Αναπαράσταση Γνώσης και Συλλογιστική στην Τεχνητή Νοημοσύνη – Κατηγορηματική Λογική Πρώτης Τάξης

Παναγιώτης Σταματόπουλος
Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Γενικά περί κατηγορηματικής λογικής

- Μέθοδος αναπαράστασης γνώσης – η πιο κλασική στην Τεχνητή Νοημοσύνη
- Αναφέρεται και σαν πρώτης τάξης ή πρωτοβάθμια λογική
- Διατύπωση γνώσης σε μορφή καλοσχηματισμένων προτάσεων
- Μετασχηματισμός προτάσεων σε συζευκτική κανονική μορφή
- Δυνατότητα εξαγωγής συμπερασμάτων – νέας γνώσης – μέσω ενός μηχανισμού αναγωγής
- Αναγωγή συνήθως διά της αντίκρουσης της αντίφασης
- Υποσύνολό της έχει υλοποιηθεί σαν μία γλώσσα προγραμματισμού (Prolog)

Ένα απλό παράδειγμα

Ο Παύλος είναι τραπεζίτης

κατηγόρημα σταθερά

$$\underbrace{banker(Paul)}_{\text{ατομική πρόταση}}$$

Τα βασικά συστατικά

- Οι **σταθερές** παριστάνουν στοιχειώδεις οντότητες, αλλά και αντικείμενα ή αφηρημένες έννοιες, του κόσμου, π.χ. $Paul$, $this_table$, $happiness$, 2562
- Τα **κατηγορήματα** συνοδεύονται και από τον **βαθμό** ή **τάξη** τους, δηλαδή το πλήθος των ορισμάτων που δέχονται – το κατηγόρημα $banker$ έχει βαθμό 1
- Ένα κατηγόρημα βαθμού 1 αποδίδει μία ιδιότητα στο όρισμά του, π.χ. $banker(Paul)$, ορίζοντας μία **ατομική πρόταση**
- Οι ατομικές προτάσεις είναι η στοιχειωδέστερη μορφή **καλοσχηματισμένων προτάσεων** (στο εξής, προτάσεων)
- Μία πρόταση, άρα και μία ατομική πρόταση, έχει μία **τιμή αλήθειας**, δηλαδή μπορεί να είναι **αληθής** (T) ή **ψευδής** (F) σε μία συγκεκριμένη κατάσταση του κόσμου

Βαθμοί κατηγορημάτων

- Κατηγορήματα βαθμού 2 ή μεγαλύτερου εκφράζουν μία σχέση μεταξύ των ορισμάτων τους
- Κατηγορήματα βαθμού 0 εκφράζουν μία πιθανή κατάσταση του κόσμου
- Παραδείγματα:

Η Άννα είναι μητέρα του Παύλου

$mother(Ann, Paul)$

κατηγορήματα σταθερές

Κάνει ζέστη

it_is_hot

- Όταν χρησιμοποιούνται μόνο κατηγορήματα βαθμού 0, έχουμε μία εξειδίκευση της κατηγορηματικής λογικής, την **προτασιακή λογική**

Προτάσεις

- Πιο σύνθετες προτάσεις (μη ατομικές) κατασκευάζονται με τη βοήθεια **λογικών συνδετικών**, τα οποία, σε φθίνουσα προτεραιότητα, είναι τα εξής:

$\neg$	άρνηση
$\wedge$	σύζευξη
$\vee$	διάζευξη
$\Rightarrow$	συνεπαγωγή
$\Leftrightarrow$	ισοδυναμία

- Παραδείγματα:

$$\neg banker(George)$$
$$mother(Ann, Paul) \wedge \neg father(Paul, George)$$
$$man(Socrates) \Rightarrow mortal(Socrates)$$
$$happy(Mary) \Leftrightarrow healthy(Mary) \wedge \neg poor(Mary)$$

Τι σημαίνουν τα παραπάνω σε φυσική γλώσσα;

Τιμή αλήθειας προτάσεων

- Η τιμή αλήθειας προτάσεων που δεν είναι ατομικές ορίζεται σύμφωνα με τον παρακάτω πίνακα, με βάση τα λογικά συνδετικά που περιέχουν και τις τιμές αλήθειας των επί μέρους προτάσεων από τις οποίες αποτελούνται

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \Leftrightarrow q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

- Επισημαίνεται η τιμή αλήθειας της συνεπαγωγής – είναι ψευδής όταν η υπόθεση είναι αληθής και το συμπέρασμα ψευδές, και μόνο τότε

Συναρτήσεις – Όροι

- Με τη βοήθεια **συναρτήσεων** κατασκευάζονται **όροι**, που αναπαριστούν με έμμεσο τρόπο οντότητες, έννοιες, κλπ.
- Παραδείγματα όρων:

$$father_of(Paul)$$
$$next(0)$$
$$date(10, nov, 2013)$$
$$number_of_legs(my_table)$$
$$mother_of(father_of(Ann))$$

- Και οι συναρτήσεις έχουν βαθμό/τάξη, όπως και τα κατηγορήματα
- Οι σταθερές είναι η ειδική περίπτωση συναρτήσεων με βαθμό 0
- <u>ΠΡΟΣΟΧΗ:</u> Οι συναρτήσεις **ΔΕΝ** είναι κατηγορήματα και οι όροι **ΔΕΝ** είναι ατομικές προτάσεις

Μεταβλητές – Ποσοδείκτες

- Στην κατηγορηματική λογική χρησιμοποιούνται **μεταβλητές** για να αναπαρασταθούν τυχαίες οντότητες, αντικείμενα, κλπ., του κόσμου
- Οι μεταβλητές εισάγονται στις προτάσεις μέσω των ποσοδεικτών:

Καθολικός ποσοδείκτης: ∀ (για κάθε)

Υπαρξιακός ποσοδείκτης: ∃ (υπάρχει)

- Παραδείγματα:

Τα πάντα είναι θνητά

$\forall x\ [mortal(x)]$

Υπάρχει κάποιος άνθρωπος

$\exists x\ [man(x)]$

Καθολικός ποσοδείκτης

- Κατά κανόνα, εφαρμόζεται σε μία συνεπαγωγή, για να διατυπωθεί ένας γενικός νόμος του κόσμου, ποσοτικοποιώντας καθολικά όλες τις μεταβλητές που περιέχονται στη συνεπαγωγή
- Παραδείγματα:

Οι άνθρωποι είναι θνητοί

$$\forall x\ [man(x) \Rightarrow mortal(x)]$$

Οι τραπεζίτες είναι πλούσιοι

$$\forall x\ [banker(x) \Rightarrow rich(x)]$$

Οι γονείς αγαπούν τα παιδιά τους

$$\forall x\ \forall y\ [parent(x, y) \Rightarrow loves(x, y)]$$

Υπαρξιακός ποσοδείκτης

- Κατά κανόνα, εφαρμόζεται σε μία σύζευξη, για να διατυπωθεί δήλωση ύπαρξης οντοτήτων, αντικειμένων, κλπ., του κόσμου, ποσοτικοποιώντας υπαρξιακά όλες τις μεταβλητές που περιέχονται στη σύζευξη
- Παραδείγματα:

Κάποιοι ερευνητές είναι πλούσιοι

$$\exists x\ [researcher(x) \wedge rich(x)]$$

Κάποιες χώρες συνορεύουν με την Ελλάδα και τη Ρουμανία

$$\exists x\ [country(x) \wedge borders(x, Greece) \wedge borders(x, Romania)]$$

Κάθε άνθρωπος έχει μητέρα μεγαλύτερη από αυτόν

$$\forall x\ [man(x) \Rightarrow \exists y\ [mother(y, x) \wedge elder(y, x)]]$$

Ισότητα

- Πολύ συχνά, χρειαζόμαστε στην κατηγορηματική λογική το κατηγόρημα της ισότητας =, ή της άρνησής της ≠, που χρησιμοποιούνται συνήθως με ενδοθεματική σύνταξη
- Παραδείγματα:

Κάθε άνθρωπος έχει ακριβώς έναν πατέρα

$$\forall x \left[man(x) \Rightarrow \exists y \left[father(y,x) \wedge \forall z \left[father(z,x) \Rightarrow z = y \right] \right] \right]$$

Υπάρχουν τουλάχιστον δύο ελέφαντες

$$\exists x \left[elephant(x) \wedge \exists y \left[elephant(y) \wedge y \neq x \right] \right]$$

Υπάρχουν ακριβώς δύο γάτες

$$\exists x \left[cat(x) \wedge \exists y \left[cat(y) \wedge y \neq x \wedge \forall z \left[cat(z) \Rightarrow z = x \vee z = y \right] \right] \right]$$

Ένα πιο σύνθετο παράδειγμα

Ο Παύλος είναι τραπεζίτης

$banker(Paul)$

Οι τραπεζίτες είναι πλούσιοι

$\forall x\ [banker(x) \Rightarrow rich(x)]$

Καθένας έχει κάποια μητέρα

$\forall x\ \exists y\ [mother(y, x)]$

Οι μητέρες των πλουσίων είναι ευτυχισμένες

$\forall x\ \forall y\ [mother(y, x) \wedge rich(x) \Rightarrow happy(y)]$

Οι ευτυχισμένες μητέρες αγαπούν τη μητέρα τους

$\forall x\ \left[happy(x) \wedge \exists y\ [mother(x, y)] \Rightarrow \forall z\ [mother(z, x) \Rightarrow loves(x, z)]\right]$

Μετασχηματισμός σε συζευκτική κανονική μορφή

- Οι ατομικές προτάσεις ή οι αρνήσεις ατομικών προτάσεων ονομάζονται **κυριολεκτικά**
- Μία πρόταση σε συζευκτική κανονική μορφή (ΣΚΜ) είναι μία σύζευξη διαζεύξεων κυριολεκτικών, όπου όλες οι μεταβλητές θεωρούνται ότι είναι καθολικά ποσοτικοποιημένες στην αρχή της πρότασης, γι' αυτό και η ποσοτικοποίηση παραλείπεται
- Μορφή κατάλληλη για την εφαρμογή της διαδικασίας αναγωγής με σκοπό την εξαγωγή συμπερασμάτων
- Βασικές ταυτότητες που χρησιμοποιούνται:

$\neg(\neg p) \equiv p$	$\neg(p \wedge q) \equiv \neg p \vee \neg q$	$\neg(p \vee q) \equiv \neg p \wedge \neg q$
$p \Rightarrow q \equiv \neg p \vee q$	$\neg\exists x\,[p(x)] \equiv \forall x\,[\neg p(x)]$	$\neg\forall x\,[p(x)] \equiv \exists x\,[\neg p(x)]$
$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$	$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	

Παραδείγματα μετασχηματισμού

- Το $banker(Paul)$ είναι ήδη σε ΣΚΜ
- Το $\forall x\ [banker(x) \Rightarrow rich(x)]$ γίνεται σε ΣΚΜ:

$$\forall x\ [\neg banker(x) \vee rich(x)]$$
$$\neg banker(x) \vee rich(x)$$

- Το $\forall x\ \exists y\ [mother(y, x)]$ χρειάζεται Skolemization για την απαλοιφή του υπαρξιακού ποσοδείκτη, οπότε έχουμε:

$$\forall x\ [mother(mother_of(x), x)]$$
$$mother(mother_of(u), u)$$

<u>!!! Προσοχή στη διαφορά του κατηγορήματος $mother$ από τη συνάρτηση $mother_of$!!!</u>

- Το $\forall x\ \forall y\ [mother(y, x) \wedge rich(x) \Rightarrow happy(y)]$ γίνεται:

$$\forall x\ \forall y\ [\neg(mother(y, x) \wedge rich(x)) \vee happy(y)]$$
$$\forall x\ \forall y\ [\neg mother(y, x) \vee \neg rich(x) \vee happy(y)]$$
$$\neg mother(y, w) \vee \neg rich(w) \vee happy(y)$$

Παραδείγματα μετασχηματισμού (συνέχεια)

- Και η τελευταία πρόταση:

$$\forall x\ [happy(x) \land \exists y\ [mother(x,y)] \Rightarrow \forall z\ [mother(z,x) \Rightarrow loves(x,z)]]$$
$$\forall x\ [\neg(happy(x) \land \exists y\ [mother(x,y)]) \lor \forall z\ [mother(z,x) \Rightarrow loves(x,z)]]$$
$$\forall x\ [\neg(happy(x) \land \exists y\ [mother(x,y)]) \lor \forall z\ [\neg mother(z,x) \lor loves(x,z)]]$$
$$\forall x\ [\neg happy(x) \lor \neg\exists y\ [mother(x,y)] \lor \forall z\ [\neg mother(z,x) \lor loves(x,z)]]$$
$$\forall x\ [\neg happy(x) \lor \forall y\ [\neg mother(x,y)] \lor \forall z\ [\neg mother(z,x) \lor loves(x,z)]]$$
$$\forall x \forall y \forall z\ [\neg happy(x) \lor \neg mother(x,y) \lor \neg mother(z,x) \lor loves(x,z)]$$
$$\neg happy(t) \lor \neg mother(t,v) \lor \neg mother(z,t) \lor loves(t,z)$$

- Τελικά, η γνώση μας σε ΣΚΜ είναι:

$$banker(Paul)$$
$$\neg banker(x) \lor rich(x)$$
$$mother(mother_of(u),u)$$
$$\neg mother(y,w) \lor \neg rich(w) \lor happy(y)$$
$$\neg happy(t) \lor \neg mother(t,v) \lor \neg mother(z,t) \lor loves(t,z)$$

Κάποιες παρατηρήσεις

- Τι διαφορά έχουν τα εξής;

$$\forall x\ \exists y\ [mother(y,x)] \text{ και } \exists y\ \forall x\ [mother(y,x)]$$

Μετά από Skolemization γίνονται, αντίστοιχα:

$$\forall x\ [mother(mother_of(x),x)] \text{ και } \forall x\ [mother(universal_mother,x)]$$

Δηλαδή:

«Οποιοσδήποτε έχει κάποια μητέρα» (συνάρτηση του εαυτού του)

«Όλοι έχουν την ίδια μητέρα» (σταθερά, ίδια για όλους)

- Όταν το αποτέλεσμα μετασχηματισμού σε ΣΚΜ δεν είναι ένα κυριολεκτικό ή μία διάζευξη, αλλά σύζευξη διαζεύξεων, παράγεται μία πρόταση για κάθε συζευκτέο, π.χ.

$$(\neg A \vee \neg B \vee C) \wedge (A \vee B) \wedge (B \vee \neg C \vee \neg D)$$

$$\Downarrow$$

$$\left\{ \begin{array}{c} \neg A \vee \neg B \vee C \\ A \vee B \\ B \vee \neg C \vee \neg D \end{array} \right\}$$

Αναγωγή και ενοποίηση

- Έστω ότι δίνονται οι προτάσεις:

$$A_1 \vee B \qquad \text{και} \qquad \neg A_2 \vee C$$

- Αν οι ατομικές προτάσεις A_1 και A_2 είναι **ενοποιήσιμες**, δηλαδή με κατάλληλες αντικαταστάσεις τιμών στις μεταβλητές τους να μπορούν να ταυτιστούν (έστω σ το σύνολο των αντικαταστάσεων), τότε, σύμφωνα με τον κανόνα συλλογισμού της **αναγωγής**, παράγεται η πρόταση:

$$(B \vee C)\sigma$$

- Ο παραπάνω συμβολισμός σημαίνει ότι εφαρμόζεται το σύνολο των αντικαταστάσεων σ στην πρόταση $B \vee C$
- Παράδειγμα: Από τις προτάσεις

$$\neg p(x) \vee q(x) \qquad \text{και} \qquad p(a) \vee r(b)$$

παράγεται η πρόταση $q(a) \vee r(b)$, με $\sigma = \{a/x\}$

- Το a/x σημαίνει ότι στη μεταβλητή x ανατίθεται η τιμή a

Παραγωγή νέας γνώσης – Συμπερασματολογία

- Δεδομένη γνώση σε ΣΚΜ:

$$banker(Paul)$$
$$\neg banker(x) \vee rich(x)$$
$$mother(mother_of(u), u)$$
$$\neg mother(y, w) \vee \neg rich(w) \vee happy(y)$$
$$\neg happy(t) \vee \neg mother(t, v) \vee \neg mother(z, t) \vee loves(t, z)$$

- Αποδεικτέο:

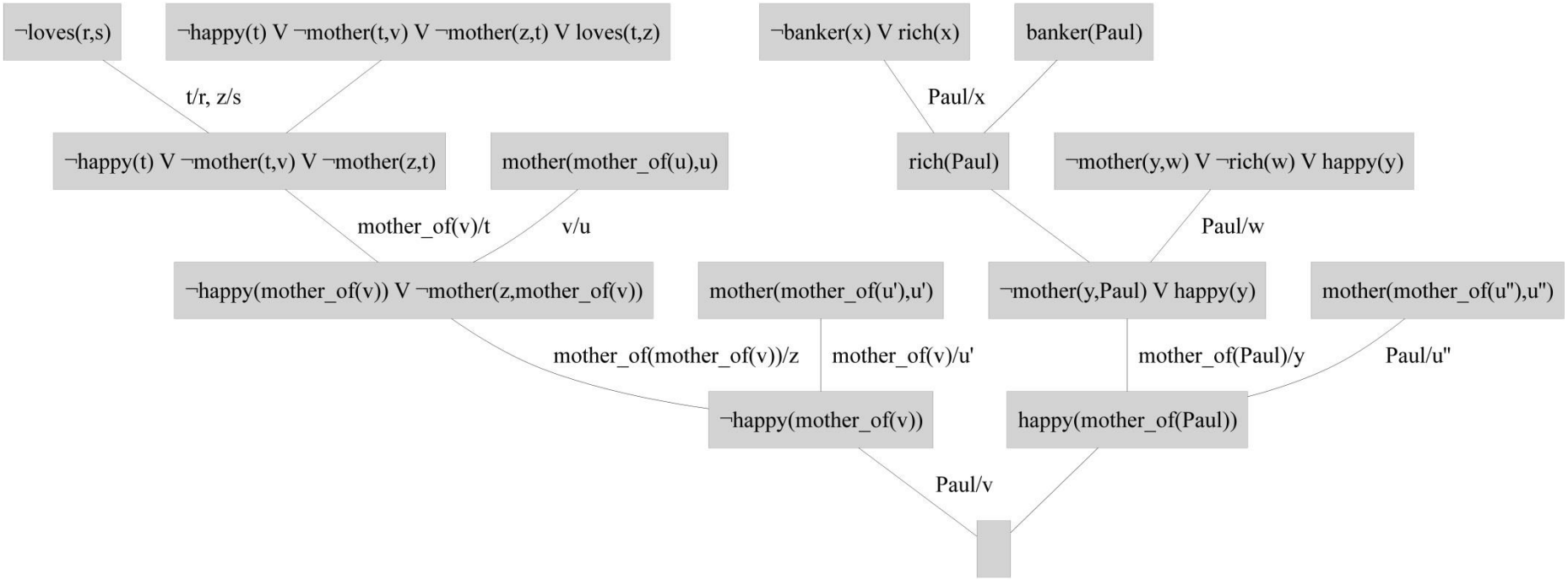
Αγαπά κάποιος κάποιον;

$$\exists r \exists s\ [loves(r, s)]$$

- Άρνηση αποδεικτέου σε ΣΚΜ:

$$\neg \exists r \exists s\ [loves(r, s)]$$
$$\forall r\ \left[\neg \exists s\ [loves(r, s)]\right]$$
$$\forall r \forall s\ [\neg loves(r, s)]$$
$$\neg loves(r, s)$$

Αναγωγή μέσω αντίκρουσης της αντίφασης



- Από τη σύνθεση των αντικαταστάσεων t/r, $mother_of(v)/t$, $Paul/v$ και z/s, $mother_of(mother_of(v))/z$, $Paul/v$, έχουμε τελικά $mother_of(Paul)/r$ και $mother_of(mother_of(Paul))/s$, δηλαδή *«η μητέρα του Παύλου αγαπά τη μητέρα της»*

Και άλλα παραδείγματα

- Όλα τα μανιτάρια είναι δηλητηριώδη

$$\forall x\ [mushroom(x) \Rightarrow poisonous(x)]$$

- Κάποιο μανιτάρι είναι δηλητηριώδες

$$\exists x\ [mushroom(x) \wedge poisonous(x)]$$

- Όλα τα μανιτάρια εκτός από ένα είναι δηλητηριώδη

$$\exists x\ \big[mushroom(x) \wedge \neg poisonous(x)$$
$$\wedge\ \forall y\ [mushroom(y) \wedge x \neq y \Rightarrow poisonous(y)]\big]$$

- Υπάρχει ακριβώς ένα μανιτάρι

$$\exists x\ [mushroom(x) \wedge \forall y\ [mushroom(y) \Rightarrow x = y]]$$

- Υπάρχουν τουλάχιστον δύο μανιτάρια

$$\exists x\ [mushroom(x) \wedge \exists y\ [mushroom(y) \wedge x \neq y]]$$

- Υπάρχουν ακριβώς δύο δηλητηριώδη μανιτάρια

$$\exists x\ [mushroom(x) \wedge poisonous(x) \wedge \exists y\ [mushroom(y) \wedge poisonous(y)$$
$$\wedge\ x \neq y \wedge \forall z\ [mushroom(z) \wedge poisonous(z) \Rightarrow z = x \vee z = y]]]$$

Και άλλα παραδείγματα (συνέχεια)

- Κάποιος άντρας αγαπά κάθε γυναίκα που μισεί τον Αντώνη

  $\exists x\ [man(x) \wedge \forall y\ [woman(y) \wedge hates(y, \text{Αντώνης}) \Rightarrow loves(x, y)]]$

- Κάθε άντρας αγαπά κάποια γυναίκα που μισεί τον Αντώνη

  $\forall x\ [man(x) \Rightarrow \exists y\ [woman(y) \wedge hates(y, \text{Αντώνης}) \wedge loves(x, y)]]$

- Ένα σύνολο είναι κενό αν και μόνο αν δεν περιέχει κανένα στοιχείο

  $$\forall s\ [empty(s) \Leftrightarrow \neg\exists x\ [in(x, s)]]$$

- Ένα στοιχείο ανήκει στη διαφορά δύο συνόλων αν και μόνο αν ανήκει στο πρώτο σύνολο και δεν ανήκει στο δεύτερο

  $$\forall x\ \forall s_1 \forall s_2\ [in\big(x, diff(s_1, s_2)\big) \Leftrightarrow in(x, s_1) \wedge \neg in(x, s_2)]$$

Παράδειγμα αναγωγής

Ο Sam, ο Clyde και ο Oscar είναι ελέφαντες. Ο Sam είναι ροζ. Ο Clyde είναι γκρι και συμπαθεί τον Oscar. Ο Oscar είναι ή ροζ ή γκρι και συμπαθεί τον Sam. Αποδείξτε ότι κάποιος γκρι ελέφαντας συμπαθεί ένα ροζ ελέφαντα.

$$elephant(Sam)$$
$$elephant(Clyde)$$
$$elephant(Oscar)$$
$$pink(Sam)$$
$$grey(Clyde) \wedge likes(Clyde, Oscar)$$
$$[pink(Oscar) \vee grey(Oscar)] \wedge likes(Oscar, Sam)$$

Αποδεικτέο:

$$\exists x\ [elephant(x) \wedge grey(x) \wedge \exists y\ [elephant(y) \wedge pink(y) \wedge likes(x, y)]]$$

Παράδειγμα αναγωγής (συνέχεια)

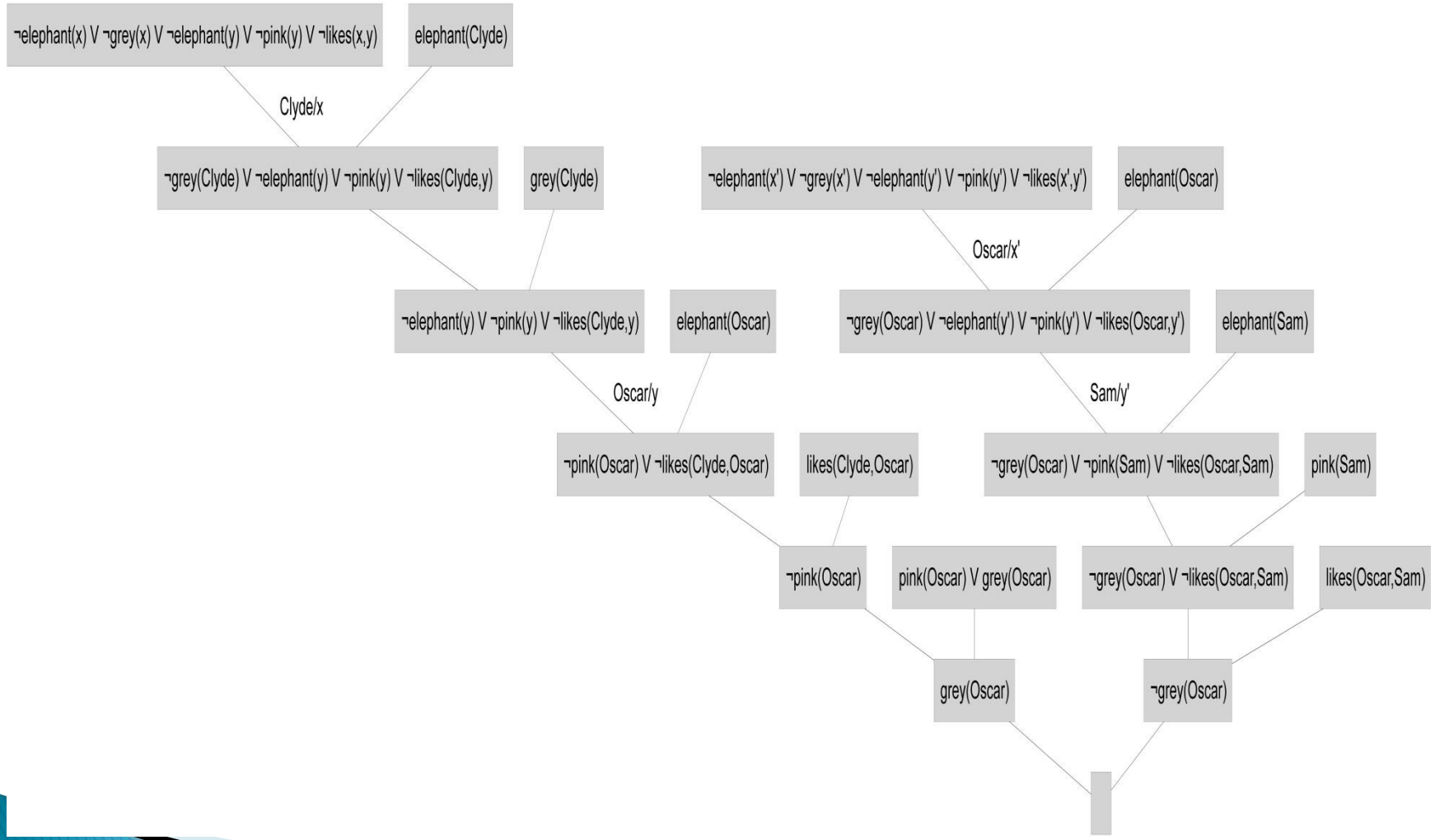
Συζευκτική κανονική μορφή (ΣΚΜ):

$$elephant(Sam)$$
$$elephant(Clyde)$$
$$elephant(Oscar)$$
$$pink(Sam)$$
$$grey(Clyde)$$
$$likes(Clyde, Oscar)$$
$$pink(Oscar) \vee grey(Oscar)$$
$$likes(Oscar, Sam)$$

Άρνηση αποδεικτέου:

$$\neg elephant(x) \vee \neg grey(x) \vee \neg elephant(y) \vee \neg pink(y) \vee \neg likes(x, y)$$

Παράδειγμα αναγωγής (συνέχεια)



Άλλο ένα παράδειγμα

Στην πόλη έγινε ένας φόνος. Το θύμα ήταν ο Victor. Η αστυνομία συνέλαβε τρεις υπόπτους, τον Abbott, τον Babbitt και τον Cabot. Ο Abbott ισχυρίζεται ότι ο Babbitt ήταν φίλος του Victor και ότι ο Cabot μισούσε τον Victor. Ο Babbitt αρνήθηκε ότι ήταν στην πόλη την ημέρα του φόνου και επίσης είπε ότι δεν γνώριζε τον Victor. Ο Cabot κατέθεσε ότι είδε τον Abbott και τον Babbitt να είναι με τον Victor λίγο πριν το έγκλημα. Η αστυνομία είναι βέβαιη ότι ακριβώς ένας από τους Abbott, Babbitt και Cabot είναι ο ένοχος. Επίσης, υποθέτει ότι οι δύο που είναι αθώοι λένε την αλήθεια. Ποιος είναι ο δολοφόνος;

Αναπαράσταση σε κατηγορηματική λογική

$$innocent(A) \rightarrow friend(B, V)$$
$$innocent(A) \rightarrow hates(C, V)$$

$$innocent(B) \rightarrow \neg in_town(B)$$
$$innocent(B) \rightarrow \neg knows(B, V)$$

$$innocent(C) \rightarrow with(A, V)$$
$$innocent(C) \rightarrow with(B, V)$$

$$\forall x\ [with(x, V) \rightarrow in_town(x)]$$
$$\forall x\ \forall y\ [friend(x, y) \rightarrow knows(x, y)]$$
$$\forall x\ \forall y\ [hates(x, y) \rightarrow knows(x, y)]$$

$$\neg innocent(A)\ \vee \neg innocent(B) \vee \neg innocent(C)$$
$$(innocent(A) \wedge\ innocent(B)) \vee (innocent(A) \wedge innocent(C)) \vee (innocent(B) \wedge innocent(C))$$

Αποδεικτέο:

$$\exists x\ [\neg innocent(x)]$$

Συζευκτική κανονική μορφή (ΣΚΜ)

$$\neg innocent(A) \vee friend(B,V)$$
$$\neg\ innocent(A) \vee hates(C,V)$$

$$\neg\ innocent(B) \vee \neg in_town(B)$$
$$\neg innocent(B) \vee \neg knows(B,V)$$

$$\neg innocent(C) \vee with(A,V)$$
$$\neg\ innocent(C) \vee with(B,V)$$

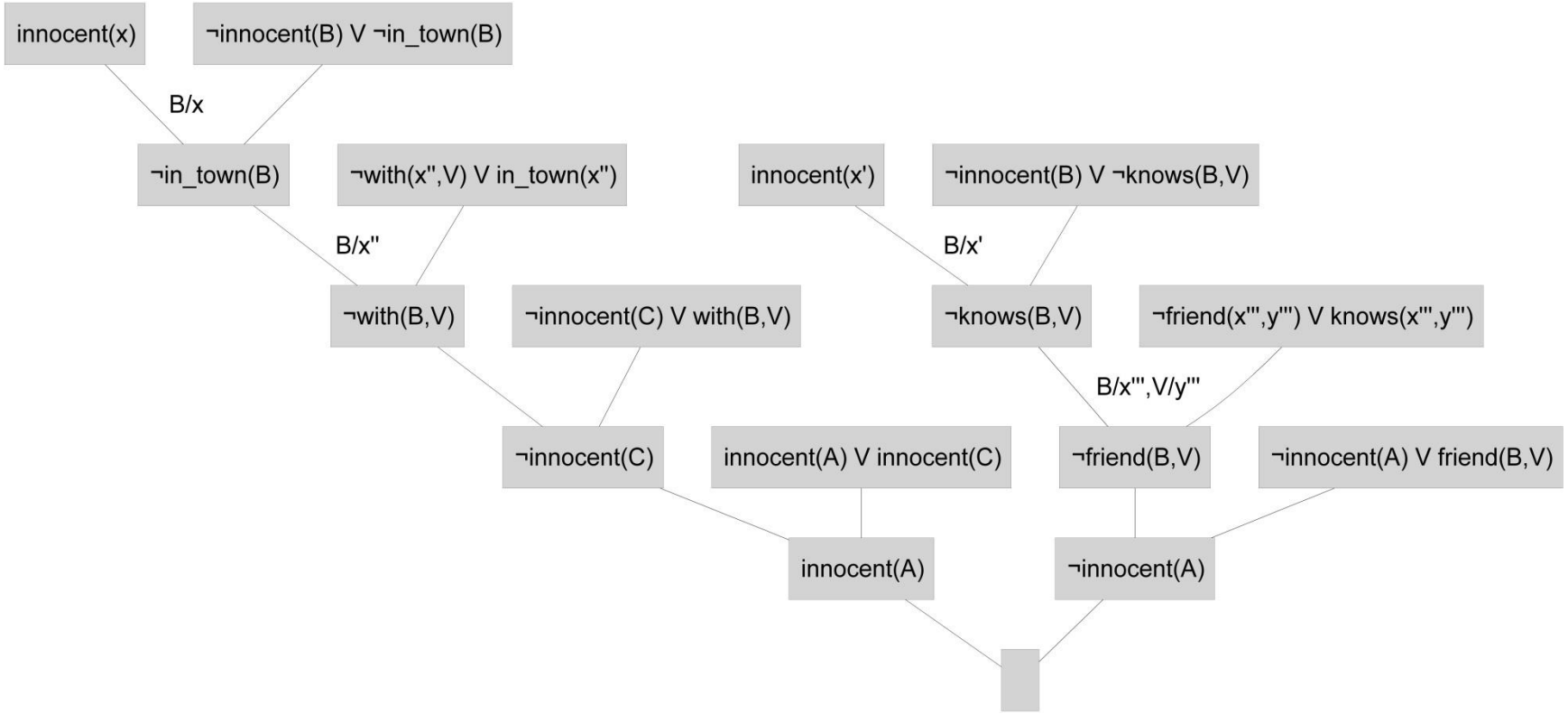
$$\neg\ with(x,V) \vee in_town(x)$$
$$\neg friend(x,y) \vee knows(x,y)$$
$$\neg\ hates(x,y) \vee knows(x,y)$$

$$\neg innocent(A)\ \vee \neg innocent(B) \vee \neg innocent(C)$$
$$innocent(A) \vee\ innocent(B)$$
$$innocent(A) \vee\ innocent(C)$$
$$innocent(B) \vee\ innocent(C)$$

Άρνηση αποδεικτέου:

$$innocent(x)$$

Αναγωγή



Άρα, ο δολοφόνος είναι ο Babbitt!