Πράκτορες Βασισμένοι στη Γνώση

- Βασικές έννοιες
- Γλώσσες αναπαράστασης γνώσης βασισμένες στη Λογική
- Προτασιακή λογική

Πράκτορες Βασισμένοι στη Γνώση

Ο σχεδιασμός ενός **πράκτορα βασισμένου στη γνώση (knowledge-based agent)** βασίζεται στις εξής υποθέσεις:

- Η έννοια της **γνώσης** είναι πρωταρχική.
- Είναι απαραίτητο να έχουμε **ρητή αναπαράσταση της γνώσης** που απαιτείται για την επίλυση ενός προβλήματος εσωτερικά στον πράκτορα.
- Η επιλογή των ενέργειων του πράκτορα γίνεται με τη βοήθεια μιας διαδικασίας **συμπερασμού (inference) ή συλλογιστικής (reasoning)** που εφαρμόζεται στην αναπαράσταση της γνώσης που υπάρχει εσωτερικά στον πράκτορα.

Βασικές Έννοιες

- Η **γλώσσα αναπαράστασης γνώσης (knowledge representation language):** είναι η γλώσσα στην οποία εκφράζεται η γνώση σχετική με τον κόσμο του πράκτορα.
- Η **βάση γνώσης (knowledge base):** είναι ένα σύνολο προτάσεων της γλώσσας αναπαράστασης γνώσης που παριστάνουν τη γνώση του πράκτορα για τον κόσμο του.

  Η βάση γνώσης είναι η εσωτερική αναπαράσταση της γνώσης που υπάρχει στον πράκτορα.

Βασικές Έννοιες

- Ο **μηχανισμός συμπερασμού ή συλλογιστικής:** είναι ένας μηχανισμός ο οποίος καθορίζει τι **έπεται λογικά** από τη γνώση στη βάση γνώσεων.

- Η διεπαφή **ενημερώσεων TELL** και **ερωτήσεων ASK**: η διεπαφή αυτή περιλαμβάνει λειτουργίες για εισαγωγή νέων προτάσεων στη βάση γνώσης και την διατύπωση ερωτήσεων σε ότι είναι ήδη γνωστό.

  Αυτή η λειτουργία είναι παρόμοια με την υποβολή ενημερώσεων και ερωτήσεων σε μια βάση δεδομένων.

  Η λειτουργία ASK χρησιμοποιεί το μηχανισμό συμπερασμού για να βρεί την απάντηση σε μια ερώτηση.

Γενικός Πράκτορας Βασισμένος στη Γνώση

function KB-AGENT(*percept*) **returns** an *action*
  static KB, βάση γνώσης
    t, μετρητής για το χρόνο, αρχικά 0

  TELL(KB,MAKE-PERCEPT-SENTENCE($percept, t$))
  $action \leftarrow$ ASK(KB,MAKE-ACTION-QUERY(t))
  TELL(KB,MAKE-ACTION-SENTENCE($action, t$))
  $t \leftarrow t + 1$
  return *action*

Πράκτορες Βασισμένοι στη Γνώση

Ένας πράκτορας βασισμένος σε γνώση μπορεί να περιγραφεί σε τρία επίπεδα:

- **Επίπεδο γνώσης (knowledge level):** Σε αυτό το επίπεδο ο πράκτορας προσδιορίζεται λέγοντας τι γνωρίζει για τον κόσμο και ποιοί είναι οι στόχοι του.
- **Λογικό επίπεδο (logical level):** Αυτό είναι το επίπεδο στο οποίο η γνώση και οι στόχοι του πράκτορα κωδικοποιούνται σε προτάσεις κάποιας λογικής γλώσσας.
- **Επίπεδο υλοποίησης (implementation level):** Αυτό είναι το επίπεδο στο οποίο οι προτάσεις υλοποιούνται από ένα πρόγραμμα που τρέχει πάνω στην αρχιτεκτονική του πράκτορα.

Σημείωση: Αντιπαραβάλετε τη **δηλωτική (declarative)** με τη **διαδικαστική (procedural)** προσέγγιση στην υλοποίηση του συστήματος.

Παράδειγμα: Αυτόματος Οδηγός Ταξί

- **Επίπεδο γνώσης:**

  Ο αυτόματος οδηγός ταξί γνωρίζει ότι η γέφυρα του Golden Gate συνδέει το San Francisco με το Marin County.

- **Λογικό επίπεδο:**

  Ο αυτόματος οδηγός ταξί έχει την πρόταση της **λογική πρώτης** τάξης *Links(GGBridge,SF,Marin)* στη βάση γνώσης του.

- **Επίπεδο υλοποίησης:**

  Η πρόταση *Links(GGBridge,SF,Marin)* υλοποιείται από μια δομή της C ή ένα γεγονός (fact) της Prolog.

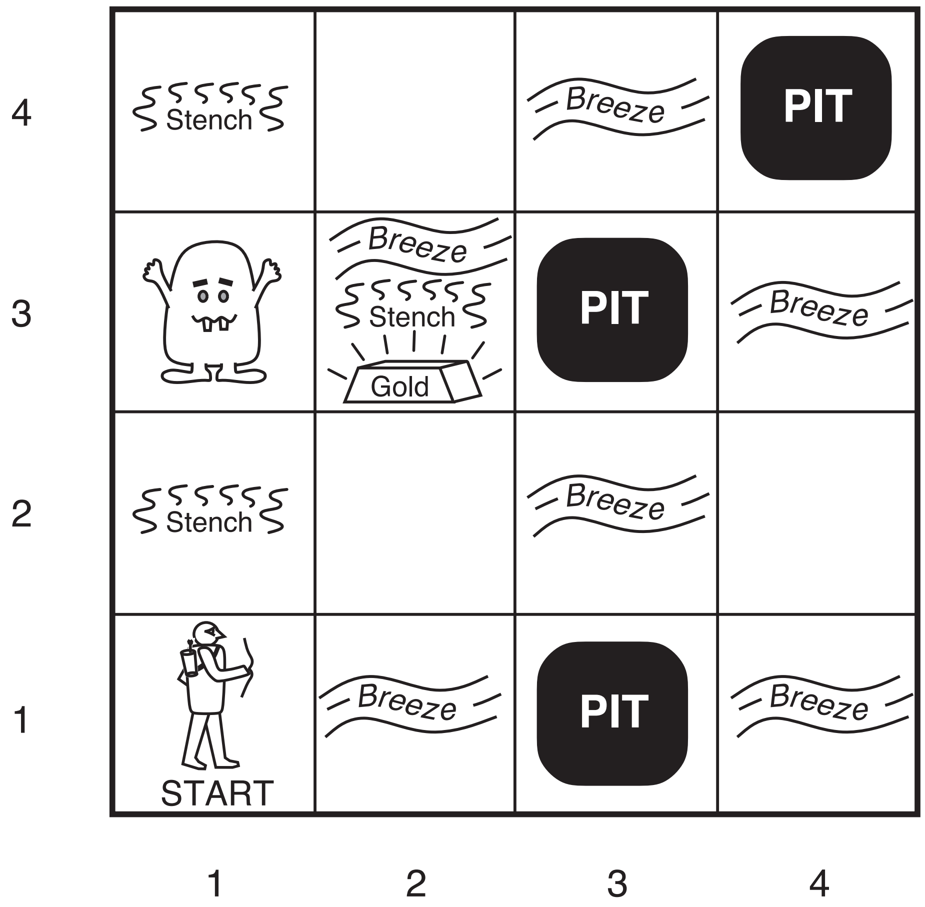
Αυτόνομοι Πράκτορες Βασισμένοι στη Γνώση

Μπορούμε να χτίσουμε έναν πράκτορα βασισμένο σε γνώση **ενημερώνοντάς** τον τι χρειάζεται να γνωρίζει πριν αρχίσει να αντιλαμβάνεται τον κόσμο (χρησιμοποιώντας την TELL).

Μπορούμε επίσης να σχεδιάσουμε έναν μηχανισμό **μάθησης** ο οποίος θα εξάγει γενική γνώση για το περιβάλλον του, έχοντας ως δεδομένα μια σειρά από πράγματα που έχει αντιληφθεί.

Αυτόνομος πράκτορας = Πράκτορας βασισμένος σε γνώση + Μηχανισμός μάθησης

Ο Κόσμος του Wumpus



Ο Κόσμος του Wumpus

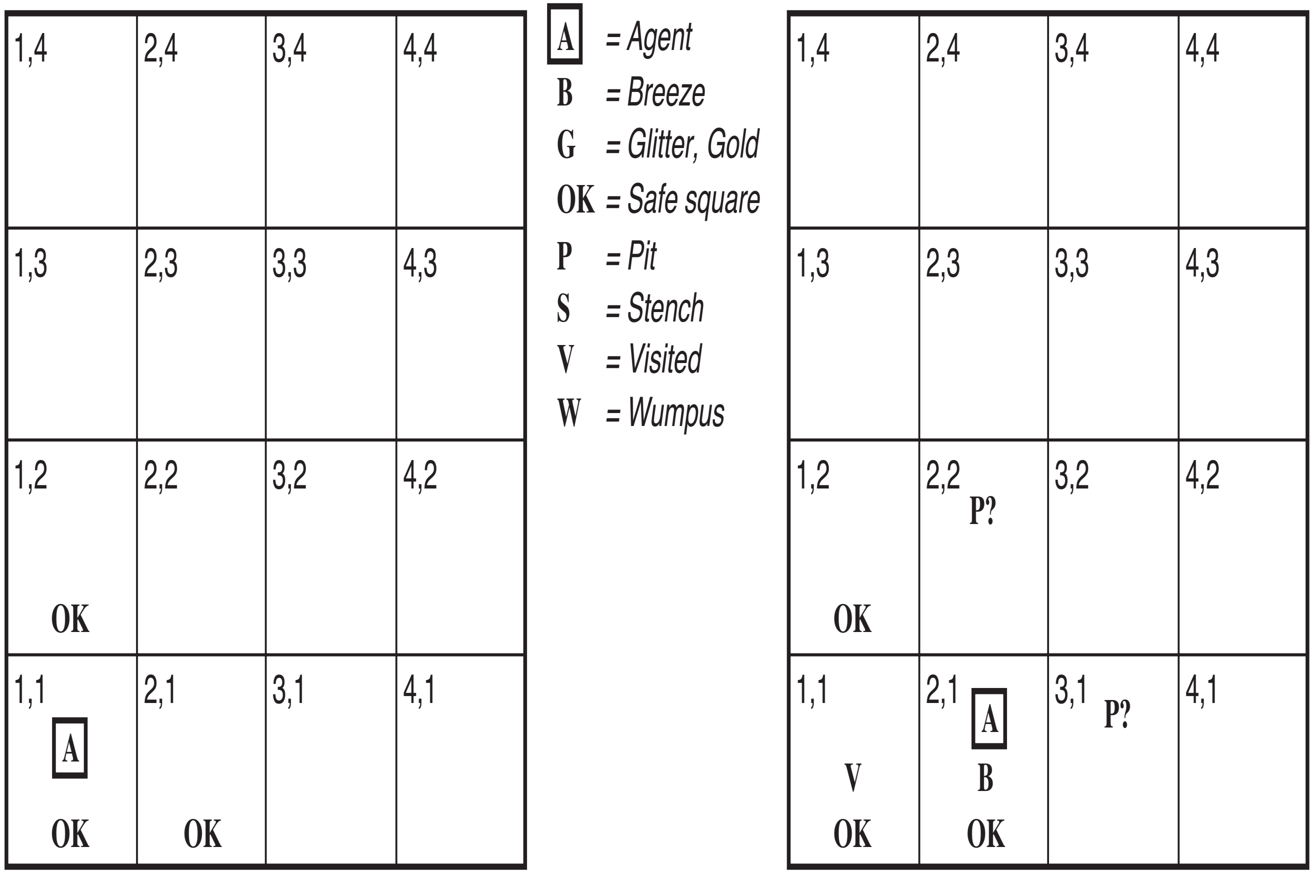
- **Περιβάλλον:** Πλαίσιο 4x4 όπου βρίσκεται ο πράκτορας, το τέρας Wumpus, μια πλάκα χρυσού και μερικά πηγάδια.
- **Μηχανισμοί δράσης:** Ο πράκτορας μπορεί να κινείται μπροστά και να στρίβει δεξιά ή αριστερά. Ο πράκτορας πεθαίνει αν μπει στο τετραγωνάκι που βρίσκεται το (ζωντανό) τέρας ή πέσει σε κάποιο πηγάδι.

  Ο πράκτορας μπορεί επίσης να εκτελέσει τις ενέργειες Grab και Shoot.
- **Αισθητήρες:** Η αντίληψη είναι μια λίστα από 5 σύμβολα:

$$(Stench,\ Breeze,\ Glitter,\ Bumb,\ Scream)$$

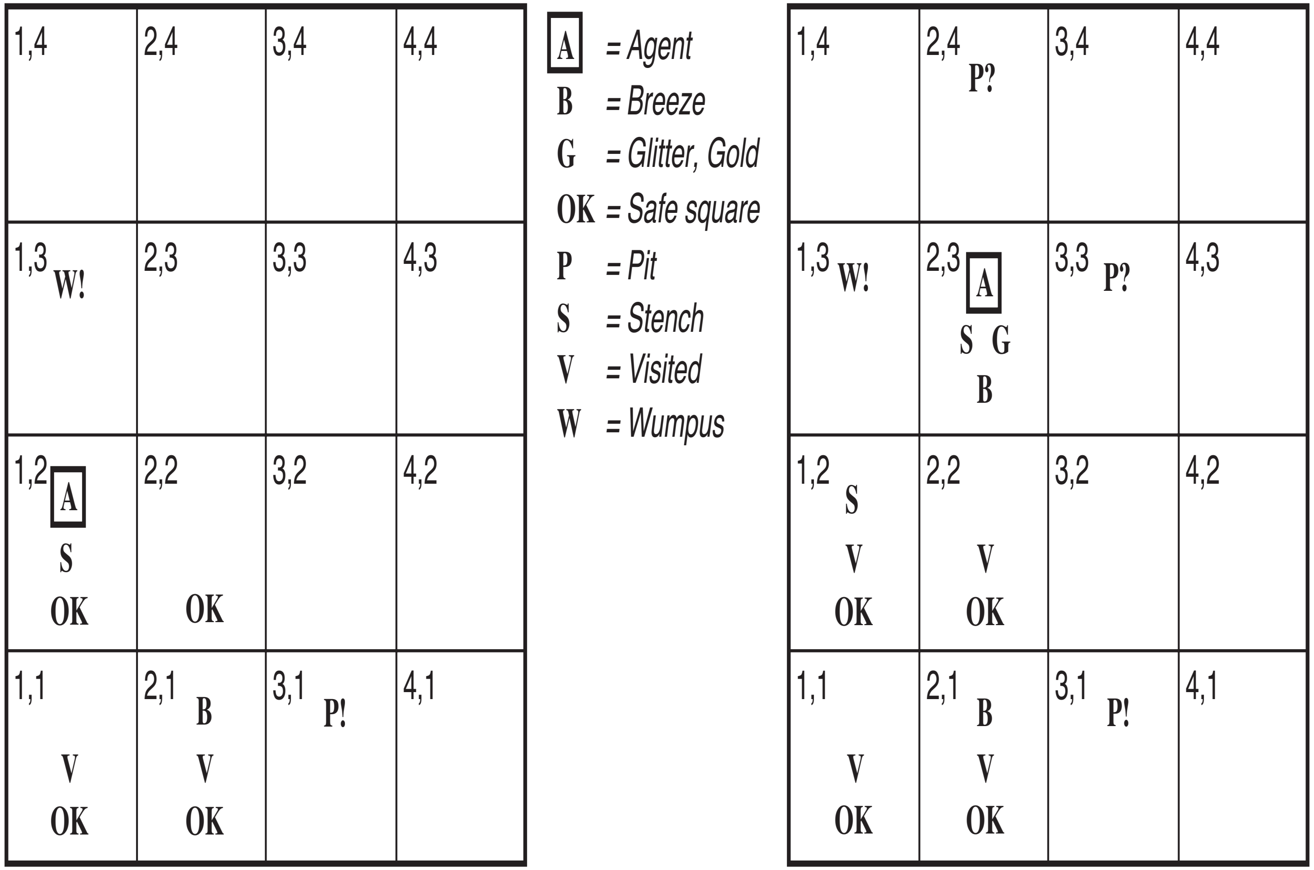
  Κάθε μια από τις παραπάνω τιμές μπορεί να είναι *None*.

Ο Κόσμος του Wumpus



(a) (b)

Ο Κόσμος του Wumpus



(a) (b)

Ο Κόσμος του Wumpus

Τι χρειάζεται ο πράκτορας για να επιζήσει και να πετύχει το σκοπό του;

- Ένα **μηχανισμό αντίληψης** του κόσμου.
- Ένα **μηχανισμό αναπαράστασης γνώσης (γεγονότων και κανόνων)** για τον κόσμο.
- Ένα **μηχανισμό συμπερασμού** ώστε από γνωστά γεγονότα να εξάγουμε άγνωστα.

Ο Κόσμος του Wumpus

Παραδείγματα γεγονότων:

- Αισθάνομαι ρεύμα στο τετραγωνάκι [2,1].
- Υπάρχει πηγάδι στο τετραγωνάκι [2,2] **ή** στο τετραγωνάκι [3,1].
- **Δεν** υπάρχει πηγάδι στο τετραγωνάκι [2,2].

Παράδειγμα κανόνα:

- Αν σε ένα τετραγωνάκι αισθάνεσαι ρεύμα, τότε σ᾽ ένα από τα διπλανά τετραγωνάκια υπάρχει πηγάδι.

Παράδειγμα συμπερασμού:

- Αφού στο τετραγωνάκι [2,1] αισθάνομαι ρεύμα, τότε υπάρχει πηγάδι στο τετραγωνάκι [2,2] **ή** στο τετραγωνάκι [3,1] (με βάση τον παραπάνω κανόνα).

Η Συνέχεια

Στη συνέχεια του μαθήματος θα μελετήσουμε **δύο γλώσσες αναπαράστασης γνώσης** και θα ορίσουμε αντίστοιχους μηχανισμούς συμπερασμού.

Οι γλώσσες αυτές θα μπορούν να χρησιμοποιηθούν για αναπαράσταση γνώσης στο κόσμο του Wumpus.

Γλώσσες Αναπαράστασης Γνώσης

Μια γλώσσα αναπαράστασης γνώσης ορίζεται προσδιορίζοντας το **συντακτικό (syntax)** και τη **σημασιολογία** της **(semantics)**.

Το **συντακτικό** μιας γλώσσας αναπαράστασης γνώσης καθορίζει με ακρίβεια τους καλά ορισμένους τύπους και προτάσεις της γλώσσας.

Η **σημασιολογία** μιας γλώσσας αναπαράστασης γνώσης ορίζει την αντιστοιχία μεταξύ των τύπων/προτάσεων της γλώσσας και των γεγονότων του κόσμου στον οποίο αναφέρονται οι τύποι/προτάσεις.

Ερμηνεία - Αληθείς Προτάσεις

Μια πρόταση μιας γλώσσας αναπαράστασης γνώσης δεν σημαίνει τίποτα από μόνη της. Η **σημασιολογία** (δηλαδή η **σημασία**) της προτασης πρέπει να παρέχεται από τον συγγραφέα της με τη μορφή μιας **ερμηνείας (interpretation).**

Αληθείς προτάσεις. Μια πρόταση θα λέγεται **αληθής (true) σε μια συγκεκριμένη ερμηνεία** αν η κατάσταση που περιγράφει η πρόταση ισχύει στην ερμηνεία.

Λογική Κάλυψη (Entailment)

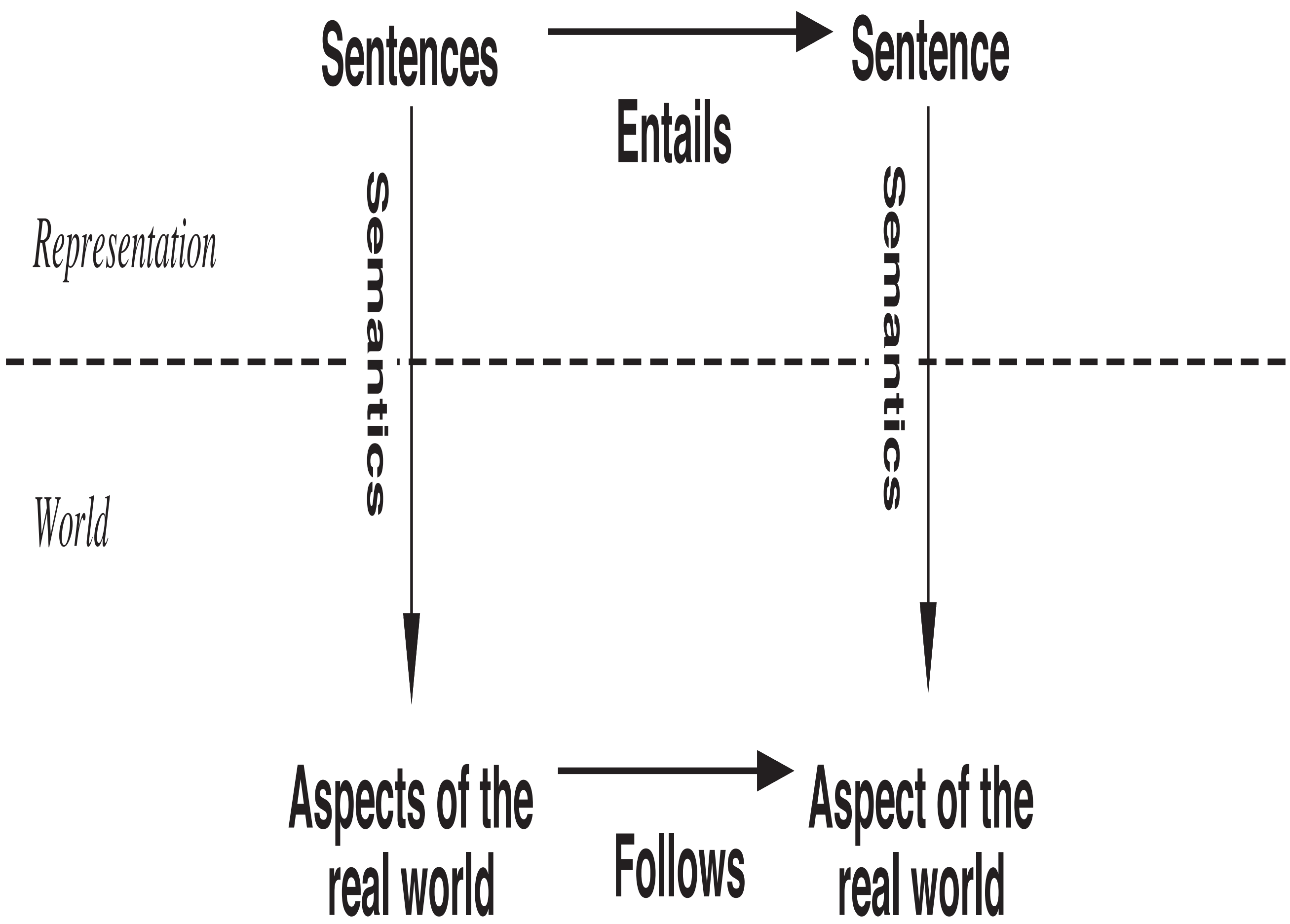
Θα γράφουμε

$$KB \models \alpha$$

για να δηλώσουμε ότι οποτεδήποτε οι προτάσεις της KB είναι αληθείς, τότε η πρόταση α είναι επίσης αληθής. Σ' αυτή την περίπτωση θα λέμε ότι οι προτάσεις της KB **καλύπτουν λογικά** την πρόταση α (ή ότι η α **έπεται λογικά** από την KB ή ότι η α είναι **λογική συνέπεια** της KB).

Δεδομένης μιας KB και μιας πρότασης α, πώς μπορούμε να σχεδιάσουμε έναν αλγόριθμο ο οποίος θα επαληθεύει αν ισχύει $KB \models \alpha$;

Ο Ρόλος της Σημασιολογίας



Συμπερασμός

Συμπερασμός (inference) είναι η διαδικασία εξαγωγής προτάσεων που έπονται λογικά από μια βάση γνώσης με **μηχανικό τρόπο**. Αν μια πρόταση α παράγεται από την KB χρησιμοποιώντας ένα μηχανισμό συμπερασμού i τότε γράφουμε $KB \vdash_i \alpha$.

Ένας μηχανισμός συμπερασμού καλείται **ορθός (sound)** αν παράγει **μόνο** προτάσεις που έπονται λογικά.

Ένας μηχανισμός συμπερασμού καλείται **πλήρης (complete)** αν παράγει **όλες** τις προτάσεις που έπονται λογικά.

Θεωρία Αποδείξεων - Κανόνες Συμπερασμού

Το σύνολο των βημάτων που ακολουθούνται για να παραχθεί μια νέα πρόταση α από το σύνολο προτάσεων μιας βάσης γνώσης KB καλείται **απόδειξη (proof).**

Μια **θεωρία αποδείξεων (proof theory)** είναι ένα σύνολο κανόνων που χρησιμοποιούνται για την εξαγωγή προτάσεων που έπονται λογικά από ένα άλλο σύνολο προτάσεων (μια βάση γνώσεων). Οι κανόνες αυτοί λέγονται **κανόνες συμπερασμού (inference rules).**

Λογική

Οι γλώσσες αναπαράστασης γνώσης που θα μελετήσουμε είναι:

- Η **προτασιακή λογική (propositional logic)** ή **λογική του Boole (Boolean logic)**.
- Η **λογική πρώτης τάξης (first-order logic).**

Έτσι οι πράκτορες βασισμένοι στη γνώση που θα δημιουργήσουμε θα ονομάζονται **λογικοί πράκτορες (logical agents).**

Λογική

Γενικά μια **λογική (logic)** είναι ένα τυπικό σύστημα που αποτελείται από:

- **Συντακτικό (Syntax)**
- **Σημασιολογία (Semantics)**
- **Θεωρία αποδείξεων (Proof theory)**

Ερώτηση: Γιατί χρησιμοποιούμε γλώσσες λογικής για την αναπαράσταση γνώσης; Γιατί να μην χρησιμοποιήσουμε φυσική γλώσσα ή κάποια γλώσσα προγραμματισμού;

Προτασιακή Λογική: Συντακτικό

Τα σύμβολα της προτασιακής λογικής είναι:

- Οι **σταθερές** $True$ και $False$.
- Ένα αριθμήσιμο απειροσύνολο **προτασιακών συμβόλων (proposition symbols)** $P_1, P_2, \ldots$ Αυτό το σύνολο θα συμβολίζεται με $\mathcal{P}$.
- Οι λογικοί σύνδεσμοι: $\neg, \wedge, \vee, \Rightarrow$ και $\Leftrightarrow$.
- Οι παρενθέσεις: (,).

Προτασιακή Λογική: Συντακτικό

Η παρακάτω γραμματική χωρίς συμφραζόμενα ορίζει τις **καλά ορισμένες προτάσεις (well-formed sentences)** της προτασιακής λογικής:

$$Sentence \rightarrow AtomicSentence \mid ComplexSentence$$

$$AtomicSentence \rightarrow \mathbf{True} \mid \mathbf{False} \mid Symbol$$

$$Symbol \rightarrow \mathbf{P_1} \mid \mathbf{P_2} \mid \cdots$$

$$\begin{aligned} ComplexSentence \rightarrow\ & (Sentence) \mid \neg Sentence \\ & \mid Sentence\ BinaryConnective\ Sentence \end{aligned}$$

$$BinaryConnective \rightarrow \wedge \mid \vee \mid \Rightarrow \mid \Leftrightarrow$$

Προτασιακή Λογική: Συντακτικό

Πρόταση: Όλοι οι δυαδικοί λογικοί σύνδεσμοι μπορούν να οριστούν με βάση τον $\neg$ και ένα από τους $\wedge, \vee, \Rightarrow$.

Προτεραιότητα

Δεν υπάρχουν αποδεκτοί απ΄ όλους κανόνες προτεραιότητας τελεστών για την προτασιακή λογική.

Η προτεραιότητα (από τη μεγαλύτερη προς τη μικρότερη) που θα χρησιμοποιήσουμε ακολουθώντας το βιβλίο AIMA είναι: $\neg, \wedge, \vee, \Rightarrow$ και $\Leftrightarrow$.

Άν έχετε αμφιβολία για την προτεραιότητα, τότε χρησιμοποιείστε **παρενθέσεις**. Επιτρέπονται και οι **αγκύλες** ώστε να εξασφαλίζεται η αναγνωσιμότητα μιας πρότασης.

Παραδείγματα

- Η πρόταση $P \wedge Q \vee R$ είναι ισοδύναμη με την $(P \wedge Q) \vee R$.
- Η πρόταση $\neg P \vee Q \wedge R \Rightarrow S$ είναι ισοδύναμη με την $(\neg(P) \vee (Q \wedge R)) \Rightarrow S$.

Η πρόταση $A \wedge B \wedge C$ (όπου η προτεραιότητα δεν βοηθάει) μπορεί να διαβαστεί σαν $(A \wedge B) \wedge C$ ή $A \wedge (B \wedge C)$ χωρίς πρόβλημα λόγω ισοδυναμίας των τριών αυτών προτάσεων (θα ορίσουμε με ακρίβεια την έννοια της ισοδυναμίας παρακάτω).

Ομοίως για τους σύνδεσμους $\vee$ και $\Leftrightarrow$, αλλά όχι για τον $\Rightarrow$. Πρέπει δηλαδή να χρησιμοποιήσουμε παρενθέσεις στην πρόταση $A \Rightarrow B \Rightarrow C$.

Προτασιακή Λογική: Οντολογικές Υποθέσεις

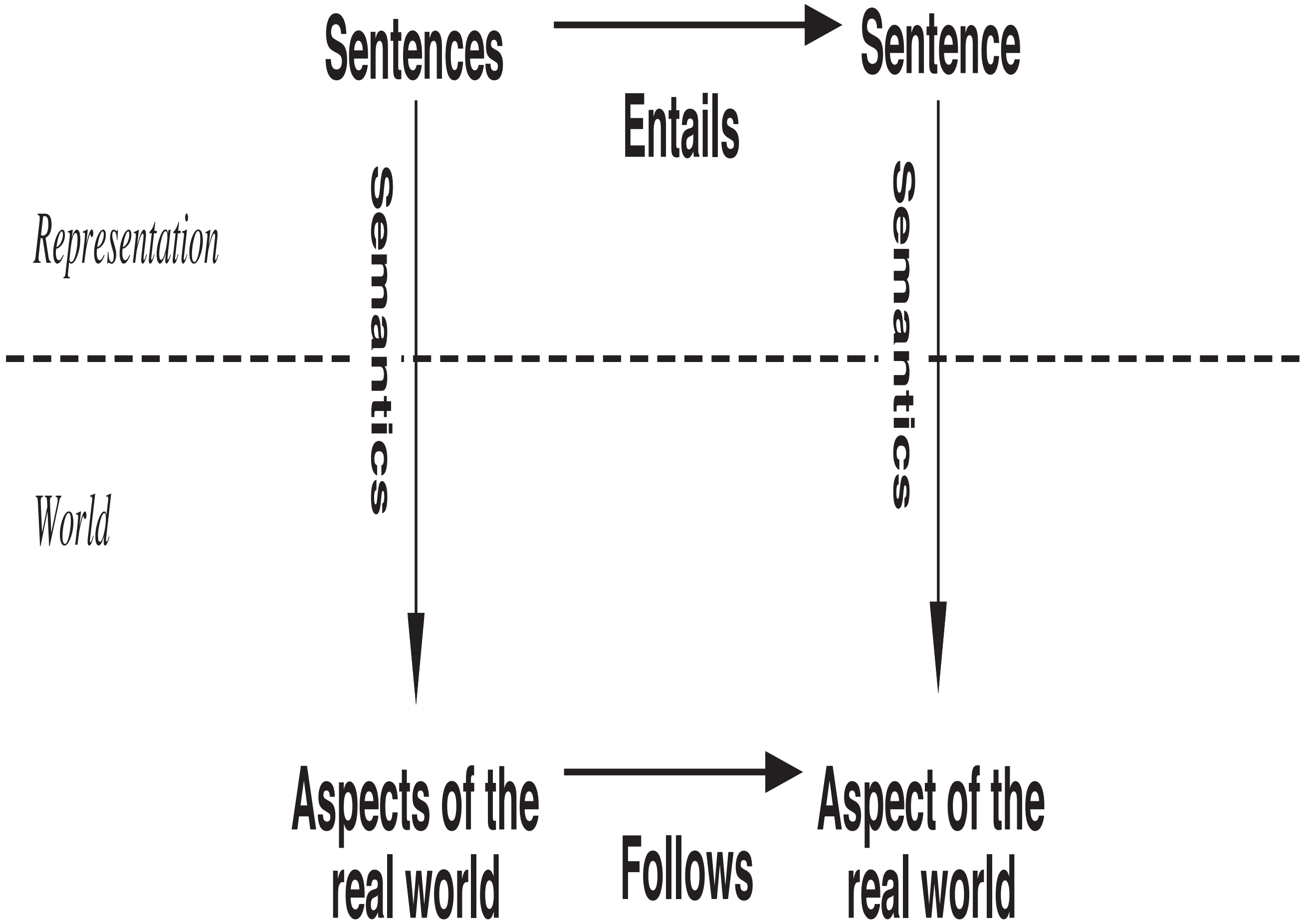
Οι **οντολογικές υποθέσεις** μιας λογικής σχετίζονται με τη φύση των κόσμων που μπορούν να αναπαρασταθούν από τη συγκεκριμένη λογική.

Η προτασιακή λογική υποθέτει ότι **ο κόσμος αποτελείται από γεγονότα (facts) τα οποία είτε αληθεύουν είτε δεν αληθεύουν.**

Η λογική πρώτης τάξης που επεκτείνει την προτασιακή λογική κάνει πιο πολύπλοκες και λεπτομερείς οντολογικές υποθέσεις.

Η βασική υπόθεση ότι ένα γεγονός είναι **αληθές ή ψευδές** δεν ισχύει σε άλλες λογικές π.χ. την **ασαφή λογική (fuzzy logic)**.

Προτασιακή Λογική: Σημασιολογία



Προτασιακή Λογική: Σημασιολογία

Ένα **προτασιακό σύμβολο (proposition symbol)** μπορεί να παριστάνει οτιδήποτε θέλουμε. Η ερμηνεία του μπορεί να είναι ένα οποιοδήποτε γεγονός. Όμως το γεγονός αυτό πρέπει να είναι είτε **αληθές** είτε **ψευδές** στον κόσμο που μοντελοποιούμε.

Τα παραπάνω διατυπώνονται τυπικά εισάγοντας την έννοια της **ερμηνείας**.

Ορισμός. Έστω $\mathcal{P}$ ένα σύνολο προτασιακών συμβόλων. Μια **ερμηνεία (interpretation)** για το $\mathcal{P}$ είναι μια απεικόνιση

$$I : \mathcal{P} \rightarrow \{false, true\}.$$

Προτασιακή Λογική: Σημασιολογία

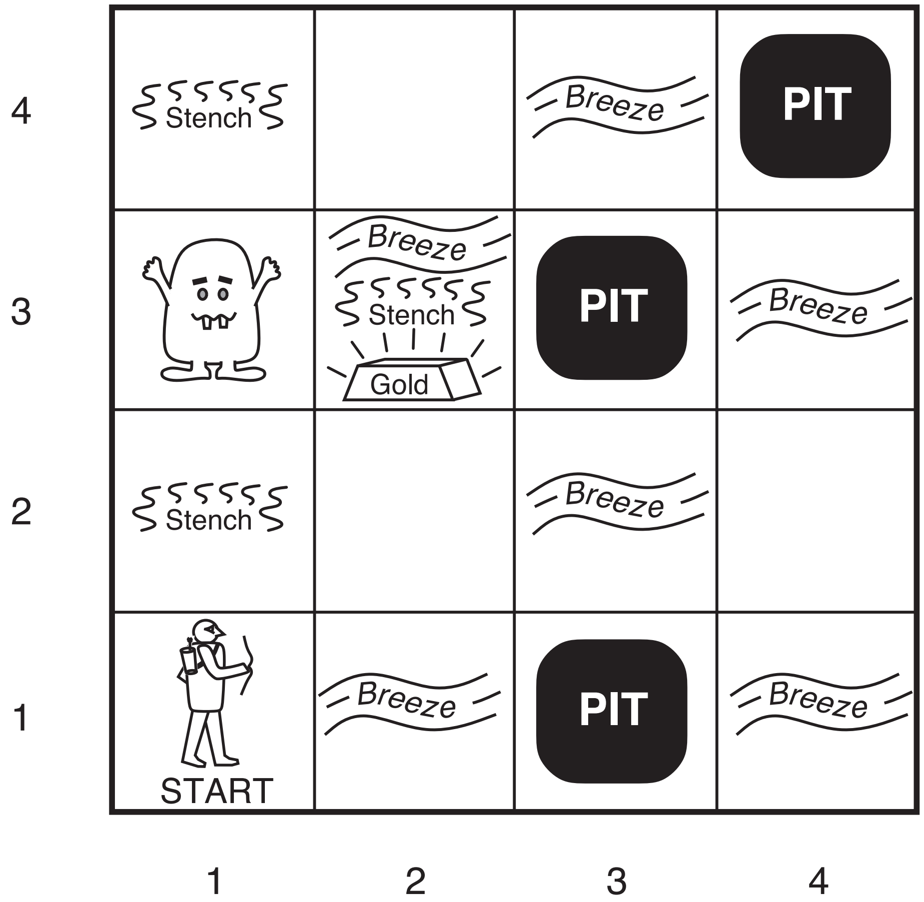
Η έννοια της ερμηνείας μπορεί να επεκταθεί σε όλες τις καλά ορισμένες προτάσεις της προτασιακής λογικής ακολουθώντας τους εξής αναδρομικούς ορισμούς:

- $I(True) = true$.
- $I(False) = false$.
- $I(\neg\phi) = true$ αν $I(\phi) = false$, διαφορετικά $I(\neg\phi) = false$.
- $I(\phi_1 \land \phi_2) = true$ αν $I(\phi_1) = true$ και $I(\phi_2) = true$, διαφορετικά $I(\phi_1 \land \phi_2) = false$.
- $I(\phi_1 \lor \phi_2) = true$ αν $I(\phi_1) = true$ ή $I(\phi_2) = true$, διαφορετικά $I(\phi_1 \lor \phi_2) = false$.

Προτασιακή Λογική: Σημασιολογία

- $I(\phi_1 \Rightarrow \phi_2) = true$ αν $I(\phi_1) = false$ ή $I(\phi_2) = true$, διαφορετικά $I(\phi_1 \Rightarrow \phi_2) = false$.
- $I(\phi_1 \Leftrightarrow \phi_2) = true$ αν $I(\phi_1) = I(\phi_2)$, διαφορετικά $I(\phi_1 \Leftrightarrow \phi_2) = false$.

Παράδειγμα: Ο Κόσμος του Wumpus



Αναπαράσταση με Προτασιακή Λογική

Θα χρησιμοποιήσουμε τα εξής προτασιακά σύμβολα για να παραστήσουμε μερικές από τις γνώσεις που έχουμε για το κόσμο του Wumpus όπως αυτές δίνονται στην προηγούμενη εικόνα:

- Ο πράκτορας είναι στο τετραγωνάκι $[x, y]$: A_{xy}
- Ο Wumpus είναι στο τετραγωνάκι $[x, y]$: W_{xy}
- Υπάρχει πηγάδι στο τετραγωνάκι $[x, y]$: P_{xy}
- Υπάρχει αύρα στο τετραγωνάκι $[x, y]$: B_{xy}

Αναπαράσταση με Προτασιακή Λογική

Στην προηγούμενη εικόνα, με βάση το συμβολισμό μας, μπορούμε διαισθητικά να πούμε ότι οι παρακάτω προτάσεις **'ισχύουν'**:

$$A_{11},\ W_{31},\ P_{13},\ P_{33},\ P_{44},\ B_{43},\ B_{32},\ B_{34},\ B_{34},\ B_{23},\ B_{12},\ B_{14},$$

$$\neg A_{12},\ \neg W_{11},\ \neg (P_{44} \wedge A_{44})$$

Μια Κατάλληλη Ερμηνεία

Μια κατάλληλη ερμηνεία I για τα προτασιακά σύμβολα μας που κωδικοποιεί ό,τι βλέπουμε σχετικά με αυτά τα σύμβολα στην προηγούμενη εικόνα καθορίζεται ως εξής:

$$I(A_{11}) = true,\ I(A_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

$$I(W_{31}) = true,\ I(W_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

$$I(P_{13}) = true,\ I(P_{33}) = true,\ I(P_{44}) = true,$$

$$I(P_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

$$I(B_{43}) = true,\ I(B_{32}) = true,\ I(B_{34}) = true,\ I(B_{34}) = true,$$

$$I(B_{23}) = true,\ I(B_{12}) = true,\ I(B_{14}) = true,$$

$$I(B_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

Ποιές Προτάσεις Ισχύουν;

Μπορούμε τώρα να δούμε ποιές προτάσεις 'ισχύουν' διαισθητικά στην εικόνα που είδαμε. Είναι αυτές που η ερμηνεία I τους δίνει την τιμή $true$. Για παράδειγμα:

$$I(\neg A_{12}) = true,\ I(\neg W_{11}) = true,\ I(\neg(P_{44} \wedge A_{44})) = true,$$

$$I(A_{11} \Rightarrow \neg W_{11}) = true,\ I(A_{13} \Rightarrow W_{11}) = true$$

Αντίθετα, οι παρακάτω προτάσεις στις οποίες η ερμηνεία δίνει την τιμή false, δεν 'ισχύουν¨:

$$I(\neg A_{11}) = false,\ I(P_{43} \vee A_{44}) = false,\ I(P_{44} \Rightarrow (W_{11} \vee A_{12})) = false$$

Άλλες Ερμηνείες

Υπάρχουν άλλες ερμηνείες για τα προτασιακά σύμβολα μας που δεν αντιστοιχούν στην προηγούμενη εικόνα. Για παράδειγμα η παρακάτω ερμηνεία J:

$$J(A_{44}) = true,\ J(A_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

$$J(W_{31}) = true,\ J(W_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

$$J(P_{13}) = true,\ J(P_{33}) = true,\ J(P_{44}) = true,$$

$$J(P_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

$$J(B_{43}) = true,\ J(B_{32}) = true,\ J(B_{34}) = true,\ J(B_{34}) = true,$$

$$J(B_{23}) = true,\ J(B_{12}) = true,\ J(B_{14}) = true,$$

$$J(B_{xy}) = false \text{ για όλα τα άλλα ζεύγη } x, y$$

Συνθετικότητα της Προτασιακής Λογικής

Μια γλώσσα καλείται **συνθετική (compositional)** όταν η σημασία μιας πρότασης της γλώσσας είναι συνάρτηση της σημασίας των τμημάτων αυτής της πρότασης.

Η συνθετικότητα είναι μια επιθυμητή ιδιότητα στις τυπικές γλώσσες.

Η προτασιακή λογική έχει την ιδιότητα της συνθετικότητας όπως μπορούμε να δούμε από τους ορισμούς της σημασιολογίας της.

Ικανοποίηση - Μοντέλο

Ορισμός. Έστω ϕ μια πρόταση της προτασιακής λογικής. Αν I είναι μια ερμηνεία τέτοια ώστε $I(\phi) = true$ τότε λέμε ότι I **ικανοποιεί (satisfies)** τη ϕ ή ότι η I είναι ένα **μοντέλο (model)** της ϕ.

Παράδειγμα

Η ερμηνεία I που δώσαμε νωρίτερα ικανοποιεί τις παρακάτω προτάσεις:

$$True,\ \neg A_{12},\ \neg W_{11},\ \neg(P_{44} \wedge A_{44}),$$

$$A_{11} \Rightarrow \neg W_{11},\ A_{13} \Rightarrow W_{11}$$

Η ερμηνεία I δεν ικανοποιεί τις παρακάτω προτάσεις:

$$False,\ \neg A_{11},\ P_{43} \vee A_{44},\ P_{44} \Rightarrow (W_{11} \vee A_{12})$$

Ικανοποιησιμότητα

Ορισμός. Μια πρόταση ϕ της προτασιακής λογικής είναι **ικανοποιήσιμη (satisfiable)** αν υπάρχει μια ερμηνεία I τέτοια ώστε $I(\phi) = true$.

Παραδείγματα: $P, \; P \vee Q, \; (P \wedge R) \vee Q$

Ορισμός. Μια πρόταση ϕ της προτασιακής λογικής είναι **μη ικανοποιήσιμη (unsatisfiable)** αν δεν υπάρχει καμία ερμηνεία I τέτοια ώστε $I(\phi) = true$.

Παράδειγμα: $P \wedge \neg P$

Εγκυρότητα (Validity)

Ορισμός. Μια πρόταση ϕ της προτασιακής λογικής είναι **έγκυρη (valid)** αν για κάθε ερμηνεία I, ισχύει $I(\phi) = true$.

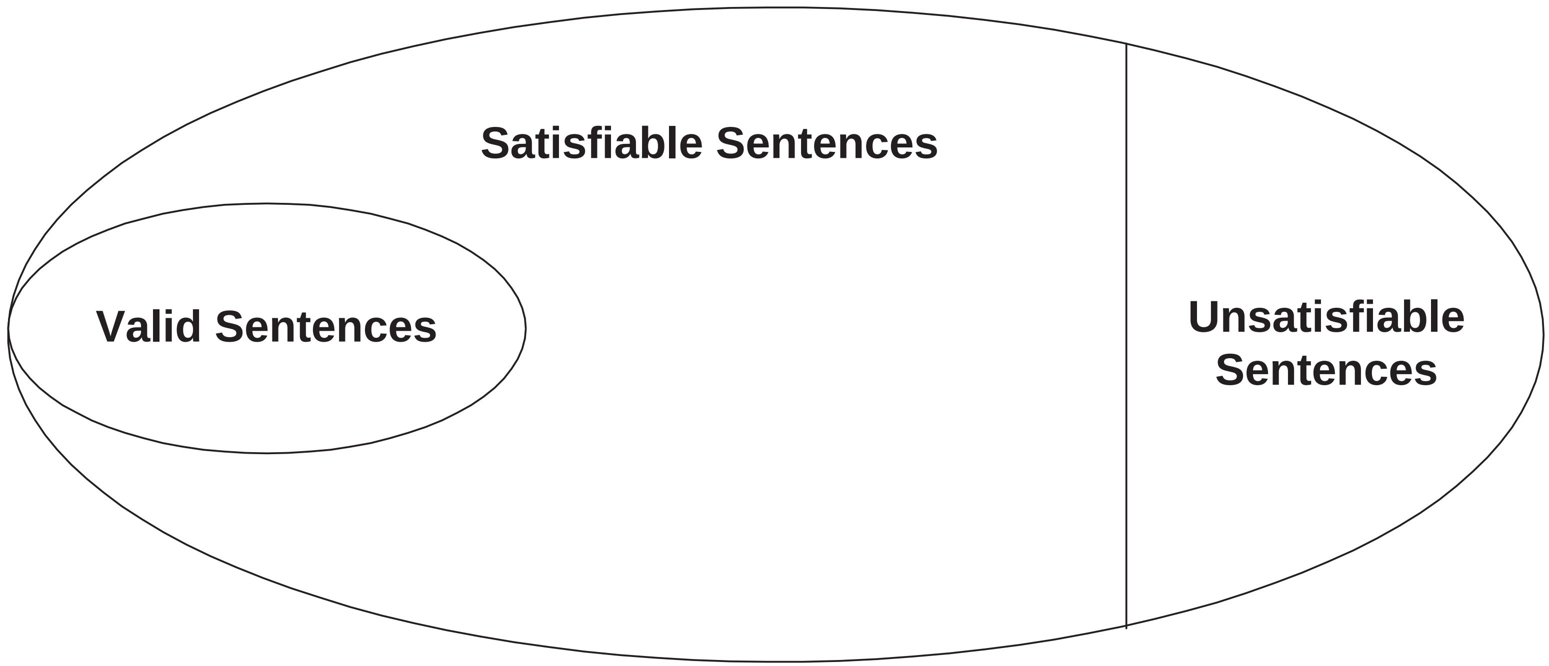
Παραδείγματα: $P \vee \neg P$, $((P \vee H) \wedge \neg H) \Rightarrow P$

Οι έγκυρες προτάσεις της προτασιακής λογικής λέγονται και **ταυτολογίες (tautologies).**

Θεώρημα. Έστω ϕ μια πρόταση της προτασιακής λογικής. Η ϕ είναι μη ικανοποιήσιμη ανν η $\neg\phi$ είναι έγκυρη. Απόδειξη;

Προσοχή: 'ανν' σημαίνει 'αν και μόνο αν'

Η Γεωγραφία της Προτασιακής Λογικής



Λογική Κάλυψη (Entailment)

Ορισμός. Έστω ϕ και ψ προτάσεις της προτασιακής λογικής. Θα λέμε ότι η ϕ **καλύπτει λογικά (entails)** την ψ ή ότι η ψ **έπεται λογικά** από την ϕ ή ότι η ψ **είναι λογική συνέπεια** της ϕ (συμβολισμός: $\phi \models \psi$) αν για κάθε ερμηνεία I τέτοια ώστε $I(\phi) = true$ ισχύει ότι $I(\psi) = true$.

Παραδείγματα: $P \wedge Q \models P,\ P \wedge (P \Rightarrow Q) \models Q$

Θεώρημα Παραγωγής (Deduction Theorem). Έστω ϕ και ψ προτάσεις της προτασιακής λογικής. Τότε $\phi \models \psi$ ανν $\phi \Rightarrow \psi$ είναι έγκυρη. Απόδειξη;

Λογική Κάλυψη και Μη Ικανοποιησιμότητα

Θεώρημα. Έστω ϕ και ψ προτάσεις της προτασιακής λογικής. Τότε $\phi \models \psi$ ανν $\phi \wedge \neg\psi$ είναι μη ικανοποιήσιμη. Απόδειξη;

Το παραπάνω θεώρημα είναι η βάση των **αποδείξεων με απαγωγή σε άτοπο** (σε χρήση από την εποχή του Αριστοτέλη).

Ισοδυναμία

Ορισμός. Έστω ϕ και ψ προτάσεις της προτασιακής λογικής. Θα λέμε ότι η ϕ είναι **ισοδύναμη (equivalent)** με την ψ (συμβολισμός: $\phi \equiv \psi$) αν $\phi \models \psi$ και $\psi \models \phi$.

Παράδειγμα: $\neg(P \wedge \neg Q) \equiv \neg P \vee Q$

Θεώρημα. $\phi \equiv \psi$ ανν η πρόταση $\phi \Leftrightarrow \psi$ είναι έγκυρη. Απόδειξη;

Μερικές Χρήσιμες Ισοδυναμίες

Έστω α, β και γ προτάσεις της προτασιακής λογικής. Τότε:

- $(\alpha \wedge \beta) \equiv (\beta \wedge \alpha)$ αντιμεταθετικότητα του $\wedge$
- $(\alpha \vee \beta) \equiv (\beta \vee \alpha)$ αντιμεταθετικότητα του $\vee$
- $((\alpha \wedge \beta) \wedge \gamma) \equiv (\alpha \wedge (\beta \wedge \gamma))$ προσεταιριστικότητα του $\wedge$
- $((\alpha \vee \beta) \vee \gamma) \equiv (\alpha \vee (\beta \vee \gamma))$ προσεταιριστικότητα του $\vee$
- $\neg(\neg\alpha) \equiv \alpha$ απαλοιφή διπλής άρνησης
- $(\alpha \Rightarrow \beta) \equiv (\neg\beta \Rightarrow \neg\alpha)$ αντιθετοαντιστροφή (contraposition)

Μερικές Χρήσιμες Ισοδυναμίες

- $(\alpha \Rightarrow \beta) \equiv (\neg\alpha \vee \beta)$ απαλοιφή συνεπαγωγής
- $(\alpha \Leftrightarrow \beta) \equiv ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha))$ απαλοιφή διπλής συνεπαγωγής
- $\neg(\alpha \wedge \beta) \equiv (\neg\alpha \vee \neg\beta)$ νόμος de Morgan
- $\neg(\alpha \vee \beta) \equiv (\neg\alpha \wedge \neg\beta)$ νόμος de Morgan
- $(\alpha \wedge (\beta \vee \gamma)) \equiv ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma))$ επιμεριστικότητα του $\wedge$ ως προς το $\vee$
- $(\alpha \vee (\beta \wedge \gamma)) \equiv ((\alpha \vee \beta) \wedge (\alpha \vee \gamma))$ επιμεριστικότητα του $\vee$ ως προς το $\wedge$

Πίνακες Αληθείας

Οι **πίνακες αληθείας (truth tables)** είναι εργαλεία που μας επιτρέπουν να βρούμε την τιμή αληθείας μιας πρότασης της προτασιακής λογικής αν γνωρίζουμε τις τιμές αληθείας των επιμέρους προτάσεων που την απαρτίζουν (με βάση τη **συνθεσιμότητα** της προτασιακής λογικής).

Πίνακες Αληθείας

A	$\neg A$
$true$	$false$
$false$	$true$

A	B	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
$false$	$false$	$false$	$false$	$true$	$true$
$false$	$true$	$false$	$true$	$true$	$false$
$true$	$false$	$false$	$true$	$false$	$false$
$true$	$true$	$true$	$true$	$true$	$true$

Πίνακες Αληθείας: Γιατί Είναι Χρήσιμοι;

- Οι πίνακες αληθείας μπορούν να χρησιμοποιηθούν για να αποδείξουμε την **αλήθεια ή το ψεύδος** μιας οποιασδήποτε πρότασης της προτασιακής λογικής σε μια δεδομένη ερμηνεία.
- Παρομοίως, οι πίνακες αληθείας μπορούν να χρησιμοποιηθούν για να αποδείξουμε την **ικανοποιησιμότητα** ή την **εγκυρότητα** μιας πρότασης, ή την **ισοδυναμία** δύο προτάσεων της προτασιακής λογικής.
- **Υπολογιστική Πολυπλοκότητα:** $O(2^n)$ όπου n είναι ο αριθμός των προτασιακών συμβόλων της πρότασης.

Παράδειγμα

Πίνακας αληθείας για να αποδείξουμε την εγκυρότητα της πρότασης $((P \vee H) \wedge \neg H) \Rightarrow P$.

P	H	$P \vee H$	$(P \vee H) \wedge \neg H$	$((P \vee H) \wedge \neg H) \Rightarrow P$
$false$	$false$	$false$	$false$	$true$
$false$	$true$	$true$	$false$	$true$
$true$	$false$	$true$	$true$	$true$
$true$	$true$	$true$	$false$	$true$

Παράδειγμα

Ένας πίνακας αληθείας που δείχνει ότι $\{ P \vee H,\ \neg H \} \models P$.

P	H	$P \vee H$	$\neg H$	P
$false$	$false$	$false$	$true$	$false$
$false$	$true$	$true$	$false$	$false$
$true$	$false$	$true$	$true$	$true$
$true$	$true$	$true$	$false$	$true$

Θεωρία Αποδείξεων και Κανόνες Συμπερασμού

Ένας **κανόνας συμπερασμού (inference rule)** είναι ένας κανόνας της μορφής

$$\frac{\alpha_1, \alpha_2, \ldots, \alpha_n}{\beta}$$

όπου $\alpha_1, \alpha_2, \ldots, \alpha_n$ είναι προτάσεις οι οποίες ονομάζονται **υποθέσεις** και β είναι μια πρόταση που ονομάζεται **συμπέρασμα**.

Ένας κανόνας συμπερασμού **εφαρμόζεται** ως εξής: οποτεδήποτε έχουμε ένα σύνολο προτάσεων που ταιριάζουν με τις υποθέσεις του κανόνα, μπορούμε να εξάγουμε την πρόταση-συμπέρασμα.

Προτασιακή Λογική: Κανόνες Συμπερασμού

- Τρόπος του θέτειν (modus ponens): $\frac{\alpha \Rightarrow \beta,\ \alpha}{\beta}$
- Απαλοιφή σύζευξης: $\frac{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}{\alpha_i}$
- Εισαγωγή σύζευξης: $\frac{\alpha_1, \alpha_2, \ldots, \alpha_n}{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}$
- Εισαγωγή διάζευξης: $\frac{\alpha_i}{\alpha_1 \vee \alpha_2 \vee \ldots \vee \alpha_n}$
- Απαλοιφή διπλής άρνησης: $\frac{\neg\neg\alpha}{\alpha}$
- Μοναδιαία ανάλυση: $\frac{\alpha \vee \beta,\ \neg\beta}{\alpha}$
- Ανάλυση: $\frac{\alpha \vee \beta,\ \neg\beta \vee \gamma}{\alpha \vee \gamma}$

Προτασιακή Λογική: Θεωρίες Αποδείξεων

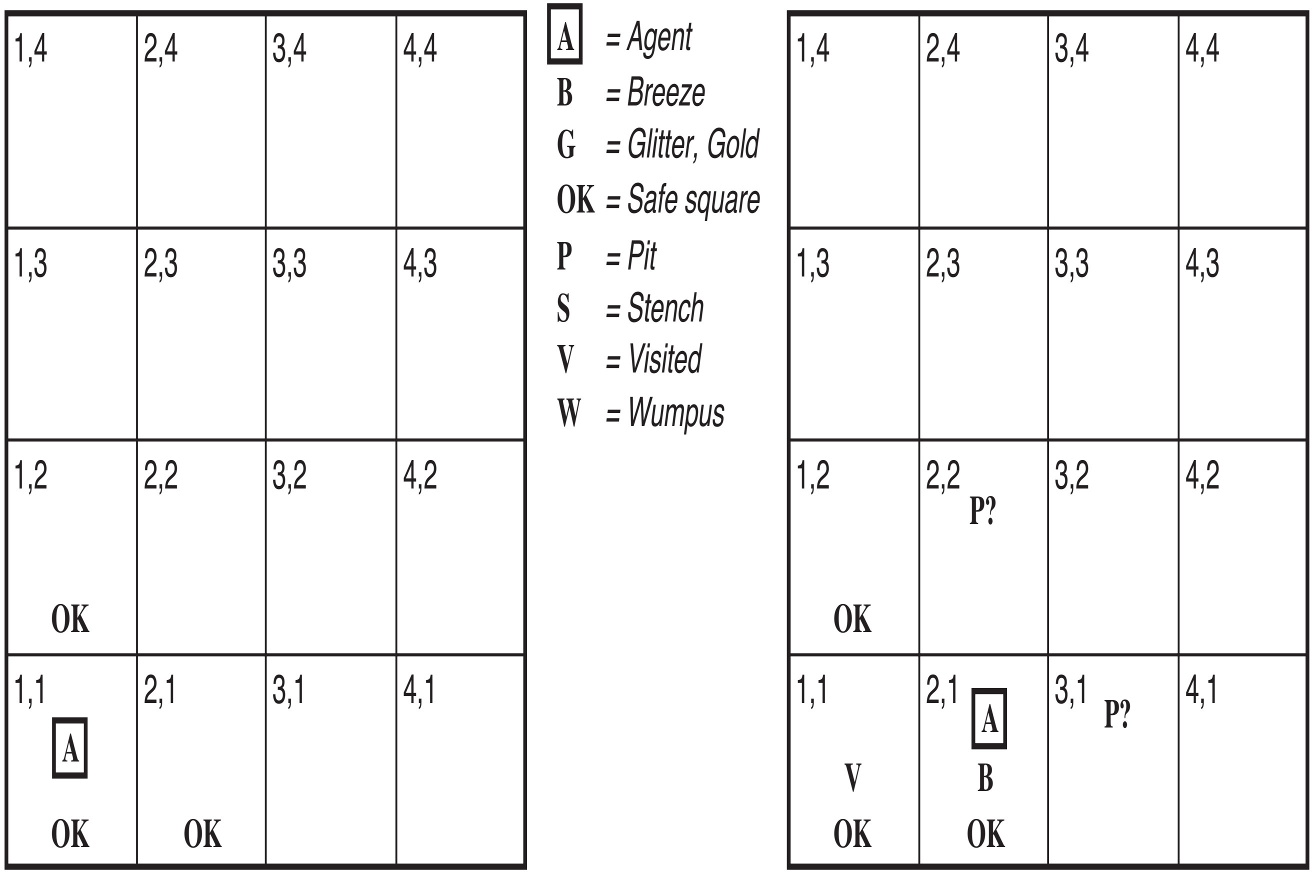
Η κλασσική βιβλιογραφία της προτασιακής λογικής μας δίνει διάφορες τυπικές θεωρίες αποδείξεων (δηλ. αποδεικτικές μεθόδους) που είναι **ορθές** και **πλήρεις** (π.χ., tableaux, Hilbert systems, natural deduction systems κλπ.). Οι μέθοδοι αυτοί ορίζουν κατάλληλους **κανόνες συμπερασμού**.

Δείτε για παράδειγμα τις μεθόδους που παρουσιάζονται στο βιβλίο:

> Melvin Fitting. *First-Order Logic and Automated Theorem Proving.* Springer, 1996.

Εμεις θα παρουσιάσουμε μια θεωρία αποδείξεων που βασίζεται στη χρήση **ενός μόνο** κανόνα συμπερασμού, του κανόνα της **ανάλυσης.**

Ο Κόσμος του Wumpus



(a) (b)

Συμπερασμός στον Κόσμο του Wumpus

Η **βάση γνώσης** για την εικόνα (a):

- Δεν υπάρχει πηγάδι στο $[1, 1]$ (αντίληψη του πράκτορα):

$$R_1 : \quad \neg P_{11}$$

- Σ' ένα τετραγωνάκι γίνεται αντιληπτή αύρα αν και μόνο αν υπάρχει πηγάδι σε γειτονικό τετραγωνάκι (κανόνας του κόσμου). Διατυπώνουμε αυτόν τον κανόνα μόνο για το τετραγωνάκι $[1, 1]$:

$$R_2 : \quad B_{11} \Leftrightarrow P_{12} \vee P_{21}$$

- Δε γίνεται αντιληπτή αύρα στο τετραγωνάκι $[1, 1]$ (αντίληψη του πράκτορα).

$$R_3 : \quad \neg B_{11}$$

Συμπερασμός στον Κόσμο του Wumpus

Ο πράκτορας μπορεί τώρα να χρησιμοποιήσει τους **κανόνες συμπερασμού της προτασιακής λογικής** και τις **λογικές ισοδυναμίες** για να αποδείξει το εξής:

Δεν υπάρχει πηγάδι στα τετραγωνάκια $[1,2]$ ή $[2,1]$.

- Εφαρμόζουμε απαλοιφή διπλής συνεπαγωγής στην R_2:

$$R_4: \quad (B_{11} \Rightarrow (P_{12} \vee P_{21})) \wedge ((P_{12} \vee P_{21}) \Rightarrow B_{11})$$

- Εφαρμόζουμε απαλοιφή σύζευξης στην R_4:

$$R_5: \quad (P_{12} \vee P_{21}) \Rightarrow B_{11}$$

Συμπερασμός στον Κόσμο του Wumpus

- Εφαρμόζουμε τη λογική ισοδυναμία για αντιθετοαντίστροφες προτάσεις στην R_5:

$$R_6 : \quad \neg B_{11} \Rightarrow \neg(P_{12} \vee P_{21})$$

- Εφαρμόζουμε modus ponens στις R_6 και R_3:

$$R_7 : \quad \neg(P_{12} \vee P_{21})$$

- Εφαρμόζουμε τον κανόνα de Morgan στην R_7 και έχουμε το **επιθυμητό συμπέρασμα**:

$$R_8 : \quad \neg P_{12} \wedge \neg P_{21}$$

Η Ιδιότητα της Μονοτονικότητας

Η προτασιακή λογική (όπως και η λογική πρώτης τάξης που θα δούμε αργότερα) είναι **μονότονη:**

Αν $KB \models \phi$ τότε $KB \wedge \psi \models \phi$.

Αυτό σημαίνει ότι αν προστεθεί νέα γνώση ψ σε μια βάση γνώσεων KB τότε μπορούμε πιθανά να συμπεράνουμε νέα γνώση, αλλά δεν μπορούμε να ακυρώσουμε γνώση που ήδη είχαμε συμπεράνει από την KB.

Ο Κανόνας Συμπερασμού της Ανάλυσης

Θα παρουσιάσουμε τώρα τον κανόνα συμπερασμού της ανάλυσης που είναι ένας **ορθός και πλήρης** κανόνας συμπερασμού για την προτασιακή λογική.

Για να εφαρμόσουμε τον κανόνα της ανάλυσης, οι προτάσεις της προτασιακής λογικής πρέπει να είναι σε συζευκτική κανονική μορφή.

Δίνουμε αμέσως τους σχετικούς ορισμούς.

Λεκτικά (Literals)

Ορισμός. Ένα **λεκτικό (literal)** είναι ένα προτασιακό σύμβολο ή η άρνηση του. Στην πρώτη περίπτωση έχουμε ένα **θετικό** λεκτικό και στη δεύτερη περίπτωση ένα **αρνητικό** λεκτικό.

Παραδείγματα:

$$P_1,\ P_2,\ \neg P_3$$

Φράσεις (Clauses)

Ορισμός. Μια **φράση (clause)** είναι μια διάζευξη λεκτικών.

Παραδείγματα:

$$P_1 \vee P_2 \vee \neg P_3, \quad P_1 \vee P_4, \quad \neg P_1 \vee \neg P_3 \vee \neg P_5$$

Συζευκτική Κανονική Μορφή

Ορισμός. Μια πρόταση της προτασιακής λογικής είναι σε **συζευκτική κανονική μορφή (conjunctive normal form, CNF)** αν είναι μια σύζευξη φράσεων (δηλ. μια σύζευξη διαζεύξεων που αποτελούνται από λεκτικά).

Θεώρημα. Κάθε πρόταση της προτασιακής λογικής μπορεί να μετατραπεί σε μια ισοδύναμη πρόταση που είναι σε CNF.

Μετατροπή σε CNF

1. **Απαλοιφή διπλών συνεπαγωγών** χρησιμοποιώντας την παρακάτω ισοδυναμία:

$$(\phi \Leftrightarrow \psi) \equiv (\phi \Rightarrow \psi \land \psi \Rightarrow \phi)$$

2. **Απαλοιφή απλών συνεπαγωγών** χρησιμοποιώντας την παρακάτω ισοδυναμία:

$$\phi \Rightarrow \psi \equiv \neg\phi \lor \psi$$

3. **Μετακίνηση των αρνήσεων ($\neg$) προς τα μέσα** ώστε κάθε άρνηση να εφαρμόζεται σε ένα ατομικό τύπο. Χρησιμοποιούμε τις παρακάτω ισοδυναμίες:

$$\neg\neg\phi \equiv \phi$$

$$\neg(\phi \lor \psi) \equiv \neg\phi \land \neg\psi$$

$$\neg(\phi \land \psi) \equiv \neg\phi \lor \neg\psi$$

Μετατροπή σε CNF

4. **Εφαρμόζουμε την επιμεριστική ιδιότητα του** $\vee$ **ως προς το** $\wedge$:

$$(\phi \wedge \psi) \vee \theta \equiv (\phi \vee \theta) \wedge (\psi \vee \theta)$$

$$\theta \vee (\phi \wedge \psi) \equiv (\theta \vee \phi) \wedge (\theta \vee \psi)$$

5. **Απλοποιούμε τις συζεύξεις και διαζεύξεις** απαλοίφοντας τις παρενθέσεις που δεν χρειάζονται.

6. **Απαλοίφουμε το σύμβολο** $\wedge$ και έχουμε μια λίστα από διαζεύξεις (φράσεις).

Παράδειγμα

Έστω η πρόταση

$$(P_1 \Leftrightarrow P_2) \wedge \neg(P_3 \Leftrightarrow P_4)$$

Βήμα 1:

$$(P_1 \Rightarrow P_2 \wedge P_2 \Rightarrow P_1) \wedge \neg(P_3 \Rightarrow P_4 \wedge P_4 \Rightarrow P_3)$$

Βήμα 2:

$$((\neg P_1 \vee P_2) \wedge (\neg P_2 \vee P_1)) \wedge \neg((\neg P_3 \vee P_4) \wedge (\neg P_4 \vee P_3))$$

Βήμα 3:

$$((\neg P_1 \vee P_2) \wedge (\neg P_2 \vee P_1)) \wedge (\neg(\neg P_3 \vee P_4) \vee \neg(\neg P_4 \vee P_3))$$

$$((\neg P_1 \vee P_2) \wedge (\neg P_2 \vee P_1)) \wedge ((P_3 \wedge \neg P_4) \vee (P_4 \wedge \neg P_3))$$

Παράδειγμα

Βήμα 4:

$$((\neg P_1 \vee P_2) \wedge (\neg P_2 \vee P_1)) \wedge ((P_3 \vee P_4) \wedge (P_3 \vee \neg P_3) \wedge (\neg P_4 \vee P_4) \wedge (\neg P_4 \vee \neg P_3))$$

Βήμα 5:

$$(\neg P_1 \vee P_2) \wedge (\neg P_2 \vee P_1) \wedge (P_3 \vee P_4) \wedge (P_3 \vee \neg P_3) \wedge (\neg P_4 \vee P_4) \wedge (\neg P_4 \vee \neg P_3)$$

Βήμα 6:

$$\neg P_1 \vee P_2, \quad \neg P_2 \vee P_1, \quad P_3 \vee P_4, \quad P_3 \vee \neg P_3, \quad \neg P_4 \vee P_4, \quad \neg P_4 \vee \neg P_3$$

Ο Κανόνας Συμπερασμού της Μοναδιαίας Ανάλυσης

Ο κανόνας συμπερασμού της **μοναδιαίας ανάλυσης (unit resolution)** για την **προτασιακή λογική** είναι ο εξής.

Αν l_i και m είναι συμπληρωματικά λεκτικά, τότε:

$$\frac{l_1 \vee \cdots \vee l_{i-1} \vee l_i \vee l_{i+1} \vee \cdots \vee l_k, \quad m}{l_1 \vee \cdots \vee l_{i-1} \vee l_{i+1} \vee \cdots \vee l_k}$$

Ορισμός. Τα λεκτικά l και m λέγονται **συμπληρωματικά (complementary)** αν το ένα είναι η άρνηση του άλλου.

Παράδειγμα

Έστω οι εξής φράσεις:

$$P_1 \vee P_2 \vee \neg P_3$$

$$\neg P_2$$

$$P_3$$

Από την φράση $P_1 \vee P_2 \vee \neg P_3$ και την P_3, με τον κανόνα της μοναδιαίας ανάλυσης μπορούμε να συμπεράνουμε την

$$P_1 \vee P_2.$$

Από την φράση $P_1 \vee P_2$ και την $\neg P_2$ με τον κανόνα της μοναδιαίας ανάλυσης μπορούμε να συμπεράνουμε την P_1.

Ο Κανόνας Συμπερασμού της Ανάλυσης

Ο κανόνας συμπερασμού της **ανάλυσης (resolution)** για την **προτασιακή λογική** είναι ο εξής.

Αν l_i και m_j είναι συμπληρωματικά λεκτικά, τότε:

$$\frac{l_1 \vee \cdots \vee l_{i-1} \vee l_i \vee l_{i+1} \vee \cdots \vee l_k, \quad m_1 \vee \cdots \vee m_{j-1} \vee m_j \vee m_{j+1} \vee \cdots \vee m_n}{l_1 \vee \cdots \vee l_{i-1} \vee l_{i+1} \vee \cdots \vee l_k \vee m_1 \vee \cdots \vee m_{j-1} \vee m_{j+1} \vee \cdots m_n}$$

Παράδειγμα

Έστω οι εξής φράσεις:

$$P_1 \vee P_2 \vee \neg P_3$$

$$\neg P_2 \vee P_1$$

$$P_3 \vee P_4$$

Από την φράση $P_1 \vee P_2 \vee \neg P_3$ και την $P_3 \vee P_4$, με τον κανόνα της ανάλυσης, μπορούμε να συμπεράνουμε την

$$P_1 \vee P_2 \vee P_4.$$

Από την φράση $P_1 \vee P_2 \vee P_4$ και την $\neg P_2 \vee P_1$ με τον κανόνα της ανάλυσης, μπορούμε να συμπεράνουμε την $P_1 \vee P_4 \vee P_1$. Η τελευταία φράση είναι ισοδύναμη με την $P_1 \vee P_4$.

Παραγοντοποίηση (Factoring)

Πρέπει πάντα να αφαιρούμε τις εμφανίσεις πολλαπλών λεκτικών από μια φράση όπως κάναμε παραπάνω.

Ισοδύναμα: μια φράση είναι ένα **σύνολο** λεκτικών.

Χρήση του Κανόνα της Ανάλυσης

Ο κανόνας της ανάλυσης χρησιμοποιείται για να δείξουμε ότι **ένα σύνολο φράσεων** S **είναι μη ικανοποιήσιμο.**

Αυτό αποδεκνύεται αν με τη χρήση του κανόνα της ανάλυσης φτάσουμε σε **αντίφαση**, δηλαδή σε δύο λεκτικά που το ένα είναι η άρνηση του άλλου. Αν εφαρμόσουμε ανάλυση σε αυτά τα λεκτικά, φτάνουμε στην **κενή φράση.**

Χρήση του Κανόνα της Ανάλυσης

Ο κανόνας της ανάλυσης μπορεί επίσης να χρησιμοποιηθεί όταν θέλουμε να δείξουμε ότι $KB \models \phi$ για μια βάση γνώσης KB και πρόταση ϕ.

Ισοδύναμα: μπορούμε να δείξουμε ότι η πρόταση $KB \land \neg\phi$ είναι μη ικανοποιήσιμη χρησιμοποιώντας ανάλυση.

Παράδειγμα

Δίνονται οι εξής προτάσεις στα Ελληνικά:

Αν ο μονόκερος είναι μυθικός τότε είναι αθάνατος, αλλά αν δεν είναι μυθικός τότε είναι θνητό θηλαστικό. Αν ο μονόκερος είναι θνητό θηλαστικό ή αθάνατος τότε είναι κερασφόρος. Αν ο μονόκερος είναι κερασφόρος τότε είναι μαγικός.

Να δείξετε: Ο μονόκερος είναι μαγικός.

Παράδειγμα: Προτασιακή Λογική

Θα χρησιμοποιήσουμε τα εξής προτασιακά σύμβολα:

- $Mythical$, θα παριστάνει την πρόταση: ο μονόκερος είναι μυθικός.
- $Mortal$, θα παριστάνει την πρόταση: ο μονόκερος είναι θνητός.
- $Mammal$, θα παριστάνει την πρόταση: ο μονόκερος είναι θηλαστικό.
- $Horned$, θα παριστάνει την πρόταση: ο μονόκερος είναι κερασφόρος.
- $Magical$, θα παριστάνει την πρόταση: ο μονόκερος είναι μαγικός.

Παράδειγμα: Προτασιακή Λογική

Οι δοσμένες προτάσεις παριστάνονται σε προτασιακή λογική ως εξής:

$$Mythical \Rightarrow \neg Mortal$$

$$\neg Mythical \Rightarrow Mortal$$

$$\neg Mythical \Rightarrow Mammal$$

$$((Mortal \wedge Mammal) \vee \neg Mortal) \Rightarrow Horned$$

$$Horned \Rightarrow Magical$$

Έχουμε να αποδείξουμε την πρόταση $Magical$.

Παράδειγμα: Συζευκτική Κανονική Μορφή

Μετατρέπουμε τις δοσμένες προτάσεις και την άρνηση της πρότασης που έχουμε να αποδείξουμε σε CNF:

$$\neg Mythical \vee \neg Mortal$$

$$Mythical \vee Mortal$$

$$Mythical \vee Mammal$$

$$\neg Mortal \vee \neg Mammal \vee Horned$$

$$Mortal \vee Horned$$

$$\neg Horned \vee Magical$$

$$\neg Magical$$

Παράδειγμα: Ανάλυση

1. Από την πρόταση $\neg Mythical \vee \neg Mortal$ και την $Mythical \vee Mammal$, έχουμε την
$$\neg Mortal \vee Mammal.$$
2. Από την πρόταση $\neg Mortal \vee Mammal$ και την $Mortal \vee Horned$, έχουμε την
$$Mammal \vee Horned.$$
3. Από την πρόταση $Mammal \vee Horned$, την $\neg Mortal \vee \neg Mammal \vee Horned$ και παραγοντοποίηση, έχουμε την
$$\neg Mortal \vee Horned.$$

Παράδειγμα: Ανάλυση

4. Από την πρόταση $\neg Mortal \lor Horned$, την $Mortal \lor Horned$ και παραγοντοποίηση, έχουμε την
$$Horned.$$
5. Από την πρόταση $Horned$ και την $\neg Horned \lor Magical$, έχουμε την
$$Magical.$$
6. Από τα λεκτικά $Magical$ και $\neg Magical$, έχουμε την κενή φράση (**αντίφαση**).

Ορθότητα και Πληρότητα του Κανόνα της Ανάλυσης

Ορθότητα. Έστω η βάση γνώσης KB. Αν η ϕ μπορεί να αποδειχθεί από την KB χρησιμοποιώντας ανάλυση, τότε $KB \models \phi$.

Πληρότητα Διάψευσης. Αν ένα σύνολο φράσεων KB είναι μη ικανοποιήσιμο, τότε υπάρχει μια απόδειξη της κενής φράσης από την KB χρησιμοποιώντας ανάλυση.

Σημείωση: Την απόδειξη αυτή μπορεί να τη βρεί οποιοσδήποτε πλήρης αλγόριθμος αναζήτησης που χρησιμοποιεί τον κανόνα της ανάλυσης.

Ικανοποιησιμότητα και CSPs

Το πρόβλημα της ικανοποιησιμότητας για την προτασιακή λογική είναι θεμελιώδες. Όπως είδαμε παραπάνω, τα προβλήματα της λογικής κάλυψης και της εγκυρότητας μπορούν να αναδιατυπωθούν ως προβλήματα ικανοποιησιμότητας.

Το πρόβλημα της ικανοποιησιμότητας για την προτασιακή λογική μπορεί να διατυπωθεί ως **πρόβλημα ικανοποίησης περιορισμών (CSP)**. Ποιες είναι οι μεταβλητές, τα πεδία και οι περιορισμοί;

Πολυπλοκότητα Ικανοποιησιμότητας

Θεώρημα. Το πρόβλημα του να αποφανθούμε αν μια πρόταση της προτασιακής λογικής είναι ικανοποιήσιμη ή όχι είναι NP-complete (Cook, 1971).

Το πρόβλημα του παραπάνω θεωρήματος είναι γνωστό σαν **SAT** (satisfiability).

Πόρισμα. Το πρόβλημα του να αποφανθούμε αν μια πρόταση της προτασιακής λογικής είναι έγκυρη ή όχι είναι co-NP-complete.

Συνεπώς οι πιθανότητες να βρεθούν αλγόριθμοι πολυωνυμικού χρόνου για την επίλυση των παραπάνω προβλημάτων είναι πολύ μικρές.

Αποδοτικοί Αλγόριθμοι

Παρά το παραπάνω φράγμα πολυπλοκότητας, υπάρχουν **αποδοτικοί αλγόριθμοι** για το πρόβλημα της ικανοποιησιμότητας και υλοποιήσεις τους σε διάφορους **SAT solvers**.

Δείτε για παράδειγμα τους αλγόριθμους DPLL και WalkSat από το βιβλίο AIMA, και τον solver SAT4J (`http://www.sat4j.org/`) που θα χρησιμοποιήσουμε στις ασκήσεις.

Φράσεις Horn

Ορισμός. Μια φράση **Horn** είναι μια φράση που έχει το πολύ ένα θετικό λεκτικό.

Δηλαδή, μια φράση Horn απαντάται στις εξής μορφές:

$$Q$$

$$\neg P_1 \vee \neg P_2 \vee \ldots \vee \neg P_n \vee Q \quad (\text{ισοδύναμα:} \quad P_1 \wedge P_2 \wedge \ldots \wedge P_n \Rightarrow Q)$$

$$\neg P_1 \vee \neg P_2 \vee \ldots \vee \neg P_n$$

όπου $P_1, \ldots, P_n, Q$ είναι προτασιακά σύμβολα.

Θεώρημα. Αν η ϕ είναι σύζευξη φράσεων Horn τότε η ικανοποιησιμότητα της ϕ μπορεί να αποφασιστεί σε πολυωνυμικό χρόνο.

Ανακεφαλαίωση

Έστω KB ένα σύνολο προτάσεων της προτασιακής λογικής και ϕ μια πρόταση της προτασιακής λογικής. Πως μπορούμε να αποφασίσουμε αν $KB \models \phi$;

- Μπορούμε να χρησιμοποιήσουμε τους σημασιολογικούς ορισμούς που δώσαμε, και να κάνουμε τη σχετική απόδειξη.
- Μπορούμε να χρησιμοποιήσουμε **πίνακες αληθείας**.
- Μπορούμε να εφαρμόσουμε τους **κανόνες συμπερασμού** που δώσαμε παραπάνω (π.χ., ανάλυση).
- Μπορούμε να χρησιμοποιήσουμε ένα **SAT solver**.

Αδυναμίες της Προτασιακή Λογικής

Θεωρείστε τον εξής κανόνα για τον κόσμο του Wumpus:

> Αν σ΄ ένα τετραγωνάκι δε γίνεται αντιληπτή μυρωδιά τότε ούτε στο ίδιο το τετραγωνάκι ούτε σε γειτονικό του δεν μπορεί να είναι ο Wumpus.

Πώς μπορούμε να αναπαραστήσουμε αυτόν τον κανόνα σε προτασιακή λογική;

Πρέπει να γράφουμε **έναν κανόνα για κάθε σχετικό τετραγωνάκι**! Για παράδειγμα:

$$\neg S_{11} \Rightarrow \neg W_{11} \wedge \neg W_{12} \wedge \neg W_{21}$$

Αδυναμίες της Προτασιακή Λογικής

- Δεν υπάρχει τρόπος στην προτασιακή λογική να εκφράσουμε κάτι για όλα τα αντικείμενα ενός συγκεκριμένου είδους (π.χ., για κάθε τετραγωνάκι).
- Δεν υπάρχει τρόπος να μιλήσουμε για ένα συγκεκριμένο αντικείμενο, τις ιδιότητες του και τις σχέσεις του με άλλα αντικείμενα.
- Δεν υπάρχει τρόπος να μιλήσουμε για την ύπαρξη (ή την μη ύπαρξη) ενός αντικειμένου με συγκεκριμένες ιδιότητες.

Μην ανησυχείτε! Αυτά μπορούμε να τα εκφράσουμε στην λογική πρώτης τάξης.

Πράκτορας Βασισμένος στη Γνώση

function PROPOSITIONAL-KB-AGENT(*percept*) **returns** an *action*
  static KB, βάση γνώσης
    t, μετρητής για το χρόνο, αρχικά 0

  TELL(KB,MAKE-PERCEPT-SENTENCE($percept, t$))
  for each *action* in the list of possible actions **do**
    if ASK(KB,MAKE-ACTION-QUERY($t, action$)) **then**
      TELL(KB,MAKE-ACTION-SENTENCE($action, t$))
      $t \leftarrow t + 1$
      return *action*
  end

Μελέτη

Κεφάλαιο 7 από το βιβλίο AIMA: Λογικοί Πράκτορες (Ενότητες 7.1 έως 7.5 και 7.7).