

Απαιτείται επίσης μία σχέση διάταξης μεταξύ των επισκέψεων first, second και third.

$T \rightarrow$ after(second, first) (6)

$T \rightarrow$ after(third, first) (7)

$T \rightarrow$ after(third, second) (8)

"Κάποια επίσκεψη σε ένα μέρος είναι η τελευταία αν δεν υπάρχει μεταγενέστερη από αυτήν"

$(\forall t)(\forall u)(\forall v)$(visit(t, u, v) $\wedge \neg(\exists w)$(after(w, v)) $\rightarrow$ last_visit(t, u, v)) = (μετασχηματισμός σε κανονική μορφή)

visit(t, u, v) $\rightarrow$ after($S_k$(t, u, v), v) $\vee$ last_visit(t, u, v) (9)

Επίσης πρέπει η σχέση after να είναι πλήρως ορισμένη. Δηλαδή πρέπει να ισχύει μόνο για τα ζευγάρια που συμμετέχουν στις (6), (7) και (8).

$(\forall m)(\forall n)(\text{after}(m, n) \rightarrow (\text{eq}(m, \text{second}) \wedge \text{eq}(n, \text{first})) \vee$

$(\text{eq}(m, \text{third}) \wedge \text{eq}(n, \text{first})) \vee (\text{eq}(m, \text{third}) \wedge \text{eq}(n, \text{second})))\ =$

(μετασχηματισμός σε κανονική μορφή)

$$\text{8 τύποι}\left\{\begin{array}{ll} \text{..............} & \\ \text{..............} & \\ \text{..............} & \\ \text{after}(m, n) \rightarrow \text{eq}(n, \text{first}) \vee \text{eq}(n, \text{second}) & (10) \end{array}\right.$$

Τέλος χρειάζεται και ο ορισμός της ισότητας eq

$T \rightarrow \text{eq}(\text{first}, \text{first})$

..............................

$\text{eq}(\text{third}, \text{first}) \rightarrow F$ (11)

$\text{eq}(\text{third}, \text{second}) \rightarrow F$ (12)

Με την ερώτηση μας προστίθεται στη βάση γνώσης

και το $\text{visit\_died}(\text{Mozart}, r) \rightarrow F$ (13)

όποτε ο κανόνας της ανάλυσης δίνει:

(13) $\text{visit\_died}(\text{Mozart}, r) \rightarrow F$

+(5) $\text{place\_died}(\text{Mozart}, y) \wedge \text{last\_visit}(\text{Mozart}, y, r) \rightarrow F$

+(4) $\text{last\_visit}(\text{Mozart}, \text{Vienna}, r) \rightarrow F$

+(9) $\text{visit}(\text{Mozart}, \text{Vienna}, r) \rightarrow \text{after}(S_k(\text{Mozart}, \text{Vienna}, r), r)$

+(10) $\text{visit}(\text{Mozart}, \text{Vienna}, r) \rightarrow \text{eq}(r, \text{first}) \vee \text{eq}(r, \text{second})$

+(3) $T \rightarrow \text{eq}(\text{third}, \text{first}) \vee \text{eq}(\text{third}, \text{second})$ (όπου $r = \text{third}$)

+(11) $T \rightarrow \text{eq}(\text{third}, \text{second})$

+(12) $T \rightarrow F$

Άρα αποδείχθηκε ότι “ο Μότσαρτ πέθανε κατά την τρίτη

(και τελευταία – τι περίεργο!) επίσκεψή του στη Βιέννη”.

## Η ΛΟΓΙΚΗ ΣΕ ΔΡΑΣΗ
## ΕΛΕΓΧΟΣ ΣΤΗ ΣΥΛΛΟΓΙΣΤΙΚΗ
## (CONTROL OF REASONING)

Έλεγχος στη συλλογιστική =
Το να συλλογίζεσαι για το πώς θα συλλογισθείς =
Σύνολο ενεργειών μετα–επιπέδου

Η συλλογιστική διαδικασία ποτέ δεν είναι στην πράξη τόσο απλή όσο στην απόδειξη του

$$(\exists x)(work(x))$$

από τα

(1) lawyer(John)

(2) lawyer($p_2$) $\rightarrow$ rich($p_2$)

(3) house(house–of($p_3$), $p_3$)

(4) house($h_4$, $p_4$) $\wedge$ rich($p_4$) $\rightarrow$ big($h_4$)

(5) big($h_5$) $\wedge$ house($h_5$, $p_5$) $\rightarrow$ work($h_5$)

Η συλλογιστική είναι ένα πρόβλημα αναζήτησης αφού ο κανόνας της ανάλυσης μπορεί να έχει πολλές δυνατότητες εφαρμογής σε δεδομένη κατάσταση της βάσης γνώσης.

Στην πράξη απαιτείται η εφαρμογή συγκεκριμένης στρατηγικής για την ανάλυση (resolution strategy), όπως

— η στρατηγική του συνόλου υποστήριξης (set of support)

— η στρατηγική της διατεταγμένης ανάλυσης (ordered resolution)

**Στρατηγική συνόλου υποστήριξης**

Για να επιλυθεί ένα πρόβλημα q:

1. Θέσε L = {¬q}. Η L είναι η λίστα των τύπων που έχουν παραχθεί από την αρχική βάση γνώσης κάθε χρονική στιγμή.
2. Διάλεξε ένα p από την L. Αν p = T →F, τότε το q έχει επιλυθεί και, συνεπώς, σήμανε επιτυχία. Αλλιώς, αν μπορεί να εφαρμοσθεί ο κανόνας της ανάλυσης μεταξύ του p και ενός τύπου από την αρχική βάση γνώσης ή από την L, εφάρμοσέ τον και βάλε το αποτέλεσμα στην L. Αν δεν υπάρχει τύπος για να συνδυασθεί με το p μέσω του κανόνα της ανάλυσης, βγάλε το p από την L.
3. Αν η L είναι κενή, σήμανε αποτυχία. Αλλιώς πήγαινε στο βήμα 2.

- "Διάλεξε ένα p από την L": Η L μπορεί να είναι μία στοίβα (πρώτα–κατά–βάθος αναζήτηση) ή μία ουρά (πρώτα–κατά–πλάτος αναζήτηση).
- "....... αν μπορεί να εφαρμοσθεί ο κανόνας της ανάλυσης μεταξύ του p και ενός τύπου .......": Ποιο άτομο του p (κάποιο $d_j$ ή $a_i$) θα είναι οδηγός για την εφαρμογή του κανόνα της ανάλυσης. Ακόμα και για δεδομένο τέτοιο άτομο είναι δυνατόν να υπάρχουν διάφοροι υποψήφιοι τύποι για να συνδυασθούν με το p.
- Τι συμβαίνει στην περίπτωση ασυνεπών βάσεων γνώσης;

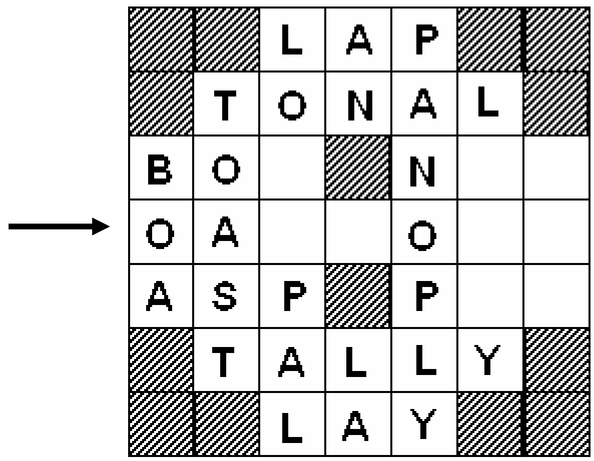
| | Αποδεικνυόμενοι τύποι | Αιτιολόγηση | Δεσμεύσεις |
|---|---|---|---|
| a | $work(x) \rightarrow F$ | άρνηση αποδεικτέου | |
| b | $big(h_5) \wedge house(h_5, p_5) \rightarrow F$ | από (a), (5) | $x = h_5$ |
| c | $big(house\text{-}of(p_3)) \rightarrow F$ | από (b), (3) | $h_5 = house\text{-}of(p_3)$, $p_5 = p_3$ |
| d | $house(house\text{-}of(p_3), p_4) \wedge rich(p_4) \rightarrow F$ | από (c), (4) | $h_4 = house\text{-}of(p_3)$ |
| e | $rich(p_3) \rightarrow F$ | από (d), (3) | $p_4 = p_3$ |
| f | $lawyer(p_3) \rightarrow F$ | από (e), (2) | $p_2 = p_3$ |
| g | $T \rightarrow F$ | από (f), (1) | $p_3 = John$ |

- Η L είναι μία στοίβα στην παραπάνω απόδειξη.
- Γιατί στο βήμα (e) συνδυάσαμε το (d) με το (3) και όχι το (d) με το (2); Για να αποδείξουμε ότι κάτι h είναι μεγάλο μέσω του κανόνα

  $$house(h, p) \wedge rich(p) \rightarrow big(h)$$

  είναι προτιμότερο να αποδείξουμε πρώτα το house(h, p) και μετά το rich(p) ή αντίστροφα;
- Η στρατηγική της διατεταγμένης ανάλυσης είναι μία εξειδίκευση της στρατηγικής του συνόλου υποστήριξης ως προς το ότι η εφαρμογή του κανόνα της ανάλυσης στο βήμα 2 προκαλεί την απαλοιφή του πρώτου κατά σειρά ατόμου στο p.
- Η στρατηγική της διατεταγμένης ανάλυσης εκμεταλλεύεται στατική πληροφορία ελέγχου συντακτικής καθαρή φύσης, που ορίζεται από αυτόν που διατυπώνει τη βάση γνώσης σε λογική πρώτης τάξης.

- Το πρόβλημα του καθορισμού της σειράς απαλοιφής των ατόμων σε τύπους, μέσω του κανόνα της ανάλυσης, είναι ανάλογο με τον καθορισμό της σειράς “γεμίσματος” των λέξεων σε ένα σταυρόλεξο.




........... – TONAL – ....... – TOAST – ....... – LAY – LAP – OA_ _O_ _

(α) Εναλλασσόμενη τοποθέτηση οριζόντιων και κάθετων λέξεων, αρχίζοντας από τις πιο επάνω και πιο αριστερά.

(β) Προκαθορισμός μίας τέτοιας σειράς που εξασφαλίζει ότι κάθε τοποθετούμενη λέξη τέμνεται με την αμέσως προηγούμενη λέξη που τοποθετήθηκε (εύρεση Hamilton μονοπατιού σε γράφο).

(γ) Τοποθέτηση κάθε φορά εκείνης της λέξης που έχει τις λιγότερες δυνατές επιλογές.

- Έλεγχος “στο στάδιο της μετάφρασης” (compile–time), όπως στα (α), (β), ή “στο στάδιο της εκτέλεσης” (run–time), όπως στο (γ), ή κάτι ανάμεσα;

- Καθορισμός της σειράς απαλοιφής των ατόμων κατά τη διαδικασία της ανάλυσης
  (α) Κατά την κατασκευή της βάσης γνώσης ("στάδιο της μετάφρασης")
  (β) Όταν δοθεί η ερώτηση που πρέπει να απαντηθεί (κάτι ανάμεσα στο "στάδιο της μετάφρασης" και στο "στάδιο της εκτέλεσης")
  (γ) Κατά τον υπολογισμό της απάντησης ("στάδιο της εκτέλεσης")

Παραδείγματα

(α) $\text{house}(h, p) \wedge \text{rich}(p) \rightarrow \text{big}(h)$

Πιθανές ερωτήσεις:

big(a–big–house)

big(a–small–house)

(β) $\text{related}(x, y) \wedge \text{loves}(x, y) \rightarrow \text{family–oriented}(x)$

Πιθανές ερωτήσεις:

family–oriented(John)

family–oriented(Mary)

όπου ο Γιάννης αγαπάει όλο σχεδόν τον κόσμο αλλά έχει λίγους συγγενείς, ενώ η Μαίρη που έχει πολλούς συγγενείς αγαπάει μόνο τη γάτα της.

(γ) $\text{rich}(x) \wedge \text{family–oriented}(x) \rightarrow \text{happy–kids}(x)$

Πιθανές ερωτήσεις:

happy–kids(x)

**Έλεγχος στο στάδιο της εκτέλεσης**

**Πρόβλεψη (Lookahead)**

$$\text{mother}(m, c) \wedge \text{lives–at}(m, h) \wedge \text{married}(c, s) \wedge \text{lives–at}(s, h) \rightarrow \text{sad}(s)$$

Ερώτηση: sad(s)

Έστω ότι είναι προκαθορισμένη η σειρά απαλοιφής των ατόμων του προηγούμενου τύπου και είναι αυτή που ορίζεται συντακτικά. Με ποιον τύπο από τους:

$$T \rightarrow \text{mother}(m_1, c_1)$$

$$T \rightarrow \text{mother}(m_2, c_2)$$

$$\vdots$$

θα πρέπει πρώτα να συνδυασθεί;

"Όταν δεσμεύεται μία μεταβλητή πρέπει αυτό να γίνεται έτσι ώστε τα υπόλοιπα άτομα που την περιέχουν να είναι επιλύσιμα"

Ακόμα καλύτερα:

"Όταν δεσμεύεται μία μεταβλητή πρέπει αυτό να γίνεται έτσι ώστε τα υπόλοιπα άτομα που την περιέχουν να έχουν όσο το δυνατόν περισσότερες λύσεις"

- Βασικό επίπεδο ⟷ Μετα–επίπεδο

- Υπάρχουν αναλογίες με το πρόβλημα του σταυρολέξου;

**Κανόνας "το–φθηνότερο–πρώτα" (cheapest–first–heuristic)**

Έστω ότι ο τύπος

mother(m, c) ∧ lives–at(m, h) ∧ married(c, s) ∧

lives–at(s, h) ⟶ sad(s)

συνδυάζεται κατά σειρά με τους τύπους:

T ⟶ mother(Jane, Chris)

T ⟶ lives–at(Jane, house–33)

Ποια είναι η καλύτερη σειρά επίλυσης των

married(Chris, s) και lives–at(s, house–33);

Ο ευριστικός κανόνας "το–φθηνότερο–πρώτα" συνιστά:
"Είναι συνήθως καλύτερο να επιλύεται πρώτα το άτομο εκείνο που έχει φθηνότερο κόστος απαρίθμησης λύσεων, επειδή αυτές είναι λιγότερες" (βλ. σταυρόλεξο – (γ)).

**Παράδειγμα επιτυχούς εφαρμογής**

Να βρεθεί ένας δημόσιος υπάλληλος που ο πατέρας του είναι βουλευτής:

public_sector_employee(x) ∧ father(y, x) ∧

member_of_parliament(y) ≡

e ∧ f ∧ p

| Άτομο | Αριθμός λύσεων |
|---|---|
| public_sector_employee(x) | 500 000 |
| member_of_parliament(y) | 300 |
| father(y, x) | 10 000 000 |
| father(y, const) | 1 |
| father(const, y) | ~ 2 |

Τι συνιστά ο ευριστικός κανόνας μας;

**$e \wedge f \wedge p$ : 500 000 · 1 · 1 = 500 000**

**$e \wedge p \wedge f$ : 500 000 · 300 · 1 = 150 000 000**

**$f \wedge e \wedge p$ : 10 000 000 · 1 · 1 = 10 000 000**

**$f \wedge p \wedge e$ : 10 000 000 · 1 · 1 = 10 000 000**

**$p \wedge e \wedge f$ : 300 · 500 000 · 1 = 150 000 000**

**$p \wedge f \wedge e$ : 300 · 2 · 1 = 600** ← **ΑΥΤΟ** ☺

**Παράδειγμα ανεπιτυχούς εφαρμογής**

north_province(x) ∧ member_of_parliament(y) ∧ from(y, x) ≡

$n \wedge p \wedge f$

| Άτομο | Αριθμός λύσεων |
|---|---|
| north_province(x) | 15 |
| member_of_parliament(y) | 300 |
| from(y, x) | 10 000 000 |
| from(y, const) | 200 000 |
| from(const, x) | 1 |

Τι συνιστά ο ευριστικός κανόνας μας;

**$n \wedge p \wedge f$ : 15 · 300 · 1 = 4 500** ← **ΑΥΤΟ** ☹

**$n \wedge f \wedge p$ : 15 · 200 000 · 1 = 3 000 000**

**$p \wedge n \wedge f$ : 300 · 15 · 1 = 4 500**

**$p \wedge f \wedge n$ : 300 · 1 · 1 = 300** **!!!**

**$f \wedge n \wedge p$ : 10 000 000 · 1 · 1 = 10 000 000**

**$f \wedge p \wedge n$ : 10 000 000 · 1 · 1 = 10 000 000**

Αλλά αν για n άτομα ελέγξουμε όλες τις δυνατές διατάξεις (n!), τότε .......

**Οπισθοδρόμηση βασισμένη στην εξάρτηση (Dependency–directed backtracking) και "άλμα προς τα πίσω" (backjumping)**

- Ο στόχος της οπισθοδρόμησης της βασισμένης στην εξάρτηση είναι σε περίπτωση αδιεξόδου να μην γίνει αναίρεση της χρονολογικά τελευταίας επιλογής αλλά εκείνης που πραγματικά προκάλεσε το πρόβλημα.

- Το κόστος της οπισθοδρόμησης της βασισμένης στην εξάρτηση είναι πολύ μεγάλο (πολλή προσπάθεια μετα–επιπέδου) με ίσως δυσανάλογα μικρό όφελος.

- Το "άλμα προς τα πίσω" είναι μια πιο καθαρή τεχνική από την την οπισθοδρόμηση τη βασισμένη στην εξάρτηση και, συνεπώς, ευκολότερη στην υλοποίηση, κατά την οποία αναιρείται εκείνη η επιλογή που "φαίνεται" να έχει σημαντικές πιθανότητες να έχει προκαλέσει το πρόβλημα.

α) Στο παράδειγμα του σταυρολέξου ποια ενέργεια προτείνουν οι δύο τεχνικές μετά την αδυναμία εύρεσης λέξης της μορφής ΟΑ_ _Ο_ _ **;**

β) Με δεδομένη τη γνώση

$$\text{parent}(p, c) \wedge \text{successful}(c, j) \wedge \text{married}(c, s) \rightarrow \text{proud}(p)$$

και το ότι ο Γιάννης και η Ελένη έχουν κόρη την ανύπαντρη επιτυχημένη χειρούργο Μαίρη, πώς συμπεριφέρονται οι δύο τεχνικές στην προσπάθεια εύρεσης κάποιου που να είναι υπερήφανος;

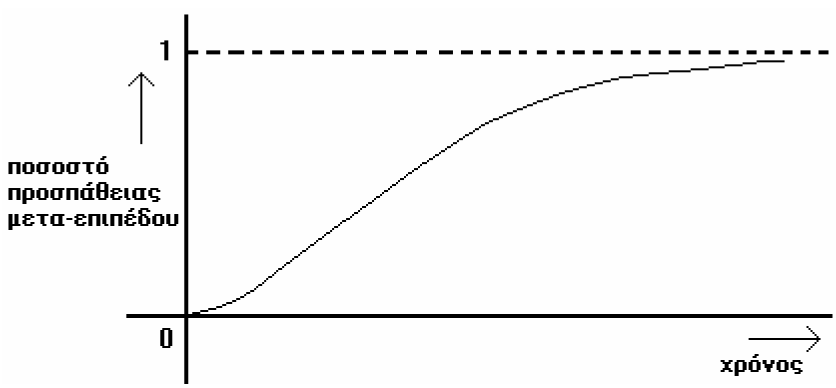
Το βασικό ερώτημα στην Τεχνητή Νοημοσύνη:

"Ποια είναι η σχέση της προσπάθειας βασικού επιπέδου με την προσπάθεια μετα–επιπέδου;"

- Οι δύο ακραίες λύσεις, δηλαδή μόνο προσπάθεια βασικού επιπέδου ή μόνο προσπάθεια μετα–επιπέδου, είναι εξίσου άσχημες.

- Σε περίπτωση αδυναμίας απόφασης για την επιλογή μίας από δύο υποψήφιες στρατηγικές επίλυσης ενός προβλήματος, μία καλή λύση είναι η διαδοχική εναλλαγή από τη μία στην άλλη, οπότε είναι εγγυημένο ότι το συνολικό χρονικό κόστος δεν θα είναι μεγαλύτερο από το διπλάσιο της καλύτερης από τις δύο στρατηγικές.

- Οι άνθρωποι εργάζονται αρχικά βάζοντας πολλή προσπάθεια βασικού επιπέδου και αν συνειδητοποιήσουν ότι το πρόβλημά τους είναι δύσκολο αρχίζουν σιγά–σιγά να προσπαθούν και στο μετα–επίπεδο.



## ΣΥΝΤΗΡΗΣΗ ΤΗΣ ΑΛΗΘΕΙΑΣ ΒΑΣΙΣΜΕΝΗ ΣΕ ΥΠΟΘΕΣΕΙΣ (ASSUMPTION–BASED TRUTH MAINTENANCE)

- Οι όροι “Συντήρηση της αλήθειας βασισμένη σε υποθέσεις” καθώς και “Συστήματα συντήρησης της αλήθειας βασισμένης σε υποθέσεις” (Assumption-based truth maintenance systems), ή σε συντομογραφία ΣΣΑΥ (ATMS), δεν έχουν σήμερα μεγάλη σχέση με το περιεχόμενό τους. Έχουν διατηρηθεί όμως για ιστορικούς λόγους.
- Σε πολλές περιπτώσεις υπάρχει ανάγκη δικαιολόγησης της ισχύος ενός συμπεράσματος με δεδομένη μια βάση γνώσης. Για παράδειγμα:

  Γιατί το σπίτι του Γιάννη χρειάζεται πολλή δουλειά για να συντηρηθεί;

  — Επειδή ο Γιάννης είναι δικηγόρος.

  — Επειδή ο Γιάννης είναι δικηγόρος και τα μεγάλα σπίτια χρειάζονται πολλή δουλειά για να συντηρηθούν.

  — Επειδή ο Γιάννης είναι δικηγόρος, οι δικηγόροι είναι πλούσιοι και τα μεγάλα σπίτια χρειάζονται πολλή δουλειά για να συντηρηθούν.

  — Επειδή ο Γιάννης είναι δικηγόρος, οι δικηγόροι είναι πλούσιοι, οι πλούσιοι έχουν μεγάλα σπίτια και τα μεγάλα σπίτια χρειάζονται πολλή δουλειά για να συντηρηθούν.

- Η εξήγηση/δικαιολόγηση της ισχύος ενός συμπεράσματος που προκύπτει από κάποια βάση γνώσης (π.χ. με εφαρμογή της ανάλυσης) συνίσταται στην παράθεση, χωρίς συγκεκριμένη σειρά, των τύπων (όλων ή ορισμένων) της βάσης που χρησιμοποιήθηκαν για την απόδειξη του συμπεράσματος.

Παράδειγμα:

Από τη γνωστή βάση γνώσης:

(1) $T \rightarrow$ lawyer(John)

(2) lawyer($p_2$) $\rightarrow$ rich($p_2$)

(3) $T \rightarrow$ house(house-of($p_3$), $p_3$)

(4) house($h_4$, $p_4$) $\wedge$ rich($p_4$) $\rightarrow$ big($h_4$)

(5) big($h_5$) $\wedge$ house($h_5$, $p_5$) $\rightarrow$ work($h_5$)

αποδεικνύεται το $(\exists x)(\text{work}(x))$ ή

(6) work(x) $\rightarrow$ F

με τη διαδικασία ανάλυσης:

| | | |
|---|---|---|
| (7) | rich($p_3$) $\rightarrow$ big(house-of($p_3$)) | (3)+(4) |
| (8) | lawyer($p_3$) $\rightarrow$ big(house-of($p_3$)) | (2)+(7) |
| (9) | $T \rightarrow$ big(house-of(John)) | (1)+(8) |
| (10) | house(house-of(John), $p_5$) $\rightarrow$work(house-of(John)) | (5)+(9) |
| (11) | $T \rightarrow$ work(house-of(John)) | (3)+(10) |
| (12) | $T \rightarrow F$ | (6)+(11) |

Οι υποθέσεις (τύποι της αρχικής βάσης γνώσης) με βάση τις οποίες αποδείχθηκε το συμπέρασμα (ή συντηρήθηκε η αλήθεια;) είναι οι (1), (2), (3), (4), (5). Τμήμα από αυτές (ή ίσως και όλες) μπορεί να δοθεί σαν εξήγηση της ισχύος του $(\exists x)(\text{work}(x))$

**Τροποποίηση της στρατηγικής συνόλου υποστήριξης για την παραγωγή των εξηγήσεων**

Για να επιλυθεί ένα πρόβλημα q:

1. Θέσε $L = \{<\neg q, \emptyset >\}$. Η L είναι η λίστα των τύπων που έχουν παραχθεί από την αρχική βάση γνώσης κάθε στιγμή. Κάθε τύπος συνοδεύεται και από τους τύπους της βάσης γνώσης που χρησιμοποιήθηκαν για την παραγωγή του.

2. Διάλεξε ένα $<p, S>$ από την L (το p είναι ένας τύπος που έχει παραχθεί και το S είναι η εξήγηση της παραγωγής). Αν $p = T \rightarrow F$, τότε το q έχει επιλυθεί και, συνεπώς, σήμανε επιτυχία και δώσε σαν αποτέλεσμα το S. Αλλιώς, αν μπορεί να εφαρμοστεί ο κανόνας της ανάλυσης μεταξύ του p και ενός τύπου d από τη βάση γνώσης και αυτή η εφαρμογή έχει σαν αποτέλεσμα το r, πρόσθεσε το $< r, S \cup \{d\}>$ στην L. Εναλλακτικά, αν η L περιέχει ένα στοιχείο $<p', S'>$ τέτοιο ώστε αν εφαρμοσθεί ο κανόνας της ανάλυσης μεταξύ των p και p΄ να προκύπτει το r, πρόσθεσε το $< r, S \cup S'>$ στην L. Αν δεν υπάρχουν ούτε στη βάση ούτε στην L τύποι για να συνδυασθούν με το p, τότε βγάλε το από την L.

3. Αν η L είναι κενή, σήμανε αποτυχία. Αλλιώς πήγαινε στο βήμα 2.

- Αποδεικνύεται ότι όταν η τροποποιημένη στρατηγική συνόλου υποστήριξης σε ένα πρόβλημα/ερώτηση q επιστρέψει το S, τότε το S είναι ένα υποσύνολο της βάσης γνώσης τέτοιο ώστε $S \models q$.

  Π.χ. Αν $q = (\exists x)(big(x))$

  τότε $S = \{(1), (2), (3), (4)\}$

- Συνήθως, δεν έχει νόημα για την εξήγηση ενός συμπεράσματος να παρατίθενται όλοι οι τύποι από τη βάση γνώσης που χρησιμοποιήθηκαν για την απόδειξή του, αλλά μόνο εκείνοι που δεν θεωρούνται "κοινή γνώση" (common knowledge). Π.χ. είναι κοινή γνώση ότι "οι πλούσιοι έχουν μεγάλα σπίτια".

**Εξήγηση (explanation) ενός p**

Αν μια βάση γνώσης D αποτελείται από την κοινή γνώση C και τις υποθέσεις (assumptions) A, δηλαδή $D = C \cup A$, τότε η εξήγηση ενός p είναι ένα ελάχιστο υποσύνολο E του A, δηλαδή $E \subseteq A$, τέτοιο ώστε $E \cup C \models p$.

- Ερωτήματα:
  - Πως διαχωρίζεται η βάση D σε C και A;
  - Τι είναι "ελάχιστο υποσύνολο";
  - Πόσες εξηγήσεις E μπορεί να υπάρχουν για ένα p;
  - Μπορεί το E να είναι ασυνεπές με το C;
- Ένα υποσύνολο του A που είναι ασυνεπές με το C ονομάζεται μη-χρήσιμο (nogood). Οι εξηγήσεις E δεν πρέπει να είναι μη-χρήσιμες.

**ATMS**

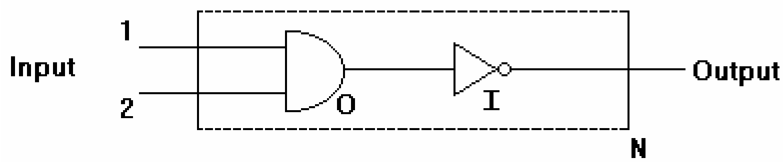
Ένα σύστημα συντήρησης της αλήθειας βασισμένης σε υποθέσεις (ATMS) με δεδομένη κάποια βάση γνώσης $D = C \cup A$ και μια ερώτηση q επιστρέφει όλες (ή μερικές) τις εξηγήσεις του q.

**Εφαρμογές των ATMS**

- Δικαιολόγηση ισχύος λογικών συμπερασμάτων
- Κατάστρωση σχεδίου (Planning)
  - Πρόβλημα σύνθεσης/δημιουργίας/κατασκευής
  - Κατασκευή της ακολουθίας των κατάλληλων ενεργειών (actions) που οδηγούν στην επίτευξη ενός στόχου (goal)
  - A = όλες οι πιθανές ενέργειες<br>C = κανόνες σύνθεσης ενεργειών<br>E = σχέδιο (plan)<br>p = στόχος
  - Παράδειγμα:<br>Με δεδομένα το σύνολο όλων των αεροπορικών δρομολογίων και το σύνολο των κανόνων συνδυασμού τους (π.χ. κάποιος μπορεί μετά από μία πτήση x να πάρει μια y, αν η x καταλήγει εκεί από όπου ξενικά η y ‘‘επαρκή’’ χρόνο πριν από την αναχώρηση της y), είναι δυνατόν, αν τη Δευτέρα βρίσκομαι στην Αθήνα, να καταστρώσω σχέδιο για δύο 48ωρες επισκέψεις σε Θεσσαλονίκη και Ηράκλειο, έτσι ώστε την Παρασκευή να είμαι στην Αθήνα;

- Σχεδίαση (Design)
  - Πρόβλημα σύνθεσης/δημιουργίας/κατασκευής
  - Κατασκευή σύνθετου αντικειμένου μέσω του συνδυασμού στοιχειωδών δομικών λίθων
  - A = όλοι οι πιθανοί δομικοί λίθοι και οι πιθανές συνδέσεις τους
  - C = κανόνες σύνδεσης δομικών λίθων
  - E = σύνθετο αντικείμενο προς κατασκευή
  - p = ιδιότητες του σύνθετου αντικειμένου
  - Παράδειγμα:

    Κατασκευή σύνθετων ψηφιακών κυκλωμάτων από στοιχειώδεις πύλες. Πώς μπορεί να κατασκευασθεί ένα κύκλωμα (πύλη) NOR 2 εισόδων από πύλες AND, OR και NOT;




A = {T $\rightarrow$ type(I, inverter),

T $\rightarrow$ type(O, or-gate),

T $\rightarrow$ connected(input(x, N), input(x, O)),

T $\rightarrow$ connected(output(O), input(I)),

T $\rightarrow$ connected(output(I), output(N)),

..................................................}

C = {type(y, or-gate) $\wedge$ value(input(x, y), 1) $\rightarrow$

value(output(y), 1),

connected(x, y) $\wedge$ value(x, v) $\rightarrow$ value(y, v),

............................................................}

$$p = (value(input(x , N), 1) \rightarrow value(output(N), 0)$$
$$\wedge (value(input(1, N), 0) \wedge value(input(2, N), 0)$$
$$\rightarrow value(output(N), 1))$$

- Διάγνωση (Diagnosis)
  - Εύρεση του αιτίου (βλάβης) που είναι υπεύθυνο για ένα σύμπτωμα κακής λειτουργίας.
  - Απαιτείται η γνώση του μοντέλου καλής λειτουργίας.
  - Παράδειγμα:

    Λειτουργία ενός αυτοκινήτου

$$\boxed{\begin{array}{l} \text{battery-ok} \wedge \text{bulbs-ok} \rightarrow \text{headlights-work} \\ \text{battery-ok} \wedge \text{starter-ok} \rightarrow \text{engine-starts} \end{array}}$$

Σε μία κατάσταση όπου ισχύουν τα

battery-ok

bulbs-ok

starter-ok

εύκολα αποδεικνύονται και τα

$$\text{headlights-work} \tag{1}$$

$$\text{engine-starts} \tag{2}$$

αλλά και το

$$\text{headlights-work} \vee \text{engine-starts} \tag{3}$$

Το (1) έχει μία εξήγηση, την:

$$\{\text{battery-ok, bulbs-ok}\} \equiv \text{battery-ok} \wedge \text{bulbs-ok} \equiv E_1$$

Το (2) έχει μία εξήγηση, την:

$$\{\text{battery-ok, starter-ok}\} \equiv \text{battery-ok} \wedge \text{starter-ok} \equiv E_2$$

Το (3) έχει δύο εξηγήσεις, τις:

{battery-ok, bulbs-ok}, {battery-ok, starter-ok} $\equiv$

{battery-ok $\wedge$ bulbs-ok} $\vee$ {battery-ok $\wedge$ starter-ok} $\equiv E_3$

Η διάγνωση/εύρεση της βλάβης που προκαλεί ένα σύμπτωμα κακής λειτουργίας f ταυτίζεται με την άρνηση της εξήγησης της κανονικής λειτουργίας $\neg$f.

Η διάγνωση για το $\neg$headlights-work είναι η

$\neg E_1 = \neg$battery-ok $\vee$ $\neg$bulbs-ok

Η διάγνωση για το $\neg$engine-starts είναι η

$\neg E_2 = \neg$battery-ok $\vee$ $\neg$starter-ok

Η διάγνωση για το $\neg$headlights-work $\wedge\neg$engine-starts είναι η

$\neg E_3 = \neg$(battery-ok $\wedge$ bulbs-ok) $\wedge$

$\neg$(battery-ok $\wedge$ starter-ok) =

= $\neg$battery-ok $\vee$ ($\neg$bulbs-ok $\wedge\neg$starter-ok)

– Διάγνωση μέσω ATMS

components-ok $\rightarrow$ behavior-ok

– Τυπική διάγνωση

components-bad $\rightarrow$ behavior-bad

Ποια διάγνωση είναι ορθότερη και γιατί;

Ποια διάγνωση υποστηρίζεται ευκολότερα και γιατί;

Δώστε περιοχές εφαρμογής τους.

- Διατήρηση συνέπειας σε βάσεις γνώσης
  - Αν μία βάση γνώσης D είναι συνεπής, η προσθήκη σε αυτήν ενός τύπου p μπορεί να διαταράξει τη συνέπειά της. Π.χ.

    $$D = \{color(walls, plaid), \quad color(x, plaid) \rightarrow \neg color(x, white)\}$$

    Μπορούμε να προσθέσουμε στη D το
    p = color(walls, white);

  - Για να εξακολουθήσει να παραμένει μία βάση γνώσης D συνεπής και μετά την προσθήκη ενός p σ' αυτήν, μπορούμε να:
    1. Χρησιμοποιήσουμε ένα ATMS για να βρούμε όλες τις εξηγήσεις $e_1, e_2, ...., e_k$ του $\neg p$ από την D.
    2. Βρούμε ένα ελάχιστο σύνολο H τέτοιο ώστε $H \cap e_i \neq \varnothing$ για κάθε $e_i$.
    3. Διαγράψουμε τα στοιχεία του H, που ονομάζεται ελάχιστο σύνολο κρούσης (minimal hitting set), από την D.

  - Αν $D = C \cup A$, τα στοιχεία του C είναι "προστατευμένα" από τυχόν προσθήκες στην D, αφού δεν περιέχονται σε εξηγήσεις.

  - Είναι δυνατόν να υπάρχουν περισσότερα του ενός ελάχιστα σύνολα κρούσης για μία προσθήκη, αλλά κανένα δεν μπορεί να είναι υποσύνολο κάποιου άλλου.

– Παράδειγμα 1:

A = {color(walls, plaid)}

C = {color(x, plaid) $\rightarrow \neg$color(x, white)}

p = color(walls, white)

H = {color(walls, plaid)}

– Παράδειγμα 2:

A = {color(walls, plaid),

$\neg$pretty(room)}

C = {color(x, plaid) $\rightarrow \neg$color(x, white),

color(walls, white) $\rightarrow$ pretty(room)}

p = color(walls, white)

H = {color(walls, plaid),

$\neg$pretty(room)}

– Παράδειγμα 3:

A = {color(walls, plaid),

$\neg$pretty(room)}

C = {color(x, plaid) $\rightarrow \neg$color(x, white)}

p = color(walls, white)

H = {color(walls, plaid)}

– Παράδειγμα 4:

A = {color(walls, plaid)}

C = {color(x, plaid) $\rightarrow \neg$color(x, white),

color(walls, plaid) $\rightarrow \neg$pretty(room)}

p = color(walls, white)

H = {color(walls, plaid)}

Μετά την προσθήκη του p, τι συμβαίνει με την ισχύ του $\neg$pretty(room);

**Υλοποίηση ενός ATMS**

Αν $D = C \cup A$ είναι μια βάση γνώσης, τότε το σύνολο X των εξηγήσεων ενός q κατασκευάζεται ως εξής:

1. Θέσε $X = \emptyset$.
2. Χρησιμοποίησε την “τροποποιημένη στρατηγική συνόλου υποστήριξης για την παραγωγή εξηγήσεων” για να βρεις ένα $U \subseteq D$ τέτοιο ώστε $U \models q$. Αν δεν υπάρχει τέτοιο U, να επιστρέψεις το X. Αλλιώς, θέσε $V = U \cap A$.
3. Αν το V δεν είναι ασυνεπές και δεν είναι υπερσύνολο κάποιου στοιχείου του X, τότε:
   α) Βγάλε από το X όλα τα στοιχεία του που είναι υπερσύνολα του V.
   β) Βάλε το V στο X.
4. Πήγαινε στο βήμα 2.

- Αποδεικνύεται ότι η προηγούμενη μέθοδος είναι σωστή.
- Ερωτήματα:
  - Πώς στο βήμα 2 θα βρίσκουμε κάθε φορά και μια νέα εξήγηση U;
  - Χρειάζεται ή είναι δυνατόν να βρούμε όλες τις εξηγήσεις ενός q;
  - Μπορεί να γίνει πιο αποδοτική η απόρριψη των ασυνεπών εξηγήσεων και των εξηγήσεων που είναι υπερσύνολα άλλων ήδη κατασκευασμένων εξηγήσεων;

# ΜΗ ΜΟΝΟΤΟΝΗ ΣΥΛΛΟΓΙΣΤΙΚΗ
# (NONMONOTONIC REASONING)

- Ένα βασικό ερώτημα είναι το εξής:

  Τι γίνεται όταν με δεδομένη κάποια πληροφορία έχουμε βγάλει ένα συμπέρασμα και κάποια στιγμή αργότερα σταματήσουμε να πιστεύουμε ότι η πληροφορία είναι σωστή ή, ακόμα περισσότερο, μάθουμε ότι ισχύει το αντίθετό της;

  Ή μέσω ρητών:

  – “Nothing is certain but death and taxes”

  Benjamin Franklin

  – “Τα πάντα ρει”

  Ηράκλειτος

- Μη μονότονη συλλογιστική: Το τι ξέρουμε για τον κόσμο δεν αυξάνει, κατ’ ανάγκη, όσο αυξάνουν οι βασικές πληροφορίες που χρησιμοποιούμε για να βγάλουμε συμπεράσματα.

- Μη μονότονη συλλογιστική = Ένας ηττοπαθής (defeasible) τρόπος εξαγωγής συμπερασμάτων.

- Η περισσότερο διαδεδομένη εκδοχή της μη μονότονης συλλογιστικής είναι η συνήθης ή “ελλείψει” συλλογιστική (default reasoning).

- Εφαρμογή της μη μονότονης συλλογιστικής στην περίπτωση κληρονομούμενων ιδιοτήτων

$$\left.\begin{array}{l} bird(x) \rightarrow flies(x) \\ bird(Tweety) \end{array}\right\} \Rightarrow \quad flies(Tweety)$$

$bird(x) \wedge normal(x) \rightarrow flies(x)$

ή

$bird(x) \wedge \neg ab(x) \rightarrow flies(x)$

$\neg ab(Tweety)$:  Ισχύει ή όχι;

---

Έστω επίσης ότι:

$ostrich(x) \rightarrow bird(x) \wedge \neg flies(x)$

$ostrich(Fred)$

$$\left\{\begin{array}{l} ostrich(x) \rightarrow bird(x) \\ ostrich(x) \rightarrow \neg flies(x) \end{array}\right\}$$

$ostrich(x) \rightarrow ab(x)$

---

Τελικά, δύο “ελλείψει” κανόνες είναι:

$bird(x) \wedge \neg ab_b(x) \rightarrow flies(x)$

$ostrich(x) \wedge \neg ab_o(x) \rightarrow \neg flies(x)$

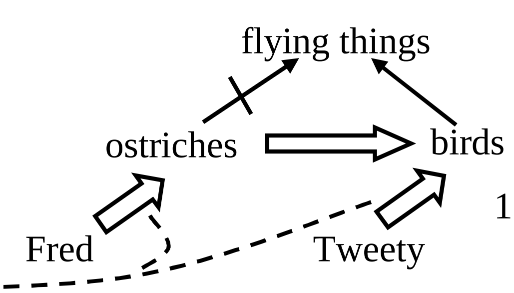
Θεωρώντας και τη γνώση:

$ostrich(x) \rightarrow bird(x)$

$bird(Tweety)$

$ostrich(Fred)$

Έχουμε το εξής:



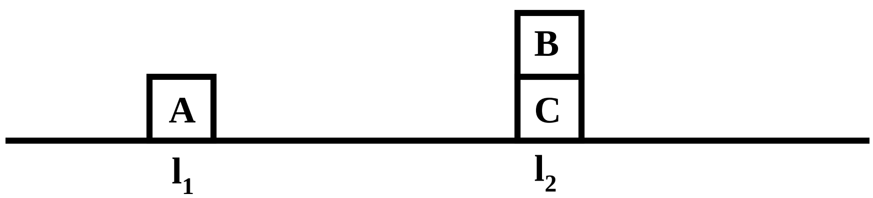

Είναι σωστά;

- Το πρόβλημα “πλαίσιο” (frame problem)

  “Πώς παριστάνεται το γεγονός ότι εν γένει τα πράγματα δεν τείνουν να αλλάζουν κατάσταση, εκτός αν υπάρχει συγκεκριμένη πληροφορία για το αντίθετο;”

Η μη μονότονη συλλογιστική είναι χρήσιμη στην αντιμετώπιση του προβλήματος “πλαίσιο”.

Παράδειγμα: Έστω ότι στον κόσμο των κύβων η κατάσταση είναι η εξής:




Εάν εκπονήσουμε ένα σχέδιο μεταφοράς του A επάνω στο C (σεβόμενοι βέβαια τους κανόνες του κόσμου των κύβων), μετά την εκτέλεση αυτού του σχεδίου το C θα εξακολουθήσει να βρίσκεται στη θέση $l_2$;

**ΑΠΑΝΤΗΣΗ:** Γιατί όχι;

Άλλο ένα πρόβλημα: Σ’ ένα δωμάτιο, κάτω από ένα παράθυρο βρίσκεται ένα τραπέζι με ένα κόκκινο βάζο επάνω του. Μετακινούμε τώρα το τραπέζι στο κέντρο του δωματίου. Πού βρίσκεται το παράθυρο, που βρίσκεται το βάζο και τι χρώμα έχει το βάζο;

**Θεωρητική πλευρά της μη μονοτονίας**

T : το σύνολο της “βεβαίας” γνώσης

A : το σύνολο της γνώσης “υπό αμφισβήτηση”

(T, A) : μη μονότονη θεωρία (nonmonotonic theory)

Ειδικότερα, ένα (T, A) είναι μία “ελλείψει” θεωρία (default theory) όταν το T περιλαμβάνει μεταξύ άλλων και “ελλείψει” κανόνες (όπως, για παράδειγμα, τον $bird(x) \land \neg ab_b(x) \rightarrow flies(x)$) και το A αποτελείται από γνώση του είδους: $\neg ab_b(Tweety), \neg ab_o(Fred)$ κλπ.

Στη συνέχεια, θα αναφερόμαστε σε “ελλείψει” θεωρίες (T, A), επειδή τα παραδείγματά μας θα είναι από κόσμους με “ελλείψει” γνώση, αλλά οι ορισμοί και τα συμπεράσματα θα ισχύουν για κάθε μη μονότονη θεωρία.

Μία επέκταση (extension) μίας “ελλείψει” θεωρίας (T, A) είναι κάθε μέγιστο υποσύνολο του A που είναι συνεπές με το T.

**ΕΡΩΤΗΣΗ:** Τι είναι μέγιστο υποσύνολο;

Παράδειγμα:

$T = \{bird(x) \land \neg ab(x) \rightarrow flies(x),$

$\quad ostrich(x) \rightarrow bird(x) \land \neg flies(x),$

$\quad bird(Tweety),$

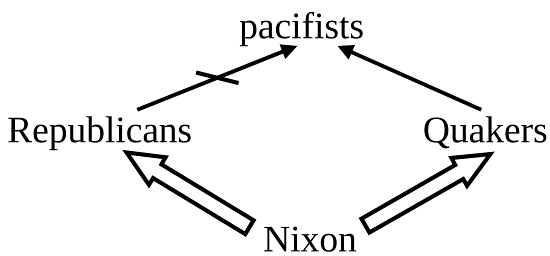
$\quad ostrich(Fred)\}$

$A = \{\neg ab(Tweety), \neg ab(Fred)\}$

επέκταση $E = \{\neg ab(Tweety)\}$

- Είναι δυνατόν σε μια “ελλείψει” θεωρία να έχουμε περισσότερες από μία επεκτάσεις; Βεβαίως.

Το διαμάντι του Νίξον:




Quaker(x) $\wedge \neg$ $ab_r$(x) $\rightarrow$ pacifist(x)

Republican(x) $\wedge \neg$ $ab_p$(x) $\rightarrow \neg$pacifist(x)

Quaker(Nixon)

Republican(Nixon)

A = {$\neg ab_r$(Nixon), $\neg ab_p$(Nixon)}

E1 = {$\neg ab_p$(Nixon)}

E2 = {$\neg ab_r$(Nixon)}

Αν συγχωνεύαμε και τη γνώση του προηγούμενου παραδείγματος, τότε:

E1 = {$\neg ab_p$(Nixon), $\neg$ab(Tweety)}

E2 = {$\neg ab_r$(Nixon), $\neg$ab(Tweety)}

Παρατηρήσεις:

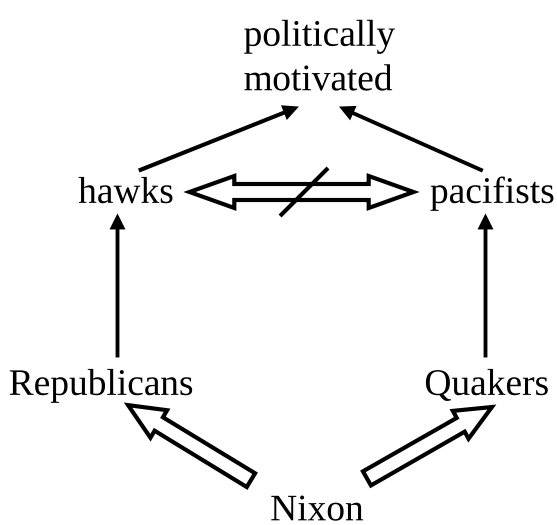
1) Τόσο με βάση το T $\cup$ E1 όσο και με βάση το T $\cup$ E2, αποδεικνύεται εύκολα το flies(Tweety).

2) Αλλά το pacifist(Nixon) μπορεί να αποδειχθεί μόνο από το T $\cup$ E2 και όχι από το T $\cup$ E1.

- Έστω μία “ελλείψει” θεωρία (T, A). Κάποιο p λέγεται ότι είναι συνετή συνέπεια (cautious consequence) της θεωρίας, όταν για κάθε επέκταση E ισχύει $T \cup E \models p$. Το p λέγεται ενθουσιώδης συνέπεια (brave consequence) της θεωρίας, όταν για κάποια επέκταση E ισχύει $T \cup E \models p$.

- Δηλαδή, για τις συνετές συνέπειες υπάρχουν μόνο υπέρ επιχειρήματα, ενώ για τις ενθουσιώδεις υπάρχουν και υπέρ και κατά επιχειρήματα (εφ’ όσον δεν είναι ταυτόχρονα και συνετές).

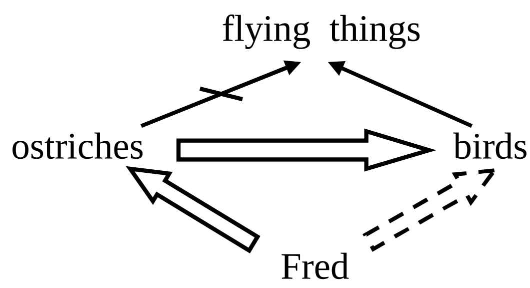
Παράδειγμα:




Αποδεικνύεται εύκολα ότι το ότι “ο Νίξον έχει πολιτικά κίνητρα” είναι μία συνετή συνέπεια της θεωρίας.

- Βασικό πρόβλημα:

  Με ποια κριτήρια μπορεί κάποιος να υιοθετήσει μία επέκταση αντί για κάποια άλλη;




$bird(x) \wedge \neg ab_b(x) \rightarrow flies(x)$

$ostrich(x) \rightarrow bird(x)$

$ostrich(x) \wedge \neg ab_o(x) \rightarrow \neg flies(x)$

$ostrich(Fred)$

$A = \{\neg ab_b(Fred), \neg ab_o(Fred)\}$

$E1 = \{\neg ab_b(Fred)\}$

$E2 = \{\neg ab_o(Fred)\}$

Τόσο το flies(Fred) όσο και το $\neg$flies(Fred) είναι ενθουσιώδεις συνέπειες της θεωρίας, γιατί το πρώτο αποδεικνύεται από το $T \cup E1$ και το δεύτερο από το $T \cup E2$.

Και στο παράδειγμα του “διαμαντιού του Νίξον”, τόσο το pacifist(Nixon) όσο και το $\neg$pacifist(Nixon) είναι ενθουσιώδεις συνέπειες της θεωρίας.

Τελικά τι ισχύουν; Το flies(Fred) ή το $\neg$flies(Fred);

Το pacifist(Nixon) ή το $\neg$pacifist(Nixon);

Η ύπαρξη του $\Longrightarrow$ μεταξύ των ostriches και birds μας επιτρέπει τελικά να αποφασίσουμε ότι ¬flies(Fred) αντί για flies(Fred), ενώ δεν έχουμε κανένα λόγο να υιοθετήσουμε κάποιο από τα pacifist(Nixon) ή ¬pacifist(Nixon).

Ουσιαστικά, για την επιλογή του ¬flies(Fred) αντί για το flies(Fred) εφαρμόσαμε έναν κανόνα: “Μεταξύ αντιφατικών κληρονομούμενων ιδιοτήτων σ’ ένα αντικείμενο που προέρχονται από σχετιζόμενες ιεραρχικά κλάσεις στις οποίες ανήκει το αντικείμενο, προτιμάται εκείνη που προκύπτει από την ειδικότερη κλάση”.

- Το υπολογιστικό πρόβλημα:
  Πώς υπολογίζουμε την απάντηση σε μία ερώτηση q δεδομένης κάποιας “ελλείψει” θεωρίας (T, A);

  Η q ισχύει σε κάποια επέκταση της (T, A) αν και μόνον αν έχει κάποια εξήγηση (θυμηθείτε τα ATMS) με δεδομένα από το A.

  Το παραπάνω, που αποδεικνύεται σχετικά εύκολα, θεμελιώνει μια στενή σχέση που υπάρχει ανάμεσα στα συστήματα συντήρησης της αλήθειας και στη μη μονότονη συλλογιστική.

  Η στενή αυτή σχέση επιβεβαιώνεται και από την εφαρμογή της μη μονότονης συλλογιστικής σε προβλήματα διάγνωσης όπως αυτό που ακολουθεί.

## Εφαρμογή της μη μονότονης συλλογιστικής για διάγνωση

$T_0 \; = \; \{$battery-ok$\wedge$bulbs-ok $\rightarrow$ headlights-work,
battery-ok$\wedge$starter-ok $\rightarrow$ engine-starts$\}$

A = {battery-ok, bulbs-ok, starter-ok}

- Αν T = {headlights-work, engine-starts} $\cup$ $T_0$
  τότε E = {battery-ok, bulbs-ok, starter-ok}
  και A – E = { }

- Αν T = {$\neg$headlights-work} $\cup$ $T_0$
  τότε $E_1$ = {bulbs-ok, starter-ok},
  $E_2$ = {battery-ok, starter-ok}
  και A – $E_1$ = {battery-ok}
  A – $E_2$ = {bulbs-ok}

- Αν T = {$\neg$engine-starts} $\cup$ $T_0$
  τότε $E_1$ = {starter-ok, bulbs-ok},
  $E_2$ = {battery-ok, bulbs-ok}
  και A – $E_1$ = {battery-ok},
  A – $E_2$ = {starter-ok}

- Αν T = {$\neg$headlights-work, $\neg$engine-starts} $\cup$ $T_0$
  τότε $E_1$ = {bulbs-ok, starter-ok},
  $E_2$ = {battery-ok}
  και A – $E_1$ = {battery-ok}
  A – $E_2$ = {bulbs-ok, starter-ok}

Η διάγνωση της βλάβης που προκαλεί ένα σύμπτωμα κακής λειτουργίας f, βρίσκεται με τη βοήθεια μη μονότονης συλλογιστικής ως εξής:

1. Θεωρούμε τη μη μονότονη θεωρία (T, A), όπου T είναι το μοντέλο λειτουργίας επαυξημένο με το σύμπτωμα κακής λειτουργίας f. Το A αποτελείται από τη γνώση που επιβάλλει την καλή λειτουργία.

2. Αν όλες οι επεκτάσεις της μη μονότονης θεωρίας (T, A) είναι οι $E_1$, $E_2$, ......., $E_n$, τότε η διάγνωση της βλάβης για το σύμπτωμα f είναι η $D_1 \vee D_2 \vee .... \vee D_n$, όπου κάθε $D_i$ είναι η σύζευξη των αρνήσεων των στοιχείων του $A - E_i$.

Στο παράδειγμά μας:

- Σύμπτωμα = $\neg$headlights-work
  Διάγνωση = $\neg$battery-ok $\vee$ $\neg$bulbs-ok

- Σύμπτωμα = $\neg$engines-starts
  Διάγνωση = $\neg$battery-ok $\vee$ $\neg$starter-ok

- Σύμπτωμα = $\neg$headlights-work $\wedge$ $\neg$engine-starts
  Διάγνωση = $\neg$battery-ok $\vee$ ($\neg$bulbs-ok $\wedge$ $\neg$starter-ok)

Οι διαφορές στις μεθοδολογίες διάγνωσης μέσω ATMS και μέσω μη μονότονης συλλογιστικής είναι μάλλον επιφανειακές παρά ουσιαστικές.

# ΣΥΛΛΟΓΙΣΤΙΚΗ ΜΕ ΠΙΘΑΝΟΤΗΤΕΣ

**(και όχι μόνο)**

# (PROBABILISTIC REASONING)

- Διατήρηση της ιδέας των ετικετών (labels)
  - Κλασική λογική πρώτης τάξης

    Ετικέτες: “αληθές” και “ψευδές”
  - Συστήματα συντήρησης της αλήθειας βασισμένης σε υποθέσεις (ATMS)

    Ετικέτες: εξηγήσεις
  - Μη μονότονη συλλογιστική

    Ετικέτες: επεκτάσεις
  - Συλλογιστική με πιθανότητες

    Ετικέτες: πραγματικοί αριθμοί (πιθανότητες, παράγοντες βεβαιότητας κλπ.)

- Συλλογιστική με πιθανότητες

  “Τα πουλιά πετούν”

  $$\text{bird}(x) \rightarrow \text{flies}(x)$$

  “Υπάρχουν όμως και εξαιρέσεις, π.χ. οι στρουθοκάμηλοι”

  $$\text{bird}(x) \wedge \neg \text{ab}(x) \rightarrow \text{flies}(x)$$

  Μία άλλη αντιμετώπιση:

  Πόσο βέβαιο είναι, ή τι πιθανότητα έχει να ισχύει, το:

  $$\text{bird}(x) \rightarrow \text{flies}(x)$$

**MYCIN**

- Ένα έμπειρο σύστημα για τη διάγνωση βακτηριακών μολύνσεων και την εισήγηση προτεινόμενων θεραπευτικών αγωγών (Buchanan, Shortliffe – 1984).
- Η γνώση του MYCIN είναι κωδικοποιημένη σε if-then κανόνες με τους οποίους είναι συνυφασμένοι παράγοντες βεβαιότητας (certainty factors), π.χ.

  If: the stain of the organism is gram-positive, and
  the morphology of the organism is coccus, and
  the growth conformation of the organism is clumps,
  then: there is suggestive evidence (0.7) that the
  identity of the organism is staphylococcus.

Παράγοντας βεβαιότητας: **CF[h, e] = MB[h, e] – MD[h, e]**

MB[h, e]: Μέτρο πίστης (measure of belief) στο ενδεχόμενο (hypothesis) h δεδομένης της μαρτυρίας/παρατήρησης (evidence) e.

MD[h, e]: Μέτρο δυσπιστίας (measure of disbelief) στο ενδεχόμενο h δεδομένης της μαρτυρίας e.

Ισχύουν τα εξής:

$$0 \leq MB[h, e] \leq 1 \quad , \quad 0 \leq MD[h, e] \leq 1$$

$$-1 \leq CF[h, e] \leq 1$$

Αν $MB[h, e] > 0$ τότε $MD[h, e] = 0$

Αν $MD[h, e] > 0$ τότε $MB[h, e] = 0$

- Ο παράγοντας βεβαιότητας σ' ένα κανόνα if-then είναι ουσιαστικά ο παράγοντας βεβαιότητας του συμπεράσματος ("then" τμήμα) με την προϋπόθεση ότι οι συνθήκες ("if" τμήμα) είναι 100% βέβαιες, δηλαδή έχουν παράγοντα βεβαιότητας 1.

- Αν ο παράγοντας βεβαιότητας του κανόνα if <συνθήκη> then <συμπέρασμα> είναι d και η <συνθήκη> έχει παράγοντα βεβαιότητας c, τότε το <συμπέρασμα> έχει παράγοντα βεβαιότητας $c \cdot d$, αν $c > 0$, ή 0, αν $c \leq 0$.

- Αν η <συνθήκη> είναι σύζευξη απλούστερων συνθηκών, τότε ο παράγοντας βεβαιότητάς της είναι το ελάχιστο των παραγόντων βεβαιότητας των επιμέρους συνθηκών.

- Αν ένα ενδεχόμενο h δεδομένης μίας μαρτυρίας $e_1$ έχει παράγοντα βεβαιότητας $CF[h, e_1] \equiv CF_1$ και δεδομένης μίας μαρτυρίας $e_2$ έχει παράγοντα βεβαιότητας $CF[h, e_2] \equiv CF_2$, τότε ο συνδυασμένος παράγοντας βεβαιότητας του h δεδομένων και των δύο μαρτυριών $e_1$ και $e_2$, δηλαδή το $CF[h, e_1 \wedge e_2] \equiv CF_{12}$, ισούται με:

$$CF_{12} = \begin{cases} CF_1 + CF_2 - CF_1 \cdot CF_2 & | \ CF_1, CF_2 \geq 0 \\ CF_1 + CF_2 + CF_1 \cdot CF_2 & | \ CF_1, CF_2 \leq 0 \\ \dfrac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)} & | \ CF_1 \cdot CF_2 < 0 \end{cases}$$

- Παράδειγμα

$$\left\{\begin{array}{ll}\text{Όταν ο Γιώργος κοιμάται, ροχαλίζει} & CF_1 = 0.7 \\ \text{Ο Γιώργος κοιμάται} & CF_2 = 0.8\end{array}\right.$$

Ο Γιώργος ροχαλίζει $CF_3 = 0.7 \cdot 0.8 = 0.56$

$$\left\{\begin{array}{ll}\text{Όταν ο Γιώργος ροχαλίζει, ο Κώστας διαμαρτύρεται} & CF_4 = 0.3 \\ \text{Ο Κώστας δεν διαμαρτύρεται} & CF_5 = 1.0\end{array}\right.$$

Ο Γιώργος δεν ροχαλίζει $CF_6 = 0.3 \cdot 1.0 = 0.3$

Ο Γιώργος ροχαλίζει $CF_7 = -\,0.3$

Ποιος είναι ο παράγοντας βεβαιότητας του "Ο Γιώργος ροχαλίζει" με βάση το συνδυασμό των μαρτυριών ότι "Ο Γιώργος κοιμάται" και ότι "Ο Κώστας δεν διαμαρτύρεται";

$$CF_8 = \frac{CF_3 + CF_7}{1 - \min(|CF_3|, |CF_7|)} = \frac{0.56 - 0.3}{1 - \min(0.56, 0.3)} = 0.37$$

- Το MYCIN έχει συγκρίσιμες επιδόσεις με εκείνες των ειδικών στον τομέα του (πείραμα Turing).

- Η αντικατάσταση του συνεχούς φάσματος [-1, 1] των παραγόντων βεβαιότητας με τις διακριτές τιμές "απόλυτα λάθος" (CF = –1.0), "σχετικά λάθος" (CF = –0.3), "σχετικά σωστό" (CF = 0.3) και "απόλυτα σωστό" (CF = 1.0) μέσω στρογγύλευσης προς αυτές των ενδιάμεσων τιμών δεν είχε σημαντική επίδραση στην απόδοση του MYCIN.

## Η θεωρία των πιθανοτήτων σαν εργαλείο στην αναπαράσταση γνώσης

**Η ΒΑΣΙΚΗ ΙΔΕΑ**: Ένας τύπος p δεν είναι απαραίτητο να είναι απόλυτα σωστός (ή απόλυτα λάθος). Ίσως είναι αρκετό να γνωρίζουμε απλώς ότι υπάρχει μία πιθανότητα pr(p) να είναι σωστός ($0 \le pr(p) \le 1$).

- Χρήσιμη έννοια είναι και η δεσμευμένη ή υπό συνθήκη (conditional) πιθανότητα pr(p | q) να ισχύει το p αν είναι γνωστό ότι ισχύει το q.

  **ΠΡΟΣΟΧΗ**: $pr(p \mid q) \neq pr(q \rightarrow p)$

- Ισχύουν τα εξής αξιώματα:

  $pr(p \wedge q) = pr(p) \cdot pr(q \mid p)$

  $pr(\neg p) = 1 - pr(p)$

  Αν $p \equiv q$ τότε $pr(p) = pr(q)$

- Ισχύει το θεώρημα του Bayes

  Η πιθανότητα pr(H | E) να είναι αληθές το ενδεχόμενο H δεδομένης της μαρτυρίας/παρατήρησης E, εφ' όσον $pr(E) \neq 0$, είναι:

$$pr(H \mid E) = \frac{pr(E \mid H) \cdot pr(H)}{pr(E)}$$

  (Επειδή $pr(H \mid E) \cdot pr(E) = pr(E \wedge H) = pr(E \mid H) \cdot pr(H)$)

- Συνέπειες του θεωρήματος του Bayes

  – Αν $pr(E \mid H) = 0$ τότε $pr(H \mid E) = 0$

  – Αν $pr(E \mid H_1) = pr(E \mid H_2)$

  τότε $$\frac{pr(H_1 \mid E)}{pr(H_2 \mid E)} = \frac{pr(H_1)}{pr(H_2)}$$

- Παράδειγμα

green = Το φανάρι είναι πράσινο

yellow = Το φανάρι είναι κίτρινο

red = Το φανάρι είναι κόκκινο

no_ticket = Δεν παίρνουμε κλήση περνώντας το φανάρι

Έστω ότι έχουμε:

pr(green) = 0.45

pr(yellow) = 0.1

pr(red) = 0.45

pr(no_ticket | green) = 1

pr(no_ticket | yellow) = 1

pr(no_ticket | red) = 0

Δεδομένου ότι τελικά περνώντας το φανάρι δεν πήραμε κλήση, ποιες είναι οι πιθανότητες να ήταν πράσινο, κίτρινο ή κόκκινο;

pr(green | no_ticket) = ;

pr(yellow | no_ticket) = ;

pr(red | no_ticket) = ;

Αφού pr(no_ticket | red) = 0 τότε pr(red | no_ticket) = 0

Αφού pr(no_ticket | green) = pr(no_ticket | yellow)

τότε $$\frac{\text{pr(yellow | no\_ticket)}}{\text{pr(green | no\_ticket)}} = \frac{0.1}{0.45} \qquad (1)$$

Επίσης pr(yellow | no_ticket) + pr(green | no_ticket) = 1 (2)

Από (1) και (2) έχουμε:

pr(yellow | no_ticket) = 2/11

pr(green | no_ticket) = 9/11

- Το H είναι ανεξάρτητο (independent) από το E όταν $pr(H \mid E) = pr(H)$.

- Όταν το H είναι ανεξάρτητο από το E ισχύει ότι $pr(H \wedge E) = pr(H) \cdot pr(E)$.

- Το H είναι υπό συνθήκη ανεξάρτητο (conditionally independent) από το E δεδομένου του F όταν $pr(H \mid E \wedge F) = pr(H \mid F)$

- Αν το H είναι ανεξάρτητο από το E, τότε και το E είναι ανεξάρτητο από το H.

$$\Big(\text{Επειδή} \quad pr(E \mid H) = \frac{pr(H \mid E) \cdot pr(E)}{pr(H)} =$$

$$= \frac{pr(H) \cdot pr(E)}{pr(H)} = pr(E)\Big)$$

- Επίσης ισχύει ότι:

$$pr(p \vee q) = pr(p) + pr(q) - pr(p \wedge q)$$

$$\Big(\text{Επειδή} \quad pr(p \vee q) = pr(\neg(\neg p \wedge \neg q)) =$$

$$= 1 - pr(\neg p \wedge \neg q) = 1 - pr(\neg p) \cdot pr(\neg q \mid \neg p) =$$

$$= 1 - pr(\neg p) \cdot (1 - pr(q \mid \neg p)) =$$

$$= 1 - pr(\neg p) + pr(\neg p) \cdot pr(q \mid \neg p) =$$

$$= pr(p) + pr(\neg p) \cdot pr(q \mid \neg p) =$$

$$= pr(p) + pr(\neg p \wedge q) = pr(p) + pr(q) \cdot pr(\neg p \mid q) =$$

$$= pr(p) + pr(q) \cdot (1 - pr(p \mid q)) =$$

$$= pr(p) + pr(q) - pr(q) \cdot pr(p \mid q) =$$

$$= pr(p) + pr(q) - pr(p \wedge q)\Big)$$

- Αν τα p και q αλληλοαποκλείονται ( $pr(p \wedge q) = 0$ ) τότε $pr(p \vee q) = pr(p) + pr(q)$.

- Νόμος της ολικής πιθανότητας (Law of total probability)

Αν $pr(w_1 \vee \ldots\ldots \vee w_n) = 1$ και

$pr(w_i \wedge w_j) = 0$ για κάθε $i \neq j$

τότε $pr(p) = \sum_i pr(p \mid w_i) \cdot pr(w_i)$

Παράδειγμα:

$$\begin{aligned} pr(no\_ticket) &= pr(no\_ticket \mid green) \cdot pr(green) + \\ &\quad pr(no\_ticket \mid yellow) \cdot pr(yellow) + \\ &\quad pr(no\_ticket \mid red) \cdot pr(red) = \\ &= 1 \cdot 0.45 + 1 \cdot 0.1 + 0 \cdot 0.45 = 0.55 \end{aligned}$$

- Γενίκευση του θεωρήματος του Bayes:
  Αν δεδομένης κάποιας μαρτυρίας E, όλα τα πιθανά και αλληλοαποκλειόμενα ανά δύο ενδεχόμενα που υπάρχουν είναι τα $H_1$, $H_2$, ......., $H_n$, τότε

$$pr(H_i \mid E) = \frac{pr(E \mid H_i) \cdot pr(H_i)}{\sum_j pr(E \mid H_j) \cdot pr(H_j)}$$

- Εκτεταμένη χρήση της συλλογιστικής με πιθανότητες και του θεωρήματος του Bayes έγινε στο έμπειρο σύστημα PROSPECTOR (Duda κλπ. – 1976) για την αναζήτηση περιοχών με αξιοποιήσιμα κοιτάσματα ορυκτών.

- Η επιτυχία του PROSPECTOR στην ανακάλυψη κοιτασμάτων ήταν εντυπωσιακή, αλλά, παρ' όλα αυτά, η χρήση πιθανοτήτων στην Τεχνητή Νοημοσύνη έχει υποστεί μεγάλη κριτική.

- Παράδειγμα:

green = το φανάρι είναι πράσινο

yellow = το φανάρι είναι κίτρινο

red = το φανάρι είναι κόκκινο

no_ticket = δεν παίρνουμε κλήση περνώντας το φανάρι

bad_mood = είμαστε κακοδιάθετοι

Έστω ότι έχουμε:

pr(green) = 0.45

pr(yellow) = 0.1

pr(red) = 0.45

pr(no_ticket | green) = 1

pr(no_ticket | yellow) = 0.95

pr(no_ticket | red) = 0.25

pr(bad_mood |¬no_ticket) = 0.9

pr(bad_mood | no_ticket) = 0.05

Αφού περάσουμε το φανάρι, ποια είναι η πιθανότητα να είμαστε κακοδιάθετοι;

pr(bad_mood) = ;

pr(bad_mood) =

= pr(bad_mood | green ∧¬no_ticket) · pr(green∧¬no_ticket) +

pr(bad_mood | green ∧ no_ticket) · pr(green ∧ no_ticket) +

pr(bad_mood | yellow∧¬no_ticket) · pr(yellow∧¬no_ticket) +

pr(bad_mood | yellow∧no_ticket) · pr(yellow∧no_ticket) +

pr(bad_mood | red ∧¬no_ticket) · pr(red∧¬no_ticket) +

pr(bad_mood | red ∧ no_ticket) · pr(red ∧ no_ticket) =

= (με ποιο δικαίωμα;)

$pr(bad\_mood \mid \neg no\_ticket) \cdot pr(green \wedge \neg no\_ticket) +$

$pr(bad\_mood \mid no\_ticket) \cdot pr(green \wedge no\_ticket) +$

$pr(bad\_mood \mid \neg no\_ticket) \cdot pr(yellow \wedge \neg no\_ticket) +$

$pr(bad\_mood \mid no\_ticket) \cdot pr(yellow \wedge no\_ticket) +$

$pr(bad\_mood \mid \neg no\_ticket) \cdot pr(red \wedge \neg no\_ticket) +$

$pr(bad\_mood \mid no\_ticket) \cdot pr(red \wedge no\_ticket) =$

$= pr(bad\_mood \mid \neg no\_ticket) \cdot pr(green) \cdot pr(\neg no\_ticket \mid green) +$

$pr(bad\_mood \mid no\_ticket) \cdot pr(green) \cdot pr(no\_ticket \mid green) +$

$pr(bad\_mood \mid \neg no\_ticket) \cdot pr(yellow) \cdot pr(\neg no\_ticket \mid yellow) +$

$pr(bad\_mood \mid no\_ticket) \cdot pr(yellow) \cdot pr(no\_ticket \mid yellow) +$

$pr(bad\_mood \mid \neg no\_ticket) \cdot pr(red) \cdot pr(\neg no\_ticket \mid red) +$

$pr(bad\_mood \mid no\_ticket) \cdot pr(red) \cdot pr(no\_ticket \mid red) =$

$= 0.9 \cdot 0.45 \cdot 0 + 0.05 \cdot 0.45 \cdot 1 +$

$0.9 \cdot 0.1 \cdot 0.05 + 0.05 \cdot 0.1 \cdot 0.95 +$

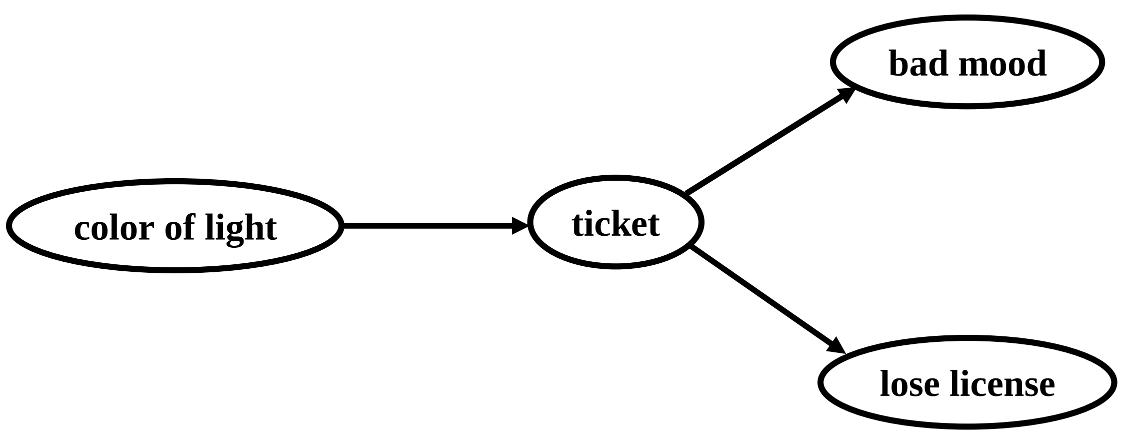
$0.9 \cdot 0.45 \cdot 0.75 + 0.05 \cdot 0.45 \cdot 0.25 =$

$= 0.341125$

- Εν γένει, για να βρεθεί η πιθανότητα ενός q που εξαρτάται (πόσο αλήθεια;) από κάποια $p_1, p_2, \ldots\ldots, p_n$ χρειάζεται να είναι γνωστές όλες οι ($2^n$ πλήθους) υπό συνθήκη πιθανότητες $pr(q \mid p_1 \wedge \ldots \wedge p_n)$, $pr(q \mid p_1 \wedge \ldots \wedge \neg p_n)$ ....... καθώς επίσης και όλες οι (επίσης $2^n$ πλήθους) πιθανότητες $pr(p_1 \wedge \ldots \wedge p_n)$, $pr(p_1 \wedge \ldots \wedge \neg p_n)$ ......., τα οποία, σε πραγματικά προβλήματα, δεν είναι ιδιαιτέρως εφικτά.

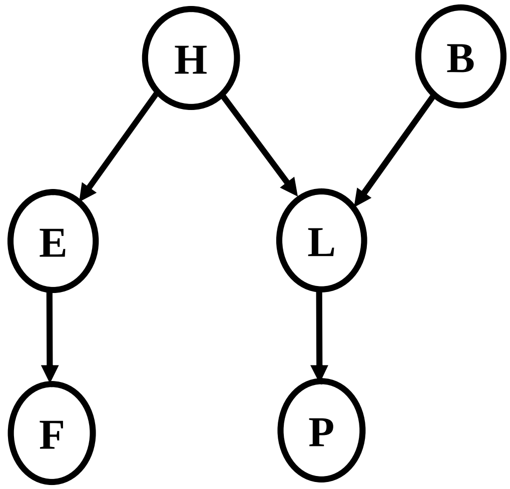
- Ένας τρόπος να μειωθεί η πολυπλοκότητα των αλληλοεξαρτήσεων μεταξύ διαφόρων ενδεχομένων ή μαρτυριών είναι η εισαγωγή υπό συνθήκη ανεξαρτησιών μέσω των λεγόμενων διαγραμμάτων επιρροής (influence diagrams) ή δικτύων Bayes (Bayesian networks).

- Ένα διάγραμμα επιρροής είναι ένας κατευθυνόμενος γράφος χωρίς κύκλους G = (V, E), όπου V είναι ένα σύνολο μεταβλητών (π.χ. “bad mood”, “ticket”, “color of light”) και E είναι ένα σύνολο ζευγών αλληλοεξαρτώμενων μεταβλητών (π.χ. το “bad mood” εξαρτάται από το “ticket” και το “ticket” εξαρτάται από το “color of light”):




- Σ’ ένα διάγραμμα επιρροής, κάθε κόμβος είναι υπό συνθήκη ανεξάρτητος από τους κόμβους που δεν είναι απόγονοί του δεδομένων των γονέων του (π.χ. το “bad mood” είναι υπό συνθήκη ανεξάρτητο από το “color of light” δεδομένου του “ticket” – με βάση αυτήν την υπόθεση είχαμε το δικαίωμα να θεωρήσουμε ότι:
pr(bad_mood | green $\wedge\neg$no_ticket) = pr(bad_mood |$\neg$no_ticket) )

- Παράδειγμα:




Το P είναι υπό συνθήκη ανεξάρτητο από τα E, F, B, H δεδομένου του L. Δηλαδή ισχύει:

$$pr(P \mid E \wedge F \wedge L \wedge B \wedge H) = pr(P \mid L)$$

- Τοπολογική ταξινόμηση (topological sort) των μεταβλητών σ' ένα διάγραμμα επιρροής είναι μία σειρά των μεταβλητών τέτοια ώστε καμία μεταβλητή να προηγείται των παιδιών της (π.χ. στο προηγούμενο διάγραμμα η P, F, L, E, B, H είναι μία τοπολογική ταξινόμηση).
- Για να υπολογισθεί η πιθανότητα να έχουν οι μεταβλητές μας συγκεκριμένες τιμές μπορεί κάποιος να εκμεταλλευθεί ένα διάγραμμα επιρροής και μία τοπολογική ταξινόμηση των μεταβλητών ως εξής:

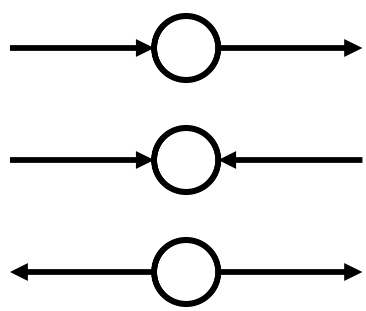
$$\begin{aligned}
& pr(P \wedge F \wedge L \wedge E \wedge B \wedge H) = \\
&= pr(P \mid F \wedge L \wedge E \wedge B \wedge H) \cdot pr(F \mid L \wedge E \wedge B \wedge H) \cdot \\
&\quad pr(L \mid E \wedge B \wedge H) \cdot pr(E \mid B \wedge H) \cdot pr(B \mid H) \cdot pr(H) = \\
&= pr(P \mid L) \cdot pr(F \mid E) \cdot pr(L \mid B \wedge H) \cdot \\
&\quad pr(E \mid H) \cdot pr(B) \cdot pr(H)
\end{aligned}$$

- Τι μπορούμε να γράψουμε για το διάγραμμα με το “ticket”;

- Η έννοια της υπό συνθήκη ανεξαρτησίας σε διαγράμματα επιρροής μπορεί να εκφρασθεί και μέσω των λεγόμενων συνδετικών-της-εξάρτησης μονοπατιών (dependency-connecting-paths).

- Ένα μη-κατευθυνόμενο (undirected) μονοπάτι σ’ ένα διάγραμμα επιρροής είναι μια ακολουθία από συνδεόμενους κόμβους χωρίς να λαμβάνεται υπόψη η κατεύθυνση της ακμής.

- Σ’ ένα μη-κατευθυνόμενο μονοπάτι, οι κόμβοι (πλην του πρώτου και του τελευταίου) διακρίνονται σε γραμμικούς (linear), συγκλίνοντες (converging) και αποκλίνοντες (diverging) ως εξής:



γραμμικός

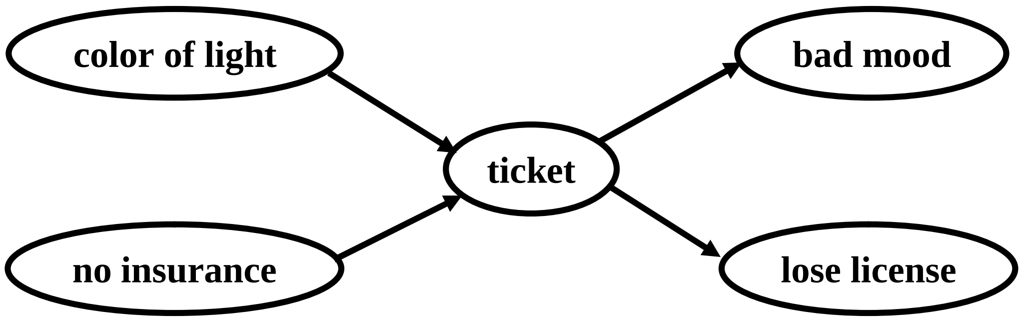
συγκλίνων

αποκλίνων

- Ένα μη-κατευθυνόμενο μονοπάτι p από έναν κόμβο X προς έναν κόμβο Y σ’ ένα διάγραμμα επιρροής $G = (V, E)$ δεδομένου ενός συνόλου $S \subset V$ λέγεται ότι είναι ένα συνδετικό-της-εξάρτησης μονοπάτι αν για κάθε κόμβο n στο p ισχύει:
  - είτε το n είναι γραμμικός ή αποκλίνων κόμβος και δεν ανήκει στο S.
  - είτε το n είναι συγκλίνων κόμβος και ή το ίδιο ή κάποιος από τους διαδόχους του ανήκει στο S.

- Ένας κόμβος X είναι υπό συνθήκη ανεξάρτητος από έναν κόμβο Y δεδομένου του S, αν δεν υπάρχει συνδετικό-της-εξάρτησης μονοπάτι που να τους συνδέει δεδομένου του S.

- Παράδειγμα




Το “color of light” δεν είναι ανεξάρτητο από το “insurance” δεδομένου του “bad mood” (γιατί;).

**Θεωρία πιθανοτήτων και τεχνητή νοημοσύνη**

- Παρά τα εξαιρετικά σε πολλές περιπτώσεις αποτελέσματα, η θεωρία των πιθανοτήτων έχει υποστεί σκληρή κριτική για τη χρησιμοποίησή της στην τεχνητή νοημοσύνη.

- Οι σπουδαιότερες αντιρρήσεις είναι: “Ποιος και πώς είναι σε θέση να εκτιμήσει σωστά τις πιθανότητες διαφόρων ενδεχομένων; Πώς μπορεί να διακριθεί η άγνοια από τη στατιστική γνώση;”

- Άλλες “αριθμοκεντρικές” αντιμετωπίσεις της αναπαράστασης γνώσης στην τεχνητή νοημοσύνη είναι:
  – Η ασαφής λογική (fuzzy logic)
  – Η θεωρία Dempster–Shafer

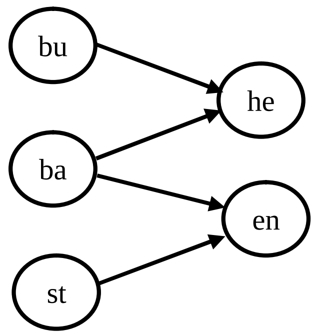
**Εφαρμογή της συλλογιστικής με πιθανότητες για διάγνωση**

Έστω ότι θεωρούνται δεδομένα τα:

$$\text{battery-ok} \wedge \text{bulbs-ok} \rightarrow \text{headlights-work}$$

$$\text{battery-ok} \wedge \text{starter-ok} \rightarrow \text{engine-starts}$$

Έστω επίσης ότι τα $\neg$bulbs-ok, $\neg$battery-ok και $\neg$starter-ok είναι ανεξάρτητα μεταξύ τους και ότι οι πιθανότητες να συμβούν είναι:




$$pr(\neg \text{bulbs-ok}) = p_h$$

$$pr(\neg \text{battery-ok}) = p_b$$

$$pr(\neg \text{starter-ok}) = p_s$$

Συνεπώς, υπάρχουν 8 πιθανές διαφορετικές καταστάσεις:

| | $\neg$bulbs-ok | $\neg$battery-ok | $\neg$starter-ok | πιθανότητα |
|---|---|---|---|---|
| (1) | T | T | T | $p_h \cdot p_b \cdot p_s$ |
| (2) | T | T | F | $p_h \cdot p_b \cdot (1 - p_s)$ |
| (3) | T | F | T | $p_h \cdot (1 - p_b) \cdot p_s$ |
| (4) | T | F | F | $p_h \cdot (1 - p_b) \cdot (1 - p_s)$ |
| (5) | F | T | T | $(1 - p_h) \cdot p_b \cdot p_s$ |
| (6) | F | T | F | $(1 - p_h) \cdot p_b \cdot (1 - p_s)$ |
| (7) | F | F | T | $(1 - p_h) \cdot (1 - p_b) \cdot p_s$ |
| (8) | F | F | F | $(1 - p_h) \cdot (1 - p_b) \cdot (1 - p_s)$ |

Θα βρούμε τις πιθανότητες να συμβούν τα $\neg$headlights-work, $\neg$engine-starts και $\neg$headlights-work $\wedge \neg$engine-starts καθώς και τις υπό συνθήκη πιθανότητες των $\neg$headlights-work (αλλά και των $\neg$engine-starts και $\neg$headlights-work $\wedge \neg$engine-starts)

Θέτουμε τα $w_i$ ως εξής:

$w_1 = \neg$ bulbs-ok $\wedge \neg$ battery-ok $\wedge \neg$ starter-ok

$w_2 = \neg$ bulbs-ok $\wedge \neg$ battery-ok $\wedge$ starter-ok

$w_3 = \neg$ bulbs-ok $\wedge$ battery-ok $\wedge \neg$ starter-ok

$w_4 = \neg$ bulbs-ok $\wedge$ battery-ok $\wedge$ starter-ok

$w_5 =$ bulbs-ok $\wedge \neg$ battery-ok $\wedge \neg$ starter-ok

$w_6 =$ bulbs-ok $\wedge \neg$ battery-ok $\wedge$ starter-ok

$w_7 =$ bulbs-ok $\wedge$ battery-ok $\wedge \neg$ starter-ok

$w_8 =$ bulbs-ok $\wedge$ battery-ok $\wedge$ starter-ok

Θέτουμε επίσης:

bu = bulbs-ok, ba = battery-ok, st = starter-ok,

he = headlights-work, en = engine-starts

Έχουμε λοιπόν:

1. pr($\neg$headlights-work) = pr($\neg$he) = 1 – pr(he) =

$$= 1 - [\, pr(he \mid w_1) \cdot pr(w_1) + pr(he \mid w_2) \cdot pr(w_2) + pr(he \mid w_3) \cdot pr(w_3) + pr(he \mid w_4) \cdot pr(w_4) + pr(he \mid w_5) \cdot pr(w_5) + pr(he \mid w_6) \cdot pr(w_6) + pr(he \mid w_7) \cdot pr(w_7) + pr(he \mid w_8) \cdot pr(w_8)\,] =$$

$$= 1 - [\, 0 + 0 + 0 + 0 + 0 + 0 + 1 \cdot (1 - p_h) \cdot (1 - p_b) \cdot p_s + 1 \cdot (1 - p_h) \cdot (1 - p_b) \cdot (1 - p_s)\,] = p_b + p_h - p_b \cdot p_h$$

Άρα: $\boxed{pr(\neg \text{headlights-work}) = p_b + p_h - p_b \cdot p_h}$

2. Με όμοιο τρόπο αποδεικνύεται εύκολα ότι:

$$\boxed{pr(\neg \text{engine-starts}) = p_b + p_s - p_b \cdot p_s}$$

3. $pr(\neg \text{headlights-work} \wedge \neg \text{engine-starts}) = pr(\neg he \wedge \neg en) =$

$= pr(\neg(he \vee en)) = 1 - pr(he \vee en) = 1 - [\, pr(he) +$

$pr(en) - pr(he \wedge en)\, ] = 1 - pr(he) - pr(en) + pr(he \wedge en) =$

$= 1 - (1 - pr(\neg he)) - (1 - pr(\neg en)) + pr(he \wedge en) =$

$= 1 - 1 + pr(\neg he) - 1 + pr(\neg en) + pr(he \wedge en) =$

$= p_b + p_h - p_b \cdot p_h + p_b + p_s - p_b \cdot p_s - 1 + pr(he \wedge en) =$

$= p_b + p_h - p_b \cdot p_h + p_b + p_s - p_b \cdot p_s - 1 +$

$pr(he \wedge en \mid w_1) \cdot pr(w_1) + pr(he \wedge en \mid w_2) \cdot pr(w_2) +$

$pr(he \wedge en \mid w_3) \cdot pr(w_3) + pr(he \wedge en \mid w_4) \cdot pr(w_4) +$

$pr(he \wedge en \mid w_5) \cdot pr(w_5) + pr(he \wedge en \mid w_6) \cdot pr(w_6) +$

$pr(he \wedge en \mid w_7) \cdot pr(w_7) + pr(he \wedge en \mid w_8) \cdot pr(w_8) =$

$= p_b + p_h - p_b \cdot p_h + p_b + p_s - p_b \cdot p_s - 1 + 0 + 0 + 0 + 0 +$

$+ 0 + 0 + 0 + 1 \cdot (1 - p_h) \cdot (1 - p_b) \cdot (1 - p_s) =$

$= p_b + p_h \cdot p_s - p_b \cdot p_h \cdot p_s$

Άρα: $\boxed{pr(\neg \text{headlights-work} \wedge \neg \text{engine-starts}) = p_b + p_h \cdot p_s - p_b \cdot p_h \cdot p_s}$

4. $pr(\neg \text{bulbs-ok} \mid \neg \text{headlights-work}) = pr(\neg bu \mid \neg he) =$

$$= \frac{pr(\neg he \mid \neg bu) \cdot pr(\neg bu)}{pr(\neg he)} = \frac{1 \cdot p_h}{p_b + p_h - p_b \cdot p_h} = \frac{p_h}{p_b + p_h - p_b \cdot p_h}$$

Άρα: $\boxed{pr(\neg \text{bulbs-ok} \mid \neg \text{headlights-work}) = \frac{p_h}{p_b + p_h - p_b \cdot p_h}}$

5. $pr(\neg \text{battery-ok} \mid \neg \text{headlights-work}) = pr(\neg ba \mid \neg he) =$

$$= \frac{pr(\neg he \mid \neg ba) \cdot pr(\neg ba)}{pr(\neg he)} = \frac{1 \cdot p_b}{p_b + p_h - p_b \cdot p_h} = \frac{p_b}{p_b + p_h - p_b \cdot p_h}$$

Άρα: $\boxed{pr(\neg \text{battery-ok} \mid \neg \text{headlights-work}) = \frac{p_b}{p_b + p_h - p_b \cdot p_h}}$

6. pr(¬starter-ok | ¬headlights-work) = pr(¬st |¬he) =

= pr(¬st) = $p_s$

Άρα: $\boxed{\text{pr}(\neg\text{starter-ok} \mid \neg\text{headlights-work}) = p_s}$

Σημείωση: Στο 4 γράψαμε pr(¬he | ¬bu) = 1 και στο 5 pr(¬he | ¬ba) = 1. Αυτά είναι σωστά εφ' όσον θεωρήσουμε ότι η βάση γνώσης μας περιέχει και τα

headlights-work → battery-ok ∧ bulbs-ok

engine-starts → battery-ok ∧ starter-ok

Επίσης στο 6 γράψαμε pr(¬st | ¬he) = pr (¬st) επειδή θεωρήσαμε ότι τα ¬st και ¬he είναι ανεξάρτητα.

7. Εύκολα αποδεικνύεται το:

$$\boxed{\text{pr}(\neg\text{bulbs-ok} \mid \neg\text{engine-starts}) = p_h}$$

8. Εύκολα αποδεικνύεται το:

$$\boxed{\text{pr}(\neg\text{battery-ok} \mid \neg\text{engine-starts}) = \frac{p_b}{p_b + p_s - p_b \cdot p_s}}$$

9. Εύκολα αποδεικνύεται το:

$$\boxed{\text{pr}(\neg\text{starter-ok} \mid \neg\text{engine-starts}) = \frac{p_s}{p_b + p_s - p_b \cdot p_s}}$$

10. pr(¬bulbs-ok | ¬headlights-work ∧¬engine-starts) =

$$= \text{pr}(\neg\text{bu} \mid \neg\text{he} \wedge \neg\text{en}) = \frac{\text{pr}(\neg\text{he} \wedge \neg\text{en} \mid \neg\text{bu}) \cdot \text{pr}(\neg\text{bu})}{\text{pr}(\neg\text{he} \wedge \neg\text{en})} =$$

$$= \frac{\text{pr}(\neg\text{en} \mid \neg\text{bu}) \cdot \text{pr}(\neg\text{bu})}{\text{pr}(\neg\text{he} \wedge \neg\text{en})} = \frac{\text{pr}(\neg\text{en}) \cdot \text{pr}(\neg\text{bu})}{\text{pr}(\neg\text{he} \wedge \neg\text{en})} =$$

$$= \frac{(p_b + p_s - p_b \cdot p_s) \cdot p_h}{p_b + p_h \cdot p_s - p_b \cdot p_h \cdot p_s}$$

Άρα: $\boxed{pr(\neg\text{bulbs-ok} \mid \neg\text{headlights-work} \wedge \neg\text{engine-starts}) = \dfrac{(p_b + p_s - p_b \cdot p_s) \cdot p_h}{p_b + p_h \cdot p_s - p_b \cdot p_h \cdot p_s}}$

11. Με όμοιο τρόπο αποδεικνύεται εύκολα ότι:

$$\boxed{pr(\neg\text{battery-ok} \mid \neg\text{headlights-work} \wedge \neg\text{engine-starts}) = \dfrac{p_b}{p_b + p_h \cdot p_s - p_b \cdot p_h \cdot p_s}}$$

12. Επίσης αποδεικνύεται και ότι:

$$\boxed{pr(\neg\text{starter-ok} \mid \neg\text{headlights-work} \wedge \neg\text{engine-starts}) = \dfrac{(p_b + p_h - p_b \cdot p_h) \cdot p_s}{p_b + p_h \cdot p_s - p_b \cdot p_h \cdot p_s}}$$

Θα εξετάσουμε τώρα κάποιες ειδικές περιπτώσεις:

A. Τι γίνεται όταν $p_h = p_b = p_s$ (= p) << 1 **;**

| | ¬bu | ¬ba | ¬st | Διάγνωση |
|---|---|---|---|---|
| ¬he | 1/2 | 1/2 | p | ? |
| ¬en | p | 1/2 | 1/2 | ? |
| ¬he ∧¬en | 2 · p | 1 – p | 2 · p | ¬ba |

B. Τι γίνεται όταν $p_b << p_h \cdot p_s$ και $p_h, p_s << 1$ **;**

| | ¬bu | ¬ba | ¬st | Διάγνωση |
|---|---|---|---|---|
| ¬he | 1 | $p_b / p_h$ | $p_s$ | ¬bu |
| ¬en | $p_h$ | $p_b / p_s$ | 1 | ¬st |
| ¬he ∧¬en | 1 | $p_b / (p_h \cdot p_s)$ | 1 | ¬bu ∧¬st |

## ΠΛΑΙΣΙΑ ΚΑΙ ΣΗΜΑΣΙΟΛΟΓΙΚΑ ΔΙΚΤΥΑ
## (FRAMES AND SEMANTIC NETS)

- Εναλλακτικές μέθοδοι χειρισμού της γνώσης

- Έστω ο "μη μονότονος" κόσμος:

$$\text{bird}(x) \rightarrow \text{feathers}(x)$$
$$\text{bird}(x) \wedge \neg ab_b(x) \rightarrow \text{flies}(x)$$
$$\text{ostrich}(x) \rightarrow \text{bird}(x)$$
$$\text{ostrich}(x) \wedge \neg ab_o(x) \rightarrow \text{walks}(x)$$
$$\text{mammal}(x) \rightarrow \text{hair}(x)$$
$$\text{mammal}(x) \wedge \neg ab_m(x) \rightarrow \text{walks}(x)$$
$$\text{tiger}(x) \rightarrow \text{mammal}(x)$$
$$\text{whale}(x) \rightarrow \text{mammal}(x)$$
$$\text{whale}(x) \wedge \neg ab_w(x) \rightarrow \text{swims}(x)$$
$$\text{bird}(\text{Tweety})$$
$$\text{ostrich}(\text{Fred})$$
$$\text{tiger}(\text{Hobbes})$$
$$\text{whale}(\text{Moby})$$

**ΕΡΩΤΗΣΗ:** Είναι δυνατόν η παραπάνω γνώση να αναπαρασταθεί πιο κατανοητά με κάποιο τρόπο βασισμένο στη λογική;

**ΑΠΑΝΤΗΣΗ:** **ΝΑΙ.** Αν κάνουμε τη λεγόμενη δεδομενοποίηση (reification).

Τα γενικά αξιώματά μας – Χειρισμός κλάσεων (classes):

$$\text{subclass}(x, y) \land \text{instance}(a, x) \rightarrow \text{instance}(a, y) \quad (1)$$

$$\text{subclass}(x, y) \land \text{subclass}(y, z) \rightarrow \text{subclass}(x, z) \quad (2)$$

$$\text{instance}(a, x) \land \text{certain}(x, p, v) \rightarrow \text{value}(a, p, v) \quad (3)$$

$$\text{instance}(a, x) \land \text{default}(x, p, v) \land \neg \text{ab}(a, x, p) \rightarrow \text{value}(a, p, v) \quad (4)$$

Η συγκεκριμένη γνώση μας:

$$\left.\begin{array}{l}\text{subclass(ostrich, bird)}\\ \text{subclass(tiger, mammal)}\\ \text{subclass(whale, mammal)}\end{array}\right] \quad (5)$$

$$\left.\begin{array}{l}\text{certain(bird, covering, feathers)}\\ \text{certain(mammal, covering, hair)}\end{array}\right] \quad (6)$$

$$\left.\begin{array}{l}\text{default(bird, locomotion, flies)}\\ \text{default(ostrich, locomotion, walks)}\\ \text{default(mammal, locomotion, walks)}\\ \text{default(whale, locomotion, swims)}\end{array}\right] \quad (7)$$

$$\left.\begin{array}{l}\text{instance(Tweety, bird)}\\ \text{instance(Fred, ostrich)}\\ \text{instance(Hobbes, tiger)}\\ \text{instance(Moby, whale)}\end{array}\right] \quad (8)$$

- Η παραπάνω αναπαράσταση πολύ εύκολα μετασχηματίζεται σε αναπαράσταση μέσω πλαισίων ή/και μέσω σημασιολογικών δικτύων.

- Ένα σύστημα πλαισίων (frame system) αποτελείται από μία συλλογή από αντικείμενα (bird, ostrich κλπ.) καθένα από τα οποία έχει σχισμές (slots) (covering, instance-of κλπ.) και τιμές γι’ αυτές τις σχισμές (feathers, άλλα αντικείμενα κλπ.).

Παράδειγμα:

```
bird
        covering:      feathers
        locomotion:    flies
ostrich
        subclass-of:   bird
        locomotion:    walks
mammal
        covering:      hair
        locomotion:    walks
tiger
        subclass-of:   mammal
whale
        subclass-of:   mammal
        locomotion:    swims
Tweety
        instance-of:   bird
Fred
        instance-of:   ostrich
Hobbes
        instance-of:   tiger
Moby
        instance-of:   whale
```

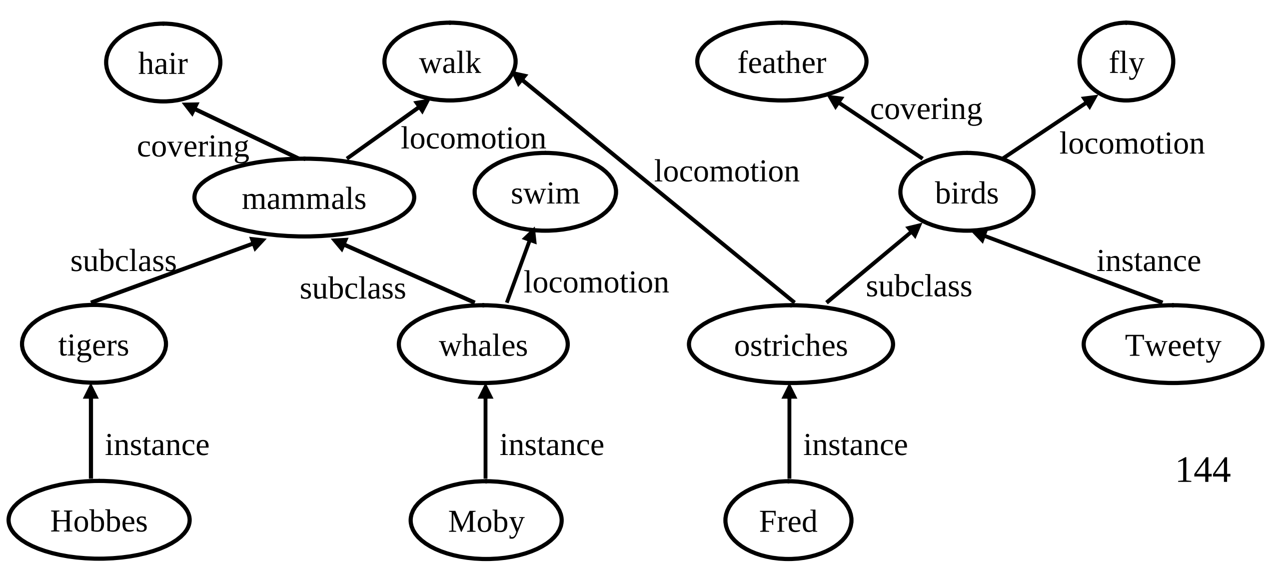
**ΠΛΕΟΝΕΚΤΗΜΑΤΑ**

- Συγκέντρωση των ιδιοτήτων αντικειμένων ή κατηγοριών αντικειμένων σ’ ένα μέρος
- Αποδοτικότερη συλλογιστική (θα δούμε γιατί)

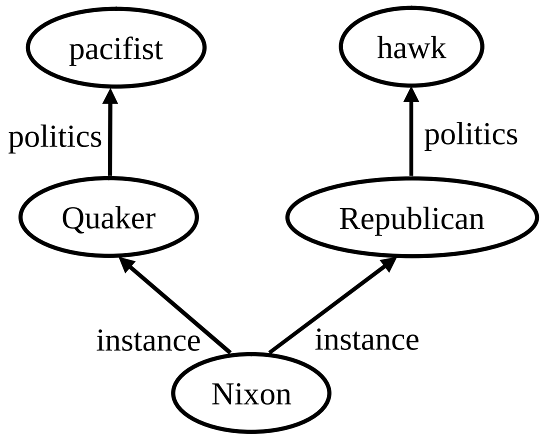
- Η μονότονη μετάφραση (monotonic translation) ενός συστήματος πλαισίων αποτελείται από τις προτάσεις (1), (2), (3) και κατάλληλες προτάσεις της μορφής (5), (6), (8).

- Η μη μονότονη μετάφραση (nonmonotonic translation) ενός συστήματος πλαισίων αποτελείται από τις προτάσεις (1), (2), (4) και κατάλληλες προτάσεις της μορφής (5), (7), (8).

- Ένα σύστημα πλαισίων F συμπεραίνει μονότονα (monotonically entails) μία πρόταση p αν η p είναι συνέπεια της μονότονης μετάφρασης του F. Μία πρόταση p την συμπεραίνει μη μονότονα (nonmonotonically entails) αν η p είναι συνέπεια της μη μονότονης μετάφρασής του.
- Τα σημασιολογικά δίκτυα είναι ένας άλλος (ίσως καλύτερος) τρόπος αναπαράστασης γνώσης που δεν διαφέρει ουσιαστικά από αυτόν μέσω συστήματος πλαισίων.



## Πολλαπλή κληρονομικότητα (Multiple inheritance)




Republican

politics: hawk

Quaker

politics: pacifist

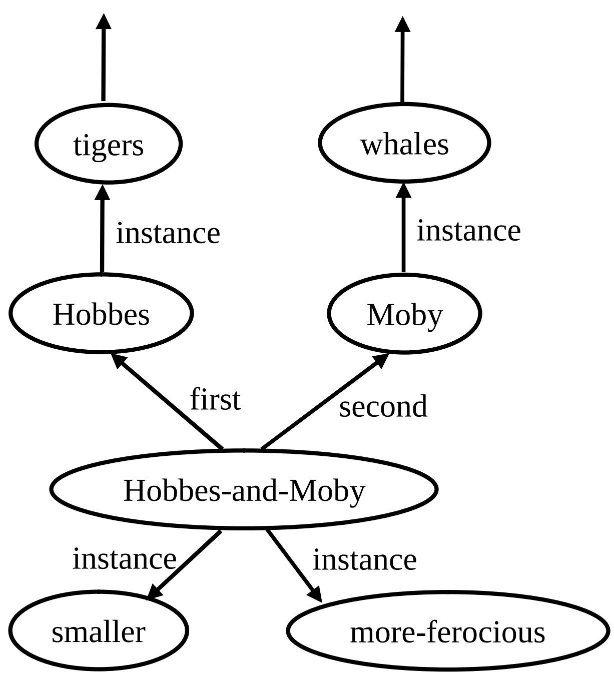
Nixon

instance-of: Republican

instance-of: Quaker

## Σχέσεις μεταξύ αντικειμένων

Πως μπορούν να εκφρασθούν τα smaller(Hobbes, Moby) και more-ferocious(Hobbes, Moby);




Hobbes-and-Moby

instance-of: smaller

instance-of: more-ferocious

first: Hobbes

second: Moby

Για τη μετάφραση σε λογική πρώτης τάξης χρειάζεται και το:

$$\text{value}(p, \text{first}, x) \wedge \text{value}(p, \text{second}, y) \wedge \text{instance}(p, r) \rightarrow \text{holds}(r, x, y)$$

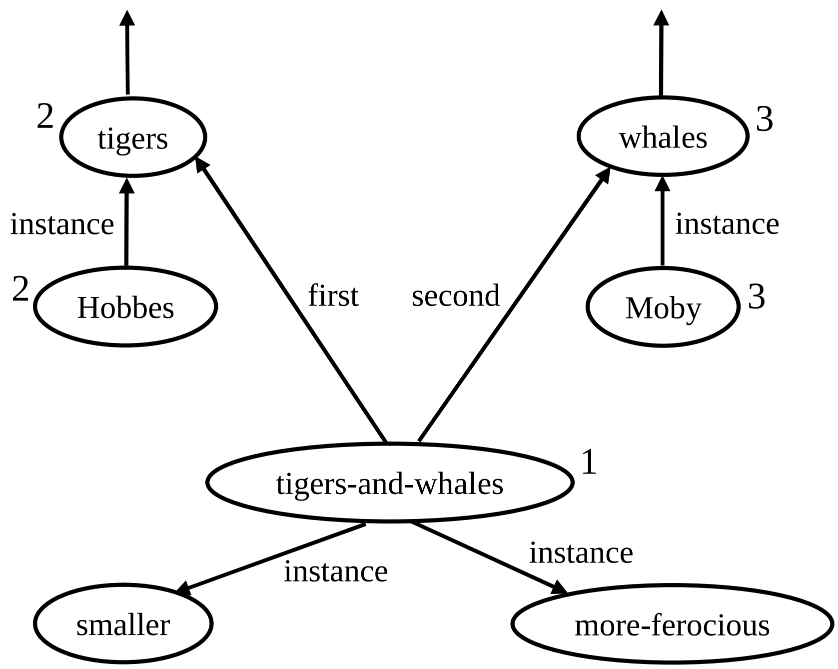
που θα συνδυασθεί με τα:

value(Hobbes-and-Moby, first, Hobbes)

value(Hobbes-and-Moby, second, Moby)

instance(Hobbes-and-Moby, smaller)

Αλλά δεν είναι ο συγκεκριμένος Hobbes που είναι smaller και more-ferocious από το συγκεκριμένο Moby. Αυτές οι σχέσεις ισχύουν για κάθε ζευγάρι tiger και whale.




Για τη μετάφραση σε λογική όμως χρειάζεται και το:

$\text{instance}(a, x) \wedge \text{instance}(b, y) \wedge \text{holds}(r, x, y) \rightarrow \text{holds}(r, a, b)$

**Συλλογιστική σε μονότονα συστήματα πλαισίων**

**ΕΡΩΤΗΣΗ:** Δεδομένου ενός συστήματος πλαισίων (ή του ισοδύναμού του σημασιολογικού δικτύου) υπάρχει κάποιος αποδοτικός τρόπος για να ελεγχθεί αν ισχύει το instance(A, X) για δεδομένα A και X**;**

**ΑΠΑΝΤΗΣΗ: ΝΑΙ.** Ο αλγόριθμος A που ακολουθεί.

**ΕΡΩΤΗΣΗ:** Μπορεί το ίδιο συμπέρασμα να εξαχθεί από κάποια εξειδίκευση της ανάλυσης με την εφαρμογή της στην αντίστοιχη μονότονη μετάφραση;

**ΑΠΑΝΤΗΣΗ: ΝΑΙ.** Είναι ο αλγόριθμος B που ακολουθεί.

**Αλγόριθμος Α**

Έστω ένα σημασιολογικό δίκτυο S, ένα αντικείμενο A, μία κλάση X και η ερώτηση instance(A, X). Για να απαντηθεί η ερώτηση αρκεί να γίνουν τα εξής:

1. Έστω C το σύνολο των κλάσεων στις οποίες καταλήγουν instance ακμές που ξεκινούν από το A.
2. Αν το C είναι κενό, απάντησε αρνητικά.
3. Διάλεξε μία κλάση c από το C. Αν c = X, απάντησε καταφατικά. Αλλιώς, πρόσθεσε στο C όλες τις κλάσεις στις οποίες καταλήγουν subclass ακμές που ξεκινούν από το c.
4. Πήγαινε στο βήμα 2.

– Πώς μπορεί να απαντηθεί το instance(Hobbes, mammals);
– Πόσο αποδοτικός είναι ο αλγόριθμος;

**Αλγόριθμος Β**

Δεδομένης της μονότονης μετάφρασης του ισοδύναμου με το σημασιολογικό δίκτυο S συστήματος πλαισίων F, για την απάντηση της ερώτησης instance(A, X) προστίθεται στη βάση γνώσης και το instance(A, X)$\rightarrow$F και στη συνέχεια ακολουθούνται τα εξής βήματα:

1. Αν το instance(A, X) βρίσκεται στη βάση, απάντησε καταφατικά. Αλλιώς, εφάρμοσε τον κανόνα της ανάλυσης μεταξύ των instance(A, X) $\rightarrow$F και

   instance(a, y) $\wedge$subclass(y, x) $\rightarrow$ instance(a, x) (9)

   για να προκύψει το:

   instance(A, y) $\wedge$subclass(y, X) $\rightarrow$F (10)

2. Ακολουθώντας τη στρατηγική του συνόλου υποστήριξης, εφάρμοσε τον κανόνα της ανάλυσης μεταξύ του (10) και όλων των προτάσεων της βάσης, εκτός από την (9), για να εξαχθούν προτάσεις της μορφής:

   $$\text{subclass}(c, X) \rightarrow F \qquad (11)$$

   για όλες τις κλάσεις c των οποίων στιγμιότυπο (instance) είναι το A.

3. Διάλεξε μία από τις προτάσεις (11). Αν το subclass(c, X) βρίσκεται στη βάση, το $T \rightarrow F$ συμπεραίνεται, άρα απάντησε καταφατικά. Αλλιώς, εφάρμοσε τον κανόνα της ανάλυσης μεταξύ της (11) και της (2) για να προκύψει το:

   $$\text{subclass}(c, y) \wedge \text{subclass}(y, X) \rightarrow F \qquad (12)$$

   Ακολουθώντας πάλι τη στρατηγική του συνόλου υποστήριξης, εφάρμοσε τον κανόνα της ανάλυσης μεταξύ του (12) και όλων των προτάσεων της βάσης, εκτός από την (2), για να εξαχθούν προτάσεις της μορφής:

   $$\text{subclass}(d, X) \rightarrow F$$

   για όλες τις κλάσεις d των οποίων υποκλάση (subclass) είναι το c.

4. Πήγαινε στο βήμα 3.

- Αποδεικνύεται ότι οι αλγόριθμοι A και B είναι ισοδύναμοι.

- Ο αλγόριθμος B είναι ουσιαστικά μία εξειδίκευση της στρατηγικής του συνόλου υποστήριξης. Η μη εφαρμογή της ανάλυσης στις δύο περιπτώσεις αποδεικνύεται ότι δεν επηρεάζει την πληρότητα του αλγορίθμου.

**ΕΡΩΤΗΣΗ:** Υπάρχει τρόπος να απαντηθεί αποδοτικά η ερώτηση value(A, P, v), π.χ. value(Hobbes, covering, v);

**ΑΠΑΝΤΗΣΗ:** **ΝΑΙ.** Με τον παρακάτω αλγόριθμο Γ.

**Αλγόριθμος Γ**

Έστω ένα σημασιολογικό δίκτυο S, ένα αντικείμενο A, μία ιδιότητα P και η ερώτηση value(A, P, v), όπου το v είναι μεταβλητή. Για να βρεθεί το v αρκούν τα εξής:

1. Αν υπάρχει τιμή για την ιδιότητα P του A, να επιστρέψεις αυτήν την τιμή.
2. Αλλιώς, θέσε C το σύνολο των κλάσεων στις οποίες καταλήγουν instance ακμές που ξεκινούν από το A.
3. Αν το C είναι κενό, απάντησε αρνητικά.
4. Διάλεξε μία κλάση c από το C. Αν υπάρχει τιμή για την ιδιότητα P του c, να επιστρέψεις αυτήν την τιμή. Αλλιώς, πρόσθεσε στο C τις κλάσεις στις οποίες καταλήγουν subclass ακμές που ξεκινούν από το c.
5. Πήγαινε στο βήμα 3.

**ΕΡΩΤΗΣΗ:** Υπάρχει τρόπος να απαντηθεί αποδοτικά η ερώτηση smaller(Hobbes, Moby)**;**

**ΑΠΑΝΤΗΣΗ:** **ΝΑΙ.** Με τη μέθοδο της διάδοσης σημαδιών (marker propagation) ή ενεργοποίησης εξαπλώσεων (spreading activation).

1 = κόμβοι στιγμιότυπα του smaller

2 = κόμβοι των οποίων στιγμιότυπο είναι ο Hobbes

3 = κόμβοι των οποίων στιγμιότυπο είναι ο Moby

Υπάρχει κόμβος 1 με first ιδιότητα κόμβο 2 και second ιδιότητα κόμβο 3**;**

**Συλλογιστική σε μη μονότονα συστήματα πλαισίων**

- Είναι ένα πολύ δύσκολο πρόβλημα που δεν έχει λυθεί μέχρι σήμερα με απόλυτα ικανοποιητικό τρόπο.
- Στις περισσότερες των περιπτώσεων, ένας αλγόριθμος σαν τον Γ έχει αποδεκτή συμπεριφορά.
  - Είναι σε θέση να αποδείξει όλες τις πιθανές ενθουσιώδεις συνέπειες της βασικής "ελλείψει" θεωρίας, π.χ. value(Nixon, politics, hawk) ή value(Nixon, politics, pacifist).
  - Βρίσκει την τιμή μίας ιδιότητας από την πιο ειδική υπερκλάση, δηλαδή value(Tweety, locomotion, flies) και value(Fred, locomotion, walks). Δεν αποδεικνύει όμως (και σωστά) το value(Fred, locomotion, flies).
- Στην αντίστοιχη μεθοδολογία μέσω λογικής χρειάζονται όμως κάποιες προσθήκες.

**ΕΡΩΤΗΣΗ:** Πώς είναι δυνατόν να προτιμηθεί το value(Fred, locomotion, walks) από το value(Fred, locomotion, flies);

**ΑΠΑΝΤΗΣΗ:** Με τις εξής δύο προθήκες:

1) $\text{value}(o, p, v) \wedge \text{different}(v, v') \rightarrow \neg \text{value}(o, p, v')$

   different(walks, flies)

   different(walks, swims)

   different(hawk, pacifist)

   ..............

2) $\text{subclass}(c_1, c_2) \wedge \text{ab}(a, c_1, p) \rightarrow \text{ab}(a, c_2, p)$

Χωρίς αυτό, υπάρχουν δύο πιθανές επεκτάσεις, οι:

$E_1=\{\neg\text{ab}(\text{Fred, ostrich, locomotion})\}$ και $E_2=\{\neg\text{ab}(\text{Fred, bird, locomotion})\}$

Με αυτό, στην $E_2$ προστίθεται και το $\neg$ab(Fred, ostrich, locomotion), οπότε η $E_2$ είναι αντιφατική τώρα, άρα δεν είναι πλέον επέκταση, δηλαδή μόνο η $E_1$ υπάρχει, συνεπώς value(Fred, locomotion, walks).