Τεχνητή Νοημοσύνη

14η διάλεξη (2024-25)

Ίων Ανδρουτσόπουλος

http://www.aueb.gr/users/ion/

Οι διαφάνειες αυτής της διάλεξης βασίζονται σε ύλη του βιβλίου *Artificial Intelligence – A Modern Approach* των S. Russel και P. Norvig, 2η έκδοση, Prentice Hall, 2003, του βιβλίου *Τεχνητή Νοημοσύνη* των Βλαχάβα κ.ά., 3η έκδοση, Β. Γκιούρδας Εκδοτική, 2006 και του βιβλίου *Machine Learning* του T. Mitchell, McGraw-Hill, 1997. Τα περισσότερα σχήματα των διαφανειών βασίζονται σε σχήματα των διαφανειών που συνοδεύουν τα πρώτα δύο βιβλία.

Τι θα ακούσετε σήμερα

- Είδη συλλογιστικής
  - Παραγωγική, απαγωγική, επαγωγική συλλογιστική.
- Μηχανική μάθηση
  - Εισαγωγή.
  - Παράσταση δεδομένων και υποθέσεων.
  - Η μάθηση ως πρόβλημα αναζήτησης.
  - Αλγόριθμος απαλοιφής υποψηφίων.
  - Στοιχεία επαγωγικού λογικού προγραμματισμού.

Είδη συλλογιστικής

- **Παραγωγική** συλλογιστική (deduction).
  - Παραγωγή ορθών συμπερασμάτων με **κανόνες λογικής**.
  - Π.χ. modus ponens, modus tollens, …
  - Π.χ. $\{\forall x_1\ (Dog(x_1) \Rightarrow Animal(x_1)), Dog(Fido)\} \vdash Animal(Fido)$
- **Απαγωγική** συλλογιστική (abduction).
  - Προσπάθεια εύρεσης **πιθανής** υπόθεσης που να εξηγεί παρατηρήσεις. Η υπόθεση μπορεί να **μην ισχύει**.
  - Π.χ. γνώση: $\forall x_1\ (Has(x_1, Grippe) \Rightarrow Fever(x_1, High))$
  - Παρατήρηση: Fever(John, High)
  - Πιθανή εξήγηση: Has(John, Grippe)
  - Πολύ σημαντικός μηχανισμός σε συστήματα **διάγνωσης**.

Είδη συλλογιστικής (συνέχεια)

- **Επαγωγική** συλλογιστική (induction).
  - Προσπάθεια εύρεσης **γενίκευσης**.
  - *Has(Patient323, Grippe), Fever(Patient323, High),* Has(Patient357, Pneumonia), Fever(Patient357, High), *Has(Patient389, Grippe), Fever(Patient389, High), Has(Patient456, Grippe), Fever(Patient456, High),* Has(Patient498, Grippe), Fever(Patient498, Medium).
  - Γενίκευση (αγνοώντας σπάνιες περιπτώσεις):

$$\forall x_1 \ (Has(x_1, Grippe) \Rightarrow Fever(x_1, High))$$

  - Ιδιαίτερα σημαντική στη **μηχανική μάθηση**.

Μηχανική μάθηση

- Η **χειρωνακτική** εισαγωγή γνώσεων σε ένα σύστημα είναι συχνά **δύσκολη**.
  - Δυσκολία/κόστος εκμαίευσης γνώσης από ειδικούς.
  - Δυσκολία προσδιορισμού των απαιτούμενων γνώσεων.
- **Αλγόριθμοι μηχανικής μάθησης**: εξάγουν αυτόματα **νέες γνώσεις** από **εμπειρικά δεδομένα**, βελτιώνοντας έτσι τη συμπεριφορά ενός συστήματος.
  - Π.χ. πρόγραμμα που μαθαίνει να κάνει **ιατρικές διαγνώσεις** από προηγούμενες διαγνώσεις ιατρών και τα αντίστοιχα αποτελέσματα εργαστηριακών εξετάσεων.
  - Π.χ. πρόγραμμα που μαθαίνει να **κατατάσσει κριτικές προϊόντων** σε θετικές, αρνητικές, μικτές, ουδέτερες.

Μερικές από τις πολλές εφαρμογές της ΜΜ

- Αναγνώριση χαρακτήρων σε **χειρόγραφα**.
  - Προγράμματα οπτικής αναγνώρισης χαρακτήρων (OCR).
  - Προγράμματα αναγνώρισης χαρακτήρων για κινητά.
- Αναγνώριση **φωνής**.
  - Αναγνώριση προφορικών εντολών (π.χ. στο αυτοκίνητο).
  - Μετατροπή συνεχούς ομιλίας σε γραπτή μορφή (υπαγόρευση).
- Προγράμματα που παίζουν **τάβλι**, **σκάκι**, **Go**, …
  - TD-GAMMON (Tesauro 1992, 95): εκπαιδεύτηκε παίζοντας πάνω από 1 εκατομμύριο παρτίδες με τον εαυτό του.
  - AlphaGo (Google DeepMind, https://en.wikipedia.org/wiki/AlphaGo).
- **Οδήγηση** αυτοκινήτου.
  - ALVINN (Pomerlau 1989): μετά από εκπαίδευση οδήγησε σε αυτοκινητόδρομο επί 90 μίλια με ταχύτητα 70 μιλίων/ώρα.
  - Self-driving cars (https://waymo.com/).
- **Μεγάλα γλωσσικά μοντέλα (LLMs)**.
  - Π.χ. Chat-GPT, Llama, Gemini, …

Επιβλεπόμενη/μη επιβλεπόμενη μάθηση

- Αλγόριθμοι **επιβλεπόμενης μάθησης**.
  - Προϋποθέτουν ότι υπάρχουν **παραδείγματα εκπαίδευσης** για τα οποία είναι γνωστές (ή μπορούν να αποκτηθούν) οι **ορθές απαντήσεις**.
  - Π.χ. σύστημα που μαθαίνει να κάνει **ιατρικές διαγνώσεις** από προηγούμενες περιπτώσεις ασθενών και διαγνώσεων.
- Αλγόριθμοι **μη επιβλεπόμενης μάθησης**.
  - Προσπαθούν να ανακαλύψουν νέες γνώσεις από δεδομένα που **δεν περιέχουν τις επιθυμητές απαντήσεις**.
  - Π.χ. οργάνωση πελατών ή ειδήσεων σε **ομάδες (clusters)** ή εύρεση συσχετισμών της μορφής «οι πελάτες που αγοράζουν το X την Παρασκευή αγοράζουν και ...».
  - Πολλές εφαρμογές στην **επιστήμη δεδομένων** (παρακολουθήστε το αντίστοιχο μάθημα).

Παραδείγματα εκπαίδευσης

- Τραπεζικό σύστημα που θα αποφασίζει αν πρέπει να δοθεί **δάνειο** σε έναν πελάτη.

ιδιότητες

Πελάτης	Οφειλές	Εισόδημα	Παντρεμένος	Καλός;
1	Υψηλές (1)	Υψηλό (1)	Ναι (1)	Καλός (1)
2	Χαμηλές (0)	Υψηλό (1)	Όχι (0)	Κακός (0)
3	Χαμηλές (0)	Υψηλό (1)	Ναι (1)	Καλός (1)
4	Υψηλές (1)	Χαμηλό (0)	Ναι (1)	Κακός (0)
5	Χαμηλές (0)	Χαμηλό (0)	Ναι (1)	Κακός (0)

- **Διανυσματική** παράσταση εμπειρίας:

  { <1, 1, 1, **1**>, <0, 1, 1, **1**>, <0, 1, 0, **0**>, <1, 0, 1, **0**>, <0, 0, 1, **0**> }

επιθυμητές απαντήσεις

Η μάθηση ως πρόβλημα αναζήτησης

- Αναζήτηση μιας **συνάρτησης** («**υπόθεσης**»).
  - h(x, y, z) = ?
  - h: {0, 1} × {0, 1} × {0, 1} → {0, 1}
  - γενικότερα **h: $D_1 \times D_2 \times \ldots \times D_n \rightarrow C$**
  - **D_i**: οι δυνατές τιμές της i-στής **ιδιότητας**.
  - **C**: οι δυνατές **απαντήσεις**.
- Αναζήτηση σε ένα **χώρο συναρτήσεων**.
  - Θεωρούμε ότι τα δεδομένα εκπαίδευσης είναι ένα **δείγμα** από έναν πληθυσμό που παράγεται σύμφωνα με μια **συγκεκριμένη άγνωστη συνάρτηση**.
  - Αναζητούμε αυτή τη συνάρτηση μέσα σε ένα **χώρο συναρτήσεων**.

Χώρος αναζήτησης

- Ο χώρος αναζήτησης εξαρτάται από τις **ιδιότητες**.
  - **Ποιες ιδιότητες** θα χρησιμοποιήσουμε;
  - Αντιστοιχούν στα ορίσματα των υποθέσεων.
  - Π.χ. οικογ. κατάσταση, οφειλές, ηλικία, επάγγελμα, ύψος;
  - **Τιμές** από πεπερασμένο σύνολο, πραγματικοί αριθμοί;
- Και από το **μοντέλο παράστασης** των υποθέσεων.
  - Π.χ. **γραμμική συνάρτηση** των ιδιοτήτων;
  - Περιορίζουμε την αναζήτηση στις (π.χ. γραμμικές) **υποθέσεις που μπορεί να παραστήσει το μοντέλο**.
  - Δεν μπορούμε να μάθουμε μια υπόθεση που δεν μπορεί να παρασταθεί από το μοντέλο που διαλέξαμε.
  - Οι περιορισμοί του μοντέλου, όμως, **μειώνουν το μέγεθος του χώρου αναζήτησης**.
  - Δυστυχώς συχνά δεν ξέρουμε τι περιορισμούς να επιβάλουμε.

Μοντέλο παράστασης υποθέσεων

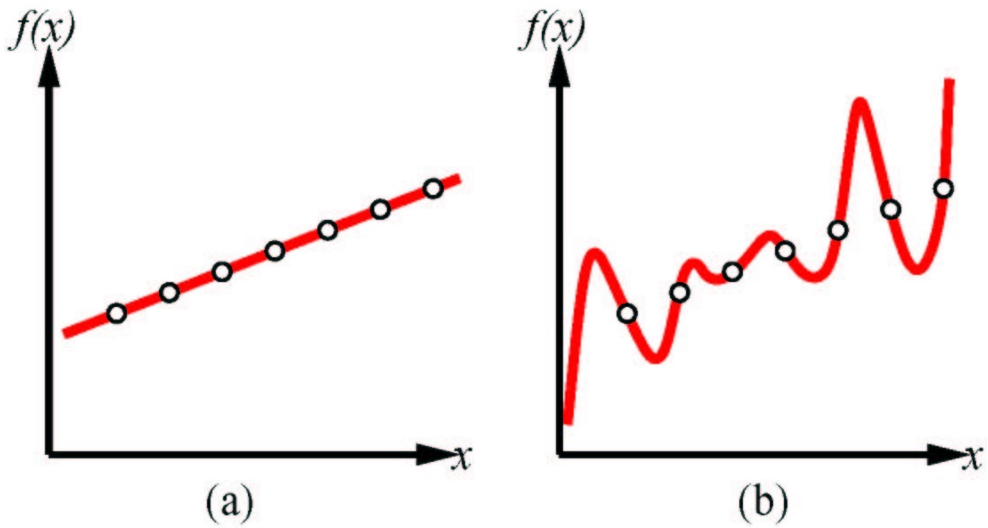
- Απλοϊκό παράδειγμα για το πρόβλημα των δανείων:
  - Παριστάνουμε κάθε συνάρτηση-υπόθεση με μια **τριάδα** που αντιστοιχεί στις περιπτώσεις που δίνουμε δάνειο.
  - Π.χ. $\mathbf{h_1 = <0, ?, 1}>$, $\mathbf{h_2 = <1, 1, 1>}$
  - «?» σημαίνει για **οποιαδήποτε** τιμή.

Αν χαμηλές οφειλές και παντρεμένος/η, δώσε δάνειο ανεξαρτήτως εισοδήματος. Διαφορετικά μη δώσεις

Αν υψηλές οφειλές και υψηλό εισόδημα και παντρεμένος/η, δώσε δάνειο. Διαφορετικά μη δώσεις.

- Αδύνατον, όμως, να μάθουμε την $(\mathbf{h_1} \vee \mathbf{h_2})$, γιατί δεν περιλαμβάνεται στο χώρο αναζήτησης.

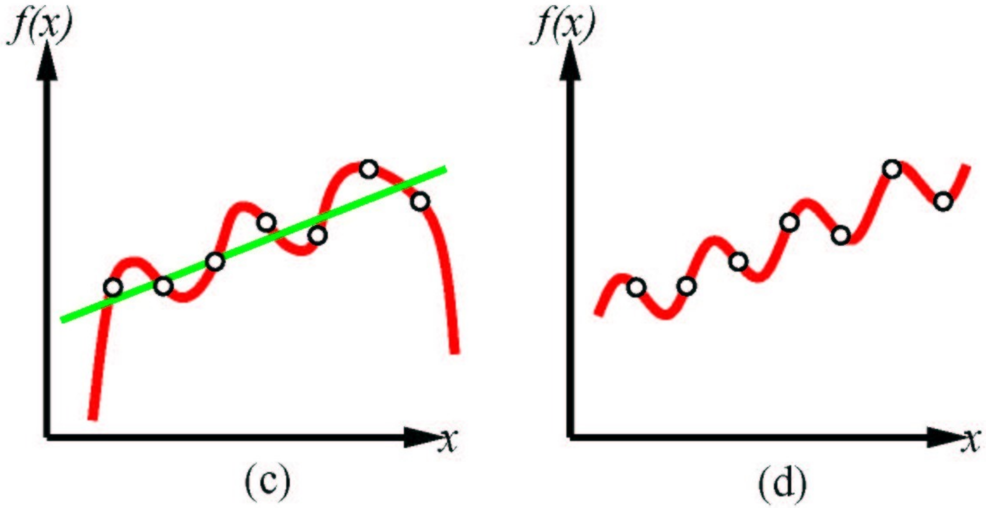
Συνεπείς υποθέσεις και γενίκευση



Οι τελείες αντιστοιχούν σε δεδομένα εκπαίδευσης. Εδώ έχουμε μία ιδιότητα και οι τιμές των υποθέσεων είναι πραγματικοί. Έστω ότι αναζητούμε πολυωνυμικές υποθέσεις.

- **Συνεπείς υποθέσεις**: συμφωνούν με τα δεδομένα εκπαίδευσης.
  - (a): πολυώνυμο 1ου βαθμού. (b): πολυώνυμο μεγαλύτερου βαθμού.
- **Ξυράφι του Ockham**: Προτιμότερη είναι η **απλούστερη** συνεπής υπόθεση, εδώ το πολυώνυμο μικρότερου βαθμού.
  - Απλούστερες υποθέσεις είναι ευκολότερο να κατασκευαστούν (π.χ. γραμμική παρεμβολή) και οι αποκρίσεις τους υπολογίζονται ευκολότερα.
- **Ικανότητα γενίκευσης**: Να δίνει σωστές απαντήσεις και για νέες περιπτώσεις που δεν ανήκουν στα παραδείγματα εκπαίδευσης.
  - Απλούστερες υποθέσεις συνήθως γενικεύουν καλύτερα τις παρατηρήσεις.

Ασυνεπείς υποθέσεις ίσως είναι προτιμότερες



Τα δεδομένα εκπαίδευσης ακολουθούν μια συνάρτηση της μορφής **a·x + b + c·sinx** (διάγραμμα d), που είναι αδύνατον να παρασταθεί ως πολυώνυμο πεπερασμένου βαθμού.

- Αν περιορίσουμε την αναζήτηση σε **πολυώνυμα πεπερασμένου βαθμού** (έστω ≤ k), η συνάρτηση-στόχος του (d) **δεν περιλαμβάνεται** στο χώρο αναζήτησης.
- Αν επιμείνουμε στην εξεύρεση **συνεπούς** υπόθεσης, καταλήγουμε στο **πολυώνυμο 6ου βαθμού** της κόκκινης γραμμής του (c), που **δεν επιτυγχάνει καλή γενίκευση**.
  - Μας ενδιαφέρει οι προβλέψεις να είναι καλές για όλα τα x.
- Οι προβλέψεις της **ευθείας γραμμής** (c) είναι εν γένει **πιο κοντά** στις επιθυμητές τιμές (d), αν και η (c) είναι **ασυνεπής**.

Γενικότερες/ειδικότερες υποθέσεις

- Επιστρέφουμε στην περίπτωση της τράπεζας, όπου οι **τιμές των υποθέσεων** είναι **0 ή 1** και οι **ιδιότητες** έχουν **διακριτές** τιμές.
- Η **h_2 είναι ειδικότερη της h_1** ανν $h_1 \neq h_2$ (διαφορετικές συναρτήσεις) και:
  - **αν $h_2(x_1, ..., x_n) = 1$, τότε $h_1(x_1, ..., x_n) = 1$.**
  - Π.χ. αν δίνει δάνειο η h_2, τότε δίνει και η h_1.
- Η h_1 είναι **γενικότερη** της h_2 ανν η h_2 είναι ειδικότερη της h_1.

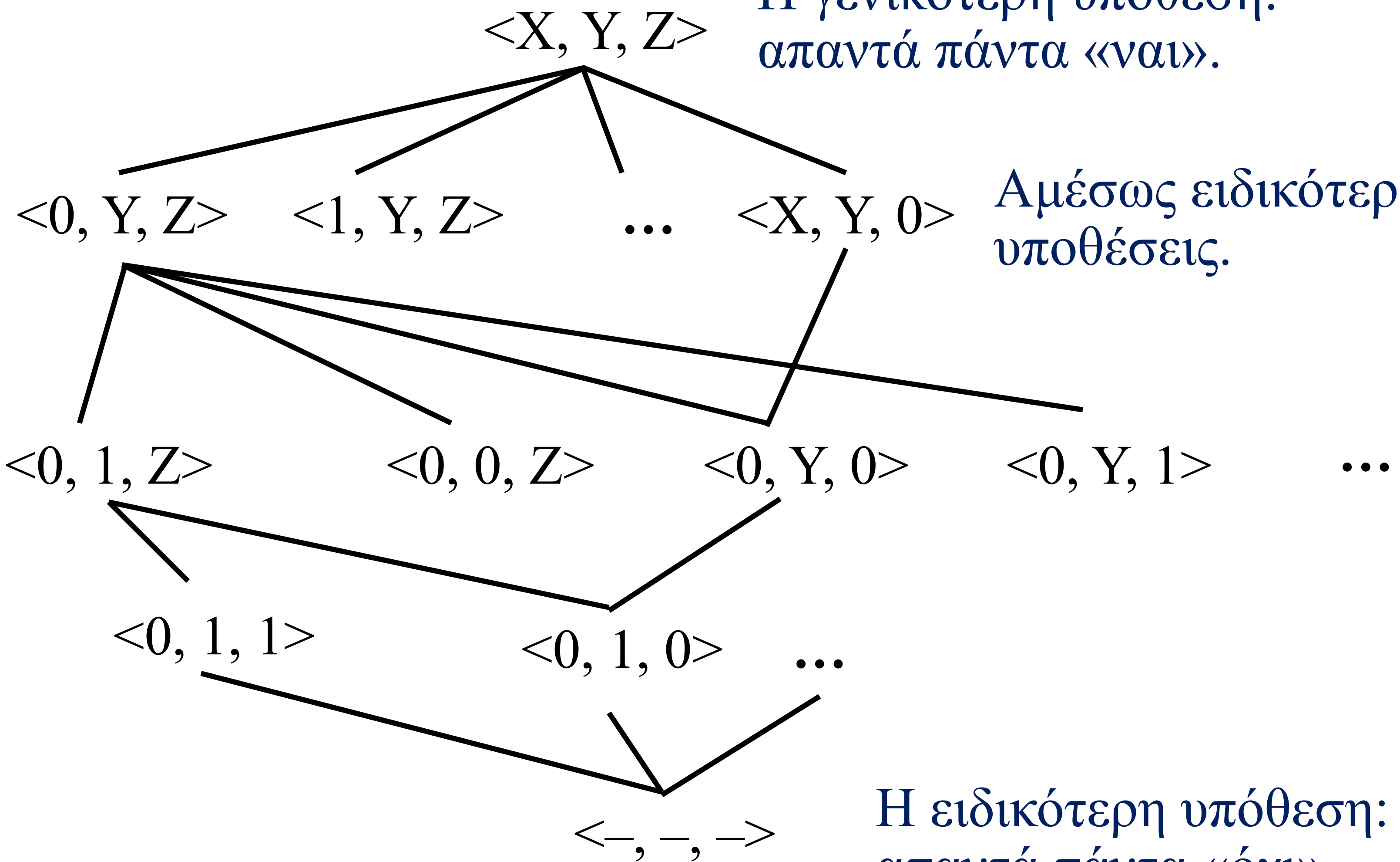
<X, Y, Z> Η γενικότερη υπόθεση: απαντά πάντα «ναι»

<0, Y, Z> <1, Y, Z> … <X, Y, 0> Αμέσως ειδικότερες υποθέσεις.

Εδώ χρησιμοποιούμε μεταβλητές αντί για «?».

Χώρος αναζήτησης

Η γενικότερη υπόθεση:
απαντά πάντα «ναι».



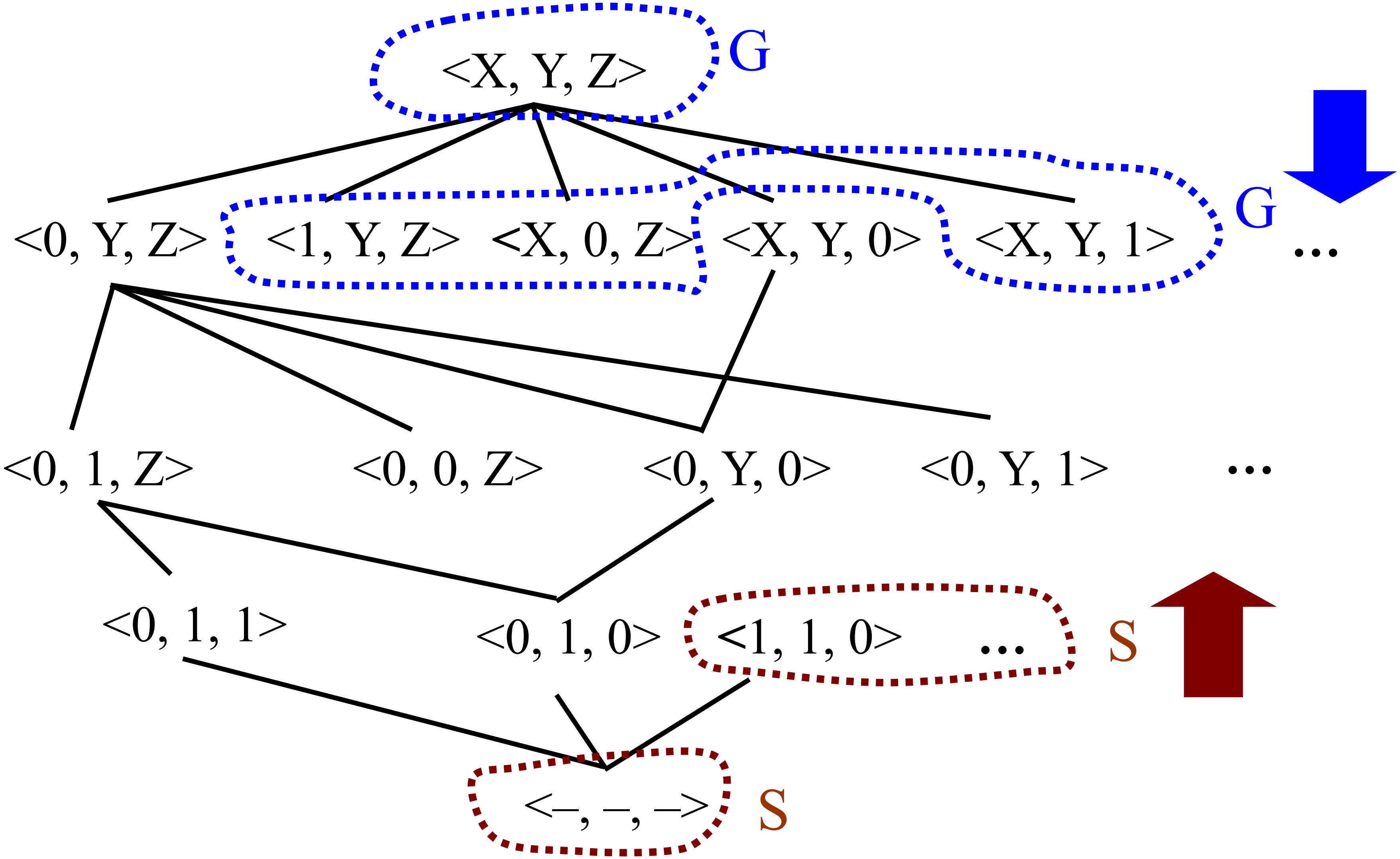
Αμέσως ειδικότερες
υποθέσεις.

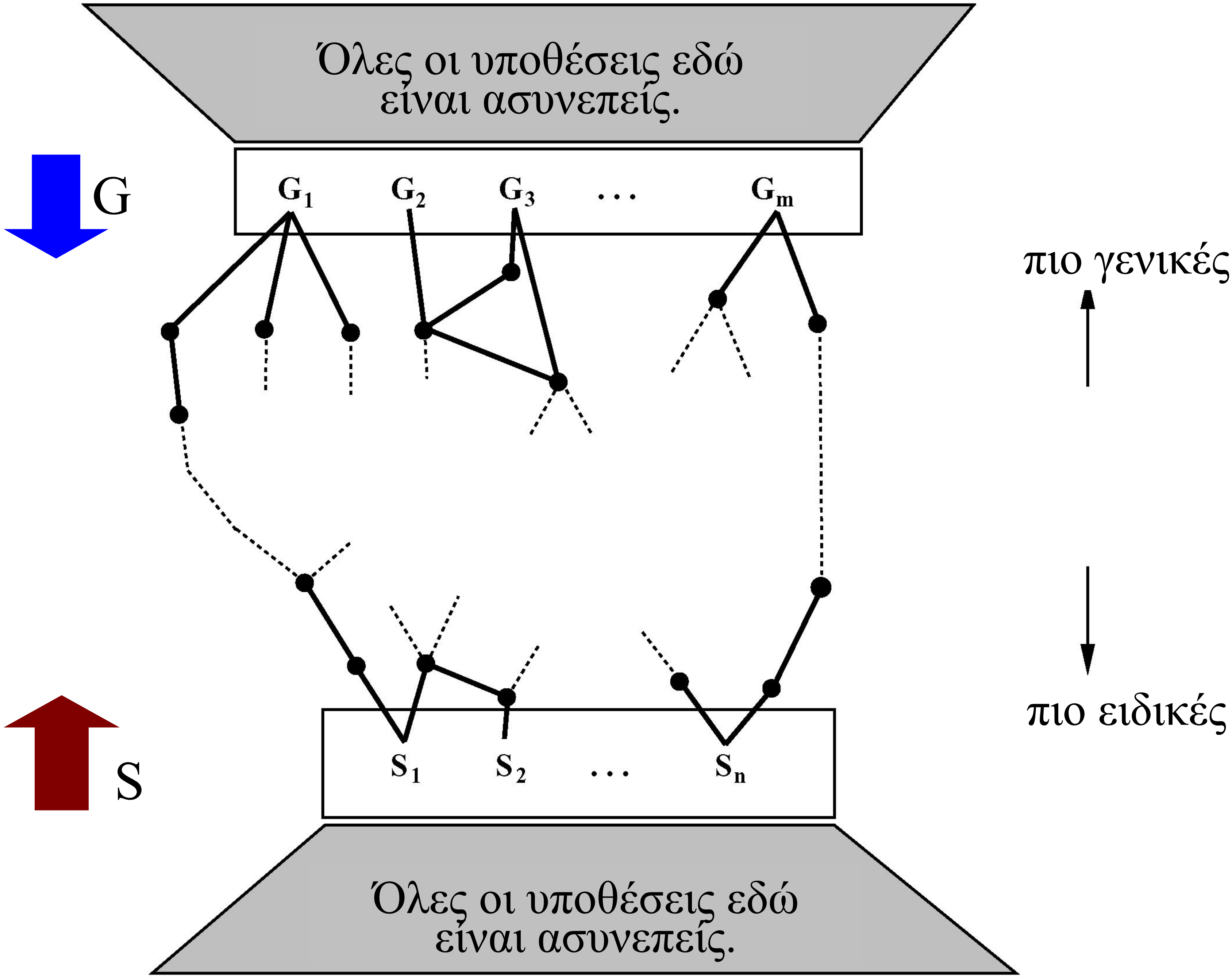
Η ειδικότερη υπόθεση:
απαντά πάντα «όχι».

Αλγόριθμος απαλοιφής υποψηφίων

- **Σύνολα G** και **S**: περιέχουν **υποθέσεις**.
- ***Αρχικά*** $G = \{ <X_1, X_2, \ldots, X_n> \}$, $S = \{ <-, \ldots, -> \}$
- Κρατάμε στο **G** τις **πιο γενικές** υποθέσεις που είναι **συνεπείς** με τα παραδείγματα εκπαίδευσης που έχουμε συναντήσει (τα κοιτάζουμε ένα-ένα).
- Κρατάμε στο **S** τις **πιο ειδικές** υποθέσεις που είναι **συνεπείς** με τα παραδείγματα εκπαίδευσης που έχουμε συναντήσει.

Εξέλιξη των G και S




Όλες οι υποθέσεις εδώ είναι ασυνεπείς.
G
G_1 G_2 G_3 … G_m
πιο γενικές
πιο ειδικές
S
S_1 S_2 … S_n
Όλες οι υποθέσεις εδώ είναι ασυνεπείς.

Αλγόριθμος απαλοιφής υποψηφίων

- Για κάθε νέο **θετικό** παράδειγμα d:
  - **Αφαίρεσε από το G** τις υποθέσεις που είναι **ασυνεπείς** με το d.
  - Για κάθε υπόθεση s του S που είναι **ασυνεπής** με το d:
    - **Αφαίρεσε** την s **από το S**.
    - **Πρόσθεσε στο S** κάθε **ελάχιστη γενίκευση** h της s, για την οποία ισχύει ότι:
      - Η h είναι **συνεπής** με το d.
      - Η h είναι **ειδικότερη ή ίδια με μια υπόθεση του G**. (Ξέρουμε ότι η υπόθεση που ψάχνουμε είτε υπάρχει ήδη στο G είτε είναι ειδικότερη μιας υπόθεσης του G.)
  - **Αφαίρεσε** από το S **κάθε υπόθεση** που είναι **γενικότερη από μια άλλη υπόθεση** του S.

Αλγόριθμος απαλοιφής υποψηφίων

- Για κάθε νέο **αρνητικό** παράδειγμα d:
  - **Αφαίρεσε από το S** τις υποθέσεις που είναι ασυνεπείς με το d.
  - Για κάθε υπόθεση g του G που είναι ασυνεπής με το d:
    - **Αφαίρεσε** την g **από το G**.
    - **Πρόσθεσε στο G** κάθε **ελάχιστη ειδίκευση** h της g, για την οποία ισχύει ότι:
      - Η h είναι **συνεπής** με το d.
      - Η h είναι **γενικότερη ή ίδια με μια υπόθεση του S**.
        (Ξέρουμε ότι η υπόθεση που ψάχνουμε είτε υπάρχει ήδη στο S είτε είναι γενικότερη μιας υπόθεσης του S.)
  - **Αφαίρεσε** από το G **κάθε υπόθεση** που είναι **ειδικότερη από μια άλλη υπόθεση** του G.

Παράδειγμα χρήσης ΑΑΥ

G: {(X,Y,Z)}
S: {<-, -, ->}

(Υψηλές, Υψηλό, Ναι) δώσε

G: {(X,Y,Z)}
S: {(Υψηλές, Υψηλό, Ναι)}

(Χαμηλές, Υψηλό, Όχι) μη δώσεις

G: {(Υψηλές,Y,Z), ~~(X, Χαμηλό,Z)~~, (X,Y, Ναι) }
S: {(Υψηλές, Υψηλό, Ναι)}

(Χαμηλές, Υψηλό, Ναι) δώσε

G: {(X,Y, Ναι) }
S: {(X, Υψηλό, Ναι)}

(Υψηλές, Χαμηλό, Ναι) μη δώσεις

G: {~~(Χαμηλές,Y, Ναι)~~, (X, Υψηλό, Ναι)}
S: {(X, Υψηλό, Ναι)}

(Χαμηλές, Χαμηλό, Ναι) μη δώσεις

G: {(X, Υψηλό, Ναι)}
S: {(X, Υψηλό, Ναι)}

Χαρακτηριστικά ΑΑΥ

- **Σύγκλιση**: G = S = { h }
  - Αν δεν υπάρχουν **ασυνεπή** παραδείγματα (ίδιες τιμές ιδιοτήτων αλλά διαφορετικές «ορθές» απαντήσεις)
  - και αν η **συνάρτηση-στόχος h μπορεί να παρασταθεί** με τον τρόπο που διαλέξαμε (και είναι μοναδική)
  - και αν υπάρχουν **αρκετά παραδείγματα** εκπαίδευσης,
  - τότε τα G και S **συγκλίνουν** στο **{ h }**.
- **Τι γίνεται** αν
  - υπάρχουν ασυνεπή παραδείγματα ή
  - η συνάρτηση στόχος δεν μπορεί να παρασταθεί με τον τρόπο που διαλέξαμε;
  - Καταλήγουμε (με αρκετά παραδείγματα): **G = S = { }**.

Παράδειγμα πρόβλεψης συμπεριφοράς

Example	Attributes										Target
	Alt	*Bar*	*Fri*	*Hun*	*Pat*	*Price*	*Rain*	*Res*	*Type*	*Est*	*WillWait*
X_1	*T*	*F*	*F*	*T*	*Some*	*$$$*	*F*	*T*	*French*	*0–10*	*T*
X_2	*T*	*F*	*F*	*T*	*Full*	*$*	*F*	*F*	*Thai*	*30–60*	*F*
X_3	*F*	*T*	*F*	*F*	*Some*	*$*	*F*	*F*	*Burger*	*0–10*	*T*
X_4	*T*	*F*	*T*	*T*	*Full*	*$*	*F*	*F*	*Thai*	*10–30*	*T*
X_5	*T*	*F*	*T*	*F*	*Full*	*$$$*	*F*	*T*	*French*	*>60*	*F*
X_6	*F*	*T*	*F*	*T*	*Some*	*$$*	*T*	*T*	*Italian*	*0–10*	*T*
X_7	*F*	*T*	*F*	*F*	*None*	*$*	*T*	*F*	*Burger*	*0–10*	*F*
X_8	*F*	*F*	*F*	*T*	*Some*	*$$*	*T*	*T*	*Thai*	*0–10*	*T*
X_9	*F*	*T*	*T*	*F*	*Full*	*$*	*T*	*F*	*Burger*	*>60*	*F*
X_{10}	*T*	*T*	*T*	*T*	*Full*	*$$$*	*F*	*T*	*Italian*	*10–30*	*F*
X_{11}	*F*	*F*	*F*	*F*	*None*	*$*	*F*	*F*	*Thai*	*0–10*	*F*
X_{12}	*T*	*T*	*T*	*T*	*Full*	*$*	*F*	*F*	*Burger*	*30–60*	*T*

- Οι **περιγραφές των παραδειγμάτων** εκπαίδευσης και οι **αποφάσεις** μπορούν να παρασταθούν **σε ΠΚΛ**:
  - Τα X_i εδώ είναι σταθερές.
  - $Alt(X_1) \wedge \neg Bar(X_1) \wedge \neg Fri(X_1) \wedge \ldots \wedge Pat(X_1, Some) \wedge \ldots$
  - $Alt(X_{12}) \wedge Bar(X_{12}) \wedge Fri(X_{12}) \wedge \ldots \wedge Pat(X_{12}, Some) \wedge \ldots$
  - $WillWait(X_1), \neg WillWait(X_2), \ldots, WillWait(X_{12})$

Μάθηση και λογική

- Αν προσπαθούμε να μάθουμε την έννοια **Q(x)**, τότε ψάχνουμε μία υπόθεση της μορφής:

$$\forall x\ (\mathbf{Q(x)} \Leftrightarrow \varphi(x))$$

  που να προβλέπει σωστά σε ποιες περιπτώσεις αληθεύει ή όχι η Q(x).

- Στο παράδειγμα του εστιατορίου:

$\forall x$ (**WillWait(x)** $\Leftrightarrow$ (Pat(x, Some) $\vee$
(Pat(x, Full) $\wedge$ Hungry(x) $\wedge$ Type(x, French)) $\vee$
(Pat(x, Full) $\wedge$ Hungry(x) $\wedge$ Type(x, Thai) $\wedge$ Fri(x)) $\vee$
(Pat(x, Full) $\wedge$ Hungry(x) $\wedge$ Type(x, Burger))))

Μια υπόθεση (υποψήφιος τύπος).

Μάθηση και λογική

- Ακριβέστερα θέλουμε να μάθουμε μια υπόθεση, ώστε:

(**Υπόθεση** ∧ **Περιγραφές**) ⊨ **Αποφάσεις**

περιγραφές και **αποφάσεις** για τα παραδείγματα εκπαίδευσης

$$(\forall x\ (\mathbf{WillWait(x)} \Leftrightarrow (Pat(x, Some) \vee (Pat(x, Full) \wedge Fri(x))))$$
$$\wedge\ Pat(X_1, Some) \wedge Pat(X_2, None) \wedge Pat(X_3, Full) \wedge Fri(X_3) \ldots) \models$$
$$(WillWait(X_1) \wedge \neg WillWait(X_2) \wedge WillWait(X_3) \ldots)$$

- Αν δεν θέσουμε κανένα περιορισμό στη λογική μορφή της υπόθεσης, τότε **κινδυνεύουμε να μάθουμε**:

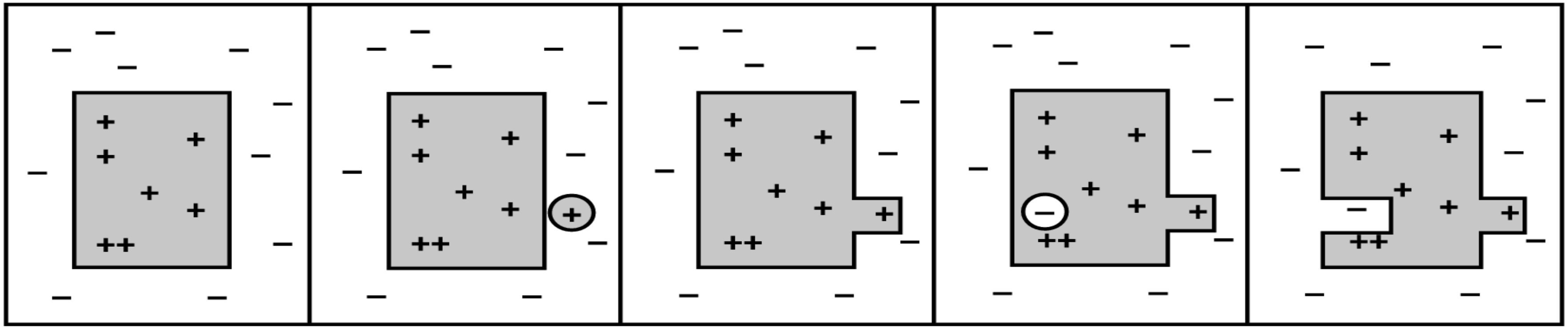
Υπόθεση = **Αποφάσεις**

Απομνημόνευση αποφάσεων εκπαίδευσης. Τεράστιος τύπος αν έχουμε πολλά παραδείγματα εκπαίδευσης.

που δεν έχει καμία ικανότητα γενίκευσης.

- Στην πράξη **περιορίζουμε το χώρο αναζήτησης** (π.χ. υποσύνολο λογικής) και **προτιμούμε τις απλούστερες** (συντομότερες) υποθέσεις (ξυράφι του Ockham).

Λαίμαργη προσέγγιση



- Έχουμε ανά πάσα στιγμή **μία μόνο υπόθεση** που συμφωνεί με τα παραδείγματα που έχουμε συναντήσει.
- Αν κατατάξουμε ένα νέο παράδειγμα **εσφαλμένα** ως **αρνητικό**, **γενικεύουμε** την υπόθεση, ώστε να είμαστε **συνεπείς** με όλα τα παραδείγματα που έχουμε συναντήσει.
  - Π.χ. $\forall x$ (**WillWait(x)** $\Leftrightarrow$ (Pat(x, Some) $\vee$ (Pat(x, Full) $\wedge$ Fri(x)))) **γενικότερη** της $\forall x$ (**WillWait(x)** $\Leftrightarrow$ Pat(x, Some))
  - $\forall x$ ($\mathbf{Q(x)} \Leftrightarrow \varphi_1(x)$) είναι **γενικότερη** από την $\forall x$ ($\mathbf{Q(x)} \Leftrightarrow \varphi_2(x)$) σημαίνει $\forall \mathbf{x}\ (\boldsymbol{\varphi_2(x) \Rightarrow \varphi_1(x)})$.
- Αν κατατάξουμε **εσφαλμένα** ως **θετικό**, **ειδικεύουμε**...

Στο παράδειγμα του εστιατορίου...

- Μια πιθανή υπόθεση **μετά το 1ο** παράδειγμα:

  $\forall x$ (**WillWait(x)** $\Leftrightarrow$ Alt(x))
- **Μετά το 2ο** παράδειγμα (κατατασσόταν εσφαλμένα ως θετικό):

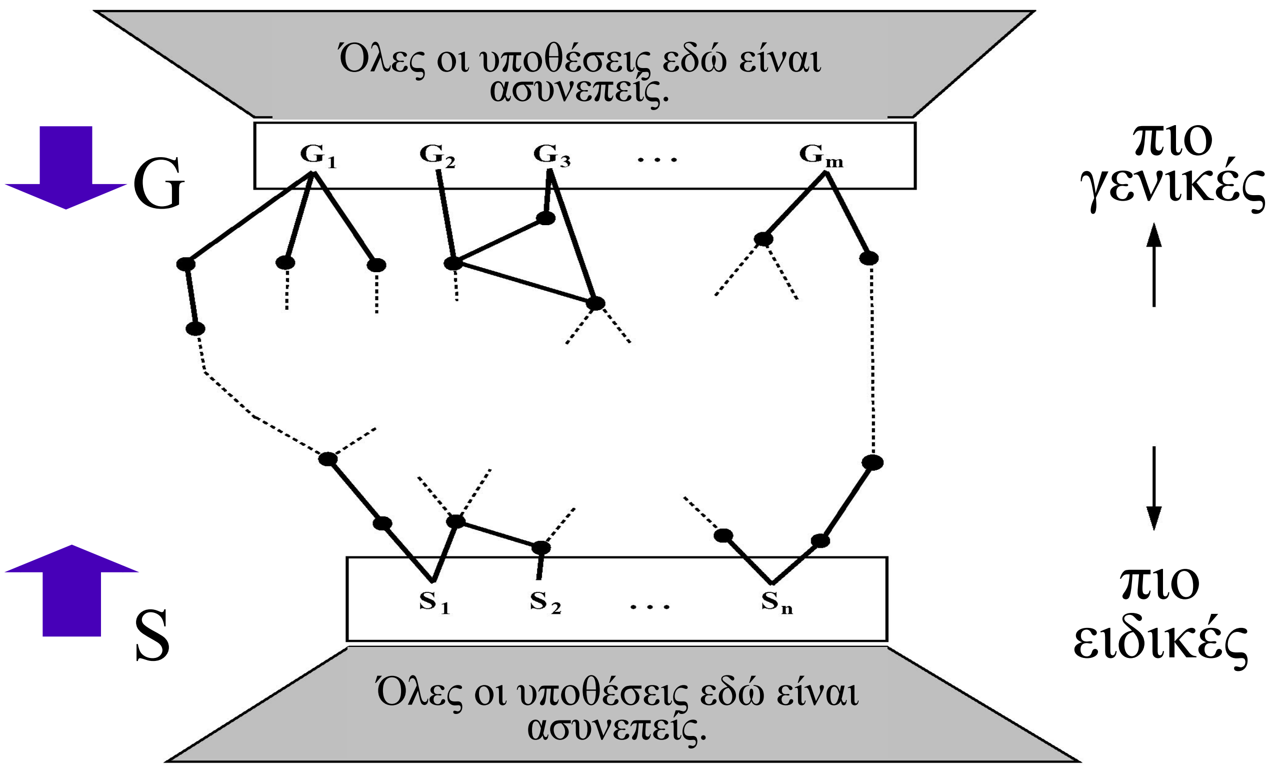
  $\forall x$ (**WillWait(x)** $\Leftrightarrow$ (Alt(x) $\wedge$ Pat(x, Some)))
- **Μετά το 3ο** παράδειγμα (εσφαλμένα ως αρνητικό):

  $\forall x$ (**WillWait(x)** $\Leftrightarrow$ Pat(x, Some))
- **Μετά το 4ο** παράδειγμα (εσφαλμένα ως αρνητικό):

  $\forall x$ (**WillWait(x)** $\Leftrightarrow$ (Pat(x, Some) $\vee$ (Pat(x, Full) $\wedge$ Fri(x))))
- Σε κάθε βήμα υπάρχουν εν γένει **πολλές επιλογές**.
- Μπορεί κάνοντας **μια επιλογή** να **παγιδευτούμε** σε **τρέχουσα υπόθεση** που να **είναι αδύνατον** να γίνει **συνεπής με επόμενο παράδειγμα** παραμένοντας **συνεπής με τα προηγούμενα**.
  - Καλύτερα να χρησιμοποιήσουμε τον ΑΑΥ (βλ. παρακάτω).

Αλγόριθμος απαλοιφής υποψηφίων

- Γενικότερα μπορούμε να χρησιμοποιήσουμε πάλι τον αλγόριθμο απαλοιφής υποψηφίων.
  - Αντί να κρατάμε μία μόνο **υπόθεση**, κρατάμε τα **άνω και κάτω «φράγματα»** G και S του **χώρου υποθέσεων**.
  - Οι **υποθέσεις** είναι τύποι της μορφής $\forall x\ (\mathbf{Q(x)} \Leftrightarrow \varphi(x))$.



Δεν κινδυνεύουμε να **παγιδευτούμε**.

Δεν χρειάζεται να επανεξετάζουμε τα **προηγούμενα** παραδείγματα.

Μπορεί όμως τα G και S να είναι πολύ **μεγάλα**.

Επαγωγικός λογικός προγραμματισμός (ILP)

Προαιρετική μελέτη

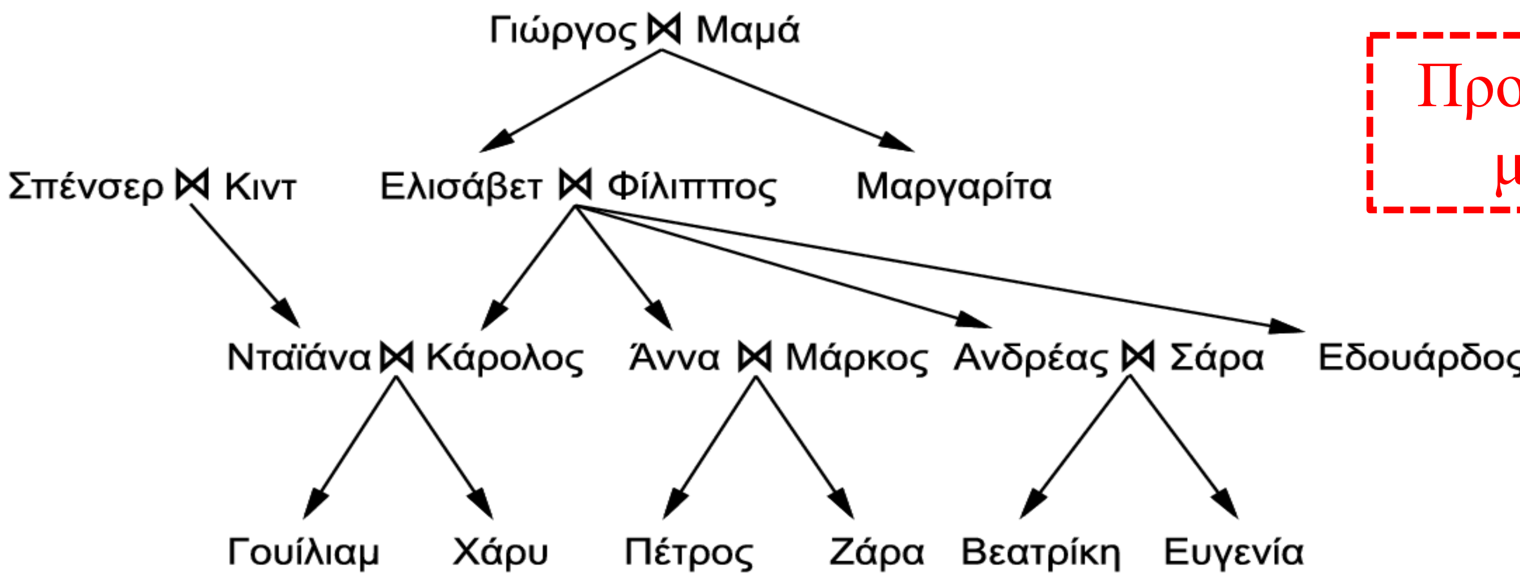
- Προσθέτει και το **Υπόβαθρο** (π.χ. γνώσεις για τον κόσμο):

  (Υπόβαθρο ∧ Υπόθεση ∧ Περιγραφές) ⊨ Αποφάσεις

- Οι υποθέσεις (και το υπόβαθρο, οι περιγραφές, οι αποφάσεις) διατυπώνονται συνήθως σε **ΠΚΛ**.
  - Οι υποθέσεις που προκύπτουν (τύποι ΠΚΛ) είναι **κατανοητές από τους ανθρώπους** (π.χ. βιολόγους), οι οποίου μπορούν να τις συζητήσουν, να τις επεκτείνουν, να τις δημοσιεύσουν κ.λπ.
  - Υπάρχουν και τρόποι να **μεταφράζουμε** αυτόματα τύπους λογικής σε **φυσική γλώσσα** (π.χ. https://arxiv.org/abs/1405.6164).
  - Είναι εύκολο να ενσωματωθεί **προηγούμενη ανθρώπινη γνώση** (π.χ. διατυπωμένη από ειδικούς σε ΠΚΛ).
- **Μειονέκτημα**: πολύ μεγάλος **χώρος αναζήτησης**.

Άλλο παράδειγμα: μάθηση έννοιας GrandParent

Προαιρετική μελέτη



- **Περιγραφές** παραδειγμάτων εκπαίδευσης:
  - Πατέρας(Φίλιππος, Κάρολος), Πατέρας(Φίλιππος, Άννα), ...
  - Μητέρα(Μαμά, Μαργαρίτα), Μητέρα(Μαμά, Ελισάβετ), ...
  - Γάμος(Νταϊάνα, Κάρολος), Γάμος(Ελισάβετ, Φίλιππος), ...
  - Αρσενικό(Φίλιππος), Αρσενικό(Κάρολος), ...
- **Αποφάσεις** για τα παραδείγματα εκπαίδευσης:
  - GrandParent(Μαμά, Κάρολος), GrandParent(Ελισάβετ, Βεατρίκη),
  - ¬GrandParent(Μαμά, Χάρυ), ¬GrandParent(Σπένσερ, Πέτρος), ...

Υπόβαθρο στην GrandParent

Προαιρετική μελέτη

- **Χωρίς** **υπόβαθρο**, ένα σύστημα θα μπορούσε π.χ. να μάθει:

$$(GrandParent(x, y) \Leftrightarrow (\exists z\, (Μητέρα(x, z) \land Μητέρα(z, y)) \lor \exists z\, (Μητέρα(x, z) \land Πατέρας(z, y)) \lor \exists z\, (Πατέρας(x, z) \land Μητέρα(z, y)) \lor \exists z\, (Πατέρας(x, z) \land Πατέρας(z, y))))$$

- **Αν το** **υπόβαθρο** περιείχε τον ορισμό:

  $(Γονέας(x, y) \Leftrightarrow (Μητέρα\ (x, y) \lor Πατέρας\ (x, y)))$

  τότε το σύστημα θα μπορούσε να μάθει:

  $(GrandParent(x, y) \Leftrightarrow \exists z\ (Γονέας\ (x, z) \land Γονέας\ (z, y)))$

- Το υπόβαθρο γνώσεων μπορεί να οδηγήσει σε **συντομότερους ορισμούς**.
- Ορισμένοι αλγόριθμοι ILP έχουν την ικανότητα να κατασκευάζουν **νέα κατηγορήματα** (π.χ. γονέας).

Αλγόριθμος FOIL

Προαιρετική μελέτη

- Μαθαίνει (λαίμαργα) ένα σύνολο **προτάσεων Horn**.

  π.χ. $GrandParent(x, y) \Leftarrow (Μητέρα(x, z) \wedge Μητέρα(z, y))$

  $GrandParent(x, y) \Leftarrow (Μητέρα(x, z) \wedge Πατέρας(z, y))$

  $GrandParent(x, y) \Leftarrow (Πατέρας(x, z) \wedge Μητέρα(z, y))$

  $GrandParent(x, y) \Leftarrow (Πατέρας(x, z) \wedge Πατέρας(z, y))$

  – Επιτρέπονται και **αρνήσεις** στο σώμα των κανόνων.
- Ξεκινάμε με έναν **πολύ γενικό** ορισμό

  π.χ. $GrandParent(x, y) \Leftarrow \text{True}$

  Κάθε x είναι πρόγονος του κάθε y.

  τον οποίο σταδιακά **ειδικεύουμε**.

FOIL – συνέχεια

Προαιρετική μελέτη

- **Προσθέτει διαρκώς συνθήκες** στην τρέχουσα πρόταση Horn, μέχρι να γίνει **συνεπής με όλα τα αρνητικά παραδείγματα** και **μερικά θετικά**.

  π.χ. $GrandParent(x, y) \Leftarrow Μητέρα(x, z)$ και μετά:

  $\boldsymbol{GrandParent(x, y) \Leftarrow (Μητέρα(x, z) \wedge Μητέρα(z, y))}$
- **Αφαιρεί τα θετικά** παραδείγματα **που κάλυψε** η προηγούμενη πρόταση Horn.
- Αρχίζει να κατασκευάζει **νέα πρόταση Horn**, ξεκινώντας πάλι από την πλέον γενική και προσθέτοντας άλλες συνθήκες, μέχρι να γίνει συνεπής με **όλα τα αρνητικά παραδείγματα** και **μερικά θετικά**.

  π.χ. $GrandParent(x, y) \Leftarrow Μητέρα(x, z)$ και μετά:

  $\boldsymbol{GrandParent(x, y) \Leftarrow (Μητέρα(x, z) \wedge Πατέρας(z, y))}$
- Μέχρι **για κάθε θετικό** να υπάρχει **πρόταση** Horn που να το **καλύπτει**.

FOIL – συνέχεια

Προαιρετική μελέτη

- Οι συνθήκες που προστίθενται μπορούν να είναι **κατηγορήματα** (ή αρνήσεις αυτών).
  - Όλα τα ορίσματά τους πρέπει να είναι **μεταβλητές**.
  - **Τουλάχιστον μία** από αυτές τις μεταβλητές πρέπει να χρησιμοποιείται σε **προηγούμενη συνθήκη** ή την **κεφαλή**.
  - Το $GrandParent(\mathbf{x}, y) \Leftarrow Μητέρα(\mathbf{x}, z)$ μπορεί να γίνει π.χ.:

    $GrandParent(x, y) \Leftarrow (Μητέρα(x, \mathbf{z}) \wedge Πατέρας(\mathbf{z}, w))$
  - Επιτρέπεται να προστεθεί ως συνθήκη και το κατηγόρημα-στόχος (ώστε να προκύπτουν και **αναδρομικοί** κανόνες).
- Οι συνθήκες μπορούν να είναι και **ισότητες**, **ανισότητες** ή αριθμητικές **συγκρίσεις**.
  - Αυτές μπορούν να περιλαμβάνουν και **σταθερές**.
  - Αν περιλαμβάνουν **μεταβλητές**, πρέπει οι μεταβλητές να εμφανίζονται και **αλλού στην πρόταση**.
- Η **επιλογή συνθήκης** γίνεται κάθε φορά με **ευρετική**.

Βιβλιογραφία

- Russel & Norvig (**ελληνική μετάφραση της 2ης έκδοσης**, υπάρχει στη βιβλιοθήκη του ΟΠΑ): ενότητες 18.1, 18.2, 19.1,19.2, 19.5 (μόνο όσα αναφέρουν οι διαφάνειες).
  - Δυστυχώς η ελληνική μετάφραση της 4ης έκδοση των R&N δεν καλύπτει το μεγαλύτερο μέρος αυτής της διάλεξης. Από την ελληνική μετάφραση της 4ης έκδοσης, σχετικές είναι οι ενότητες 19.1, 19.2. Θα καλύψουμε αρκετές άλλες υπο-ενότητες του κεφαλαίου 19 σε επόμενες διαλέξεις.
  - Εναλλακτικά μπορείτε να συμβουλευτείτε την αγγλική 4η έκδοση (υπάρχει στη βιβλιοθήκη του ΟΠΑ), ενότητες 19.1, 19.2, 20.1, 20.2, 20.5 (μόνο όσα αναφέρουν οι διαφάνειες).
- Βλαχάβας κ.ά: ενότητα 8.2, κεφάλαιο 18 ως και ενότητα 18.2, ενότητες 18.4, 18.6 (μόνο όσα αναφέρουν οι διαφάνειες).
  - Όσοι ενδιαφέρονται μπορούν να διαβάσουν προαιρετικά (εκτός εξεταστέας ύλης) και τις υπόλοιπες ενότητες του κεφαλαίου 8.