ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Τεχνητή Νοημοσύνη

9η διάλεξη (2023-24)

Ίων Ανδρουτσόπουλος

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

Οι διαφάνειες αυτής της διάλεξης βασίζονται εν μέρει στο βιβλίο *Artificial Intelligence – A Modern Approach* των S. Russel και P. Norvig, 2η και 4η έκδοση, Prentice Hall, 2003 και 2020.

Τι θα ακούσετε σήμερα

- Εξαγωγή συμπερασμάτων με **προτάσεις Horn προτασιακής λογικής**.
- **Πρωτοβάθμια κατηγορηματική λογική**.
  - Συντακτικό ΠΚΛ.
  - Παράσταση γνώσεων με ΠΚΛ.

Προτάσεις Horn προτασιακής λογικής

- **Πρόταση Horn ΠΛ**: $(l_1 \vee \ldots \vee l_k)$
  - Κάθε l_i είναι σύμβολο (π.χ. P) ή άρνηση συμβόλου (π.χ. ¬P).
  - Το **πολύ ένα από τα** l_i **δεν είναι άρνηση** συμβόλου («θετικό»).
  - Π.χ. $(\neg L_{1,1} \vee \neg Breeze \vee B_{1,1})$, $(\neg W_{1,1} \vee \neg W_{1,2})$, $B_{1,1}$
- Μετατρέπεται ισοδύναμα σε «**κανόνα**» $((\ldots \wedge \ldots \wedge \ldots) \Rightarrow l_i)$
  - $((L_{1,1} \wedge Breeze) \Rightarrow B_{1,1})$, $((W_{1,1} \wedge W_{1,2}) \Rightarrow False)$, $(True \Rightarrow B_{1,1})$
  - Αν δεν έχει **κανένα αρνητικό** σύμβολο, «**γεγονός**» (fact).
  - Αν **δεν έχει θετικό** σύμβολο, «**περιορισμός ακεραιότητας**».
  - Δε θα ασχοληθούμε με περιορισμούς ακεραιότητας.
  - Αν **έχει (ένα) θετικό** σύμβολο, «**οριστική πρόταση**» (definite clause). (Τα γεγονότα είναι και αυτά οριστικές προτάσεις.)
- **ΒΓ προτάσεων Horn ΠΛ**:
  - Περιέχει μόνο προτάσεις Horn ΠΛ.
  - Άρα χρησιμοποιούμε **υποσύνολο της προτασιακής λογικής**.

Εξαγωγή συμπερασμάτων με προτάσεις Horn ΠΛ

- Η εξαγωγή συμπερασμάτων μπορεί να γίνει σε **χρόνο γραμμικό** προς το μέγεθος της ΒΓ.
  - Πιο εύκολη από ό,τι αν χρησιμοποιούμε την πλήρη προτασιακή λογική και τον κανόνα της ανάλυσης.
  - Όμως χρησιμοποιούμε υποσύνολο της προτασιακής λογικής.
- Εξαγωγή συμπερασμάτων **προς τα εμπρός** (forward chaining).
  - Διαδοχική «πυροδότηση» (κατά την κατεύθυνση των βελών) κανόνων των οποίων οι υποθέσεις αληθεύουν (με Modus Ponens), μέχρι να καταλήξουμε στο επιθυμητό συμπέρασμα.
- Εξαγωγή συμπερασμάτων **προς τα πίσω** (backward chaining).
  - Χρήση κανόνων κατά την αντίστροφη κατεύθυνση των βελών, ξεκινώντας από το συμπέρασμα μέχρι να καταλήξουμε σε υποθέσεις που γνωρίζουμε (υπάρχουν στη ΒΓ).
- Η βάση του **λογικού προγραμματισμού**.
  - Π.χ. τα προγράμματα **Prolog** είναι συλλογές προτάσεων Horn (αλλά πρωτοβάθμιας κατηγορηματικής λογικής).

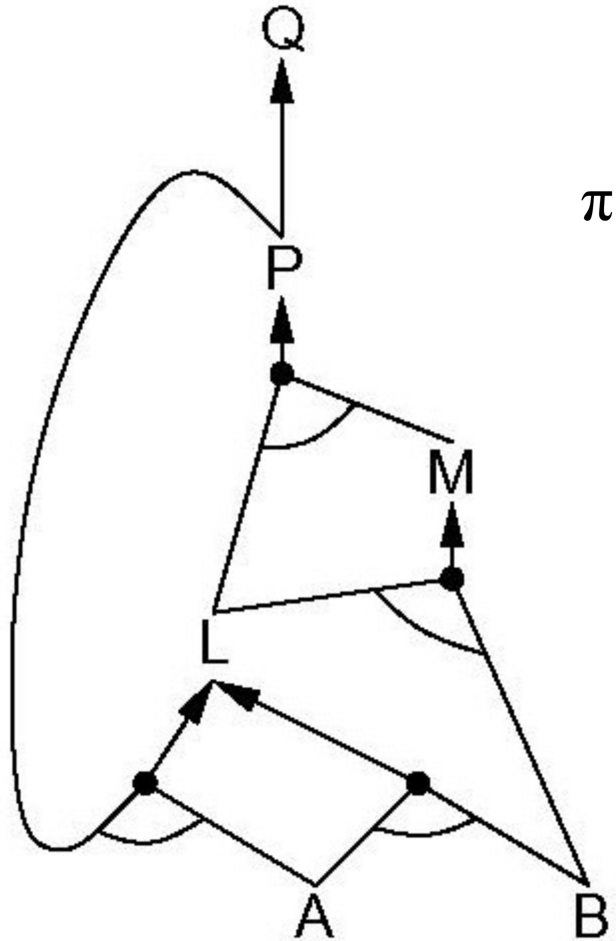
Γράφοι σύζευξης-διάζευξης (AND-OR)

ΒΓ σε μορφή προτάσεων Horn ΠΛ:

$$P \Rightarrow Q$$
$$L \wedge M \Rightarrow P$$
$$B \wedge L \Rightarrow M$$
$$A \wedge P \Rightarrow L$$
$$A \wedge B \Rightarrow L$$
$$A$$
$$B$$

ισοδύναμος γράφος AND-OR:

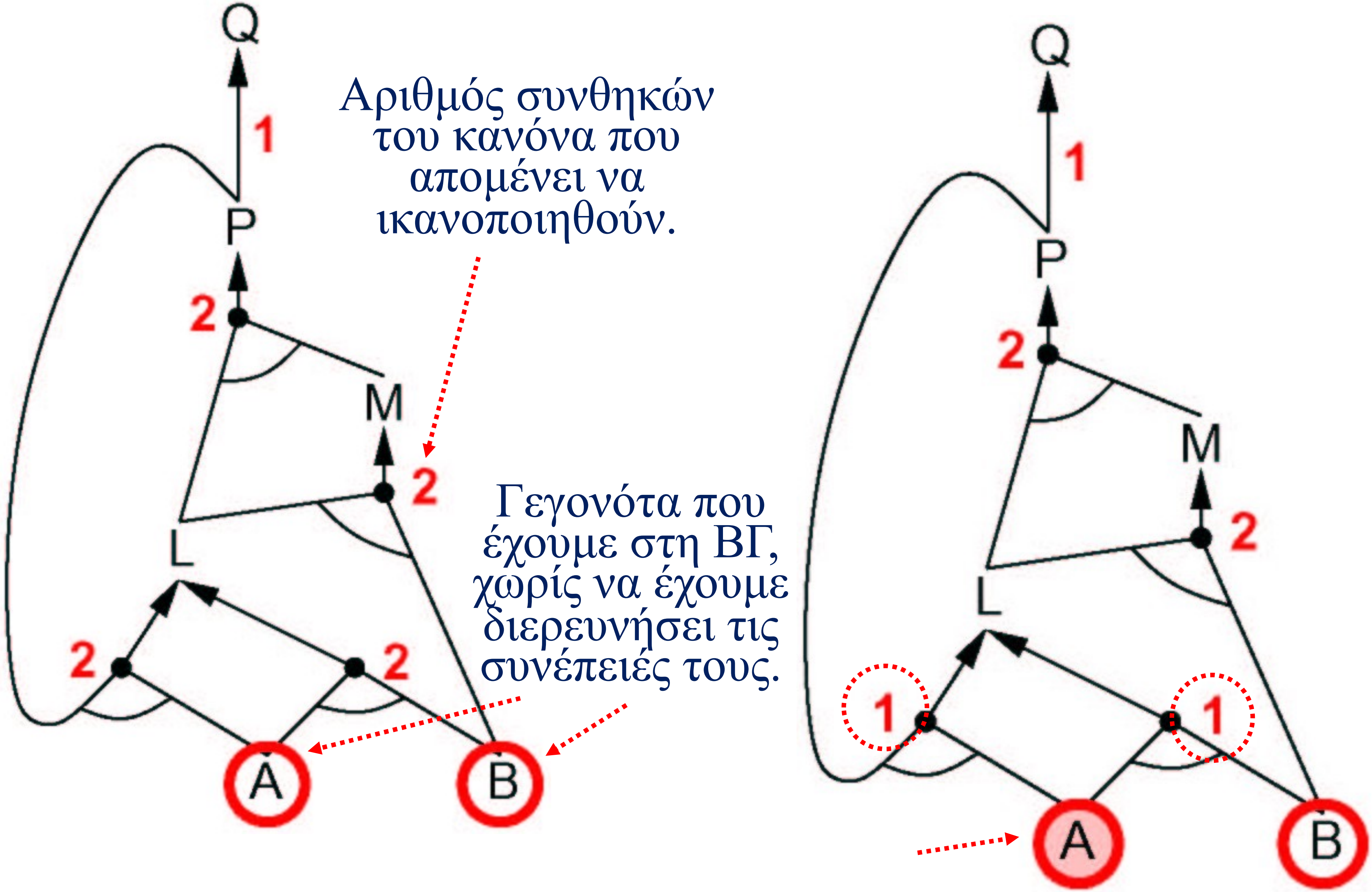
προς τα εμπρός



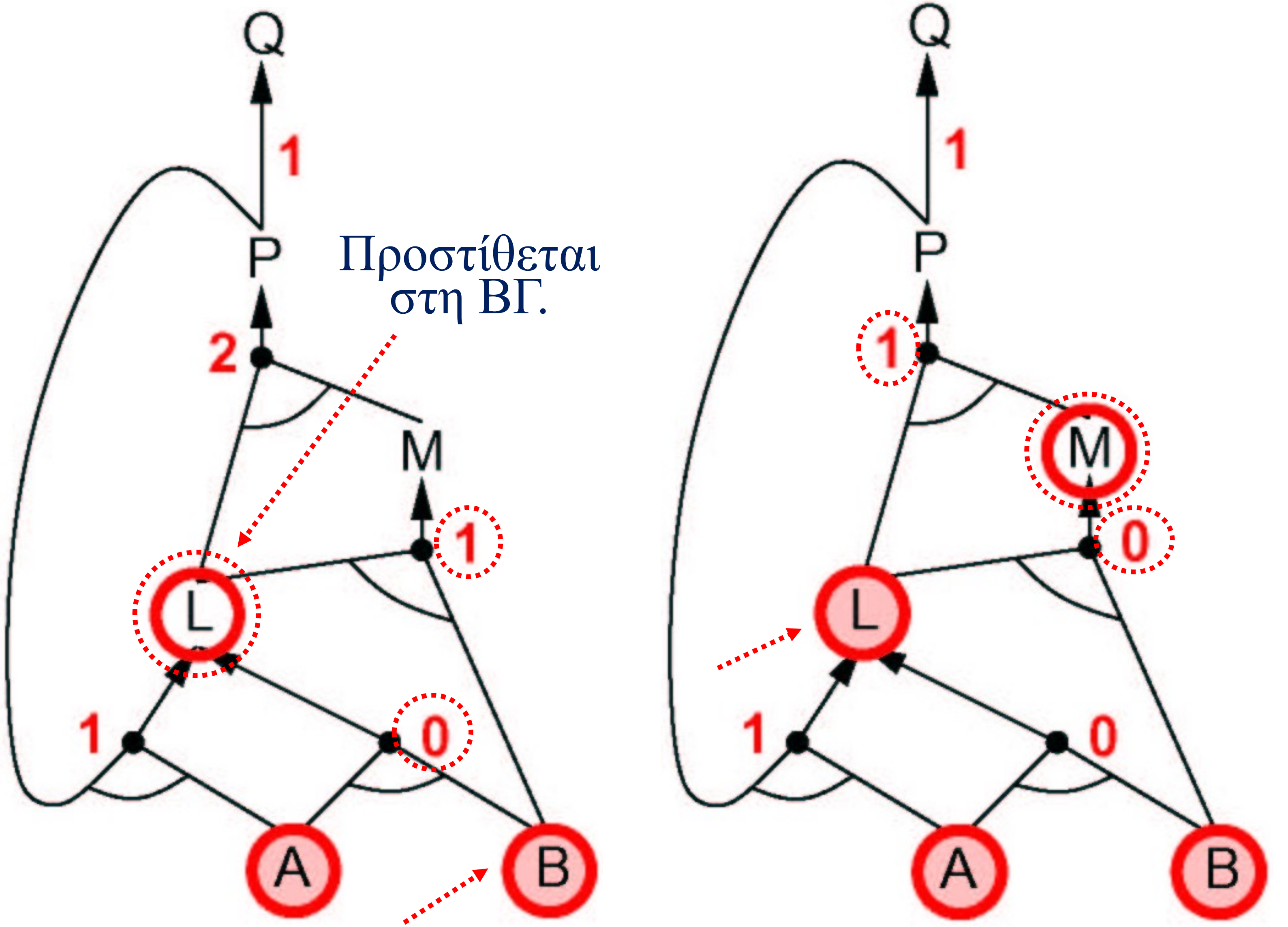
προς τα πίσω

Εξαγωγή συμπεράσματος προς τα εμπρός

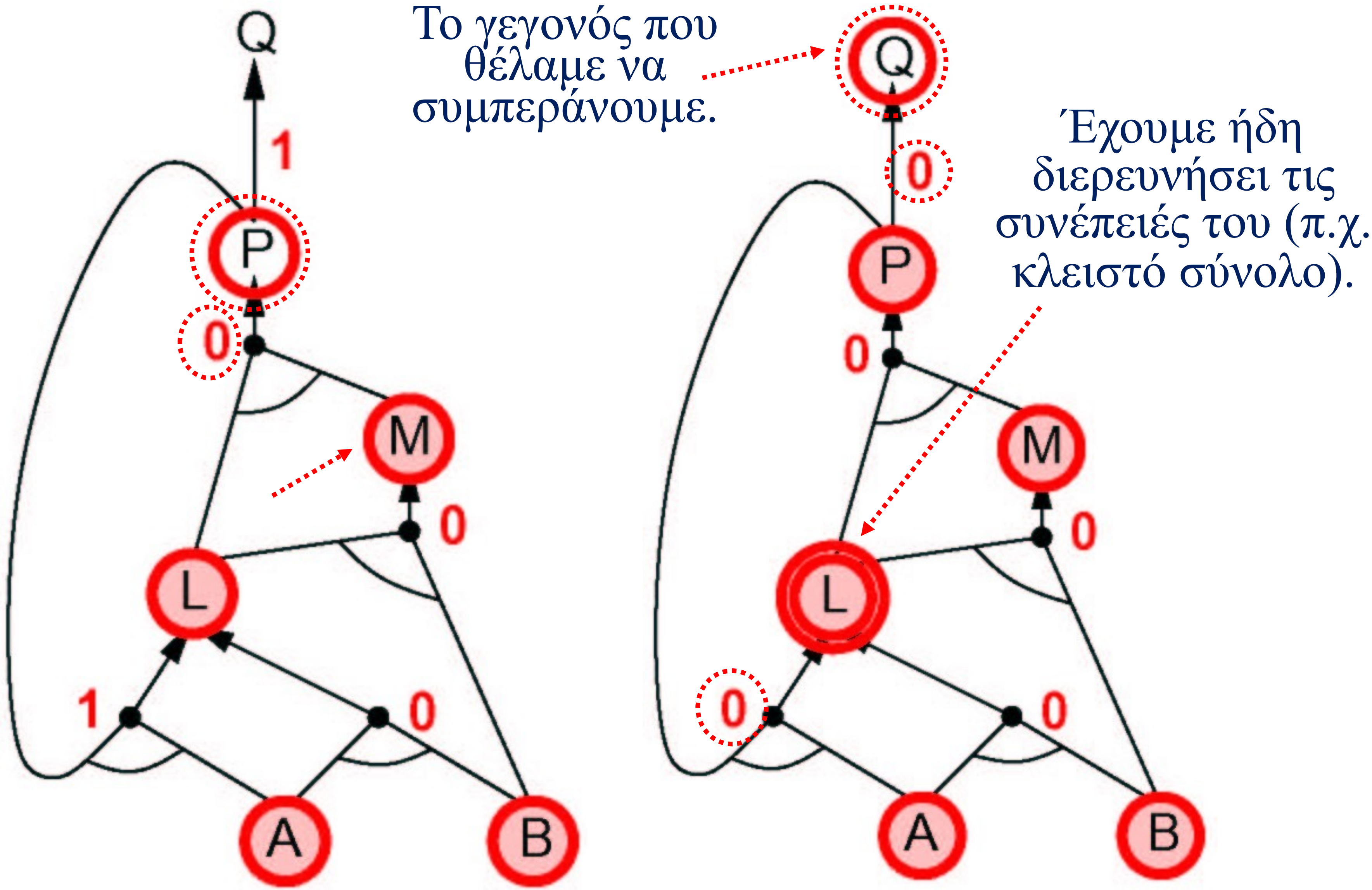
Θέλουμε να δούμε αν μπορούμε να συμπεράνουμε Q από τη ΒΓ.



Συνέχεια εξαγωγής συμπεράσματος προς τα εμπρός



Συνέχεια εξαγωγής συμπεράσματος προς τα εμπρός



Εξαγωγή συμπερασμάτων (με ΒΓ προτάσεων Horn προτασιακής λογικής) προς τα **εμπρός**

```
function PL-FC-ENTAILS?(KB, q) returns true or false
  local variables: count, a table, indexed by clause, initially the number of premises
                   inferred, a table, indexed by symbol, each entry initially false
                   agenda, a list of symbols, initially the symbols known to be true

  while agenda is not empty do
      p ← POP(agenda)
      unless inferred[p] do
          inferred[p] ← true
          for each Horn clause c in whose premise p appears do
              decrement count[c]
              if count[c] = 0 then do
                  if HEAD[c] = q then return true
                  PUSH(HEAD[c], agenda)
  return false
```

Γεγονότα που έχουμε στη ΒΓ χωρίς να έχουμε διερευνήσει τις συνέπειές τους (μέτωπο).

inferred[p]: Δείχνει αν έχουμε συμπεράνει το γεγονός p ή όχι.

Πυροδότηση κανόνα.

Χαρακτηριστικά του PL-FC-Entails?

- **Τερματίζει**.
  - Η ατζέντα μικραίνει κατά ένα σε κάθε επανάληψη του while.
  - Μεγαλώνει όταν πυροδοτείται ένας κανόνας.
  - Κάθε κανόνας πυροδοτείται το πολύ μία φορά (όταν μηδενιστεί ο μετρητής του) και έχουμε πεπερασμένο αριθμό κανόνων.
  - Άρα τελικά η ατζέντα αδειάζει και ο αλγόριθμος τερματίζει.
- **Χρονική πολυπλοκότητα: γραμμική ως προς το μέγεθος της ΒΓ.**
  - Η απόδειξη παραλείπεται.
- **Ορθότητα**: Χρησιμοποιούμε μόνο Modus Ponens, που είναι ορθός κανόνας εξαγωγής συμπερασμάτων.
- **Πληρότητα** (αν ΒΓ $\models$ q, τότε ΒΓ $\vdash_i$ q, <u>για q γεγονός</u>).
  - Η απόδειξη παραλείπεται.

Μειονεκτήματα προτασιακής λογικής

- Δύσκολη παράσταση **γενικεύσεων**.
  - Π.χ. «Υπάρχει ρεύμα αέρος σε ένα τετράγωνο ανν υπάρχει όρυγμα σε διπλανό τετράγωνο».
  - Χρειαζόμαστε μια διαφορετική πρόταση για κάθε τετράγωνο:

    $(B_{2,2} \Leftrightarrow (P_{2,1} \vee P_{2,3} \vee P_{1,2} \vee P_{3,2}))$

    $(B_{3,2} \Leftrightarrow (P_{3,1} \vee P_{3,3} \vee P_{2,2} \vee P_{4,2}))$ κ.ο.κ.
- Ενώ εύκολη παράσταση στην **πρωτοβάθμια κατηγορηματική λογική** (ΠΚΛ).
  - Ένας μόνο τύπος για όλα τα μη ακραία τετράγωνα:

    $$\forall x\ \forall y\ (\ B(x, y) \Leftrightarrow$$
    $$(\ P(x, y-1) \vee P(x, y+1) \vee P(x-1, y) \vee P(x+1, y)\ )\ )$$
  - Και ειδικοί τύποι για τα ακραία τετράγωνα.

Παραδείγματα τύπων ΠΚΛ

Συμβουλή: μην παραλείπετε ποτέ **παρενθέσεις**!

- *Σε όλες τις γάτες αρέσει το γάλα.*

  $\forall x\ (IsCat(x) \Rightarrow Likes(x, Milk))$

- *Υπάρχει μια γάτα που της αρέσει το γάλα.*

  $\exists x\ (IsCat(x) \land Likes(x, Milk))$

  – Προσοχή: ο τύπος $\exists x\ (IsCat(x) \Rightarrow Likes(x, Milk))$ λέει «Υπάρχει ένα x που: (i) δεν είναι γάτα ή (ii) αν είναι γάτα του αρέσει το γάλα».

- *Η Ψίτα συμπαθεί όλους τους σκύλους.*

  $\forall x\ (IsDog(x) \Rightarrow Likes(Psita, x))$

  – Προσοχή: ο τύπος $\forall x\ (IsDog(x) \land Likes(Psita, x))$ λέει «Τα πάντα είναι σκύλοι και αρέσουν στην Ψίτα».

- *Υπάρχει μια γάτα που συμπαθεί όλους τους σκύλους.*

  $\exists x\ (IsCat(x) \land \forall y\ (IsDog(y) \Rightarrow Likes(x, y)))$

Παραδείγματα τύπων ΠΚΛ – συνέχεια

- *Ο Μίλος αντιπαθεί όλες τις γάτες.*

  $\forall x\ (\mathrm{IsCat}(x) \Rightarrow \neg\ \mathrm{Likes}(\mathrm{Milos}, x))$

- *Όλοι οι σκύλοι αντιπαθούν όλες τις γάτες.*

  $\forall x\ (\mathrm{IsDog}(x) \Rightarrow \forall y\ (\mathrm{IsCat}(y) \Rightarrow \neg\ \mathrm{Likes}(x, y)))$

  ή ισοδύναμα:

  $\forall x\ \forall y\ ((\mathrm{IsDog}(x) \wedge \mathrm{IsCat}(y)) \Rightarrow \neg\ \mathrm{Likes}(x, y))$

- *Κάθε άνθρωπος συμπαθεί τον πατέρα του.*

  $\forall x\ \forall y\ ((\mathrm{IsHuman}(x) \wedge \mathrm{IsFatherOf}(y, x)) \Rightarrow \mathrm{Likes}(x, y))$

  ή: $\forall x\ (\mathrm{IsHuman}(x) \Rightarrow \mathrm{Likes}(x, \mathbf{FatherOf(x)}))$

Παραδείγματα τύπων ΠΚΛ – συνέχεια

- *Κάθε σκύλος που γαβγίζει φοβάται μια (πιθανώς διαφορετική ή την ίδια) γάτα.*

  $\forall x\ ((IsDog(x) \wedge Barks(x)) \Rightarrow$

  $\exists y\ (IsCat(y) \wedge IsAfraidOf(x, y)))$

- *Κάθε γάτα συμπαθεί <u>ακριβώς έναν</u> (πιθανώς διαφορετικό ή τον ίδιο όλες) σκύλο.*

  $\forall y\ (IsCat(y) \Rightarrow$

  $\exists x\ (IsDog(x) \wedge Likes(y, x) \wedge$

  $\forall z\ ((IsDog(z) \wedge Likes(y, z)) \Rightarrow z = x)))$

Συντακτικό ΠΚΛ

τύπος → *ατομικός_τύπος*
| (*τύπος σύνδεσμος τύπος*)
| *ποσοδείκτης μεταβλητή τύπος*
| ¬*τύπος*

ατομικός_τύπος → *σύμβολο_σχέσης*(*όρος*, ...) | *όρος* = *όρος*

όρος → *σταθερά* | *μεταβλητή* |
σύμβολο_συνάρτησης(*όρος*, ...)

σύνδεσμος → $\wedge$ | $\vee$ | $\Rightarrow$ | $\Leftrightarrow$

ποσοδείκτης → $\forall$ | $\exists$

σταθερά → A | X_1 | John | Mary | ...

μεταβλητή → a | x | s | ...

σύμβολο_σχέσης → IsFatherOf | HasColor | IsKing | ...

σύμβολο_συνάρτησης → FatherOf | LeftLeg | ...

Τα σύνολα των σταθερών, μεταβλητών, συμβόλων σχέσεων, συμβόλων συναρτήσεων θεωρούμε ότι είναι ανά δύο ξένα.

Βιβλιογραφία

- Russel & Norvig (4η έκδοση): ενότητες 7.5.3, 7.5.4, εισαγωγή ενότητας 8.1.
    - Όσοι ενδιαφέρονται μπορούν να διαβάσουν προαιρετικά και τις υπόλοιπες ενότητες του κεφαλαίου 7.
- Βλαχάβας κ.ά: εισαγωγή ενότητας 9.2 (χωρίς την υπο-ενότητα 9.2.1).