ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΤΗΝ ΛΟΓΙΚΗ ΚΑΙ
ΘΕΩΡΙΑ ΑΛΓΟΡΙΘΜΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΜΟΥ

μΠλ∀

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΟΣ

ΜΙΧΑΛΗΣ ΝΙΚΟΛΑΟΥ

ΜΠΛΑ 200603

Επιβλέπων:

Ευάγγελος Ράπτης, Αναπληρωτής Καθηγητής ΕΚΠΑ

ΑΘΗΝΑ
ΙΟΥΛΙΟΣ 2010

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΟΣ

ΜΙΧΑΛΗΣ ΝΙΚΟΛΑΟΥ
Α.Μ 200603

ΕΠΙΒΛΕΠΩΝ:
ΕΥΑΓΓΕΛΟΣ ΡΑΠΤΗΣ, Αναπληρωτής Καθηγητής ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ:
ΕΥΑΓΓΕΛΟΣ ΡΑΠΤΗΣ, Αναπληρωτής Καθηγητής ΕΚΠΑ
ΚΩΝΣΤΑΝΤΙΝΟΣ ΔΗΜΗΤΡΑΚΟΠΟΥΛΟΣ, Καθηγητής ΕΚΠΑ
ΔΗΜΗΤΡΙΟΣ ΘΗΛΥΚΟΣ, Επίκουρος Καθηγητής ΕΚΠΑ

ΑΘΗΝΑ
ΙΟΥΛΙΟΣ 2010

Ευχαριστίες

Πρώτα απ'όλα θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου Ευάγγελο Ράπτη για την εμπιστοσύνη που έδειξε στο πρόσωπο μου ως προς την διεκπεραίωση αυτού του δύσκολου θέματος για την διπλωματική μου εργασία, την συνεργατικότητα του και το προσωπικό ενδιαφέρον που δείχνει απροσωπόληπτα σε όποιον φοιτητή τον προσεγγίσει. Επιπλέον θα ήθελα να ευχαριστήσω και τους υπόλοιπους καθηγητές, των οποίων τα μαθήματα παρακολούθησα κατά την διάρκεια της φοίτησης μου στο ΜΠΛΑ, για την υπομονή που έδειξαν ως προς το πρόσωπο μου σε σχέση με την ασυνέπεια που ίσως επέδειξα κάποιες φορές λόγω των επαγγελματικών μου υποχρεώσεων.

Επιπλέον θα ήθελα να ευχαριστήσω την οικογένεια μου, χωρίς την ψυχολογική υποστήριξη των οποίων δεν θα μπορούσα να αντέξω τις απαιτήσεις για την επιτυχή διεκπεραίωση αυτού του μεταπτυχιακού.

“Cogito ergo sum”

Descartes, 1637

ΠΕΡΙΛΗΨΗ

Προσπαθώντας να βρω το θέμα της διπλωματικής μου εργασίας και έχοντας κατά νου να κάνω – όσο ήταν δυνατόν – μερική χρήση των βασικότερων εννοιών που πραγματεύτηκα κατά την διάρκεια της φοίτησης μου στο ΜΠΛΑ, ήρθα αντιμέτωπος με το πεδίο της Τεχνητής Νοημοσύνης. Φαινόταν πολύ τρομακτικό και αρχικά πίστευα ότι δεν θα έβγαζα άκρη από την τεράστια βιβλιογραφία που έχει γραφτεί για αυτό το θέμα.

Έχοντας κατά νου ότι τα Μαθηματικά, αν μη τι άλλο, είναι η τέχνη του Συμπερασμού, προσπάθησα να περιορίσω αυτή την αχανή βιβλιογραφία σε ένα αρκετά μεγάλο υποσύνολο που αφορούσε όμως τον Συμπερασμό που θα μπορούσε ίσως να εκτελέσει μια Τεχνητή Νοημοσύνη. Έτσι με υπομονή και επιμονή, πιστεύω ότι κατάφερα να γράψω ένα συνεκτικό κομμάτι που προσεγγίζει με τρεις διαφορετικούς τρόπους την "προσπάθεια" μιας Τεχνητής Νοημοσύνης να βγάλει ένα συμπέρασμα, έχοντας στην διάθεση της μια βάση γνώσης.

Και οι τρεις τρόποι Συμπερασμού έχω φροντίσει να γεφυρωθούν υπό μια έννοια. Η προσέγγιση εδώ είναι δηλωτικού χαρακτήρα, εννοώντας ότι ο "πράκτορας" μας βασίζει τα συμπεράσματα του πάνω στην γνώση που έχει για ένα περιβάλλον. Δεν ασχολήθηκα με διαδικαστικές προσεγγίσεις όπου κωδικοποιούν τις επιθυμητές συμπεριφορές ενός "πράκτορα" απευθείας ως κώδικα προγράμματος ανάλογα με το ερέθισμα που λαμβάνουν την δεδομένη στιγμή, και αυτό γιατί θα είχα πλατιάσει τόσο πολύ που θα είχα ξεφύγει από τους σκόπους αυτής της διπλωματικής.

Ο Σκοπός αυτής της διπλωματικής εργασίας είναι να κάνει μια πρακτική εφαρμογή όλων όσων έμαθα στο ΜΠΛΑ από τα πεδία της Λογικής, της Υπολογιστικής άλγεβρας, των αλγορίθμων και της Υπολογισιμότητας. Και μάλιστα να τα εφαρμόσει προς όφελος της πιο σημαντικής ικανότητας που διαθέτουμε ως άνθρωποι έναντι των άλλων έμβιων όντων. Της Σκέψης!

ΠΕΡΙΕΧΟΜΕΝΑ σελ.

0. ΔΗΜΙΟΥΡΓΩΝΤΑΣ ΜΙΑ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Για χιλιάδες χρόνια προσπαθούμε να κατανοήσουμε το πώς σκεπτόμαστε· δηλαδή, πώς μια χούφτα ύλης μπορεί να αντιλαμβάνεται, να κατανοεί, να προβλέπει και να χειρίζεται έναν κόσμο πολύ μεγαλύτερο και πολύ πιο πολύπλοκο από τον εαυτό της. Το πεδίο της τεχνητής νοημοσύνης (artificial intelligence), ή για συντομία ΤΝ, πηγαίνει ακόμα πιο πέρα: Επιχειρεί όχι μόνο να κατανοήσει αλλά και να *κατασκευάσει* νοήμονες οντότητες.

Η ΤΝ συστηματοποιεί και αυτοματοποιεί τις διανοητικές εργασίες, γι'αυτό και μπορεί να έχει εφαρμογή σε οποιαδήποτε σφαίρα της ανθρώπινης δραστηριότητας.

Μιας και **Πράκτορας** (agent) είναι κάτι που "πράττει" (ο αγγλικός όρος επίσης προέρχεται από το λατινικό agere, που σημαίνει "πράττω"), στον παρόν κείμενο αυτές οι κατασκευασμένες νοήμονες οντότητες θα ονομάζονται με τη λέξη *Πράκτορες*.

0.1 ΠΡΑΚΤΟΡΕΣ ΒΑΣΙΣΜΕΝΟΙ ΣΤΗΝ ΓΝΩΣΗ

Η Κεντρική συνιστώσα ενός βασισμένου στη γνώση πράκτορα είναι η **βάση γνώσης** του (knowledge base ή KB). Άτυπα, βάση γνώσης είναι ένα σύνολο προτάσεων (sentences). Κάθε πρόταση είναι εκφρασμένη σε μια γλώσσα που ονομάζεται **γλώσσα αναπαράστασης γνώσης** και αντιπροσωπεύει κάποιον ισχυρισμό για τον κόσμο.

Θα πρέπει να υπάρχει ένας τρόπος να προστίθενται νέες προτάσεις στη βάση γνώσης και ένας τρόπος να τίθενται ερωτήματα για ότι είναι γνωστό. Τα καθιερωμένα ονόματα για τις εργασίες αυτές είναι TELL και ASK , αντίστοιχα. Και οι δύο εργασίες μπορεί να περιλαμβάνουν **συμπερασμό** (inference) - δηλαδή , εξαγωγή νέων προτάσεων από τις υπάρχουσες. Στους **λογικούς πράκτορες** (logical agents) , οι οποίοι είναι το κύριο αντικείμενο μελέτης μας εδώ, ο συμπερασμός πρέπει να υπακούει στην θεμελιώδη απαίτηση ότι,όταν κάποιος θέτει ένα ερώτημα (ASK) στην βάση γνώσης , η απάντηση θα πρέπει να προκύπτει από τις γνώσεις με τις οποίες έχει προηγουμένως ενημερωθεί (TELL) η βάση γνώσης.

Στην εικόνα 0.1 σκιαγραφείται ένα πρόγραμμα πράκτορα βασισμένου στην γνώση. Όπως όλοι οι πράκτορες μας , παίρνει ως είσοδο μια αντίληψη (percept) και επιστρέφει μια ενέργεια. Ο πράκτορας συντηρεί μια βάση γνώσης , KB, η οποία μπορεί αρχικά να περιέχει κάποιο **γνωστικό υπόβαθρο** (background knowledge). Κάθε φορά που καλείται , το πρόγραμμα του πράκτορα κάνει δύο πράγματα. Πρώτον ,ενημερώνει (TELL) τη βάση γνώσης για το τι αντιλαμβάνεται . Δεύτερον, ρωτά (ASK) τη βάση γνώσης τι ενέργεια θα πρέπει να κάνει. Κατά τη διαδικασία της απάντησης στο ερώτημα αυτό, μπορούν να πραγματοποιούνται εκτεταμένοι συλλογισμοί για την τρέχουσα κατάσταση του κόσμου, για τα αποτελέσματα των δυνατών ακολουθιών ενεργειών κ.ο.κ. Αφού επιλεγεί η ενέργεια , ο πράκτορας καταγράφει την επιλογή του με μια πράξη TELL και εκτελεί την ενέργεια. Η δεύτερη TELL είναι απαραίτητη για να γνωρίζει η βάση γνώσης ότι η υποθετική ενέργεια έχει πραγματικά εκτελεστεί.

```
function KB-Agent ( αντίληψη) returns μια ενέργεια
    static : KB , βάση γνώσης
             t, μετρητής για το χρόνο,αρχικά 0

    TELL(KB, MAKE-PERCEPT-SENTENCE( αντίληψη, t) )
    ενέργεια ← ASK ( KB, MAKE-ACTION-QUERY(t) )
    TELL( KB , MAKE-ACTION-SENTENCE( ενέργεια , t) )
    t ← t+1
    return ενέργεια
```

Εικόνα 0.1. Γενικός πράκτορας βασισμένος στη γνώση

Οι λεπτομέρειες της γλώσσας αναπαράστασης είναι κρυμμένες μέσα στις τρεις συναρτήσεις που υλοποιούν τη διασύνδεση μεταξύ των αισθητήρων και των μηχανισμών δράσης από τη μία και του συστήματος εσωτερικής αναπαράστασης και συλλογιστικής από την άλλη. Η MAKE-PERCEPT-SENTENCE παίρνει μια αντίληψη (percept) και ένα χρόνο και επιστρέφει μια πρόταση που ισχυρίζεται ότι ο πράκτορας προσέλαβε την αντίληψη στο δεδομένο χρόνο. Η MAKE-ACTION-QUERY παίρνει ως είσοδο ένα χρόνο και επιστρέφει μια πρόταση που ρωτά τι ενέργεια θα πρέπει να γίνει στον χρόνο αυτό. Οι λεπτομέρειες των μηχανισμών συμπερασμού είναι κρυμμένες μέσα στις TELL και ASK . Σε επόμενες ενότητες θα αποκαλύψουμε αυτές τις λεπτομέρειες.

Λόγω των ορισμών της TELL και της ASK , ο βασισμένος στην γνώση πράκτορας δεν είναι ένα τυχαίο πρόγραμμα για τον υπολογισμό ενεργειών. Επιτρέπει μια περιγραφή στο **επίπεδο γνώσης** , όπου χρειάζεται να καθορίσουμε μόνο τι γνωρίζει ο πράκτορας και ποιοι είναι οι στόχοι του προκειμένου να κανονίσει την συμπεριφορά του. Για παράδειγμα , ένα αυτόματο ταξί θα μπορούσε να έχει στόχο να μεταφέρει έναν επιβάτη στο Marin County , και να γνωρίζει ότι βρίσκεται στο San Francisco και ότι η γέφυρα Golden Gate είναι ο μόνος σύνδεσμος μεταξύ των δύο τοποθεσιών. Τότε μπορούμε να αναμένουμε ότι το ταξί θα διασχίσει τη γέφυρα Golden Gate , *επειδή γνωρίζει ότι έτσι θα επιτύχει το στόχο του*. Η ανάλυση αυτή είναι ανεξάρτητη απο το πώς δουλεύει το ταξί σε **επίπεδο υλοποίησης .** Δεν έχει σημασία αν η γεωγραφική του γνώση είναι υλοποιημένη ως συνδεδεμένες λίστες ή ως χάρτες εικονοστοιχείων , ούτε αν συλλογίζεται με το χειρισμό ακολουθιών συμβόλων αποθηκευμένων σε καταχωρητές ή με τη διάδοση ενθόρυβων σημάτων σε ένα δίκτυο νευρώνων.

Όπως αναφέραμε στην αρχή , *μπορεί κανείς να κατασκευάσει ένα βασισμένο στην γνώση πράκτορα απλώς λέγοντας του (TELL) ότι χρειάζεται να γνωρίζει.* Το αρχικό πρόγραμμα του πράκτορα , πριν αρχίσει να προσλαμβάνει αντιλήψεις , κατασκευάζεται με την προσθήκη μίας προς μία των προτάσεων που αντιπροσωπεύουν τη γνώση του περιβάλλοντος που έχει ο σχεδιαστής . Η σχεδίαση της γλώσσας αναπαράστασης ώστε να μπορεί εύκολα να εκφράζεται αυτή η γνώση με τη μορφή προτάσεων απλουστεύει εξαιρετικά το πρόβλημα της κατασκευής. Η προσέγγιση αυτή ονομάζεται **δηλωτική** (declarative) προσέγγιση για την κατασκευή συστημάτων. Αντίθετα, η **διαδικαστική** (procedural) προσέγγιση κωδικοποιεί τις επιθυμητές συμπεριφορές απευθείας ως κώδικα προγράμματος , όπου η ελαχιστοποίηση του ρόλου της ρητής αναπαράστασης και συλλογιστικής μπορεί να δώσει ως αποτέλεσμα ένα πολύ πιο αποδοτικό σύστημα. Στις δεκαετίες του 1970 και του 1980 οι υποστηρικτές των δύο προσεγγίσεων επιδόθηκαν σε έντονες διαμάχες. Σήμερα καταλαβαίνουμε ότι ένας επιτυχημένος πράκτορας πρέπει να συνδυάζει και δηλωτικά και διαδικαστικά στοιχεία στη σχεδίαση του.

Εκτός από το να πληροφορούμε (TELL) έναν βασισμένο στη γνώση πράκτορα με ότι χρειάζεται να γνωρίζει , μπορούμε να τον εφοδιάσουμε με μηχανισμούς που του επιτρέπουν να μαθαίνει μόνος του. Αυτή η γνώση μπορεί να ενσωματωθεί στη βάση γνώσης του πράκτορα και να χρησιμοποιείται για τη λήψη αποφάσεων. Με αυτόν τον τρόπο ο πράκτορας μπορεί να είναι εντελώς αυτόνομος.

Όλες αυτές οι δυνατότητες – αναπαράσταση , συλλογιστική και μάθηση – βασίζονται στους αιώνες εξέλιξης της θεωρίας και της τεχνολογίας της λογικής.

0.2 Ο ΚΟΣΜΟΣ ΤΟΥ WUMPUS

Ο **κόσμος του wumpus** (wumpus world) είναι ένα σπήλαιο που αποτελείται από δωμάτια που συνδέονται με διαδρόμους. Κάπου μέσα στο σπήλαιο παραμονεύει το wumpus , ένα τέρας που τρώει όποιον μπει στο δωμάτιο του. Ένας πράκτορας μπορεί να σκοτώσει το wumpus, αλλά έχει μόνο ένα βέλος. Μερικά δωμάτια έχουν άπατες γούβες που παγιδεύουν όποιον τύχει να μπει σ'αυτά (εκτός από το wumpus ,το οποίο είναι πολύ μεγάλο για να χωρέσει στη γούβα). Το μόνο θετικό που έχει η ζωή σ'αυτό το περιβάλλον είναι η δυνατότητα να βρει κανείς ένα σωρό από χρυσάφι. Αν και ο κόσμος του wumpus είναι μάλλον ήμερος για τα σημερινά πρότυπα των παιχνιδιών για υπολογιστές , είναι ένα εξαιρετικό περιβάλλον δοκιμών για τους ευφυείς πράκτορες.

Ένα παράδειγμα κόσμου του wumpus παρουσιάζεται στην εικόνα 1.2 . Ο ακριβής ορισμός του περιβάλλοντος εργασιών δίνεται με την ακόλουθη περιγραφή:

➔ **Μέτρο της απόδοσης :** +1000 αν πάρει ο πράκτορας το χρυσό , - 1000 αν πέσει σε γούβα ή φαγωθεί από το wumpus , -1 για κάθε ενέργεια που κάνει και -10 αν ξοδέψει το βέλος.

➔ **Περιβάλλον :** Πλέγμα δωματίων 4 x 4. Ο πράκτορας ξεκινά πάντα από το τετράγωνο ΑΦΕΤΗΡΙΑ που βρίσκεται στο [1,1], βλέποντας προς τα δεξιά. Οι θέσεις του χρυσού και του wumpus επιλέγονται τυχαία με ομοιόμορφη κατανομή , μεταξύ όλων των τετραγώνων εκτός του τετραγώνου εκκίνησης.Επίσης , κάθε τετράγωνο εκτός από την αφετηρία μπορεί να είναι γούβα , με πιθανότητα 0,2.

➔ **Μηχανισμοί δράσης :** Ο πράκτορας μπορεί να μετακινείται προς τα εμπρός , να στρίβει αριστερά κατά 90° ή να στρίβει δεξιά κατά 90°. Ο πράκτορας βρίσκει οικτρό θάνατο αν μπει σε τετράγωνο που περιέχει γούβα ή ένα ζωντανό wumpus .(Ένα τετράγωνο με νεκρό wumpus είναι ασφαλές , αν και δύσοσμο.) Η μετακίνηση προς τα εμπρός δεν έχει επίδραση αν υπάρχει τοίχος μπροστά από τον πράκτορα. Η ενέργεια *Αρπαγή* μπορεί να χρησιμοποιηθεί για να πάρει ο πράκτορας ένα αντικείμενο που βρίσκεται στο ίδιο τετράγωνο με αυτόν.Η ενέργεια *Εξακόντιση* μπορεί να χρησιμοποιηθεί για να ρίξει ο πράκτορας ένα βέλος προς την κατεύθυνση που κοιτάζει.Το βέλος συνεχίζει την πορεία του μέχρι να χτυπήσει (και να σκοτώσει) το wumpus ή μέχρι να χτυπήσει σε τοίχο. Ο πράκτορας έχει μόνο ένα βέλος , γι'αυτό μόνο η πρώτη ενέργεια *Εξακόντιση* έχει επίδραση.

➔ **Αισθητήρες :** Ο πράκτορας έχει πέντε αισθητήρες , κάθε ένας από τους οποίους δίνει μία πληροφορία :

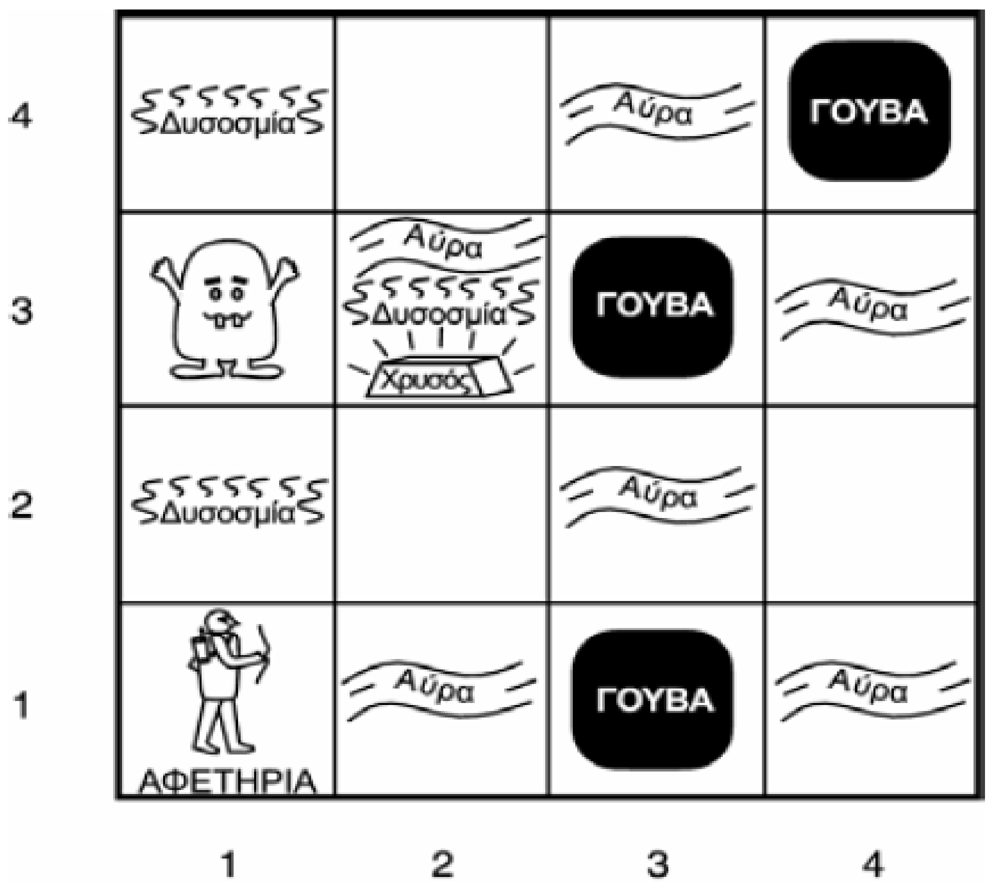
✔ Στο τετράγωνο που περιέχει το wumpus και στα άμεσα (όχι διαγώνια) γειτονικά τετράγωνα ο πράκτορας θα αντιλαμβάνεται δυσοσμία .

✔ Στα τετράγωνα που γειτονεύουν άμεσα με γούβα ο πράκτορας θα αντιλαμβάνεται αύρα.

✔ Στο τετράγωνο όπου βρίσκεται ο χρυσός ο πράκτορας θα αντιλαμβάνεται μια λάμψη.

- ✔ Όταν ένας πράκτορας πέφτει πάνω σε τοίχο θα αντιλαμβάνεται ένα γδούπο.
- ✔ Όταν σκοτωθεί το wumpus εκπέμπει μια θρηνώδη κραυγή η οποία γίνεται αντιληπτή από οπουδήποτε μέσα στο σπήλαιο.

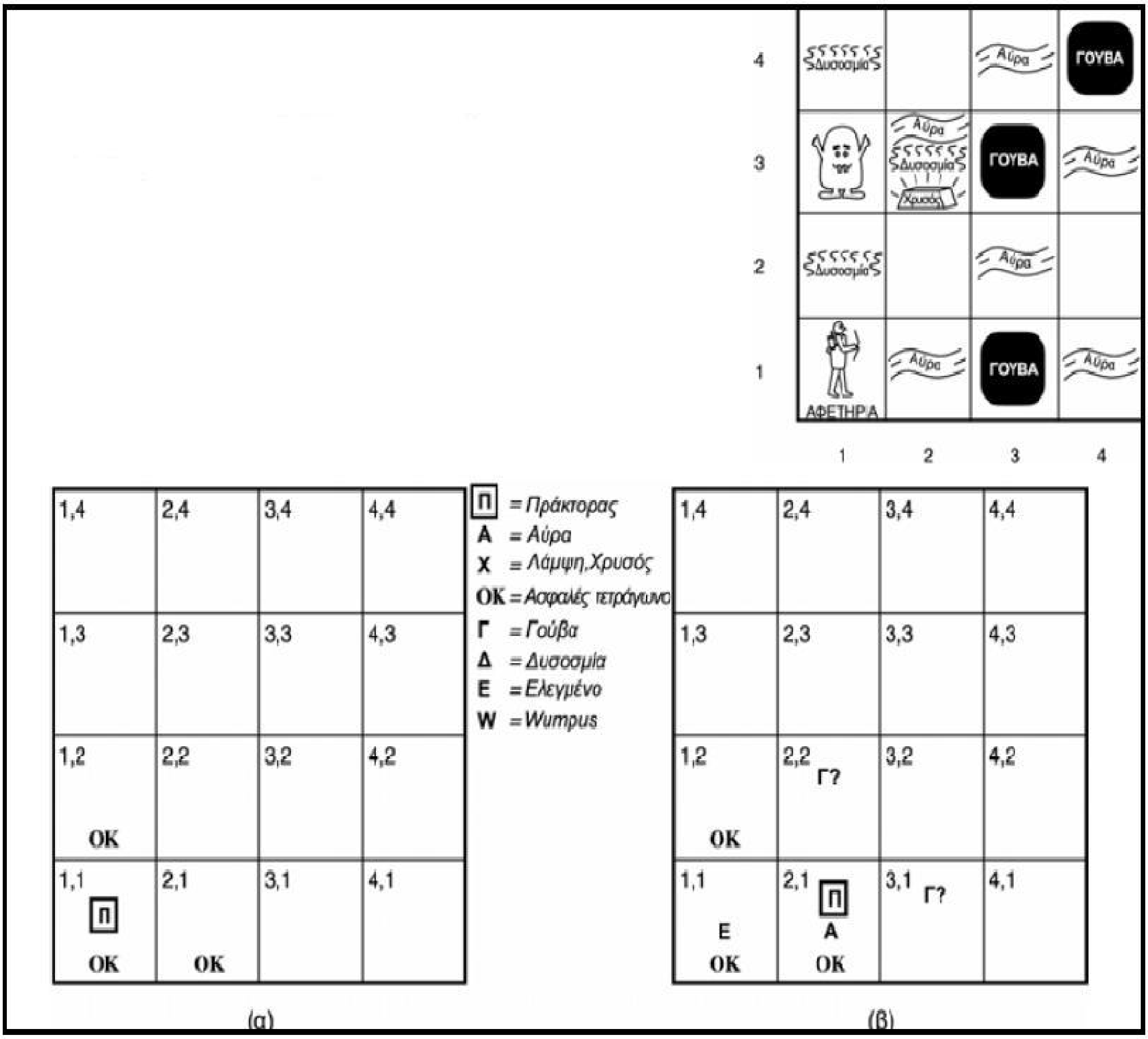
Οι αντιλήψεις θα παρέχονται στον πράκτορα με την μορφή μιας λίστας πέντε συμβόλων· για παράδειγμα , αν υπάρχει δυσοσμία και αύρα , αλλά δεν υπάρχει λάμψη , γδούπος ή κραυγή, ο πράκτορας θα λάβει την αντίληψη [*Δυσοσμία, Αύρα , Τίποτα, Τίποτα, Τίποτα*]



Εικόνα 0.2. Αντιπροσωπευτικός κόσμος του wumpus . Ο πράκτορας βρίσκεται στην κάτω αριστερή γωνία.

Η βασική δυσκολία για έναν πράκτορα είναι η αρχική του άγνοια για τη διαμόρφωση του περιβάλλοντος· το ξεπέρασμα αυτής της άγνοιας φαίνεται να απαιτεί λογική συλλογιστική. Στις περισσότερες περιπτώσεις του κόσμου του wumpus είναι δυνατό για τον πράκτορα να πάρει το χρυσό με ασφάλεια. Μερικές φορές ο πράκτορας πρέπει να επιλέξει αν θα γυρίσει πίσω με άδεια χέρια ή αν θα διακινδυνεύσει να σκοτωθεί για να βρει το χρυσό. Περίπου το 21% των περιβαλλόντων είναι εντελώς άδικα , επειδή ο χρυσός βρίσκεται σε γούβα ή περιβάλλεται από γούβες.

Ας παρατηρήσουμε ένα βασισμένο στη γνώση πράκτορα να εξερευνά το περιβάλλον της Εικόνας 0.2 . **Η αρχική βάση γνώσης** του πράκτορα περιέχει τους κανόνες του περιβάλλοντος τους οποίους απαριθμήσαμε παραπάνω· ειδικότερα ο πράκτορας γνωρίζει ότι βρίσκεται στο [1,1] και ότι το [1,1] είναι ασφαλές τετράγωνο. Θα δούμε πως εξελίσσεται η γνώση του καθώς δέχεται νέες αντιλήψεις και πραγματοποιεί ενέργειες.



Εικόνα 0.3. Το πρώτο βήμα που κάνει ο πράκτορας στον κόσμο του wumpus. (α) Η αρχική κατάσταση μετά την αντίληψη [Τίποτα,Τίποτα,Τίποτα,Τίποτα,Τίποτα]. (β) Μετά από μια κίνηση, με την αντίληψη [Τίποτα , Αύρα , Τίποτα , Τίποτα , Τίποτα] .

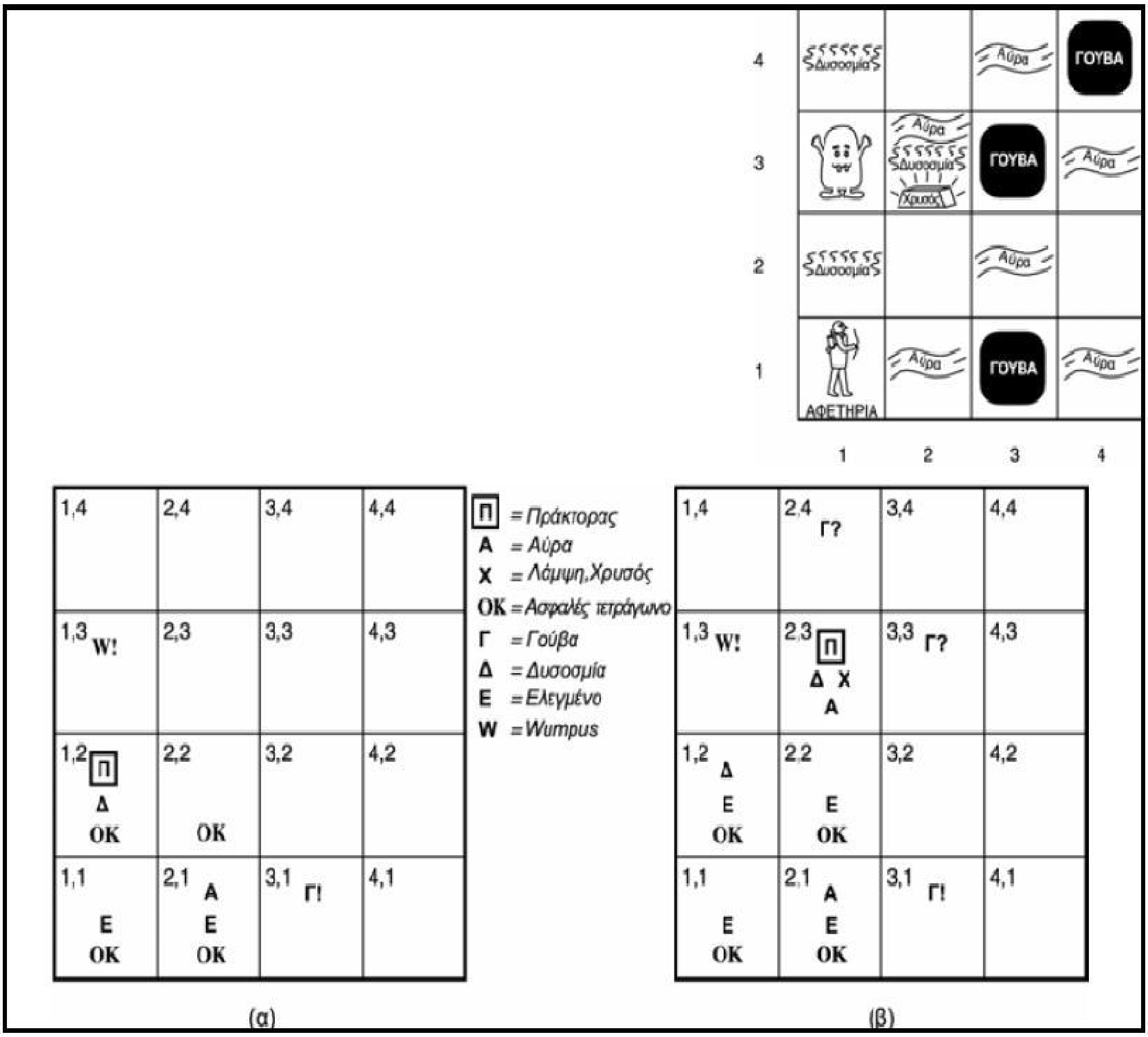
Η πρώτη αντίληψη είναι [Τίποτα, Τίποτα , Τίποτα , Τίποτα , Τίποτα] , από την οποία ο πράκτορας μπορεί να συμπεράνει ότι τα γειτονικά του τετράγωνα είναι ασφαλή. Η εικόνα 0.3(α) δείχνει την κατάσταση γνώσης του πράκτορα σε αυτό το σημείο. Παραθέτουμε (μερικές από) τις προτάσεις της βάσης γνώσης αναπαριστώντας τις με σύμβολα όπως Α(αύρα) και ΟΚ (ασφαλές , ούτε γούβα ούτε wumpus) σημειωμένα στα αντίστοιχα τετράγωνα. Η εικόνα 0.2 , από την άλλη απεικονίζει τον ίδιο τον κόσμο.

Από το γεγονός ότι δεν υπάρχει ούτε δυσοσμία ούτε αύρα στο [1,1] , ο πράκτορας μπορεί να συμπεράνει ότι τα [1,2] και [2,1] είναι ακίνδυνα. Σημειώνονται λοιπόν με ΟΚ , που υποδηλώνει αυτό το γεγονός. Ένας συνετός πράκτορας θα μετακινείται μόνο σε τετράγωνα που γνωρίζει ότι είναι ΟΚ. Ας υποθέσουμε ότι ο πράκτορας αποφασίζει να μετακινηθεί προς τα εμπρός στο [2,1] , καταλήγοντας στη σκηνή της εικόνας 0.3 (β) .

Ο πράκτορας ανιχνεύει αύρα στο [2,1] , επομένως θα πρέπει να υπάρχει γούβα σε ένα γειτονικό τετράγωνο . Η γούβα δεν μπορεί να είναι στο [1,1] , λόγω των κανόνων του παιχνιδιού, και επομένως θα πρέπει να υπάρχει γούβα στο [2,2] ή στο [3,1] ή και στα δύο. Το σύμβολο Γ? στην εικόνα 0.3(β) υποδηλώνει μια ενδεχόμενη γούβα σε αυτά τα τετράγωνα. Στο σημείο αυτό , υπάρχει μόνο ένα γνωστό τετράγωνο που είναι ΟΚ και δεν έχει ελεγχθεί ακόμα. Έτσι ο συνετός πράκτορας θα κάνει στροφή, θα γυρίσει πίσω στο [1,1] και μετά θα προχωρήσει στο [1,2].

Η νέα αντίληψη στο [1,2] είναι [Δυσοσμία, Τίποτα , Τίποτα , Τίποτα , Τίποτα] , με αποτέλεσμα την κατάσταση γνώσης που φαίνεται στην εικόνα 0.4(α) . Η δυσοσμία στο [1,2] σημαίνει ότι θα πρέπει να υπάρχει wumpus κοντά. Όμως το wumpus δεν μπορεί να βρίσκεται στο [1,1], λόγω των κανόνων του παιχνιδιού, ούτε μπορεί να βρίσκεται στο [2,2] (αλλιώς ο πράκτορας θα είχε αντιληφθεί δυσοσμία όταν βρισκόταν στο [2,1]). Επομένως ο πράκτορας μπορεί να συμπεράνει ότι ο wumpus βρίσκεται στο [1,3]. Το σύμβολο W! Υποδηλώνει αυτό το γεγονός. Επίσης, η απουσία *Αύρας* στο [1,2] συνεπάγεται ότι δεν υπάρχει γούβα στο [2,2]. Έχουμε όμως συμπεράνει ήδη ότι θα πρέπει να υπάρχει γούβα είτε στο [2,2] είτε στο [3,1] , οπότε αυτό σημαίνει ότι θα πρέπει να βρίσκεται στο [3,1] . Ο συμπερασμός αυτός είναι αρκετά δύσκολος , επειδή συνδυάζει γνώση που αποκτήθηκε σε διαφορετικούς χρόνους και σε διαφορετικές θέσεις και βασίζεται στην απουσία μιας αντίληψης για να πραγματοποιήσει ένα κρίσιμο βήμα. Ο συμπερασμός υπερβαίνει τις ικανότητες των περισσότερων ζώων , αλλά είναι αντιπροσωπευτικός για το είδος συλλογιστικής που χρησιμοποιεί ένας ευφυής πράκτορας.

Ο πράκτορας έχει τώρα αποδείξει στον εαυτό του ότι δεν υπάρχει ούτε γούβα ούτε wumpus στο [2,2] και επομένως είναι ασφαλές (ΟΚ) να μετακινηθεί εκεί. Δε θα εξετάσουμε την κατάσταση γνώσης του πράκτορα στο [2,2] · θεωρούμε απλώς ότι ο πράκτορας στρίβει και μετακινείται στο [2,3] , καταλήγοντας στην Εικόνα 0.4(β) . Στο [2,3] ο πράκτορας ανιχνεύει λάμψη , επομένως θα πρέπει να αρπάξει το χρυσό και έτσι να τελειώσει το παιχνίδι.



Εικόνα 0.4. Δύο επόμενα στάδια στην πρόοδο του πράκτορα. (α) Μετά την τρίτη κίνηση , με την αντίληψη [Δυσοσμία , Τίποτα , Τίποτα , Τίποτα , Τίποτα] . (β) Μετά την πέμπτη κίνηση , με την αντίληψη [Δυσοσμία , Αύρα , Λάμψη , Τίποτα , Τίποτα] .

Σε κάθε περίπτωση που ο πράκτορας παράγει ένα συμπέρασμα από τις διαθέσιμες πληροφορίες , το συμπέρασμα αυτό είναι εγγυημένα σωστό αν οι διαθέσιμες πληροφορίες είναι σωστές. Αυτό είναι θεμελιώδης ιδιότητα της λογικής συλλογιστικής . Στο υπόλοιπο αυτού του κεφαλαίου περιγράφουμε πώς μπορούν να κατασκευαστούν λογικοί πράκτορες που μπορούν να αναπαριστούν τις απαιτούμενες πληροφορίες και να παράγουν τα συμπεράσματα που περιγράψαμε στις προηγούμενες παραγράφους.

0.3 ΛΟΓΙΚΗ

Αυτή η ενότητα κάνει μια επισκόπηση όλων των θεμελιωδών εννοιών της λογικής αναπαράστασης και συλλογιστικής. Θα αναβάλλουμε τις τεχνικές λεπτομέρειες της οποιασδήποτε συγκεκριμένης μορφής λογικής μέχρι την επόμενη ενότητα. Προς το παρόν θα χρησιμοποιήσουμε άτυπα παραδείγματα από το γνωστό πεδίο της αριθμητικής και τον κόσμο του wumpus. Υιοθετούμε αυτή τη μάλλον ασυνήθιστη προσέγγιση επειδή οι ιδέες της λογικής είναι πολύ πιο γενικές και όμορφες από ότι πιστεύεται συνήθως.

Είδαμε στην προηγούμενη ενότητα ότι οι βάσεις γνώσης αποτελούνται από προτάσεις (sentences). Οι προτάσεις αυτές διατυπώνονται σύμφωνα με τη **σύνταξη** της γλώσσας αναπαράστασης , η οποία καθορίζει όλες τις προτάσεις που είναι καλά διατυπωμένες.Η έννοια της σύνταξης είναι αρκετά σαφής στη συνήθη αριθμητική : Η "x+y=4" είναι μια καλά διατυπωμένη πρόταση ενώ η "x2y+=" δεν είναι. Η σύνταξη των λογικών γλωσσών (και της αριθμητικής,βέβαια) είναι συνήθως σχεδιασμένη για χρήση σε συγγράματα και βιβλία. Υπάρχουν,χωρίς υπερβολή,δεκάδες διαφορετικές συντάξεις,μερικές με πολλά ελληνικά γράμματα και εξωτικά μαθηματικά σύμβολα και μερικές με μάλλον οπτικά ευχάριστα διαγράμματα με βέλη και φυσαλίδες. Σε όλες τις περιπτώσεις όμως,οι προτάσεις της βάσης γνώσης ενός πράκτορα είναι πραγματικές φυσικές διαμορφώσεις του πράκτορα (ή των συνιστωσών του). Η συλλογιστική συνίσταται στην παραγωγή και το χειρισμό αυτών των διαμορφώσεων.

Μια λογική πρέπει επίσης να ορίζει τη **σημασιολογία** (semantics) της γλώσσας. Μπορούμε να πούμε ότι η σημασιολογία έχει να κάνει με το "νόημα" των προτάσεων. Στη λογική ο ορισμός είναι πιο αυστηρός. Η σημασιολογία της γλώσσας ορίζει την **αλήθεια** της κάθε πρότασης σε σχέση με τον κάθε **δυνατό κόσμο** (possible world). Για παράδειγμα, η συνήθης σημασιολογία που υιοθετείται στην αριθμητική ορίζει ότι η πρόταση "x+y=4" είναι αληθής σε έναν κόσμο όπου το x είναι 2 και το y είναι 2 , αλλά ψευδής σε έναν κόσμο όπου το x είναι 1 και το y είναι 1. Στις συνήθεις λογικές , κάθε πρόταση πρέπει να είναι ή αληθής ή ψευδής σε κάθε δυνατό κόσμο – δεν υπάρχει "μέση οδός".

Όπου χρειάζεται να είμαστε ακριβείς, θα χρησιμοποιήσουμε τον όρο **μοντέλο** αντί του όρου "δυνατός κόσμος". (Θα χρησιμοποιούμε επίσης τη φράση "το m είναι ένα μοντέλο της α" εννοώντας ότι η πρόταση α είναι αληθής στο μοντέλο m.). Ενώ οι δυνατοί κόσμοι μπορούν να θεωρούνται ως (δυνητικά) πραγματικά περιβάλλοντα στα οποία μπορεί να βρίσκεται ή να μην βρίσκεται ο πράκτορας, τα μοντέλα είναι μαθηματικές αφαιρέσεις κάθε μία από τις οποίες ρυθμίζει την αλήθεια ή το ψεύδος όλων των σχετικών προτάσεων. Άτυπα μπορούμε να θεωρήσουμε, για παράδειγμα, ότι τα x και y είναι οι αριθμοί των ανδρών και των γυναικών που κάθονται σε ένα τραπέζι και παίζουν μπριτζ και ότι η πρόταση x+y=4 είναι αληθής όταν υπάρχουν συνολικά τέσσερα άτομα· από τυπική άποψη, τα δυνατά μοντέλα είναι απλώς όλες οι δυνατές αναθέσεις αριθμητικών τιμών στις μεταβλητές x και y. Κάθε τέτοια ανάθεση τιμών ρυθμίζει την αλήθεια οποιασδήποτε πρότασης της αριθμητικής, της οποίας μεταβλητές είναι τα x και y.

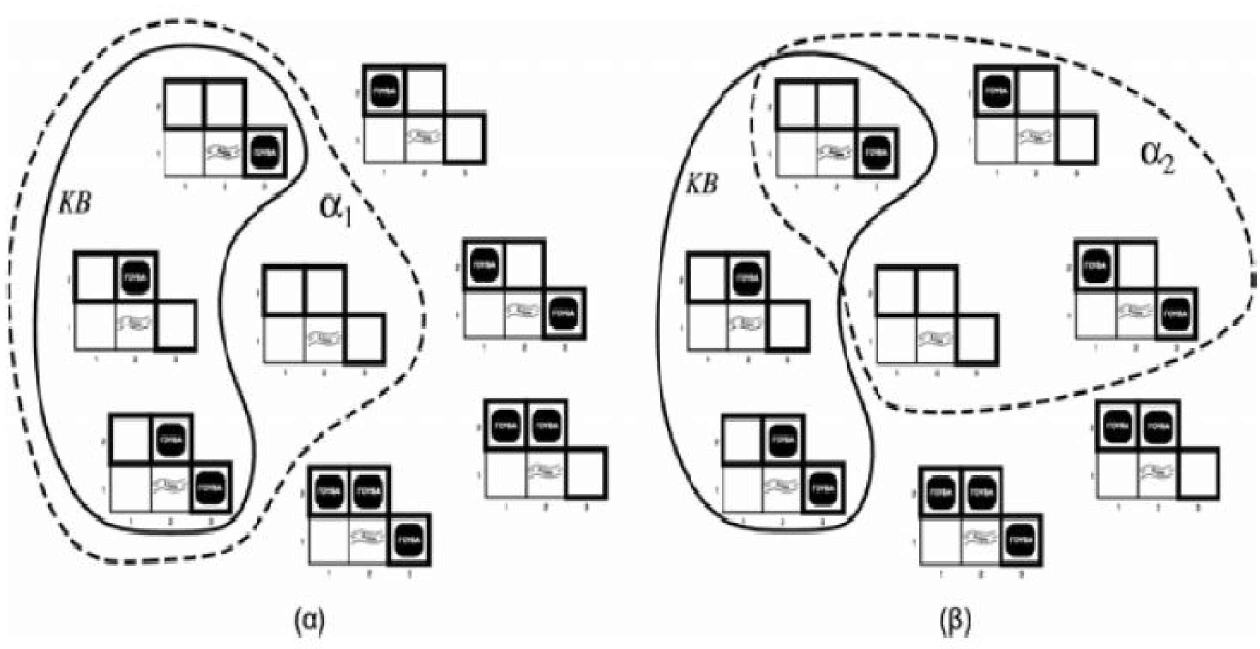
Τώρα έχουμε στη διάθεσή μας μια έννοια της αλήθειας είμαστε έτοιμοι να μιλήσουμε για τη λογική συλλογιστική. Αυτή περιλαμβάνει τη σχέση της **λογικής κάλυψης** (entailment) μεταξύ προτάσεων - την ιδέα ότι μια πρόταση *προκύπτει λογικά* από μια άλλη πρόταση.Στη μαθηματική σημειογραφία, γράφουμε

$$\alpha \models \beta$$

το οποίο σημαίνει ότι η πρόταση α καλύπτει την πρόταση β. Ο τυπικός ορισμός της λογικής κάλυψης είναι ο εξής : $\alpha \models \beta$ εάν και μόνο εάν , σε κάθε μοντέλο στο οποίο η α είναι αληθής, η β είναι επίσης αληθής. Ένας άλλος τρόπος για να πούμε το ίδιο πράγμα είναι ότι αν η α είναι αληθής τότε η β *πρέπει* επίσης να είναι αληθής. Άτυπα μπορούμε να πούμε ότι η αλήθεια της β "περιέχεται" στην αλήθεια της α . Η σχέση κάλυψης μας είναι γνωστή από την αριθμητική· είμαστε ικανοποιημένοι με την ιδέα ότι η πρόταση x+y=4 καλύπτει την πρόταση 4=x+y. Προφανώς , σε οποιοδήποτε μοντέλο όπου x+y=4 - όπως το μοντέλο στο οποίο το x είναι 2 και το y είναι 2 - ισχύει ότι 4=x+y. Θα δούμε σύντομα ότι μια βάση γνώσης μπορεί να θεωρηθεί ως μια δήλωση και συχνά λέμε ότι μια βάση γνώσης καλύπτει μια πρόταση.

Μπορούμε να εφαρμόσουμε το ίδιο είδος ανάλυσης στο παράδειγμα της συλλογιστικής του κόσμου του wumpus. Ας εξετάσουμε την κατάσταση της Εικόνας 1.3 (β) : <u>Ο πράκτορας δεν αντιλαμβάνεται τίποτα στο [1,1] και αντιλαμβάνεται αύρα στο [2,1]. Οι αντιλήψεις αυτές σε συνδυασμό με την γνώση των κανόνων του κόσμου του wumpus απο τον πράκτορα αποτελούν την βάση γνώσης.</u> Ο πράκτορας ενδιαφέρεται (μεταξύ άλλων) για το αν τα γειτονικά τετράγωνα [1,2],[2,2] και [3,1] περιέχουν γούβες. Κάθε ένα από τα τρία τετράγωνα μπορεί να έχει ή να μην έχει γούβα και επομένως (για τους σκοπούς αυτού του παραδείγματος) υπάρχουν $2^3 = 8$ δυνατα μοντέλα. Αυτά παρουσιάζονται στην Εικόνα 0.5.

Η βάση γνώσης είναι ψευδής στα μοντέλα που αντιβαίνουν σε ότι γνωρίζει ο πράκτορας – για παράδειγμα , η βάση γνώσης είναι ψευδής σε οποιοδήποτε μοντέλο στο οποίο το [1,2] περιέχει γούβα , επειδή δεν υπάρχει αύρα στο [1,1]. Για την ακρίβεια υπάρχουν μόνο τρία μοντέλα στα οποία η βάση γνώσης είναι αληθής, και αυτά παρουσιάζονται ως υποσύνολο των μοντέλων της Εικόνας 0.5.



Εικόνα 0.5. Δυνατά μοντέλα για την ύπαρξη γουβών στα τετράγωνα [1,2],[2,2] και [3,1] , με δεδομένο ότι δεν παρατηρήθηκε τίποτα στο [1,1] ενώ παρατηρήθηκε αύρα στο [2,1]. (α) Μοντέλα της βάσης γνώσης και της πρότασης a_1 (δεν υπάρχει γούβα στο [1,2]) . (β) Μοντέλα της βάσης γνώσης και της πρότασης a_2 (δεν υπάρχει γούβα στο [2,2]).

Ας εξετάσουμε τώρα δύο δυνατά συμπεράσματα:

a_1 = "Δεν υπάρχει γούβα στο [1,2]."
a_2 = "Δεν υπάρχει γούβα στο [2,2]."

Έχουμε σημειωμένα τα μοντέλα των προτάσεων a_1 και a_2 στις Εικόνες 0.5(α) και 0.5(β), αντίστοιχα. Παρατηρώντας την εικόνα βλέπουμε τα εξής:

σε κάθε μοντέλο όπου η βάση γνώσης KB είναι αληθής, η πρόταση a_1 είναι επίσης αληθής.

Επομένως, KB $\models$ a_1 : δεν υπάρχει γούβα στο [1,2]. Μπορούμε επίσης να δούμε ότι ,

σε μερικά μοντέλα όπου η βάση γνώσης KB είναι αληθής, η πρόταση a_2 είναι ψευδής.

Επομένως , KB $\not\models$ a_2 : ο πράκτορας *δεν μπορεί* να συμπεράνει ότι δεν υπάρχει γούβα στο [2,2]. (Ούτε μπορεί βέβαια να συμπεράνει ότι υπάρχει γούβα στο [2,2].)

Το προηγούμενο παράδειγμα δεν επιδεικνύει απλώς τη λογική κάλυψη, αλλά δείχνει επίσης πώς μπορεί να εφαρμοστεί ο ορισμός της κάλυψης για την εξαγωγή συμπερασμάτων - δηλαδή , για την πραγματοποίηση **λογικού συμπερασμού** (logical inference). Ο αλγόριθμος συμπερασμού που είδαμε στην Εικόνα 0.5 ονομάζεται **έλεγχος μοντέλων** (model checking), <u>επειδή απαριθμεί όλα τα δυνατά μοντέλα για να ελέγξει ότι η α είναι αληθής σε όλα τα μοντέλα στα οποία η βάση γνώσης ΚΒ είναι αληθής.</u>

Για να γίνει κατανοητή η λογική κάλυψη και ο συμπερασμός, ίσως είναι χρήσιμο να φανταστούμε το σύνολο όλων των συνεπειών της βάσης γνώσης ΚΒ σαν μια θημωνιά και την πρόταση α σαν μια βελόνα. Η λογική κάλυψη μοιάζει με βελόνα μέσα στ'άχυρα· ο συμπερασμός μοιάζει με την εύρεση της. Η διάκριση αυτή είναι ενσωματωμένη σε κάποια τυπική σημειογραφία: Αν ένας αλγόριθμος συμπερασμού i μπορεί να παράγει την πρόταση α από τη βάση γνώσης ΚΒ , τότε γράφουμε

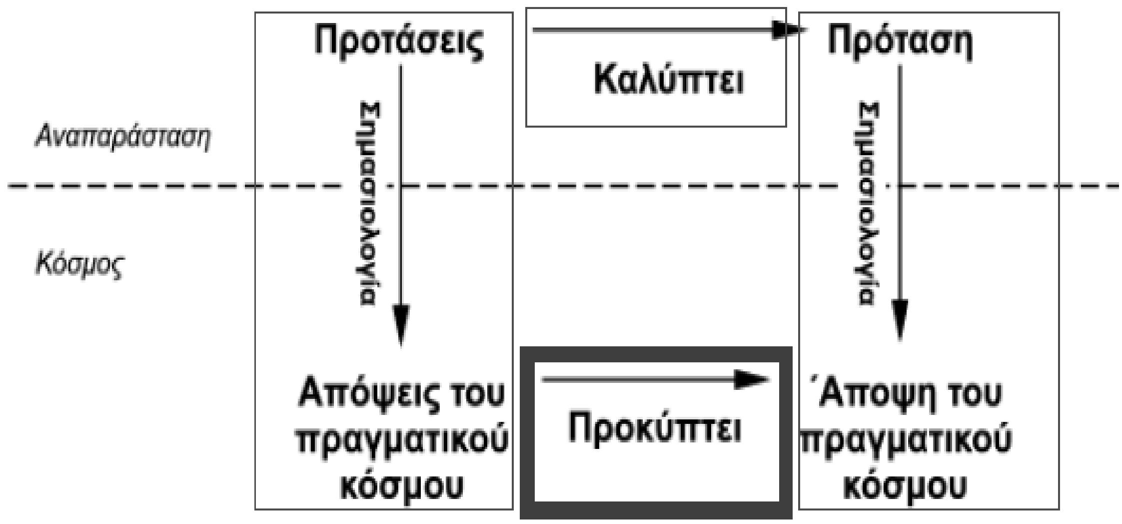
$$KB \vdash_i \alpha \quad ,$$

το οποίο διαβάζεται "η α προκύπτει από την ΚΒ με τον i" ή "ο i παράγει την α από την ΚΒ".

Ένας αλγόριθμος συμπερασμού που παράγει μόνο καλυπτόμενες (entailed) προτάσεις ονομάζεται **ορθός** (sound) ή συμπερασμός που **διατηρεί την αλήθεια** (truth-preserving). Η ορθότητα είναι πολύ επιθυμητή ιδιότητα. Μια μη ορθή διαδικασία συμπερασμού ουσιαστικά επινοεί πράγματα στην πορεία της – ανακοινώνει την ανακάλυψη ανύπαρκτων βελονών. Μπορούμε εύκολα να δούμε ότι ο έλεγχος μοντέλων είναι ορθή διαδικασία όταν είναι εφαρμόσιμος. Αυτό συμβαίνει όταν ο χώρος των μοντέλων είναι πεπερασμένος – για παράδειγμα, σε κόσμους του wumpus με καθορισμένο μέγεθος. Στην αριθμητική από την άλλη , ο χώρος των μοντέλων είναι άπειρο: Άκομα και αν περιοριστούμε στους ακεραίους , υπάρχουν άπειρα ζεύγη τιμών για τα x και y στην πρόταση x+y=4.

Η ιδιότητα της **πληρότητας** (completeness) είναι επίσης επιθυμητή: Ένας αλγόριθμος συμπερασμού είναι πλήρης αν μπορεί να παραγάγει οποιαδήποτε πρόταση που είναι λογικά καλυπτόμενη (entailed).Για τις πραγματικές θημωνιές, οι οποίες έχουν πεπερασμένη έκταση, φαίνεται προφανές ότι μια συστηματική εξέταση μπορεί πάντα να αποφανθεί αν η βελόνα βρίσκεται στα άχυρα.Για πολλές βάσεις γνώσης όμως η θημωνιά των συνεπειών είναι άπειρη, και η πληρότητα γίνεται σημαντικό ζήτημα.Ευτυχώς υπάρχουν πλήρεις διαδικασίες συμπερασμού για τις λογικές που έχουν αρκετά μεγάλη εκφραστική ισχύ ώστε να χειρίζονται πολλές βάσεις γνώσης.

Περιγράψαμε μια διαδικασία συλλογιστικής που τα συμπεράσματα της είναι εγγυημένα αληθή σε οποιονδήποτε κόσμο όπου οι προυποθέσεις είναι αληθείς· ειδικότερα , *αν η βάσης γνώσης ΚΒ είναι αληθής στον πραγματικό κόσμο, τότε οποιαδήποτε πρόταση α που παράγεται από την ΚΒ με μια ορθή διαδικασία συμπερασμού είναι επίσης αληθής στον πραγματικό κόσμο*. Ενώ λοιπόν μια διαδικασία συμπερασμού επενεργεί στη "σύνταξη" - στις εσωτερικές φυσικές διαμορφώσεις όπως τα bit των καταχωρητών ή οι διατάξεις ηλεκτρικών παλμών του εγκεφάλου – η διαδικασία *αντιστοιχεί* στη σχέση με τον πραγματικό κόσμο σύμφωνα με την οποία κάποια άποψη του πραγματικού κόσμου είναι η προκείμενη περίπτωση λόγω του ότι άλλες απόψεις του πραγματικού κόσμου ισχύουν γενικότερα.Αυτή η αντιστοιχία μεταξύ κόσμου και αναπαράστασης παρουσιάζεται στην Εικόνα 0.6.



Εικόνα 0.6 Οι προτάσεις είναι φυσικές διαμορφώσεις του πράκτορα και η συλλογιστική είναι η διαδικασία κατασκευής νέων φυσικών διαμορφώσεων από τις παλιές. Η λογική συλλογιστική θα πρέπει να εξασφαλίζει ότι οι νέες διαμορφώσεις αντιπροσωπεύουν απόψεις του κόσμου που προκύπτουν πραγματικά από τις απόψεις που αντιπροσωπεύουν οι παλιές διαμορφώσεις.

Το τελευταίο ζήτημα που πρέπει να αντιμετωπιστεί σε μια παρουσίαση των λογικών πρακτόρων είναι το ζήτημα της **θεμελίωσης** (grounding) - η σύνδεση , αν υπάρχει , μεταξύ των διαδικασιών λογικής συλλογιστικής και του πραγματικού περιβάλλοντος μέσα στο οποίο υπάρχει ο πράκτορας. Ειδικότερα , *πως γνωρίζουμε ότι η βάση γνώσης ΚΒ είναι αληθής στον πραγματικό κόσμο;* (Στο κάτω-κάτω , η βάση γνώσης είναι απλώς "σύνταξη" μέσα στο κεφάλι του πράκτορα.) Αυτό είναι ένα φιλοσοφικό ερώτημα για το οποίο έχουν γραφεί πάρα πολλά βιβλία. Μια απλή απάντηση είναι ότι τη σύνδεση τη δημιουργούν οι αισθητήρες του πράκτορα. Για παράδειγμα, ο πράκτορας μας από τον κόσμο του wumpus έχει έναν αισθητήρα οσμής. Το πρόγραμμα του πράκτορα δημιουργεί μια κατάλληλη πρόταση όποτε υπάρχει οσμή. Τότε, όποτε η πρόταση αυτή υπάρχει στη βάση γνώσης, είναι αληθής και στον πραγματικό κόσμο. Έτσι το νόημα και η αλήθεια των προτάσεων αντίληψης ορίζονται από τις διαδικασίες αίσθησης και κατασκευής προτάσεων που τις παράγουν. Τι γίνεται όμως με την υπόλοιπη γνώση του πράκτορα, όπως η πεποίθηση του ότι τα wumpus προκαλούν οσμές στα γειτονικά τετράγωνα; Αυτό δεν είναι άμεση αναπαράσταση μιας μεμονωμένης αντίληψης αλλά γενικός κανόνας – που έχει προκύψει ίσως από αισθητηριακή εμπειρία αλλά δεν είναι ταυτόσημος με μια δήλωση αυτής της εμπειρίας. Γενικοί κανόνες σαν αυτόν παράγονται από μια διαδικασία κατασκευής προτάσεων που ονομάζεται **μάθηση** (learning) ,η οποία είναι από μόνη της ένας κλάδος της τεχνητής νοημοσύνης, αλλά δεν θα αναπτυχθεί επί του παρόντος. Η μάθηση δεν είναι αλάνθαστη . Θα μπορούσε να συμβαίνει τα wumpus να αναδίδουν οσμές *πάντα εκτός από την 29η Φεβρουαρίου των δίσεκτων ετών*, επειδή τότε κάνουν το μπάνιο τους. Επομένως η βάση γνώσης KB μπορεί να μην είναι αληθής στον πραγματικό κόσμο, αλλά με καλές διαδικασίες μάθησης υπάρχουν λόγοι για αισιοδοξία.

1. ΠΡΟΤΑΣΙΑΚΗ ΛΟΓΙΚΗ

1.1 ΠΡΟΤΑΣΙΑΚΗ ΛΟΓΙΚΗ : ΜΙΑ ΠΟΛΥ ΑΠΛΗ ΛΟΓΙΚΗ

Θα παρουσιάσουμε τώρα μια πολύ απλή λογική, που ονομάζεται **προτασιακή λογική** (propositional logic). Θα καλύψουμε τη σύνταξη της προτασιακής λογικής και τη σημασιολογία της – τον τρόπο με τον οποίο προσδιορίζεται η αλήθεια των προτάσεων. Έπειτα θα εξετάσουμε τη **λογική κάλυψη** (entailment) – τη σχέση μεταξύ μιας πρότασης και μιας άλλης πρότασης που προκύπτει από αυτήν – και θα δούμε πώς αυτό οδηγεί σε έναν απλό αλγόριθμο για το λογικό συμπερασμό. Φυσικά τα πάντα συμβαίνουν στον κόσμο του wumpus.

Σύνταξη

Η σύνταξη της προτασιακής λογικής ορίζει τις επιτρεπτές προτάσεις. Οι **ατομικές προτάσεις** (atomic sentences) – τα αδιαίρετα συντακτικά στοιχεία – αποτελούνται από ένα μεμονωμένο **προτασιακό σύμβολο** (proposition symbol). Κάθε τέτοιο σύμβολο αντιπροσωπεύει μια ατομική πρόταση που μπορεί να είναι αληθής ή ψευδής. Θα χρησιμοποιούμε κεφαλαία γράμματα για τα σύμβολα: P, Q, R, κ.ο.κ. Τα ονόματα είναι τυχαία αλλά συχνά επιλέγονται έτσι ώστε να έχουν μνημονική αξία για τον αναγνώστη. Για παράδειγμα, θα μπορούσαμε να χρησιμοποιήσουμε το όνομα $W_{1,3}$ για την ατομική πρόταση ότι το wumpus βρίσκεται στο [1,3]. Υπάρχουν δυο προτασιακά σύμβολα με προκαθορισμένο νόημα : Η *Αληθές* είναι η πάντοτε αληθής πρόταση και η *Ψευδές* είναι η πάντοτε ψευδής πρόταση.

Οι σύνθετες προτάσεις (complex sentences) κατασκευάζονται από απλούστερες με τη χρήση λογικών συνδετικών (logical connectives). Τα συνδετικά που χρησιμοποιούνται είναι πέντε:

- $\neg$ (not – όχι). Μια πρόταση όπως η $\neg W_{1,3}$ λέγεται **άρνηση** (negation) της $W_{1,3}$. Ένα **λεκτικό** (literal) είναι είτε ατομική πρόταση (**θετικό λεκτικό**) είτε ατομική πρόταση με άρνηση (**αρνητικό λεκτικό**).

- $\wedge$ (and – και). Μια πρόταση που το κύριο συνδετικό της είναι το $\wedge$, όπως η $W_{1,3} \wedge \Gamma_{3,1}$,λέγεται **σύζευξη** (conjunction)· τα μέλη της είναι οι **συζευκτέοι** της (conjuncts) .

- $\vee$ (or – ή). Μια πρόταση που χρησιμοποιεί το $\vee$, όπως η $(W_{1,3} \wedge \Gamma_{3,1}) \vee W_{2,2}$ είναι μια **διάζευξη** (disjunction) των **διαζευκτέων** (disjuncts) $(W_{1,3} \wedge \Gamma_{3,1})$ και $W_{2,2}$.(Ιστορικά, το $\vee$ προέρχεται από το λατινικό "vel" , που σημαίνει "είτε".)

⇒ (συνεπάγεται). Μια πρόταση όπως η $(W_{1,3} \wedge \Gamma_{3,1}) \Rightarrow \neg W_{2,2}$ λέγεται **συνεπαγωγή** (implication) ή υποθετική πρόταση (conditional). Η **προυπόθεση** (premise) ή το **προηγούμενο** (antecedent) της είναι το $(W_{1,3} \wedge \Gamma_{3,1})$ και το **συμπέρασμα** (conclusion) ή το **επακόλουθο** (consequent) της είναι το **¬** $W_{2,2}$ **.**Οι συνεπαγωγές είναι επίσης γνωστές ως κανόνες (rules) ή προτάσεις εάν – τότε (if -then).

⇔ (εάν και μόνο εάν). Η πρόταση $W_{1,3}$ <=> **¬** $W_{2,2}$ είναι μια **αμφίδρομη υποθετική πρόταση** (biconditional).

Η Εικόνα 1.1 παρουσιάζει μια τυπική γραμματική τη προτασιακής λογικής .

Πρόταση	→	ΑτομικήΠρόταση ∣ ΣύνθετηΠρόταση
ΑτομικήΠρόταση	→	**Αληθές** ∣ **Ψευδές** ∣ Σύμβολο
Σύμβολο	→	**P** ∣ **Q** ∣ **R** ∣ …
ΣύνθετηΠρόταση	→	**¬** Πρόταση
		∣ (Πρόταση ∧ Πρόταση)
		∣ (Πρόταση ∨ Πρόταση)
		∣ (Πρόταση ⇒ Πρόταση)
		∣ (Πρόταση ⇔ Πρόταση)

Εικόνα 1.1 Μια γραμματική BNF (Backus – Naur Form) των προτάσεων της προτασιακής λογικής.

Πρόσεξτε ότι η γραμματική είναι πολύ αυστηρή με τις παρενθέσεις : Κάθε πρόταση που κατασκευάζεται με δυαδικά συνδετικά πρέπει να περικλείεται σε παρενθέσεις. Αυτό εξασφαλίζει ότι η σύνταξη είναι εντελώς μονοσήμαντη. Σημαίνει επίσης ότι πρέπει να γράφουμε , για παράδειγμα ((A∧B)=> C) και όχι A∧B => C . Για να βελτιώσουμε την αναγνωσιμότητα συχνά παραλείπουμε παρενθέσεις, βασιζόμενοι στη σειρά προτεραιότητας των συνδετικών. Η προτεραιότητα αυτή είναι όμοια με την προτεραιότητα που χρησιμοποιείται στην αριθμητική — για παράδειγμα , το ab+c διαβάζεται ως ((ab)+c) και όχι ως a(b+c), επειδή ο πολλαπλασιασμός έχει μεγαλύτερη προτεραιότητα από την πρόσθεση. Η σειρά προτεραιότητας στην προτασιακή λογική είναι (από τη μεγαλύτερη προς τη μικρότερη) : $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$.

Σημασιολογία

Αφού ορίσαμε τη σύνταξη της προτασιακής λογικής, θα ορίσουμε τώρα και τη σημασιολογία της. Η σημασιολογία καθορίζει τους κανόνες για τον προσδιορισμό της αλήθειας μιας πρότασης σε σχέση με ένα συγκεκριμένο μοντέλο. Στην προτασιακή λογική ένα μοντέλο απλώς καθορίζει την τιμή αληθείας — Αληθές ή Ψευδές — για κάθε προτασιακό σύμβολο. Για παράδειγμα , αν οι προτάσεις της βάσης γνώσης χρησιμοποιούν τα προτασιακά σύμβολα $\Gamma_{1,2}$, $\Gamma_{2,2}$ και $\Gamma_{3,1}$, τότε ένα από τα δυνατά μοντέλα είναι

$$m_1 = \{ \Gamma_{1,2} = \text{Ψευδές}, \quad \Gamma_{2,2} = \text{Ψευδές}, \quad \Gamma_{3,1} = \text{Αληθές} \}.$$

Για τρία προτασιακά σύμβολα υπάρχουν $2^3 = 8$ δυνατά μοντέλα — ακριβώς αυτά που εμφανίζονται στην Εικόνα 0.5. Προσέξτε όμως ότι , επειδή έχουμε προσδιορίσει επακριβώς τη σύνταξη , τα μοντέλα γίνονται καθαρά μαθηματικά αντικείμενα χωρίς καμία απαιτούμενη σύνδεση με κόσμους του wumpus. Το $\Gamma_{1,2}$ είναι απλώς ένα σύμβολο · μπορεί να σημαίνει "υπάρχει γούβα στο [1,2]" ή "θα είμαι στο Παρίσι σήμερα και αύριο".

Η σημασιολογία της προτασιακής λογικής πρέπει να καθορίζει πώς υπολογίζεται η τιμή αληθείας οποιασδήποτε πρότασης με δεδομένο ένα μοντέλο. Αυτό γίνεται με αναδρομή. Όλες οι προτάσεις κατασκευάζονται από ατομικές προτάσεις και από τα πέντε συνδετικά · χρειάζεται λοιπόν να καθορίσουμε πώς υπολογίζεται η αλήθεια των ατομικών προτάσεων και πώς υπολογίζεται η αλήθεια των προτάσεων που σχηματίζονται με κάθε έναν από τα πέντε συνδετικά. Οι ατομικές προτάσεις είναι εύκολες :

- Η Αληθές είναι αληθής σε κάθε μοντέλο και η Ψευδές είναι ψευδής σε κάθε μοντέλο.
- Η τιμή αληθείας κάθε άλλου προτασιακού συμβόλου πρέπει να καθορίζεται άμεσα στο μοντέλο. Για παράδειγμα , στο μοντέλο m_1 που είδαμε προηγουμένως η $\Gamma_{1,2}$ είναι ψευδής.

Για τις σύνθετες προτάσεις, έχουμε κανόνες όπως οι παρακάτω :

- Για οποιαδήποτε πρόταση s και για οποιοδήποτε μοντέλο m , η πρόταση **¬ s** είναι αληθής στο m εάν και μόνο εάν η s είναι ψευδής στο m.

Τέτοιοι κανόνες ανάγουν την αλήθεια μιας σύνθετης πρότασης στην αλήθεια απλούστερων προτάσεων. Οι κανόνες για το κάθε συνδετικό μπορούν να συνοψιστούν σε έναν πίνακα αληθείας (truth table) που καθορίζει την τιμή αληθείας μιας σύνθετης πρότασης για κάθε δυνατή ανάθεση τιμών αληθείας στις συνιστώσες της. Οι πίνακες αληθείας για τα πέντε λογικά συνδετικά παρουσιάζονται στην Εικόνα 1.2. Με τη χρήση αυτών των πινάκων , μπορεί να υπολογιστεί η τιμή αληθείας οποιασδήποτε πρότασης s σε σχέση με οποιοδήποτε μοντέλο m με μια απλή διαδικασία αναδρομικής αποτίμησης.

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
Ψευδές	Ψευδές	Αληθές	Ψευδές	Ψευδές	Αληθές	Αληθές
Ψευδές	Αληθές	Αληθές	Ψευδές	Αληθές	Αληθές	Ψευδές
Αληθές	Ψευδές	Ψευδές	Ψευδές	Αληθές	Ψευδές	Ψευδές
Αληθές	Αληθές	Ψευδές	Αληθές	Αληθές	Αληθές	Αληθές

Εικόνα 1.2 Πίνακες αληθείας για τα πέντε λογικά συνδετικά. Για να χρησιμοποιήσετε τον πίνακα για να υπολογίσετε, για παράδειγμα , την τιμή της $P \vee Q$ όταν η P είναι αληθής και η Q είναι ψευδής, κοιτάξτε πρώτα στα αριστερά για τη γραμμή όπου το P είναι *Αληθές* και το Q είναι *Ψευδές* (τρίτη γραμμή). Έπειτα κοιτάξτε στην ίδια γραμμή στη στήλη $P \vee Q$ και θα βρείτε το αποτέλεσμα Αληθές. Ένας άλλος τρόπος να το δείτε είναι να θεωρήσετε ότι η κάθε γραμμή είναι ένα μοντέλο και ότι οι καταχωρήσεις στην κάθε στήλη σε αυτή τη γραμμή λένε αν η αντίστοιχη πρόταση είναι αληθής σε αυτό το μοντέλο.

Αναφέραμε προηγουμένως ότι μια βάση γνώσης αποτελείται από ένα σύνολο προτάσεων. Μπορούμε να δούμε τώρα ότι μια λογική βάση γνώσης είναι μια σύζευξη αυτών των προτάσεων. Δηλαδή, αν ξεκινήσουμε με μια κενή βάση γνώσης KB και πραγματοποιήσουμε τις πράξεις TELL(KB , S_1) … TELL (KB , S_n) τότε έχουμε KB = $S_1 \wedge \ldots \wedge S_n$. Αυτό σημαίνει ότι μπορούμε να αντιμετωπίζουμε τις βάσεις γνώσης και τις προτάσεις με τον ίδιο τρόπο.

Οι πίνακες αληθείας για τα συνδετικά "και", "ή" και "όχι" είναι σύμφωνοι με τη διαισθητική μας αντίληψη για τις αντίστοιχες λέξεις της φυσικής γλώσσας. Το κύριο σημείο όπου μπορεί να υπάρξει σύγχυση είναι το ότι η πρόταση $P \vee Q$ είναι αληθής όταν η P είναι αληθής ή η Q είναι αληθής ή και οι δύο. Υπάρχει ένα άλλο συνδετικό που ονομάζεται "αποκλειστικό ή" ("exclusive or" ή για συντομία "xor") , το οποίο δίνει τιμή Ψευδές όταν και οι δύο διαζευκτέοι είναι αληθείς. Δεν υπάρχει ομοφωνία για το σύμβολο του αποκλειστικού ή .

1.2 Μια απλή βάση γνώσης

Αφού ορίσαμε τη σημασιολογία της προτασιακής λογικής μπορούμε να κατασκευάσουμε μια βάση γνώσης για τον κόσμο του wumpus. Για να κρατήσουμε τα πράγματα απλά θα ασχοληθούμε μόνο με τις γούβες. Θα την εφοδιάσουμε με επαρκή γνώση για την πραγματοποίηση του συμπερασμού που έγινε άτυπα στην Ενότητα 0.3.

Πρώτα χρειάζεται να επιλέξουμε το λεξιλόγιό μας για τα σύμβολα των προτάσεων. Για κάθε i , j :

- Έστω $\Gamma_{i,j}$ αληθές αν υπάρχει γούβα στο [i , j] .
- Έστω $A_{i,j}$ αληθές αν υπάρχει αύρα στο [i , j] .

Η βάση γνώσης περιλαμβάνει τις παρακάτω προτάσεις , κάθε μία με την ετικέτα της για ευκολία:

- Δεν υπάρχει γούβα στο [1,1]:

  $R_1 : \neg \Gamma_{1,1}$.

- Ένα τετράγωνο έχει αύρα εάν και μόνο εάν υπάρχει γούβα σε γειτονικό τετράγωνο . Αυτό πρέπει να δηλωθεί για κάθε τετράγωνο · για την ώρα , θα συμπεριλάβουμε μόνο τα τετράγωνα που μας αφορούν:

  $R_2 : A_{1,1} \Leftrightarrow (\Gamma_{1,2} \vee \Gamma_{2,1})$

  $R_3 : A_{2,1} \Leftrightarrow (\Gamma_{1,1} \vee \Gamma_{2,2} \vee \Gamma_{3,1})$

- Οι προηγούμενες προτάσεις είναι αληθείς σε όλους τους κόσμους wumpus. Τώρα θα συμπεριλάβουμε τις αντιλήψεις αύρας για τα δύο πρώτα τετράγωνα που επισκεφτήκαμε στο συγκεκριμένο κόσμο που βρίσκεται ο πράκτορας, το οποίο μας οδηγεί στην κατάσταση της Εικόνας 0.3(β).

  $R_4 : \neg A_{1,1}$

  $R_5 : A_{2,1}$.

Η βάση γνώσης αποτελείται τώρα από τις προτάσεις R_1 μέχρι R_5 . Μπορεί επίσης να θεωρηθεί ως μία και μόνο πρόταση – ως σύζευξη $R_1 \wedge R_2 \wedge R_3 \wedge R_4 \wedge R_5$ – επειδή ισχυρίζεται ότι όλες οι μεμονωμένες προτάσεις είναι αληθείς.

1.3 Συμπερασμός με απαρίθμηση μοντέλων.

Θυμηθείτε ότι σκοπός του λογικού συμπερασμού είναι να αποφασίζεται αν ΚΒ |= α για κάποια πρόταση α. Για παράδειγμα, καλύπτεται η $\Gamma_{2,2}$; Ο πρώτος μας αλγόριθμος συμπερασμού θα είναι μια άμεση υλοποίηση του ορισμού της λογικής κάλυψης : Απαριθμεί τα μοντέλα και ελέγχει αν η α είναι αληθής σε κάθε μοντέλο στο οποίο η βάση γνώσης ΚΒ είναι αληθής. Για την προτασιακή λογική , τα μοντέλα είναι αναθέσεις τιμών Αληθές ή Ψευδές σε κάθε προτασιακό σύμβολο. Επιστρέφοντας στον κόσμο του wumpus του παραδείγματος μας, τα σχετικά προτασιακά σύμβολα είναι τα $A_{1,1}, A_{2,1}, \Gamma_{1,1}, \Gamma_{2,1}, \Gamma_{2,2}$ και $\Gamma_{3,1}$. Με επτά σύμβολα, υπάρχουν $2^7 = 128$ δυνατά μοντέλα· στα τρία από αυτά η ΚΒ είναι αληθής (Εικόνα 1.3). Σε αυτά τα τρία μοντέλα η $\neg\Gamma_{1,2}$ είναι αληθής, και επομένως δεν υπάρχει γούβα στο [1,2]. Από την άλλη , η $\Gamma_{2,2}$ είναι αληθής στα δύο από τα τρία μοντέλα και ψευδής στο ένα, γι'αυτό δεν μπορούμε ακόμα να πούμε αν υπάρχει γούβα στο [2,2].

Η Εικόνα 1.3 αναπαριστά σε πιο ακριβή μορφή τη συλλογιστική που είδαμε στην Εικόνα 0.5. Ένας γενικός αλγόριθμος που αποφασίζει για τη λογική κάλυψη στην προτασιακή λογική παρουσιάζεται στην Εικόνα 1.4. Ο αλγόριθμος αυτός TT-ENTAILS (Truth Table Entails) πραγματοποιεί αναδρομική απαρίθμηση σε έναν πεπερασμένο χώρο αναθέσεων τιμών σε μεταβλητές. Ο αλγόριθμος είναι **ορθός** (sound), επειδή υλοποιεί άμεσα τον ορισμό της λογικής κάλυψης, και **πλήρης** (complete) , επειδή λειτουργεί για οποιεσδήποτε προτάσεις ΚΒ και α και πάντα τερματίζει – υπάρχει πεπερασμένο πλήθος μοντέλων προς εξέταση.

$A_{1,1}$	$A_{2,1}$	$\Gamma_{1,1}$	$\Gamma_{1,2}$	$\Gamma_{2,1}$	$\Gamma_{2,2}$	$\Gamma_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
false	*false*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*false*	*false*
false	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*false*	*true*	*false*	*false*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
false	*true*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*false*	*true*	*true*	*false*
false	*true*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*false*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*true*	*false*	*false*	*true*	*false*	*false*	*true*	*true*	*false*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
true	*true*	*true*	*true*	*true*	*true*	*true*	*false*	*true*	*true*	*false*	*true*	*false*

Εικόνα 1.3 Πίνακας αληθείας κατασκευασμένος για τη βάση γνώσης που περιγράψαμε στο κείμενο. Η ΚΒ είναι αληθής αν οι προτάσεις R_1 εώς R_5 είναι αληθείς, το οποίο συμβαίνει σε μόλις 3 από τις 128 γραμμές. Και στις 3 γραμμές η $\Gamma_{1,2}$ είναι ψευδής και επομένως δεν υπάρχει γούβα στο [1,2]. Από την άλλη, μπορεί να υπάρχει (ή να μην υπάρχει) γούβα στο [2,2].

function TT-ENTAILS? (KB , *α*) **returns** *Αληθές ή Ψευδές*
inputs: KB , η βάση γνώσης , μια πρόταση της προτασιακής λογικής *α* ,
το ερώτημα , μια πρόταση της προτασιακής λογικής

σύμβολα ← λίστα των προτασιακών συμβόλων της KB και της *α*
return TT-CHECK-ALL(KB, *α* , σύμβολα, [ ])

function TT-CHECK-ALL (KB , *α* , σύμβολα , μοντέλο) **returns** *Αληθές ή Ψευδές*
if EMPTY? (*σύμβολα*) **then**
if PL-TRUE? (KB , μοντέλο) **then return** PL-TRUE? (*α* , μοντέλο)
else return Αληθές
else do
P ← FIRST(*σύμβολα*); υπόλοιπα ← REST (*σύμβολα*)
return TT-CHECK-ALL(KB, *α* , υπόλοιπα, EXTEND(P , Αληθές , μοντέλο)
and
TT-CHECK-ALL(KB, *α* , υπόλοιπα, EXTEND(P , Ψευδές , μοντέλο)

Εικόνα 1.4 Αλγόριθμος απαρίθμησης πίνακα αληθείας που αποφασίζει για την προτασιακή λογική κάλυψη. TT είναι ο πίνακας αληθείας. Η συνάρτηση PL-TRUE? επιστρέφει *Αληθές* αν μια πρόταση ισχύει μέσα στα πλαίσια ενός μοντέλου. Η μεταβλητή μοντέλο αντιπροσωπεύει ένα μερικό μοντέλο – μια ανάθεση τιμών σε μερικές μόνο από τις μεταβλητές. Η κλήση της συνάρτησης EXTEND(P , *Αληθές* , *μοντέλο*) επιστρέφει ένα νέο μερικό μοντέλο στο οποίο η πρόταση P έχει την τιμή *Αληθές*.

Φυσικά, "πεπερασμένο πλήθος" δε σημαίνει πάντα "λίγα". Αν οι προτάσεις KB και *α* περιέχουν **n** σύμβολα συνολικά , τότε υπάρχουν 2^n μοντέλα.
Επομένως η χρονική πολυπλοκότητα του αλγορίθμου είναι O(2^n) .
(Η χωρική πολυπλοκότητα είναι μόνο O(n) , επειδή η απαρίθμηση γίνεται πρώτα σε βάθος.)

Δυστυχώς , κάθε γνωστός αλγόριθμος συμπερασμού για την προτασιακή λογική έχει μια πολυπλοκότητα **χειρότερης περίπτωσης** που είναι **εκθετική** ως προς το μέγεθος της εισόδου. Δεν αναμένουμε να πετύχουμε τίποτα καλύτερο από αυτό, επειδή η προτασιακή λογική κάλυψη είναι co-NP πλήρης .

Λίγα περί κλάσεων πολυπλοκότητας.

Η κλάση P είναι η κλάση των προβλημάτων που μπορούν να επιλυθούν σε πολυωνυμικό χρόνο .
Η κλάση NP είναι η κλάση των μη αιτιοκρατικών πολυωνυμικών προβλημάτων. Ένα πρόβλημα εμπίπτει σε αυτή την κλάση εάν υπάρχει κάποιος αλγόριθμος ο οποίος μπορεί να μαντέψει μια λύση και στη συνέχεια να εξακριβώσει εάν ισχύει αυτή η πρόβλεψη σε πολυωνυμικό χρόνο.
Η κλάση **co-NP** είναι το συμπλήρωμα της NP, με την έννοια ότι για κάθε πρόβλημα απόφασης της κλάσης NP, υπάρχει ένα αντίστοιχο πρόβλημα στην κλάση co-NP όπου οι απαντήσεις "ναι" και "όχι" είναι αντεστραμμένες. Γνωρίζουμε ότι η κλάση P είναι υποσύνολο τόσο της κλάσης NP όσο και της κλάσης co-NP, και πιστεύεται ότι υπάρχουν προβλήματα της κλάσης co-NP που δεν ανήκουν στην P.
Με τον όρο co-NP πλήρης εννοούμε τα δυσκολότερα προβλήματα της κλάσης co-NP.

Ισοδυναμία , εγκυρότητα και ικανοποιησιμότητα

Πριν προχωρήσουμε στις λεπτομέρειες των αλγορίθμων λογικού συμπερασμού , θα χρειαστούμε μερικές επιπλέον έννοιες σχετικές με τη λογική κάλυψη (entailment).

Η πρώτη έννοια είναι η **λογική ισοδυναμία** (logical equivalence) : Δύο προτάσεις α και β είναι λογικά ισοδύναμες αν είναι αληθείς στο ίδιο σύνολο μοντέλων. Αυτό γράφεται $\alpha \Leftrightarrow \beta$

– $(a \wedge \beta) \equiv (\beta \wedge a)$	αντιμεταθετικότητα του ∧
– $(a \vee \beta) \equiv (\beta \vee a)$	αντιμεταθετικότητα του ∨
– $((a \wedge \beta) \wedge \gamma) \equiv (a \wedge (\beta \wedge \gamma))$	προσεταιριστικότητα του ∧
– $((a \vee \beta) \vee \gamma) \equiv (a \vee (\beta \vee \gamma))$	προσεταιριστικότητα του ∨
– $\neg(\neg a) \equiv a$	απαλοιφή διπλής άρνησης
– $(a \Rightarrow \beta) \equiv (\neg\beta \Rightarrow \neg a)$	αντιθετοαντιστροφή
– $(a \Rightarrow \beta) \equiv (\neg a \vee \beta)$	απαλοιφή συνεπαγωγής
– $(a \Leftrightarrow \beta) \equiv ((a \Rightarrow \beta) \wedge (\beta \Rightarrow a))$	απαλοιφή αμφίδρομης συνεπαγωγής
– $\neg(a \wedge \beta) \equiv (\neg a \vee \neg\beta)$	νόμος De Morgan
– $\neg(a \vee \beta) \equiv (\neg a \wedge \neg\beta)$	νόμος De Morgan
– $(a \wedge (\beta \vee \gamma)) \equiv ((a \wedge \beta) \vee (a \wedge \gamma))$	επιμεριστικότητα του ∧ ως προς το ∨
– $(a \vee (\beta \wedge \gamma)) \equiv ((a \vee \beta) \wedge (a \vee \gamma))$	επιμεριστικότητα του ∨ ως προς το ∧

Εικόνα 1.5 Συνήθεις λογικές ισοδυναμίες, Τα σύμβολα *α*, β και *γ* αντιπροσωπεύουν οποιεσδήποτε προτάσεις της προτασιακής λογικής.

Η δεύτερη έννοια που θα χρειαστούμε είναι η **εγκυρότητα** (validity). Μια πρόταση είναι έγκυρη αν είναι αληθής σε *όλα* τα μοντέλα. Για παράδειγμα, η πρόταση $P \vee \neg P$ είναι έγκυρη. Οι έγκυρες προτάσεις είναι γνωστές και ως **ταυτολογίες –** είναι *αναγκαία* αληθείς και επομένως κενές από περιεχόμενο, Επειδή η πρόταση *Αληθές* είναι αληθής σε όλα τα μοντέλα, κάθε έγκυρη πρόταση είναι λογικά ισοδύναμη με την *Αληθές*.

Τι χρησιμότητα έχουν λοιπόν οι έγκυρες προτάσεις ; Από τον ορισμό που δώσαμε για τη λογική κάλυψη μπορεί να προκύψει το **θεώρημα της παραγωγής** (deduction theorem) , το οποίο ήταν ήδη γνωστό στους αρχαίους Έλληνες :

Για οποιεσδήποτε προτάσεις α και β , α $\models$ *β εάν και μόνο εάν η πρόταση (α ⇒ β) είναι έγκυρη .*

Μπορούμε να θεωρήσουμε ότι ο αλγόριθμος συμπερασμού της Εικόνας 1.4 εξετάζει την εγκυρότητα της (KB ⇒ α). Αντίστροφα , κάθε έγκυρη πρόταση συνεπαγωγής περιγράφει ένα νόμιμο συμπερασμό.

Η τελευταία έννοια που θα χρειαστούμε είναι η **ικανοποιησιμότητα** (satisfiability) . Μια πρόταση είναι ικανοποιήσιμη αν είναι αληθής σε κάποια μοντέλα. Για παράδειγμα , η βάση γνώσης που εξετάσαμε προηγουμένως , η $R_1 \wedge R_2 \wedge R_3 \wedge R_4 \wedge R_5$, είναι ικανοποιήσιμη επειδή υπάρχουν τρία μοντέλα στα οποία είναι αληθής, όπως είδαμε στον πίνακα της Εικόνας 0.3. Αν μια πρόταση α είναι αληθής σε ένα μοντέλο m , τότε λέμε ότι το m **ικανοποιεί** την α ή ότι το m **είναι μοντέλο της** α. Η ικανοποιησιμότητα μπορεί να ελεγχθεί με απαρίθμηση των δυνατών μοντέλων, μέχρι να βρεθεί κάποιο που να ικανοποιεί την πρόταση. <u>Ο προσδιορισμός της ικανοποιησιμότητας των προτάσεων στην προτασιακή λογική ήταν το πρώτο πρόβλημα που αποδείχθηκε ότι είναι NP -πλήρες .</u>

Πολλά προβλήματα της επιστήμης των υπολογιστών είναι στην πραγματικότητα προβλήματα ικανοποιησιμότητας.

Η εγκυρότητα και η ικανοποιησιμότητα είναι φυσικά αλληλένδετες : η α είναι έγκυρη εάν και μόνο εάν η $\neg\alpha$ είναι μη ικανοποιήσιμη · με αντιθετοαντιστροφή (contraposition) , η α είναι ικανοποιήσιμη εάν και μόνο εάν η $\neg\alpha$ δεν είναι έγκυρη. Έχουμε επίσης το παρακάτω χρήσιμο συμπέρασμα :

α $\models$ *β εάν και μόνο εάν η πρόταση (* $\alpha \wedge \neg\beta$ *) είναι μη ικανοποιήσιμη.*

Η απόδειξη της β από την α με έλεγχο της μη ικανοποιησιμότητας της ($\alpha \wedge \neg\beta$) αντιστοιχεί ακριβώς στην καθιερωμένη μαθηματική αποδεικτική τεχνική της **απαγωγής σε άτοπο** (δηλαδή αναγωγή σε κάτι παράλογο). Λέγεται επίσης απόδειξη μέσω **διάψευσης** ή απόδειξη μέσω **αντίφασης**. Υποθέτουμε ότι μια πρόταση β είναι ψευδής και αποδεικνύουμε ότι αυτό οδηγεί σε αντίφαση με κάποια γνωστά αξιώματα α. Αυτήν ακριβώς την αντίφαση εννοούμε όταν λέμε ότι η πρόταση ($\alpha \wedge \neg\beta$) είναι μη ικανοποιήσιμη .

1.4 <u>ΣΥΜΠΕΡΑΣΜΟΣ ΜΕ ΑΠΟΔΕΙΞΗ</u>

<u>ΠΡΟΤΥΠΑ ΣΥΛΛΟΓΙΣΤΙΚΗΣ ΣΤΗΝ ΠΡΟΤΑΣΙΑΚΗ ΛΟΓΙΚΗ</u>

Η ενότητα αυτή καλύπτει καθιερωμένα πρότυπα συμπερασμού που μπορούν να εφαρμόζονται για να προκύπτουν αλυσίδες συμπερασμάτων που οδηγούν στον επιθυμητό στόχο. Αυτά τα πρότυπα συμπερασμού λέγονται **κανόνες συμπερασμού** (inference rules). Ο πιο γνωστός κανόνας ονομάζεται **Τρόπος του Θέτειν** (modus ponens) και γράφεται έτσι :

$$\frac{\alpha \Rightarrow \beta \quad , \alpha}{\beta}$$

Ο συμβολισμός αυτός σημαίνει ότι , όποτε μας δοθούν δύο οποιεσδήποτε προτάσεις της μορφής
$\alpha \Rightarrow \beta$ και α , μπορούμε να συμπεράνουμε την πρόταση β.

Ένας άλλος χρήσιμος κανόνας συμπερασμού είναι η **απαλοιφή του Και** (And-Elimination), ο οποίος λέει ότι από μια σύζευξη μπορούμε να συμπεράνουμε οποιονδήποτε από τους όρους της :

$$\frac{\alpha \wedge \beta}{\alpha}$$

Εξετάζοντας τις δυνατές τιμές αληθείας των α και β , μπορεί κανείς εύκολα να αποδείξει μια για πάντα ότι ο Modus Ponens και η απαλοιφή του Και είναι ορθοί κανόνες. <u>Οι κανόνες αυτοί μπορούν λοιπόν να χρησιμοποιούνται σε οποιαδήποτε συγκεκριμένα στιγμιότυπα όπου είναι εφαρμόσιμοι , παράγοντας ορθούς συμπερασμούς χωρίς να χρειάζεται να απαριθμούνται μοντέλα.</u>

Όλες οι λογικές ισοδυναμίες της Εικόνας 1.5 μπορούν να χρησιμοποιούνται ως κανόνες συμπερασμού.
Όμως δεν λειτουργούν όλοι οι κανόνες συμπερασμού και προς τις δύο κατευθύνσεις όπως οι λογικές ισοδυναμίες. Για παράδειγμα, δεν μπορούμε να εφαρμόσουμε το Modus Ponens προς την αντίθετη κατεύθυνση για να πάρουμε τις προτάσεις $\alpha \Rightarrow \beta$ και α από τη β .

Ας δούμε πώς μπορούν να χρησιμοποιηθούν αυτοί οι κανόνες συμπερασμού και οι ισοδυναμίες στον κόσμο του wumpus. Θα ξεκινήσουμε από τη βάση γνώσης που περιέχει τις προτάσεις R_1 εώς R_5 και θα δείξουμε πώς μπορούμε να αποδείξουμε την $\neg\Gamma_{1,2}$, δηλαδή ότι δεν υπάρχει γούβα στο [1,2]. Πρώτα εφαρμόζουμε απαλοιφή αμφίδρομης υποθετικής πρότασης στην R_2 και παίρνουμε :

$$R_6 : (A_{1,1} \Rightarrow (\Gamma_{1,2} \vee \Gamma_{2,1})) \wedge ((\Gamma_{1,2} \vee \Gamma_{2,1}) \Rightarrow A_{1,1}) \quad .$$

Έπειτα εφαρμόζουμε απαλοιφή του Και στην R_6 και παίρνουμε :

$$R_7 : ((\Gamma_{1,2} \vee \Gamma_{2,1}) \Rightarrow A_{1,1})$$

Η λογική ισοδυναμία για τις αντιθετοαντίστροφες προτάσεις μας δίνει :

$$R_8 : (\neg A_{1,1} \Rightarrow \neg(\Gamma_{1,2} \vee \Gamma_{2,1}))$$

Τώρα μπορούμε να εφαρμόσουμε το Modus Ponens με την R_8 και την αντίληψη R_4 (δηλαδή την $\neg A_{1,1}$) και να πάρουμε :

$$R_9 : \neg(\Gamma_{1,2} \vee \Gamma_{2,1}) \quad .$$

Τέλος , εφαρμόζουμε το νόμο του De Morgan και καταλήγουμε στο συμπέρασμα :

$$R_{10} : \neg\Gamma_{1,2} \wedge \neg\Gamma_{2,1} \quad .$$

Δηλαδή , ούτε το [1,2] ούτε το [2,1] έχουν γούβα.

Η παραπάνω συλλογιστική διαδικασία — μια ακολουθία εφαρμογών των κανόνων συμπερασμού — ονομάζεται **απόδειξη** (proof). Η εύρεση αποδείξεων είναι ακριβώς σαν την εύρεση λύσεων σε προβλήματα αναζήτησης. Αν μάλιστα η συνάρτηση διαδόχων οριστεί έτσι ώστε να παράγει όλες τις δυνατές εφαρμογές κανόνων συμπερασμού , τότε όλοι οι αλγόριθμοι αναζήτησης μπορούν να εφαρμόζονται και για την εύρεση αποδείξεων. **Έτσι, η αναζήτηση αποδείξεων είναι μια μέθοδος εναλλακτική της απαρίθμησης μοντέλων.** Η αναζήτηση μπορεί να προχωρά από την αρχική βάση γνώσης προς τα εμπρός εφαρμόζοντας κανόνες συμπερασμού για να παράγει την πρόταση-στόχο , ή μπορεί να προχωρά από την πρόταση στόχο προς τα πίσω προσπαθώντας να βρει μια αλυσίδα κανόνων συμπερασμού που οδηγεί στην αρχική βάση γνώσης.

Το γεγονός ότι ο συμπερασμός στην προτασιακή λογική είναι NP-πλήρης υποδηλώνει ότι, στην χειρότερη περίπτωση , η αναζήτηση αποδείξεων δε θα είναι περισσότερο αποδοτική από την απαρίθμηση μοντέλων. Σε πολλές πρακτικές *περιπτώσεις όμως* , *η εύρεση μιας απόδειξης μπορεί να είναι πολύ αποδοτική απλώς επειδή μπορεί να αγνοήσει τις άσχετες ατομικές προτάσεις, όσο πολλές και να υπάρχουν*. Για παράδειγμα, στην απόδειξη που δώσαμε παραπάνω , η οποία οδηγεί στην $\neg\Gamma_{1,2} \wedge \neg\Gamma_{2,1}$, οι προτάσεις $A_{2,1}, \Gamma_{1,1}, \Gamma_{2,2}$ ή $\Gamma_{3,1}$ δεν αναφέρονται. Μπορούν να αγνοηθούν επειδή η πρόταση-στόχος $\Gamma_{1,2}$ εμφανίζεται μόνο στην πρόταση R_4 · οι άλλες ατομικές προτάσεις της R_4 εμφανίζονται μόνο στις R_4 και R_2 · έτσι, οι R_1, R_3 *και* R_5 δεν έχουν καμία επίδραση στην απόδειξη. Το ίδιο θα ίσχυε ακόμα και αν προσθέταμε άλλο ένα εκατομμύριο προτάσεις στη βάση γνώσης· ο απλός αλγόριθμος του πίνακα αληθείας, από την άλλη, θα κατέρρεε από την εκθετική έκρηξη του αριθμού των μοντέλων.

Αυτή η ιδιότητα των λογικών συστημάτων στην πραγματικότητα προκύπτει από μια πολύ πιο θεμελιώδη ιδιότητα που ονομάζεται **μονοτονικότητα** (monotonicity). Η μονοτονικότητα λέει ότι το σύνολο των λογικά καλυπτόμενων προτάσεων μπορεί μόνο να *αυξάνει* καθώς προστίθενται πληροφορίες στη βάση γνώσης. (Οι **μη μονοτονικές** (nonmonotonic) λογικές, οι οποίες παραβιάζουν την ιδιότητα της μονοτονικότητας, εμπεριέχουν μια κοινή ιδιότητα της ανθρώπινης συλλογιστικής: την αλλαγή γνώμης.) Για οποιεσδήποτε προτάσεις α και β:

εάν KB $\models$ α τότε $KB \wedge \beta \models \alpha$.

Για παράδειγμα, ας υποθέσουμε ότι η βάση γνώσης περιέχει τον πρόσθετο ισχυρισμό β που λέει ότι υπάρχουν ακριβώς οκτώ γούβες στον κόσμο. Η γνώση αυτή θα μπορούσε να βοηθήσει τον πράκτορα να αντλήσει *πρόσθετα* συμπεράσματα , δεν μπορεί όμως να αναιρέσει οποιοδήποτε συμπέρασμα α που έχει ήδη εξαχθεί – όπως το συμπέρασμα ότι δεν υπάρχει γούβα στο [1,2]. Η μονοτονικότητα σημαίνει ότι οι κανόνες συμπερασμού μπορούν να εφαρμόζονται οποτεδήποτε βρίσκονται κατάλληλες προυποθέσεις στη βάση γνώσης – το συμπέρασμα του κανόνα πρέπει να προκύπτει *άσχετα από το τι άλλο υπάρχει στη βάση γνώσης*.

1.5 ΣΥΜΠΕΡΑΣΜΟΣ ΜΕ ΑΝΑΛΥΣΗ

Υποστηρίξαμε ότι οι κανόνες συμπερασμού που καλύψαμε μέχρι εδώ είναι ορθοί (sound), αλλά δε συζητήσαμε το ζήτημα της πληρότητας των αλγορίθμων συμπερασμού που τους χρησιμοποιούν. Οι αλγόριθμοι αναζήτησης όπως η επαναληπτική εκβάθυνση είναι πλήρεις, με την έννοια ότι βρίσκουν οποιονδήποτε προσπελάσιμο κόμβο στόχου, αν όμως οι διαθέσιμοι κανόνες συμπερασμού είναι ανεπαρκείς τότε ο κόμβος στόχου δεν είναι προσπελάσιμος – δεν υπάρχει απόδειξη που να χρησιμοποιεί μόνο αυτούς τους κανόνες συμπερασμού. Για παράδειγμα,αν αφαιρούσαμε τον κανόνα απαλοιφής αμφίδρομης υποθετικής πρότασης, η απόδειξη της προηγούμενης ενότητας δε θα έφτανε σε ικανοποιητικό τέλος. Η τρέχουσα ενότητα παρουσιάζει ένα μόνο κανόνα συμπερασμού, την **ανάλυση** (resolution), που δίνει έναν πλήρη αλγόριθμο συμπερασμού όταν συνδυάζεται με οποιονδήποτε πλήρη αλγόριθμο αναζήτησης.

Θα ξεκινήσουμε χρησιμοποιώντας μια απλή εκδοχή του κανόνα της ανάλυσης στον κόσμο του wumpus. Ας εξετάσουμε τα βήματα που οδηγούν στην Εικόνα 0.4(α) : Ο πράκτορας επιστρέφει από το [2,1] στο [1,1] και μετά πηγαίνει στο [1,2], όπου αντιλαμβάνεται δυσοσμία, αλλά όχι αύρα. Προσθέτουμε τα παρακάτω γεγονότα στη βάση γνώσης:

$$R_{11}: \neg A_{1,2} \quad .$$

$$R_{12}: A_{1,2} \Leftrightarrow (\Gamma_{1,1} \vee \Gamma_{2,2} \vee \Gamma_{1,3}) \quad .$$

Με την ίδια διαδικασία που μας οδήγησε στην R_{10} παραπάνω, μπορούμε τώρα να συμπεράνουμε ότι δεν υπάρχουν γούβες στα [2,2] και [1,3] (θυμηθείτε ότι είναι ήδη γνωστό ότι το [1,1] δεν έχει γούβα):

$$R_{13}: \neg \Gamma_{2,2} \quad .$$

$$R_{14}: \neg \Gamma_{1,3} \quad .$$

Μπορούμε επίσης να εφαρμόσουμε την απαλοιφή αμφίδρομης υποθετικής πρότασης στην R_3 και έπειτα το modus ponens στην R_5 για να συμπεράνουμε το γεγονός ότι υπάρχει γούβα σε κάποιο από τα [1,1], [2,2] ή [3,1]:

$$R_{15}: \Gamma_{1,1} \vee \Gamma_{2,2} \vee \Gamma_{3,1}$$

Τώρα φτάνουμε στην πρώτη εφαρμογή του κανόνα ανάλυσης : Το λεκτικό $\neg \Gamma_{2,2}$ στην R_{13} αναλύεται με το λεκτικό $\Gamma_{2,2}$ της R_{15} και μας δίνει :

$$R_{16}: \Gamma_{1,1} \vee \Gamma_{3,1} \quad .$$

Σε φυσική γλώσσα: Αν υπάρχει γούβα σε ένα από τα [1,1],[2,2] και [3,1] και δεν είναι στο [2,2] τότε είναι στο [1,1] ή στο [3,1]. Παρόμοια, το λεκτικό $\neg\Gamma_{1,1}$ της R_1 αναλύεται με το λεκτικό $\Gamma_{1,1}$ της R_{16} και μας δίνει :

$$R_{17}: \Gamma_{3,1} \quad .$$

Σε φυσική γλώσσα : Αν υπάρχει γούβα στο [1,1] ή στο [3,1] και δεν είναι στο [1,1] τότε είναι στο [3,1]. Αυτά τα δύο τελευταία βήματα του συμπερασμού είναι παραδείγματα του κανόνα συμπερασμού της **μοναδιαίας ανάλυσης** (unit resolution)

$$\frac{l_1 \vee \ldots \vee l_k ,\, m}{l_1 \vee \ldots \vee l_{i-1} \vee l_{i+1} \vee \ldots \vee l_k}$$

όπου κάθε l είναι ένα λεκτικό και τα l_i και m είναι **συμπληρωματικά λεκτικά** (complementary literals – δηλαδή το ένα είναι άρνηση του άλλου). Η μοναδιαία ανάλυση παίρνει λοιπόν μια **διαζευκτική πρόταση** (clause) – μια διάζευξη λεκτικών – και ένα λεκτικό και παράγει μια νέα διαζευκτική πρόταση. Σημειώστε ότι ένα μεμονωμένο λεκτικό μπορεί να θεωρηθεί διάζευξη ενός μόνο λεκτικού, γνωστή και ως **μοναδιαία διαζευκτική πρόταση** (unit clause) .

Ο κανόνας της μοναδιαίας ανάλυσης μπορεί να γενικευτεί στον πλήρη κανόνα της **ανάλυσης** (resolution)

$$\frac{l_1 \vee \ldots \vee l_k ,\, m_1 \vee \ldots \vee m_n}{l_1 \vee \ldots \vee l_{i-1} \vee l_{i+1} \vee \ldots \vee l_k \vee m_1 \vee \ldots \vee m_{j-1} \vee m_{j+1} \vee \ldots \vee m_n}$$

όπου τα l_i και m_j είναι συμπληρωματικά λεκτικά. Αν είχαμε να κάνουμε μόνο με διαζευκτικές προτάσεις με μήκος δύο, θα μπορούσαμε να το γράψουμε έτσι:

$$\frac{l_1 \vee l_2 ,\, \neg l_2 \vee l_3}{l_1 \vee l_3}$$

Δηλαδή , η ανάλυση παίρνει δύο διαζευκτικές προτάσεις και παράγει μια νέα διαζευκτική πρόταση που περιέχει όλα τα λεκτικά των δύο αρχικών διαζευκτικών προτάσεων εκτός από τα δύο συμπληρωματικά λεκτικά. Για παράδειγμα, έχουμε :

$$\frac{\Gamma_{1,1} \vee \Gamma_{3,1}, \qquad \neg\Gamma_{1,1} \vee \neg\Gamma_{2,2}}{\Gamma_{3,1} \vee \neg\Gamma_{2,2}}$$

Υπάρχει ένα ακόμα τεχνικό ζήτημα στον κανόνα της ανάλυσης: Η διαζευκτική πρόταση που προκύπτει θα πρέπει να περιέχει μόνο ένα αντίγραφο του κάθε λεκτικού. Η αφαίρεση των πολλών αντιγράφων των λεκτικών λέγεται **παραγοντοποίηση** (factoring). Για παράδειγμα, αν αναλύσουμε την ($A \vee B$) με την ($A \vee \neg B$) παίρνουμε την ($A \vee A$), η οποία καταλήγει απλώς σε Α.

Η ορθότητα του κανόνα της ανάλυσης φαίνεται εύκολα αν εξετάσουμε το λεκτικό l_i . Αν το l_i είναι αληθές, τότε το m_j είναι ψευδές και επομένως η $m_1 \vee \ldots \vee m_{j-1} \vee m_{j+1} \vee \ldots \vee m_n$ πρέπει να είναι αληθής, επειδή η $m_1 \vee \ldots \vee m_n$ είναι δοσμένη. Αν το l_i είναι ψευδές, τότε η $l_1 \vee \ldots \vee l_{i-1} \vee l_{i+1} \vee \ldots \vee l_k$ πρέπει να είναι αληθής επειδή η $l_1 \vee \ldots \vee l_k$ είναι δοσμένη. Τώρα το l_i είναι ή αληθές ή ψευδές, επομένως ισχύει το ένα από αυτά τα δύο συμπεράσματα – ακριβώς όπως δηλώνει ο κανόνας της ανάλυσης.

Το πιο εντυπωσιακό με τον κανόνα της ανάλυσης είναι ότι παρέχει τη βάση για μια οικογένεια διαδικασιών *πλήρους* συμπερασμού. *Οποιοσδήποτε πλήρης αλγόριθμος αναζήτησης που εφαρμόζει μόνο τον κανόνα της ανάλυσης μπορεί να συνάγει οποιοδήποτε συμπέρασμα που καλύπτεται λογικά από οποιαδήποτε βάση γνώσης της προτασιακής λογικής.* Υπάρχει μια επιφύλαξη: Η ανάλυση είναι πλήρης κατά μια ειδική ένοια. Με δεδομένο ότι η A είναι αληθής, δεν μπορούμε να χρησιμοποιήσουμε την ανάλυση για να παραγάγουμε αυτόματα τη συνέπεια $A \vee B$. Μπορούμε όμως να χρησιμοποιήσουμε την ανάλυση για να απαντήσουμε στην ερώτηση αν η $A \vee B$ είναι αληθής. Αυτό λέγεται **πληρότητα της διάψευσης** (refutation completeness), που σημαίνει ότι η ανάλυση μπορεί πάντα να χρησιμοποιείται είτε για την επιβεβαίωση είτε για την διάψευση μιας πρότασης, αλλά δεν μπορεί να χρησιμοποιείται για την απαρίθμηση αληθών προτάσεων. Οι επόμενες δύο υποενότητες εξηγούν πώς επιτυγχάνεται αυτό με την ανάλυση.

Συζευκτική κανονική μορφή

Ο κανόνας της ανάλυσης έχει εφαρμογή μόνο στις διαζεύξεις λεκτικών και επομένως φαίνεται να αφορά μόνο τις βάσεις γνώσης και τα ερωτήματα που αποτελούνται από τέτοιες διαζεύξεις. Πώς λοιπόν μπορεί να οδηγήσει σε μια πλήρη διαδικασία συμπερασμού για ολόκληρη την προτασιακή λογική; Η απάντηση είναι ότι κάθε πρόταση της προτασιακής λογικής είναι λογικά ισοδύναμη με μια σύζευξη διαζεύξεων λεκτικών. Μια πρόταση εκφρασμένη ως σύζευξη διαζεύξεων λεκτικών λέμε ότι είναι σε **συζευκτική κανονική μορφή** (conjunctive normal from **CNF**). Θα μας φανεί επίσης χρήσιμο αργότερα να εξετάσουμε την περιορισμένη οικογένεια των προτάσεων **k-CNF**. Μια πρόταση σε μορφή k-CNF έχει ακριβώς k λεκτικά ανά διαζευκτική πρόταση:

$$(l_{1,1} \vee \ldots \vee l_{1,k}) \wedge \ldots \wedge (l_{n,1} \vee \ldots \vee l_{n,k})$$

Αποδεικνύεται ότι κάθε πρόταση μπορεί να μετασχηματιστεί σε μια πρόταση 3-CNF που έχει ένα ισοδύναμο σύνολο μοντέλων.

Αντί να αποδείξουμε αυτούς τους ισχυρισμούς, θα περιγράψουμε μια απλή διαδικασία μετατροπής. Θα επιδείξουμε αυτή τη διαδικασία μετατρέποντας την R_2, την πρόταση $A_{1,1} \Leftrightarrow (\Gamma_{1,2} \vee \Gamma_{2,1})$, σε μορφή CNF. Τα βήματα είναι τα εξής :

1. Απαλοιφή του $\Leftrightarrow$, αντικατάσταση της $\alpha \Leftrightarrow \beta$ με την $(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$

$$(A_{1,1} \Rightarrow (\Gamma_{1,2} \vee \Gamma_{2,1})) \wedge ((\Gamma_{1,2} \vee \Gamma_{2,1}) \Rightarrow A_{1,1})$$

2. Απαλοιφή του $\Rightarrow$, αντικατάσταση της $\alpha \Rightarrow \beta$ με την $\neg \alpha \vee \beta$:

$$(\neg A_{1,1} \vee \Gamma_{1,2} \vee \Gamma_{2,1}) \wedge (\neg(\Gamma_{1,2} \vee \Gamma_{2,1}) \vee A_{1,1})$$

3. Η μορφή CNF απαιτεί να εμφανίζεται το $\neg$ μόνο σε λεκτικά, γι'αυτό "μετακινούμε το $\neg$ πρός τα μέσα" με επαναληπτική εφαρμογή των παρακάτω ισοδυναμιών από τον πίνακα της Εικόνας 1.5:

$\neg(\neg\alpha) \equiv \alpha$ (απαλοιφή διπλής άρνησης)
$\neg(\alpha \wedge \beta) \equiv (\neg\alpha \vee \neg\beta)$ (De Morgan)
$\neg(\alpha \vee \beta) \equiv (\neg\alpha \wedge \neg\beta)$ (De Morgan)

Στο παράδειγμα απαιτούμε να εφαρμοστεί ο τελευταίος κανόνας μόνο μία φορά:

$$(\neg A_{1,1} \vee \Gamma_{1,2} \vee \Gamma_{2,1}) \wedge ((\neg \Gamma_{1,2} \wedge \neg \Gamma_{2,1}) \vee A_{1,1}) \quad .$$

4. Τώρα έχουμε μια πρόταση που περιέχει ένθετους τελεστές $\wedge$ και $\vee$ που εφαρμόζονται σε λεκτικά. Εφαρμόζουμε τον επιμεριστικό νόμο του πίνακα της Εικόνας 1.5, επιμερίζοντας τον τελεστή $\vee$ στον $\wedge$ όπου είναι δυνατό.

$$(\neg A_{1,1} \vee \Gamma_{1,2} \vee \Gamma_{2,1}) \wedge ((\neg \Gamma_{1,2} \vee A_{1,1}) \wedge (\neg \Gamma_{2,1}) \vee A_{1,1}) \quad .$$

Η αρχική πρόταση είναι τώρα σε μορφή CNF , μια σύζευξη τριών διαζευκτικών προτάσεων. Είναι πολύ δυσκολότερο να διαβαστεί αλλά μπορεί να χρησιμοποιηθεί ως είσοδος σε μια διαδικασία ανάλυσης.

Ένας αλγόριθμος ανάλυσης

Οι διαδικασίες συμπερασμού που βασίζονται στην ανάλυση λειτουργούν με τη χρήση της αρχής απόδειξης μέσω αντίφασης την οποία συζητήσαμε στο τέλος της ενότητας 1.4. Δηλαδή, για να αποδείξουμε ότι KB $\models$ α , αποδεικνύουμε ότι η $(KB \wedge \neg \alpha)$ είναι μη ικανοποιήσιμη. Αυτό το κάνουμε επιχειρώντας να αποδείξουμε μια αντίφαση.

Ένας αλγόριθμος ανάλυσης παρουσιάζεται στην Εικόνα 1.6. Πρώτα η $(KB \wedge \neg \alpha)$ μετατρέπεται σε μορφή CNF. Έπειτα εφαρμόζεται ο κανόνας της ανάλυσης στις διαζευκτικές προτάσεις που προκύπτουν. Κάθε ζεύγος που περιέχει συμπληρωματικά λεκτικά αναλύεται για να παραχθεί μια νέα διαζευκτική πρόταση, η οποία προστίθεται στο σύνολο αν δεν υπάρχει ήδη. Η διαδικασία συνεχίζεται μέχρι να συμβεί ένα από τα δύο :

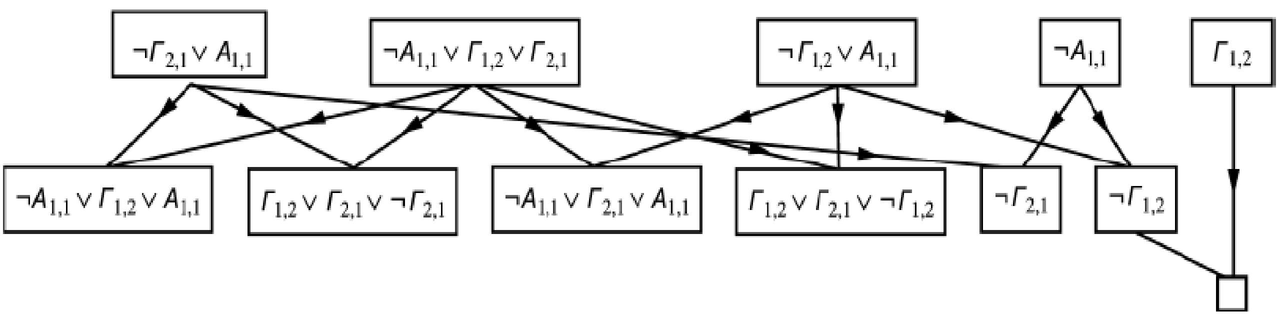
- να μην υπάρχουν νέες διαζευκτικές προτάσεις που μπορούν να προστεθούν, οπότε η KB δεν καλύπτει λογικά την α , ή
- μια εφαρμογή του κανόνα της ανάλυσης να παράγει την κενή πρόταση, οπότε η KB καλύπτει λογικά την α.

Η κενή πρόταση – μια διάζευξη χωρίς διαζευκτέους – ισοδυναμεί με την πρόταση *Ψευδές*, επειδή μια διάζευξη είναι αληθής μόνο αν τουλάχιστον ένας από τους διαζευκτέους είναι αληθής. Ένας άλλος τρόπος για να δούμε ότι μια κενή πρόταση αντιπροσωπεύει μια αντίφαση είναι να παρατηρήσουμε ότι προκύπτει μόνο από την ανάλυση δύο συμπληρωματικών μοναδιαίων διαζευκτικών προτάσεων , όπως οι P *και* $\neg P$.

function PL-RESOLUTION (KB, α) **returns** *Αληθές* ή *Ψευδές*
inputs: KB, η βάση γνώσης, πρόταση της προτασιακής λογικής α,
το ερώτημα , πρόταση της προτασιακής λογικής

διαζ_προτάσεις ← το σύνολο των διαζευκτικών προτάσεων στην αναπαράσταση CNF της
$(KB \wedge \neg\alpha)$
νέες ← { }
loop do
for each C_i , C_j **in** *διαζ_προτάσεις* **do**
αναλυθέντα ← PL-RESOLVE(C_i , C_j)
if *αναλυθέντα* περιέχει την κενή πρόταση **then return** *Αληθές*
νέες ← $νέες \cup αναλυθέντα$
if $νέες \subseteq διαζ - προτάσεις$ **then return** *Ψευδές*
διαζ_προτάσεις ← $διαζ_{προτάσεις} \cup νέες$

Εικόνα 1.6 Απλός αλγόριθμος ανάλυσης για την προτασιακή λογική. Η συνάρτηση PL-RESOLVE επιστρέφει το σύνολο όλων των δυνατών διαζευκτικών προτάσεων που προκύπτουν από την ανάλυση των δύο εισόδων της.



Εικόνα 1.7 Μερική εφαρμογή της PL-RESOLUTION σε έναν απλό συμπερασμό στον κόσμο του wumpus. Όπως φαίνεται, η $\neg\Gamma_{1,2}$ προκύπτει από τις τέσσερις πρώτες διαζευκτικές προτάσεις στην επάνω σειρά.

Μπορούμε να εφαρμόσουμε τη διαδικασία της ανάλυσης σε έναν πολύ απλό συμπερασμό στον κόσμο του wumpus. Όταν ο πράκτορας είναι στο [1,1] δεν υπάρχει αύρα, επομένως δεν μπορούν να υπάρχουν γούβες σε γειτονικά τετράγωνα. Η σχετική βάση γνώσης είναι

$$KB = R_2 \wedge R_4 = (A_{1,1} \Leftrightarrow (\Gamma_{1,2} \vee \Gamma_{2,1})) \wedge \neg A_{1,1}$$

και θέλουμε να αποδείξουμε την *α* η οποία είναι, ας πούμε, η $\neg \Gamma_{1,2}$. Όταν μετατρέψουμε την $(KB \wedge \neg \alpha)$ σε μορφή CNF, παίρνουμε τις διαζευκτικές προτάσεις που εμφανίζονται στο επάνω μέρος της Εικόνας 1.7. Η δεύτερη σειρά της εικόνας δείχνει όλες τις διαζευκτικές προτάσεις που προκύπτουν από την ανάλυση των ζευγών της πρώτης σειράς. Όταν λοιπόν η $\Gamma_{1,2}$ αναλυθεί με την $\neg \Gamma_{1,2}$, παίρνουμε την κενή πρόταση που εμφανίζεται ως μικρό τετράγωνο. Εξετάζοντας την Εικόνα 1.7 διαπιστώνουμε ότι πολλά βήματα ανάλυσης είναι άσκοπα. Για παράδειγμα, η διαζευκτική πρόταση $A_{1,1} \vee \neg A_{1,1} \vee \Gamma_{1,2}$ είναι ισοδύναμη με την $A\lambda\eta\theta\acute{\varepsilon}\varsigma \vee \Gamma_{1,2}$ η οποία είναι ισοδύναμη με την *Αληθές*. Το να συναγάγουμε ότι η *Αληθές* είναι αληθής δε μας βοηθά και πολύ. Γι'αυτό, οποιαδήποτε διαζευκτική πρόταση στην οποία εμφανίζονται δύο συμπληρωματικά λεκτικά μπορεί να αποβάλλεται.

Πληρότητα της ανάλυσης

Για την απόδειξη πληρότητας της ανάλυσης μπορείτε να ανατρέξετε στην βιβλιογραφία μου στο [1].

1.6 ΣΥΜΠΕΡΑΣΜΟΣ ΜΕ ΠΡΟΤΑΣΕΙΣ HORN

Αλυσίδα εκτέλεσης προς τα εμπρός και προς τα πίσω

Η πληρότητα της ανάλυσης την κάνει πολύ σημαντική μέθοδο συμπερασμού. Σε πολλές πρακτικές περιπτώσεις, όμως, η πλήρης ισχύς της ανάλυσης δεν είναι απαραίτητη. Οι πραγματικές βάσεις γνώσης συχνά περιέχουν μόνο διαζευκτικές προτάσεις ενός περιορισμένου τύπου, που ονομάζονται **προτάσεις Horn** (Horn clauses). Πρόταση Horn είναι μια διάζευξη λεκτικών από τα οποία το πολύ ένα είναι θετικό. Για παράδειγμα, η πρόταση $(\neg \Theta_{1,1} \vee \neg Αύρα \vee A_{1,1})$, όπου η $\Theta_{1,1}$ σημαίνει ότι η θέση του πράκτορα είναι [1,1], είναι πρόταση Horn, ενώ η $(\neg A_{1,1} \vee \Gamma_{1,2} \vee \Gamma_{2,1})$ δεν είναι.

Ο περιορισμός σε ένα μόνο θετικό λεκτικό ίσως φαίνεται κάπως αυθαίρετος και αδιάφορος, αλλά στην πραγματικότητα είναι πολύ σημαντικός για τρεις λόγους:

1. Κάθε πρόταση Horn μπορεί να γραφεί ως συνεπαγωγή που η προυπόθεση της είναι μια σύζευξη θετικών λεκτικών και το συμπέρασμα της είναι ένα μεμονωμένο θετικό λεκτικό. Για παράδειγμα, η πρόταση Horn $(\neg \Theta_{1,1} \vee \neg Αύρα \vee A_{1,1})$ μπορεί να γραφεί ως συνεπαγωγή $(\Theta_{1,1} \wedge Αύρα) \Rightarrow A_{1,1}$. Στην τελευταία μορφή, η πρόταση διαβάζεται πολύ ευκολότερα: λέει ότι αν ο πράκτορας είναι στο [1,1] και υπάρχει αύρα, τότε το [1,1] έχει αύρα. Ο άνθρωπος το βρίσκει εύκολο να διαβάζει και να γράφει προτάσεις αυτής της μορφής σε πολλά γνωστικά πεδία.
   Οι προτάσεις Horn σαν αυτή, με ακριβώς ένα θετικό λεκτικό, ονομάζονται **οριστικές προτάσεις** (definite clauses). Το θετικό λεκτικό λέγεται **κεφαλή** (head) και τα αρνητικά λεκτικά αποτελούν το **σώμα** (body) της οριστικής πρότασης. Μια οριστική πρόταση χωρίς αρνητικά λεκτικά απλώς βεβαιώνει μια δεδομένη ατομική πρόταση – μερικές φορές λέγεται **γεγονός** (fact). Οι οριστικές προτάσεις αποτελούν τη βάση του **λογικού προγραμματισμού** (logic programming). Μια πρόταση Horn χωρίς θετικά λεκτικά μπορεί να γραφεί ως συνεπαγωγή που το συμπέρασμα της είναι το λεκτικό *Ψευδές*. Για παράδειγμα , η διαζευκτική πρόταση $(\neg W_{1,1} \vee \neg W_{1,2})$ – το wumpus δεν μπορεί να είναι και στο [1,1] και στο [1,2] – είναι ισοδύναμη με την $W_{1,1} \wedge W_{1,2} \Rightarrow Ψευδές$. Τέτοιες προτάσεις ονομάζονται **περιορισμοί ακεραιότητας** (integrity constraints) στον κόσμο των βάσεων δεδομένων, όπου χρησιμοποιούνται για να σηματοδοτούν σφάλματα στα δεδομένα. Στους αλγόριθμους που ακολουθούν θα θεωρήσουμε για απλούστευση ότι η βάση γνώσης περιέχει μόνο οριστικές προτάσεις και δεν έχει περιορισμούς ακεραιότητας. Αυτές οι βάσεις γνώσης λέμε ότι είναι σε μορφή Horn.

2. Συμπερασμός με προτάσεις Horn μπορεί να γίνεται μέσω αλγορίθμων της **προς τα εμπρός αλυσίδας εκτέλεσης** (forward chaining) και της **προς τα πίσω**

αλυσίδας εκτέλεσης (backward chaining), από τους οποίους θα παραθέσουμε απλώς μόνο τον πρώτο στην συνέχεια. Και οι δύο αλγόριθμοι είναι πολύ φυσικοί κατά το ότι τα βήματα του συμπερασμού είναι προφανή και ο άνθρωπος μπορεί να τα ακολουθεί εύκολα. Για περισσότερες λεπτομέρειες μπορείτε να ανατρέξετε στην βιβλιογραφία.

3. <u>Ο προσδιορισμός της λογικής κάλυψης με προτάσεις Horn μπορεί να γίνεται σε χρόνο **γραμμικό** ως προς το μέγεθος της βάσης γνώσης.</u>

Αυτό το τελευταίο είναι ευχάριστη έκπληξη. Σημαίνει ότι ο λογικός συμπερασμός έχει πολύ χαμηλό κόστος για πολλές προτασιακές βάσεις γνώσης που συναντώνται στην πράξη.

function PL-FC-ENTAILS?(KB, q) **returns** *Αληθές* ή *Ψευδές*
  inputs: KB, η βάση γνώσης, ένα σύνολο προτάσεων Horn ,
    q το ερώτημα , ένα προτασιακό σύμβολο
  local variables: *μετρητές*, πίνακας ευρετηριασμένος κατά διαζευκτική πρόταση,
    αρχικά ο αριθμός των προυποθέσεων
    συμπεράσματα, πίνακας ευρετηριασμένος κατά σύμβολο,
    κάθε καταχώρηση αρχικά Ψευδές
    ατζέντα, λίστα συμβόλων, αρχικά τα γνωστά αληθή σύμβολα στην KB

while το *ατζέντα* δεν είναι κενό **do**
  p ← Pop(*ατζέντα*)
  unless *συμπεράσματα*[p] **do**
    συμπεράσματα[p] ← *Αληθές*
    for each πρόταση Horn c στης οποίας τις προυποθέσεις εμφανίζεται το p **do**
      μείωση του *μετρητές*[c]
      if *μετρητές*[c] = 0 **then do**
        if HEAD[c] = q **then return** *Αληθές*
        PUSH(HEAD[c], *ατζέντα*)
return *Ψευδές*

Εικόνα 1.14 Ο αλγόριθμος της προς τα εμπρός αλυσίδας εκτέλεσης για την προτασιακή λογική. Ο πίνακας *ατζέντα* παρακολουθεί τα σύμβολα που είναι γνωστό ότι είναι αληθή αλλά δεν έχει ακόμα γίνει η “επεξεργασία” τους. Ο πίνακας *μετρητές* παρακολουθεί πόσες προυποθέσεις της κάθε συνεπαγωγής είναι ακόμα άγνωστες. Όποτε ένα νέο σύμβολο p από τον πίνακα *ατζέντα* υπόκειται σε επεξεργασία, ο μετρητής μειώνεται κατά ένα για κάθε συνεπαγωγή στης οποίας τις προυποθέσεις εμφανίζεται το p. (Αυτά μπορούν να προσδιορίζονται σε σταθερό χρόνο αν η ΚΒ είναι κατάλληλα ευρετηριασμένη.) Αν ένας μετρητής φτάσει στο μηδέν, όλες οι προυποθέσεις της συνεπαγωγής είναι γνωστές, και επομένως το συμπερασμά της μπορεί να προστεθεί στον πίνακα *ατζέντα*. Τέλος, χρειάζεται να παρακολουθούμε ποια σύμβολα έχον υποστεί επεξεργασία· ένα σύμβολο του πίνακα *συμπεράσματα* δε χρειάζεται να προστεθεί στον πίνακα *ατζέντα* αν έχει υποστεί ήδη επεξεργασία. Έτσι αποφεύγεται η πλεονάζουσα δουλειά· αυτό αποτρέπει επίσης τους ατέρμονους βρόχους που θα μπορούσαν να προκληθούν από συνεπαγωγές όπως οι P=>Q και Q=>P.

Ο παραπάνω αλγόριθμος παρατέθηκε στην παρούσα εργασία, όχι για να περιγραφεί εκτενώς, αλλά απλώς για να δείξει στον αναγνώστη ότι υπό κατάλληλες προυποθέσεις μπορεί ο συμπερασμός να γίνει με πολύ πιο “μικρές” Πολυπλοκότητες έναντι μιας NP-complete προσέγγισης. Αυτές οι μέθοδοι όμως, είναι αντικείμενο μιας πιο διευρημένης και εξιδεικευμένης μελέτης του συμπερασμού γιαυτό και ξεφεύγουν λίγο από τους σκοπούς αυτής της διπλωματικής εργασίας, που είναι να δείξει μια γενικότερη προσέγγιση του θέματος και να είναι η αφορμή για περισσότερη έρευνα.

1.7 ΠΡΟΒΛΗΜΑΤΙΣΜΟΣ

Αρχίζοντας με έναν πράκτορα που συλλογίζεται λογικά για τη θέση των γούβων, των wumpus και των ασφαλών τετραγώνων, θα πρέπει να ξεκινήσει με μια βάση γνώσης που διατυπώνει τη "φυσική" του κόσμου του wumpus..

Μετά να έχει στη βάση γνώσης το πρόταση που να υποδηλώνει ότι το [1,1] κελί δεν έχει γούβα ή wumpus. Επίσης για κάθε τετράγωνο [x,y] να γνωρίζει μια πρόταση που δηλώνει πώς προκύπτει η αύρα. Για κάθε τετράγωνο [x,y] να γνωρίζει μια πρόταση που δηλώνει πώς προκύπτει η δυσοσμία. Τέλος να γνωρίζει ότι υπάρχει ακριβώς ένα wumpus. Αυτό εκφράζεται με δύο τμήματα. Πρώτα θα πρέπει να πούμε ότι υπάρχει τουλάχιστον ένα wumpus και έπειτα θα πρέπει να πούμε ότι υπάρχει το πολύ ένα wumpus. Αυτό γίνεται με το να πούμε ότι, για δύο οποιαδήποτε τετράγωνα, το ένα από αυτά θα πρέπει να μην έχει wumpus.

Για έναν κόσμο 4 x 4 , λοιπόν, ξεκινάμε με 155 προτάσεις συνολικά, που περιέχουν 64 διακεκριμένα σύμβολα. Για ένα τέτοιο περιβάλλον ο αλγόριθμος της λογικής κάλυψης TT-ENTAILS? (Εικόνα 1.4) είναι προφανώς ακατάλληλος, αφού θα χρειαζόταν να απαριθμήσει 2^{64} γραμμές.

Και ο προβληματισμός συνεχίζεται. Για να χρησιμοποιεί ο πράκτορας μας όντως λογική συλλογιστική, θα χρειαζόμασταν προτάσεις για τις θέσεις και όχι μόνο απλώς για τις θέσεις αλλά και για κάθε διαφορετική χρονική στιγμή της μετακίνησης του μέσα σε αυτό το περιβάλλον. Χρειάζονται προτάσεις προσανατολισμού και ενέργειας που και αυτές θα εξαρτώνται από το χρόνο.

Ακόμα και αν θέσουμε ένα άνω όριο στον αριθμό των χρονικών βημάτων που επιτρέπονται —ίσως 100 — θα καταλήξουμε με δεκάδες χιλιάδες προτάσεις.

Τα γρήγορα προτασιακά προγράμματα επίλυσης μπορούν ακομα να χειρίζονται έναν κόσμο wumpus 4 x 4 με ευκολία (φτάνουν στα όρια τους γύρω στα 100 x 100), ωστόσο υπάρχει η ανάγκη ενός τρόπου αναπαράστασης πιο εκφραστικού από την Προτασιακή λογική έτσι ώστε σύνολα προτάσεων της να μπορούν να αντικατασταθούν απλώς από μια πρόταση. Το επόμενο κεφάλαιο θα περιγράψει αυτή την εκφραστική δυνατότητα.

2 . ΛΟΓΙΚΗ ΠΡΩΤΗΣ ΤΑΞΗΣ

2.1 ΑΝΑΠΑΡΑΣΤΑΣΗ

Οι φυσικές γλώσσες (όπως η Αγγλική ή η Ελληνική) είναι πραγματικά πολύ εκφραστικές. Υπάρχει μια μακρόχρονη παράδοση στη γλωσσολογία και τη φιλοσοφία της γλώσσας που βλέπει τη φυσική γλώσσα ουσιαστικά ως μια δηλωτική γλώσσα αναπαράστασης της γνώσης και προσπαθεί να προσδιορίσει επακριβώς την τυπική της σημασιολογία. Ένα τέτοιο ερευνητικό πρόγραμμα, αν κατέληγε σε επιτυχία , θα είχε μεγάλη αξία για την τεχνητή νοημοσύνη επειδή θα επέτρεπε να χρησιμοποιείται μια φυσική γλώσσα (ή κάποια παράγωγη) στα συστήματα αναπαράστασης και συλλογιστικής.

Η σύγχρονη άποψη για τη φυσική γλώσσα είναι ότι εξυπηρετεί έναν κάπως διαφορετικό σκοπό· συγκεκριμένα, είναι μάλλον μέσο **επικοινωνίας** παρά απλό μέσο αναπαράστασης. Όταν ένας ομιλητής δείχνει και λέει "Κοίτα!" ο ακροατής καταλήγει να γνωρίζει ότι, ας πούμε, ο Σούπερμαν εμφανίστηκε τελικά πάνω από τις στέγες. Όμως, δε θα ήταν σωστό να πούμε ότι η πρόταση "Κοίτα!" κωδικοποιούσε αυτό το γεγονός. Το νόημα της πρότασης εξαρτάται τόσο από την ίδια την πρόταση όσο και από τα συμφραζόμενα όπου ειπώθηκε. Οπωσδήποτε, δε θα μπορούσε κανείς να αποθηκεύσει μια πρόταση όπως η "Κοίτα!" σε μια βάση γνώσης και να αναμένει να ανακτήσει το νόημα της χωρίς να έχει αποθηκεύσει επίσης μια αναπαράσταση των συμφραζομένων — το οποίο μας φέρνει στο ερώτημα του πώς μπορούν να αναπαρασταθούν τα συμφραζόμενα. Οι φυσικές γλώσσες είναι επίσης μη συνθετικές (noncompositional) — το νόημα μιας πρότασης όπως η "Τότε το είδε" μπορεί να εξαρτάται από το γλωσσικό περιβάλλον που ορίζεται από πολλές προηγούμενες και επόμενες προτάσεις. Τέλος, οι φυσικές γλώσσες πάσχουν από **αμφισημία** (ambiguity), η οποία θα μπορούσε να προκαλέσει δυσκολίες στη σκέψη. Όπως το θέτει ο Pinker (1995): "Όταν οι άνθρωποι σκέπτονται τη λέξη "spring" , σίγουρα δεν έχουν αμφιβολία για το αν σκέπτονται την Άνοιξη ή ένα ελατήριο που αναπηδά — και αφού μία λέξη μπορεί να αντιστοιχεί σε δύο σκέψεις, οι σκέψεις δεν μπορεί να είναι λέξεις."

Η προσέγγιση μας θα είναι να υιοθετήσουμε τις βασικές αρχές της προτασιακής λογικής — μια δηλωτική, συνθετική σημασιολογία που είναι ανεξάρτητη από τα συμφραζόμενα και χωρίς αμφισημίες — και να οικοδομήσουμε μια πιο εκφραστική λογική πάνω σ'αυτά τα θεμέλια, δανειζόμενοι ιδέες αναπαράστασης από τη φυσική γλώσσα ενώ αποφεύγουμε τα μειονεκτήματα της. Όταν κοιτάζουμε τη σύνταξη της φυσικής γλώσσας, τα πιο φανερά στοιχεία είναι τα ουσιαστικά και οι ονοματικές φράσεις που αναφέρονται σε **αντικείμενα** (objects — τετράγωνα, γούβες, wumpus) και τα ρήματα και οι ρηματικές φράσεις που αναφέρονται σε **σχέσεις** (relations) μεταξύ αντικειμένων (έχει αύρα, γειτονεύει με το, εξακοντίζει). Μερικές από αυτές τις σχέσεις είναι **συναρτήσεις** (functions) — σχέσεις στις οποίες υπάρχει μόνο μία "τιμή" για μια δεδομένη "είσοδο". Είναι εύκολο να αρχίσουμε να απαριθμούμε παραδείγματα αντικειμένων, σχέσεων και συναρτήσεων:

- Αντικείμενα: άνθρωποι, σπίτια, αριθμοί, θεωρίες, χρώματα, παιχνίδια, πόλεμοι, αιώνες...
- Σχέσεις: μπορούν να υπάρχουν μοναδιαίες σχέσεις ή **ιδιότητες** (properties) όπως κόκκινο, στρογγυλό, πλαστό, πρώτο, πολυώροφο..., ή γενικότερες n-αδικές σχέσεις όπως αδελφός του, μεγαλύτερο από, εσωτερικό, μέρους του, έχει χρώμα, συνέβη μετά, κατέχει, βρίσκεται μεταξύ,...
- Συναρτήσεις: πατέρας του, ο καλύτερος φίλος, τρίτη σειρά του, ένα περισσότερο από το, αφετηρία του...

Πραγματικά, σχεδόν οποιοσδήποτε ισχυρισμός μπορεί να θεωρηθεί ότι αναφέρεται σε αντικείμενα και ιδιότητες ή σχέσεις. Ακολουθούν μερικά παραδείγματα:

- "Ένα συν δύο ίσον τρία"
  Αντικείμενα: ένα, δύο, τρία, ένα συν δύο· Σχέση: ίσον· Συνάρτηση: συν. (Το "Ένα συν δύο" είναι ένα όνομα για το αντικείμενο που προκύπτει από την εφαρμογή της συνάρτησης "συν" στα αντικείμενα "ένα" και "δύο". Το "τρία" είναι ένα άλλο όνομα γι'αυτό το αντικείμενο.)
- "Τα τετράγωνα που γειτονεύουν με το wumpus έχουν δυσοσμία."
  Αντικείμενα: wumpus, τετράγωνα· Ιδιότητα: δυσοσμία· Σχέση: γειτονεύουν.
- "Ο κακός Βασιλιάς Ιωάννης κυβερνούσε την Αγγλία το 1200."
  Αντικείμενα: Ιωάννης, Αγγλία, 1200· Σχέσεις: κυβερνούσε· Ιδιότητες: κακός, βασιλιάς.

Η γλώσσα της **λογικής πρώτης τάξης**, της οποίας τη σύνταξη και τη σημασιολογία θα ορίσουμε στην επόμενη ενότητα, είναι δομημένη γύρω από αντικείμενα και σχέσεις. Έχει αποδειχτεί τόσο σημαντική για τα μαθηματικά, τη φιλοσοφία και την τεχνητή νοημοσύνη ακριβώς επειδή αυτά τα πεδία –– και μεγάλο μέρος της καθημερινής ανθρώπινης ύπαρξης – είναι χρήσιμο να θεωρείται ότι ασχολούνται με αντικείμενα και με τις μεταξύ τους σχέσεις. Η λογική πρώτης τάξης μπορεί επίσης να εκφράζει γεγονότα για μερικά ή για όλα τα αντικείμενα του σύμπαντος. Αυτό μας επιτρέπει να αναπαριστούμε γενικούς νόμους ή κανόνες, όπως η δήλωση "Τα τετράγωνα που γειτονεύουν με το wumpus είναι δύσοσμα".

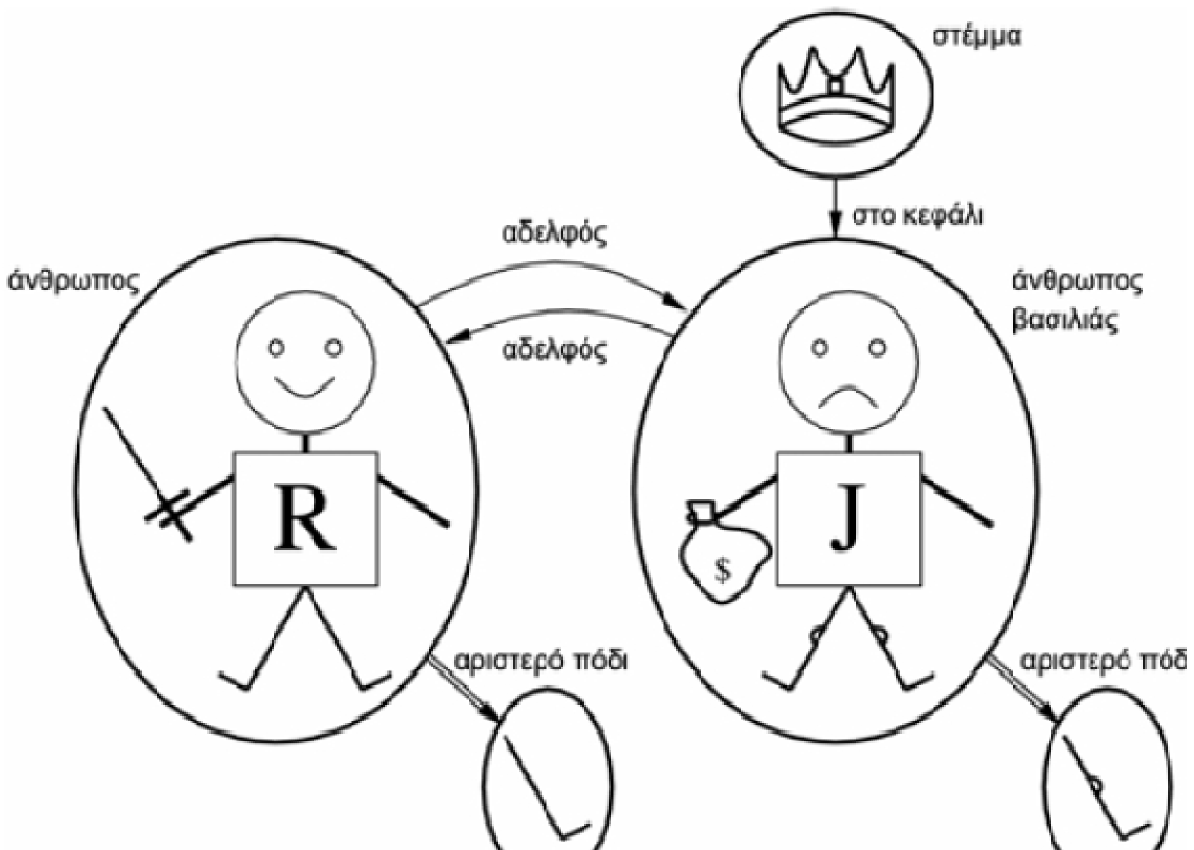
Η κύρια διαφορά μεταξύ προτασιακής λογικής και της λογικής πρώτης τάξης έγκειται στην **οντολογική δέσμευση** (ontological commitment) που γίνεται από την κάθε γλώσσα – δηλαδή στο τι παραδοχές γίνονται για τη φύση της *πραγματικότητας*. Για παράδειγμα, η προτασιακή λογική δέχεται ότι υπάρχουν γεγονότα που ή ισχύουν ή δεν ισχύουν στον κόσμο. Κάθε γεγονός μπορεί να είναι σε μία από τις δύο καταστάσεις: αληθές ή ψευδές. Η λογική πρώτης τάξης κάνει περισσότερες παραδοχές· συγκεκριμένα, δέχεται ότι ο κόσμος αποτελείται από αντικείμενα με κάποιες σχέσεις μεταξύ τους οι οποίες ισχύουν ή δεν ισχύουν.

Υπάρχουν ειδικές λογικές που κάνουν περισσότερες οντολογικές παραδοχές, αλλά δεν αποτελούν αντικείμενο της περιγραφής μας εδώ.

2.2 ΣΥΝΤΑΞΗ ΚΑΙ ΣΗΜΑΣΙΟΛΟΓΙΑ ΤΗΣ ΛΟΓΙΚΗΣ ΠΡΩΤΗΣ ΤΑΞΗΣ

Μοντέλα για τη λογική πρώτης τάξης

Όπως θα θυμάστε από το 0° κεφάλαιο , τα μοντέλα μιας λογικής γλώσσας είναι οι τυπικές δομές που αποτελούν τους δυνατούς κόσμους που εξετάζονται. Τα μοντέλα της προτασιακής λογικής είναι απλώς σύνολα τιμών αληθείας για τα σύμβολα τα προτάσεων. Τα μοντέλα της λογικής πρώτης τάξης είναι πιο ενδιαφέροντα. Πρώτον, περιέχουν αντικείμενα! Το **πεδίο** (domain) ενός μοντέλου είναι το σύνολο των αντικειμένων που περιέχει· τα αντικείμενα αυτά λέγονται μερικές φορές **στοιχεία πεδίου** (domain elements). Η Εικόνα 2.1 δείχνει ένα μοντέλο με πέντε αντικείμενα: ο Ριχάρδος ο Λεοντόκαρδος, βασιλιάς της Αγγλίας από το 1189 μέχρι το 1199· ο νεώτερος αδελφός του, ο κακός Βασιλιάς Ιωάννης, που βασίλευσε από το 1199 μέχρι το 1215· τα αριστερά πόδια του Ριχάρδου και του Ιωάννη· ένα στέμμα.



Εικόνα 2.1 Μοντέλο που περιέχει πέντε αντικείμενα, δύο δυαδικές σχέσεις, τρεις μοναδιαίες σχέσεις (σημειωμένες με ετικέτες στα αντικείμενα) και μία μοναδιαία συνάρτηση, την "αριστερό πόδι".

Τα αντικείμενα του μοντέλου μπορεί να σχετίζονται με διάφορους τρόπους, Στην εικόνα, ο Ριχάρδος και ο Ιωάννης είναι αδελφοί. Από τυπική άποψη, μια σχέση είναι απλώς το σύνολο των πλειάδων (tuples) των σχετιζόμενων αντικειμένων. (Πλειάδα είναι μια συλλογή αντικειμένων τακτοποιημένων σε μια σταθερή διάταξη και γράφεται με γωνιακές αγκύλες γύρω από τα αντικείμενα.). Έτσι, η σχέση της αδελφότητας σε αυτό το μοντέλο είναι το παρακάτω σύνολο:

{< Ριχάρδος Λεοντόκαρδος, Βασιλιάς Ιωάννης > , <Βασιλιάς Ιωάννης, Ριχάρδος Λεοντόκαρδος>}.

(2.1)

(Εδώ ονομάσαμε τα αντικείμενα σε φυσική γλώσσα, αλλά μπορείτε αν θέλετε να αντικαταστήσετε νοερά τα ονόματα με τις εικόνες.) Το στέμμα βρίσκεται στο κεφάλι του Βασιλιά Ιωάννη, γι'αυτό η σχέση "στο κεφάλι" περιέχει μόνο μία πλειάδα, <το στέμμα, Βασιλιάς Ιωάννης>. Οι σχέσεις "αδελφός" και "στο κεφάλι" είναι δυαδικές σχέσεις — δηλαδή συσχετίζουν ζεύγη αντικειμένων. Το μοντέλο περιέχει επίσης μοναδιαίες σχέσεις, ή ιδιότητες: η ιδιότητα "άνθρωπος" είναι αληθής και για το Ριχάρδο και για τον Ιωάννη· η ιδιότητα "βασιλιάς" είναι αληθής μόνο για τον Ιωάννη (ίσως επειδή ο Ριχάρδος είναι νεκρός σ'αυτό το σημείο)· η ιδιότητα "στέμμα" είναι αληθής μόνο για το στέμμα.

Ορισμένα είδη σχέσεων είναι καλύτερο να θεωρούνται συναρτήσεις, με την έννοια ότι ένα δεδομένο αντικείμενο πρέπει να σχετίζεται με ακριβώς ένα αντικείμενο με αυτόν τον τρόπο. Για παράδειγμα, κάθε άνθρωπος έχει ένα αριστερό πόδι, γι'αυτό το μοντέλο έχει μια μοναδιαία συνάρτηση "αριστερό πόδι" που περιλαμβάνει τις παρακάτω απεικονίσεις:

<Ριχάρδος ο Λεοντόκαρδος> → αριστερό πόδι του Ριχάρδου
<Βασιλιάς Ιωάννης> → αριστερό πόδι του Ιωάννη (2.2)

Σύμβολα και ερμηνείες

Ερχόμαστε τώρα στη σύνταξη της γλώσσας. Ο ανυπόμονος αναγνώστης μπορεί να βρει μια πλήρη περιγραφή της στην τυπική γραμματική της λογικής πρώτης τάξης στην Εικόνα 2.2.

Τα βασικά συντακτικά στοιχεία της λογικής πρώτης τάξης είναι τα σύμβολα που αντιπροσωπεύουν αντικείμενα, σχέσεις και συναρτήσεις. Τα σύμβολα αυτά λοιπόν είναι τριών ειδών: **σύμβολα σταθερών** (constant symbols), τα οποία αντιπροσωπεύουν αντικείμενα· **σύμβολα κατηγορημάτων** (predicate symbols), τα οποία αντιπροσωπεύουν σχέσεις· **σύμβολα συναρτήσεων** (function symbols), τα οποία αντιπροσωπεύουν συναρτήσεις. Υιοθετούμε τη σύμβαση ότι αυτά τα σύμβολα θα αρχίζουν με κεφαλαίο γράμμα. Για παράδειγμα, θα μπορούσαμε να χρησιμοποιήσουμε τα σύμβολα σταθερών *Ριχάρδος* και *Ιωάννης*· τα σύμβολα κατηγορημάτων *Αδελφός, ΣτοΚεφάλι, Άνθρωπος, Βασιλιάς*, και *Στέμμα*· το σύμβολο συνάρτησης *ΑριστερόΠόδι*. Όπως συμβαίνει και με τα προτασιακά σύμβολα, η επιλογή των ονομάτων είναι αποκλειστικά ζήτημα του χρήστη. Κάθε κατηγόρημα και σύμβολο συνάρτησης έχει την **τάξη** του (arity), η οποία καθορίζει τον αριθμό των ορισμάτων.

Πρόταση	$\rightarrow$	*ΑτομικήΠρόταση*
	|	(*Πρόταση Συνδετικό Πρόταση*)
	|	*Ποσοδείκτης Μεταβλητή, ... Πρόταση*
	|	$\neg$*Πρόταση*
ΑτομικήΠρόταση	$\rightarrow$	*Κατηγόρημα(Όρος,...)* | *Όρος = Όρος*
Όρος	$\rightarrow$	*Συνάρτηση (Όρος,...)*
	|	*Σταθερά*
	|	*Μεταβλητή*
Συνδετικό	$\rightarrow$	$\Rightarrow$ | $\wedge$ | $\vee$ | $\Leftrightarrow$
Ποσοδείκτης	$\rightarrow$	$\forall$ | $\exists$
Σταθερά	$\rightarrow$	*A* | X_1 | *Ιωάννης* | ...
Μεταβλητή	$\rightarrow$	α | χ | s | …
Κατηγόρημα	$\rightarrow$	*ΠρινΑπό* | *ΈχειΧρώμα* | *Βρέχει* | ...
Συνάρτηση	$\rightarrow$	*Μητέρα* | *ΑριστερόΠόδι* | ...

Εικόνα 2.2 Η σύνταξη της λογικής πρώτης τάξης με ισότητα, όπως ορίζεται σε μορφή Backus-Naur. Η σύνταξη είναι αυστηρή με τις παρενθέσεις· τα σχόλια για τις παρενθέσεις και την προτεραιότητα των τελεστών στην ενότητα 1.4 ισχύουν και για τη λογική πρώτης τάξης.

Η σημασιολογία πρέπει να συσχετίζει τις προτάσεις με μοντέλα για να προσδιορίζει την αλήθεια. Για να γίνει αυτό, χρειαζόμαστε μια **ερμηνεία** (interpetation) που να καθορίζει ακριβώς σε ποια αντικείμενα, σχέσεις και συναρτήσεις αναφέρονται τα σύμβολα των σταθερών, των κατηγορημάτων και των συναρτήσεων. Μία δυνατή ερμηνεία για το παράδειγμα μας – την οποία θα ονομάσουμε **επιδιωκόμενη ερμηνεία** (intended interpretation) – είναι η παρακάτω :

- Το σύμβολο Ριχάρδος αναφέρεται στο Ριχάρδο το Λεοντόκαρδο και το Ιωάννης αναφέρεται στον κακό Βασιλιά Ιωάννη.
- Το σύμβολο Αδελφός αναφέρεται στην αδελφική σχέση, δηλαδή στο σύνολο των πλειάδων αντικειμένων που δίνεται στην Εξίσωση (2.1) · το σύμβολο *ΣτοΚεφάλι* αναφέρεται στη σχέση "στο κεφάλι" που ισχύει μεταξύ του

στέμματος και του Βασιλιά Ιωάννη· τα σύμβολα *Άνθρωπος*, *Βασιλιάς* και *Στέμμα* αναφέρονται στα σύνολα αντικειμένων που είναι άνθρωποι, βασιλιάδες και στέμματα.
- Το σύμβολο *ΑριστερόΠόδι* αναφέρεται στη συνάρτηση "αριστερό πόδι", δηλαδή στην απεικόνιση που δίνεται από την Εξίσωση (2.2).

Υπάρχουν πολλές άλλες δυνατές ερμηνείες που συσχετίζουν αυτά τα σύμβολα με το συγκεκριμένο μοντέλο. Για παράδειγμα, μία ερμηνεία μπορεί να αντιστοιχίζει το σύμβολο *Ριχάρδος* στο στέμμα και το σύμβολο *Ιωάννης* στο αριστερό πόδι του Βασιλιά Ιωάννη. Το μοντέλο περιέχει πέντε αντικείμενα, και επομένως υπάρχουν 25 δυνατές ερμηνείες μόνο για τα σύμβολα σταθερών *Ριχάρδος* και *Ιωάννης*. Προσέξτε ότι δε χρειάζεται όλα τα αντικείμενα να έχουν όνομα – για παράδειγμα, η επιδιωκόμενη ερμηνεία δεν ονομάζει το στέμμα ή τα πόδια. Είναι επίσης δυνατό να έχει ένα αντικείμενο πολλά ονόματα· υπάρχει μια ερμηνεία σύμφωνα με την οποία και το σύμβολο *Ριχάρδος* και το *Ιωάννης* αναφέρονται στο στέμμα. Αν βρίσκετε ότι αυτή η δυνατότητα προκαλεί σύγχυση, θυμηθείτε ότι στην προτασιακή λογική είναι εντελώς δυνατό να έχουμε ένα μοντέλο στο οποίο τα σύμβολα *Συννεφιασμένος* και *Ηλιόλουστος* είναι και τα δύο αληθή· είναι δουλειά της βάσης γνώσης να εξαλείψει τα μοντέλα που είναι ασυνεπή με τη γνώση μας.

Η αλήθεια οποιασδήποτε πρότασης προσδιορίζεται από ένα μοντέλο και μια ερμηνεία για τα σύμβολα της πρότασης. Έτσι, η λογική κάλυψη, η εγκυρότητα και τα λοιπά ορίζονται με βάση <u>*όλα τα δυνατά μοντέλα*</u> και <u>*όλες τις δυνατές ερμηνείες*</u>. Μια σημαντική παρατήρηση είναι ότι ο αριθμός των στοιχείων πεδίου στο κάθε μοντέλο μπορεί να είναι απεριόριστος – για παράδειγμα, τα στοιχεία πεδίου μπορεί να είναι ακέραιοι ή πραγματικοί αριθμοί. Γι'αυτό ο αριθμός των δυνατών μοντέλων είναι απεριόριστος, όπως και ο αριθμός των ερμηνειών. <u>Ο έλεγχος της λογικής κάλυψης με την απαρίθμηση όλων των δυνατών μοντέλων, ο οποίος είναι εφαρμόσιμος στην προτασιακή λογική, δεν είναι εφαρμόσιμος στη λογική πρώτης τάξης</u>. Ακόμα και αν ο αριθμός των αντικειμένων είναι περιορισμένος, ο αριθμός των συνδυασμών μπορεί να είναι πολύ μεγάλος. Με τα σύμβολα του παραδείγματος μας, υπάρχουν περίπου 10^{25} συνδυασμοί για ένα πεδίο με πέντε αντικείμενα.

Όροι

Ένας όρος (term) είναι μια λογική έκφραση που αναφέρεται σε ένα αντικείμενο, Τα σύμβολα σταθερών είναι λοιπόν όροι, αλλά δεν είναι πάντα βολικό να έχουμε ένα ξεχωριστό σύμβολο για να ονομάζουμε κάθε αντικείμενο. Για παράδειγμα, σε φυσική γλώσσα θα μπορούσαμε να χρησιμοποιήσουμε την έκφραση "αριστερό πόδι του Βασιλιά Ιωάννη" αντί να δώσουμε ένα όνομα στο πόδι του. Αυτό το σκοπό εξυπηρετούν τα σύμβολα συναρτήσεων: Αντί να χρησιμοποιούμε ένα σύμβολο σταθεράς, χρησιμοποιούμε το σύμβολο *ΑριστερόΠόδι(Ιωάννης)*. Στη γενική περίπτωση, ένας σύνθετος όρος σχηματίζεται από ένα σύμβολο συνάρτησης ακολουθούμενο από μια λίστα όρων σε παρένθεση ως ορίσματα για το σύμβολο συνάρτησης. Είναι σημαντικό να θυμάστε ότι ένας σύνθετος όρος είναι απλώς ένα περίπλοκο είδος ονόματος. Δεν είναι "κλήση υπορουτίνας" που "επιστρέφει μια τιμή". Δεν υπάρχει υπορουτίνα *ΑριστερόΠόδι* που δέχεται ως είσοδο έναν άνθρωπο και επιστρέφει ένα πόδι. Μπορούμε να συλλογιζόμαστε σχετικά με τα αριστερά πόδια (π.χ. να διατυπώσουμε το γενικό κανόνα ότι όλοι έχουν από ένα και μετά να εξαγάγουμε το συμπέρασμα ότι και ο Ιωάννης θα πρέπει να έχει) χωρίς ούτε καν να δώσουμε ορισμό για το σύμβολο *ΑριστερόΠόδι*. Αυτό είναι κάτι που δεν μπορεί να γίνει με υπορουτίνες σε γλώσσες προγραμματισμού.

Η τυπική σημασιολογία των όρων είναι απλή. Έστω ένας όρος $f(t_1, \dots, t_n)$. Το σύμβολο συνάρτησης f αναφέρεται σε κάποια συνάρτηση του μοντέλου (ας την ονομάσουμε F)· οι όροι των ορισμάτων αναφέρονται σε αντικείμενα του πεδίου (ας τα ονομάσουμε $d_1, \dots, d_n$)· ολόκληρος ο όρος αναφέρεται στο αντικείμενο το οποίο είναι η τιμή της συνάρτησης F όταν εφαρμόζεται στα $d_1, \dots, d_n$. Για παράδειγμα, αν θεωρήσουμε ότι το σύμβολο συνάρτησης *ΑριστερόΠόδι* αναφέρεται στη συνάρτηση της Εξίσωσης (2.2) και το σύμβολο *Ιωάννης* αναφέρεται στο Βασιλιά Ιωάννη, τότε το *ΑριστερόΠόδι(Ιωάννης)* αναφέρεται στο αριστερό πόδι του Βασιλιά Ιωάννη. Με αυτόν τον τρόπο η ερμηνεία καθορίζει το αναφερόμενο του κάθε όρου.

Ατομικές προτάσεις

Τώρα που έχουμε τόσο τους όρους για να αναφερόμαστε σε αντικείμενα όσο και τα σύμβολα κατηγορημάτων για να αναφερόμαστε σε σχέσεις, μπορούμε να τα συνδυάζουμε για να δημιουργούμε **ατομικές προτάσεις** (atomic sentences) που δηλώνουν γεγονότα. Μια ατομική πρόταση σχηματίζεται από ένα σύμβολο κατηγορήματος ακολουθούμενο από μια λίστα όρων σε παρενθέσεις:

Αδελφός(Ριχάρδος, Ιωάννης).

Η πρόταση αυτή λέει ότι, με την επιδιωκόμενη ερμηνεία που δόθηκε παραπάνω, ο Ριχάρδος ο Λεοντόκαρδος είναι αδελφός του Βασιλιά Ιωάννη. Οι ατομικές προτάσεις μπορούν να έχουν ως ορίσματα σύνθετους όρους. Έτσι, η πρόταση

Σύζυγος(Πατέρας(Ριχάρδος), Μητέρα(Ιωάννης))

λέει ότι ο πατέρας του Ριχάρδου του Λεοντόκαρδου είναι σύζυγος της μητέρας του Βασιλιά Ιωάννη (και πάλι υπό μια κατάλληλη ερμηνεία).

*Μια ατομική πρόταση είναι **αληθής** σε ένα δεδομένο μοντέλο, υπό μια δεδομένη ερμηνεία, αν η σχέση στην οποία αναφέρεται το σύμβολο κατηγορήματος ισχύει μεταξύ των αντικειμένων στα οποία αναφέρονται τα ορίσματα.*

Σύνθετες προτάσεις

Μπορούμε να χρησιμοποιούμε **λογικά συνδετικά** για να κατασκευάζουμε πιο σύνθετες προτάσεις, όπως ακριβώς στον προτασιακό λογισμό. Η σημασιολογία των προτάσεων που σχηματίζονται με λογικά συνδετικά είναι ταυτόσημη με εκείνη του προτασιακού λογισμού. Ας δούμε τέσσερις προτάσεις που είναι αληθείς στο μοντέλο της Εικόνας 2.1 με την επιδιωκόμενη ερμηνεία μας:

$$\begin{array}{l} \neg\, Αδελφός(ΑριστερόΠόδι(Ριχάρδος), Ιωάννης) \\ Αδελφός(Ριχάρδος, Ιωάννης) \wedge Αδελφός(Ιωάννης, Ριχάρδος) \\ Βασιλιάς(Ριχάρδος) \vee Βασιλιάς(Ιωάννης) \\ \neg\, Βασιλιάς(Ριχάρδος) \Rightarrow Βασιλιάς(Ιωάννης) \quad . \end{array}$$

Ποσοδείκτες

Από τη στιγμή που έχουμε μια λογική που επιτρέπει αντικείμενα, είναι φυσικό να θέλουμε να εκφράζουμε ιδιότητες ολόκληρων συλλογών αντικειμένων, αντί να απαριθμούμε τα αντικείμενα ονομαστικά. Αυτό μας το επιτρέπουν οι **ποσοδείκτες** (quantifiers). Η λογική πρώτης τάξης διαθέτει δύο καθιερωμένους ποσοδείκτες, που ονομάζονται *καθολικός* (*universal*) και *υπαρξιακός* (*existential*).

Καθολική ποσοτικοποίηση ($\forall$)

Θυμηθείτε τη δυσκολία που είχαμε στο Κεφάλαιο 1 στο να εκφράζουμε γενικούς κανόνες στην προτασιακή λογική. Κανόνες όπως "Τα τετράγωνα που γειτονεύουν με το wumpus είναι δύσοσμα" και "Όλοι οι βασιλιάδες είναι άνθρωποι" είναι πολύ κοινοί στη λογική πρώτης τάξης. Θα ασχοληθούμε με τον δεύτερο κανόνα. Στη λογική πρώτης τάξης γράφεται έτσι:

$$\forall \chi\; Βασιλιάς(\chi) \Rightarrow Άνθρωπος(\chi) \quad .$$

Το $\forall$ διαβάζεται συνήθως "για κάθε...". Έτσι, η πρόταση λέει: "Για κάθε χ, αν το χ είναι βασιλιάς, τότε το χ είναι άνθρωπος". Το σύμβολο χ λέγεται **μεταβλητή**. Κατά σύμβαση, οι μεταβλητές είναι πεζά γράμματα. Μια μεταβλητή είναι από μόνη της όρος, και ως τέτοιος, μπορεί επίσης να εξυπηρετεί ως όρισμα μιας συνάρτησης – για παράδειγμα, *ΑριστερόΠόδι(χ)*. Ένας όρος χωρίς καμία μεταβλητή λέγεται **βασικός όρος** (ground term).

Διαισθητικά, η πρόταση $\forall\ \chi\ P$, όπου P είναι οποιαδήποτε λογική έκφραση, λέει ότι η P είναι αληθής για κάθε αντικείμενο χ. Ακριβέστερα, η $\forall\ \chi\ P$ είναι αληθής σε ένα δεδομένο μοντέλο υπό μια δεδομένη ερμηνεία αν η P είναι αληθής σε όλες τις δυνατές **εκτεταμένες ερμηνείες** (extended interpretations) που κατασκευάζονται από τη δεδομένη ερμηνεία, όπου η κάθε εκτεταμένη ερμηνεία καθορίζει ένα στοιχείο πεδίου στο οποίο αναφέρεται το χ.

Αυτό ακούγεται περίπλοκο, αλλά είναι απλώς ένας προσεκτικός τρόπος διατύπωσης του διαισθητικού νοήματος της καθολικής ποσοτικοποίησης. Ας εξετάσουμε το μοντέλο που παρουσιάζεται στην Εικόνα 2.1 και την επιδιωκόμενη ερμηνεία που το συνοδεύει. Μπορούμε να επεκτείνουμε την ερμηνεία με πέντε τρόπους:

χ → Ριχάρδος ο Λεοντόκαρδος,
χ → Βασιλιάς Ιωάννης,
χ → αριστερό πόδι του Ριχάρδου,
χ → αριστερό πόδι του Ιωάννη,
χ → το στέμμα.

Η καθολικά ποσοτικοποιημένη πρόταση $\forall \chi\ Βασιλιάς(\chi) \Rightarrow Άνθρωπος(\chi)$ είναι αληθής υπό την αρχική ερμηνεία αν η πρόταση $Βασιλιάς(\chi) \Rightarrow Άνθρωπος(\chi)$ είναι αληθής σε κάθε μία από τις πέντε εκτεταμένες ερμηνείες. Δηλαδή, η πρόταση με τον καθολικό ποσοδείκτη είναι ισοδύναμη με τον ισχυρισμό της αλήθειας των παρακάτω πέντε προτάσεων:

ο Ριχάρδος ο Λεοντόκαρδος είναι βασιλιάς ⇒ ο Ριχάρδος ο Λεοντόκαρδος είναι άνθρωπος.
ο Βασιλιάς Ιωάννης είναι βασιλιάς ⇒ ο Βασιλιάς Ιωάννης είναι άνθρωπος.
το αριστερό πόδι του Ριχάρδου είναι βασιλιάς ⇒ το αριστερό πόδι του Ριχάρδου είναι άνθρωπος.
το αριστερό πόδι του Ιωάννη είναι βασιλιάς ⇒ το αριστερό πόδι του Ιωάννη είναι άνθρωπος.
το στέμμα είναι βασιλιάς ⇒ το στέμμα είναι άνθρωπος.

Ας εξετάσουμε προσεκτικά αυτό το σύνολο ισχυρισμών. Αφού στο μοντέλο μας ο Βασιλιάς Ιωάννης είναι ο μόνος βασιλιάς, η δεύτερη πρόταση ισχυρίζεται ότι είναι άνθρωπος, όπως θα περιμέναμε. Τι γίνεται όμως με τις άλλες τέσσερις προτάσεις, οι οποίες φαίνεται να κάνουν ισχυρισμούς για πόδια και στέμματα; Είναι μέρος του νοήματος της πρότασης "Όλοι οι βασιλιάδες είναι άνθρωποι"; Στην πραγματικότητα οι άλλοι τέσσερις ισχυρισμοί είναι αληθείς στο μοντέλο, αλλά δεν κάνουν απολύτως κανέναν ισχυρισμό για την ανθρώπινη εξειδίκευση των ποδιών, των στεμμάτων, ή ακόμα και του Ριχάρδου. Αυτό συμβαίνει επειδή κανένα από αυτά τα αντικείμενα δεν είναι βασιλιάς. Κοιτάζοντας στον πίνακα αληθείας του συνδετικού ⇒ (Εικόνα 1.8), βλέπουμε ότι η συνεπαγωγή είναι αληθής όποτε η προυπόθεση του είναι ψευδής – *άσχετα* από την αλήθεια του συμπεράσματος. Έτσι, ισχυριζόμενοι την αλήθεια της πρότασης με τον καθολικό ποσοδείκτη, το οποίο είναι ισοδύναμο με τον ισχυρισμό της αλήθειας μιας ολόκληρης λίστας μεμονωμένων συνεπαγωγών, καταλήγουμε να ισχυριζόμαστε την αλήθεια του συμπεράσματος του κανόνα μόνο για εκείνα τα αντικείμενα για τα οποία αληθεύει η προυπόθεση και να μη λένε τίποτα απολύτως για εκείνες τις ατομικές προτάσεις για τις οποίες η προυπόθεση είναι ψευδής. Έτσι, οι καταχωρήσεις του πίνακα αληθείας του συνδετικού $\Rightarrow$ αποδεικνύονται άριστες για το γράψιμο γενικών κανόνων με καθολικούς ποσοδείκτες.

Υπαρξιακή ποσοτικοποίηση ($\exists$)

Η καθολική ποσοτικοποίηση κάνει δηλώσεις για κάθε αντικείμενο. Παρόμοια, μπορούμε να κάνουμε μια δήλωση για μερικά αντικείμενα του σύμπαντος χωρίς να τα κατονομάσουμε, χρησιμοποιώντας έναν υπαρξιακό ποσοδείκτη (existential quantifier).Για να πούμε, για παράδειγμα, ότι ο Βασιλιάς Ιωάννης έχει στέμμα στο κεφάλι του, γράφουμε:

$$\exists x\ Στέμμα(\chi) \wedge ΣτοΚεφάλι(\chi\,,\,Ιωάννης)\ .$$

Το $\exists x$ διαβάζεται "Υπάρχει τουλάχιστον ένα χ τέτοιο ώστε..." ή "Για κάποιο χ ...".

Διαισθητικά, η πρόταση $\exists x P$ λέει ότι η P είναι αληθής για τουλάχιστον ένα αντικείμενο χ. Ακριβέστερα, η $\exists x P$ είναι αληθής σε ένα δεδομένο μοντέλο υπό μια δεδομένη ερμηνεία αν η P είναι αληθής *σε τουλάχιστον μία* εκτεταμένη ερμηνεία που απονέμει το χ σε ένα στοιχείο πεδίου. Στο παράδειγμα μας, αυτό σημαίνει ότι τουλάχιστον ένα από τα παρακάτω πρέπει να είναι αληθές:

$$ο\ Ριχάρδος\ ο\ Λεοντόκαρδος\ είναι\ στέμμα \wedge ο\ Ριχάρδος\ ο\ Λεοντόκαρδος\ είναι\ στο\ κεφάλι\ του\ Ιωάννη$$

$$ο\ Βασιλιάς\ Ιωάννης\ είναι\ στέμμα \wedge ο\ Βασιλιάς\ Ιωάννης\ είναι\ στο\ κεφάλι\ του\ Ιωάννη$$

$$το\ αριστερό\ πόδι\ του\ Ριχάρδου\ είναι\ στέμμα \wedge το\ αριστερό\ πόδι\ του\ Ριχάρδου\ είναι\ στο\ κεφάλι\ του\ Ιωάννη$$

$$το\ αριστερό\ πόδι\ του\ Ιωάννη\ είναι\ στέμμα \wedge το\ αριστερό\ πόδι\ του\ Ιωάννη\ είναι\ στο\ κεφάλι\ του\ Ιωάννη$$

$$το\ στέμμα\ είναι\ στέμμα \wedge το\ στέμμα\ είναι\ στο\ κεφάλι\ του\ Ιωάννη$$

Ο πέμπτος ισχυρισμός είναι αληθής στο μοντέλο, επομένως η υπαρξιακά ποσοτικοποιημένη πρόταση είναι αληθής στο μοντέλο. Προσέξτε ότι, κατά τον ορισμό μας, η πρόταση θα ήταν επίσης αληθής σε ένα μοντέλο όπου ο Βασιλιάς Ιωάννης φορούσε δύο στέμματα. Αυτό είναι εντελώς συνεπές με την αρχική πρόταση "Ο Βασιλιάς Ιωάννης έχει στέμμα στο κεφάλι του".

Όπως ακριβώς το $\Rightarrow$ φαίνεται να είναι το φυσικό συνδετικό για να χρησιμοποιείται με το $(\forall)$, το $\wedge$ είναι το φυσικό συνδετικό για να χρησιμοποιείται με το $\exists$. Η χρήση του $\wedge$ ως κύριου συνδετικού με το $(\forall)$ οδήγησε σε μια υπερβολικά ισχυρή δήλωση στο παράδειγμα της προηγούμενης ενότητας· η χρήση του συνδετικού $\Rightarrow$ με το $\exists$ συνήθως οδηγεί σε μια πραγματικά πολύ ασθενή δήλωση. Ας εξετάσουμε την πρόταση:

$$\exists x\ Στέμμα(χ) \Rightarrow ΣτοΚεφάλι(χ\ ,\ Ιωάννης)\ .$$

Επιφανειακά, αυτό μπορεί να φαίνεται ως μια εύλογη διατύπωση της πρότασης μας. Εφαρμόζοντας τη σημασιολογία, βλέπουμε ότι η πρόταση λέει ότι τουλάχιστον ένας από τους παρακάτω ισχυρισμούς είναι αληθής:

ο Ριχάρδος ο Λεοντόκαρδος είναι στέμμα ⇒ ο Ριχάρδος ο Λεοντόκαρδος είναι στο κεφάλι του Ιωάννη

ο Βασιλιάς Ιωάννης είναι στέμμα ⇒ ο Βασιλιάς Ιωάννης είναι στο κεφάλι του Ιωάννη

το αριστερό πόδι του Ριχάρδου είναι στέμμα ⇒ το αριστερό πόδι του Ριχάρδου είναι στο κεφάλι του Ιωάννη

και τα λοιπά. Τώρα μια συνεπαγωγή είναι αληθής αν και η προυπόθεση και το συμπέρασμα είναι αληθή *ή αν η προυπόθεση της είναι ψευδής*. Αν λοιπόν ο Ριχάρδος ο Λεοντόκαρδος δεν είναι στέμμα, τότε ο πρώτος ισχυρισμός είναι αληθής και ο υπαρξιακός προσδιορισμός ικανοποιείται. Έτσι, μια πρόταση συνεπαγωγής με υπαρξιακό ποσοδείκτη είναι αληθής σε οποιοδήποτε μοντέλο περιέχει ένα αντικείμενο για το οποίο η προυπόθεση της συνεπαγωγής είναι ψευδής· επομένως, τέτοιες προτάσεις πραγματικά δε λένε και πολλά πράγματα.

Ένθετοι ποσοδείκτες

Συχνά θα χρειαστεί να εκφράζουμε πιο σύνθετες προτάσεις που χρησιμοποιούν πολλούς ποσοδείκτες. Η απλούστερη περίπτωση είναι όταν οι ποσοδείκτες είναι του ιδίου τύπου. Για παράδειγμα, η πρόταση "Οι αδελφοί είναι αδέλφια" μπορεί να γραφεί έτσι:

$$\forall x \forall y\ Αδελφός(x\ ,\ y) \Rightarrow Αδέλφι(x\ ,\ y)\ .$$

Οι διαδοχικοί ποσοδείκτες του ίδιου τύπου μπορούν να γράφονται ως ένας ποσοδείκτης με πολλές μεταβλητές. Για παράδειγμα, για να πούμε ότι η ιδιότητα της αδελφότητας είναι συμμετρική σχέση, μπορούμε να γράψουμε:

$$\forall x \forall y\ Αδέλφι(x\ ,\ y) \Leftrightarrow Αδέλφι(y\ ,\ x)\ .$$

Σε άλλες περιπτώσεις θα έχουμε μίγματα ποσοδεικτών. "Καθένας αγαπά κάποιον" σημαίνει ότι για κάθε άνθρωπο υπάρχει κάποιος τον οποίο αγαπά:

$$\forall x \exists y\ Αγαπά(x\ ,\ y)$$

Από την άλλη, για να πούμε "Υπάρχει κάποιος που τον αγαπούν όλοι" γράφουμε:

$$\exists y \forall x\ Αγαπά(x\ ,\ y)$$

Η σειρά των ποσοδεικτών είναι λοιπόν πολύ σημαντική. Αυτό γίνεται φανερό αν βάλουμε παρενθέσεις. Η πρόταση $\forall x (\exists y\, Αγαπά(x,y))$ λέει ότι *καθένας* έχει μια συγκεκριμένη ιδιότητα, και συγκεκριμένα την ιδιότητα ότι αγαπά κάποιον. Από την άλλη, η πρόταση $\exists y (\forall x\, Αγαπά(x,y))$ λέει ότι *κάποιος* στον κόσμο έχει μια συγκεκριμένη ιδιότητα, και συγκεκριμένα την ιδιότητα ότι τον αγαπούν όλοι.

Κάποια σύγχυση μπορεί να προκύψει όταν χρησιμοποιηθούν δύο ποσοδείκτες με το ίδιο όνομα μεταβλητής. Ας εξετάσουμε την πρόταση :

$$\forall x [Στέμμα(\chi) \vee (\exists x\, Αδελφός(Ριχάρδος, \chi))] \ .$$

Εδώ το χ στην Αδελφός(Ριχάρδος, χ) έχει *υπαρξιακό* ποσοδείκτη. Ο κανόνας είναι ότι η μεταβλητή ανήκει στον πιο εσωτερικό ποσοδείκτη που αναφέρεται σ'αυτή· έτσι δεν υπόκειται σε κανέναν άλλο ποσοδείκτη. Ένας άλλος τρόπος να το δούμε είναι ο εξής: η $\exists x\, Αδελφός(Ριχάρδος, \chi)$ είναι μια πρόταση για το Ριχάρδο (που λέει ότι έχει αδελφό) και όχι για το χ· έτσι, το να θέσουμε ένα $\forall x$ έξω από αυτή δεν έχει καμία επίδραση. Θα μπορούσε εξίσου καλά να γραφεί $\exists z\, Αδελφός(Ριχάρδος, z)$. Επειδή αυτό μπορεί να γίνει πηγή σύγχυσης, θα χρησιμοποιούμε πάντα διαφορετικές μεταβλητές.

Συνδέσεις μεταξύ των $\forall$ και $\exists$

Οι δύο ποσοδείκτες είναι στην πραγματικότητα στενά συνδεδεμένοι μεταξύ τους, μέσω της άρνησης. Ο ισχυρισμός ότι κανένας δεν αγαπά τις γλυκοπατάτες είναι το ίδιο σαν να ισχυριζόμαστε ότι δεν υπάρχει κάποιος που τις αγαπά, και αντίστροφα:

$\forall x \neg Αγαπά(\chi, Γλυκοπατάτες)$ είναι ισοδύναμη με την $\neg \exists x\, Αγαπά(\chi, Γλυκοπατάτες)$.

Μπορούμε να προχωρήσουμε ένα βήμα ακόμα: "Καθένας αγαπά το παγωτό" σημαίνει ότι δεν υπάρχει κανένας που δεν αγαπά το παγωτό:

$\forall x\, Αγαπά(\chi, Παγωτό)$ είναι ισοδύναμη με την $\neg \exists x \neg Αγαπά(\chi, Παγωτό)$.

Επειδή το $\forall$ είναι στην πραγματικότητα μια σύζευξη που καλύπτει το σύμπαν των αντικειμένων και το $\exists$ είναι μια διάζευξη, δε θα πρέπει να προκαλεί έκπληξη ότι υπακούουν στους νόμους του De Morgan. Οι νόμοι του De Morgan για τις ποσοτικοποιημένες και τις μη ποσοτικοποιημένες προτάσεις είναι οι παρακάτω:

$\forall x \neg P \equiv \neg \exists x P$ $\qquad\qquad$ $\neg P \wedge \neg Q \equiv \neg(P \vee Q)$

$\neg\forall x P \equiv \exists x \neg P$ $\qquad\qquad$ $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$

$\forall x P \equiv \neg\exists x \neg P$ $\qquad\qquad$ $P \wedge Q \equiv \neg(\neg P \vee \neg Q)$

$\exists x P \equiv \neg\forall x \neg P$ $\qquad\qquad$ $P \vee Q \equiv \neg(\neg P \wedge \neg Q)$.

Δε χρειαζόμαστε λοιπόν πραγματικά και το $\forall$ και το $\exists$, όπως δε χρειαζόμαστε πραγματικά και το $\wedge$ και το $\vee$. Όμως η αναγνωσιμότητα είναι σημαντικότερη από την φειδωλία, γι'αυτό θα κρατήσουμε και τους δύο ποσοδείκτες.

Ισότητα

Η λογική πρώτης τάξης περιλαμβάνει έναν ακόμα τρόπο δημιουργίας ατομικών προτάσεων, εκτός από τη χρήση ενός κατηγορήματος και όρων όπως περιγράψαμε παραπάνω. Μπορούμε να χρησιμοποιούμε το σύμβολο της ισότητας για να κάνουμε δηλώσεις που λένε ότι δύο όροι αναφέρονται στο ίδιο αντικείμενο. Για παράδειγμα, η πρόταση

Πατέρας (Ιωάννης) = Ερίκος

λέει ότι το αντικείμενο στο οποίο αναφέρεται ο όρος *Πατέρας(Ιωάννης)* και το αντικείμενο στο οποίο αναφέρεται ο όρος *Ερίκος* ταυτίζονται. Επειδή μια ερμηνεία καθορίζει το αναφερόμενο οποιουδήποτε όρου, ο προσδιορισμός της αλήθειας μιας πρότασης ισότητας είναι απλώς ζήτημα του να διαπιστωθεί ότι τα αναφερόμενα δύο όρων είναι το ίδιο αντικείμενο.

Το σύμβολο της ισότητας μπορεί να χρησιμοποιείται για να δηλώνονται γεγονότα σχετικά με μια δεδομένη συνάρτηση, όπως κάναμε μόλις τώρα με το σύμβολο *Πατέρας*. Μπορεί επίσης να χρησιμοποιείται με άρνηση για να δηλώνεται ότι δύο όροι δεν είναι το ίδιο αντικείμενο. Για να πούμε ότι ο Ριχάρδος έχει τουλάχιστον δύο αδελφούς, θα γράφαμε:

$$\exists x, y \; Αδελφός(\chi, Ριχάρδος) \wedge Αδελφός(y, Ριχάρδος) \wedge \neg(x = y) \quad .$$

Η πρόταση

$$\exists x, y \; Αδελφός(\chi, Ριχάρδος) \wedge Αδελφός(y, Ριχάρδος)$$

δεν έχει την επιδιωκόμενη σημασία. Ειδικότερα, είναι αληθής στο μοντέλο της Εικόνας 2.1, όπου ο Ριχάρδος έχει μόνο έναν αδελφό. Για να το δούμε, ας εξετάσουμε την εκτεταμένη ερμηνεία στην οποία και το x και το y έχουν αποδοθεί στο Βασιλιά Ιωάννη. Η προσθήκη της $\neg(x = y)$ αποκλείει αυτά τα μοντέλα. Μερικές φορές χρησιμοποιείται ο συμβολισμός x ≠ y ως συντομογραφία της $\neg(x = y)$.

2.3 ΧΡΗΣΗ ΤΗΣ ΛΟΓΙΚΗΣ ΠΡΩΤΗΣ ΤΑΞΗΣ ΣΤΟΝ ΚΟΣΜΟ ΤΟΥ WUMPUS

Είδαμε μερικά αξιώματα της προτασιακής λογικής για τον κόσμο του wumpus στο Κεφάλαιο 1.
Τα αξιώματα της λογικής πρώτης τάξης που θα δούμε σε αυτή την ενότητα είναι πολύ πιο εμπεριστατωμένα, συλλαμβάνοντας με πολύ φυσικό τρόπο ό,τι ακριβώς θέλουμε να πούμε.

Θυμηθείτε ότι ο πράκτορας του κόσμου του wumpus λαμβάνει ένα διάνυσμα αντιλήψεων με πέντε στοιχεία. Η αντίστοιχη πρόταση της λογικής πρώτης τάξης που αποθηκεύεται στη βάση γνώσης πρέπει να περιλαμβάνει και την αντίληψη και το χρόνο κατά τον οποίο εμφανίστηκε, αλλιώς ο πράκτορας θα βρίσκεται σε σύγχυση για το πότε είδε τι. Θα χρησιμοποιήσουμε ακεραίους για τα χρονικά βήματα. Μια τυπική πρόταση αντίληψης θα μπορούσε να είναι η παρακάτω:

Αντίληψη ([*Δυσοσμία, Αύρα, Λάμψη, Τίποτα, Τίποτα*], 5) .

Εδώ, η *Αντίληψη* είναι ένα δυαδικό κατηγόρημα και η *Δυσοσμία* και τα λοιπά είναι σταθερές τοποθετημένες σε μια λίστα. Οι ενέργειες στον κόσμο του wumpus μπορούν να αναπαρασταθούν με λογικούς όρους:

Στροφή(Δεξιά) , Στροφή(Αριστερά), Εμπρός, Εξακόντιση , Αρπαγή, Άφεση, Αναρρίχηση.

Για να αποφασίσει ποια είναι η καλύτερη ενέργεια, το πρόγραμμα πράκτορα κατασκευάζει ένα ερώτημα όπως το παρακάτω:

$$\exists \alpha\, \text{ΚαλύτερηΕνέργεια}(\alpha, 5) \quad .$$

Η ASK θα επέλυε αυτό το ερώτημα και θα επέστρεφε μια λίστα δεσμεύσεων όπως η {α / *Αρπαγή*}. Το πρόγραμμα πράκτορα μπορεί τότε να επιστρέψει την *Αρπαγή* ως ενέργεια που θα πραγματοποιηθεί, αλλά πρώτα πρέπει να ενημερώσει (TELL) τη βάση γνώσης του ότι εκτελεί *Αρπαγή*.

Τα ανεπεξέργαστα δεδομένα αντίληψης συνεπάγονται ορισμένα γεγονότα για την τρέχουσα κατάσταση. Για παράδειγμα:

$$\forall t, s, g, m, c\; \text{Αντίληψη}([s, \text{Αύρα}, g, m, c], t) \Rightarrow \text{Αύρα}(t),$$

$$\forall t, s, b, m, c\; \text{Αντίληψη}([s, b, \text{Λάμψη}, m, c], t) \Rightarrow \text{Λάμψη}(t),$$

και τα λοιπά. Οι κανόνες αυτοί εκφράζουν μια τετριμμένη μορφή της συλλογιστικής διαδικασίας που ονομάζεται αντίληψη (perception). Παρατηρήστε την ποσοτικοποίηση ως προς το χρόνο t. <u>Στην προτασιακή λογική θα χρειαζόμασταν αντίγραφα της κάθε πρότασης για κάθε χρονικό βήμα.</u>

Η απλή "αντανακλαστική" συμπεριφορά μπορεί επίσης να υλοποιηθεί με ποσοτικοποιημένες προτάσεις συνεπαγωγής. Για παράδειγμα, έχουμε:

$$\forall t\ Λάμψη(t) \Rightarrow ΚαλύτερηΕνέργεια(Αρπαγή, t).$$

Με δεδομένα την αντίληψη και τους κανόνες των προηγούμενων παραγράφων, αυτό θα έδινε το επιθυμητό συμπέρασμα *ΚαλύτερηΕνέργεια(Αρπαγή, 5)* – δηλαδή, ότι η *Αρπαγή* είναι η σωστή ενέργεια.

Μέχρι αυτό το σημείο σε αυτή την ενότητα οι προτάσεις που αναφέρονταν στο χρόνο ήταν **συγχρονικές** (synchronic) , δηλαδή συσχέτιζαν ιδιότητες μιας κατάστασης του κόσμου με άλλες ιδιότητες της ίδιας κατάστασης του κόσμου. Οι προτάσεις που επιτρέπουν συλλογιστική "δια μέσου του χρόνου" ονομάζονται **διαχρονικές** (diachronic)· για παράδειγμα, ο πράκτορας χρειάζεται να γνωρίζει πώς να συνδυάζει πληροφορίες για την προηγούμενη θέση του με πληροφορίες για την ενέργεια που μόλις έκανε για να προσδιορίζει την τρέχουσα θέση του.

Έχουμε αναπαραστήσει τις αντιλήψεις και τις ενέργειες· είναι ώρα να αναπαραστήσουμε και το ίδιο το περιβάλλον. Ας ξεκινήσουμε με τα αντικείμενα. Προφανείς υποψήφιοι είναι τα τετράγωνα, οι γούβες και το wumpus. Θα μπορούσαμε να ονομάσουμε κάθε τετράγωνο – $Τετράγωνο_{1,2}$ κ.ο.κ. – αλλά τότε το γεγονός ότι το $Τετράγωνο_{1,2}$ και το $Τετράγωνο_{1,3}$ είναι γειτονικά θα έπρεπε να είναι ένα "επιπλέον" γεγονός και θα χρειαζόμασταν ένα τέτοιο γεγονός για κάθε ζεύγος τετραγώνων. Είναι καλύτερα να χρησιμοποιήσουμε ένα σύνθετο όρο στον οποίο η γραμμή και η στήλη να εμφανίζονται ως ακέραιοι αριθμοί· για παράδειγμα, μπορούμε απλώς να χρησιμοποιήσουμε τον όρο λίστας [1,2]. Η γειτνίαση δύο οποιωνδήποτε τετραγώνων μπορεί να οριστεί έτσι:

$$\forall x, y, a, b\ Γειτονικός([x, y], [a, b]) \Leftrightarrow [a, b] \in \{[x+1, y], [x-1, y], [x, y+1], [x, y-1]\} \quad .$$

Θα μπορούσαμε επίσης να ονομάσουμε κάθε γούβα, αλλά αυτό δε θα ήταν ενδεδειγμένο για διαφορετικό λόγο: δεν υπάρχει λόγος να κάνουμε διάκριση μεταξύ γούβων. Είναι πολύ απλούστερο να χρησιμοποιείται ένα μοναδιαίο κατηγόρημα *Γούβα* το οποίο είναι αληθές για τα τετράγωνα πο
έχουν γούβες. Τέλος, αφού υπάρχει ένα και μόνο wumpus, μια σταθερά *Wumpus* είναι εξίσου καλή με ένα μοναδιαίο κατηγόρημα (και ίσως πιο αξιοπρεπής από την άποψη του wumpus). Το wumpus ζει σε ένα ακριβώς τετράγωνο, γι'αυτό είναι καλή ιδέα να χρησιμοποιείται μια συνάρτηση όπως η *Κατοικία(Wumpus)* για να κατονομάζεται αυτό το τετράγωνο. Έτσι αποφεύγεται εντελώς το δύσχρηστο σύνολο προτάσεων που απαιτείται στην προτασιακή λογική για να λέμε ότι ακριβώς ένα τετράγωνο περιέχει wumpus.

Η θέση του πράκτορα αλλάζει με το χρόνο γι'αυτό γράφουμε *Σε(Πράκτορας, s, t)* που σημαίνει ότι ο πράκτορας είναι στο τετράγωνο *s* τη χρονική στιγμή *t*. Με δεδομένη την τρέχουσα θέση του, ο πράκτορας μπορεί να συμπεραίνει ιδιότητες του τετραγώνου από ιδιότητες της τρέχουσας αντίληψης του. Για παράδειγμα, αν ο πράκτορας βρίσκεται σε ένα τετράγωνο και αντιλαμβάνεται αύρα, τότε αυτό το τετράγωνο έχει αύρα:

$$\forall s,t\ Σε(Πράκτορας, s, t) \wedge Αύρα(t) \Rightarrow ΈχειΑύρα(s).$$

Είναι χρήσιμο να γνωρίζουμε ότι ένα *τετράγωνο* έχει αύρα επειδή γνωρίζουμε ότι οι γούβες δεν μπορούν να μετακινούνται. Προσέξτε ότι το *ΈχειΑύρα* δεν έχει χρονικό όρισμα.

Έχοντας ανακαλύψει ποιες θέσεις έχουν αύρα (ή δυσοσμία) και, ακόμα σημαντικότερο, ποιες δεν έχουν αύρα (ή δεν έχουν δυσοσμία), ο πράκτορας μπορεί να συμπεράνει πού βρίσκονται οι γούβες (και πού βρίσκεται το wumpus). Υπάρχουν δύο είδη συγχρονικών κανόνων που μπορούν να επιτρέπουν τέτοιες εξαγωγές συμπερασμάτων:

- **Διαγνωστικοί κανόνες** (diagnostic rules)

  Οι διαγνωστικοί κανόνες οδηγούν από παρατηρούμενα αποτελέσματα σε κρυφά αίτια. Για να βρίσκονται οι γούβες, οι προφανείς διαγνωστικοί κανόνες λένε ότι αν ένα τετράγωνο έχει αύρα, κάποιο γειτονικό τετράγωνο πρέπει να έχει γούβα, ή

  $$\forall s\ ΈχειΑύρα(s) \Rightarrow \exists r\ Γειτονικό(r,s) \wedge Γούβα(r) \quad ,$$

  και ότι αν ένα τετράγωνο δεν έχει αύρα κανένα γειτονικό τετράγωνο δεν έχει γούβα:

  $$\forall s \neg ΈχειΑύρα(s) \Rightarrow \neg \exists r\ Γειτονικό(r,s) \wedge Γούβα(r) \quad .$$

  Συνδυάζοντας αυτούς τους δύο κανόνες, παίρνουμε την αμφίδρομη υποθετική πρόταση

  $$\forall s\ ΈχειΑύρα(s) \Leftrightarrow \exists r\ Γειτονικό(r,s) \wedge Γούβα(r) \qquad (2.3)$$

- **Αιτιολογικοί κανόνες** (casual rules)

  Οι αιτιολογικοί κανόνες αντανακλούν την παραδεκτή κατεύθυνση της αιτιότητας στον κόσμο: κάποια κρυφή ιδιότητα του κόσμου κάνει να παραχθούν ορισμένες αντιλήψεις, Για παράδειγμα, μια γούβα κάνει όλα τα γειτονικά τετράγωνα να έχουν αύρα:

  $$\forall r\ Γούβα(r) \Rightarrow [\forall s\ Γειτονικό(r,s) \Rightarrow ΈχειΑύρα(s)]$$

και αν όλα τα τετράγωνα που γειτονεύουν με ένα δεδομένο τετράγωνο δεν έχουν γούβες, το τετράγωνο δεν θα έχει αύρα:

$$\forall s [\forall r\ Γειτονικό(r, s) \Rightarrow \neg Γούβα(r)] \Rightarrow \neg ΈχειΑύρα(s) \quad .$$

Με λίγη δουλειά, μπορούμε να δείξουμε ότι οι δύο αυτές προτάσεις μαζί είναι λογικά ισοδύναμες με την αμφίδρομη υποθετική πρόταση της Εξίσωσης (2.3). Η αμφίδρομη υποθετική πρόταση μπορεί και η ίδια να θεωρηθεί αιτιολογική, επειδή δηλώνει το πώς παράγεται η τιμή αληθείας της *ΈχειΑύρα* από την κατάσταση του κόσμου.

Τα συστήματα που συλλογίζονται με αιτιολογικούς κανόνες ονομάζονται συστήματα **συλλογιστικής βασισμένης σε μοντέλο** (model-based reasoning), επειδή οι αιτιολογικοί κανόνες αποτελούν ένα μοντέλο του πώς λειτουργεί το περιβάλλον. Η διάκριση μεταξύ της βασισμένης σε μοντέλο και διαγνωστικής συλλογιστικής είναι σημαντική σε πολλούς τομείς της ΤΝ. Η ιατρική διάγνωση, ειδικότερα, είναι ένας ενεργός ερευνητικός τομέας όπου οι προσεγγίσεις που βασίζονται σε άμεσες συσχετίσεις μεταξύ συμπτωμάτων και ασθενειών (μια διαγνωστική προσέγγιση) έχουν βαθμιαία αντικατασταθεί από προσεγγίσεις που χρησιμοποιούν ένα ρητό μοντέλο της διαδικασίας ασθένειας και του πώς εκδηλώνεται σε συμπτώματα.

Όποιο είδος αναπαράστασης και να χρησιμοποιεί ο πράκτορας, *αν τα αξιώματα περιγράφουν σωστά και ολοκληρωμένα τον τρόπο που λειτουργεί ο κόσμος και τον τρόπο που παράγονται οι αντιλήψεις, τότε οποιαδήποτε πλήρης διαδικασία λογικού συμπερασμού θα παράγει την ισχυρότερη δυνατή περιγραφή της κατάστασης του κόσμου, με δεδομένες τις διαθέσιμες αντιλήψεις.* Έτσι, ο σχεδιαστής πρακτόρων μπορεί να συγκεντρώνει την προσοχή του στη σωστή γνώση, χωρίς να ασχολείται πολύ με τις διαδικασίες του παραγωγικού συμπερασμού. Επίσης, έχουμε δει ότι η λογική πρώτης τάξης μπορεί να αναπαριστά τον κόσμου του wumpus χωρίς να υστερεί σε συνέπεια από την αρχική περιγραφή σε φυσική γλώσσα που δώσαμε στο Κεφάλαιο 1.

2.4 ΣΥΜΠΕΡΑΣΜΟΣ ΣΤΗ ΛΟΓΙΚΗ ΠΡΩΤΗΣ ΤΑΞΗΣ

Στο Κεφάλαιο 1 ορίσαμε την έννοια του συμπερασμού (inference) και δείξαμε πώς μπορεί να επιτευχθεί ορθός και πλήρης συμπερασμός για την προτασιακή λογική. Στη παρούσα ενότητα θα επεκτείνουμε αυτά τα αποτελέσματα για να κατασκευάσουμε αλγόριθμους οι οποίοι μπορούν να απαντήσουν σε οποιαδήποτε ερώτηση διατυπωθεί σε λογική πρώτης τάξης και είναι δυνατό να απαντηθεί. Αυτό είναι σημαντικό, επειδή λίγο-πολύ τα πάντα μπορούν να διατυπωθούν σε λογική πρώτης τάξης αν προσπαθήσουμε αρκετά.

Η Ενότητα 2.4.1 παρουσιάζει τους κανόνες συμπερασμού για τους ποσοδείκτες και δείχνει πώς ο συμπερασμός πρώτης τάξης μπορεί να αναχθεί σε προτασιακό συμπερασμό, αν και με μεγάλο κόστος. Η Ενότητα 2.4.2 περιγράφει την ιδέα της ενοποίησης (unification), δείχνοντας πώς μπορεί να χρησιμοποιηθεί για την κατασκευή κανόνων συμπερασμού οι οποίοι λειτουργούν απευθείας με προτάσεις πρώτης τάξης. Κατόπιν στην 2.4.3 περιγράφουμε μια βασική οικογένεια αλγορίθμων συμπερασμού πρώτης τάξης: τα συστήματα απόδειξης θεωρημάτων με βάση την μέθοδο της ανάλυσης. Η συλλογιστική με πλήρως γενικές προτάσεις πρώτης τάξης με τη χρήση της ανάλυσης συνήθως είναι λιγότερο αποδοτική από τη συλλογιστική με άλλες μεθόδους που εφαρμόζονται σε πιο ειδικές προτάσεις (όπως οι Horn), όμως είναι μια πλήρη συλλογιστική.

2.4.1 ΠΡΟΤΑΣΙΑΚΟΣ ΣΥΜΠΕΡΑΣΜΟΣ ΚΑΙ ΣΥΜΠΕΡΑΣΜΟΣ ΠΡΩΤΗΣ ΤΑΞΗΣ

Αυτή και η επόμενη ενότητα παρουσιάζουν τις ιδέες που βρίσκονται πίσω από τα σύγχρονα συστήματα λογικού συμπερασμού. Θα ξεκινήσουμε με μερικούς απλούς κανόνες συμπερασμού, οι οποίοι μπορούν να εφαρμοστούν σε προτάσεις με ποσοδείκτες για τη λήψη προτάσεων χωρίς ποσοδείκτες. Αυτοί οι κανόνες οδηγούν φυσιολογικά στην ιδέα ότι ο συμπερασμός πρώτης τάξης μπορεί να επιτευχθεί με τη μετατροπή της βάσης γνώσης σε προτασιακή λογική και τη χρήση προτασιακού συμπερασμού, κάτι που ήδη γνωρίζουμε πως γίνεται. Η επόμενη ενότητα επισημαίνει μια προφανή συντόμευση, οδηγώντας σε μεθόδους συμπερασμού που χειρίζονται απευθείας προτάσεις πρώτης τάξης.

Κανόνες συμπερασμού για ποσοδείκτες

Ας ξεκινήσουμε με τους καθολικούς ποσοδείκτες. Υποθέτουμε ότι η βάση γνώσης μας περιέχει το κοινό λαικό αξίωμα που δηλώνει ότι όλοι οι άπληστοι βασιλείς είναι κακοί:

$$\forall\, x\; Βασιλιάς(\chi) \wedge Άπληστος(\chi) \Rightarrow Κακός(\chi) \quad .$$

Τότε φαίνεται απόλυτα αποδεκτό να συμπεράνουμε οποιαδήποτε από τις ακόλουθες προτάσεις

$$Βασιλιάς(Ιωάννης) \wedge Άπληστος(Ιωάννης) \Rightarrow Κακός(Ιωάννης) \quad .$$

$$Βασιλιάς(Ριχάρδος) \wedge Άπληστος(Ριχάρδος) \Rightarrow Κακός(Ριχάρδος) \quad .$$

$$Βασιλιάς(Πατέρας(Ιωάννης)) \wedge Άπληστος(Πατέρας(Ιωάννης)) \Rightarrow Κακός(Πατέρας(Ιωάννης)) \quad .$$

.
.
.

Ο κανόνας του **Καθολικού Προσδιορισμού** (Universal Instantiation, UI για συντομία) λέει ότι μπορούμε να συμπεράνουμε οποιαδήποτε πρόταση προκύπτει από την αντικατάσταση μιας μεταβλητής με ένα **βασικό όρο** (έναν όρο χωρίς μεταβλητές, ground term). Μη συγχέετε αυτές τις αντικαταστάσεις με τις εκτεταμένες ερμηνείες που χρησιμοποιούνται για τον καθορισμό της σημασιολογίας των ποσοδεικτών. Η αντικατάσταση (substitution) αντικαθιστά μια μεταβλητή με έναν όρο (ένα τμήμα συντακτικού) για την παραγωγή μια νέας πρότασης, ενώ μια ερμηνεία (interpretation) αντιστοιχίζει μια μεταβλητή σε ένα αντικείμενο στο πεδίο.

Για να γράψουμε τον κανόνα συμπερασμού με τυπική μορφή, θα χρησιμοποιήσουμε την έννοια των **αντικαταστάσεων** . Θεωρούμε ότι SUBST(θ, α) δηλώνει το αποτέλεσμα της αντικατάστασης θ στην πρόταση α. Τότε ο κανόνας γράφεται

$$\frac{\forall\, \nu\, \alpha}{\text{SUBST}(\{\nu/g\}, \alpha)}$$

για οποιαδήποτε μεταβλητή ν και βασικό όρο g. Για παράδειγμα, οι τρεις προτάσεις που δόθηκαν νωρίτερα προκύπτουν από τις αντικαταστάσεις { χ / Ιωάννης } , { χ /

Ριχάρδος }, και { χ / Πατέρας(Ιωάννης) }.

Ο αντίστοιχος κανόνας του **Υπαρξιακού Προσδιορισμού** (Existential Instantiation) για τον υπαρξιακό ποσοδείκτη είναι λίγο πιο πολύπλοκος. Για οποιαδήποτε πρόταση α, μεταβλητή ν, και σταθερό σύμβολο k που δεν εμφανίζεται πουθενά αλλού στη βάση γνώσης,

$$\frac{\exists \nu \; \alpha}{\text{SUBST}(\{\nu/k\}, \alpha)}$$

Για παράδειγμα, από την πρόταση

$$\exists \chi \; Στέμμα(\chi) \wedge ΣτοΚεφάλι(\chi, Ιωάννης)$$

μπορούμε να συμπεράνουμε την πρόταση

$$Στέμμα(C_1) \wedge ΣτοΚεφάλι(C_1, Ιωάννης)$$

όσο το C_1 δεν εμφανίζεται σε κανένα άλλο σημείο της βάσης γνώσης. Βασικά, η υπαρξιακή πρόταση λέει ότι υπάρχει ένα αντικείμενο που ικανοποιεί μια συνθήκη, ενώ η διαδικασία του προσδιορισμού είναι απλώς η απόδοση ενός ονόματος σε αυτό το αντικείμενο. Φυσιολογικά, το όνομα αυτό δεν πρέπει να ανήκει ήδη σε κάποιο άλλο αντικείμενο. Τα μαθηματικά μάς δίνουν ένα καλό παράδειγμα: υποθέστε ότι ανακαλύπτουμε πως υπάρχει ένας αριθμός ο οποίος είναι λίγο μεγαλύτερος από 2,71828 και ικανοποιεί την εξίσωση $d(x^y)/dy = x^y$ για χ. Μπορούμε να δώσουμε σε αυτόν τον αριθμό ένα όνομα, ας πούμε *e*, αλλά θα ήταν λάθος να του δώσουμε το όνομα ενός υπάρχοντος αντικειμένου, όπως το π. Στην λογική, το νέο όνομα καλείται **σταθερά Skolem.** Ο Υπαρξιακός Προσδιορισμός αποτελεί ειδική περίπτωση μιας πιο γενικής διαδικασίας που ονομάζεται **μετατροπή κατά Skolem** (skolemization) , την οποία θα περιγράψουμε στην Ενότητα 2.4.3.

Εκτός του ότι ο Υπαρξιακός Προσδιορισμός είναι πιο σύνθετος από τον Καθολικό Προσδιορισμό, διατηρεί επίσης και έναν ελαφρώς διαφορετικό ρόλο στο συμπερασμό. Ενώ ο Καθολικός Προσδιορισμός μπορεί να εφαρμοστεί πολλές φορές για να παράγει πολλά διαφορετικά επακόλουθα, ο Υπαρξιακός Προσδιορισμός μπορεί να εφαρμοστεί μόνο μία φορά, και μετά η υπαρξιακά ποσοτικοποιημένη πρόταση μπορεί να απορριφθεί. Για παράδειγμα, από τη στιγμή που θα προσθέσουμε την πρόταση *Φόνος(Δολοφόνος, Θύμα)* δε χρειαζόμαστε πλέον την πρόταση $\exists x \, Φόνος(\chi, Θύμα)$. Αν θέλαμε να χρησιμοποιήσουμε πιο αυστηρή ορολογία θα λέγαμε ότι η νέα βάση γνώσης δεν είναι λογικά ισοδύναμη με την παλιά, αλλά μπορεί να αποδειχθεί ότι είναι **συμπερασματικά ισοδύναμη** με την έννοια ότι είναι ικανοποιήσιμη μόνο όταν είναι ικανοποιήσιμη και η αρχική βάση γνώσης.

Αναγωγή στον προτασιακό συμπερασμό

Από την στιγμή που έχουμε κανόνες για το συμπερασμό μη ποσοτικοποιημένων προτάσεων από ποσοτικοποιημένες προτάσεις, γίνεται εφικτή η μετατροπή του συμπερασμού πρώτης τάξης σε προτασιακό συμπερασμό. Στην ενότητα αυτή θα δώσουμε τις βασικές ιδέες· οι λεπτομέρειες περιγράφονται στην Ενότητα 2.4.3.

Η πρώτη ιδέα είναι ότι, όπως μια υπαρξιακά ποσοτικοποιημένη πρόταση μπορεί να αντικατασταθεί από μια συγκεκριμένη περίπτωση, έτσι και μια καθολικά ποσοτικοποιημένη πρόταση μπορεί να αντικατασταθεί από το σύνολο όλων των δυνατών περιπτώσεων. Για παράδειγμα, υποθέστε ότι η βάση γνώσης μας περιέχει μόνο τις προτάσεις

$$\forall\, x\; Βασιλιάς(\chi) \wedge Άπληστος(\chi) \Rightarrow Κακός(\chi) \qquad (2.4)$$

Βασιλιάς(Ιωάννης)
Άπληστος(Ιωάννης)
Αδελφός(Ριχάρδος, Ιωάννης).

Κατόπιν εφαρμόζουμε καθολικό προσδιορισμό στην πρώτη πρόταση χρησιμοποιώντας όλες τις δυνατές αντικαταστάσεις <u>βασικών όρων</u> από το λεξιλόγιο της βάσης γνώσης – σε αυτή την περίπτωση, { *χ* / *Ιωάννης* } και { *χ* / *Ριχάρδος* }. Παίρνουμε

$$Βασιλιάς(Ιωάννης) \wedge Άπληστος(Ιωάννης) \Rightarrow Κακός(Ιωάννης),$$
$$Βασιλιάς(Ριχάρδος) \wedge Άπληστος(Ριχάρδος) \Rightarrow Κακός(Ριχάρδος),$$

και απορρίπτουμε την καθολικά ποσοτικοποιημένη πρόταση. Τώρα, η βάση γνώσης είναι ουσιαστικά προτασιακή, αν δούμε τις βασικές ατομικές προτάσεις – *Βασιλιάς(Ιωάννης)*, *Άπληστος(Ιωάννης)* , κ.ο.κ. – ως σύμβολα πρότασης. Έτσι, μπορούμε να εφαρμόσουμε οποιονδήποτε από τους πλήρεις προτασιακούς αλγορίθμους του Κεφαλαίου 1 για να εξάγουμε συμπεράσματα όπως το *Κακός(Ιωάννης)*.

Όπως θα δείξουμε στην Ενότητα 2.4.3., αυτή η τεχνική μετατροπής σε προτασιακή μορφή (propositionalization) μπορεί να γενικευτεί πλήρως · δηλαδή, κάθε βάση γνώσης και ερώτημα πρώτης τάξης μπορούν να μετατραπούν σε προτασιακά έτσι ώστε να διατηρηθεί η λογική κάλυψη. Κατ' αυτόν τον τρόπο έχουμε μια ολοκληρωμένη διαδικασία αποφάσεων για την κάλυψη...ή ίσως όχι. Υπάρχει πρόβλημα: <u>Όταν η βάση γνώσης περιλαμβάνει ένα συναρτησιακό σύμβολο, το σύνολο των δυνατών αντικαταστάσεων βασικών όρων είναι άπειρο! Για παράδειγμα, αν η βάση γνώσης αναφέρει το σύμβολο *Πατέρας*, τότε μπορούν να κατασκευαστούν άπειροι ένθετοι όροι όπως *Πατέρας(Πατέρας(Πατέρας(Ιωάννης)))*. Οι προτασιακοί μας αλγόριθμοι θα έχουν πρόβλημα με ένα άπειρα μεγάλο σύνολο προτάσεων.</u>

Ευτυχώς, υπάρχει ένα διάσημο θεώρημα του Jacques Herbrand (1930), το οποίο λέει ότι αν μια πρόταση καλύπτεται από την αρχική βάση γνώσης πρώτης τάξης, τότε υπάρχει απόδειξη που αφορά μόνο ένα πεπερασμένο υποσύνολο της προτασιακής βάσης δεδομένων. Εφόσον οποιοδήποτε τέτοιο υποσύνολο διαθέτει ένα μέγιστο βάθος ένθεσης στους βασικούς όρους, μπορούμε να το εντοπίσουμε δημιουργώντας όλες τις περιπτώσεις του με σταθερά σύμβολα (*Ριχάρδος* και *Ιωάννης*), μετά όλους τους όρους με βάθος 1 (*Πατέρας(Ριχάρδος)* και *Πατέρας(Ιωάννης)*), μετά όλους τους όρους με βάθος 2, κ.ο.κ., μέχρι να μπορέσουμε να κατασκευάσουμε μια προτασιακή απόδειξη της λογικά καλυπτόμενης πρότασης.

Η προσέγγιση στο συμπερασμό πρώτης τάξης μέσω της μετατροπής σε προτασιακή μορφή που περιγράψαμε είναι πλήρης – δηλαδή, οποιαδήποτε καλυπτόμενη πρόταση είναι αποδείξιμη. Αυτό αποτελεί μια τεράστια επιτυχία, με δεδομένο ότι ο χώρος των δυνατών μοντέλων είναι άπειρος. Από την άλλη πλευρά, μέχρι να κατασκευαστεί η απόδειξη δε γνωρίζουμε ότι η πρόταση είναι καλυπτόμενη! Τι συμβαίνει αν η πρόταση δεν είναι καλυπτόμενη; Μπορούμε να το ξέρουμε; Όσον αφορά τη λογική πρώτης τάξης, φαίνεται ότι δεν μπορούμε. Η αποδεικτική διαδικασία μας μπορεί να συνεχίζεται επ' άπειρον, δημιουργώντας ένθετους όρους με όλο και μεγαλύτερο βάθος, όμως εμείς δεν μπορούμε να γνωρίζουμε αν θα παγιδευτεί σε ένα μάταιο βρόχο ή ξαφνικά θα εμφανιστεί η απόδειξη. Αυτό μοιάζει πάρα πολύ με το πρόβλημα τερματισμού των μηχανών Turing. Ο Alan Turing (1936) και ο Alonzo Church (1936) απέδειξαν, ο καθένας με τον δικό του τρόπο, το αναπόφευκτο της κατάστασης. *Το ερώτημα για την κάλυψη στη λογική πρώτης τάξης είναι* ***ημιαποφασίσιμο*** *(semidecidable) – δηλαδή, ενώ υπάρχουν αλγόριθμοι που απαντούν θετικά σε κάθε καλυπτόμενη πρόταση, δεν υπάρχει κάποιος αλγόριθμος που να απαντά αρνητικά σε κάθε μη καλυπτόμενη πρόταση.*

2.4.2 <u>ΕΝΟΠΟΙΗΣΗ ΚΑΙ ΑΝΥΨΩΣΗ</u>

Στην προηγούμενη ενότητα περιγράφηκε η αντίληψη σχετικά με το συμπερασμό πρώτης τάξης που υπήρχε στις αρχές της δεκαετίας του '60. Οι παρατηρητικοί αναγνώστες (και σίγουρα οι υπολογιστικοί επιστήμονες της λογικής εκείνης της εποχής) θα πρόσεξαν ότι η προσέγγιση της μετατροπής σε προτασιακή μορφή είναι μάλλον αναποδοτική. Για παράδειγμα, με δεδομένο το ερώτημα *Κακός(χ)* και τη βάση γνώσης της Εξίσωσης (2.4), η δημιουργία προτάσεων όπως $Βασιλιάς(Ριχάρδος) \wedge Άπληστος(Ριχάρδος) \Rightarrow Κακός(Ριχάρδος)$ φαίνεται παράξενη. Πράγματι, το συμπέρασμα *Κακός(Ιωάννης)* από τις προτάσεις

$$\forall x \, Βασιλιάς(x) \wedge Άπληστος(x) \Rightarrow Κακός(x),$$
$$Βασιλιάς(Ιωάννης)$$
$$Άπληστος(Ιωάννης)$$

φαίνεται απόλυτα προφανές σε έναν άνθρωπο. Τώρα θα δούμε πώς μπορούμε να το κάνουμε απόλυτα προφανές και σε έναν υπολογιστή.

Ένας κανόνας συμπερασμού πρώτης τάξης

Η παραγωγή του συμπεράσματος ότι ο Ιωάννης είναι κακός λειτουργεί ως εξής: βρες ένα χ τέτοιο ώστε το χ να είναι βασιλιάς και το χ να είναι άπληστος, και κατόπιν συμπέρανε ότι το χ είναι κακός. Γενικότερα, αν υπάρχει κάποια αντικατάσταση *θ* η οποία εξομοιώνει την προυπόθεση της συνεπαγωγής με τις προτάσεις που βρίσκονται ήδη στη βάση γνώσης, τότε μπορούμε να ισχυριστούμε το συμπέρασμα της συνεπαγωγής, μετά την εφαρμογή της *θ*. Στην περίπτωση αυτή, η αντικατάσταση { *χ* / *Ιωάννης* } επιτυγχάνει αυτόν το στόχο.

Στην πραγματικότητα το βήμα του συμπερασμού θα μπορούσε να κάνει ακόμα περισσότερα πράγματα. Υποθέστε ότι αντί να γνωρίζουμε ότι Άπληστος(Ιωάννης), γνωρίζουμε ότι όλοι είναι άπληστοι:

$$\forall y \, Άπληστος(y) \qquad (2.5)$$

Σε αυτή την περίπτωση θα θέλαμε να μπορούμε και πάλι να συμπεράνουμε ότι *Κακός(Ιωάννης)*, επειδή γνωρίζουμε ότι ο Ιωάννης είναι βασιλιάς (δεδομένο) και ότι ο Ιωάννης είναι άπληστος (επειδή όλοι είναι άπληστοι). Αυτό που χρειαζόμαστε για να πετύχουμε κάτι τέτοιο είναι να βρούμε μια αντικατάσταση τόσο για τις μεταβλητές στην πρόταση της συνεπαγωγής όσο και για τις μεταβλητές στις προτάσεις που θα ταιριάξουν. Στην περίπτωση αυτή, η εφαρμογή της αντικατάστασης {*χ* / *Ιωάννης* , *y* / *Ιωάννης* } στις προυποθέσεις συνεπαγωγής *Βασιλιάς(χ)* και *Άπληστος(χ)* και τις προτάσεις της βάσης γνώσης *Βασιλιάς(Ιωάννης)* και *Άπληστος(y)* θα τις εξομοιώσει. Έτσι μπορούμε να συνάγουμε το συμπέρασμα της συνεπαγωγής.

Αυτή η διαδικασία συμπερασμού μπορεί να εκφραστεί ως ένας κανόνας συμπερασμού ο οποίος ονομάζεται **Γενικευμένος Τρόπος του Θέτειν** (Generalized Modus Ponens) : Για ατομικές προτάσεις p_i, p'_i και q , όπου υπάρχει μια αντικατάσταση θ τέτοια ώστε SUBST(θ, p'_i) = SUBST(θ, p_i), για όλα τα i ,

$$\frac{p'_1, p'_2, \ldots, p'_n, (p_1 \wedge p_2 \wedge \ldots \wedge p_n \Rightarrow q)}{\text{SUBST}(\theta, q)}$$

Για αυτόν τον κανόνα υπάρχουν n+1 προυποθέσεις: οι n ατομικές προτάσεις p'_i και η μία συνεπαγωγή. Το συμπέρασμα είναι το αποτέλεσμα της αντικατάστασης θ στο επακόλουθο q. Στο παράδειγμα μας:

p'_1 είναι *Βασιλιάς(Ιωάννης)* p_1 είναι *Βασιλιάς(χ)*
p'_2 είναι *Άπληστος(y)* p_2 είναι *Άπληστος(χ)*
θ είναι *{χ/Ιωάννης, y/Ιωάννης}* q είναι *Κακός(χ)*
SUBST(θ, q) είναι *Κακός(Ιωάννης)*

Είναι εύκολο να αποδείξουμε ότι ο Γενικευμένος Τρόπος του Θέτειν αποτελεί έναν ορθό κανόνα συμπερασμού. Πρώτα παρατηρούμε ότι, για κάθε πρόταση p (της οποίας οι μεταβλητές θεωρούνται καθολικά ποσοτικοποιημένες), και για κάθε αντικατάσταση θ,

$$p \models \text{SUBST}(\theta, p) .$$

Αυτό ισχύει για τους ίδιους λόγους που ισχύει και ο κανόνας του Καθολικού Προσδιορισμού. Ισχύει, πιο συγκεκριμένα, για ένα θ που ικανοποιεί τις συνθήκες του κανόνα του Γενικευμένου Τρόπου του Θέτειν. Συνεπώς, από $p'_1, p'_2, \ldots, p'_n$ μπορούμε να συμπεράνουμε ότι

$$SUBST(\theta, p'_1) \wedge \ldots \wedge SUBST(\theta, p'_n)$$

και από την συνεπαγωγή $p_1 \wedge p_2 \wedge \ldots \wedge p_n \Rightarrow q$ μπορούμε να συμπεράνουμε ότι

$$SUBST(\theta, p_1) \wedge \ldots \wedge SUBST(\theta, p_n) \Rightarrow SUBST(\theta, q).$$

Τώρα, το θ στον Γενικευμένο τρόπο του Θέτειν είναι μια **ανυψωμένη** (lifted) έκδοση του Τρόπου του Θέτειν – ανυψώνει τον Τρόπο του Θέτειν από την προτασιακή λογική στη λογική πρώτης τάξης. Στο υπόλοιπο του κεφαλαίου θα δούμε ότι μπορούμε να αναπτύξουμε ανυψωμένες εκδόσεις των αλγορίθμων ανάλυσης που παρουσιάστηκαν στο Κεφάλαιο 1. Το βασικό πλεονέκτημα των ανυψωμένων κανόνων συμπερασμού σε σχέση με τη μετατροπή σε προτασιακή μορφή είναι ότι πραγματοποιούν μόνο εκείνες τις αντικαταστάσεις που απαιτούνται για την πρόοδο συγκεκριμένων διαδικασιών εξαγωγής συμπερασμάτων. Ένα ενδεχομένως μπερδεμένο σημείο είναι ότι κατά μία έννοια ο Γενικευμένος Τρόπος του Θέτειν είναι λιγότερο γενικός από τον Τρόπο του Θέτειν (ενότητα 1.4): Ο Τρόπος του Θέτειν επιτρέπει κάθε *α* στην αριστερή πλευρά της συνεπαγωγής, ενώ ο Γενικευμένος Τρόπος του Θέτειν απαιτεί μια συγκεκριμένη μορφή αυτής της πρότασης. Είναι γενικευμένος με την έννοια του ότι επιτρέπει οποιοδήποτε πλήθος από P'_i .

Ενοποίηση

Οι ανυψωμένοι κανόνες συμπερασμού απαιτούν την εύρεση αντικαταστάσεων οι οποίες κάνουν τις διάφορες λογικές εκφράσεις να φαίνονται ίδιες. Η διαδικασία αυτή ονομάζεται **ενοποίηση** (unification) και αποτελεί βασικό στοιχείο όλων των αλγορίθμων συμπερασμού πρώτης τάξης. Ο αλγόριθμος UNIFY δέχεται δύο προτάσεις και επιστρέφει έναν **ενοποιητή** (unifier) για αυτές, αν υπάρχει κάποιος:

UNIFY(p , q) = θ όπου *SUBST(θ, p) = SUBST(θ, q).*

Ας δούμε μερικά παραδείγματα για το πώς πρέπει να συμπεριφέρεται ο αλγόριθμος UNIFY. Υποθέστε ότι έχουμε ένα ερώτημα *Γνωρίζει(Γιάννης, χ)* : ποιον γνωρίζει ο Γιάννης; Κάποιες απαντήσεις σε αυτό το ερώτημα μπορούν να βρεθούν αν αναζητήσουμε όλες τις προτάσεις στη βάση γνώσης μας που ενοποιούνται με το *Γνωρίζει(Γιάννης, χ)*. Ακολουθούν τα αποτελέσματα της ενοποίησης με τέσσερις διαφορετικές προτάσεις που μπορεί να υπάρχουν στη βάση γνώσης.

UNIFY(*Γνωρίζει(Γιάννης, χ), Γνωρίζει(Γιάννης, Μαρία)*) = { *χ / Μαρία* }
UNIFY(*Γνωρίζει(Γιάννης, χ), Γνωρίζει(y, Βασίλης)*) = { *χ / Βασίλης, y / Γιάννης* }
UNIFY(*Γνωρίζει(Γιάννης, χ), Γνωρίζει(y, Μητέρα(y))*) = {*y/ Γιάννης, χ/ Μητέρα(Γιάννης)*}
UNIFY(*Γνωρίζει(Γιάννης, χ) , Γνωρίζει(χ, Ελισάβετ)*) = αποτυχία.

Η τελευταία ενοποίηση αποτυγχάνει επειδή το χ δεν μπορεί να έχει τις τιμές *Γιάννης* και *Ελισάβετ* την ίδια στιγμή. Τώρα, θυμηθείτε ότι *Γνωρίζει(χ, Ελισάβετ)* σημαίνει "Οποιοσδήποτε γνωρίζει την Ελισάβετ", *άρα θα πρέπει* να μπορούμε να εξάγουμε το συμπέρασμα ότι ο Γιάννης γνωρίζει την Ελισάβετ. Το πρόβλημα παρουσιάζεται μόνο επειδή οι δύο προτάσεις τυχαίνει να χρησιμοποιούν το ίδιο όνομα μεταβλητής, χ. Το πρόβλημα μπορεί να αποφευχθεί με την **κατά μέρος προτυποποίηση** (standardizing apart) της μίας εκ των δύο προτάσεων, δηλαδή με τη μετονομασία των μεταβλητών της ώστε να αποφευχθούν οι συγκρούσεις ονομάτων. Για παράδειγμα, μπορούμε να μετονομάσουμε το *χ* στο *Γνωρίζει(χ, Ελισάβετ)* σε z_{17} (ένα νέο όνομα μεταβλητής) χωρίς να αλλάξουμε τη σημασία του. Τώρα η ενοποίηση θα λειτουργήσει:

UNIFY(*Γνωρίζει(Γιάννης, χ), Γνωρίζει(* z_{17} *, Ελισάβετ)*) = { *χ/ Ελισάβετ,* z_{17} / *Γιάννης* }.

Υπάρχει ένα ακόμα πρόβλημα: είπαμε ότι ο UNIFY θα πρέπει να επιστρέφει μια αντικατάσταση η οποία κάνει τα δύο ορίσματα να φαίνονται το ίδιο. Ωστόσο, μπορεί να υπάρχουν περισσότεροι του ενός ενοποιητές, Για παράδειγμα, η πρόταση : UNIFY(*Γνωρίζει(Γιάννης, χ), Γνωρίζει(y, z)*)

θα μπορούσε να επιστρέψει

{*y/ Γιάννης, χ/z*} ή { *y/Γιάννης, χ/Γιάννης, z/ Γιάννης* }.

Ο πρώτος ενοποιητής δίνει ως αποτέλεσμα της ενοποίησης *Γνωρίζει(Γιάννης, z)*, ενώ ο δεύτερος δίνει *Γνωρίζει(Γιάννης, Γιάννης)*. Το δεύτερο αποτέλεσμα μπορεί να επιτευχθεί από το πρώτο με μια επιπλέον αντικατάσταση { *z / Γιάννης* }· λέμε ότι ο πρώτος ενοποιητής είναι πιο γενικός από τον δεύτερο, επειδή θέτει λιγότερους περιορισμούς στις τιμές των μεταβλητών. Αποδεικνύεται ότι, για κάθε ενοποιήσιμο ζεύγος παραστάσεων, υπάρχει ένας και μοναδικός πιο γενικός ενοποιητής (Most General Unifier, MGU) που είναι μοναδικός για τη μετονομασία των μεταβλητών. Σε αυτή την περίπτωση είναι { *y/ Γιάννης, x/z* }.

Στην Εικόνα 2.3 μπορείτε να δείτε ένα αλγόριθμο για τον υπολογισμό των πιο γενικών ενοποιητών. Η διαδικασία είναι πολύ απλή: αναδρομική εξερεύνηση των δύο παραστάσεων ταυτόχρονα “πλευρά-πλευρά”, κατασκευή ενός ενοποιητή κατά τη διάρκεια της διαδικασίας, αλλά αποτυχία σε περίπτωση που δύο αντίστοιχα σημεία των δομών δεν ταιριάζουν. Υπάρχει ένα δαπανηρό (από άποψη απόδοσης) βήμα: κατά το ταίριασμα μιας μεταβλητής με ένα σύνθετο όρο θα πρέπει να γίνεται έλεγχος για το αν η ίδια η μεταβλητή περιέχεται μέσα στον όρο· αν περιέχεται, το ταίριασμα αποτυγχάνει επειδή δεν μπορεί να κατασκευαστεί κάποιος συνεπής ενοποιητής. Αυτός ο λεγόμενος **έλεγχος ύπαρξης** (occur check) κάνει την πολυπλοκότητα όλου του αλγορίθμου τετραγωνική ως προς το μέγεθος των παραστάσεων που ενοποιούνται. Μερικά συστήματα, μεταξύ των οποίων όλα τα συστήματα λογικού προγραμματισμού, απλώς παραλείπουν τον έλεγχο ύπαρξης με αποτέλεσμα μερικές φορές να πραγματοποιούν μη έγκυρους συμπερασμούς· άλλα συστήματα χρησιμοποιούν πιο σύνθετους αλγόριθμους με πολυπλοκότητα γραμμικού χρόνου.

```
function UNIFY(x, y, θ) returns μια αντικατάσταση ώστε το χ και το y να εξομοιωθούν
    inputs: x, μεταβλητή, σταθερά, λίστα, ή σύνθετος όρος
            y, μεταβλητή, σταθερά, λίστα, ή σύνθετος όρος
            θ, η αντικατάσταση που έχει κατασκευαστεί μέχρι τώρα (προαιρετικό,
               προεπιλεγμένη τιμή κενή)

    if θ = αποτυχία then return αποτυχία
    else if x = y then return θ
    else if VARIABLE?( x ) then return UNIFY-VAR( x, y, θ )
    else if VARIABLE?( y ) then return UNIFY-VAR( y, x, θ )
    else if COMPOUND?( x ) and COMPOUND?( y ) then
        return UNIFY( ARGS[x], ARGS[y], UNIFY(OP[x], OP[y], θ ))
    else if LIST?(x) and LIST?(y) then
        return UNIFY( REST[x], REST[y], UNIFY(FIRST[x], FIRST[y], θ))
    else return αποτυχία



function UNIFY-VAR(var, x, θ) returns μια αντικατάσταση
    inputs: var, μεταβλητή
            x, οποιαδήποτε παράσταση
            θ, η αντικατάσταση που έχει κατασκευαστεί μέχρι τώρα

    if (var/val)∈θ then return UNIFY( val, x, θ )
    else if (x/val)∈θ then return UNIFY( var, val, θ )
    else if OCCUR-CHECK?( var, x ) then return αποτυχία
    else return πρόσθεσε {var/x} στο θ
```

Εικόνα 2.3 Ο αλγόριθμος ενοποίησης. Ο αλγόριθμος λειτουργεί συγκρίνοντας τις δομές των εισόδων στοιχείο προς στοιχείο. Η αντικατάσταση θ που αποτελεί το όρισμα στον αλγόριθμο κατασκευάζεται κατά τη διάρκεια της εκτέλεσης και χρησιμοποιείται για να εξασφαλιστεί ότι μεταγενέστερες συγκρίσεις είναι συνεπείς με τις δεσμεύσεις που θεσπίστηκαν νωρίτερα. Σε μια σύνθετη παράσταση, όπως η F(A, B), η συνάρτηση OP απομονώνει το συναρτησιακό σύμβολο F και η συνάρτηση ARGS απομονώνει τη λίστα ορισμάτων (A, B).

Αποθήκευση και ανάκτηση

Πίσω από τις συναρτήσεις TELL και ASK που χρησιμοποιούνται για να πληροφορούν και να θέτουν ερωτήματα σε μια βάση γνώσης βρίσκονται οι πιο στοιχειώδεις συναρτήσεις STORE και FETCH. Η συνάρτηση STORE(s) αποθηκεύει μια πρόταση s στη βάση γνώσης και η συνάρτηση FETCH(q) επιστρέφει όλους τους ενοποιητές με τους οποίους το ερώτημα q ενοποιείται με κάποια πρόταση της βάσης γνώσης. Το πρόβλημα που χρησιμοποιήσαμε για να παρουσιάσουμε την ενοποίηση − την εύρεση όλων των γεγονότων που ενοποιούνται με το *Γνωρίζει(Γιάννης, x)* − είναι μια περίπτωση "προσκόμισης" (fetching).

Ο απλούστερος τρόπος υλοποίησης των συναρτήσεων STORE και FETCH είναι η τοποθέτηση όλων των γεγονότων της βάσης γνώσης σε μια μεγάλη λίστα· κατόπιν, όταν δίνεται ένα ερώτημα *q*, καλείται η UNIFY(*q*, *s*) για κάθε πρόταση *s* στη λίστα αυτή. Μια τέτοια διαδικασία είναι αναποδοτική, αλλά λειτουργεί, και είναι το μόνο που χρειάζεστε για να κατανοήσετε το υπόλοιπο του κεφαλαίου.

2.4.3 **ΑΝΑΛΥΣΗ**

Είδαμε στο Κεφάλαιο 1 ότι η προτασιακή ανάλυση είναι μια πλήρης διαδικασία συμπερασμού με διάψευση για την προτασιακή λογική. Σε αυτή την ενότητα θα δούμε τον τρόπο επέκτασης της ανάλυσης για τη λογική πρώτης τάξης.

Το ερώτημα σχετικά με την ύπαρξη πλήρους διαδικασίας απόδειξης έχει άμεσο ενδιαφέρον για τους μαθηματικούς. Εάν μπορεί να βρεθεί μια πλήρης διαδικασία απόδειξης για μαθηματικές δηλώσεις, υπάρχουν δύο συνέπειες: πρώτον, όλες οι εικασίες μπορούν να αποδειχθούν με μηχανικό τρόπο· δεύτερον, το σύνολο των μαθηματικών μπορεί να οριστεί ως λογική συνέπεια ενός συνόλου θεμελιωδών αξιωμάτων. Έτσι το ερώτημα της πληρότητας έχει προκαλέσει πολλές από τις σημαντικότερες εργασίες που έχουν πραγματοποιηθεί τον 20ό αιώνα. Το 1930 ο Γερμανός μαθηματικός Kurt Godel απέδειξε το πρώτο **θεώρημα πληρότητας** (completeness theorem) για τη λογική πρώτης τάξης, δείχνοντας ότι κάθε καλυπτόμενη πρόταση έχει πεπερασμένη απόδειξη. (Δεν είχε βρεθεί καμία πραγματικά *πρακτική* διαδικασία απόδειξης μέχρι την έκδοση του αλγόριθμου ανάλυσης από τον J. A. Robinson το 1965.) Το 1931 ο Godel απόδειξε ένα ακόμα πιο διάσημο **θεώρημα μη πληρότητας** (incompleteness theorem). Το θεώρημα ορίζει ότι ένα λογικό σύστημα το οποίο περιλαμβάνει την αρχή της επαγωγής – χωρίς την οποία μπορεί να κατασκευαστεί μόνο ένα πολύ μικρό μέρος των διακριτών μαθηματικών – είναι κατανάγκην μη πλήρες. Συνεπώς, θα υπάρχουν προτάσεις που θα είναι λογικά καλυπτόμενες αλλά δεν θα έχουν πεπερασμένη απόδειξη εντός του συστήματος. Μιλώντας μεταφορικά, η βελόνα μπορεί να βρίσκεται μέσα στα άχυρα, αλλά καμία διαδικασία δεν μπορεί να εγγυηθεί ότι θα βρεθεί.

Παρά το θεώρημα Godel, έχουν εφαρμοστεί σε ευρεία έκταση αποδείκτες θεωρημάτων που βασίζονται στην ανάλυση για την εύρεση μαθηματικών θεωρημάτων, συμπεριλαμβανομένων πολλών θεωρημάτων για τα οποία δεν υπήρχε προηγουμένως γνωστή απόδειξη. Οι αποδείκτες θεωρημάτων έχουν επίσης χρησιμοποιηθεί, μεταξύ άλλων, για την επαλήθευση σχεδίων υλικού και για τη δημιουργία λογικά ορθών προγραμμάτων.

Συζευκτική κανονική μορφή για τη λογική πρώτης τάξης

Όπως και στην προτασιακή περίπτωση, η ανάλυση πρώτης τάξης απαιτεί ότι οι προτάσεις θα είναι σε συζευκτική κανονική μορφή (conjuctive normal form, CNF) – δηλαδή μια σύζευξη προτάσεων, όπου κάθε πρόταση είναι διάζευξη λεκτικών. Τα λεκτικά μπορούν να περιέχουν μεταβλητές, οι οποίες θεωρούνται ότι είναι καθολικά ποσοτικοποιημένες. Για παράδειγμα, η πρόταση

$$\forall x\, Αμερικανός(\chi) \wedge Όπλο(y) \wedge Πώληση(x, y, z) \wedge Αφιλόξενος(z) \Rightarrow Εγκληματίας(x)$$

γίνεται στη μορφή CNF

$$\neg Αμερικανός(x) \vee \neg Όπλο(y) \vee \neg Πώληση(x, y, z) \vee \neg Αφιλόξενος(z) \vee Εγκληματίας(x).$$

Κάθε πρόταση λογικής πρώτης τάξης μπορεί να μετατραπεί σε μια ισοδύναμη όσον αφορά το συμπερασμό πρόταση CNF. Πιο συγκεκριμένα, η πρόταση CNF δεν θα ικανοποιείται όταν η αρχική πρόταση δεν θα ικανοποιείται, έτσι έχουμε τη βάση για την εύρεση απόδειξης με την αντίφαση των προτάσεων CNF.

Η διαδικασία μετατροπής σε μορφή CNF μοιάζει πολύ με την προτασιακή περίπτωση, την οποία είδαμε στην Ενότητα 1.5. <u>Η κύρια διαφορά προκύπτει από την ανάγκη εξάλειψης των υπαρξιακών ποσοδεικτών</u>. Θα παρουσιάσουμε τη διαδικασία αυτή μεταφράζοντας την πρόταση "Όλοι όσοι αγαπούν όλα τα ζώα αγαπούνται από κάποιον", ή

$$\forall x [\forall y\, Ζώο(y) \Rightarrow Αγαπά(x, y)] \Rightarrow [\exists y\, Αγαπά(y, x)].$$

Τα βήματα είναι τα εξής:

- **Απαλοιφή των συνεπαγωγών:**

$$\forall x [\neg \forall y \neg Ζώο(y) \vee Αγαπά(x, y)] \vee [\exists y\, Αγαπά(y, x)].$$

- **Κίνηση του $\neg$ προς τα μέσα:** Εκτός από τους συνηθισμένους κανόνες για τα αρνητικά συνδετικά, χρειαζόμαστε κανόνες και για τους αρνητικούς ποσοδείκτες. Έτσι έχουμε

το $\neg \forall x\, p$ γίνεται $\exists x \neg p$

το $\neg \exists x\, p$ γίνεται $\forall x \neg p$.

Η πρόταση μας υφίσταται τους παρακάτω μετασχηματισμούς:

$$\forall x[\exists y \neg(\neg Ζώο(y) \vee Αγαπά(x, y))] \vee [\exists y\, Αγαπά(y, x)].$$

$$\forall x[\exists y \neg\neg Ζώο(y) \wedge \neg Αγαπά(x, y)] \vee [\exists y\, Αγαπά(y, x)].$$

$$\forall x[\exists y\, Ζώο(y) \wedge \neg Αγαπά(x, y)] \vee [\exists y\, Αγαπά(y, x)].$$

Προσέξτε πώς ο καθολικός ποσοδείκτης (forall $\forall$ y) στην αρχή της συνεπαγωγής έγινε υπαρξιακός ποσοδείκτης. Η πρόταση μας τώρα έχει γίνει "Είτε υπάρχει κάποιο ζώο που ο χ δεν αγαπά, ή (εάν δεν συμβαίνει αυτό) κάποιος αγαπά τον χ". Το μήνυμα της αρχικής πρότασης έχει εμφανώς διατηρηθεί.

- **Προτυποποίηση των μεταβλητών:** Για προτάσεις όπως η $(\forall x\, P(x)) \vee (\exists x\, Q(x))$ που χρησιμοποιούν δύο φορές το ίδιο όνομα μεταβλητής, αλλάζουμε το όνομα της μίας από τις μεταβλητές. Έτσι θα αποφύγουμε αργότερα τη σύγχυση, μετά την απαλοιφή των ποσοδεικτών. Άρα έχουμε

  $$\forall x[\exists y\, Ζώο(y) \wedge \neg Αγαπά(x, y)] \vee [\exists z\, Αγαπά(z, x)].$$

- **Μετατροπή κατά Skolem:** Η **μετατροπή κατά Skolem** (Skolemization) είναι η διαδικασία αφαίρεσης των υπαρξιακών ποσοδεικτών μέσω της απαλοιφής. Στην απλή περίπτωση, είναι ακριβώς όπως ο κανόνας του Υπαρξιακού Προσδιορισμού της Ενότητας 2.4.1: μεταφράζουμε το $\exists x\, P(x)$ σε P(A), όπου A είναι μια νέα σταθερά. Εάν όμως εφαρμόσουμε αυτόν τον κανόνα στο παράδειγμα πρότασής μας, παίρνουμε

  $$\forall x[Ζώο(A) \wedge \neg Αγαπά(x, A)] \vee [Αγαπά(B, x)].$$

  που έχει τελείως λανθασμένο νόημα: λέει ότι είτε όλοι δεν καταφέρνουν να αγαπήσουν ένα συγκεκριμένο ζώο Α, ή ότι αγαπιούνται από μια συγκεκριμένη ύπαρξη Β. Στην πραγματικότητα η αρχική μας πρόταση επιτρέπει σε κάθε πρόσωπο να μην αγαπά κάποιο διαφορετικό ζώο ή να αγαπιέται από ένα διαφορετικό πρόσωπο. Συνεπώς θέλουμε οι οντότητες Skolem να εξαρτώνται από το x :

  $$\forall x[Ζώο(F(x)) \wedge \neg Αγαπά(x, F(x))] \vee [Αγαπά(G(x), x)].$$

  Εδώ τα F και G είναι **συναρτήσεις Skolem** (Skolem Functions). Ο γενικός κανόνας είναι ότι τα ορίσματα της συνάρτησης Skolem είναι όλες οι καθολικά ποσοτικοποιημένες μεταβλητές στην εμβέλεια των οποίων εμφανίζεται ο υπαρξιακός ποσοδείκτης. Όπως συμβαίνει και με τον Υπαρξιακό

Προσδιορισμό, η πρόταση Skolem ικανοποιείται ακριβώς όταν ικανοποιείται και η αρχική πρόταση.

- **Κατάργηση καθολικών ποσοτικοποιητών:** Σε αυτό το σημείο όλες οι εναπομείνασες μεταβλητές πρέπει να είναι καθολικά ποσοτικοποιημένες. Επιπλέον, η πρόταση είναι ισοδύναμη με μια πρόταση στην οποία όλοι οι καθολικοί ποσοδείκτες έχουν μετακινηθεί στα αριστερά. Άρα μπορούμε να καταργήσουμε τους καθολικούς ποσοδείκτες:

  $$[Ζώο(F(x)) \wedge \neg Αγαπά(x, F(x))] \vee [Αγαπά(G(x), x)].$$

- **Κατανομή του $\vee$ ως προς το $\wedge$:**

  $$[Ζώο(F(x)) \vee Αγαπά(G(x), x)] \wedge [\neg Αγαπά(x, F(x)) \vee Αγαπά(G(x), x)].$$

  Αυτό το βήμα μπορεί να απαιτεί επίσης την απλοποίηση των ένθετων συζεύξεων και διαζεύξεων.

Η πρόταση είναι πλέον σε μορφή CNF και αποτελείται από δύο προτάσεις. Είναι αρκετά δυσανάγνωστη. (Ίσως βοηθά να εξηγήσουμε ότι η συνάρτηση Skolem F(x) αναφέρεται στο ζώο που πιθανώς δεν αγαπά ο χ, ενώ η G(x) αναφέρεται σε κάποιον που μπορεί να αγαπά τον χ.) Ευτυχώς, σπανίως οι άνθρωποι χρειάζεται να δουν προτάσεις CNF – η διαδικασία μετάφρασης μπορεί εύκολα να αυτοματοποιηθεί.

Ο κανόνας συμπερασμού για την ανάλυση

Ο κανόνας ανάλυσης για τις προτάσεις της λογικής πρώτης τάξης είναι απλώς μια γενικευμένη εκδοχή του κανόνα προτασιακής ανάλυσης που δώσαμε στην Ενότητα 1.5. Δύο προτάσεις οι οποίες θεωρούνται ότι είναι κατά μέρος προτυποποιημένες, έτσι ώστε να μην έχουν κοινές μεταβλητές, μπορούν να αναλυθούν εάν περιέχουν συμπληρωματικά λεκτικά. Τα προτασιακά λεκτικά είναι συμπληρωματικά εάν το ένα είναι η άρνηση του άλλου. Τα λεκτικά πρώτης τάξης είναι συμπληρωματικά εάν το ένα <u>ενοποιείται</u> με την άρνηση του άλλου. Έτσι έχουμε

$$\frac{l_1 \vee \ldots \vee l_k, \; m_1 \vee \ldots \vee m_n}{SUBST(\theta, l_1 \vee \ldots \vee l_{i-1} \vee l_{i+1} \vee \ldots \vee l_k \vee m_1 \vee \ldots \vee m_{j-1} \vee m_{j+1} \vee \ldots \vee m_n)}$$

όπου $UNIFY(l_i, \neg m_j) = \theta$.

Για παράδειγμα μπορούμε να αναλύσουμε τις δύο προτάσεις

$[Ζώο(F(x)) \vee Αγαπά(G(x), x)]$ και $[\neg Αγαπά(u, v) \vee \neg Σκότωσε(u, v)]$

απαλείφοντας τα συμπληρωματικά λεκτικά $Αγαπά(G(x), x)$ και $\neg Αγαπά(u, v)$ με τον ενοποιητή θ = { u/G(x), v/x } για να πάρουμε την **αναλυθείσα** (resolvent) πρόταση

$[Ζώο(F(x)) \vee \neg Σκότωσε(G(x), x)]$.

Ο κανόνας που μόλις αναφέραμε ονομάζεται κανόνας **δυαδικής ανάλυσης** (binary resolution), επειδή αναλύει δύο ακριβώς λεκτικά. Ο κανόνας δυαδικής ανάλυσης από μόνος του δεν μας δίνει μια πλήρη διαδικασία συμπερασμού. Ο πλήρης κανόνας ανάλυσης αναλύει υποσύνολα λεκτικών σε κάθε πρόταση τα οποία μπορούν να ενοποιηθούν. Μια εναλλακτική προσέγγιση είναι να επεκτείνουμε την **παραγοντοποίηση** (factoring) – την κατάργηση περιττών λεκτικών – για τη λογική πρώτης τάξης. Η προτασιακή παραγοντοποίηση απλοποιεί δύο λεκτικά σε ένα εάν είναι *πανομοιότυπα*· η παραγοντοποίηση πρώτης τάξης απλοποιεί δύο λεκτικά σε ένα εάν είναι *ενοποιήσιμα*. Ο ενοποιητής θα πρέπει να εφαρμοστεί σε ολόκληρη την πρόταση. Ο συνδυασμός δυαδικής ανάλυσης και της παραγοντοποίησης είναι πλήρης.

Παραδείγματα απόδειξης

Η ανάλυση αποδεικνύει ότι ΚΒ $\models \alpha$ αποδεικνύοντας ότι το $KB \wedge \neg \alpha$ δεν είναι ικανοποιήσιμο, δηλαδή παράγοντας την κενή πρόταση. Η αλγοριθμική προσέγγιση είναι πανομοιότυπη με την προτασιακή περίπτωση, που περιγράψαμε στην Εικόνα 1.6, έτσι δεν θα την επαναλάβουμε εδώ. Αντί γι'αυτό θα δώσουμε δύο παραδείγματα απόδειξης.

Δείτε το ακόλουθο πρόβλημα:

- Ο νόμος λέει ότι η πώληση όπλων σε αντίπαλα (των Η.Π.Α) κράτη από Αμερικανούς πολίτες αποτελεί κακούργημα. Η χώρα Νόνο, εχθρός της Αμερικής, διαθέτει μερικούς πυραύλους, και όλοι οι πύραυλοι πουλήθηκαν σε αυτήν από τον Συνταγματάρχη Γουέστ, ο οποίος είναι Αμερικανός.

Θα αποδείξουμε ότι ο Γουέστ είναι εγκληματίας. Πρώτα θα παρουσιάσουμε αυτά τα γεγονότα ως προτάσεις πρώτης τάξης.

"...η πώληση όπλων σε αντίπαλα (των Η.Π.Α) κράτη από Αμερικανούς πολίτες αποτελεί κακούργημα...":

$$Αμερικανός(χ) \wedge Όπλο(y) \wedge Πώληση(x, y, z) \wedge Αφιλόξενος(z) \Rightarrow Εγκληματίας(x) \quad (2.6)$$

"...Η χώρα Νόνο.. διαθέτει μερικούς πυραύλους..." :

$\exists x\, Κατέχει(Νόνο, x) \wedge Πύραυλος(x)$ η οποία μετατρέπεται σε δύο προτάσεις από την Υπαρξιακή Απαλοιφή (Existential Elimination), με την παρουσία μιας νέας σταθεράς M_1 :

$$Κατέχει(Νόνο, M_1) \quad (2.7)$$

$$Πύραυλος(M_1) \quad (2.8)$$

"...Όλοι οι πύραυλοι πουλήθηκαν σε αυτήν από τον Συνταγματάρχη Γουέστ..":

$$Πύραυλος(χ) \wedge Κατέχει(Νόνο, χ) \Rightarrow Πώληση(Γουέστ, χ, Νόνο) \quad (2.9)$$

Πρέπει επίσης να γνωρίζουμε ότι οι πύραυλοι είναι όπλα:

$$Πύραυλος(χ) \Rightarrow Όπλο(χ) \quad (2.10)$$

καθώς και ότι ένας εχθρός της Αμερικής θεωρείται "αφιλόξενος":

$$Εχθρός(χ, Αμερική) \Rightarrow Αφιλόξενος(χ) \quad (2.11)$$

"...από τον Γουέστ, ο οποίος είναι Αμερικανός....":

$$Αμερικανός(Γουέστ) \quad (2.12)$$

"Η χώρα Νόνο, εχθρός της Αμερικής...":

$$Εχθρός(Νόνο, Αμερική) \quad (2.13)$$

Αυτή η βάση γνώσης δεν περιέχει συναρτησιακά σύμβολα, και αποτελεί έτσι μια περίπτωση των βάσεων γνώσης της κατηγορίας **Datalog** – δηλαδή, σύνολα από οριστικές προτάσεις πρώτης τάξης χωρίς συναρτησιακά σύμβολα. Η απουσία συναρτησιακών συμβόλων κάνει το συμπερασμό πολύ ευκολότερο.

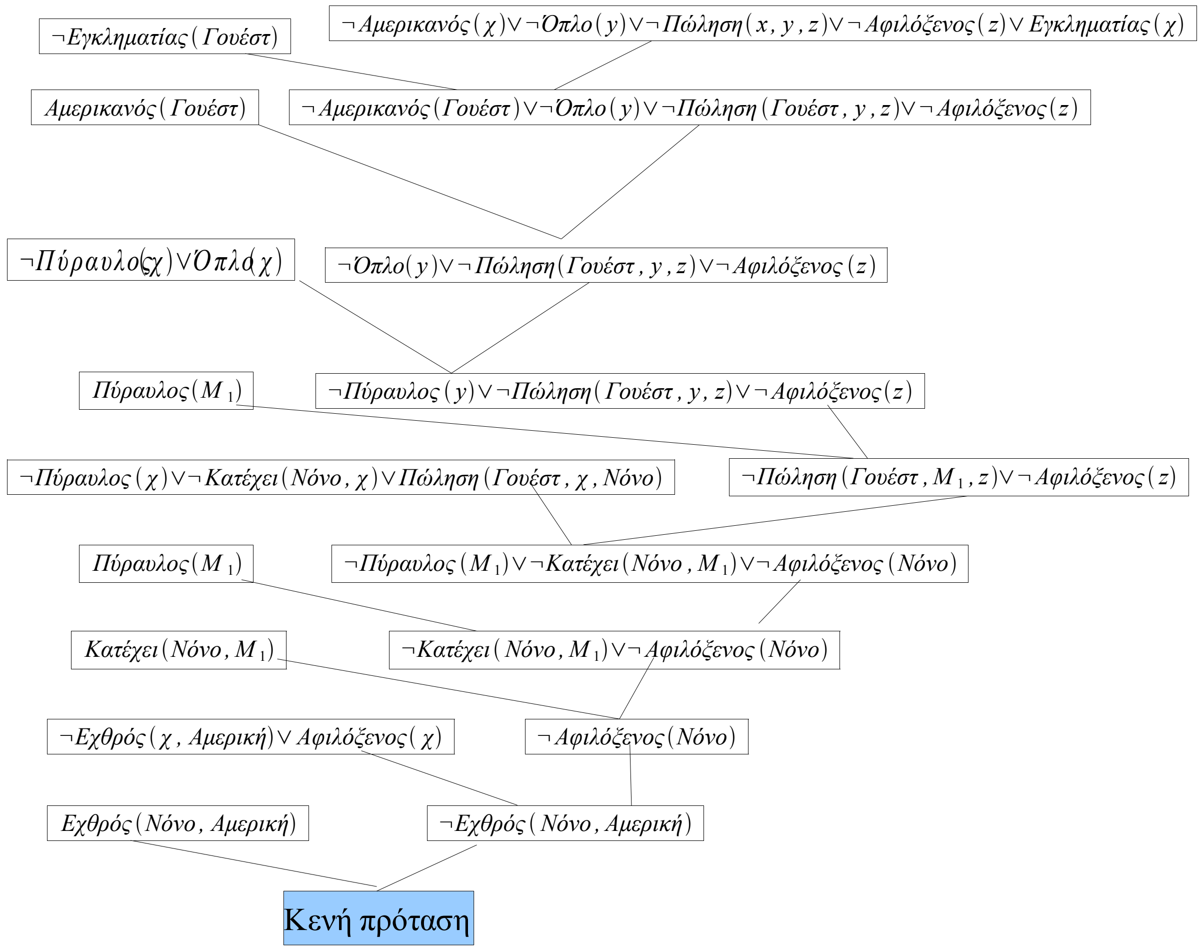
Οι προτάσεις σε μορφή CNF είναι

$$\neg Αμερικανός(χ) \vee \neg Όπλο(y) \vee \neg Πώληση(x, y, z) \vee \neg Αφιλόξενος(z) \vee Εγκληματίας(x).$$
$$\neg Πύραυλος(χ) \vee \neg Κατέχει(Νόνο, χ) \vee Πώληση(Γουέστ, χ, Νόνο).$$
$$\neg Εχθρός(χ, Αμερική) \vee Αφιλόξενος(χ).$$

¬*Πύραυλος*(*χ*) ∨ *Όπλο*(*χ*).
Κατέχει(*Νόνο*, M_1).
Πύραυλος(M_1).
Αμερικανός(*Γουέστ*).
Εχθρός(*Νόνο*, *Αμερική*).

Συμπεριλαμβάνουμε επίσης τον αρνητικό στόχο ¬*Εγκληματίας*(*Γουέστ*) .

Η απόδειξη μέσω ανάλυσης φαίνεται στην Εικόνα 2.4. Προσέξτε τη δομή: έναν "κορμό" που ξεκινά με την πρόταση στόχου και αναλύει ως προς τις προτάσεις από τη βάση γνώσης μέχρι να δημιουργηθεί μια κενή πρόταση.



Εικόνα 2.4 Μια απόδειξη μέσω ανάλυσης ότι ο Γουέστ είναι εγκληματίας.

Το δεύτερο παράδειγμα μας χρησιμοποιεί τη μετατροπή κατά Skolem και περιλαμβάνει προτάσεις που δεν είναι οριστικές. Αυτό οδηγεί σε μια κάπως πιο πολύπλοκη δομή απόδειξης. Σε απλά Ελληνικά, το πρόβλημα είναι το εξής:

- Όλους όσους αγαπούν όλα τα ζώα τους αγαπά κάποιος.
  Όποιον σκοτώνει κάποιο ζώο δεν τον αγαπά κανείς.
  Ο Τζακ αγαπά όλα τα ζώα.
  Είτε ο Τζακ είτε ένα Παράξενο πρόσωπο σκότωσε τη γάτα, που ονομάζεται Τούνα.
  Μήπως το Παράξενο πρόσωπο σκότωσε τη γάτα;

Αρχικά εκφράζουμε σε λογική πρώτης τάξης τις αρχικές προτάσεις, κάποιες γνώσεις υποβάθρου, και τον αρνητικό στόχο G:

Α. $\forall x [\forall y\, Ζώο(y) \Rightarrow Αγαπά(x, y)] \Rightarrow [\exists y\, Αγαπά(y, x)]$

Β. $\forall x [\exists y\, Ζώο(y) \land Σκότωσε(x, y)] \Rightarrow [\forall z \neg Αγαπά(z, x)]$

Γ. $\forall x\, Ζώο(x) \Rightarrow Αγαπά(Τζακ, x)$

Δ. $Σκότωσε(Τζακ, Τούνα) \lor Σκότωσε(Παράξενο, Τούνα)$

Ε. $Γάτα(Τούνα)$

ΣΤ. $\forall x\, Γάτα(x) \Rightarrow Ζώο(x)$

$\neg$Ζ. $\neg Σκότωσε(Παράξενο, Τούνα)$

Τώρα εφαρμόζουμε τη διαδικασία μετατροπής για να μετατρέψουμε την κάθε πρόταση σε μορφή CNF:

Α1. $Ζώο(F(x)) \lor Αγαπά(G(x), x)$

Α2. $\neg Αγαπά(x, F(x)) \lor Αγαπά(G(x), x)$

Β. $\neg Ζώο(\chi) \lor \neg Σκότωσε(x, y) \lor \neg Αγαπά(z, x)$

Γ. $\neg Ζώο(\chi) \lor Αγαπά(Τζακ, x)$

Δ. $Σκότωσε(Τζακ, Τούνα) \lor Σκότωσε(Παράξενο, Τούνα)$

Ε. $\mathit{Γάτα}(\mathit{Τούνα})$

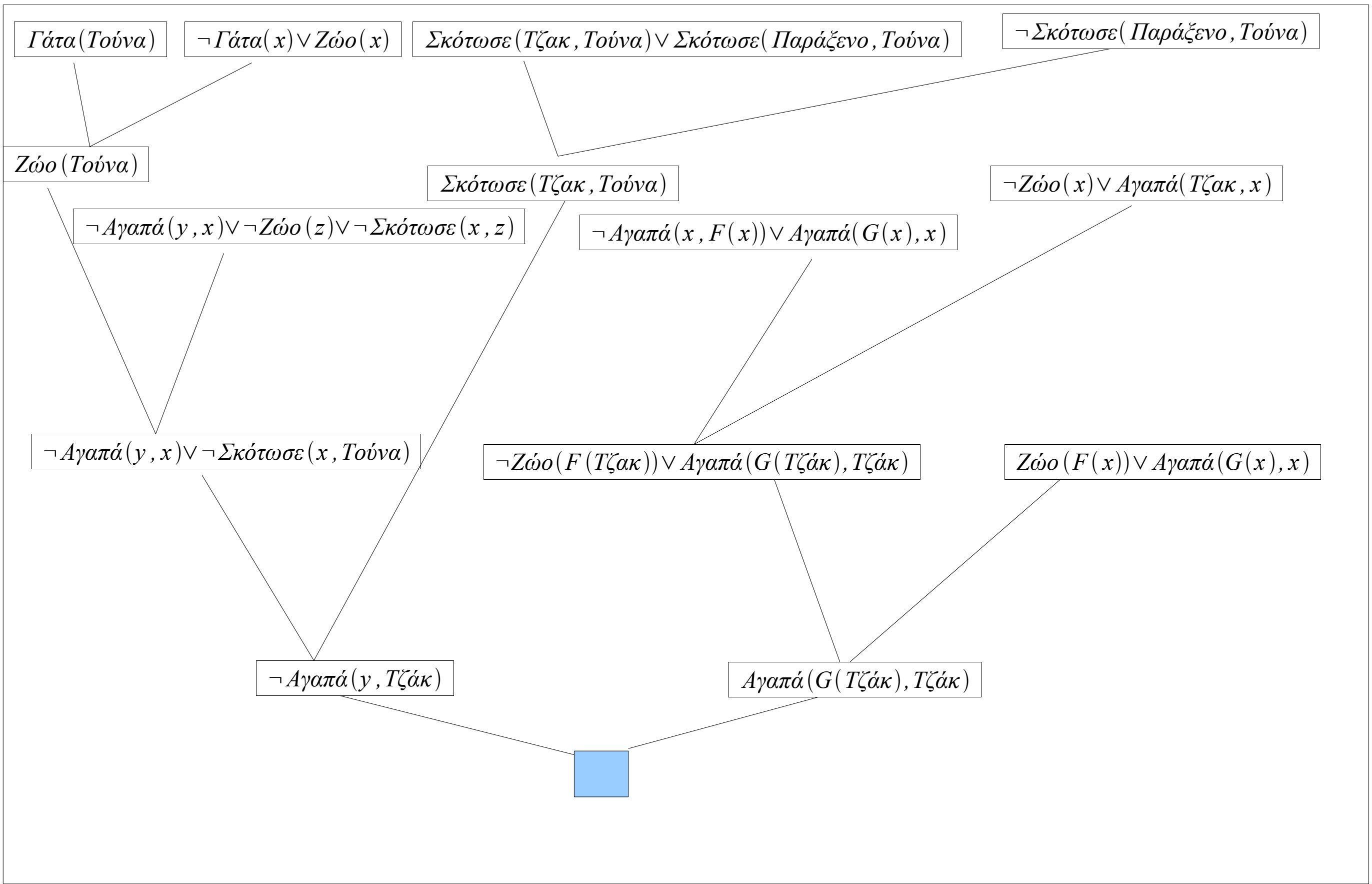
ΣΤ. $\neg \mathit{Γάτα}(x) \vee \mathit{Ζώο}(x)$

¬Ζ. $\neg \mathit{Σκότωσε}(\mathit{Παράξενο}, \mathit{Τούνα})$

Η απόδειξη μέσω ανάλυσης ότι το Παράξενο πρόσωπο σκότωσε τη γάτα φαίνεται στην Εικόνα 2.5. Σε απλά Ελληνικά, η απόδειξη μπορεί να παραφραστεί ως εξής:

- Έστω ότι το Παράξενο πρόσωπο δεν σκότωσε την Τούνα. Γνωρίζουμε ότι το έκανε είτε ο Τζακ είτε το Παράξενο πρόσωπο· συνεπώς πρέπει να το έκανε ο Τζακ. Όμως η Τούνα είναι γάτα, και οι γάτες είναι ζώα, άρα η Τούνα είναι ζώο. Επειδή όποιον σκοτώνει κάποιο ζώο δεν τον αγαπά κανείς, ξέρουμε ότι κανένας δεν αγαπά τον Τζακ. Από την άλλη πλευρά, ο Τζακ αγαπά όλα τα ζώα, έτσι κάποιος τον αγαπά· συνεπώς έχουμε αντίφαση. Άρα το Παράξενο πρόσωπο σκότωσε τη γάτα.

Η απόδειξη απαντά στο ερώτημα "Μήπως το Παράξενο πρόσωπο σκότωσε τη γάτα;" αλλά συχνά θέλουμε να θέσουμε πιο γενικά ερωτήματα όπως "Ποιος σκότωσε τη γάτα;" Η ανάλυση μπορεί να το κάνει αυτό, αλλά χρειάζεται περισσότερη δουλειά για την εύρεση απάντησης. Ο στόχος είναι ο $\exists w\, \mathit{Σκότωσε}(w, \mathit{Τούνα})$, ο οποίος στην άρνηση του γίνεται $\neg \mathit{Σκότωσε}(w, \mathit{Τούνα})$ σε μορφή CNF. Επαναλαμβάνοντας την απόδειξη της Εικόνας 2.5 με το νέο αρνητικό στόχο βρίσκουμε ένα παρόμοιο δένδρο απόδειξης, αλλά με την αντικατάσταση *{w/Παράξενο}* σε ένα από τα βήματα. Έτσι , σε αυτή την περίπτωση, η εύρεση του δολοφόνου της γάτας είναι απλώς ζήτημα παρακολούθησης των δεσμεύσεων των μεταβλητών ερωτήματος στην απόδειξη.



Εικόνα 2.5 Μια απόδειξη μέσω ανάλυσης ότι το Παράξενο πρόσωπο σκότωσε τη γάτα. Προσέξτε τη χρήση της παραγοντοποίησης για την παραγωγή της πρότασης Αγαπά(G(Τζακ) ,Τζακ)

Δυστυχώς, η ανάλυση μπορεί να δημιουργήσει **μη κατασκευαστικές αποδείξεις** (non-constructive proofs) για τους υπαρξιακούς στόχους. Για παράδειγμα, το $\neg Σκότωσε(w, Τούνα)$ αναλύεται με το $Σκότωσε(Τζακ, Τούνα) \vee Σκότωσε(Παράξενο, Τούνα)$ για να μας δώσει το $Σκότωσε(Τζακ, Τούνα)$, το οποίο αναλύεται ξανά με το $\neg Σκότωσε(w, Τούνα)$ για να μας δώσει την κενή πρόταση. Προσέξτε ότι το *w* έχει δύο διαφορετικές δεσμεύσεις σε αυτή την απόδειξη· η ανάλυση μας λέει ότι, πράγματι, κάποιος σκότωσε την Τούνα – είτε ο Τζακ είτε το Παράξενο πρόσωπο. Αυτό δεν μας προκαλεί έκπληξη! Μια λύση είναι να περιορίσουμε τα επιτρεπόμενα βήματα ανάλυσης έτσι ώστε οι μεταβλητές ερωτήματος να μπορούν να δεσμευτούν μόνο μια φορά σε κάθε δεδομένη απόδειξη· στη συνέχεια θα πρέπει να μπορούμε να υπαναχωρήσουμε στις δυνατές δεσμεύσεις μεταβλητών. Μια άλλη λύση είναι να προσθέσουμε ένα **λεκτικό απάντησης** (answer literal) στον αρνητικό στόχο, ο οποίος γίνεται $\neg Σκότωσε(w, Τούνα) \vee Απάντηση(w)$. Τώρα η διαδικασία ανάλυσης παράγει μια απάντηση όταν παράγεται μια πρόταση που περιέχει ένα μόνο λεκτικό απάντησης. Για την απόδειξη της Εικόνας 2.5, αυτό είναι το $Απάντηση(Παράξενο)$. Η μη κατασκευαστική απόδειξη θα δημιουργούσε την πρόταση $Απάντηση(Παράξενο) \vee Απάντηση(Τζακ)$, που δεν αποτελεί απάντηση.

Χειρισμός της ισότητας

Καμία από τις μεθόδους συμπερασμού που έχουν περιγραφεί μέχρι τώρα σε αυτό το κεφάλαιο δεν χειρίζεται την ισότητα. Υπάρχουν τρεις διαφορετικές προσεγγίσεις που μπορούμε να ακολουθήσουμε. Η πρώτη προσέγγιση είναι να ορίσουμε αξιώματα για την ισότητα — να καταγράψουμε στη βάση γνώσης προτάσεις για τη σχέση ισότητας. Πρέπει να πούμε ότι η ισότητα είναι ανακλαστική, συμμετρική, και μεταβατική, και πρέπει επίσης να πούμε ότι μπορούμε να αντικαταστήσουμε μεταξύ τους τα ίσα στοιχεία σε οποιοδήποτε κατηγόρημα ή συνάρτηση. Έτσι χρειαζόμαστε τρία βασικά αξιώματα, και ακολούθως ένα για κάθε κατηγόρημα και συνάρτηση:

$$\forall x \quad x = x$$
$$\forall x, y \quad x = y \Rightarrow y = x$$
$$\forall x, y, z \quad x = y \land y = z \Rightarrow x = z$$
$$\forall x, y \quad x = y \Rightarrow (P_1(x) \Leftrightarrow P_1(y))$$
$$\forall x, y \quad x = y \Rightarrow (P_2(x) \Leftrightarrow P_2(y))$$

•
•
•

$$\forall w, x, y, z \quad w = y \land x = z \Rightarrow (F_1(w, x) \Leftrightarrow F_1(y, z))$$
$$\forall w, x, y, z \quad w = y \land x = z \Rightarrow (F_2(w, x) \Leftrightarrow F_2(y, z))$$

•
•
•

Με δεδομένες αυτές τις προτάσεις οι συνήθεις διαδικασίες συμπερασμού, όπως η ανάλυση, μπορούν να εκτελέσουν εργασίες που απαιτούν συλλογιστική σχετικά με την ισότητα, όπως είναι η επίλυση μαθηματικών εξισώσεων.

Ένας άλλος τρόπος αντιμετώπισης της ισότητας είναι με έναν πρόσθετο κανόνα συμπερασμού. Ο απλούστερος κανόνας, η **αποδιαμόρφωση** (demodulation), λαμβάνει μια μοναδιαία πρόταση x=y και αντικαθιστά με y κάθε όρο που ενοποιείται με το x σε κάποια άλλη πρόταση.
Σε πιο τυπική μορφή, έχουμε

- **Αποδιαμόρφωση :** Για οποιουσδήποτε όρους x,y, και z, όπου η UNIFY(x,z)=θ και $m_n[z]$ είναι ένα λεκτικό που περιέχει το z :

$$\frac{x = y, \qquad m_1 \lor \ldots \lor m_n[z]}{m_1 \lor \ldots \lor m_n[SUBST(\theta, y)]}.$$

Η αποδιαμόρφωση χρησιμοποιείται συνήθως για την απλοποίηση των παραστάσεων χρησιμοποιώντας συλλογές δηλώσεων όπως οι $x+0=x$, $x^1=x$, και ούτω καθεξής. Ο κανόνας μπορεί επίσης να επεκταθεί έτσι ώστε να χειρίζεται μη μοναδιαίες προτάσεις στις οποίες εμφανίζεται λεκτικό ισότητας:

- **Παραδιαμόρφωση** (paramodulation): Για οποιουσδήποτε όρους x,y, και z, όπου UNIFY(x,z)=θ,

$$\frac{l_1 \vee \ldots \vee l_k \vee x = y, \qquad m_1 \vee \ldots \vee m_n[z]}{SUBST(\theta, l_1 \vee \ldots \vee l_k \vee m_1 \vee \ldots \vee m_n[y])}$$

Αντίθετα με την αποδιαμόρφωση, η παραδιαμόρφωση δίνει μια πλήρη διαδικασία συμπερασμού για τη λογική πρώτης τάξης με ισότητα.

Μια τρίτη προσέγγιση χειρίζεται πλήρως τη συλλογιστική για την ισότητα μέσα στα πλαίσια ενός εκτεταμένου αλγόριθμου ενοποίησης. Αυτό σημαίνει ότι οι όροι είναι ενοποιήσιμοι εάν αποδεικνύεται ότι είναι ίσοι μετά από κάποια αντικατάσταση, όπου το "αποδεικνύεται" επιτρέπει κάποια συλλογιστική σχετικά με την ισότητα. Για παράδειγμα, οι όροι 1 + 2 και 2 + 1 υπο κανονικές συνθήκες δεν είναι ενοποιήσιμοι, αλλά ένας αλγόριθμος ενοποίησης που γνωρίζει ότι x + y = y + x θα μπορούσε να τους ενοποιήσει με την κενή αντικατάσταση. Η **ενοποίηση με εξισώσεις** (equational unification) αυτού του τύπου μπορεί να εκτελεστεί με αποδοτικούς αλγορίθμους που είναι σχεδιασμένοι για τα συγκεκριμένα αξιώματα τα οποία χρησιμοποιούνται (αντιμεταθετικότητα, προσεταιριστικότητα, και ούτω καθεξής), αντί μέσω του ρητού συμπερασμού με αυτά τα αξιώματα.

3. ΑΛΓΕΒΡΑ
ΛΟΓΙΚΕΣ ΑΠΟΔΕΙΞΕΙΣ =ΕΓΚΛΕΙΣΜΟΙ ΙΔΕΩΔΩΝ

Η βασική ιδέα αυτού του κεφαλαίου είναι να χρησιμοποιήσουμε αλγεβρικές μεθόδους για να κάνουμε (προτασιακές) λογικές αποδείξεις. Συγκεκριμένα θα δείξουμε ότι μπορούν προτάσεις του προτασιακού λογισμού να μετασχηματιστούν σε πολυώνυμα με τέτοιο τρόπο ώστε οι τιμές αληθείας των προτάσεων να διατηρούνται. Με αυτήν την νέα αναπαράσταση ως πολυώνυμα , μπορούμε να πραγματοποιήσουμε λογικές αποδείξεις και με τους δύο τρόπους που έχουμε μάθει στον προτασιακό λογισμό, τον Σημασιολογικό και τον Αποδεικτικό τρόπο.

Το γενικό σχήμα αυτού του κεφαλαίου συνοψίζεται στην Εικόνα 3.1

$$\begin{tikzcd}
\Pi \models \phi \arrow[d, leftrightarrow] & \Pi \vdash \phi \arrow[d, leftrightarrow] \\
V(g_1 \dots g_m) \supseteq V(f) \arrow[r, leftrightarrow] & \langle g_1, \dots, g_m \rangle \subseteq \langle f \rangle
\end{tikzcd}$$

Εικόνα 3.1 Συσχετίζοντας την προτασιακή λογική με την άλγεβρα. Έστω $\Pi = \{\gamma_1, \dots, \gamma_n\}$ το σύνολο των προτάσεων και φ μια πρόταση, όπου σε όλες τις περιπτώσεις χρησιμοποιούμε προτασιακές μεταβλητές τις $\chi_1, \dots, \chi_n$. Έστω $g_i = P(\gamma_i)$ και $f = P(\varphi)$ όπου P είναι η συνάρτηση μετασχηματισμού των προτάσεων σε πολυώνυμα με τέτοιο τρόπο που οι τιμές αληθείας διατηρούνται (Δείτε τους παρακάτω ορισμούς). Εδώ το V(α) είναι η πολλαπλότητα (Variety) που συσχετίζεται με το ιδεώδες α και $\langle g_1, \dots, g_m \rangle$ είναι το ιδεώδες που παράγεται από τα $g_1, \dots, g_m$.

3.1 Περί Άλγεβρας

Τα βασικά αντικείμενα της άλγεβρας που θα χρειαστούμε σε αυτό το κεφάλαιο είναι τα πολυώνυμα πολλών μεταβλητών. Έστω ότι δουλεύουμε σε πολυώνυμα με μεταβλητές τις $\chi_1, \ldots, \chi_n$, τότε ένα **μονώνυμο** (monomial) είναι μια έκφραση της μορφής :

$$\chi_1^{\alpha_1} \ldots \chi_n^{\alpha_n} = \chi^{\alpha}, \alpha = [\alpha_1, \ldots, \alpha_n] \in N^n$$

Χρησιμοποιώντας τον σύμβολισμό των αθροισμάτων και του παραπάνω ορισμού των μονωνύμων, μπορούμε εύκολα να γράψουμε ένα πολυώνυμο πολλών μεταβλητών p ως :

$$p = \sum_{\alpha \in N^n} c_a x^a$$

όπου μόνο πεπερασμένα το πλήθος c_a είναι μη μηδενικά. Δεν έχουμε ακόμα ορίσει από πιο σύνολο θα παίρνουμε τους συντελεστές c_a , αλλά σε αυτό το κεφάλαιο θα προέρχονται από ένα σώμα K .

Με την λέξη **σώμα** εννοούμε ένα σύνολο αριθμών (αντικειμένων γενικότερα) K μαζί με δύο δυαδικές πράξεις + και * ορισμένες στο K, οι οποίες απεικονίζουν σε 2 στοιχεία a και b που ανήκουν στο K, άλλα στοιχεία, a+b και a*b, πάλι στο K, και επιπλέον ισχύουν οι ιδιότητες

1) (a+b)+c=a+(b+c)
2) Υπάρχει στοιχείο 0 που ανήκει στο K τέτοιο ώστε
   i) a+0=a για κάθε a που ανήκει στο K, και
   ii) Για κάθε a που ανήκει στο K υπάρχει b που ανήκει στο K τέτοιο ώστε a+b=0.
3) a+b=b+a Δηλαδή να ισχύει η αντιμεταθετική ιδιότητα στο K
4) (a*b)*c=a*(b*c)
5) Υπάρχει αριθμός 1 που ανήκει στο K τέτοιος ώστε

   (i).a*1=a

   (ii). Και να υπάρχει, για κάθε a διάφορο του μηδενός, ένα b, τέτοιο ώστε a*b=1.
6) a*b=b*a
7) a*(b+c)=a*b+a*c

Το σύνολο τέτοιων πολυωνύμων το συμβολίζουμε

K[$\chi_1, \ldots, \chi_n$] = { $\sum_{\alpha \in N^n} c_a x^a$: $c_a \in K$ με πεπερασμένο το πλήθος $c_a \neq 0$ }

Ορισμός 3.1 Έστω $p_1, \ldots, p_l \in K[x_1, \ldots, x_n]$.

Ιδεώδες (Ideal) : Το ιδεώδες α που παράγεται από τα $p_1, \cdots, p_l$ είναι το σύνολο των πολυωνύμων της μορφής

$$\alpha = \{ \sum_{1 \le i \le l} f_i p_i : f_i \in K[x_1, \ldots, x_n]\}$$

Πολλαπλότητα "Μηδενισμού" (Variety) : Η πολλαπλότητα A των $p_1, \cdots, p_l$ είναι το σύνολο :

$$A = \{u \in K^n : p_i(u) = 0, i = 1, \ldots, l\}$$

Συμβολίζουμε το **σύνολο των ιδεωδών** στο K[$\chi_1, \cdots, \chi_n$] με **Ideals**(K[$\chi_1, \cdots, \chi_n$]). Επίσης συμβολίζουμε το ιδεώδες που παράγεται από τα $p_1, \ldots, p_l$ με $\langle p_1, \ldots, p_l \rangle$.
Ίσως η πιο αναπτυγμένη μορφή του ορισμού του Ιδεώδους να μας βοηθήσει να καταλάβουμε καλύτερα την βασική ιδιότητα του.
Το Ιδεώδες α που παράγεται από τα $p_1, \ldots, p_l$ ουσιαστικά είναι τα πολυώνυμα της μορφής
$a = f_1(\underline{x}) p_1(\underline{x}) + \ldots + f_l(\underline{x}) p_l(\underline{x}) \quad \underline{x} = (x_1, \ldots, x_n), \quad \forall f_i \in K[x_1, \ldots, x_n]$
Όπως διαπιστώνεται , οι κοινές λύσεις των $p_1, \ldots, p_l$ είναι και λύσεις αυτών των πολυωνύμων, γιαυτό και μπορούμε να πούμε ότι η πολλαπλότητα που σχετίζεται με τα $p_1, \ldots, p_l$ είναι η ίδια με την πολλαπλότητα που σχετίζεται με το $\langle p_1, \ldots, p_l \rangle$.
Αυτή η πολλαπλότητα συνήθως συμβολίζεται με V($\langle p_1, \ldots, p_l \rangle$).
Με τους παραπάνω συμβολισμούς, μπορούμε να δημιουργήσουμε τους παρακάτω μετασχηματισμούς

$$V : Ideals(K[x_1, \ldots, x_n]) \to S(K^n)$$

όπου $S(K^n)$ είναι το σύνολο όλων των υποσυνόλων του K^n .
Συμμετρικά μπορούμε να ορίσουμε τον μετασχηματισμό

$I : S(K^n) \to Ideals(K[x_1, \ldots, x_n])$ ο οποίος αποδίδει το

$$I(A) = \{p \in K[x_1, \ldots, x_n] : p(u) = 0 \, \forall u \in A\}$$

η εικόνα αυτού του μετασχηματισμού αποδεικνύεται εύκολα ότι είναι ένα ιδεώδες. Το παρακάτω θεώρημα είναι περίπου άμεσες συνέπειες των παραπάνω ορισμών.

Θεώρημα 3.1 Έστω $\alpha_1, \alpha_2 \in Ideals(K[x_1, \ldots, x_n])$ τότε

$$V(\langle 0 \rangle) = K^n \quad (1)$$

$$V(\langle 1 \rangle) = \emptyset \quad (2)$$

$$V(a_1 \cap a_2) = V(a_1) \cup V(a_2) \quad (3)$$

$$V(a_1 + a_2) = V(a_1) \cap V(a_2) \quad (4)$$

$$a_1 \subseteq a_2 \Rightarrow V(a_1) \supseteq V(a_2) \quad (5)$$

Έστω $A_1, A_2 \in S(K^n)$ τότε

$$I(K^n) = \langle 0 \rangle \quad (6)$$

$$I(\emptyset) = K[x_1, \ldots, x_n] \quad (7)$$

$$A_1 \subseteq A_2 \Rightarrow I(A_1) \supseteq I(A_2) \quad (8)$$

και συνδυάζοντας και τους δύο χώρους συνόλων
$(\alpha \in Ideals(K[x_1, \ldots, x_n]), A \in S(K^n))$ έχουμε :

$$I(V(a)) \supseteq a \quad (9)$$

$$V(I(a)) \supseteq A \quad (10)$$

$$I(V(I(A))) = I(A) \quad (11)$$

$$V(I(V(a))) = V(a) \quad (12)$$

Είναι πολύ χρήσιμο να έχουμε υπόψη μας ότι

$$\langle 1 \rangle = K[x_1, \ldots, x_n]$$

Σε αυτό το κεφαλαίο θα περιοριστούμε στο σώμα K = Z_2 , δηλαδή τους ακέραιους modulo 2.
(Μιας και οι τιμές αληθείας των Προτάσεων στον Προτασιακό Λογισμό μπορούν να είναι δύο , αλήθεια ή ψέμα , 1 ή 0)

Θεώρημα 3.2 (Κύριο θεώρημα) Έστω p πρώτος τότε

$$\forall A \in S(Z_p^n) \qquad V(I(A))=A \tag{13}$$

$$\forall a \in Ideals(Z_p[x_1, \dots, x_n]) \qquad I(V(a)) = a+\langle x_1^p - x_1, \dots, x_n^p - x_n \rangle \tag{14}$$.

Η (13) εξίσωση υποδηλώνει <u>ότι κάθε υποσύνολο του</u> Z_2^n μπορεί να περιγραφεί ως Πολλαπλότητα "μηδενισμού" (Variety) κάπου ιδεώδους ή ισοδύναμα <u>ως το σύνολο των κοινών λύσεων ενός συνόλου πολυωνύμων.</u>

Με όμοιο τρόπο, η εξίσωση (14) δείχνει ότι μπορούμε να ανακτήσουμε το αρχικό μας ιδεώδες από την πολλαπλότητα "μηδενισμού" του.

Μιας και έχουμε περιοριστεί στο σώμα K= Z_2 , δηλαδή τους ακέραιους modulo 2 ενδιαφέρον παρουσιάζει το "μικρό θεώρημα" του Fermat

$$x^{(p-1)}=1 \, mod \, p \quad \text{για p πρώτο}$$

το οποίο είναι ισοδύναμο με $x^p - x = 0 \, mod \, p \, \forall x \in Z_p$

στην περίπτωση μας όπου p=2 , έχουμε $x^2 - x = 0 \, mod \, 2 \qquad \forall x \in Z_2 = \{0,1\}$
δηλ. $x^2 = x \, mod \, 2 \qquad \forall x \in Z_2 = \{0,1\}$

Ερώτηση:

Πώς μπορούμε να υπολογίσουμε τα πολυώνυμα που αντιστοιχούν σε ένα δεδομένο υποσύνολο του
Z_2^n (δηλαδή που έχουν ως πολλαπλότητα μηδενισμού (Variety) ένα υποσύνολο του Z_2^n) ;

(Είναι σημαντικό αυτό για να μετασχηματίσουμε λογικές προτασιακές εκφράσεις σε πολυώνυμα, γιατί ο σκοπός είναι οι τιμές αληθείας των λογικών εκφράσεων (υποσύνολα του Z_2^n) να διατηρούνται από τα αντίστοιχα πολυώνυμα)

Απάντηση: πολυώνυμα παρεμβολής Lagrange.

Έστω

$$L_i(x)=\frac{x(x-1)\dots(x-(i-1))(x-(i+1))\dots(x-(p-1))}{(i-0)(i-1)\dots(i-(i-1))(i-(i+1))\dots(i-(p-1))} \in Z_p[x]$$

τότε

$$L_i(x) = \begin{cases} 1, & x = i \\ 0, & x \neq i \end{cases}$$

Από αυτό μπορούμε να ορίσουμε τα

$$L_a(x_1, \ldots, x_n) = L_{a_1}(x_1) \ldots L_{a_1}(x_n), \qquad a \in Z_p^n$$ με την ιδιότητα

$$L_a(x_1, \ldots, x_n) = \begin{cases} 1 & ,[x_1, \ldots, x_n] = a \\ 0 & ,[x_1, \ldots, x_n] \neq a \end{cases}$$

αν υποθέσουμε ότι μας δίνεται μια συνάρτηση $f : Z_p^n \rightarrow Z_p$ τότε μπορούμε να δημιουργήσουμε ένα πολυώνυμο q τέτοιο ώστε $q(a) = f(a) \;\; \forall a \in Z_p^n$. Αυτό το πολυώνυμο είναι το:

$$q = \sum_{a \in Z_p^n} L_a(x_1 \ldots x_n) f(a) \tag{15}$$

3.2 Βάσεις Gröbner

Μερικές φορές το να βρούμε τις κοινές λύσεις ενός συστήματος πολυωνύμων $p_1, \dots, p_l$ δεν είναι και τόσο εύκολη υπόθεση. Είδαμε όμως τις ίδιες λύσεις θα έχει όμως και το Ιδεώδες που παράγεται από τα $p_1, \dots, p_l$. Η δομή του Ιδεώδους όμως έχει κάποιες ενδιαφέρουσες προοπτικές.

Αποδεικνύεται ότι αυτό το σύνολο πολυωνύμων που παράγεται από τα $p_1, \dots, p_l$ (Ιδεώδες) , είναι ένα σύνολο που θα μπορούσε να παραχθεί και από άλλα σύνολα πολυωνύμων (όπως και σε ένα διανυσματικό χώρο μπορούμε να έχουμε διαφορετικούς γεννήτορες που να τον παράγουν) για τα οποία θα μπορούσαμε να βρούμε πιο εύκολα τις κοινές τους λύσεις και κατα συνέπεια τις λύσεις του Ιδεώδους (συνεπώς την Πολλαπλότητα του) άρα και τις λύσεις του αρχικού "γεννήτορα" του Ιδεώδους.

Ο σκοπός αυτής της Ενότητας είναι να παρουσιάσουμε ένα αλγόριθμο υπολογισμού αυτού του άλλου συνόλου γεννητόρων ενός Ιδεώδους για το οποίο ισχύουν τα παραπάνω. Μάλιστα αυτό το σύνολο γεννητόρων ενός Ιδεώδους θα το καλέσουμε **Βάση Gröbner .**

Σημαντική έννοια στην Ενότητα αυτή είναι και η διαίρεση πολυωνύμων πολλών μεταβλητών. Όπως και στην διαίρεση πολυωνύμων μιας μεταβλητής στο Λυκείο , σημαντικοί είναι οι μεγιστοβάθμιοι όροι ενός πολυωνύμου. Αν και διακρίνονται εύκολα σε πολυώνυμα μιας μεταβλητής, στην περίπτωση πολυωνύμων πολλών μεταβλητών , χρειαζόμαστε να ορίσουμε μια διάταξη μονωνύμων. Υπάρχουν πολλά είδη διατάξεων μονωνύμων, εμείς όμως μπορούμε να χρησιμοποιήσουμε την οικεία **λεξικογραφική διάταξη**.

Ο πυρήνας του αλγόριθμου που θα παρουσιάσουμε είναι το **πολυώνυμο S** δύο πολυωνύμων f, g.

$$S(f, g) = \frac{ΕΚΠ(μεγ(f), μεγ(g))}{μεγ(f)} * f - \frac{ΕΚΠ(μεγ(f), μεγ(g))}{μεγ(g)} * g$$

όπου μεγ(f) είναι ο μεγιστοβάθμιος όρος του f , όμοια για το g, και ΕΚΠ είναι το Ελάχιστο Κοινό Πολλαπλάσιο αυτών.

function Grobner (F) **returns** Βάση Gröbner
inputs: F ={το σύνολο των αρχικών γεννητόρων του Ιδεώδους}

G ← F
for $\forall$ *ζεύγος πολυωνύμων* $\in G$ **do**

Υπολόγισε το S-πολυώνυμο του Ζεύγους και
βρες το υπόλοιπο h της διαίρεσης του με το σύνολο G.

if h=0 **then**
συνέχισε με το άλλο ζεύγος
else
πρόσθεσε το h στο σύνολο G και επανάλαβε

return G

Εικόνα 3.2 Αλγόριθμος **Buchberger .**

<u>**Παράδειγμα 3.1**</u>

Έστω το Ιδεώδες $\langle x^3-2xy\,,\ x^2y-2y^2+x\rangle$ με γεννήτορες προφανώς τα πολυώνυμα $f_1=x^3-2xy$ και $f_2=x^2y-2y^2+x$. Άρα το αρχικό μας σύνολο γεννητόρων του Ιδεώδους είναι το G = { f_1 , f_2 } .
Σύμφωνα με την λεξικογραφική διάταξη οι μεγιστοβάθμιοι όροι είναι οι $\mu\varepsilon\gamma(f_1)=x^3$ και $\mu\varepsilon\gamma(f_2)=x^2y$.

Το S-πολυώνυμο των f_1 και f_2 είναι

$$S(f_1,f_2)=\frac{ΕΚΠ(x^3,x^2y)}{x^3}*(x^3-2xy)-\frac{ΕΚΠ(x^3,x^2y)}{x^2y}*(x^2y-2y^2+x)=\ -x^2$$

Αν διαίρεσουμε το $S(f_1,f_2)=-x^2$ με το σύνολο G έχουμε :

$$\begin{array}{l|l} \underline{-x^2} & \underline{-x^3}-2xy \\ \hline & 0 \end{array}$$

$= \ -\chi^2$

$$\begin{array}{l|l} \underline{-x^2} & \underline{x^2y}-2y^2+x \\ \hline & 0 \end{array}$$

$= \ -\chi^2$

Το υπόλοιπο της διαίρεσης του $S(f_1, f_2)$ από το σύνολο των { f_1 , f_2 } είναι $-x^2$ άρα θα το προσθέσουμε στο αρχικό μας σύνολο των πολυωνύμων G, οπότε θα έχουμε ότι το $G=\{f_1, f_2, f_3\}=\{x^3-2\mathrm{xy}, x^2 y-2\mathrm{y}^2+x, -x^2\}$.

Υπολογίζουμε τώρα (παραλείποντας τις λεπτομέρειες) τα

$$S(f_1, f_2)=f_3$$

και διαιρώντας το με το G έχουμε $\overbrace{S(f_1, f_2)}^{G}=0$, εφόσον όμως πρέπει να πάρουμε ανα δύο όλα τα πολυώνυμα του G υπολογίζουμε και το

$$S(f_1, f_3)=(x^3-2\mathrm{xy})-(-x)(-x^2)=-2\mathrm{xy} \quad \alpha\lambda\lambda\acute{\alpha}$$

$$\overbrace{S(f_1, f_3)}^{G}=-2\mathrm{xy}\neq 0 \quad .$$

Συνεπώς αναγκαζόμαστε να προσθέσουμε στο G και το πολυώνυμο $f_4=-2\mathrm{xy}$. Έχουμε δηλαδή τώρα $G=\{f_1, f_2, f_3, f_4\}=\{x^3-2\mathrm{xy}, x^2 y-2\mathrm{y}^2+x, -x^2, -2\mathrm{xy}\}$

αρχίζουμε πάλι να υπολογίζουμε τα S-πολυώνυμα και τα υπόλοιπα τους από την διαίρεση τους από το G.

$$\overbrace{S(f_1, f_2)}^{G}=0 \quad \text{και} \quad \overbrace{S(f_1, f_3)}^{G}=0$$

$$S(f_1, f_4)=y(x^3-2\mathrm{xy})-(-1/2)(x^2)(-2\mathrm{xy})=-2\mathrm{xy}^2=yf_4 \quad \text{όποτε}$$

$$\overbrace{S(f_1, f_4)}^{G}=0 \quad ,$$

$$S(f_2, f_3)=(x^2 y-2\mathrm{y}^2+x)-(-y)(-x^2)=-2\mathrm{y}^2+x \quad \text{όποτε}$$

$\overbrace{S(f_2, f_3)}^{G}=-2\mathrm{y}^2+x\neq 0$. Συνεπώς αναγκαζόμαστε να προσθέσουμε ξανά στο G και το πολυώνυμο $f_5=-2\mathrm{y}^2+x$.

Έχουμε τώρα $G=\{f_1, f_2, f_3, f_4, f_5\}=\{x^3-2\mathrm{xy}, x^2 y-2\mathrm{y}^2+x, -x^2, -2\mathrm{xy}, -2\mathrm{y}^2+x\}$

Αν αποφασίσουμε τώρα να υπολογίσουμε όλα τα S-πολυώνυμα τους και τα υπόλοιπα τους από τις διαιρέσεις με το G, θα διαπιστώσουμε ότι

$\overbrace{S(f_i, f_j)}^{G}=0\ \forall\ 1\leq i\leq j\leq 5$, άρα σύμφωνα με τον Αλγόριθμο **Buchberger** το νέο επιθυμητό σύνολο γεννητόρων του αρχικού μας Ιδεώδους είναι το

$$G=\{f_1, f_2, f_3, f_4, f_5\}=\{x^3-2xy, x^2 y-2y^2+x, -x^2, -2xy, -2y^2+x\}$$

Πολυπλοκότητα

Έχει αποδειχθεί ότι η πολυπλοκότητα του παραπάνω αλγορίθμου είναι "Εκθετική" χρονικά και χωρικά κάτι που καθιστά απαγορευτικό μερικές φορές τη χρήση βάσεων Gröbner. Ωστόσο η πολυπλοκότητα μιας μεθόδου υπολογίζεται με βάση τη χειρότερη περίπτωση εφαρμογής της για την επίλυση προβλήματος. Παρόλα αυτά έχει επισημανθεί ότι σε πολλά ενδιαφέροντα προβλήματα οι βάσεις Gröbner μπορούν να υπολογιστούν σε σχετικά σύντομο χρονικό διάστημα.
Ειδικά όμως στην περίπτωση των πολυωνύμων στο Z_2 , έχει βρεθεί ένας P-SPACE αλγόριθμος ο οποίος κάνοντας χρήση με bit-πράξεις όπως XOR και άλλες τεχνικές, ρίχνει σαφώς κατα πολύ την χωρική πολυπλοκότητα και μιας και η σχέση μεταξύ των βασικών κλάσεων πολυπλοκότητας είναι
$P\subseteq NP\subseteq P-SPACE = NP-SPACE\subseteq EXP-TIME\subseteq EXP-SPACE$ αυτό αφήνει περιθώρια να ευελπιστούμε για καλύτερες χρονικές πολυπλοκότητες μελλοντικά.
(Για περισσότερα δείτε το [3] της βιβλιογραφίας μου).
Στο [4] της βιβλιογραφίας μου μπορείτε να δείτε πως ο Groebner basis algorithm βρίσκει μια Groebner απόδειξη μιας ταυτότητας σε χρόνο πολυωνυμικό ως προς το μέγεθος της μικρότερης δυνατής απόδειξης. Επίσης δείχνει πως θα μπορούσε τελικά ο Groebner αλγόριθμος να αντικαταστήσει την μέθοδο της ανάλυσης στον προτασιακό λογισμό με σημαντικές βελτιώσεις στις πολυπλοκότητες.

3.3 Μετασχηματίζοντας τις λογικές εκφράσεις σε πολυώνυμα

Ορισμός 3.2 Έστω $\Pi[\chi_1, \ldots, \chi_n]$ το σύνολο των προτάσεων του Προτασιακού Λογισμού με προτασιακές μεταβλητές τις $\chi_1, \ldots, \chi_n$.
Μια απεικόνιση $\upsilon\colon\ \Pi[\chi_1, \ldots, \chi_n] \to \{0,1\}$ είναι μια αποτίμηση εάν για
$\varphi, \psi \in \Pi[\chi_1, \ldots, \chi_n]$
έχουμε

$$\upsilon(\varphi \wedge \psi) = min(\upsilon(\varphi), \upsilon(\psi))$$

$$\upsilon(\varphi \vee \psi) = max(\upsilon(\varphi), \upsilon(\psi))$$

$$\upsilon(\varphi \to \psi) = \upsilon(\neg\varphi \vee \psi)$$

$$\upsilon(\varphi \Leftrightarrow \psi) = \upsilon((\varphi \to \psi) \wedge (\psi \to \varphi))$$

$$\upsilon(\neg\varphi) = 1 - \upsilon(\varphi)$$

$$\upsilon(\chi_i) \in \{0,1\}$$

Με βάση τις τιμές αληθείας των λογικών τελεστών της Εικόνας 1.8 και την παρεμβολή Lagrange που αναφέραμε πριν έχουμε τους κάτωθι μετασχηματισμούς :

$$p_{\neg}(x) = 1 + x \qquad (16)$$

$$p_{\wedge}(x, y) = xy \qquad (17)$$

$$p_{\vee}(x, y) = x + y + xy \qquad (18)$$

$$p_{\to}(x, y) = 1 + x + xy \qquad (19)$$

$$p_{\Leftrightarrow}(x, y) = 1 + x + y \qquad (20)$$

Τώρα είμαστε σε θέση με αναδρομικό ορισμό να μετασχηματίσουμε οποιαδήποτε λογική έκφραση σε πολυώνυμο.

Ορισμός 3.3 Η απεικόνιση $P:\ \Pi[\chi_1,\dots,\chi_n]\rightarrow Z_2[\chi_1,\dots,\chi_n]$ ορίζεται από τις :

$$P(x_i) = x_i,\ \ i=1,\dots,n$$

$$P(\varphi\circ\psi) = p_{\circ}(P(\varphi),P(\psi)), \circ\in\{\wedge,\vee,\rightarrow,\Leftrightarrow\}$$

$$P(\neg\varphi) = p_{\neg}(P(\varphi))$$

Το επόμενο θεώρημα δικαιολογεί την προσπάθεια μας να ταυτίσουμε τις προτάσεις του προτασιακού λογισμού με τα πολυώνυμα. <u>"Η αλήθεια διατηρείται".</u>

Θεώρημα 3.3 Έστω $\upsilon:\ \Pi[\chi_1,\dots,\chi_n]\rightarrow\{0,1\}$ μια αποτίμηση τέτοια ώστε $\upsilon(\chi_i)=\alpha_i\ \in\ \{0,1\}$, τότε έχουμε

$$\forall\varphi\in\Pi[x_1,\dots,x_n]\qquad \upsilon(\varphi)=P(\varphi)(\alpha_1,\dots,\alpha_n)$$

Είναι χρήσιμο τώρα να αναφέρθουμε και στον αντίστροφο μετασχηματισμό του P, τον Λ.

$\Lambda:\ Z_2[x_1,\dots,x_n]\rightarrow\Pi[x_1,\dots,x_n]$

$$\Lambda(0) = -$$

$$\Lambda(pq) = \Lambda(p)\wedge\Lambda(q)$$

$$\Lambda(p+q) = ((\Lambda(p)\rightarrow\Lambda(q))\wedge(\Lambda(q)\rightarrow\Lambda(p)))\rightarrow -$$

3.4 Συμπερασμός με την βοήθεια της Άλγεβρας (Σημασιολογικός και Αποδεικτικός)

Ιδέα :
Μιας και είδαμε ότι η αλήθεια διατηρείται από μια λογική έκφραση και στο αντίστοιχο πολυώνυμο της, δηλαδή αν είναι "αληθής" μια λογική έκφραση τότε το αντίστοιχο πολυώνυμο της θα έχει την τιμή 1 για την δεδομένη αποτίμηση στις προτασιακές μεταβλητές της, τότε θα μπορούμε να εκφράσουμε τους αποδεικτικούς μηχανισμούς της προτασιακής λογικής με όρους αλγεβρικούς.
Επίσης δεν ξεχνάμε ότι λόγω της πληρότητας του προτασιακού λογισμού και οι δυο τρόποι συμπερασμού, Σημασιολογικός και Αποδεικτικός είναι ισοδύναμοι.

Σημασιολογικά :

Θεώρημα 3.4 Έστω $\gamma_1, \ldots, \gamma_m, \varphi \in \Pi[x_1, \ldots, x_n]$ τότε

$$\{\gamma_1, \ldots, \gamma_m\} \models \phi \quad \Leftrightarrow \quad V(P(\gamma_1)) \cup \ldots \cup V(P(\gamma_m)) \supseteq V(P(\varphi)) \ .$$

Το παραπάνω θεώρημα έχει μια απλότητα στον συλλογισμό του με την έννοια ότι αν κάθε αποτίμηση που ικανοποιεί απο κοινού τις $\{\gamma_1, \ldots, \gamma_m\}$ (τις κάνει αληθείς ταυτόχρονα) ικανοποιεί και την ϕ τότε αυτό με όρους αλγεβρικούς σημαίνει ότι κάθε αποτίμηση στα πολυώνυμα των $\{\gamma_1, \ldots, \gamma_m\}$ που αποδίδει την τιμή 1 ταυτόχρονα (το αντίστοιχο της αλήθειας στον κόσμο των πολυωνύμων), θα αποδίδει επίσης και την τιμή 1 και στο πολυώνυμο της ϕ .
Όποτε αν βαφτίσουμε ως Πολλαπλότητα της "Αλήθειας" την T(P(γ)) κάθε υποσύνολο του Z_2^n που αποδίδει την τιμή 1 στο πολυώνυμο P(γ), συνολοθεωρητικά θα έχουμε να ισχύει ότι

$$T(P(\gamma_1)) \cap \ldots \cap T(P(\gamma_m)) \subseteq T(P(\varphi))$$

Εφαρμόζοντας όμως τον γνωστό συνολοθεωρητικό κανόνα του De Morgan

$$(T(P(\gamma_1)) \cap \ldots \cap T(P(\gamma_m)))' \supseteq (T(P(\varphi)))'$$

$$(T(P(\gamma_1)))' \cup \ldots \cup (T(P(\gamma_m)))' \supseteq (T(P(\varphi)))'$$

$V(P(\gamma_1)) \cup \ldots \cup V(P(\gamma_m)) \supseteq V(P(\varphi))$ μιας και το συμπλήρωμα του T(P(γ)) στον Z_2^n είναι το V(P(γ)) (γιατί ένα διάνυσμα του Z_2^n ή θα προκαλεί την τιμή 1 στο πολυώνυμο ή την τιμή 0)

Παράδειγμα 3.2 Θέλουμε να δείξουμε

$$\models \neg(x \wedge y) \Leftrightarrow (\neg x \vee \neg y) \quad \text{δηλ.} \quad \varnothing \models \neg(x \wedge y) \Leftrightarrow (\neg x \vee \neg y)$$

Αν το λύναμε με την βοήθεια του προτασιακού λογισμού θα μπορούσαμε να δημιουργήσουμε τον Πίνακα αληθείας του και να παρατηρήσουμε τις τιμές αληθείας του.

x	y	$\neg(x \wedge y) \Leftrightarrow (\neg x \vee \neg y)$
0	0	1
0	1	1
1	0	1
1	1	1

Όποτε θα βλέπαμε ότι η παραπάνω πρόταση είναι μια ταυτότητα.

Αν προσπαθούσαμε να κάνουμε την απόδειξη με την βοήθεια πολυωνύμων, τότε

$$P(\neg(x \wedge y) \Leftrightarrow (\neg x \vee \neg y)) = 3 + x + xy + x(1+y) = 3 + x + xy + x + xy = 3 + 2\text{x} + 2\text{xy} = 1 \; \textit{στο} \; Z_2[x, y]$$

όμως το V(1) = $\varnothing$ (γιατί το σταθερό πολυώνυμο 1 δεν έχει λύσεις μηδενισμού)

άρα τετριμμένα ισχύει $\varnothing \supseteq V(1) = \varnothing$ όποτε εφαρμόζοντας το θεώρημα 3.4 επιβεβαιώνεται ο ισχυρισμός μας.

Αποδεικτικά :

Θεώρημα 3.5 Έστω $\gamma_1, \ldots, \gamma_m, \varphi \in \Pi[x_1, \ldots, x_n]$ τότε

$$\{\gamma_1, \ldots, \gamma_m\} \vdash \phi \quad \Leftrightarrow \quad \langle P(\gamma_1) \rangle \cap \ldots \cap \langle P(\gamma_m) \rangle \subseteq \langle P(\varphi) \rangle$$

όπου $\langle p \rangle$ είναι το ιδεώδες που παράγεται από τα $p \in R_2^n$.

Θεώρημα 3.6

$$\langle P(\gamma_1)\rangle \cap \ldots \cap \langle P(\gamma_m)\rangle \subseteq \langle P(\varphi)\rangle$$

$$\Leftrightarrow$$

$$\text{Υπόλοιπο}(P(\varphi), GrobnerBasis(P(\gamma_1), \ldots, P(\gamma_m), x_1^2 - x_1, \ldots, x_n^2 - x_n)) = 0$$

όπου $x_1, \ldots, x_n$ οι προτασιακές μεταβλητές των προτάσεων.

Πόρισμα 3.6 Η πρόταση φ είναι θεώρημα δηλ.

$$\vdash \phi \quad \Leftrightarrow \quad \langle P(\varphi), x_1^2 - x_1, \ldots, x_n^2 - x_n\rangle = \langle 1\rangle$$

Το παραπάνω πόρισμα έχει επίσης ένα προφανή συλλογισμό. Μια λογική έκφραση είναι θεώρημα όταν αληθεύει για κάθε αποτίμηση, το οποίο στον κόσμο των πολυωνύμων σημαίνει ότι κάθε στοιχείο του Z_2^n αποδίδει την τιμή 1 στο αντίστοιχο πολυώνυμο της φ. Το οποίο με την σειρά του σημαίνει ότι κανένα στοιχείο του Z_2^n δεν μηδενίζει το αντίστοιχο πολυώνυμο της φ. Αυτό σημαίνει ότι η πολλαπλότητα "μηδενισμού" αυτού του πολυωνύμου είναι το κενό.
Μπορεί να είναι δύσκολο να βρούμε την πολλαπλότητα "μηδενισμού" ενός πολυωνύμου, όμως μιας και ξέρουμε ότι αυτή η πολλαπλότητα είναι και πολλαπλότητα μηδενισμού και για το ιδεώδες που παράγεται από αυτό το πολυώνυμο, οι ελπίδες μας προσανατολίζονται στο ιδεώδες.
Εφόσον το Ιδεώδες ενός πολυωνύμου "συμπεριφέρεται" όπως και ένας διανυσματικός χώρος, η προσπάθεια να βρούμε "εναλλακτικές" βάσεις για τις οποίες η εύρεση λύσεων είναι πιο εύκολη είναι και η λύση του προβλήματος.
Γιαυτό και αν βρούμε ότι μια τέτοια βάση είναι και το 1 (βάση Gröbner), τότε εφόσον το πολυώνυμο 1 δεν μηδενίζεται για κανένα στοιχείο του Z_2^n , μπορούμε να υποθέσουμε και το ίδιο και για την αρχική βάση του Ιδεώδους μας, που είναι το αρχικό αντίστοιχο πολυώνυμο της φ. Αυτό σημαίνει ότι κάθε στοιχείο του Z_2^n θα του αποδίδει την τιμή 1. Συνεπώς η φ θα είναι αληθής πάντα, για οποιαδήποτε αποτίμηση.
Οι επιπλέον εξισώσεις $x_1^2 - x_1, \ldots, x_n^2 - x_n$ στην βάση του Ιδεώδους , απλώς περιορίζουν τις λύσεις των πολυωνύμων μας μόνο στον Z_2^n όπως άλλωστε επιδιώκουμε και όχι σε μεγαλύτερους χώρους.

Παράδειγμα 3.3

$$\vdash \quad (\varphi \to (\psi \to \sigma)) \to ((\varphi \wedge \psi) \to \sigma)$$

Υποθέτοντας $P(\varphi) = x_1$, $P(\psi) = x_2$ και $P(\sigma) = x_3$ τότε έχουμε

$P((\varphi \to (\psi \to \sigma)) \to ((\varphi \wedge \psi) \to \sigma)) =$

$$= 2 + x_1 + x_1(1 + x_2 + x_2 x_3) + (1 + x_1 x_2 + x_1 x_2 x_3)(1 + x_1 + x_1(1 + x_2 + x_2 x_3))$$

= 1 αφού δουλεύουμε στον $Z_2[x_1, x_2, x_3]$ (Μην ξεχνάμε πάντα ότι εδώ ισχύει $x^2 = x \bmod 2$ $\forall x \in Z_2 = \{0,1\}$)

άρα είναι θεώρημα , που είναι ήδη γνωστό από τον Προτασιακό λογισμό.

Παράδειγμα 3.4

Δείξτε ότι

$$(\varphi \to \psi) \to (\varphi \to \sigma) \quad \vdash \quad \varphi \to (\psi \to \sigma)$$

Ας θεωρήσουμε θ = $(\varphi \to \psi) \to (\varphi \to \sigma)$ και ω = $\varphi \to (\psi \to \sigma)$

Σύμφωνα με το θεώρημα 3.5 αρκεί να δείξουμε ότι

$\langle P(\theta) \rangle \subseteq \langle P(\omega) \rangle$

Υποθέτοντας $P(\varphi) = x_1$, $P(\psi) = x_2$ και $P(\sigma) = x_3$ τότε έχουμε

$P(\theta) = 1 + (1 + x_1 + x_1 x_2) + (1 + x_1 + x_1 x_2)(1 + x_1 + x_1 x_3) = 1 + x_1 x_2 + x_1 x_2 x_3$ αφού δουλεύουμε στον $Z_2[x_1, x_2, x_3]$ (έχουμε υπόψη ότι $x^2 = x \bmod 2$ $\forall x \in Z_2 = \{0,1\}$)

ομοίως

$$P(\omega) = 1 + x_1 + x_1(1 + x_2 + x_2 x_3) = 1 + x_1 x_2 + x_1 x_2 x_3$$

Στην περίπτωση μας αυτή είναι τετριμμένο ότι $\langle P(\theta) \rangle \subseteq \langle P(\omega) \rangle$

άρα έπεται το συμπέρασμα.

Παράδειγμα 3.5

Ελέγξτε αν

$$\phi \vdash (\varphi \rightarrow \psi)$$

Υποθέτοντας $P(\varphi)=x_1$, $P(\psi)=x_2$

οπότε $P(\varphi \rightarrow \psi)=1+x_1+x_1 x_2$

Το υπόλοιπο της διαίρεσης του $1+x_1+x_1 x_2$ με την Gröbner βάση του $\left\langle P(\varphi), x_1^2-x_1, x_2^2-x_2 \right\rangle$ που είναι η { $x_1, x_2^2-x_2$ } είναι 1 ! διαφορετικό του 0 συνεπώς δεν ισχύει το παραπάνω , κάτι αναμενόμενο από τον προτασιακό λογισμό.

Αν προσπαθούσαμε επίσης να το δείξουμε σημασιολογικά, σύμφωνα με το θεώρημα 3.4, τότε βρίσκοντας τις πολλαπλότητες μηδενισμού της υπόθεσης και του συμπεράσματος, θα είχαμε τα υποσύνολα του Z_2^2 μιας και έχουμε μόνο δύο προτασιακές μεταβλητές

$V(x_1)$ ={ (0,0) , (0,1) } και

$V(1+x_1+x_1 x_2)$ = { (1,0) }

Όπως φαίνεται $V(x_1) \not\supseteq V(1+x_1+x_1 x_2)$ άρα δεν ισχύει το ζητούμενο.

Παράδειγμα 3.6

Έστω ότι θέλαμε να λύσουμε το πρόβλημα στον κόσμο του Wumpus που περιγράφεται στην Εικόνα 1.7, όπου
η σχετική βάση γνώσης είναι

$$KB = R_2 \wedge R_4 = (A_{1,1} \Leftrightarrow (\Gamma_{1,2} \vee \Gamma_{2,1})) \wedge \neg A_{1,1}$$

και θέλουμε να αποδείξουμε ότι $\neg \Gamma_{1,2}$

Αν συμβολίσουμε με $x_1 = P(A_{1,1})$, $x_2 = P(\Gamma_{1,2})$ και $x_3 = P(\Gamma_{2,1})$ τότε οι **υποθέσεις** μας είναι τα πολυώνυμα

$$1 + x_1 + (x_2 + x_3 + x_2 x_3) \quad \textit{και} \quad 1 + x_1$$

ενώ το προς απόδειξη **συμπέρασμα** μας είναι το πολυώνυμο $1 + x_2$.

Η Gröbner βάση του $\langle 1 + x_1 + (x_2 + x_3 + x_2 x_3), 1 + x_1, x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle$ = {1} (με χρήση υπολογιστικού πακέτου)

συνεπώς το *Υπόλοιπο*($1 + x_2, 1$)= 0 , *άρα* όντως αποδεικνύεται ότι $\neg \Gamma_{1,2}$!

Αν προσπαθούσαμε **σημασιολογικά** να κάνουμε την απόδειξη, δηλαδή με την βοήθεια του θεωρήματος 3.4 τότε θα έπρεπε να υπολογίσουμε τις πολλαπλότητες μηδενισμού των πολυωνύμων των υποθέσεων και του συμπεράσματος.

Εφόσον έχουμε 3 προτασιακές μεταβλητές είναι υποσύνολο του Z_2^3 .

$V(1 + x_1 + (x_2 + x_3 + x_2 x_3))$ = { (1,0,0) , (0,1,0), (0,0,1), (0,1,1) }

$V(1 + x_1)$ = { (1,0,0), (1,1,0), (1,0,1), (1,1,1) }

και η πολλαπλότητα μηδενισμού του συμπεράσματος είναι η

$V(1 + x_2)$ = { (0,1,0), (0,1,1) , (1,1,0) , (1,1,1) }

οπότε προφανώς διαπιστώνουμε ότι

$$V(1 + x_1 + (x_2 + x_3 + x_2 x_3)) \cup V(1 + x_1) \supseteq V(1 + x_2)$$

Άρα $(A_{1,1} \Leftrightarrow (\Gamma_{1,2} \vee \Gamma_{2,1}))$, $\neg A_{1,1}$ $\models$ $\neg \Gamma_{1,2}$

οπότε λόγω πληρότητας του προτασιακού λογισμού το ζητούμενο απεδείχθη.

Εδώ μπορεί να σημειωθεί ότι έχουμε γλυτώσει την προσπάθεια μετατροπής των λογικών εκφράσεων σε CNF (κανονική μορφή) όπως χρειάζονταν στην περίπτωση του αλγορίθμου της ανάλυσης.

3.5 **ΑΠΟΔΕΙΚΝΥΟΝΤΑΣ ΑΥΤΟΜΑΤΑ ΓΕΩΜΕΤΡΙΚΑ ΘΕΩΡΗΜΑΤΑ**

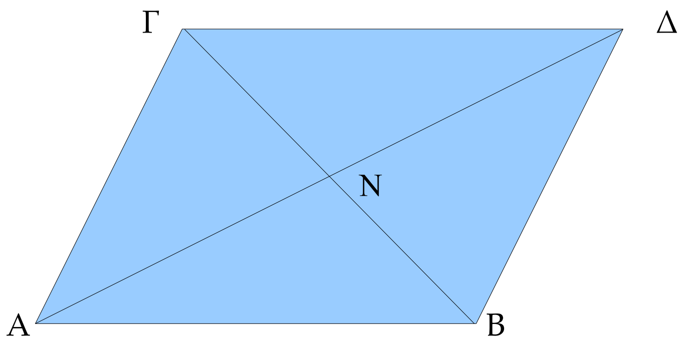
Από την στιγμή που ένας εξελιγμένος πράκτορας μπορεί να έρθει σε επαφή με πιο φυσικά περιβάλλοντα από αυτό του κόσμου του Wumpus , είναι αναμενόμενο οι "οπτικοί" αισθητήρες του να έρθουν σε επαφή με γεωμετρικές κανονικότητες που ενυπάρχουν σε πολλά σχήματα του κόσμου μας. Συνεπώς η αντίληψη που θα έχει εκείνη την στιγμή να μην είναι απλώς μια συλλογή προτάσεων για τον κόσμο την δεδομένη στιγμή , αλλά μια διδιάσταση γεωμετρική εικόνα που οι "οπτικοί" αισθητήρες του θα έχουν αντιληφθεί εκείνη την στιγμή.

Άραγε θα μπορούσε να αποδείξει αυτές τις γεωμετρικές κανονικότητες χωρίς κάποια βάση δεδομένων, απλώς έχοντας τις συντεταγμένες των αντικειμένων ως προς κάποιο σύστημα συντεταγμένων;

Σε αυτή την Ενότητα, θα περιγράψουμε την θεωρία και τον αλγόριθμο μέσα από ένα παράδειγμα.

Ας θεωρήσουμε ένα απλό γεωμετρικό θεώρημα

Έστω Α,Β,Γ και Δ οι κορυφές ενός παραλληλογράμμου στο επίπεδο, όπως στο σχήμα



Είναι γνωστό θεώρημα ότι οι δύο διαγώνιες ΑΔ και ΒΓ οποιουδήποτε παραλληλογράμμου ΑΒΓΔ θα τέμνονται σε σημείο Ν που διχοτομεί και τις δύο διαγωνίους. Πώς όμως θα μπορούσε να αποδειχθεί αυτό αλγεβρικά και αυτόματα ; Για να περάσουμε στην άλγεβρα , θα πρέπει αυτό το σχήμα , συγκεκριμένα οι κορυφές του να αποκτήσουν συντεταγμένες. Συμφέρει το Α να το βάλουμε στην αρχή του συστήματος συντεταγμένων και αυτό για οικονομία μεταβλητών και πράξεων.

Έτσι έχουμε λοιπόν

Α(0,0) , Β(u_1 ,0) Γ(u_2 , u_3) όπου u_2 , u_3 είναι μεταβλητές ανεξάρτητες του u_1 . Οι συντεταγμένες του Δ (x_1 , x_2) μπορούν άνετα να καθορισθούν από τις συντεταγμένες των άλλων κορυφών (αφού το ΑΒΓΔ είναι παραλληλόγραμμο) συνεπώς φαίνεται σωστός ο χαρακτηρισμός ως εξαρτημένων μεταβλητών , ενώ η ελευθερία κινήσεων των u_1 , u_2 , u_3 τις χαρακτηρίζει δικαίως ανεξάρτητες. Ο παραπάνω διαχωρισμός θα παίξει σημαντικό ρόλο στην συνέχεια.
Βέβαια σύμφωνα με τον παραπάνω συμβολισμό, αν θέλουμε να μιλάμε για τετράπλευρο ΑΒΓΔ γενικά, θα πρέπει $u_1 \neq 0$ και $u_3 \neq 0$.

Καιρός να βρούμε τις εξαρτήσεις αυτές και να τις εκφράσουμε αλγεβρικά.

Μιας και είναι παραλληλόγραμμο θα έχουμε

$$\overline{AB} \parallel \overline{\Gamma\Delta} \quad : \quad 0 = \frac{x_2 - u_3}{x_1 - u_2}$$

$$\overline{A\Gamma} \parallel \overline{B\Delta} \quad : \quad \frac{u_3}{u_2} = \frac{x_2}{x_1 - u_1} \quad .$$

Κάνοντας απαλοιφή παρανομαστών , παίρνουμε τις πολυωνυμικές εξισώσεις

$$h_1 = x_2 - u_3 = 0 \quad ,$$

$$h_2 = (x_1 - u_1) u_3 - x_2 u_2 = 0$$

Στη συνέχεια θα πρέπει να εκφράσουμε αλγεβρικά τις συντεταγμένες του σημείου Ν, αλλά και τις αλγεβρικές εξαρτήσεις που φαίνεται να έχει με τα άλλα σημεία του σχήματος.

Το Ν (x_3 , x_4) είναι σημείο τομής των δύο διαγωνίων ΑΔ και ΒΓ, αυτό σημαίνει ότι τα Α, Ν, Δ είναι συνευθειακά, όπως και τα Β, Ν, Γ.

Στην γλώσσα των συντελεστών διευθύνσεων έχουμε λοιπόν:

$$A, N, \Delta \text{ συνευθειακά} \quad : \quad \frac{x_4}{x_3} = \frac{u_3}{x_1} \quad ,$$

$$B, N, \Gamma \text{ συνευθειακά} \quad : \quad \frac{x_4}{x_3 - u_1} = \frac{u_3}{u_2 - u_1} \quad .$$

Απαλείφοντας τους παρανομαστές, έχουμε τις πολυωνυμικές εξισώσεις

$$h_3 = x_4 x_1 - x_3 u_3 = 0$$

$$h_4 = x_4(u_2 - u_1) - (x_3 - u_1) u_3 = 0$$

Άρα μέχρι στιγμής έχουμε καταφέρει να εκφράσουμε αλγεβρικά τις **υποθέσεις** μας. Σειρά έχει τώρα να εκφράσουμε αλγεβρικά και το συμπέρασμα μας!

Το **συμπέρασμα** μας είναι ότι θέλουμε να δείξουμε ότι οι διαγώνιες διχοτομούνται, δηλαδή ότι το Ν είναι το μέσο και της ΑΔ και της ΒΓ.

Από τον Ευκλείδιο τύπο των αποστάσεων σημείων έχουμε:

$$AN = N\Delta \quad \rightarrow \quad x_3^2 + x_4^2 = (x_3 - x_1)^2 + (x_4 - x_2)^2$$

$$BN = N\Gamma \quad \rightarrow \quad (x_3 - u_1)^2 + x_4^2 = (x_3 - u_2)^2 + (x_4 - u_3)^2$$

Εκτελώντας κάποιες αλγεβρικές πράξεις καταλήγουμε στις εξείς πολυωνυμικές εξισώσεις

$$g_1 = x_1^2 - 2\mathrm{x}_1 x_3 - 2\mathrm{x}_4 x_2 + x_2^2 = 0$$

$$g_2 = 2\mathrm{x}_3 u_1 - 2\mathrm{x}_3 u_2 - 2\mathrm{x}_4 u_3 - u_1^2 + u_2^2 + u_3^2 = 0$$

Πρόταση 3.7 : Το πολυώνυμο "συμπέρασμα" G προκύπτει γενικά από τα πολυώνυμα "υποθέσεις" h_1, h_2, h_3 *και* h_4 *εάν*
εάν η $\{1\}$ *είναι η Grobner βάση του ιδεώδους που παράγεται από τα* $\{h_1, h_2, h_3, h_4, 1 - pG\}$
όπου p είναι μεταβλητή.

Ξέρουμε ήδη από την γεωμετρία ότι όντως οι διαγώνιες ενός παραλληλογράμμου διχοτομούνται, οπότε περιμένουμε αν εφαρμόσουμε την παραπάνω πρόταση στις πολυωνυμικές μορφές των υποθέσεων και των συμπερασμάτων μας κάνοντας χρήση κάποιου υπολογιστικού πακέτου όπως το Mathematica ή πολλές άλλες java applets που υπάρχουν στο διαδίκτυο , πρέπει να βρούμε {1} ως βάση Gröbner .

Αν όμως στην περίπτωση μας θελήσουμε να δούμε αν το πολυώνυμο g_1 όντως προκύπτει από τις υποθέσεις h_1, h_2, h_3 *και* h_4 , θα βρούμε βάση Gröbner για το $\{h_1, h_2, h_3, h_4, 1-pg_1\}$
την

GROEBNER BASIS =
{ u3*u1 ,
x4*p*u2^2 +x4*p*u3^2 -1/2*p*u3*u2^2 -1/2*p*u3^3 +1/2*u3 ,
x4*u1 ,
x3*u3 -x4*u2 ,
x4*x3*p*u2 +x4^2*p*u3 +1/2*x4 -1/4*p*u3*u2^2 -1/4*p*u3^3 +1/4*u3 ,
x2 -u3 ,
x1*u3 -u3*u2 ,
x4*x1 -x4*u2 ,
x1^2*p -2*x3*x1*p -2*x4*p*u3 +p*u3^2 -1 }

Σίγουρα αυτή η βάση δεν μοιάζει με {1}!!

Τι κάναμε λάθος;

Το πρόβλημα είναι με απλά λόγια σε ποιο σύνολο (δακτύλιο) πολυωνύμων δουλεύουμε.

Κανονικά, νομίσαμε ότι τα πολυώνυμα μας βρίσκονται στον δακτύλιο

$R[x_1, x_2, x_3, x_4, u_1, u_2, u_3, p]$

όμως αν παρατηρήσουμε τον αρχικό μας συμβολισμό παρατηρούμε ότι αν δεν έχουμε $u_1 \neq 0$ και $u_3 \neq 0$ τότε προκύπτουν κάποιες ενδιαφέρουσες παρενέργειες.
Συγκεκριμένα αν $u_1 = 0$ τότε υπάρχει ταύτιση του Β και του Α! Δεν έχουμε καν τετράπλευρο τότε.
Επίσης αν $u_3 = 0$ τότε το Γ με το Α και το Β είναι συνευθειακά ! Πάλι δεν έχουμε τετράπλευρο.

Αυτές οι ιδιάζουσες περιπτώσεις που μπορούν να πάρουν οι “ανεξάρτητες” μεταβλητές μας που αναφερθήκαμε παραπάνω πρέπει να ληφθούν υπόψη.

Αυτό το πετυχαίνουμε με το να δουλέψουμε στον δακτύλιο

$R(u_1, u_3)[x_1, x_2, x_3, x_4, u_2, p]$. Πιο συγκεκριμένα με το να θεωρήσουμε την δυνατότητα να έχουμε και ρητές εκφράσεις των u_1 και u_3 , όποτε τότε δεν θα μπορούν να μηδενιστούν.
Αυτό το πετυχαίνουμε με το θεωρήσουμε αυτές τις ανεξάρτητες μεταβλητές ώς παραμέτρους στον υπολογισμό της βάσης Gröbner . Όποτε δεν θα συμμετέχουν στην λεξικογραφική διάταξη τους ως μεταβλητές.

Τότε η βάση Gröbner για το $\{h_1, h_2, h_3, h_4, 1-pg_1\}$ με λεξικογραφική διάταξη των μεταβλητών $x_1, x_2, x_3, x_4, u_2, p$ και όχι των $x_1, x_2, x_3, x_4, u_1, u_2, u_3, p$, θα είναι:

GROEBNER BASIS =
{ 1 }

Επιβεβαίωντας έτσι την Πρόταση 3.7.

Όμοιως προκύπτει το ίδιο και για το συμπέρασμα g_2 .

4 ΣΥΜΠΕΡΑΣΜΑΤΑ

4.1 ΣΥΝΟΨΗ

Παρουσιάσαμε τους βασισμένους στη γνώση πράκτορες και δείξαμε πώς μπορεί να οριστεί μια λογική με την οποία αυτοί οι πράκτορες μπορούν να συλλογίζονται για τον κόσμο. Τα κύρια σημεία είναι τα παρακάτω:

- Οι ευφυείς πράκτορες χρειάζονται γνώση για τον κόσμο για να καταλήγουν σε καλές αποφάσεις.
- Η γνώση περιέχεται στους πράκτορες με τη μορφή **προτάσεων** (sentences) σε μια **γλώσσα αναπαράστασης γνώσης**, οι οποίες αποθηκεύονται σε μια **βάση γνώσης** (knowledge base).
- Ένας βασισμένος στη γνώση πράκτορας αποτελείται από μια βάση γνώσης και από ένα μηχανισμό συμπερασμού. Λειτουργεί αποθηκεύοντας προτάσεις για τον κόσμο στη βάση της γνώσης του, χρησιμοποιώντας το μηχανισμό συμπερασμού για να συνάγει νέες προτάσεις και χρησιμοποιώντας αυτές τις προτάσεις για να αποφασίζει τι ενέργειες θα κάνει.
- Μια γλώσσα αναπαράστασης ορίζεται από τη **σύνταξή** της, η οποία καθορίζει τη δομή των προτάσεων, και από τη **σημασιολογία** της, η οποία ορίζει την **αλήθεια** της καθε πρότασης στον κάθε **δυνατό κόσμο** ή **μοντέλο**.
- Η σχέση **λογικής κάλυψης** (entailment) μεταξύ προτάσεων είναι καθοριστική για την κατανόησή μας για τη συλλογιστική. Μια πρόταση α καλύπτει λογικά μια άλλη πρόταση β αν η β είναι αληθής σε όλους τους κόσμους όπου η α είναι αληθής. Υπάρχουν ισοδύναμοι ορισμοί που βασίζονται στην **εγκυρότητα** (validity) της πρότασης $\alpha \Rightarrow \beta$ και στη **μη ικανοποιησιμότητα** (unsatisfiability) της πρότασης $\alpha \wedge \neg \beta$.
- Συμπερασμός είναι η διαδικασία με την οποία προκύπτουν νέες προτάσεις από τις παλιές. Οι **ορθοί** (sound) αλγόριθμοι συμπερασμού συνάγουν μόνο προτάσεις που καλύπτονται λογικά· οι **πλήρεις** (complete) αλγόριθμοι συνάγουν όλες τις προτάσεις που καλύπτονται λογικά.
- Η **προτασιακή λογική** (propositional logic) είναι μια πολύ απλή γλώσσα που αποτελείται από **προτασιακά σύμβολα** και **λογικά συνδετικά**. Μπορεί να χειρίζεται ατομικές προτάσεις που είναι γνωστές αληθείς, γνωστές ψευδείς, ή εντελώς άγνωστες.
- Το σύνολο των δυνατών μοντέλων, με δεδομένο ένα σταθερό προτασιακό λεξιλόγιο, είναι πεπερασμένο, επομένως η λογική κάλυψη μπορεί να ελέγχεται με απαρίθμηση μοντέλων. Υπάρχουν αποδοτικοί αλγόριθμοι συμπερασμού με **έλεγχο μοντέλων** για την προτασιακή λογική και μπορούν συχνά να επιλύουν μεγάλα προβλήματα πολύ γρήγορα.
- Οι **κανόνες συμπερασμού** (inference rules) είναι πρότυπα ορθού συμπερασμού

που μπορούν να χρησιμοποιούνται για την εύρεση αποδείξεων. Ο κανόνας της **ανάλυσης** (resolution) μας δίνει έναν πλήρη αλγόριθμο συμπερασμού για βάσεις γνώσης εκφρασμένες σε **συζευκτική κανονική μορφή** (conjunctive normal form). Το πρόβλημα γενικά του συμπερασμού είναι **NP-complete** ωστόσο για ειδικές μορφές βάσεων γνώσης όπως μορφής **Horn** υπάρχουν αλγόριθμοι συλλογιστικής που λειτουργούν ακόμα και σε χρόνο **γραμμικό** ως προς το μέγεθος της βάσης γνώσης.

- Η προτασιακή λογική είναι σχετικά αποτελεσματική για κάποιες εργασίες μέσα σε έναν πράκτορα, αλλά δεν κλιμακώνεται σε περιβάλλοντα απεριόριστου μεγέθους επειδή δεν έχει την απαιτούμενη εκφραστική ισχύ για να χειρίζεται συνοπτικά το χρόνο, το χώρο και τα γενικά πρότυπα των σχέσεων μεταξύ αντικειμένων.
- Ενώ η προτασιακή λογική δεσμεύεται μόνο για την ύπαρξη γεγονότων, η **λογική πρώτης τάξης** δεσμεύεται για την ύπαρξη αντικειμένων και σχέσεων και έτσι υπερέχει σε εκφραστική ισχύ.
- Ένας **δυνατός κόσμος** ή **μοντέλο** στη λογική πρώτης τάξης ορίζεται από ένα σύνολο αντικειμένων, από τις σχέσεις μεταξύ τους και από τις συναρτήσεις που μπορούν να εφαρμόζονται σε αυτά.
- Τα **σύμβολα σταθερών** ονομάζουν αντικείμενα, τα **σύμβολα κατηγορημάτων** ονομάζουν σχέσεις και τα **σύμβολα συναρτήσεων** ονομάζουν συναρτήσεις. Μια ερμηνεία καθορίζει μια απεικόνιση από τα σύμβολα στο μοντέλο. Οι **σύνθετοι όροι** εφαρμόζουν σύμβολα συναρτήσεων σε όρους για να ονομάσουν ένα αντικείμενο. Με δεδομένη μια ερμηνεία και ένα μοντέλο, προσδιορίζεται η αλήθεια μιας πρότασης.
- Μια **ατομική πρόταση** αποτελείται από ένα κατηγόρημα που εφαρμόζεται σε έναν ή περισσότερους όρους· είναι αληθής μόνο όταν η σχέση που κατονομάζεται από το κατηγόρημα ισχύει μεταξύ των αντικειμένων που κατονομάζονται από τους όρους. Οι **σύνθετες προτάσεις** χρησιμοποιούν συνδετικά όπως ακριβώς και στην προτασιακή λογική, και οι **ποσοτικοποιημένες προτάσεις** επιτρέπουν να εκφράζονται γενικοί κανόνες.
- Η ανάπτυξη μιας βάσης γνώσης στη λογική πρώτης τάξης απαιτεί μια προσεκτική διαδικασία ανάλυσης του πεδίου, επιλογής λεξιλογίου και κωδικοποίησης των αξιωμάτων που απαιτούνται για την υποστήριξη των επιθυμητών συμπερασμών.
- Ως προς τον συμπερασμό στη λογική πρώτης τάξης, μια πρώτη προσέγγιση χρησιμοποιεί κανόνες συμπερασμού για τον προσδιορισμό των ποσοδεικτών, έτσι ώστε να δώσουμε προτασιακή μορφή στο πρόβλημα συμπερασμού. Συνήθως αυτή η διαδικασία είναι πολύ αργή.
- Η χρήση της **ενοποίησης** για την εύρεση κατάλληλων αντικαταστάσεων για τις μεταβλητές απαλείφει το βήμα προσδιορισμού στιγμιοτύπων στις αποδείξεις πρώτης τάξης, καθιστώντας τη διαδικασία πολύ πιο αποδοτική.
- Μια ανυψωμένη εκδοχή του **Τρόπου του Θέτειν** χρησιμοποιεί την ενοποίηση για να παράσχει ένα φυσικό και ισχυρό κανόνα συμπερασμού, το **γενικευμένο Τρόπο του Θέτειν**. Ο Γενικευμένος Τρόπος του Θέτειν είναι πλήρης για τις

οριστικές προτάσεις, αν και το πρόβλημα λογικής κάλυψης είναι **ημιαποφασίσιμο**. Για τα προγράμματα **Datalog**, που αποτελούνται από οριστικές προτάσεις χωρίς συναρτήσεις, η λογική κάλυψη είναι αποφασίσιμη.

- Ο γενικευμένος κανόνας συμπερασμού για την **ανάλυση** παρέχει ένα πλήρες σύστημα απόδειξης για τη λογική πρώτης τάξης, χρησιμοποιώντας τις βάσεις γνώσης σε συζευκτική κανονική μορφή.
- Υπάρχουν διάφορες στρατηγικές για την ελάττωση του χώρου αναζήτησης ενός συστήματος ανάλυσης χωρίς να ζημιώνεται η πληρότητα.
- Η χρήση της άλγεβρας ως αποδεικτικό εργαλείο για τον προτασιακό λογισμό με την χρήση των **βάσεων Gröbner**, υπόσχεται πολλές διευκολύνσεις ειδικά από την στιγμή που η εκθετική χωρική πολυπλοκότητα του αλγορίθμου υπολογισμού αυτών των βάσεων για τον **χώρο των πολυωνύμων στο** Z_2 , έχει αντικατασταθεί από έναν P-SPACE αλγόριθμο.
- Οι βάσεις Gröbner μπορούν να γίνουν πολύ χρήσιμο εργαλείο για την απόδειξη **γεωμετρικών θεωρημάτων**.

4.2 <u>**ΤΟ ΕΡΩΤΗΜΑ**</u>

Όπως αναφέραμε στο Κεφάλαιο 0, φιλόσοφοι υπάρχουν εδώ και πολύ περισσότερο καιρό από τους υπολογιστές και έχουν προσπαθήσει να επιλύσουν μερικά ζητήματα τα οποία σχετίζονται με την ΤΝ: Πώς δουλεύει το μυαλό; Είναι δυνατό για τις μηχανές να ενεργήσουν έξυπνα με τον τρόπο που ενεργούν οι άνθρωποι, και εάν ναι θα έχουν νόηση; Ποιες είναι οι ηθικές επιπτώσεις των ευφυών μηχανών;

Καταρχήν χρειαζόμαστε κάποια ορολογία: ο ισχυρισμός ότι οι μηχανές μπορούν να ενεργήσουν ευφυώς (ή, ακόμα καλύτερα, να ενεργήσουν *σαν* να ήταν ευφυείς) ονομάζεται από τους φιλοσόφους υπόθεση της **ασθενούς ΤΝ** (weak AI), ενώ ο ισχυρισμός ότι οι μηχανές που ενεργούν έτσι είναι *πραγματικά* σκεπτόμενες (σε αντίθεση με την *προσομοίωση* της σκέψης) ονομάζεται υπόθεση της **ισχυρής ΤΝ** (strong AI).

Οι περισσότεροι ερευνητές ΤΝ θεωρούν δεδομένη την ασθενή ΤΝ και δεν ενδιαφέρονται για την υπόθεση της ισχυρής ΤΝ — εφόσον δουλεύει το πρόγραμμα τους, δεν ενδιαφέρονται αν το αποκαλεί ο κόσμος προσομοίωση ευφυίας ή πραγματική ευφυία.

Ερώτημα:

Ασθενής ΤΝ: Μπορούν να ενεργήσουν ευφυώς οι μηχανές ;

Η μαθηματική ένσταση

Είναι γνωστό, από την εργασία του Turing (1936) και του Gödel (1931), ότι συγκεκριμένα μαθηματικά ερωτήματα είναι κατ'αρχήν αδύνατο να απαντηθούν από ειδικά τυπικά συστήματα. Το θεώρημα μη πληρότητας του Gödel είναι το πιο γνωστό παράδειγμα. Εν συντομία, για οποιοδήποτε τυπικό αξιωματικό σύστημα F το οποίο είναι αρκετά ισχυρό ώστε να μπορεί να κάνει αριθμητικές πράξεις, μπορούμε να κατασκευάσουμε μια "πρόταση Gödel " G(F) με τις παρακάτω ιδιότητες:

- Η G(F) είναι μια πρόταση του F, αλλά δεν μπορεί να αποδειχθεί μέσα στο F.
- Εάν το F είναι συνεπές, τότε η G(F) είναι αληθής.

Οι φιλόσοφοι, όπως ο J.R. Lucas (1961), υποστήριξαν ότι αυτό το θεώρημα αποδεικνύει πως οι μηχανές είναι νοητικά κατώτερες από τους ανθρώπους, επειδή οι μηχανές είναι τυπικά συστήματα που περιορίζονται από το θεώρημα της μη πληρότητας — δεν μπορούν να διαπιστώσουν την αλήθεια της δικιάς τους πρότασης Gödel — ενώ οι άνθρωποι δεν έχουν τέτοιο περιορισμό. Αυτός ο ισχυρισμός προκάλεσε δεκαετίες διαφωνιών, και δημιούργησε μια τεράστια βιβλιογραφία η οποία περιλαμβάνει και δύο βιβλία από το μαθηματικό Sir Roger Penrose (1989, 1994) που επαναλαμβάνουν τον ισχυρισμό με κάποιες νέες παραλλαγές (όπως η υπόθεση ότι οι άνθρωποι διαφέρουν επειδή ο εγκέφαλος τους λειτουργεί με κβαντική βαρύτητα). Θα εξετάσουμε μόνο τρία από τα προβλήματα αυτού του ισχυρισμού.

Πρώτον, το θεώρημα μη πληρότητας του Gödel ισχύει μόνο για τα τυπικά συστήματα που είναι αρκετά ισχυρά ώστε να εκτελούν αριθμητικές πράξεις. Σε αυτά περιλαμβάνονται οι μηχανές Turing, και ο ισχυρισμός του Lucas βασίζεται εν μέρει στην υπόθεση ότι οι υπολογιστές είναι μηχανές Turing. Αυτό είναι μια καλή προσέγγιση, αλλά δεν είναι τελείως αληθές. Οι μηχανές Turing είναι άπειρες ενώ οι υπολογιστές είναι πεπερασμένοι, και έτσι οποιοσδήποτε υπολογιστής μπορεί να περιγραφεί ως ένα (πολύ μεγάλο) σύστημα σε προτασιακή λογική, το οποίο δεν υπόκειται στο θεώρημα μη πληρότητας του Gödel.

Δεύτερον, ένας πράκτορας δεν θα πρέπει να ντρέπεται αν δεν μπορεί να προσδιορίσει την αλήθεια μιας πρότασης, ενώ άλλοι πράκτορες μπορούν. Έστω η πρόταση

Ο J.R. Lucas δεν μπορεί να επιβεβαιώσει με συνέπεια ότι αυτή η πρόταση είναι αληθής.

Εάν ο Lucas επιβεβαιώσει αυτή την πρόταση θα αντιφάσκει με τον εαυτό του, έτσι ο Lucas δεν μπορεί να την επιβεβαιώσει με συνέπεια και έτσι η πρόταση πρέπει να είναι αληθής. (Η πρόταση δεν μπορεί να είναι ψευδής, επειδή εάν ήταν τότε ο Lucas δεν θα μπορούσε να την επιβεβαιώσει με συνέπεια και έτσι θα ήταν αληθής.) Αποδείξαμε λοιπόν ότι υπάρχει μια πρόταση την οποία δεν μπορεί να επιβεβαιώσει με συνέπεια ο Lucas ενώ άλλοι άνθρωποι (και μηχανές) μπορούν. Αυτό όμως δεν μας κάνει να σκεφτόμαστε αρνητικά για τον Lucas. Για να εξετάσουμε ένα άλλο παράδειγμα, κανένας άνθρωπος δεν θα μπορούσε να υπολογίσει το άθροισμα 10 δισεκατομμυρίων 10ψήφιων αριθμών σε όλη τη ζωή του, ενώ ένας υπολογιστής θα μπορούσε να το κάνει σε δευτερόλεπτα. Και πάλι, αυτό δεν το θεωρούμε ως θεμελιώδη περιορισμό ως προς τη δυνατότητα του ανθρώπου να σκέπτεται. Οι άνθρωποι συμπεριφέρονταν ευφυώς για χιλιάδες χρόνια πριν την ανακάλυψη των μαθηματικών, έτσι είναι απίθανο να έχει η μαθηματική συλλογιστική ρόλο μεγαλύτερο του περιφερειακού σε αυτό που θεωρείται ευφυϊα.

Τρίτον, και σημαντικότερο, ακόμα και εάν δεχθούμε ότι οι υπολογιστές έχουν περιορισμούς σε αυτά που μπορούν να αποδείξουν, δεν υπάρχει απόδειξη ότι οι άνθρωποι είναι άτρωτοι σε αυτούς τους περιορισμούς. Είναι πολύ εύκολο να αποδείξουμε ότι ένα τυπικό σύστημα δεν μπορεί να κάνει το Χ και στη συνέχεια να ισχυριστούμε ότι οι άνθρωποι *μπορούν να το κάνουν* χρησιμοποιώντας της δική τους άτυπη μέθοδο, χωρίς να δώσουμε καμία απόδειξη για αυτόν τον ισχυρισμό. Πράγματι, είναι αδύνατο να αποδείξουμε ότι οι άνθρωποι δεν υπόκεινται στο θεώρημα μη πληρότητας του Gödel, επειδή κάθε ακριβής απόδειξη θα περιείχε η ίδια μια τυποποίηση του οριζομένου ως μη τυποποιήσιμου ανθρώπινου ταλέντου, και συνεπώς θα αναιρούσε τον εαυτό της. Έτσι αυτό που μας μένει είναι να προσφύγουμε στη διαίσθηση ότι οι άνθρωποι μπορούν με κάποιον τρόπο να κάνουν υπεράνθρωπους άθλους μαθηματικής διορατικότητας. Αυτή η προσφυγή εκφράζεται με επιχειρήματα όπως "οφείλουμε να υποθέσουμε τη συνέπειά μας, εάν είναι δυνατή η σκέψη" (Lucas, 1976). Αν μη τι άλλο, όμως, οι άνθρωποι είναι γνωστό ότι δεν είναι συνεπείς. Αυτό είναι γνωστό από την καθημερινή μας εμπειρία, ισχύει όμως και για προσεκτική μαθηματική σκέψη. Ένα διάσημο παράδειγμα, είναι το πρόβλημα του τετράχρωμου χάρτη. Ο Alfred Kempe δημοσίευσε μια απόδειξη το 1879, η οποία έγινε ευρέως αποδεκτή και συνεισέφερε στην εκλογή του ως μέλους της Βασιλικής Εταιρίας. Το 1890, όμως, ο Percy Heawood εντόπισε ένα σφάλμα και το θεώρημα παρέμεινε αναπόδεικτο μέχρι το 1977.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Artificial Intelligence A Modern Approach Second Edition Stuart J. Russell and Peter Norvig.

[2] Ideals, Varieties and Algorithms, David Cox, John Little, Donal O'Shea

[3] A P-SPACE Algorithm for Groebner Bases Computation in Boolean Rings, Quoc-Nam Tran 2008

[4] Clegg, M., Edmonds, J., and Impagliazzo, R. "Using the Groebner basis algorithm to find proofs of unsatisfiability," In Procceedings of the ACM Symposium on Theory of Computing, 1996.

[5] R. Germundsson. Basic results on ideals and varieties in finite fields. Report LiTH-ISY-I-1259, Dept of EE, 1991

[6] R. Germundsson. Logic Proofs ≡ Ideal Inclusions, 1991

[7] Κ.Ι.Δημητρακόπουλος ΣΗΜΕΙΩΣΕΙΣ ΜΑΘΗΜΑΤΙΚΗΣ ΛΟΓΙΚΗΣ 1999