ΣΤΡΑΤΙΩΤΙΚΗ ΣΧΟΛΗ ΕΥΕΛΠΙΔΩΝ
Τμήμα Στρατιωτικών Επιστημών

Ε Λ Λ Η Ν Ι Κ Η Δ Η Μ Ο Κ Ρ Α Τ Ι Α
ΔΙΙΔΡΥΜΑΤΙΚΟ ΔΙΑΤΜΗΜΑΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΑΚΑΔΗΜΑΪΚΟΥ ΕΤΟΥΣ 2018-20

ΕΦΑΡΜΟΣΜΕΝΗ
ΕΠΙΧΕΙΡΗΣΙΑΚΗ ΕΡΕΥΝΑ & ΑΝΑΛΥΣΗ

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
Σχολή Μηχανικών Παραγωγής & Διοίκησης

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΕΠΑΓΩΓΙΚΗ ΑΛΓΟΡΙΘΜΙΚΗ ΓΝΩΣΗ

Διατριβή που υπεβλήθη για την μερική ικανοποίηση των απαιτήσεων για την απόκτηση Μεταπτυχιακού Διπλώματος Ειδίκευσης

Υπό:

ΓΚΟΥΝΤΟΥΡΑ ΚΟΣΜΑ

Α.Μ.: 2018018016

ΑΠΡΙΛΙΟΣ 2021

ΣΕΛΙΔΑ ΣΚΟΠΙΜΑ ΚΕΝΗ

Η Μεταπτυχιακή Διατριβή του Γκουντούρα Κοσμά εγκρίνεται:

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Καθηγητής Δάρας Νικόλαος (Επιβλέπων) ……

Καθηγητής Ματσατσίνης Νικόλαος ……..

Nikolaos Matsatsinis
Digitally signed by Nikolaos Matsatsinis
Date: 2021.04.03 09:23:51 +03'00'

Αναπληρωτής Καθηγητής Παπαδάκης Νικόλαος …Νικόλαος Παπαδάκης
..

ΣΕΛΙΔΑ ΣΚΟΠΙΜΑ ΚΕΝΗ

Έτος 2021

Η φαντασία είναι πιο σημαντική από τη γνώση. Η γνώση είναι περιορισμένη. Η φαντασία περικλείει τον κόσμο.

"Albert Einstein"

ΣΕΛΙΔΑ ΣΚΟΠΙΜΑ ΚΕΝΗ

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ……9

ΚΕΦΑΛΑΙΟ 1Ο

1. Εισαγωγή……10

1.1 Η Έννοια της Λογικής……..10

1.2 Παραγωγικός και Επαγωγικός Συλλογισμός – Γνώση…… ..12

1.3 Ιδιότητες Συστημάτων Λογικής……....13

1.4 Προκείμενες……13

1.5 Αξιώματα……14

1.6 Υπολογιστική Λογική…….....15

1.7 Προτασιακή Λογική……16

1.8 Τεχνητή Νοημοσύνη…….....17

1.9 Μηχανική Μάθηση - Αλγοριθμική Γνώση…….....18

ΚΕΦΑΛΑΙΟ 2Ο

2. Επαγωγική Αλγοριθμική Γνώση……...18

2.1 Το Θεωρητικό Πλαίσιο της Επαγωγικής Αλγοριθμικής Γνώσης……....18

2.2 Επαγωγικά Συστήματα…….....25

2.3 Ένα Μοντέλο Επαγωγικής Αλγοριθμικής Γνώσης……29

2.4 Συστήματα Αξιωμάτων……37

2.5 Διαδικασίες Επιβεβαίωσης «Αποφασιστικότητας» Τύπων που Ανήκουν σε Δομές……42

2.6 Απόδοση Συλλογιστικής Πολλών Πρακτόρων στο Ίδιο Σύστημα…….45

ΚΕΦΑΛΑΙΟ 3Ο

3. Επαγωγική Αλγοριθμική Γνώση και Συστήματα Ασφάλειας Δεδομένων……50

3.1 Η Εφαρμογή της Επαγωγικής Αλγοριθμικής Γνώσης στα Συστήματα Διαχείρισης Ασφάλειας Προσωπικών Δεδομένων…….50

3.2 Το Πλαίσιο Ανάλυσης…….53

3.3 Ένα απόσπασμα μιας Προτασιακής Λογικής για την Περιγραφή Συστημάτων Παροχής Ασφάλειας Προσωπικών Δεδομένων…….....57

3.4 Σημασιολογία και Συστήματα Αξιωμάτων……60

3.5 Η Χρησιμότητα του Μοντέλου Τυπικής Λογικής…….62

3.6 Ένα Παράδειγμα Εφαρμογής για Ενίσχυση της Ασφάλειας του Πρωτοκόλλου Επικοινωνίας μεταξύ δύο Συστημάτων……...64

ΚΕΦΑΛΑΙΟ 4Ο

4. Επίλογος…….69

ΠΑΡΑΡΤΗΜΑ «Α»……71

Αστερίσκοι……71

ΒΙΒΛΙΟΓΡΑΦΙΑ……...75

Επιστημονικές Αναφορές……...75

Επιστημονικά Άρθρα……80

ΠΕΡΙΛΗΨΗ

Η επαγωγική αλγοριθμική γνώση αποτελεί πτυχή της επιστήμης της μηχανικής μάθησης, η κατανόηση και χρήση της οποίας δύναται να συνεισφέρει τα μέγιστα στην εξέλιξη των σχετικών με αυτή μηχανισμών. Το πλαίσιο ανάλυσης της αλγοριθμικής γνώσης λαμβάνει ως δεδομένο, ότι οι μηχανισμοί μηχανικής μάθησης χρησιμοποιούν αλγόριθμους, προκειμένου να καταλήξουν σε δεδομένα. Αντί αυτών θα μπορούσαμε να χρησιμοποιήσουμε μία λογική θεωρία, για να αποδώσουμε τη συλλογιστική των εν λόγω μηχανισμών και να υπολογίσουμε τα σχετικά δεδομένα. Με την εργασία αυτή παραθέτουμε ένα μοντέλο τυπικής προτασιακής λογικής προς απόδοση της συλλογιστικής διαδικασίας μηχανισμών μηχανικής μάθησης, η οποία εμπλέκει τόσο την παραγωγική, όσο και την επαγωγική γνώση, όταν η δεύτερη προσδιορίζεται μέσα από ένα επαγωγικό σύστημα. Η εξαιρετικά δομημένη φύση των επαγωγικών συστημάτων, οδηγεί σε αξιώματα της χρησιμοποιούμενης λογικής, μόνο όταν αυτά αποδεικνύονται μέσα σε ένα σταθερό επαγωγικό σύστημα. Η λύση ενός τέτοιου προβλήματος μέσω μιας τέτοιας λογικής, πρέπει να δύναται να διεκπεραιωθεί σε πολυωνυμικό χρόνο (NP "nondeterministic polynomial time"- complete) ***[1] (Παράρτημα «Α»)**

Η λύση του προβλήματος δύναται να διεκπεραιωθεί σε πολυωνυμικό χρόνο (NP-complete), μόνο όταν σε αυτό προσαρμόσουμε ένα επαγωγικό σύστημα, το οποίο και αυτό δύναται να διεκπεραιωθεί, σε πολυωνυμικό χρόνο. Τα αποτελέσματα αυτά ισχύουν με ανάλογο τρόπο και σε περιπτώσεις προβλημάτων με την παρουσία πολλών μηχανισμών ταυτόχρονα στο ίδιο σύστημα. Η εργασία μας ξεκινά με την παράθεση κάποιων βασικών εννοιών και όρων, η κατανόηση των οποίων καθίσταται απαραίτητη για την μετέπειτα κατανόηση της έννοιας της επαγωγικής αλγοριθμικής γνώσης και του πλαισίου ανάλυσής της. Κλείνει με την παρουσίαση μιας εφαρμογής της επαγωγικής αλγοριθμικής γνώσης, για την ενίσχυση της ασφάλειας συστημάτων που διαχειρίζονται προσωπικά δεδομένα.

ΚΕΦΑΛΑΙΟ 1ο

§1. Εισαγωγή

§1.1 Η Έννοια της Λογικής

Προκειμένου να γίνει πιο εύκολα κατανοητό το πλαίσιο ανάλυσης της παρούσας εργασίας, κρίνεται σκόπιμο να παρατεθούν κάποιες έννοιες, οι οποίες σχετίζονται με τη συλλογιστική διαδικασία και τη λογική εν γένει. Η λογική έχει δύο έννοιες. Αρχικά είναι η μελέτη των τρόπων συλλογισμού (εκείνων που ισχύουν, καθώς και των εσφαλμένων), αλλά και η χρήση των έγκυρων συλλογισμών. Στα μαθηματικά είναι η μελέτη των έγκυρων συμπερασμάτων στο πλαίσιο μιας τυπικής γλώσσας. Η λογική συνήθως χωρίζεται σε τρία μέρη, τον παραγωγικό συλλογισμό τον υποθετικό-παραγωγικό συλλογισμό και τον επαγωγικό συλλογισμό.

Η έννοια της λογικής μορφής είναι ουσιαστικής σημασίας για τη λογική και ορίζει, ότι η εγκυρότητα ενός επιχειρήματος καθορίζεται από την λογική του μορφή και όχι από το περιεχόμενο του. Γενικές μορφές λογικής είναι οι παρακάτω:

- **Άτυπη λογική**: είναι η μελέτη των επιχειρημάτων της φυσικής γλώσσας. Η μελέτη των λογικών πλάνων είναι ένας ιδιαίτερα σημαντικός κλάδος της άτυπης λογικής. Οι «διάλογοι του Πλάτωνα» αποτελούν χαρακτηριστικά παραδείγματα άτυπης λογικής.

- **Τυπική λογική:** είναι η μελέτη της συμπερασματολογίας με καθαρά τυπικό περιεχόμενο. Ένα συμπέρασμα διαθέτει ένα καθαρά τυπικό περιεχόμενο, εάν αυτό μπορεί να εκφραστεί ως ιδιαίτερη εφαρμογή ενός εντελώς αφηρημένου κανόνα, ενός κανόνα δηλαδή που δεν είναι για κάποιο συγκεκριμένο πράγμα ή υπάρχον αντικείμενο. Τα έργα του Αριστοτέλη περιέχουν την αρχαιότερη γνωστή μελέτη τυπικής λογικής. Σε πολλούς ορισμούς της λογικής, το λογικό συμπέρασμα και το συμπέρασμα με καθαρά τυπικό

περιεχόμενο είναι το ίδιο. Αυτό δεν καθιστά την έννοια της άτυπης λογικής κενή, καθώς καμία τυπική λογική δεν μπορεί να αποτυπώσει όλες τις διαβαθμίσεις της φυσικής γλώσσας.

- **Συμβολική λογική:** είναι η μελέτη της χρήσης των συμβόλων, τα οποία χρησιμοποιούνται για την απόδοση των τυπικών γνωρισμάτων ενός λογικού συμπεράσματος.

- **Μαθηματική λογική:** είναι η επέκταση της συμβολικής λογικής σε άλλα πεδία, ιδίως στη μελέτη της θεωρίας μοντέλων, της θεωρίας αποδείξεων, της θεωρίας αναδρομής και της θεωρίας συνόλων.

Η λογική είναι εν γένει τυπική, όταν έχει ως στόχο να αναλύσει και να παρουσιάσει τη μορφή οποιουδήποτε έγκυρου τύπου επιχειρήματος. Η μορφή ενός επιχειρήματος εμφανίζεται παρουσιάζοντας τις προτάσεις του στην τυπική γραμματική και τον συμβολισμό μιας λογικής γλώσσας, προκειμένου να φέρουν το περιεχόμενο του σε μορφή που να μπορεί να χρησιμοποιηθεί σε τυπικό συμπέρασμα. Αυτό στην ουσία αποδεικνύει τη λογική μορφή του επιχειρήματος. Καθίσταται απαραίτητη, καθότι πολλές προτάσεις των κοινών γλωσσών (Ordinary Languages), εμφανίζουν ποικίλες ιδιομορφίες, γεγονός που καθιστά τη χρήση τους στην συμπερασματολογία μη πρακτική. Λειτουργεί αγνοώντας τα γραμματικά χαρακτηριστικά, τα οποία δεν σχετίζονται με τη λογική, όπως το γένος (Gender) ή η πτώση (Declension), αντικαθιστώντας συνδέσμους που δεν έχουν σχέση με τη λογική (όπως το "αλλά") με λογικούς συνδέσμους και αντικαθιστώντας διφορούμενες ή εναλλακτικές λογικές εκφράσεις, με εκφράσεις ενός συγκεκριμένου τύπου όπως για παράδειγμα ο καθολικός ποσοδείκτης ∀. Επιπλέον, ορισμένα τμήματα της πρότασης αντικαθίστανται με σύμβολα. Η έννοια της μορφής (The Concept of Form), θεμελιώδης έννοια στην μελέτη της λογικής, ήταν ήδη γνωστή στην αρχαιότητα. Ο Αριστοτέλης χρησιμοποιούσε μεταβλητά γράμματα που παρίσταναν έγκυρα συμπεράσματα στο έργο του «*Αναλυτικά Πρότερα»,* για να έρθει χιλιάδες χρόνια μετά να δηλώσει ο Jan Lucacevic, ότι η εισαγωγή μεταβλητών στον γραπτό λόγο, ήταν «μία από τις μεγαλύτερες εφευρέσεις του Αριστοτέλη». Σύμφωνα με τους ακόλουθους του Αριστοτέλη (όπως ο Αμμώνιος), μόνο οι αρχές της λογικής που αποδίδονται με σχηματικούς όρους - σύμβολα (Schematic Terms) ανήκουν στη λογική και όχι εκείνες που αποδίδονται με λεκτικούς όρους. Όροι όπως «άνδρας», «κτίριο» κ.τ.λ. είναι

αντικαταστάσιμοι από σχηματικούς όρους – σύμβολα όπως 'Α', 'Β', 'Γ', οι οποίοι λέγονται αλλιώς και η «ουσία» του συμπεράσματος.

§1.2 Παραγωγικός και Επαγωγικός Συλλογισμός – Γνώση

Ο παραγωγικός συλλογισμός αφορά σε αυτό που προκύπτει αναγκαστικά από δοθείσες, προκείμενες, προτάσεις (αν Α, τότε Β). Ο επαγωγικός συλλογισμός (η διαδικασία δηλαδή κατά την οποία προκύπτει μια αξιόπιστη γενίκευση από παρατηρήσεις) αρκετές φορές εξετάζεται και αυτός στο πλαίσιο μελέτης της λογικής. Είναι σημαντικό να γίνει διάκριση μεταξύ της παραγωγικής εγκυρότητας και της επαγωγικής εγκυρότητας (που ονομάζεται «Cogency», στα ελληνικά «Πειστικότητα»). Η εξαγωγή ενός συμπεράσματος είναι παραγωγικά έγκυρη, αν δεν υπάρχει πιθανή κατάσταση στην οποία, όλες οι προτάσεις να είναι αληθινές, αλλά το συμπέρασμα ψευδές. Ένα επαγωγικό επιχείρημα μπορεί να μην είναι ούτε έγκυρο ούτε άκυρο, καθώς οι προτάσεις μπορεί να καθορίζουν συγκεκριμένες πιθανότητες εγκυρότητάς του. Η παραγωγική εγκυρότητα διέπεται από αυστηρούς κανόνες στη λειτουργία των συστημάτων της τυπικής λογικής της, σε σχέση με στις έννοιες της σημασιολογίας. Από την άλλη πλευρά ωστόσο, η επαγωγική εγκυρότητα ζητά από εμάς να καθορίσουμε ένα σύνολο αξιόπιστων γενικεύσεων, προκειμένου να λειτουργήσει το σύστημα λογικής που αναλύουμε.

§1.3 Ιδιότητες Συστημάτων Λογικής

Μεταξύ των σημαντικών ιδιοτήτων που τα συστήματα λογικής μπορούν να έχουν είναι:

- Η **Συνέπεια «Consistency»**, που σημαίνει ότι κανένα θεώρημα του συστήματος δεν πρέπει να έρχεται σε αντίθεση με ένα άλλο.

- Η **Εγκυρότητα «Validy»**, που σημαίνει ότι οι αποδεικτικοί κανόνες του συστήματος δεν θα επιτρέψουν ποτέ ένα ψευδό συμπέρασμα να εξαχθεί από αληθινές προκείμενες.

- Η **Πληρότητα «Completeness»** (του συστήματος λογικής), που σημαίνει ότι αν ένας τύπος είναι ορθός, τότε θα δύναται να αποδειχθεί (αν είναι αλήθεια, αποτελεί θεώρημα του εν λόγω συστήματος).

- Η **Αξιοπιστία «Soundness»**, ο όρος αυτός έχει πολλές διαφορετικές έννοιες, γεγονός το οποίο δημιουργεί κάποια σύγχυση στη βιβλιογραφία. Πιο συχνά, η αξιοπιστία αναφέρεται σε λογικά συστήματα, γεγονός που σημαίνει ότι αν ένας τύπος δύναται να αποδειχθεί σε ένα σύστημα λογικής, τότε είναι αληθής στο αντίστοιχο μοντέλο/ή δομή (αν Α είναι ένα θεώρημα, τότε το Α είναι αλήθεια). Αυτό στην ουσία αποτελεί το αντίστροφο της πληρότητας. Μια διαφορετική χρήση της έννοιας της αξιοπιστίας αναφέρεται στα επιχειρήματα, όπου σημαίνει ότι οι προκείμενες προτάσεις ενός έγκυρου επιχειρήματος είναι αληθείς και στον πραγματικό κόσμο.

Κάποια λογικά συστήματα δεν πληρούν και τις τέσσερις αυτές ιδιότητες. Για παράδειγμα, τα θεωρήματα μη πληρότητας Kurt Gödel, αποδεικνύουν ότι αρκετά σύνθετα τυπικά συστήματα της αριθμητικής δεν μπορούν να είναι και συνεπή και πλήρη.

§1.4 Προκείμενες

Μια προκείμενη είναι μια δήλωση, την οποία ένα επιχείρημα θα προβάλλει, για να επάγει ή να αιτιολογήσει το συμπέρασμα του. Συνεπώς: η προκείμενη είναι η παραδοχή πως κάτι είναι αλήθεια. Στη λογική, ένα επιχείρημα απαιτεί μια σειρά από τουλάχιστον δύο δηλωτικές φράσεις (ή "προτάσεις") γνωστές και ως προκείμενες μαζί με μια άλλη δηλωτική φράση (ή "πρόταση") γνωστή ως συμπέρασμα. Η εν λόγω δομή, των δύο προκείμενων και

του ενός συμπεράσματος, αποτελεί τη βασική επιχειρηματολογική δομή. Τα πιο πολύπλοκα επιχειρήματα δύνανται να χρησιμοποιήσουν μια σειρά από κανόνες, προκειμένου να συνδέσουν διάφορες προκείμενες σε ένα συμπέρασμα, ή να αντλήσουν ορισμένα συμπεράσματα από τις αρχικές προκείμενες που θα λειτουργήσουν εν συνεχεία ως προκείμενες για περεταίρω συμπεράσματα.

§1.5 Αξιώματα

Τα αξιώματα συνιστούν στην ουσία προκείμενες. Ένα αξίωμα, είναι μια παραδοχή ή αλλιώς μία αφετηρία ενός συλλογισμού. Όπως γίνεται αντιληπτό, ένα αξίωμα είναι μια υπόθεση τόσο εμφανής, ώστε να γίνει αποδεκτή ως αληθινή άνευ αμφισβήτησης. Όπως χρησιμοποιείται στη σύγχρονη λογική, ένα αξίωμα είναι επίσης μια παραδοχή ή μία αφετηρία για ένα συλλογισμό. Τα αξιώματα καθορίζουν και οριοθετούν το πλαίσιο της ανάλυσης. Η σχετική ισχύς ενός αξιώματος θεωρείται δεδομένη σε ένα συγκεκριμένο πλαίσιο ανάλυσης, και χρησιμεύει ως σημείο εκκίνησης για την εξαγωγή συμπερασμάτων και άλλων σχετικών αληθειών. Δεν υπάρχει ξεκάθαρη τοποθέτηση, όσον αφορά στην αιτιολόγηση της απόλυτης ισχύος των αξιωμάτων που χρησιμοποιούνται, στο πλαίσιο των σύγχρονων μαθηματικών, καθώς κάτι τέτοιο θα ήταν μια επουσιώδης και δίχως νόημα διαδικασία.

Στα μαθηματικά, ένα σύστημα αξιωμάτων, που χρησιμοποιείται για τη θεμελίωση ενός κλάδου, πρέπει να πληροί τρεις συνθήκες, τα αξιώματα του συστήματος θα πρέπει: 1) να στερούνται αντιφάσεων (είναι αδύνατον, με τη χρήση αξιωμάτων, να αποδειχθούν δύο προτάσεις που έρχονται σε αντίθεση μεταξύ τους), 2) να πληρούν το σύστημα (δηλαδή, κάθε θεώρημα του κλάδου, που θέλουμε να θεμελιώσουμε αξιωματικά, να μπορεί να προκύψει ως συνέπεια των αξιωμάτων του συστήματος, και 3) πρέπει να είναι ανεξάρτητα μεταξύ τους (δηλαδή, κανένα από αυτά δεν πρέπει να προκύπτει συνέπεια των υπολοίπων και συνεπώς να αποδεικνύεται ως θεώρημα).

§1.6 Υπολογιστική Λογική

Η λογική εισήλθε στην καρδιά της επιστήμης των υπολογιστών, καθώς αναδείχθηκε ως τομέας της. Η εργασία του Alan Touring στο «*Entscheidungsproblem*» (Πρόβλημα Απόφασης) πάνω στα θεωρήματα μη πληρότητας, και η ιδέα των υπολογιστών γενικής χρήσης που προήλθε από την τελευταία, ήταν μέγιστης σημασίας για τους σχεδιαστές του μηχανισμού των ηλεκτρονικών υπολογιστών τη δεκαετία του 1940. Στις δεκαετίες του 1950 και του 1960, οι ερευνητές πρόβλεψαν, πως όταν η ανθρώπινη γνώση θα μπορούσε να εκφράζεται χρησιμοποιώντας τη λογική με μαθηματική σημειογραφία, θα ήταν δυνατόν να δημιουργήσουμε ένα μηχάνημα το οποίο θα συλλογίζεται, ή αλλιώς να δημιουργήσουμε **Τεχνητή Νοημοσύνη «Artificial Intelligence»**. Αποδείχθηκε αργότερα, ότι είναι πιο δύσκολο από ό, τι αναμενόταν, λόγω της πολυπλοκότητας του ανθρώπινου τρόπου σκέψης. Στον λογικό προγραμματισμό, ένα πρόγραμμα αποτελείται από ένα σύνολο αξιωμάτων και κανόνων. Τα συστήματα του λογικού προγραμματισμού, υπολογίζουν τις συνέπειες των αξιωμάτων και των κανόνων, για να απαντήσουν σε ερωτήματα.

Στις μέρες μας, η λογική εφαρμόζεται εκτενώς στους τομείς της τεχνητής νοημοσύνης και της επιστήμης των υπολογιστών, καθότι οι τομείς αυτοί αποτελούν πλούσια πηγή προβλημάτων για την τυπική και την άτυπη λογική. Η θεωρία της επιχειρηματολογίας αποτελεί ένα καλό παράδειγμα, του πώς η λογική εφαρμόζεται στον τομέας της τεχνητής νοημοσύνης.

§1.7 Προτασιακή Λογική

Ένας προτασιακός λογισμός ή λογική είναι ένα τυπικό σύστημα, στο οποίο οι τύποι που αναπαριστούν προτάσεις, δύνανται να σχηματίζονται συνδυάζοντας ατομικές προτάσεις χρησιμοποιώντας λογικούς συνδέσμους, και στον οποίο ένα σύστημα τυπικών κανόνων, δίνει

τη δυνατότητα σε ορισμένους τύπους να θεμελιώνονται ως «θεωρήματα». Η Προτασιακή λογική είναι μία υποκατηγορία της τυπικής λογικής. Προτασιακή είναι η μορφή της λογικής που θα χρησιμοποιήσουμε, ως πλαίσιο ανάλυσης της επαγωγικής αλγοριθμικής γνώσης.

§1.8 Τεχνητή Νοημοσύνη και Αλγοριθμική Γνώση

Ο όρος Τεχνητή Νοημοσύνη (Artificial Intelligence) αναφέρεται στον κλάδο της πληροφορικής, ο οποίος ασχολείται με τη σχεδίαση και την υλοποίηση υπολογιστικών συστημάτων, που μιμούνται στοιχεία της ανθρώπινης συμπεριφοράς, τα οποία δηλαδή ενέχουν έστω και στοιχειώδη δείγματα ευφυΐας: μάθηση, εξαγωγή συμπερασμάτων, κατανόηση από συμφραζόμενα, επίλυση προβλημάτων, προσαρμοστικότητα κλπ. Ο John McCarthy όρισε τον τομέα αυτόν ως «επιστήμη και μεθοδολογία της δημιουργίας νοημόνων μηχανών».

Η τεχνητή νοημοσύνη αποτελεί σημείο τομής μεταξύ πολλών επιστημών όπως της πληροφορικής, της φιλοσοφίας, της γλωσσολογίας, της νευρολογίας, της ψυχολογίας και της επιστήμης μηχανικών, με στόχο τη δημιουργία ευφυούς συμπεριφοράς, με στοιχεία μάθησης συλλογιστικής και προσαρμογής στο περιβάλλον, ενώ κατά το πλείστον εφαρμόζεται σε μηχανές ή υπολογιστές ειδικής κατασκευής. Επιμερίζεται στη συμβολική τεχνητή νοημοσύνη, η οποία προσπαθεί να εξομοιώσει την ανθρώπινη νοημοσύνη αλγοριθμικά χρησιμοποιώντας σύμβολα και λογικούς κανόνες υψηλού επιπέδου, και στην υποσυμβολική τεχνητή νοημοσύνη, η οποία επιχειρεί να αναπαράγει την ανθρώπινη ευφυΐα χρησιμοποιώντας βασικά αριθμητικά μοντέλα που συνθέτουν επαγωγικά νοήμονες συμπεριφορές μέσω της διαδοχικής αυτό-οργάνωσης απλούστερων δομικών συστατικών («συμπεριφορική τεχνητή νοημοσύνη»). Επιπρόσθετα προσομοιώνουν ορισμένες βιολογικές διαδικασίες όπως η εξέλιξη των ειδών και η λειτουργία του εγκεφάλου («υπολογιστική νοημοσύνη»), ή συντελούν στην εφαρμογή στατιστικών μεθοδολογιών σε προβλήματα τεχνητής νοημοσύνης.

Η διάκριση σε συμβολικές και υποσυμβολικές προσεγγίσεις αφορά στον χαρακτήρα των χρησιμοποιούμενων εργαλείων, ενώ είναι συχνή η σύζευξη πολλαπλών προσεγγίσεων (διαφορετικών συμβολικών, υποσυμβολικών, ή ακόμα συμβολικών και υποσυμβολικών μεθόδων) στην προσπάθεια αντιμετώπισης ενός προβλήματος. Με βάση τον επιθυμητό επιστημονικό στόχο η τεχνητή νοημοσύνη κατηγοριοποιείται και σε άλλου είδους ευρείς τομείς, όπως η επίλυση προβλημάτων, η ανακάλυψη γνώσης, η μηχανική μάθηση, τα συστήματα γνώσης κλπ. Η τεχνητή νοημοσύνη αποτελεί ένα από τα πλέον «μαθηματικοποιημένα» και ταχέως εξελισσόμενα πεδία της πληροφορικής. Η μηχανική μάθηση είναι ο τομέας της τεχνητής νοημοσύνης, στο πλαίσιο της οποίας, ξεδιπλώνεται η παρούσα εργασία.

§1.9 Μηχανική Μάθηση

Η Μηχανική μάθηση αποτελεί υποπεδίο της επιστήμης των υπολογιστών, η οποία αναπτύχθηκε από τη μελέτη της αναγνώρισης προτύπων και της υπολογιστικής θεωρίας μάθησης στην σύγχρονη τεχνητή νοημοσύνη. Το 1959, ο Arthur Samuel ορίζει τη μηχανική μάθηση ως το "Πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί". Η μηχανική μάθηση διερευνά τη μελέτη και την κατασκευή αλγορίθμων, οι οποίοι δύνανται να μαθαίνουν από τα δεδομένα και να κάνουν προβλέψεις σε σχέση με αυτά. Τέτοιοι αλγόριθμοι λειτουργούν δημιουργώντας μοντέλα προερχόμενα από πειραματικά δεδομένα, με σκοπό να πραγματοποιήσουν προβλέψεις βασισμένες στα δεδομένα ή να εξάγουν αποφάσεις, οι οποίες εκφράζονται ως το αποτέλεσμα.

Η μηχανική μάθηση είναι στενά συνδεδεμένη και πολλές φορές συγχέεται με την υπολογιστική στατιστική, ένας κλάδος, που και αυτός επικεντρώνεται στην πρόβλεψη με τη βοήθεια της χρήσης των υπολογιστών. **Έχει ισχυρούς δεσμούς με τη μαθηματική βελτιστοποίηση**, η οποία την τροφοδοτεί με μεθόδους, τη θεωρία και ποικίλους τομείς

εφαρμογής. Η μηχανική μάθηση εφαρμόζεται σε μια πλειάδα από υπολογιστικές εργασίες, όπου τόσο ο σχεδιασμός όσο και ο σαφής προγραμματισμός αλγορίθμων δεν είναι εφικτός.

Η μηχανική μάθηση πολλές φορές συγχέεται επίσης με την εξόρυξη δεδομένων, (Data Mining), όπου η τελευταία επικεντρώνεται περισσότερο στην εξερευνητική ανάλυση των δεδομένων, γνωστή και ως "**μη επιτηρούμενη μάθηση**". Στο πεδίο της ανάλυσης δεδομένων, η μηχανική μάθηση είναι μια μέθοδος που χρησιμοποιείται για την επινόηση σύνθετων μοντέλων και αλγορίθμων που οδηγούν στην πρόβλεψη. Τα αναλυτικά μοντέλα επιτρέπουν στους ερευνητές, τους μηχανικούς, τους αναλυτές και τους επιστήμονες δεδομένων, να παράγουν αξιόπιστες αποφάσεις και αποτελέσματα και να αναδείξουν αλληλοσυσχετίσεις, μέσω της μάθησης από ιστορικές σχέσεις και από τάσεις στα δεδομένα. Στο πλαίσιο της μηχανικής μάθησης η επαγωγική αλγοριθμική γνώση, είναι στην ουσία μία μέθοδος δημιουργίας νέας γνώσης από παλιά. Δημιουργείται δηλαδή μία μέθοδος συλλογισμού του υπολογιστή προς εξαγωγή συμπερασμάτων και διατύπωση προβλέψεων από δεδομένα.

ΚΕΦΑΛΑΙΟ 2^{Ο}

§2. Επαγωγική Αλγοριθμική Γνώση

§2.1 Το Θεωρητικό Πλαίσιο της Επαγωγικής Αλγοριθμικής Γνώσης

Είναι γνωστό, ότι το κανονικό μοντέλο της γνώσης, το οποίο βασίζεται στη θεωρία των πιθανών κόσμων, υπόκειται στο πρόβλημα της λογικής παντογνωσίας, κατά το οποίο **οι**

πράκτορες γνωρίζουν όλες τις λογικές επιπτώσεις της γνώσης που κατέχουν (**στο εξής πράκτορα θα αποκαλούμε** οποιονδήποτε μηχανισμό μηχανικής μάθησης – τεχνητής νοημοσύνη, είναι σε θέση να διαχειρίζεται δεδομένα και να εξάγει συμπεράσματα - αποτελέσματα από αυτά) **[Halpern, Moses, and Vardi 1995, Chapter 9] [1]**. Εν ολίγοις, η ιδέα των πιθανών κόσμων χρησιμοποιείται, προκειμένου να εκφράσουμε ισχυρισμούς μέσα από μία βοηθητική λογική. Θεωρούμε δηλαδή, ότι ο πραγματικός κόσμος είναι αυτός στον οποίο ζούμε και είναι ένας από τους πολλούς πιθανούς που υπάρχουν. Η βοηθητική λογική που χρησιμοποιούμε διέπει και ορίζει στην ουσία τον πραγματικό αυτό κόσμο. Η βοηθητική λογική είναι ένα είδος τυπικής λογικής, η οποία μας επιτρέπει να αποδίδουμε προτασιακά ισχυρισμούς με διάφορους τελεστές. Υπάρχει στενή σχέση μεταξύ των προτάσεων και των πιθανών κόσμων. Επισημαίνουμε ότι κάθε πρόταση είναι είτε αληθής είτε εσφαλμένη σε κάθε πιθανό κόσμο. Η βοηθητική λογική ορίζει έπειτα, σε ποιους κόσμους μία πρόταση είναι αληθής και σε ποιους εσφαλμένη. Έτσι ο ορισμός της γνώσης σε σχέση με τη θεωρία των πιθανών κόσμων, καθιστά δύσκολη την διατύπωση της γνώσης, την οποία οι πράκτορες χρειάζεται να υπολογίζουν με βεβαιότητα, προκειμένου να πάρουν αποφάσεις και να δράσουν ή να αντιληφθούν καταστάσεις, τις οποίες χρειάζεται με τη σειρά τους να επεξεργασθούν άλλοι πράκτορες, προκειμένου και αυτοί να δράσουν ανάλογα.

Αυτή η παρατήρηση οδηγεί σε μία διαφοροποίηση μεταξύ δύο ειδών γνώσης, την παραγωγική γνώση και την επαγωγική γνώση (ή αλλιώς γνώση βασισμένη σε πηγές), μία διαφοροποίηση καιρό τώρα διατυπωμένη από τον **[Rosenschein 1985] [2]**. Η κλασσική μέθοδος τεχνητής νοημοσύνης (Artificial Intelligence "AI"), γνωστή και ως μέθοδος δομών μεταφρασμένων συμβόλων, στην οποία η γνώση βασίζεται σε πληροφορίες αποθηκευμένες σε βάσεις δεδομένων, μπορεί να θεωρηθεί μία περίπτωση παραγωγικής γνώσης. Αντίθετα η μέθοδος των προκαθορισμένων αυτοματισμών (Situated Automata Approach), η οποία μεταφράζει τη γνώση, με βάση τις πληροφορίες που αντλούνται από την κατάσταση στην οποία βρίσκεται ένας μηχανισμός, μπορεί να θεωρηθεί μία περίπτωση επαγωγικής γνώσης. Ο **[Levesque 1984] [3]** ορίζει μία παρόμοια διαφοροποίηση μεταξύ των παραγωγικών και των επαγωγικών πεποιθήσεων. Χρησιμοποιώντας τον όρο «**πεποίθηση**» αναγνωρίζουμε το

γεγονός ότι αυτό που γνωρίζει ένας πράκτορας, δεν είναι απαραίτητα αληθές και μπορεί στην πραγματικότητα να μεταβληθεί στο μέλλον.

Ενώ η θεωρία των πιθανών κόσμων θεωρείται το κανονικό μοντέλο εξέτασης της επαγωγικής γνώσης, δεν υπάρχει κανονικό μοντέλο εξέτασης της παραγωγικής γνώσης. Μία γενική προσέγγιση για πολλές περιπτώσεις παραγωγικής γνώσης , είναι η αλγοριθμική γνώση (Algorithmic Knowledge) **[Fagin, Halpern, Moses, and Vardi 1994] [4]**. Στο πλαίσιο της αλγοριθμικής γνώσης, η παραγωγική γνώση ενός πράκτορα δίδεται από έναν αλγόριθμο, τον οποίο ο πράκτορας χρησιμοποιεί, προκειμένου να αξιολογήσει αν γνωρίζει ένα συγκεκριμένο δεδομένο. Η αλγοριθμική γνώση μπορεί βοηθητικά να μελετηθεί, προκειμένου να κατανοήσουμε μία σειρά από μοντέλα απόδοσης συλλογιστικής βασισμένων σε πηγές δεδομένων, τα οποία παρατίθενται στη σχετική βιβλιογραφία **[Levesque 1984, Konolige 1986, Elgot - Drapkin and Perlis 1990][5] *[2] (Παράρτημα «Α»).**

Η γενικότητα της προσέγγισης της αλγοριθμικής γνώσης, την καθιστά ιδανική σαν πλαίσιο μοντελοποίησης. Ένα πρόβλημα της γενικότητας αυτής ωστόσο είναι, ότι δεν υπάρχουν αξιοσημείωτες λογικές ιδιότητες στην υπό εξέταση αλγοριθμική γνώση, εκτός αν εστιάσουμε σε συγκεκριμένες κατηγορίες γνωσιακών αλγορίθμων. Το γεγονός αυτό είναι σημαντικό, όταν το εν λόγω πλαίσιο μοντελοποίησης χρησιμοποιείται ως γλώσσα επικοινωνίας, για τον προσδιορισμό ιδιοτήτων συστημάτων πολλών πρακτόρων, υποκείμενων σε αυτόματες διαδικασίες ταυτοποίησης. Σε μία τέτοια κατάσταση, θα ήταν θεμιτή η χρησιμοποίηση μίας κατηγορίας γνωσιακών αλγορίθμων, οι οποίοι θα εκμαιεύουν ιδιότητες, μέσα από μία διαδικασία ταυτοποίησης, ενώ θα διατηρούν δομή τέτοια, που να δημιουργούν ένα ανιχνεύσιμο και αναλύσιμο πλαίσιο μοντελοποίησης.

Αυτή η δομή τυπικά περιλαμβάνει μία σειρά από ιδιότητες με αντίστοιχους τελεστές αλγοριθμικής γνώσης, οι οποίοι μπορούν να χρησιμοποιηθούν για τη μελέτη ιδιοτήτων συστημάτων πολλών πρακτόρων επαγωγικά. Αυτή η γενική παρατήρηση, μας οδηγεί σε μία διαδικασία μελέτης κατηγοριών γνωσιακών αλγορίθμων με εξαιρετικό ενδιαφέρον. Σε αυτή την εργασία εξετάζουμε μία μορφή αλγοριθμικής γνώσης, την επαγωγική αλγοριθμική

γνώση, κατά την οποία η παραγωγική γνώση πρακτόρων που προέρχεται από μία λογική θεωρία, εκφράζεται με τη βοήθεια ενός επαγωγικού συστήματος δομημένου με δικούς του επαγωγικούς κανόνες, μέσα από το οποίο οι πράκτορες παρουσιάζουν τη συλλογιστική τους, σε σχέση με δεδομένα που γνωρίζουν. Πολλές χρήσιμες μορφές παραγωγικής γνώσης δύνανται να τυποποιηθούν με τη χρήση επαγωγικών συστημάτων. Για παράδειγμα οι θεωρίες των χοανών (Horn Theories) **[Selman and Kautz 1996] [6]**, οι οποίες έχουν χρησιμοποιηθεί για να προσεγγίσουν πιο γενικά γνωσιακά υπόβαθρα, ταιριάζουν περισσότερο στο πλαίσιο που εξετάζουμε. Η παραγωγική γνώση μέσω ενός επαγωγικού συστήματος, μπορεί να θεωρηθεί ως μία μορφή αλγοριθμικής γνώσης, όπου ο γνωσιακός αλγόριθμος που χρησιμοποιείται από έναν πράκτορα, προσπαθεί να συμπεράνει αν ένα δεδομένο δύναται να προκύψει από επαγωγικούς κανόνες, οι οποίοι υπάρχουν στο πλαίσιο ενός επαγωγικού συστήματος. Πλεονέκτημα αποτελεί και το γεγονός, ότι εξετάζοντας την επαγωγική αλγοριθμική γνώση ως μία περίπτωση αλγοριθμικής γνώσης, μπορούμε να μοντελοποιήσουμε την «εφαρμόσιμη» παραγωγική γνώση, μέσα από επαγωγικά συστήματα, των οποίων ο αντίστοιχος γνωσιακός αλγόριθμος είναι αποτελεσματικός (δύναται δηλαδή να περατωθεί σε πολυωνυμικό χρόνο). Η προσέγγιση μοντελοποίησης της παραγωγικής γνώσης μέσα από επαγωγικά συστήματα είναι στη βάση της του ερευνητή [**Konolige 1986] [7**]. Η βασική ιδέα, την οποία θα παραθέσουμε στη δεύτερη παράγραφο ορίζει, ότι ένα επαγωγικό σύστημα είναι ένα σύνολο κανόνων, οι οποίοι περιγράφουν το πώς μπορούμε να συμπεράνουμε καινούρια δεδομένα από παλιά. Θεωρούμε ότι τα αρχικά δεδομένα είναι στην ουσία οι παρατηρήσεις που έχει κάνει ο πράκτορας(ή τις οποίες δύναται να κάνει), απλά εξετάζοντας την τοπική κατάσταση στην οποία βρίσκεται.

Στην τρίτη παράγραφο, περιγράφουμε ένα μοντέλο απόδοσης επαγωγικής γνώσης και επαγωγικής αλγοριθμικής γνώσης, με τον τελεστή $K\varphi$ να εκφράζει την επαγωγική γνώση της φ και τον τελεστή $X\varphi$ να εκφράζει την παραγωγική γνώση της φ. Προς απλούστευση, η λογική που παρουσιάζουμε είναι προτασιακή, αν και δεν θα υπήρχε κάποιου είδους δυσκολία στο να την παρουσιάσουμε ως μαθηματική λογική με αντίστοιχες συναρτήσεις. (Φυσικά, ερωτήματα εγείρονται για τον τρόπο με τον οποίο ποσοτικοποιούνται οι χρησιμοποιούμενοι

βοηθητικοί τελεστές (σε σχέση με τις συναρτήσεις). Γιατί όμως χρειάζεται να ερμηνεύσουμε και τις δύο αυτές μορφές γνώσης; Σε αρκετές περιπτώσεις προκύπτει, ότι χρειάζονται και οι δύο. Κι αυτό γιατί η επαγωγική γνώση είναι χρήσιμη για απόδοση υφιστάμενων δεδομένων και περιγράφει την «κατ' εκδοχή» γνώση, ενώ η επαγωγική αλγοριθμική γνώση είναι χρήσιμη για την εκμαίευση γνώσης, που οι πράκτορες μπορούν πρακτικά να υπολογίσουν και να χρησιμοποιήσουν. Ας πάρουμε ένα παράδειγμα. Έχει αποδειχθεί στο παρελθόν, το πώς το πλαίσιο της αλγοριθμικής γνώσης θα μπορούσε να ερμηνεύσει την επικοινωνία μεταξύ πρακτόρων, μέσα από κρυπτογραφημένα πρωτόκολλα **[Halpern and Pucella 2002] [8]**. Η αλγοριθμική γνώση είναι χρήσιμη για τη μοντελοποίηση ενός αντιπάλου, ο οποίος έχει συγκεκριμένες δυνατότητες προς αποκωδικοποίηση μηνυμάτων τα οποία υποκλέπτει. Υπάρχουν φυσικά περιορισμοί στις δυνατότητες ενός τέτοιου αντιπάλου. Για παράδειγμα, ο αντίπαλος μπορεί να μην γνωρίζει σαφώς, ότι έχει στην κατοχή του ένα μήνυμα, αν το μήνυμα αυτό είναι κρυπτογραφημένο με μία κλείδα που ο αντίπαλος δεν γνωρίζει. Προκειμένου να τυποποιηθούν αυτοί οι περιορισμοί, οι **[Dolev and Yao 1983] [9]**περιέγραψαν κάποιες σταθερές δυνατότητες των αντιπάλων. Γενικά, ένας αντίπαλος Dolev – Yao (όπως ονομάσθηκε) δύναται να αποσυνθέσει τα μηνύματα σε επιμέρους τμήματα ή να τα αποκρυπτογραφήσει, αν γνωρίζει τις κατάλληλες κλείδες, διαφορετικά δεν μπορεί να «σπάσει» την κρυπτογράφηση. Ο αντίπαλος επίσης, μπορεί να κατασκευάσει νέα μηνύματα συγκολλώντας γνωστά ή αποκρυπτογραφώντας τα με μία γνωστή κλείδα αποκρυπτογράφησης. Είναι σαφές λοιπόν ότι πραγματοποιείται η τυποποίηση ενός αντιπάλου «Dolev – Yao», με τη χρήση ενός επαγωγικού συστήματος, το οποίο περιγράφει το περιεχόμενο των μηνυμάτων του, με βάση μηνύματα που αυτός έχει αποκωδικοποιήσει και που μπορεί να κατασκευάσει από τα υπάρχοντα δεδομένα. Αυτό μας δίνει τη δυνατότητα να περιγράψουμε ιδιότητες όπως: «Ο αντίπαλος μπορεί να υπολογίσει (γνωρίζει με σιγουριά δηλαδή), τη διαβίβαση του κωδικοποιημένου μηνύματος, μέσα από ένα πρωτόκολλο επικοινωνίας. Για να αποδώσουμε το γεγονός, ότι κάποιος άλλος πράκτορας γνωρίζει, ότι η διαβίβαση ενός κωδικοποιημένου μηνύματος μέσα από ένα πρωτόκολλο επικοινωνίας

παραμένει κρυφή από τον αντίπαλο, μπορούμε να χρησιμοποιήσουμε την επαγωγική γνώση όπως: «Ο Πράκτορας Α γνωρίζει (επαγωγικά), ότι ο αντίπαλος δεν μπορεί να υπολογίσει (να γνωρίζει δηλαδή σαφώς) το απόρρητο μήνυμα που διαβιβάσθηκε». Μέσα από ένα τέτοιο μοτίβο, τα ζητούμενα μπορούν να περιγράφονται και με τα δύο είδη γνώσης. (Τα παραπάνω ζητούμενα χρειάζονται ορισμένες προσαρμογές, προκειμένου να εφαρμοσθούν σε συστήματα ανάλυσης με πολλούς πράκτορες, ζήτημα το οποίο θα αναλυθεί στην παράγραφο 6).

Ένας βασικός τελεστής στη λογική που αναλύουμε είναι ο τελεστής Ob, ο οποίος συμβολίζει τις παρατηρήσεις που πραγματοποιούνται στην παρούσα κατάσταση. Αυτός ο τελεστής αποτελεί στην ουσία τη σύνδεση μεταξύ επαγωγικής και παραγωγικής γνώσης: η επαγωγική αλγοριθμική γνώση κάνει χρήση των παρατηρήσεων ως αρχικά δεδομένα, από τα οποία δευτερεύοντα δεδομένα προκύπτουν, ενώ η επαγωγική γνώση χρησιμοποιεί τις παρατηρήσεις, προκειμένου να διαφοροποιήσει καταστάσεις. Δύο καταστάσεις δεν γίνεται να διαφοροποιηθούν, όταν ο πράκτορας κάνει τις ίδιες παρατηρήσεις και στις δύο καταστάσεις. Έτσι, μπορούμε να μελετήσουμε την αλληλεπίδραση μεταξύ παραγωγικής και επαγωγικής γνώσης. Με αυτόν τον τρόπο ξεφεύγουμε από προηγούμενα μοντέλα, τα οποία προσπαθούν να μοντελοποιήσουν μόνο την παραγωγική γνώση **[Giunchiglia, Serafini, Giunchiglia, and Frixione 1993] [10]**. Η εργασία αυτή παρουσιάζει έναν τρόπο να συνδυάσουμε την επαγωγική γνώση που υπάρχει σε έναν πιθανό κόσμο, με ένα επαγωγικό σύστημα, το οποίο αναπαριστά την παραγωγική γνώση των πρακτόρων και παράλληλα να προσδιορίσουμε την μεταξύ τους σχέση.

Ένας θεμελιώδης σκοπός της εργασίας αυτής είναι, να μελετήσει τις τεχνικές ιδιότητες της προκύπτουσας λογικής, όπως τα αξιώματα που τη διέπουν, καθώς και την πολυπλοκότητα των προβλημάτων απόφασης που προκύπτουν. Έπειτα να αναδείξει τον τρόπο με τον οποίο τα παραπάνω, συνδέονται με τα επαγωγικά συστήματα. Στην παράγραφο 4, μελετούμε αξιώματα που αποδίδονται σε συγκεκριμένα επαγωγικά συστήματα. Είναι επόμενο, ότι αν δεν κάνουμε καμία υπόθεση στο πλαίσιο των επαγωγικών συστημάτων, υπάρχουν ελάχιστες ιδιότητες που προκύπτουν από σχετικά αξιώματα, το οποίο στην ουσία

αποτελεί μία μικρή επέκταση των γνωστών αξιωμάτων, που αφορούν και στην επαγωγική γνώση. Ωστόσο, αν επικεντρωθούμε σε ένα συγκεκριμένο επαγωγικό σύστημα, οι ιδιότητες ενός τελεστή που μπορεί να χρησιμοποιεί ένας πράκτορας (π.χ. X), εξαρτώνται από το σύστημα αυτό. Συνεπώς, είμαστε σε θέση να εξάγουμε τις ιδιότητες του τελεστή αυτού, μέσα από τους επαγωγικούς κανόνες του μοντέλου που χρησιμοποιούμε. Αποδεικνύουμε λοιπόν, ότι μπορούμε να εξάγουμε πλήρη και έγκυρα αξιώματα για τη λογική που αναπτύσσουμε, σε σχέση με ένα συγκεκριμένο επαγωγικό σύστημα και τα οποία προκύπτουν μηχανικά μέσα από του κανόνες του συστήματος.

Στην πέμπτη παράγραφο, αναλύουμε την πολυπλοκότητα ενός προβλήματος προς απόφαση, σε σχέση με τη χρησιμοποιούμενη λογική, υπό την παρουσία ενός πράκτορα. Αναλύουμε δηλαδή, αν μία συνάρτηση της λογικής αυτής, ανήκει σε κάποια συγκεκριμένη κατηγορία μοντέλων. **Εδώ είναι σημαντικό να υπενθυμίσουμε τις έννοιες της «ικανοποιητικότητας» και της «αποφασιστικότητας» μίας συνάρτησης, που μπορεί να βρεθεί στο πλαίσιο μιας υπό εξέταση λογικής.** Μία συνάρτηση είναι «**ικανοποιητική**», εάν είναι δυνατόν να βρούμε ένα μοντέλο ερμηνείας της, ώστε όλα τα στοιχεία της να είναι αληθή στην χρησιμοποιούμενη λογική. Στη λογική, ένα αληθές/ψευδές πρόβλημα είναι «**αποφασιστικό**», εάν υπάρχει μία αποτελεσματική μέθοδος - αλγόριθμος προς εξαγωγή μίας ορθής απάντησης αυτού σε ορισμένο πολυωνυμικό χρονικό διάστημα. Χωρίς καθόλου υποθέσεις σε σχέση με τα επαγωγικά συστήματα, το να επιβεβαιώσουμε την «ικανοποιητικότητα» μιας συνάρτησης, είναι ένα πλήρες πρόβλημα σε πολυωνυμικό χρόνο (NP - complete). Αυτό είναι επόμενο, αφού η λογική του πλαισίου αυτού ανήκει στην επαγωγική γνώση, η οποία είναι γνωστό ότι είναι «αποφασιστική» σε πολυωνυμικό χρόνο (NP - complete) **[Ladner 1977] [11]**, στην περίπτωση του ενός πράκτορα. Εφόσον καθορίσουμε ένα συγκεκριμένο επαγωγικό σύστημα, η «ικανοποιητικότητα» μιας συνάρτησης που ανήκει σε αυτό, σε σχέση με μία κατηγορία μοντέλου, η οποία υπόκειται στο σύστημα αυτό, είναι ένα πρόβλημα δύσκολα επιλύσιμο σε πολυωνυμικό χρόνο (NP – hard). Επίσης είναι δυνατό να αποδείξουμε την πληρότητα του προβλήματος (NP-

completeness), όταν το επαγωγικό σύστημα από μόνο του είναι «αποφασιστικό» σε πολυωνυμικό χρόνο και τα υπό εξέταση μοντέλα περιέχουν έναν μικρό αριθμό παρατηρήσεων σε κάθε κατάσταση.

Στην έκτη παράγραφο, θεωρούμε μία φυσική επέκταση του πλαισίου μας και αναλύουμε την περίπτωση των πολλών πρακτόρων. Για να γίνει αυτό πρέπει να έχουμε μελετήσει το παράδειγμα Dolev-Yao, το οποίο δίδεται παραπάνω γενικευμένο ή κάθε σχετική με τα συστήματα πολλών πρακτόρων βιβλιογραφία. Η επέκταση αυτή γίνεται αναλογικά και πολλά από τα αποτελέσματα είναι προφανή. Αυτός είναι και ο λόγος που ασχολούμαστε κυρίως με τη περίπτωση του ενός πράκτορα. Για παράδειγμα, η πολυπλοκότητα ενός προβλήματος «αποφασιστικότητας» όταν υπάρχουν πολλοί πράκτορες, είναι πλήρες σε πολυωνυμικό χώρο, αφού οι οικογένειες λογικών της επαγωγικής γνώσης έχουν προβλήματα «αποφασιστικά» σε πολυωνυμικό χώρο (PSPACE-complete) **[Halpern and Moses 1992] [12]**. Εδώ επισημαίνεται ότι ένα πρόβλημα είναι «αποφασιστικό» σε πολυωνυμικό **χώρο**, εάν δύναται να λυθεί με τη χρήση μνήμης, η οποία είναι χωροταξικά τοποθετημένη πολυωνυμικά. Οι αποδείξεις των τεχνικών μας αποτελεσμάτων παρατίθενται στα παραρτήματα στο τέλος.

§2.2 Επαγωγικά Συστήματα

Θα ξεκινήσουμε καθορίζοντας το πλαίσιο μέσα στο οποίο εξετάζονται οι λογικές σκέψεις όλων των πρακτόρων, δηλαδή οι επαγωγικές και συμπερασματικές τους δυνατότητες. Το μοντέλο που παρουσιάζουμε παρακάτω, ενοποιεί στην ουσία τα ξεχωριστά μοντέλα λογικής που μπορεί να χρησιμοποιούν δύο ή περισσότεροι πράκτορες για να επικοινωνούν μεταξύ τους. Με αυτόν τον τρόπο μπορούμε να μοντελοποιήσουμε πράκτορες με διαφορετικές συμπερασματικές ικανότητες, χωρίς να επηρεάζεται η λογική απόδοση των πραγμάτων.

Ακολουθώντας την κοινή τακτική, θεωρούμε ότι τα επαγωγικά συστήματα επιδρούν πάνω σε κάποιους αλγεβρικούς όρους. Συγκεκριμένα, θεωρούμε ένα πεπερασμένων σύνολο συναρτήσεων, αλλιώς και υπογραφή, $\Sigma = (f_1, \ldots\ldots, f_n)$, όπου κάθε f_i είναι το σύμβολο της συνάρτησης με δείκτηi. Στην άλγεβρα μία υπογραφή αποτελεί μία λίστα από συναρτήσεις που χαρακτηρίζουν μία αλγεβρική δομή. Οι συναρτήσεις με βαθμό 0 καλούνται σταθερές. Θεωρούμε επίσης ένα μετρήσιμο σύνολο μεταβλητών $Vars$. Καθορίζουμε το αλγεβρικό σύνολο T_Σ ως το ελάχιστο σύνολο όρων, έτσι ώστε $Vars \subseteq T_\Sigma$, και για όλες τις συναρτήσεις $f \in \Sigma$ με δείκτη n και για όλους τους όρους $t_1, \ldots\ldots, t_n \in T_\Sigma$, τότε $f(t_1, \ldots. t_n) \in T_\Sigma$. Επομένως το T_Σ περιέχει όλα τα αποτελέσματα που μπορούν να προκύψουν από το σύνολο των μεταβλητών, των σταθερών και των συναρτήσεων αυτών στην υπογραφή Σ. Λέμε ότι ένας όρος είναι βασικός, όταν δεν περιέχει μεταβλητές. Έτσι ορίζουμε το T_Σ^g ως το σύνολο των βασικών όρων στο T_Σ. Ένα βασικό υποκατάστατο ρ είναι η προβολή των μεταβλητών που ανήκουν στο $Vars$ στους βασικούς όρους. Η εφαρμογή ενός βασικού υποκατάστατου ρ σε έναν όροt, συμβολίζεται με το $\rho(t)$ και προκύπτει όταν αντικαταστήσουμε κάθε t με τον βασικό όρο που αντιστοιχεί στο εκάστοτεt στο ρ. Ξεκάθαρα λοιπόν, η εφαρμογή ενός βασικού υποκατάστατου σε έναν όρο παράγει έναν βασικό όρο.

Ένα Σ-επαγωγικό σύστημα D είναι ένα υποσύνολο $\delta fin(T_\Sigma) X\, T_\Sigma$. Συμβολίζουμε με $\delta(X)$ το σύνολο των υποσυνόλων του X, και $\delta fin(X)$ το σύνολο των πεπερασμένων υποσυνόλων του X. Συχνά αντιστοιχούμε μία υπογραφή Σ σε αυτό, όταν ορίζεται ξεκάθαρα στο υπό εξέταση πλαίσιο. Ένας επαγωγικός κανόνας $(\{t_1, \ldots. t_n\}, t)$ ενός επαγωγικού συστήματος D, συμβολίζεται με τον τύπο $t_1, \ldots. t_n \rhd t$ και σημαίνει, ότι ο όροςt μπορεί να προκύψει επαγωγικά άμεσα από τους $t_1, \ldots. t_n$. Μία επαγωγή του t από ένα σύνολο όρων Γ, είναι μία αλληλουχία από βασικούς όρους $t_1, \ldots, t_n$ έτσι ώστε να είναι $t_n = t$ και κάθε t_i είναι είτε:

(1) ένας όρος $\rho(t')$, για κάποιο βασικό υποκατάστατο ρ και κάποιον όρο $t' \in \Gamma$,

(2) ένας όρος $\rho(t')$, για κάποιο βασικό υποκατάστατο ρ και κάποιον όρο t' για τον οποίο υπάρχει ένας επαγωγικός κανόνας $t'_{i_1}, \ldots\ldots, t'_{i_n} \rhd t$ στο D τέτοιος ώστε $\rho(t'_{i_j}) = t'_{i_j}$ για όλα τα j και κάθε $i_1, \ldots . i_j < i$.

Συμβολίζουμε με $\Gamma \vdash_{D} t$, αν υπάρχει κάποια επαγωγή από το Γ στο tμέσω επαγωγικών κανόνων στο D. Εξ ορισμού λοιπόν έχουμε $t \vdash_{D} t$ για όλους τους όρους t. ***[3](Παράρτημα «Α»)**

Θα ασχοληθούμε κυρίως με επαγωγικά συστήματα τα οποία είναι «αποφασιστικά», δηλαδή τα συστήματα αυτά στα οποία, το πρόβλημα ύπαρξης μίας επαγωγής από τον t στο Γ είναι «αποφασιστικό» για έναν όρο tκαι για ένα πεπερασμένο σύνολο όρων Γ. Επιπλέον, είναι ξεκάθαρο από τους ορισμούς ότι τα επαγωγικά συστήματα είναι μονοτονικά. Με άλλα λόγια, αν $\Gamma \vdash_{D} t$, τότε $\Gamma' \vdash_{D} t$, όταν $\Gamma \subseteq \Gamma'$. Τέλος παρατηρούμε ότι δεν επιβάλλουμε κανέναν περιορισμό στον σχηματισμό των όρων. Η όλη θεωρία που αναπτύσσουμε σε αυτή την εργασία, θα μπορούσε να λάβει ως σημείο εκκίνησης την θεωρία των αλγεβρικών ταξινομημένων όρων (Sorted Term Algebras) **[Higgins 1963] [13]**, με μικρές διαφοροποιήσεις. Αυτό θα επέτρεπε την επιβολή περιορισμών με έναν φυσιολογικό τρόπο.

Παράδειγμα 2.2.1. Θεωρούμε ένα επαγωγικό σύστημα, το οποίο έχει τις δυνατότητες του αντιπάλου "Dolev – Yao" όπως έχει ήδη περιγραφεί. Καθορίζουμε το ακόλουθο Σ-επαγωγικό σύστημα DY, με την υπογραφή $\Sigma = (recv, has, encr, conc, inv)$, όπου το $recv(m)$ συμβολίζει το γεγονός ότι ο αντίπαλος έχει λάβει τον όρο m, το $has(m)$ ότι ο αντίπαλος κατέχει τον όρο m (ότι δηλαδή είναι ικανός να εξάγει το μήνυμα m μέσω του συνόλου των μηνυμάτων που έχει λάβει), το $encr(m, k)$ συμβολίζει την κρυπτογράφηση του όρου m με την κλείδα k, το $conc(m_1, m_2)$ την συγκόλληση των όρων m_1 και m_2 και το $inv(k)$ το αντίστροφο της κλείδας k(είδος κρυπτογράφησης). Έτσι ισχύουν τα παρακάτω:

$$recv(m) \rhd has(m)$$
$$has(inv(k)), has(encr(m,k)) \rhd has(m)$$
$$has(conc(m_1,m_2)) \rhd has(m_1)$$
$$has(conc(m_1,m_2)) \rhd has(m_2)$$

Υποθέτουμε περαιτέρω ότι η υπογραφή Σ περιέχει τους όρους, k_1 και k_2. Έτσι προκύπτει:

$$recv(encr(m,k_1)), recv(encr((inv(k_1),k_2)), recv(inv(k_2)) \vdash_{DY} has(m).$$

Με άλλα λόγια, είναι δυνατό για έναν αντίπαλο "Dolev – Yao" να εξάγει ένα μήνυμα m, αν έχει λάβει το μήνυμα αυτό με μία κλείδα, k_1, η αντίστροφη της οποίας έχει ληφθεί μαζί με μία κλείδα k_2, της οποία η αντίστροφη έχει ληφθεί. Επισημαίνουμε εδώ ότι ο επιθετικός προσδιορισμός «αντίστροφος» χρησιμοποιείται για να περιγράψει το κρυπτογραφημένο.

Προκειμένου να αιτιολογήσουμε την κατασκευή νέων μηνυμάτων, θεωρούμε μία υπογραφή Σ', η οποία επεκτείνει την Σ με έναν τελεστή κατασκευής $constr$, όπου το $constr(m)$ συμβολίζει το γεγονός ότι ο αντίπαλος δύναται να κατασκευάσει τον όρο m. Αιτιολογούμε λοιπόν αυτόν τον νέο τελεστή, προσθέτοντας τους παρακάτω επαγωγικούς κανόνες στο επαγωγικό σύστημα DY.

$$has(m) \rhd constr(m)$$
$$constr(k), constr(m) \rhd constr(encr(m,k))$$
$$constr(m_1), constr(m_2) \rhd constr(conc(m_1,m_2))$$

Για παράδειγμα έχουμε:

$$recv(encr(m,k_1)), recv(inv(k_1)), recv(k_2) \vdash_{DY} constr(encr(m,k_2)).$$

Με ποια έννοια όμως μπορούμε να χρησιμοποιήσουμε επαγωγικά συστήματα, προκειμένου να μοντελοποιήσουμε παραγωγική γνώση; Τα στοιχεία των αλγεβρικών όρων αναπαριστούν δεδομένα, τα οποία μπορούν ενίοτε να επαχθούν από άλλα δεδομένα και οι επαγωγικοί κανόνες του συστήματος μοντελοποίησης τις προκύπτουσες ικανότητες του πράκτορα. Όλα αυτά εγείρουν ένα άλλο ερώτημα, σε σχέση με το ποια δεδομένα μπορούν να θεωρηθούν βασικά για έναν πράκτορα, χωρίς να προκύψουν από επαγωγή. Ποια είναι αυτά δηλαδή που υπάρχουν από την αρχή, δίχως να προκύψουν από άλλα. Ο **[Konolige 1986] [14]** ονόμασε τα γεγονότα αυτά «βασικές πεποιθήσεις» και υπάρχουν αρκετές προσεγγίσεις που έχουν διατυπωθεί ως προς αυτό. Μία από αυτές είναι απλά να θέσουμε ένα σύνολο βασικών δεδομένων - γεγονότων, τα οποία θα είναι από την αρχή γνωστά στον πράκτορα (διαφορετικά βασικά δεδομένα είναι λογικό να υπάρχουν για διαφορετικές καταστάσεις). Αυτό δεν λύνει ολοκληρωτικά το πρόβλημα, καθώς ακόμη απαιτείται ο καθορισμός των δεδομένων που θα είναι γνωστά σε κάθε μία από τις επιμέρους υπάρχουσες καταστάσεις. Εδώ προτείνεται μία διαφορετική προσέγγιση, συμβατή με πολλές από τις γνωστές περιγραφές στη σχετική βιβλιογραφία. Θεωρούμε ότι οι βασικές πεποιθήσεις του πράκτορα είναι οι παρατηρήσεις που κάνει. Επομένως μία παρατήρηση είναι ένα δεδομένο, το οποίο ο πράκτορας κατά περίπτωση καθορίζει σε σχέση με την υφιστάμενη τοπική κατάσταση. Διαχωρίζουμε τις παρατηρήσεις από άλλα δεδομένα χρησιμοποιώντας τον μοναδιαίο κατασκευαστή ob στην υπογραφή των επαγωγικών συστημάτων που εξετάζουμε, όπως στις επόμενες παραγράφους. Επισημαίνεται στο σημείο αυτό, ότι στους τομείς της μαθηματικής λογικής και της πληροφορικής, ένας κατασκευαστής είναι ένας τελεστής, με τη βοήθεια του οποίου δομούνται νέοι τύποι από παλιούς.

§2.3 Ένα Μοντέλο Επαγωγικής Αλγοριθμικής Γνώσης

Θα εισάγουμε τώρα, μία βοηθητική προτασιακή λογική, για την απόδοση επαγωγικής και παραγωγικής γνώσης από έναν πράκτορα, όπου η παραγωγική γνώση τυποποιείται με τη χρήση ενός επαγωγικού συστήματος. Σε αυτή την παράγραφο, εστιάζουμε στην περίπτωση του ενός πράκτορα. Επεκτείνουμε τη μελέτη μας σε περιπτώσεις πολλών πρακτόρων στην έκτη παράγραφο.

Καθορίζουμε λοιπόν τη λογική $L^{KD}(\Sigma)$, η οποία διέπεται από μία υπογραφή Σ. Λαμβάνουμε έπειτα κάποιες προκείμενες, οι οποίες αποτελούν βασικούς όρους T_{Σ}^{g}, της υπογραφής Σ. Λέμε ότι το p χαρακτηρίζει το σύνολο των όρων T_{Σ}^{g}, προκειμένου να τονίσουμε το γεγονός, ότι είναι προτάσεις (περιγραφική απόδοση όρων) και να τις διαχωρίσουμε από όρους (μαθηματικούς) που υπάρχουν σε άλλες υπογραφές που θα περιγράψουμε παρακάτω. Η γλώσσα που χρησιμοποιούμε για την εν λόγω λογική, δομείται στην ουσία, αρχικά χρησιμοποιώντας τις προκείμενες στο T_{Σ}^{g} και εν συνεχεία με άρνηση και συνδυαστικά τον τελεστή K, τον τελεστή X και τον τελεστή Ob (ο οποίος εφαρμόζεται μόνο στις πρωταρχικές προτάσεις). Επομένως το $K\varphi$ σημαίνει «ο πράκτορας γνωρίζει επαγωγικά το φ», το $X\varphi$ σημαίνει «ο πράκτορας γνωρίζει σαφώς το φ, σύμφωνα με το επαγωγικό σύστημα που χρησιμοποιεί» και το $Ob(p)$ σημαίνει «ο πράκτορας παρατηρεί το p». Καθορίζουμε τις συνήθεις συντομογραφίες, $\varphi \vee \psi$ για το $\neg(\neg\varphi \wedge \neg\psi)$ και $\varphi \Rightarrow \psi$ για το $\neg\varphi \vee \psi$. Καθορίζουμε "$\alpha\lambda\eta\theta\acute{\varepsilon}\varsigma$" την συντομογραφία μιας αυθαίρετης αλλά σταθερής ταυτολογίας και "$\psi\varepsilon\upsilon\delta\acute{\varepsilon}\varsigma$" την συντομογραφία του «$\neg\alpha\lambda\eta\theta\acute{\varepsilon}\varsigma$».

Προκειμένου να μεταφράσουμε την επαγωγική αλγοριθμική γνώση ενός πράκτορα, θα πρέπει να τον τροφοδοτήσουμε με ένα επαγωγικό σύστημα, μέσα στο οποίο, θα δύναται να πραγματοποιήσει τις σχετικές επαγωγές. Όπως έχουμε ήδη αναφέρει, θέλουμε ο πράκτορας να μπορεί να αιτιολογήσει τις παρατηρήσεις που διαπιστώνει στην κατάσταση που βρίσκεται. Επιπλέον επειδή θα πρέπει να μπορεί να ερμηνεύει τύπους στο πλαίσιο της λογικής L^{KD}, ως όρους τους οποίους το επαγωγικό σύστημα δύναται να αποδώσει,

χρησιμοποιούμε επαγωγικά συστήματα με την υπογραφή Σ και με μία σχετική της επέκταση Σ^{KD}, η οποία περιέχει κατασκευαστές αντίστοιχους με τους χρησιμοποιούμενους στη λογική μας τελεστές. Έτσι έχουμε: $\Sigma^{KD} = \{ob, true, false, not, and, know, xknow\}$, όπου οι τελεστές $true$ και $false$ έχουν βαθμό 0 (δεν επεξεργάζονται κανέναν όρο), οι τελεστές $ob, not, know, xknow$ έχουν βαθμό 1 (επεξεργάζονται έναν όρο) και ο τελεστής and έχει βαθμό 2 (επεξεργάζεται δύο όρους).

Η σημασιολογία της λογικής μας, ακολουθεί την θεωρία του απλού μοντέλου των πιθανών κόσμων για βοηθητικές λογικές **[Hintika 1962] [15]**. Μία δομή (ή αλλιώς και αρχιτεκτονική) επαγωγικής αλγοριθμικής γνώσης είναι μία πλειάδα συνόλων $M = (S, \pi, D)$, όπου το S είναι ένα σύνολο από καταστάσεις, το π είναι η ερμηνεία των προκείμενων p σε μία κατάσταση s και το D είναι ένα επαγωγικό σύστημα με υπογραφή το $\Sigma \cup \Sigma^{KD}$. Κάθε κατάσταση sτου συνόλου καταστάσεων S είναι της μορφής (e, O), όπου το eσυμβολίζει τη γενική κατάσταση του επαγωγικού συστήματος και το O είναι ένα σύνολο από πεπερασμένο αριθμού παρατηρήσεων. Κάθε παρατήρηση O είναι στη ουσία μία προκείμενη και αναπαριστά τις παρατηρήσεις που ο πράκτορας έκανε στην παρούσα κατάσταση ***[4](Παράρτημα «Α»).** Οι λεπτομέρειες που χαρακτηρίζουν τις καταστάσεις είναι ουσιαστικά αδιάφορες, καθώς χρησιμοποιούνται μόνο ως τρόπος ερμηνείας της αλήθειας των προκείμενων παρατηρήσεων. Γενικά δεν μοντελοποιούμε το πώς οι πράκτορες κάνουν τις παρατηρήσεις ή τη χρονική συσχέτιση μεταξύ των καταστάσεων. Μία κατάσταση απλά αναπαριστά ένα στιγμιότυπο του υπό εξέταση συστήματος. Η ερμηνεία π σχετίζεται με κάθε κατάσταση του συνόλου των προκείμενων, οι οποίες είναι αληθείς στην κατάσταση αυτή, έτσι ώστε για κάθε προκείμενη $p \in T_{\Sigma}^{g}$, να έχουμε $\pi(s)(p) \in \{true, false\}$.

Διαχωρίζουμε ένα δεδομένο, το οποίο συμβολίζεται μέσω μίας προκείμενης p και μιας παρατήρησης του δεδομένου με τον τύπο $Ob(p)$. Για παράδειγμα, το γεγονός ότι η Αλίκη κρατά ένα μήλο, μπορεί να συμβολισθεί με την προκείμενη $holds(alice, apple)$, το οποίο μπορεί να είναι αληθές ή όχι σε μία κατάσταση. Όμως το γεγονός ότι ο πράκτορας παρατήρησε ότι η Αλίκη κρατά ένα μήλο, συμβολίζεται με τον τύπο

$Ob(holds(alice, apple))$, το οποίο είναι αληθές μόνο και μόνο αν η παρατήρηση έγινε στην υφιστάμενη του πράκτορα κατάσταση. Δεν είναι απαραίτητο, ότι αν το $Ob(p)$ ισχύει στην παρούσα κατάσταση, το ίδιο θα ισχύει και για το p. Για τον λόγο αυτό, θεωρούμε ότι είναι πιθανό ένας πράκτορας να κάνει αναξιόπιστες παρατηρήσεις. Μπορούμε φυσικά να υποθέσουμε ότι οι παρατηρήσεις είναι αξιόπιστες, επιβάλλοντας έναν περιορισμό στην ερμηνεία π, θεωρώντας ότι $\pi(e, O)(p) = true$, όποτε το $p \in O$. Πιο γενικά, μπορούμε να επιβάλλουμε περιορισμούς στα υπό εξέταση μοντέλα, όπως οι παρατηρήσεις να περιορίζονται από ένα συγκεκριμένο υποσύνολο των προκείμενων κ.α. ***[5](Παράρτημα «Α»)**.

Παράδειγμα 2.3.1. Τα μοντέλα είναι αναπαραστάσεις καταστάσεων που θέλουμε να αναλύσουμε, για παράδειγμα, ταυτοποιώντας ότι μία κατάσταση ικανοποιεί μία ιδιότητα. Υποθέτουμε ότι θέλουμε να αποδώσουμε τη γνώση ενός αντιπάλου Dolev – Yao, όπως έχει αναφερθεί στην εισαγωγή, σε ένα πλαίσιο μέσα στο οποίο, υπάρχουν αρχές ανταλλαγής μηνυμάτων σύμφωνα με ένα πρωτόκολλο.

Το επαγωγικό σύστημα που χρησιμοποιείται από έναν αντίπαλο, είναι μία επέκταση του DY επαγωγικού συστήματος του παραδείγματος 2.1, έτσι ώστε να επεξεργάζεται παρατηρήσεις. Στη βιβλιογραφία της ασφάλειας πληροφοριών μία σχέση ενός υποόρου των ανταλλασσόμενων μηνυμάτων στο πλαίσιο του πρωτοκόλλου, μπορεί να αξιολογηθεί και να συσχετισθεί. Ορίζουμε το σύμβολο $\sqsubseteq$ στο T^{g}_{DY}, ως τη μικρότερη δυνατή συσχέτιση που ικανοποιεί τα παρακάτω:

$$t \sqsubseteq t$$

$$\text{αν } t \sqsubseteq t_1 \text{ τότε } t \sqsubseteq conc(t_1, t_2)$$

$$\text{αν } t \sqsubseteq t_2 \text{ τότε } t \sqsubseteq conc(t_1, t_2)$$

$$\text{αν } t \sqsubseteq t_1 \text{ τότε } t \sqsubseteq encr(t_1, t_2)$$

Θεωρούμε λοιπόν μία δομή $M = (S, \pi, DY')$, όπου καταγράφουμε σε κάθε κατάσταση όλα τα μηνύματα, στα οποία παρεμβάλλεται ένας πράκτορας στην υφιστάμενη κάθε φορά κατάσταση. Περιορίζουμε τις παρατηρήσεις σε μία κατάσταση, ώστε να είναι της μορφής $recv(t)$ για τους βασικούς όρους t, στους οποίους το has δεν προκύπτει. Το επαγωγικό σύστημα DY' είναι επέκταση του DY με έναν κανόνα που κάνει τις παρατηρήσεις διαθέσιμες στο επαγωγικό σύστημα όπως παρακάτω:

$$Ob(recv(t)) \rhd recv(t)$$

Η ερμηνεία π καθορίζεται, έτσι ώστε $\pi(e, O)(has(t)) = true$, αν και μόνο αν υπάρχει ένας όρος $t' \in T^g_{DY}$, τέτοιος ώστε $recv(t') \in O$ και $t \subseteq t'$. Με άλλα λόγια, το $has(t)$ ισχύει σε μία κατάσταση, αν ο t είναι υποόρος του μηνύματος στο οποίο παρεμβάλλεται ο αντίπαλος. Για παράδειγμα, σε μία κατάσταση s_1 με παρατηρήσεις ισχύει:

$$\{recv(encr(m, k_1)), recv(encr(inv(k_1), k_2))\}$$

και σε μία κατάσταση s_2 με παρατηρήσεις ισχύει:

$$\{recv(encr(m, k_1)), recv(encr(inv(k_1), k_2)), recv(inv(k_2))\}$$

Αυτές οι καταστάσεις αναπαριστούν καταστάσεις, όπου ο αντίπαλος παρεμβλήθηκε σε συγκεκριμένα μηνύματα. Θα δούμε πώς συγκεκριμενοποιούμε ιδιότητες της δομής M, όπως το γεγονός ότι ο αντίπαλος δεν γνωρίζει σαφώς στην κατάσταση s_1 το μήνυμα m, παρόλο που $\pi(s_1)(has(m)) = true$.

Θεωρούμε $M(\Sigma)$ το σύνολο όλων των δομών επαγωγικής αλγοριθμικής γνώσης που χρησιμοποιούν επαγωγικά συστήματα με υπογραφή $\Sigma \cup \Sigma^{KD}$. Για ένα συγκεκριμένο

επαγωγικό σύστημα D με υπογραφή $\Sigma \cup \Sigma^{KD}$, θεωρούμε $M_D(\Sigma)$το σύνολο όλων των δομών επαγωγικής αλγοριθμικής γνώσης που χρησιμοποιούν το επαγωγικό σύστημα D.

Καθορίζουμε με ποιόν τρόπο ένας τύποςφ είναι αληθής στην κατάσταση s της M, το οποίο γράφεται $(M,s) \models \varphi$ επαγωγικά ως ακολούθως. Για το προτασιακό κομμάτι της λογικής, οι κανόνες εφαρμόζονται άμεσα.

$$(M,s) \models p \text{ αν } \pi(s)(p) = true$$
$$(M,s) \models \neg\varphi \text{ αν } (M,s) \not\models \varphi$$
$$(M,s) \models \varphi \wedge \psi \text{ αν } (M,s) \models \varphi \text{ και } (M,s) \models \psi$$

Προκειμένου να καθορίσουμε τη σημασιολογία της γνώσης, ακολουθούμε την κανονική προσέγγιση που εισήγαγε ο **[Hintikka 1962] [16]**. Καθορίζουμε μία σχέση μεταξύ των καταστάσεων, την οποία ο πράκτορας δεν μπορεί να διαχωρίσει με βάση τις παρατηρήσεις που κάνει. Θεωρούμε ότι $s \sim s'$ αν και μόνο αν $s = (e,O)$ και $s' = (e',O)$ για κάποια e, e' και ένα σύνολο παρατηρήσεων O. Ξεκάθαρα το σύμβολο $\sim$ δηλώνει ισοδυναμία μεταξύ των δύο καταστάσεων.

$$(M,s) \models K\varphi \text{ αν } (M,s') \models \varphi \text{ για όλα τα } s' \sim s$$

Προκειμένου να καθορίσουμε τη σημασιολογία του τελεστή X, θα χρειαστεί να εμπλέξουμε το επαγωγικό σύστημα. Για να το κάνουμε αυτό, πρέπει πρώτα να καθορίσουμε την ερμηνεία του τύπουφ της λογικής $L^{KD}(\Sigma)$ σε έναν αλγεβρικό όρο του τύπου φ^T, με ένα τελείως προφανή τρόπο: το p^T είναι p για κάθε προκείμενη p (οι προκείμενες είναι απλά όροι στο T_Σ^g), το $(\neg\varphi^T)$ ισούται με το not(φ^T), $(\varphi \wedge \psi)^T$ ισούται με το and(φ^T, ψ^T), το $(K\varphi)^T$ με το *know*(φ^T), το $(X\varphi)^T$ με το *xknow*(φ^T) και το $(Ob(p))^T$με το *ob(p)*.

$$(M,s) \models X\varphi \text{ αν } s = (e,O) \text{ και } \{ob(p) \mid p \in O\} \vdash_D \varphi^T$$

Η μονοτονία των επαγωγικών συστημάτων σημαίνει ότι για μία δομή Mσε καταστάσεις $s\ s\ =\ (e,O),s'\ =(\ e',O')$ και με$O\ O\ \subseteq O'$, θα ισχύει ότι το $(M,s)\mid=\ X\varphi$ συνεπάγεται το $(M,s')\mid=\ X\varphi$. Έτσι η παραγωγική γνώση δεν χάνεται ποτέ, όταν νέες παρατηρήσεις γίνονται. Τέλος, ερμηνεύουμε το $ob(p)$ ελέγχοντας αν το p είναι μία από τις παρατηρήσεις που έγιναν από τον πράκτορα:

$$(M,s)\mid=\ Ob(p)\ \textit{αν}\ s\ =\ (e,O)\ \text{και}\ \mid p\ \in\ O$$

Ως συνήθως , λέμε ότι ο τύπος φ είναι έγκυρος στην M αν το $(M,s)\mid=\ \varphi$ για κάθε $s\ \in S$ και «ικανοποιητικός» στην M αν $(M,s)\mid=\ \varphi$ για κάποιος $\in S$. Αν το $\mathcal{M}$ είναι ένα σύνολο από δομές επαγωγικής αλγοριθμικής γνώσης, λέμε ότι ο τύπος φ είναι έγκυρος στην M αν ο τύπος φ είναι έγκυρος για κάθε $M\in\mathcal{M}$ και «ικανοποιητικός» στο $\mathcal{M}$ αν οφ είναι «ικανοποιητικός» σε μερικά $M\in\mathcal{M}$. Ένας τύποςφ στη λογική $L^{KD}(\Sigma)$ είναι «έγκυρος», αν είναι «έγκυρος» σε όλο το σύνολο δομών $\mathcal{M}(\Sigma)$ και «ικανοποιητικός», αν είναι «ικανοποιητικός» σε όλο το σύνολο δομών $\mathcal{M}(\Sigma)$.

Παράδειγμα 2.3.2. Θεωρούμε δεδομένο το Παράδειγμα 3.1. Από τον ορισμό του π, $(M,s_1)\mid=K(has(m))$ και $(M,s_2)\mid=K(has(m))$, έτσι ώστε και στις δύο καταστάσεις, ο αντίπαλος επαγωγικά γνωρίζει ότι κατέχει το μήνυμα m. Ωστόσο από τα αποτελέσματα του Παραδείγματος 2.1, μπορούμε να καταλάβουμε ότι στο $(M,s_2)\mid=X(has(m))$, ενώ στο $(M,s_1)\mid=\neg X(has(m))$. Με άλλα λόγια ο αντίπαλος γνωρίζει σαφώς, ότι κατέχει το μήνυμα m στην δεύτερη κατάσταση (όπου έχει παρεμβληθεί με τους κατάλληλους όρους), αλλά όχι στην πρώτη κατάσταση.

Παράδειγμα 2.3.3. Οι ακόλουθοι επαγωγικοί κανόνες μπορούν να προστεθούν σε κάθε επαγωγικό σύστημα, έτσι ώστε να δημιουργηθεί ένα νέο, το οποίο θα επάγει ένα υποσύνολο συμπερασμάτων, τα οποία δύνανται να αποδοθούν με προτασιακή λογική όπως

παρακάτω:

$$t \rhd not(not(t))$$
$$not(and(t, not(t'))), t \rhd t'$$
$$and(t, t') \rhd t'$$
$$not(not(t)) \rhd t$$
$$not(and(t, not(t'))), not(t') \rhd not(t)$$
$$t, not(t) \rhd false$$
$$t \rhd not(and(not(t), not(t')))$$
$$t, t' \rhd and(t, t')$$
$$t' \rhd not(and(not(t), not(t')))$$
$$false \rhd t$$
$$and(t, t') \rhd t.$$

Ένα πλεονέκτημα των κανόνων αυτών, παρόλο που είναι ημιτελείς, είναι ότι μπορούν να χρησιμοποιηθούν για να παρουσιάσουν πολύ αποτελεσματικά (χρονικά και γραμμικά) προτασιακά συμπεράσματα **[McAllester 1993] [17]**.

Παράδειγμα 2.3.4. Μπορούμε να επιτρέψουμε εύκολα στον πράκτορα, να αιτιολογήσει σαφώς την επαγωγική αλγοριθμική του γνώση, προσθέτοντας έναν κανόνα.

$$t \rhd xknow(t) \qquad (1)$$

στο επαγωγικό του σύστημα D. Έτσι, αν η M είναι μία δομή επαγωγικής αλγοριθμικής γνώσης, η οποία πλαισιώνει το D και ισχύει $(M, s) \models X\varphi$, τότε έχουμε $s = (e, O)$, με $O \vdash_D \varphi^T$ και με τον παραπάνω κανόνα, μέσω του επαγωγικού συστήματος D εξάγουμε, ότι $O \vdash_D xknow(\varphi^T)$. Έτσι $(M, s) \models X(X\varphi)$, όπως απαιτείται. Είναι δυνατό να περιορίσουμε την επαγωγική αλγοριθμική γνώση που κατέχει ένας πράκτορας, τροποποιώντας κατάλληλα τον κανόνα (1), περιορίζοντας δηλαδή την εφαρμογή του σε ένα μόνο υποσύνολο των διαθέσιμων όρων.

Κάθε λογική με βοηθητικούς τελεστές, πρέπει να την χειριζόμαστε με λεπτομέρεια, έτσι ώστε να μην εμπίπτει στο πρόβλημα της λογικής παντογνωσίας. Επομένως μια τέτοια λογική εξαναγκάζει κάποιον, να είναι προσεκτικός σε σχέση με το ποια σύμβολα καθορίζονται με συντομογραφίες, και ποια όχι. Προηγουμένως σε αυτή την παράγραφο καθορίσαμε τις συντομογραφίες $true, false,$ ∨ και ⇒, πράγμα το οποίο σημαίνει ότι κάθε τύπος που περιέχει ∨ ή ⇒, είναι στην πραγματικότητα μία συνάρτηση που επίσης περιέχει τα ∧ και ¬. Έτσι ο πράκτορας δεν μπορεί να διαχωρίσει σαφώς μεταξύ τους για παράδειγμα τα φ ∨ ψ και $\neg(\neg\varphi \wedge \neg\psi)$, καθότι στην παρούσα λογική έχουν ακριβώς την ίδια έννοια και η «εγκυρότητα» της σχέσης $\models X(\varphi \vee \psi) \Leftrightarrow X\big(\neg(\neg\varphi \wedge \neg\psi)\big)$ αντανακλά την εν λόγω ταυτότητα. Ένα τέτοιο αποτέλεσμα φαίνεται να αντιβαίνει στις αρχές της παραγωγικής γνώσης. Αυτό όμως γίνεται, προκειμένου να εξασφαλισθεί ότι η γνώση δεν εγκλωβίζεται κάτω από συγκεκριμένες μονοσήμαντες ταυτολογίες. Μέρος του προβλήματος αυτού, είναι ότι καθορίζοντας τελεστές με συντομογραφίες, αναπόφευκτα προκύπτουν ισότητες. Ένας εύκολος τρόπος να παρακάμψουμε το πρόβλημα αυτό, είναι να χρησιμοποιήσουμε μία σύνταξη, η οποία χρησιμοποιείται με τα ∨ και ⇒ και ίσως άλλους συνδέσμους, παρά να τα εισάγουμε στη λογική μας μέσω συντομογραφιών. Θα προσθέταμε σε αυτή την περίπτωση παρόμοιους κατασκευαστές στην υπογραφή Σ^{KD} και θα επεκτείναμε την ερμηνεία του τύπου φ^T κατάλληλα. Αυτό θα μας έδινε τον απόλυτο έλεγχο ως προς το ποιες ταυτολογίες είναι έγκυρες μέσω της παραγωγικής γνώσης και ποιες όχι.

Τέλος είναι σημαντικό να προσδιορίσουμε, τη σχέση του δικού μας πλαισίου ανάλυσης, με αυτό του **[Konolige 1986][18]**. Το πλαίσιο ανάλυσης του Konolige, αντιστοιχεί σε επαγωγικά συστήματα που περιγράψαμε στην προηγούμενη παράγραφο. Παρέχει δηλαδή μία λογική θεωρία απόδοσης δεδομένων με μαθηματική λογική σε έναν πράκτορα (λογική θεωρία πρώτου βαθμού). Η λογική θεωρία που αναλύσαμε στην παρούσα παράγραφο, μπορεί να χρησιμοποιηθεί για να αποδώσει την συλλογιστική, σε σχέση με τη γνώση πρακτόρων, οι οποίοι χρησιμοποιούν τις λογικές θεωρίες του Konolige. Υπό αυτή την έννοια, η λογική μας είναι συμβατή με το πλαίσιο ανάλυσης της λογικής του Konolige.

Επιπλέον όμως, το δικό μας πλαίσιο θεσπίζει σημασιολογικά τις βασικές πεποιθήσεις, των οποίων την ύπαρξη ο Konolige απλά υπέθεσε, λέγοντας ότι: οι βασικές πεποιθήσεις αντιστοιχούν σε παρατηρήσεις, οι οποίες μπορούν να στοιχειοθετηθούν, ανεξάρτητα από τις πεποιθήσεις του πράκτορα.

§2.4 Συστήματα Αξιωμάτων

Σε αυτή την παράγραφο, παρουσιάζουμε μία ορθή και πλήρη δημιουργία αξιωμάτων, προκειμένου να αιτιολογήσουμε την άμεση γνώση, η οποία προκύπτει μέσω ενός επαγωγικού συστήματος. Εδώ να σημειώσουμε, ότι ένας τύπος φ είναι «**αποδείξιμος**» σε ένα σύστημα αξιωμάτων, αν δύναται να αποδειχθεί με τη χρήση αξιωμάτων και συμπερασματικών κανόνων του συστήματος αυτού. Ένα τέτοιο σύστημα αξιωμάτων είναι ορθό σε σχέση με μία κατηγορία δομών– αρχιτεκτονική $\mathcal{M}$, αν κάθε τύπος «αποδείξιμος» στο σύστημα είναι έγκυρος στην $\mathcal{M}$. Ένα σύστημα αξιωμάτων είναι πλήρες σε σχέση με την $\mathcal{M}$, αν κάθε τύπος έγκυροςστην $\mathcal{M}$ είναι «αποδείξιμος» στο σύστημα αυτό.

Ξεκάθαρα για ένα συγκεκριμένο επαγωγικό σύστημα, οι ιδιότητες του τελεστή X, εξαρτώνται από το επαγωγικό σύστημα αυτό. Επομένως θα μπορούσαμε να εξάγουμε τις ιδιότητες του X μέσω των ίδιων των επαγωγικών κανόνων. Οι ιδιότητες των γνωσιακών αλγορίθμων στο πλαίσιο της αλγοριθμικής γνώσης, μεταφράζονται άμεσα σε ιδιότητες του τελεστή X. Προκειμένου να προσαρμόσουμε ένα παράδειγμα των **[Halpern, Moses, and Vardi 1994][19],** αν ένας γνωσιακός αλγόριθμος είναι ορθός, δηλαδή ο πράκτορας γνωρίζει σαφώς τον τύποφ σε μία κατάσταση, εφόσον φυσικά ο φ είναι αληθής στην κατάσταση αυτή, τότε το $X\varphi \Rightarrow \varphi$είναι έγκυρο σε κάθε δομή, η οποία χρησιμοποιεί τον γνωσιακό αλγόριθμο αυτό. Αξίζει να σημειωθεί, ότι στο πλαίσιο της επαγωγικής αλγοριθμικής γνώσης, μπορούμε να προσδιορίσουμε πλήρως τις ιδιότητες του X, εκμεταλλευόμενοι τη δομή των επαγωγικών

συστημάτων. Το υπόλοιπο της παραγράφου αυτής συγκεκριμενοποιεί στην ουσία αυτή την πρόταση.

Σαν πρώτο βήμα παρουσιάζουμε γενικά ένα σύστημα αξιωμάτων, ανεξάρτητα από τους υπάρχοντες επαγωγικούς κανόνες του συστήματος αυτού. Αυτό θα δημιουργήσει τη βάση για περεταίρω διαδικασίες σχηματισμού αξιωμάτων. Αρχικά χρειαζόμαστε αξιώματα και συμπερασματικούς κανόνες, οι οποίοι θα αποδίδουν την προτασιακή συλλογιστική μας στην υπάρχουσα λογική:

Taut. Όλες οι περιπτώσεις προτασιακών ταυτολογιών.

MP. Από τον φ και το $\varphi \Rightarrow \psi$ συμπεραίνουμε τον ψ.

Το αξίωμα Taut μπορεί να αντικατασταθεί από μία διαδικασία δημιουργίας αξιωμάτων προτασιακών ταυτολογιών **[Enderton 1972][20]**. Τα ακόλουθα γνωστά αξιώματα και συμπερασματικοί κανόνες αποδίδουν τις ιδιότητες ενός γνωσιακού τελεστή.

K1. $(K\varphi \wedge K(\varphi \Rightarrow \psi)) \Rightarrow K\psi$.

K2. Από τονφσυμπεραίνουμε το $K\varphi$.

K3. $K\varphi \Rightarrow \varphi$

K4.$K\varphi \Rightarrow KK\varphi$

K5. $\neg K\varphi \Rightarrow K\neg K\varphi$

Μεταπίπτουμε τώρα σε επαγωγική αλγοριθμική γνώση. Προφανώς ο $X\varphi$ δεν ικανοποιεί πολλές ιδιότητες, διότι δεν έγινε καμία υπόθεση, σε σχέση με τα επαγωγικά συστήματα. Η επαγωγική αλγοριθμική γνώση, ερμηνεύεται σε σχέση με παρατηρήσεις στην παρούσα κατάσταση και δύο καταστάσεις δεν μπορούν να διαχωριστούν από έναν πράκτορα, αν οι ίδιες παρατηρήσεις γίνονται και στις δύο καταστάσεις. Για τον λόγο αυτό, οι πράκτορες

μπορούν να αντιλαμβάνονται, αν γνωρίζουν με βεβαιότητα ένα δεδομένο ή όχι. Αυτό αποδίδεται με το ακόλουθο αξίωμα:

X1. $X\varphi \Rightarrow KX\varphi$.

Από τα K1 – K5 αξιώματα και το X1 είναι εύκολο να συνάγουμε το $\neg X\varphi \Rightarrow K\neg X\varphi$. Επιπλέον όλες οι παρατηρήσεις είναι σαφώς γνωστές. Το γεγονός αυτό εκφράζεται από το ακόλουθο αξίωμα.

X2. $Ob(p) \Rightarrow XOb(p)$.

Το αξίωμα X2 απλά τυποποιεί την ακόλουθη ιδιότητα όπως καθορίστηκε στην 2ηπαράγραφο της εργασίας: για όλους του όρους tενός επαγωγικού συστήματος D, ισχύει $t \vdash_D t$. Τέλος πρέπει να αποδώσουμε το γεγονός ότι δυσδιάκριτες καταστάσεις μπορεί να έχουν ακριβώς τις ίδιες παρατηρήσεις. Έτσι έχουμε το ακόλουθο:

X3. $Ob(p) \Rightarrow KOb(p)$.

Είναι εύκολο λοιπόν να καταλάβουμε, ότι ο τύπος $\neg Ob(p) \Rightarrow K\neg Ob(p)$είναι «αποδείξιμος» μέσω του X3 και των K1-K5.

Θεωρούμε τώρα το σύνολο αξιωμάτων AX, αποτελείται από τα αξιώματα Taut, MP, K1-K5 και X1-X3. Χωρίς περεταίρω υποθέσεις επί των επαγωγικών συστημάτων, κατανοούμε ότι το AX χαρακτηρίζει τη συλλογιστική της επαγωγικής αλγοριθμικής γνώσης.

Θεώρημα 2.4.1. Η διαδικασία δημιουργίας των αξιωμάτων που ανήκουν στο AX είναι ορθή και πλήρης στη λογική $L^{KD}(\Sigma)$, σε σχέση με το σύνολο δομών– αρχιτεκτονική $\mathcal{M}(\Sigma)$.

Μπορούμε δε να πούμε περισσότερα για τις δομές επαγωγικής αλγοριθμικής γνώσης, οι οποίες ορίζονται από ένα συγκεκριμένο επαγωγικό σύστημα. Μπορούμε ουσιαστικά να αποδώσουμε τη συλλογιστική, σε σχέση με το σύστημα αυτό της λογικής μας. Η βασική ιδέα είναι να μεταφράσουμε του επαγωγικούς κανόνες του συστήματος, σε τύπους της λογικής $L^{KD}(\Sigma)$. Ένας επαγωγικός κανόνας της μορφής $t_1, \dots . t_n \rhd t$ στο σύστημα Dμεταφράζεται στον τύπο $(Xt_1^R \wedge \dots \wedge Xt_n^R) \Rightarrow Xt^R$, με δεδομένο ότι ένας κενός σύνδεσμος είναι κάτι το αληθές. Έπειτα καθορίζουμε τον τύπο t^R, ο οποίοςαντιστοιχείστον όρο tεπαγωγικά στη δομή M που αυτή ανήκει: το $true^R$ αντιστοιχεί στο $\alpha\lambda\eta\theta\acute{\varepsilon}\varsigma$ (true), το $false^R$στο $\psi\varepsilon\upsilon\delta\acute{\varepsilon}\varsigma$ (false), το $(not(t))^R$ στο $\neg(t)^R$, το $(and(t_1,t_2))^R$ στο $t_1^R \wedge t_2^R$, το $(know(t))^R$ στο $K(t^R)$, το $(xknow(t))^R$ στο $X(t^R)$, το $(Ob(t))^R$ στο $Ob(t^R)$ (αν το $t \in T_\Sigma^g$) και το t^R στο t για όλους τους υπόλοιπους όρους t. Λαμβάνουμε λοιπόν τα αποτελέσματα της ερμηνείας αυτής ως ένα σύστημα αξιωμάτων, όπου οι μεταβλητές των όρων $t_1, \dots . t_n, t$, λειτουργούν ως ένα σύστημα περαιτέρω συμβολικών μεταβλητών (μετά-μεταβλητές), οι οποίες αναπαρίστανται με κατάλληλα αλγεβρικά σύμβολα ***[6]** **(Παράρτημα «Α»)**

Είναι εύκολο λοιπόν να αντιληφθούμε, ότι το $(t^T)^R = t$ για όλους τους όρους t. Επιπλέον δεν ερμηνεύουμε τους κατασκευαστές στο Σ^{KD}, οι οποίοι εμφανίζονται με άλλους κατασκευαστές στο Σ εντός ενός όρου. (Επομένως, αυτοί οι κατασκευαστές δεν θα μπορούν πάντα να ερμηνευθούν στους δοθέντες κάθε φορά τύπους). Θεωρούμε AX^D ένα σύστημα αξιωμάτων, το οποίο προκύπτει με αυτόν τον τρόπο για το επαγωγικό σύστημα $\Sigma \cup \Sigma^{KD}$.

Ένα απλό επιχείρημα δείχνει ότι το σύστημα αξιωμάτων AX, ενισχυμένο με το σύστημα AX^D δεν είναι πλήρες στο $\mathcal{M}_D(\Sigma)$, εφόσον υπάρχουν τύποι της μορφής $X\psi$, οι οποίοι δεν μπορούν να είναι αληθείς σε κάθε δομή στο $\mathcal{M}_D(\Sigma)$, όπου το ψ^T, δεν προκύπτει από κάθε σύνολο παρατηρήσεων που χρησιμοποιεί το επαγωγικό σύστημα D. Έτσι το $\neg X\psi$ είναι έγκυρο για τα ψ, αλλά τα παραπάνω αξιώματα δεν μπορούν να αποδείξουν το $\neg X\psi$. Με άλλα λόγια, τα αξιώματα στο AX^Dαποδίδουν την επαγωγικότητα στο D, αλλά όχι

το αντίστροφο. Μπορούμε ωστόσο να επιτύχουμε την πληρότητα του, σε σχέση με μία πιο γενική κατηγορία δομών, δηλαδή με τις δομές που χρησιμοποιούν ένα επαγωγικό σύστημα, το οποίο περιέχει τουλάχιστον όλους τους επαγωγικούς κανόνες του συστήματος D. Θεωρούμε λοιπόν το $\mathcal{M}_{D\subseteq}(\Sigma)$ ως το σύνολο των δομών M, έτσι ώστε να υπάρχει ένα D' με $D \subseteq D'$και $M \in \mathcal{M}_D{}'$.

Θεώρημα 2.4.2. Το σύστημα αξιωμάτων AX ενισχυμένο με τα αξιώματα του συστήματος AX^D είναι έγκυρο και πλήρες στην λογική $L^{KD}(\Sigma)$, σε σχέση με τη δομή $\mathcal{M}_{D\subseteq}(\Sigma)$.

Αν περιορίσουμε τους υπό εξέταση τύπους, μπορούμε να επιτύχουμε πληρότητα σε σχέση με το $\mathcal{M}_D(\Sigma)$. Αυτό προκύπτει απευθείας από το γεγονός, ότι ένα σύστημα αξιωμάτων μπορεί να αποδείξει όλες τις επαγωγές στο D, αλλά όχι το αντίστροφο. Εδώ πρέπει να καθορίσουμε ότι: μία μέγιστης σημασίας παρουσία ενός τύπου επαγωγικής αλγοριθμικής γνώσηςψ, εντός ενός άλλου τύπου φ είναι μία παρουσία, η οποία δεν προκύπτει μέσω του τελεστή X. Η παρουσία του τύπου ψ, χαρακτηρίζεται θετική, αν υφίσταται υπό την παρουσία ισάριθμων αρνητικών εμφανίσεων του ιδίου.

Θεώρημα 2.4.3. Θεωρούμε τον τύπο φ της λογικής $L^{KD}(\Sigma)$, στον οποίο κάθε μέγιστης σημασίας παρουσία ενός τύπου ψ είναι θετική. Τότε ο τύπος φ είναι έγκυρος στο $\mathcal{M}_D(\Sigma)$, αν και μόνο ανοφ είναι «αποδείξιμος» στο σύστημα αξιωμάτων AX, ενισχυμένο με το σύστημα AX^D. Συγκεκριμένα το Θεώρημα 2.4.3 ορίζει, ότι ο τύπος της μορφής $X\varphi$ είναι έγκυρος στο $\mathcal{M}_D(\Sigma)$, αν και μόνο αν το $X\varphi$ είναι αποδείξιμο στο σύστημα αξιωμάτων AX, ενισχυμένο με το σύστημα AX^D.

§2.5 Διαδικασίες Επιβεβαίωσης «Αποφασιστικότητας» Τύπων που Ανήκουν σε Δομές

Σε αυτή την παράγραφο θα μελετήσουμε το πρόβλημα «αποφασιστικότητας» ενός τύπου στη λογική $L^{KD}(\Sigma)$, δηλαδή το πρόβλημα της επιβεβαίωσης για το αν έναν δοθείς τύπος είναι «ικανοποιητικός» ή όχι σε αυτή. Και πάλι τονίζουμε εδώ, ότι στην περίπτωσή μας υπάρχει μόνο ένας πράκτορας. Η περίπτωση της παρουσίας πολλών πρακτόρων μεταβάλει την πολυπλοκότητα, όπως θα δούμε στην Παράγραφο 6. Εφόσον η λογική $L^{KD}(\Sigma)$ επεκτείνει τη λογική της γενικής γνώσης, όπου ο γνωσιακός τελεστής ερμηνεύεται με μία ισοδύναμη σχέση (στην περίπτωσή μας το σύμβολο $\sim$ σημαίνει ότι δύο καταστάσεις περιέχουν τις ίδιες παρατηρήσεις) και εφόσον η «αποφασιστικότητα» του προβλήματος (για τη γενική γνώση) είναι δύσκολη, η «πολυπλοκότητα» στο να ελέγξουμε αν το πρόβλημα είναι και «ικανοποιητικό» στη λογική $L^{KD}(\Sigma)$ είναι τουλάχιστον εξίσου μεγάλη. **Με τον όρο «πολυπλοκότητα» εννοούμε το πόσο αργά ή γρήγορα αποδίδει ένα συγκεκριμένος αλγόριθμος.** Μπορούμε να εκμεταλλευτούμε το γεγονός ότι υπάρχει μία στενή σχέση μεταξύ των δύο λογικών αυτών, προκειμένου να προσδιορίσουμε την «πολυπλοκότητα» της λύσης αυτής.

Μετράμε την «πολυπλοκότητα» σε σχέση με το μέγεθος των τύπων. Καθορίζουμε το μέγεθος $|t|$ ενός όρου t, ότι είναι ένας αριθμός συμβόλων, τα οποία απαιτούνται για να αποδώσουμε τον όροt, όπου κάθε αριθμητικός τελεστής, προσμετράται ως ξεχωριστή μονάδα.

Μπορούμε λοιπόν να παραθέσουμε τα αποτελέσματα της πολυπλοκότητας. Προκύπτει ότι προσθέτοντας έναν τελεστή επαγωγικής αλγοριθμικής γνώσης στη λογική της γενικής γνώσης με μία σχέση ισότητας, δεν μεταβάλλεται η πολυπλοκότητα του προβλήματος «αποφασιστικότητας». Το να αποφασίσουμε την «ικανοποιητικότητα» ενός τύπου στη λογική $L^{KD}(\Sigma)$ είναι ουσιαστικά το ίδιο, όπως το να αποφασίσουμε την «ικανοποιητικότητα» ενός τύπου της λογικής στη γενική γνώση με μία σχέση ισότητας. Η

διαφορά έγκειται στο ότι προκειμένου να αιτιολογήσουμε την επαγωγική αλγοριθμική γνώση, χρειάζεται να κατασκευάσουμε ένα επαγωγικό σύστημα με συγκεκριμένους επαγωγικούς κανόνες, οι οποίοι θα επαρκούν ώστε να ικανοποιούνται οι υπο-τύποι– τελεστές $X\varphi$ στον κυρίως τύπο.

Θεώρημα 2.5.1. Το πρόβλημα για το αν ένας τύπος φ της λογικής $L^{KD}(\Sigma)$ είναι «ικανοποιητικός» στο $\mathcal{M}(\Sigma)$, είναι «αποφασιστικό» σε πολυωνυμικό χρόνο (NP-complete).

Τι συμβαίνει όμως αν προσαρμόσουμε ένα συγκεκριμένο επαγωγικό σύστημα και θέλουμε να επιτύχουμε ένας τύπος φ να είναι «ικανοποιητικός» σε μια δομή, η οποία διέπεται από το σύστημα αυτό; Η δυσκολία του προβλήματος αυτού, εξαρτάται εγγενώς από τη δυσκολία της επιβεβαίωσης για το αν η επαγωγή $\Gamma \vdash_D t$ ισχύει στο σύστημα D. Εφόσον το πρόβλημα αυτό μπορεί πιθανόν να είναι δύσκολο για συγκεκριμένα επαγωγικά συστήματα D, και η απόδοση των δεδομένων στη λογική μας πιθανόν να είναι επίσης δύσκολη στα επαγωγικά συστήματα αυτά. Η λογική $L^{KD}(\Sigma)$ περιλαμβάνει προτασιακούς συνδέσμους, γεγονός το οποίο μας δίδει εύκολα ένα κατώτερο όριο.

Θεώρημα 2.5.2. Για κάθε δοθέν επαγωγικό σύστημα D με υπογραφή $\Sigma \cup \Sigma^{KD}$, το πρόβλημα απόφασης για το αν ένας τύπος φ της λογικής $L^{KD}(\Sigma)$ είναι «ικανοποιητικός» στο σύνολο δομών $M_D(\Sigma)$, είναι δύσκολα «αποφασιστικός» σε πολυωνυμικό χρόνο (NP-hard).

Από την άλλη, αν το επαγωγικό σύστημα είναι «αποφασιστικό» σε πολυωνυμικό χρόνο (αν δηλαδή το πρόβλημα απόφασης για το αν μία επαγωγή $\Gamma \vdash_D t$ ισχύει στο D, μπορεί να επιλυθεί από μία μηχανή Τούρινγκ[7], σε χρόνο πολυωνυμικά διαμορφωμένο στα $|\Gamma|$ και $|t|$), τότε το πρόβλημα απόφασης για την λογική $L^{KD}(\Sigma)$ παραμένει σχετικά εύκολο, τουλάχιστον σε σχέση με μοντέλα λογικού μεγέθους. Για την ακρίβεια θεωρούμε το $M_D^n(\Sigma)$

ως μία κατηγορία, στην οποία ανήκουν όλες οι δομές επαγωγικής αλγοριθμικής γνώσης που χρησιμοποιούν το επαγωγικό σύστημα D, όπου ο αριθμός των παρατηρήσεων σε κάθε κατάσταση είναι το πολύ n.

Θεώρημα 2.5.3. Για κάθε δοθέν επαγωγικό σύστημα D με υπογραφή $\Sigma \cup \Sigma^{KD}$, το οποίο είναι «αποφασιστικό» σε πολυωνυμικό χρόνο και για κάθε πολυώνυμο $P(x)$, το πρόβλημα επιβεβαίωσης για το αν ένας τύπος φ μιας λογικής $L^{KD}(\Sigma)$ είναι «ικανοποιητικός» στο σύνολο δομών $M_D^{P(|\varphi|)}(\Sigma)$ είναι «αποφασιστικό» σε πολυωνυμικό χρόνο (NP-complete).

Υπάρχει μία κατηγορία επαγωγικών συστημάτων, η οποία είναι αποτελεσματικά «αποφασιστική» (σε πολυωνυμικό χρόνο) και έτσι μέσω του Θεωρήματος 2.5.3., οδηγούμαστε σε μία διαχειρίσιμη πολυπλοκότητα της λογικής $L^{KD}(\Sigma)$, η οποία ερμηνεύεται μέσω των συστημάτων αυτών. Ονομάζουμε μία επαγωγή τοπική σε ένα επαγωγικό σύστημα D, αν κάθε υπο-όρος ενός όρου στην επαγωγή αυτή, είναι κατάλληλος υπο-όρος του t, κατάλληλος όρος του Γ ή εμφανίζεται ως υπο-όρος ενός επαγωγικού κανόνα στο D.

Για κάθε επαγωγικό σύστημα D, το αν μία τοπική επαγωγή του όρου t ενός συνόλου Γ ισχύει, δύναται να αποφασισθεί σε πολυωνυμικό χρόνο στα $|\Gamma|$ και $|t|$. Ένα επαγωγικό σύστημα Dείναι τοπικό, αν όποτε ισχύει η επαγωγή$\Gamma \vdash_D t$, υπάρχει και μία τοπική επαγωγή από τον όρο tστο σύνολο όρων Γ **[McAllester 1993] [21]**. Έτσι αν το D είναι ένα τοπικό επαγωγικό σύστημα, η ύπαρξη μιας επαγωγής, σημαίνει την ύπαρξη μιας τοπικής επαγωγής και επομένως η επαγωγική σχέση $\vdash_D$ είναι «αποφασιστική» σε πολυωνυμικό χρόνο. Το επαγωγικό σύστημα στο Παράδειγμα 2.2.1. είναι τοπικό, ενώ προσθέτοντας τους επαγωγικούς κανόνες στο Παράδειγμα 2.3.3 σε κάθε τοπικό επαγωγικό σύστημα, προκύπτουν παντού τοπικά επαγωγικά συστήματα.

Θεώρημα 2.5.4. Για κάθε επαγωγικό σύστημα D με υπογραφή $\Sigma \cup \Sigma^{KD}$ και για κάθε πολυώνυμο $P(x)$, το πρόβλημα απόφασης για το αν ένας τύποςφ στη λογική $L^{KD}(\Sigma)$ είναι «ικανοποιητικός» στο σύνολο δομών $M_D^{P(|\varphi|)}(\Sigma)$ είναι «αποφασιστικό» σε πολυωνυμικό χρόνο (NP-complete).

§2.6 Απόδοση Συλλογιστικής Πολλών Πρακτόρων στο Ίδιο Σύστημα

Το πλαίσιο που έχουμε περιγράψει μέχρι στιγμής, επεκτείνεται στην περίπτωση των πολλαπλών πρακτόρων με αναλογικό τρόπο. Η επέκταση αυτή είναι παρόμοια με την επέκταση των βοηθητικών λογικών της γνώσης σε πολλούς πράκτορες **[Fagin, Halpern, Moses, and Vardi 1995] [22]**. Η μόνη προσθήκη που πρέπει να κάνουμε, είναι να εξοπλίσουμε τον κάθε πράκτορα με ένα δικό του επαγωγικό σύστημα.

Θεωρούμε μία ομάδα πρακτόρων $1, \ldots . n$γιαλόγους απλότητας. Καθορίζουμε τη λογική $L_n^{KD}(\Sigma)$όπως κάναμε με τη λογική $L^{KD}(\Sigma)$, με τη διαφορά ότι οι τελεστές K_i, X_iκαι Ob_iδημιουργούνται από έναν πράκτορα. Εκ των προτέρων, δεν υπάρχει κάποια δυσκολία στην απόδοση της συλλογιστικής σε αυτή τη λογική όπως κάναμε και στην Παράγραφο 3. Δυστυχώς, αυτό δεν επιτρέπει στον πράκτορα να αποδώσει σαφώς τη γνώση ενός άλλου πράκτορα. Προκειμένου να γίνει αυτό, χρειάζεται να τροποποιήσουμε και να επεκτείνουμε το πλαίσιο ανάλυσης. Όπως προηγουμένως, θεωρούμε επαγωγικά συστήματα με υπογραφή Σ, τα οποία επεκτείνονται με ένα σύνολο κατασκευαστών $\Sigma_n^{KD} = \{true, false, not, and\} \cup \bigcup_{i=1}^{n}\{ob_i, know_i, xknow_i\}$, όπου τα $true, false$ έχουν βαθμό 0, τα $ob_i,\ know_i,\ xknow_i,\ not$ έχουν βαθμό 1 και το and έχει βαθμό 2.

Μία δομή επαγωγικής αλγοριθμικής γνώσης με nπράκτορες είναι μία πλειάδα συνόλων $M = (S, \pi, D_1, \ldots..D_n)$, όπου το S είναι ένα σύνολο καταστάσεων, το π είναι η ερμηνεία των πρωταρχικών προτάσεων και το D_iείναι ένα επαγωγικό σύστημα με υπογραφή $\Sigma \cup \Sigma^{KD}$. Κάθε κατάσταση sστο Sείναι της μορφής $(e, O_1, \ldots..O_n)$,όπου το e αποδίδει τη γενική κατάσταση του συστήματος και το O_iείναι ένα πεπερασμένο σύνολο παρατηρήσεων από T_Σ^g(βασικών όρων), που αναπαριστούν τις παρατηρήσεις που ο πράκτορας iέκανε σε εκείνη την κατάσταση. Η ερμηνεία π σχετίζεται με κάθε κατάσταση, στην οποία το σύνολο

των πρωταρχικών προτάσεων είναι αληθές, έτσι ώστε για όλες τις πρωταρχικές προτάσεις $p \in T_{\Sigma}^{g}$, να έχουμε $\pi(s)(p) \in \{true, false\}$.

Ας θεωρήσουμε το $M_D(\Sigma)$ το σύνολοόλων των δομών επαγωγικής αλγοριθμικής γνώσης που χρησιμοποιούν επαγωγικά συστήματα με υπογραφή $\Sigma \cup \Sigma^{KD}$για κάθε πράκτορα. Για τα συγκεκριμένα επαγωγικά συστήματα $D_1,\ldots..D_n$ με υπογραφή $\Sigma \cup \Sigma^{KD}$, ας είναι το $M_{D_1,\ldots..D_n}(\Sigma)$ το σύνολο όλων των δομών επαγωγικής αλγοριθμικής γνώσης για nπράκτορες με επαγωγικά συστήματα $D_1,\ldots..D_n$(δηλαδή ο πράκτορας i χρησιμοποιεί το επαγωγικό σύστημα D_i.

Οι υπόλοιποι ορισμοί γενικεύονται με παρόμοιο τρόπο. Καθορίζουμε για κάθε πράκτορα μία σχέση, η οποία αποδίδει τις καταστάσεις αυτές που ο πράκτορας δεν δύναται να διαχωρίσει μεταξύ τους, με βάση τις παρατηρήσεις του. Πιο συγκεκριμένα, θεωρούμε ότι ισχύει το $s \sim_i s'$ αν και μόνο αν $s = (e, O_1,\ldots..O_n)$ και $s = (e', O'_1, \ldots..O'_n)$για κάποια e, e'και σύνολα παρατηρήσεων $O_1,\ldots..O_n$, $O'_1,\ldots..O'_n$ με $O_i = O'_i$. Και πάλι εδώ το $\sim_I$ είναι μία σχέση ισοδυναμίας μεταξύ των καταστάσεων.

Η ερμηνεία του τύπου φ σε έναν άλλο τύπο φ^T του επαγωγικού συστήματος, εδώ λαμβάνει υπόψη της τα ονόματα των πρακτόρων. Όπως αναμένονταν, το p^Tισούται με το pγια κάθε προκείμενη p, το $(\neg\varphi^T)$ ισούται με το $\text{not}(\varphi^T)$, $(\varphi \wedge \psi)^T$ ισούται με το $\text{and}(\varphi^T, \psi^T)$, το $(K_i\varphi)^T$ με το $know_i(\varphi^T)$, το $(X_i\varphi)^T$ με το$know_i(\varphi^T)$ και το $(Ob_i(p))^T$με το $ob_i(p)$.

Η σημασιολογία είναι ακριβώς ίδια με αυτή της Παραγράφου 3, εκτός από τους ακόλουθους κανόνες για τα $K_i\varphi$, $X_i\varphi$ και $Ob_i(p)$:

$(M, s) \models K_i\varphi$ αν $(M, s') \models \varphi$ *για όλα τα* $s' \sim_i s$

$(M, s) \models X_i\varphi$ *αν* $s = (e, O_1, \ldots..O_n)$*και* $\{ob(p) \mid p \in O_i\} \vdash_D \varphi^T$

$(M, s) \models Ob_i(p)$ *αν* $s = (e, O_1, \ldots..O_n)$*και* $p \in O_i$

Παράδειγμα 2.6.1. Οι **[Kaplan και Schubert 2000][23]** μελέτησαν ένα φαινόμενο, το οποίο αποκαλούν «μιμητικό συμπέρασμα», όπου κατά προσέγγιση, ένας πράκτορας δύναται να ανακατασκευάσει τη συλλογιστική ενός άλλου πράκτορα. Μπορούμε να αποδώσουμε το φαινόμενο αυτό, κάνοντας κατάλληλες υποθέσεις στο επαγωγικό σύστημα ενός πράκτορα. (Οι Kaplan και Schubert δουλεύουν με διαφορετικά δεδομένα – υποθέτουν ότι στον μηχανισμό εξαγωγής αποτελεσμάτων δίδονται σαφείς συναρτήσεις και έτσι δουλεύουν με δεδομένα παρόμοια με αυτά της «αναθεώρησης πεποιθήσεων» **[Alchourro´n, Gardenfors, and Makinson 1985] [24]**.) Θεωρούμε λοιπόν, ότι ένα επαγωγικό σύστημα D_i ενός πράκτορα i επιτρέπει σε έναν πράκτορα j να κάνει ένα μιμητικό συμπέρασμα με ένα επαγωγικό σύστημα D_j αν το D_i περιέχει έναν κανόνα $ob_j(t) \rhd xknow_j(ob_j(t))$ και για κάθε κανόνα $t_1,\ldots..t_k \rhd t$ του D_j, υπάρχει ένας αντίστοιχος κανόνας $xknow_j(t_1),\ldots..,xknow_j(t_k) \rhd xknow_j(t)$ στο D_i. Είναι εύκολο τότε να ελέγξουμε, ότι αν ισχύει $(M,s) \models X_j\varphi$ για κάποια κατάσταση $s = (e, O_1,\ldots..O_n)$ με $\{p_1,\ldots..p_k\} \subseteq O_j$, και $(M,s) \models X_i(ob_j(p_1)) \wedge\cdots\wedge X_i(ob_j(p_k))$, τότε $(M,s) \models X_iX_j\varphi$. Να σημειώσουμε εδώ ότι αυτή η απαγωγή υποθέτει ότι ο πράκτορας i δύναται σαφώς να προσδιορίσει, ότι ο πράκτορας j έχει παρατηρήσει τα $p_1,\ldots..p_k$.

Όσον αφορά στα συστήματα αξιωμάτων, μπορούμε να χρησιμοποιήσουμε τα αποτελέσματα της παραγράφου 4. Αρκεί να θεωρήσουμε ένα σύστημα αξιωμάτων, όπου τα K1-K5 πλέον αναφέρονται στον τελεστή K_i και όχι απλά στον K. Για παράδειγμα, το K1 γίνεται $K_i\varphi \wedge K_i(\varphi \Rightarrow \psi) \Rightarrow K_i\psi$. Για τα X2 και X3, χρειάζεται να περιορίσουμε περεταίρω τις παρατηρήσεις, ώστε αυτές να είναι οι υπό εξέταση του πράκτορα: $Ob_i(p) \Rightarrow XOb_i(p)$ και $Ob_i(p) \Rightarrow KOb_i(p)$. Θεωρούμε το AX_n το προκύπτον σύστημα αξιωμάτων.

Θεώρημα 2.6.2. Το σύστημα αξιωμάτων AX_n είναι έγκυρο και πλήρες σε μία λογική $L_n^{KD}(\Sigma)$, σε σχέση με το σύνολο δομών $M_n(\Sigma)$.

Όπως και στην περίπτωση του ενός πράκτορα, μπορούμε να αποδώσουμε τη συλλογιστική σε σχέση με συγκεκριμένα επαγωγικά συστήματα (ένα για κάθε πράκτορα) εντός της λογικής μας. Και πάλι ερμηνεύουμε τους επαγωγικούς κανόνες των επαγωγικών συστημάτων σε τύπους της λογικής $L_n^{KD}(\Sigma)$ που χρησιμοποιούμε. Θεωρούμε λοιπόν ένα επαγωγικό σύστημα D_i ενός πράκτορα i. Ένας επαγωγικός κανόνας της μορφής $t_1,\ldots..t_k \rhd t$στο D_i ερμηνεύεται σε μία συνάρτηση $(X_i t_1^R \wedge \ldots \wedge X_i t_n^R) \Rightarrow X_i t^R$. Καθορίζουμε τη συνάρτηση t^R, η οποία αντιστοιχεί στον όρο tμε απαγωγή στη δομή του t: το $true^R$ σημαίνει $\alpha\lambda\eta\theta\acute{\epsilon}\varsigma(true)$, το $false^R$ σημαίνει $\psi\epsilon\upsilon\delta\acute{\epsilon}\varsigma(false)$, το $\big(not(t)\big)R$είναι το $\neg(t)^R$, το $(and(t_1,t_2))^R$ είναι το $t_1^R \wedge t_2^R$, το $(know_i(t))^R$ είναι το $K_i(t^R)$, το $(xknow_i(t))^R$ είναι το $X_i(t^R)$, το $(Ob_i(t))^R$ είναι το $Ob_i(t^R)$ (αν το $t \in T_\Sigma^g$) και το t^R είναι το t για όλους τους υπόλοιπους όρους t. Όπως και στην Παράγραφο 4, από μια τέτοια ερμηνεία προκύπτει ένα σύστημα αξιωμάτων, όπου βλέπουμε τις μεταβλητές στο $t_1,\ldots..t_k$, t ως σχήμα μετα-μεταβλητών, οι οποίες αντικαθίστανται από κατάλληλα στοιχεία αλγεβρικών όρων. Θεωρούμε το $AX_n^{D_i}$ ένα σύνολο αξιωμάτων, το οποίο προκύπτει με αυτόν τον τρόπο για το επαγωγικό σύστημα D_i του πράκτορα i.

Όπως και στην περίπτωση του ενός πράκτορα, δεν μπορούμε να αποδώσουμε ακριβώς τη συλλογιστική στις δομές, όπου ο πράκτορας i χρησιμοποιεί το επαγωγικό σύστημα D_i, εφόσον δεν μπορούμε να αποδώσουμε την αντίστροφη επαγωγικότητα στην παρούσα λογική. Για τον λόγο αυτό η πληρότητα μπορεί να επιτευχθεί, σε σχέση με μεγαλύτερες κατηγορίες δομών. Θεωρούμε λοιπόν ότι το $\mathcal{M}_{D_1,\ldots..D_n\subseteq}(\Sigma)$ είναι όλες οι δομές M μιας κατηγορίας, έτσι ώστε να υπάρχουν $D'_1,\ldots..D'_n$ με τα $D_1 \subseteq D'_1,\ldots..D_n \subseteq D'_n$ και $M \in \mathcal{M}_{D'_1,\ldots..D'_n}$.

Θεώρημα 2.6.3. Το σύστημα αξιωμάτων AX_nενισχυμένο με τα αξιώματα $AX_n^{D_1},\ldots.,$ $AX_n^{D_n}$είναι έγκυρο και πλήρες στη λογική $L_n^{KD}(\Sigma)$, σε σχέση με το σύνολο δομών $\mathcal{M}_{D_1,\ldots..D_n\subseteq}(\Sigma)$.

Η πολυπλοκότητα ενός προβλήματος απόφασης στην περίπτωση των πολλαπλών πρακτόρων, αντανακλά την πολυπλοκότητα του προβλήματος απόφασης σε μία βοηθητική λογική με πολλούς πράκτορες. Η λογική $L_n^{KD}(\Sigma)$ επεκτείνει τη λογική της γνώσης με σχέσεις ισότητας για nπράκτορες και είναι γνωστό, ότι το πρόβλημα απόφασης για τη λογική αυτή είναι πλήρες σε πολυωνυμικό χώρο (PSPACE-complete)**[Halpern and Moses 1992][25]** .

Θεώρημα 2.6.4. Αν $n \geq 2$, το πρόβλημα απόφασης για το αν μία συνάρτηση φ μιας λογικής $L_n^{KD}(\Sigma)$ είναι «ικανοποιητική» στο $\mathcal{M}_n(\Sigma)$ είναι «αποφασιστικό» εξεταζόμενο σε πολυωνυμικό χώρο (PSPACE-complete).

Δεν υπάρχει ένα ξεκάθαρο υποψήφιο θεώρημα, αντίστοιχο του 5.3 για την περίπτωση των πολλών πρακτόρων. Υποθέτοντας ότι κάθε πράκτορας χρησιμοποιεί ανιχνεύσιμα επαγωγικά συστήματα, προκύπτει ένα πρόβλημα ανώτερου ορίου με χρονικό όριο στη λογική $L_n^{KD}(\Sigma)$, ενώ το κατώτερο όριο που μπορούμε να λάβουμε, είναι το ίδιο όπως αυτό στο Θεώρημα 2.6.4, δηλαδή ότι το πρόβλημα είναι «αποφασιστικό» εξεταζόμενο σε πολυωνυμικό χώρο (PSPACE-complete).

ΚΕΦΑΛΑΙΟ 3^{Ο}

§3. Επαγωγική Αλγοριθμική Γνώση και Συστήματα Ασφάλειας Δεδομένων

§3.1 Η Εφαρμογή της Επαγωγικής Αλγοριθμικής Γνώσης στα Συστήματα Διαχείρισης Ασφάλειας Προσωπικών Δεδομένων.

Στο σημείο αυτό κρίνεται σκόπιμο να παρουσιάσουμε έναν τομέα της επιστήμης της πληροφορικής, όπου η επαγωγική αλγοριθμική γνώση έχει ουσιαστική εφαρμογή. Η ανάπτυξη συστημάτων διαχείρισης ασφάλειας προσωπικών δεδομένων, βασίζεται σε σύνολα δομών (αρχιτεκτονικές), οι οποίες χρησιμοποιούν μοντέλα τυπικής λογικής και συνεπώς την επαγωγική αλγοριθμική γνώση προκειμένου να επιτύχουν την ομαλή και αποτελεσματική λειτουργία τους.

Η παροχή ασφάλειας δεδομένων μέσω της κατάλληλης σχεδίασης των συστημάτων επεξεργασίας τους, θεωρείται τα τελευταία χρόνια απαραίτητη προϋπόθεση για τη βελτίωση της προστασίας των απορρήτων προσωπικών δεδομένων, στη συνεχώς αυξανόμενη σε πληθυσμό διεθνή ψηφιακή κοινότητα **[M. Langheinrich 2001, Y. Poullet 2010] [26] [27]**. Στο μέλλον πιστεύεται, ότι η προστασία αυτή θα αποτελεί νόμιμη υποχρέωση, στο πλαίσιο της Ευρωπαϊκής Ένωσης, αν το παρόν προσχέδιο του Κανονισμού Προστασίας Προσωπικών Δεδομένων ψηφισθεί από το ευρωπαϊκό κοινοβούλιο **[E.C. European Commission 2013] [28]**. Το γεγονός ότι μελλοντικά κανονιστικά πλαίσια προωθούν και επιβάλλουν την ασφάλεια των προσωπικών δεδομένων, μέσω της σχεδίασης των συστημάτων

επεξεργασίας τους, είναι αναμφίβολα ένα θετικό βήμα προς αυτή την κατεύθυνση. Η εφαρμογή της ωστόσο, θα χρειαστεί χρόνο και σίγουρα θα καθυστερήσει σε πολλές χώρες του κόσμου, όπου η ερμηνεία των αρχών και της ηθικής, σε θέματα ασφάλειας προσωπικών δεδομένων, διαφέρουν σημαντικά. Για τον λόγο αυτό, η «υιοθεσία» τέτοιου είδους νόμων από τις ανά των κόσμο κυβερνήσεις, δεν πρέπει να θεωρηθεί ως τον μόνο τρόπο εφαρμογής των μέτρων. Υπάρχουν και άλλοι παράγοντες που μπορούν να επιβάλλουν την εφαρμογή τους, όπως οι κοινωνικές απαιτήσεις, οι οικονομικές συνθήκες και φυσικά η διαθεσιμότητα των τεχνικών επί του θέματος λύσεων. Αν και όλες αυτές οι διαστάσεις ανάλυσης του θέματος είναι απαραίτητες και όλα τα δυνατά μέσα θα έπρεπε να επιστρατευτούν για την εφαρμογή των μέτρων σε αυτή την κατεύθυνση, εμείς επιλέγουμε να εστιάσουμε (ως επιστημονικοί ερευνητές) στις τεχνικές λεπτομέρειες του θέματος αυτού.

Όπως έχει συζητηθεί από πολλούς συγγραφείς **[S. F. Gurses, C. Troncoso, and C. Diaz 2001, F. Kerschbaum 2012] [29] [30]**, ένα πλήθος από τεχνολογίες ενίσχυσης ασφάλειας προσωπικών δεδομένων (Privacy Enhancing Technologies "PETs") είναι πλέον διαθέσιμες, πράγμα το οποίο μπορεί να προσδώσει ισχυρές εγγυήσεις ασφαλείας σε αρκετές περιπτώσεις εφαρμογών. Αν και πάντα επιπλέον έρευνα στον τομέα των «PETs» θα απαιτείται, το ερώτημα που εγείρεται είναι ως προς το πώς θα προωθηθεί η εφαρμογή τους στις διάφορες επιχειρήσεις, που θα μπορούσαν να επωφεληθούν από αυτές. Η θέση που θέλουμε να αναδείξουμε εδώ έχει τρεις πτυχές:

1. Η υπάρχουσα στον τομέα γνώση, εστιάζει περισσότερο σε τεχνολογίες παρά σε μεθοδολογίες και σε συστατικά παρά σε δομές. Υποστηρίζουμε την ιδέα, ότι η παρεχόμενη ασφάλεια μέσω της σχεδίασης θα έπρεπε να εξετάζεται στο δομικό επίπεδο και να συσχετίζεται με κατάλληλες μεθοδολογίες. Μεταξύ άλλων, ένα πλεονέκτημα των δομικών περιγραφών είναι, ότι καθιστούν δυνατή μία πιο συστηματική εξερεύνηση του χώρου σχεδίασης.

2. Επειδή η απαίτηση για ασφάλεια προσωπικών δεδομένων είναι εγγενώς μία σύνθετη έννοια, η οποία επιπλέον συχνά καταλήγει (ή τουλάχιστον έτσι φαίνεται) να έρχεται σε αντίθεση με άλλες απαιτήσεις, οι μέθοδοι εφαρμογής μαθηματικής λογικής θα έπρεπε να διαδραματίζουν σημαντικό ρόλο στον τομέα αυτό. Ένα επιπλέον πλεονέκτημα των μεθόδων περιγραφής με μοντέλα μαθηματικής λογικής είναι, ότι μπορούν να καθορίσουν επακριβώς τις χρησιμοποιούμενες έννοιες και τις παρεχόμενες μιας λύσης εγγυήσεις, οι οποίες χρησιμοποιούνται για να αποδώσουν τις σχεδιαστικές επιλογές και ταυτόχρονα να αιτιολογήσουν την ύπαρξή τους, συντελώντας με αυτόν τον τρόπο και στην λογικά καθολική επεξήγηση της λειτουργίας του συστήματος.

3. Επειδή οι σχεδιαστές δεν είναι εν γένει ειδήμονες στη χρήση μεθόδων λογικής απόδοσης των δεδομένων, φιλικά προς αυτούς περιβάλλοντα θα έπρεπε να σχεδιάζονται, ώστε να απλοποιούν τα μοντέλα και να τα καθιστούν εύκολα χρησιμοποιήσιμα. Τέτοια περιβάλλοντα θα πρέπει να επιτρέπουν στους σχεδιαστές να εκφράζουν τις απαιτήσεις τους και να αλληλεπιδρούν με το σύστημα, ώστε να τα βελτιώνουν και να τα προσαρμόζουν στα δεδομένα τους, μέχρι να βρούνε την καταλληλότερη για αυτούς δομή λειτουργίας.

§3.2 Το Πλαίσιο Ανάλυσης

Ένα ευρύ φάσμα από τεχνικές έχουν προταθεί ή ακόμα και εφαρμοσθεί προς καθορισμό νέων τεχνολογιών ενίσχυσης ασφάλειας προσωπικών δεδομένων (Privacy Enhancing Technologies "PETs") κατά τις τελευταίες δεκαετίες **[Y. Deswarte and C. A. Melchor 2006, I. Goldberg 2007] [31] [32]** μεταξύ των οποίων και οι τεχνολογίες: αποδείξεων μηδενικής γνώσης, ασφαλών υπολογισμών πολλαπλών ομάδων ατόμων, ομοιόμορφης κρυπτογράφησης, δεσμεύσεων, ανάκτησης προσωπικών δεδομένων (Private Information Retrieval "PIR"), ανώνυμων διαπιστευτηρίων, ανώνυμων δίαυλων επικοινωνίας και έμπιστων συστημάτων σχεδίασης (Trusted Platform Modules "TPM"). Αυτές οι τεχνολογίες έχουν χρησιμοποιηθεί για να παράσχουν ισχυρές εγγυήσεις προστασίας

προσωπικών δεδομένων, σε έναν μεγάλο αριθμό συστημάτων ασφαλείας, όπως οι έξυπνοι μετρητές, οι αξιολογητές ηλεκτρονικής συμφόρησης, τα συστήματα αδιάλειπτων υπολογισμών και τα συστήματα παροχής υπηρεσιών με βάση την τοποθεσία. Αν και περισσότερες τεχνικές πάντα θα χρειάζονται, για να αντιμετωπισθούν νέες επιθέσεις από "hackers" στο αέναο κυνηγητό μεταξύ κλεφτών και αστυνόμων, υπάρχει ωστόσο πλέον ανάγκη για μία σοβαρή θεώρηση των συνθηκών εφαρμογής των υπαρχόντων τεχνολογιών ενίσχυσης προσωπικών δεδομένων στη βιομηχανία. Το γεγονός αυτό ξεκίνησε από ανακύπτουσες ανάγκες εξέλιξης κατάλληλων τεχνολογιών ή εργαλείων εξέλιξης της ασφάλειας των προσωπικών δεδομένων στο επίπεδο της σχεδίασης **[F. Kerschbaum 2012] [33]**. Έτσι συνάγουμε ότι, η ασφάλεια προσωπικών δεδομένων θα έπρεπε να εξετάζεται στο επίπεδο της σχεδίασης των αρχιτεκτονικών και ταυτόχρονα να υποστηρίζεται από κατάλληλα εργαλεία, τα οποία θα βασίζονται σε μοντέλα που χρησιμοποιούν μοντέλα μαθηματικής λογικής.

Ας εξετάσουμε αρχικά την περίπτωση της σχεδίασης κατάλληλων αρχιτεκτονικών. Πληθώρα ορισμών της αρχιτεκτονικής έχουν παρατεθεί στη σχετική βιβλιογραφία. Εδώ όμως θα χρησιμοποιήσουμε τον ορισμό που δίνουν οι **[L.Bass, P. Clements and R. Kazman 2013] [34]**. Αρχιτεκτονική ενός συστήματος είναι ένα σύνολο από δομές, οι οποίες χρειάζονται, προκειμένου αυτό να αποδοθεί με μαθηματική λογική, και το οποίο περιλαμβάνει στοιχεία υπολογιστικού λογισμικού και υπολογιστικών εξαρτημάτων, μεταξύ τους συσχετίσεις και ιδιότητες και των δύο. Τα συστατικά μίας αρχιτεκτονικής είναι διακριτές οντότητες, όπως δομοστοιχεία, δομικά συστατικά ή σύνδεσμοι. Στο πλαίσιο εξέτασης της ασφάλειας προσωπικών δεδομένων, τα δομικά συστατικά είναι ουσιαστικά οι τεχνολογίες ενίσχυσης προσωπικών δεδομένων (PETs) και σκοπός μιας αρχιτεκτονικής είναι ο συνδυασμός τους, προκειμένου να ικανοποιηθούν οι απαιτήσεις ασφαλείας του συστήματος. Για τον λόγο αυτό, μία αρχιτεκτονική είναι πρωταρχικά ένα αφηρημένο σχέδιο και το σχέδιο αυτό, είναι απαραίτητο για τον μετριασμό την πολυπλοκότητας ενός συστήματος, την οποία δεν μπορούμε να βρίσκουμε σαν εμπόδιο κάθε φορά που χρησιμοποιούμε το σύστημα.

Οι περισσότεροι λόγοι οι οποίοι προσδιορίζονται από τους ερευνητές, ώστε να επεξηγηθεί το γιατί οι αρχιτεκτονικές είναι σημαντικές στην κατεύθυνση της ασφάλειας προσωπικών δεδομένων μέσω της κατάλληλης σχεδίασης του συστήματος παρατίθενται παρακάτω:

- Η αρχιτεκτονική είναι φορέας των πιο πρώιμων και έτσι των πιο θεμελιωδών και δύσκολων να αλλάξουν αποφάσεων σχεδίασης: το να παραβλέψουμε αρχιτεκτονικές επιλογές, μπορεί να παρεμποδίσει σημαντικά την ενοποίηση απαιτήσεων ασφαλείας στο σχεδιαζόμενο σύστημα.

- Οι αρχιτεκτονικές καθοδηγούν τη δημιουργικότητα των σχεδιαστών, μειώνοντας σημαντικά την πολυπλοκότητα σχεδιασμού των συστημάτων: επειδή καθιστούν δυνατή την απόρριψη περιττών λεπτομερειών και ταυτόχρονα την εστίαση σε σημαντικά θέματα, οι αρχιτεκτονικές βοηθούν τους σχεδιαστές να προσδιορίσουν τις απαιτήσεις ασφαλείας και να συνδυάσουν τεχνολογίες ενίσχυσης προσωπικών δεδομένων (PETs) προς αυτή την κατεύθυνση.

- Μία αποτυπωμένη αρχιτεκτονική ενισχύει την επικοινωνία μεταξύ των εμπλεκομένων στο σύστημα: ένα αποτυπωμένο σχέδιο επιλογών πρέπει οπωσδήποτε να πληροί τις προδιαγραφές, οι οποίες προκύπτουν από εκτιμήσεις για την επίδρασή τους στην ασφάλεια προσωπικών δεδομένων (Privacy Impact Assessments "PIA") και από νομικά θέματα στο κομμάτι της ανάληψης ευθυνών σε περίπτωση διαρροής δεδομένων (E.C. European Commission 2013).

- Μία αρχιτεκτονική μπορεί να δομηθεί ώστε να είναι δυνατόν να μεταφερθεί και να επαναχρησιμοποιηθεί: για τον λόγο αυτό, οι αρχιτεκτονικές διαδραματίζουν έναν πολύ σημαντικό ρόλο στην βελτίωση εύρεσης λύσεων σε θέματα ασφαλείας, μέσω της δυνατότητας επαναχρησιμοποίησης τους. Το γεγονός αυτό θα μπορούσε να οδηγήσει σε σημαντική μείωση του κόστους σχεδίασης και στην ευκολότερη διασπορά της γνώσης μεταξύ των σχεδιαστών. Η τυποποίηση μιας τέτοιας διαδικασίας, θα μπορούσε να οδηγήσει σε μία θα λέγαμε «βιομηχανοποίηση» της παραγωγής των αρχιτεκτονικών.

Αν η επιλογή γίνει ώστε το σύστημα να λειτουργεί στο «αρχιτεκτονικό» επίπεδο, η επόμενη ερώτηση προς απάντηση είναι: «πώς ορίζονται, απεικονίζονται και χρησιμοποιούνται οι αρχιτεκτονικές; Πρακτικά, οι αρχιτεκτονικές συχνά περιγράφονται οπτικά, με χρήση διαφόρων ειδών γραφημάτων, τα οποία έχουν σχετικά υπομνήματα που καθορίζουν τις έννοιες των κόμβων και των διανυσμάτων και των διάφορων διαγραμμάτων (διαγράμματα επικοινωνιών, διαγράμματα ακολουθιών κ.α.). Το δεύτερο σημείο που θέλουμε να επισημάνουμε εδώ, είναι ότι αν και οι οπτικές απεικονίσεις είναι πολύ χρήσιμες (ειδικά όταν η χρήση τους τυποποιηθεί), η απόδοση των απαιτήσεων ασφαλείας προσωπικών δεδομένων είναι ένα τόσο λεπτό και πολύπλοκο θέμα, που πρέπει να καθορίζεται λεπτομερώς με όρους μαθηματικής λογικής. Οι ιδιότητες αυτές της χρησιμοποιούμενης μαθηματικής λογικής, πρέπει να δύναται να αποδειχθούν. Η απόδοση με μαθηματική λογική των θεμάτων ασφαλείας προσωπικών δεδομένων είναι πολύπλοκη για διάφορους λόγους: πρώτα απ' όλα είναι μία πολυσχιδής έννοια, η οποία πηγάζει από μία σειρά από αρχές, οι οποίες δεν προσδιορίζονται επακριβώς. Για τον λόγο αυτό είναι πρώτιστης σημασίας να οριοθετήσουμε το πρόβλημα, προκειμένου να καθορίσουμε επακριβώς τις πτυχές του αυτές που λαμβάνονται υπόψη και το πώς μεταφράζονται σε σχεδιαστικές επιλογές. Το γεγονός ότι όλες οι πτυχές της ασφάλειας προσωπικών δεδομένων δεν είναι εύκολο να αποδοθούν με μαθηματική λογική, δεν αποτελεί ένα ανυπέρβλητο εμπόδιο. Το σημαντικό είναι να δομηθούν κατάλληλα μοντέλα για τα δεδομένα του θέματος (όπως η ελαχιστοποίηση των χρησιμοποιούμενων δεδομένων), τα οποία θα είναι εύκολα τυποποιήσιμα λογικά και θα εμπλέκουν μια πιο σύνθετη συλλογιστική. Μια άλλη πηγή πολυπλοκότητας επί του θέματος αυτού, είναι το γεγονός ότι η ασφάλεια προσωπικών δεδομένων συχνά φαίνεται να έρχεται σε αντίθεση με άλλες απαιτήσεις όπως οι εγγυήσεις αυθεντικότητας ή ορθότητας, η αποτελεσματικότητα, χρηστικότητα κ.τ.λ. Συνοψίζοντας οι μέθοδοι χρήσης μαθηματικής λογικής θα έπρεπε:

- Να καθιστούν δυνατό τον επακριβή προσδιορισμό των υπαρχόντων εννοιών (απαιτήσεων, υποθέσεων, εγγυήσεων κ.τ.λ.).

- Να βοηθούν τους σχεδιαστές να εξερευνούν τον χώρο σχεδίασης και να αποδίδουν κατανοητά τις πιθανές επιλογές που διαθέτουν **[D. Le Metayer 2010] [35]**: πολλαπλές επιλογές είναι γενικά διαθέσιμες, ώστε να επιτυγχάνονται ένα σύνολο από λειτουργίες, μερικές εκ των οποίων σχετίζονται με θέματα ασφαλείας. Η μεγαλύτερη πρόκληση για τον σχεδιαστή αποτελεί η κατανόηση όλων αυτών των διατιθέμενων επιλογών, καθώς και των πλεονεκτημάτων και μειονεκτημάτων που αυτές επισύρουν.

- Να παρέχουν αρχειοθετημένο και σαφή τρόπο αιτιολόγησης του συνόλου των σχεδιαστικών επιλογών, οι οποίες θα συνεισέφεραν τα πλείστα στο κομμάτι των απαιτήσεων λογοδοσίας. Ο όρος λογοδοσία επεξηγείται, με βάση τον Νέο Κανονισμό Προστασίας Προσωπικών Δεδομένων (Data Protection Regulation) ως απορρέουσα υποχρέωση των διαχειριστών απορρήτων πληροφοριών: «Ο κάθε φορέας που διαθέτει και επεξεργάζεται πληροφορίες, θα πρέπει να υιοθετεί κατάλληλες πολιτικές και να εφαρμόζει κατάλληλα, τεχνικά εύκολα παρουσιάσιμα και οργανωτικά μέτρα, ώστε να διασφαλίζει την διαφάνεια της επεξεργασίας και διαχείρισης των προσωπικών δεδομένων που έχει στην κατοχή του».

Θα πρέπει να είναι ξεκάθαρο ωστόσο, ότι οι σχεδιαστές εν γένει, δεν είναι ειδήμονες στη χρήση μεθόδων απόδοσης μαθηματικής λογικής και έτσι φιλικά προς αυτούς περιβάλλοντα θα έπρεπε να σχεδιάζονται, ώστε να απλοποιούν τα μοντέλα και να τα καθιστούν εύκολα χρησιμοποιήσιμα. Τέτοια περιβάλλοντα θα πρέπει να επιτρέπουν στους σχεδιαστές να εκφράζουν τις απαιτήσεις τους και να αλληλεπιδρούν με το σύστημα, ώστε να τα βελτιώνουν και να τα προσαρμόζουν στα δεδομένα τους, μέχρι να βρούνε την καταλληλότερη για αυτούς δομή λειτουργίας.

§3.3 Ένα Απόσπασμα μιας Προτασιακής Λογικής για την Περιγραφή Συστημάτων Παροχής Ασφάλειας Προσωπικών Δεδομένων

Επειδή η ασφάλεια προσωπικών δεδομένων είναι στενά συνδεδεμένη με την έννοια της γνώσης, οι γνωσιακές λογικές σχηματίζουν μία ιδανική βάση για να αιτιολογήσουμε τα χαρακτηριστικά της ασφάλεια προσωπικών δεδομένων. Οι γνωσιακές λογικές είναι μία κατηγορία βοηθητικών λογικών, οι οποίες χρησιμοποιούν τη γνώση βοηθητικά, συνήθως με τη χρήση του τελεστή $K_i(\varphi)$, προκειμένου να εκφράσουν το γεγονός ότι ο πράκτορας i γνωρίζει τον όρο – συνάρτηση φ. Ωστόσο οι κανονικές γνωσιακές λογικές, οι οποίες βασίζονται στη σημασιολογία της θεωρία των πιθανών κόσμων, έχουν μία αδυναμία, η οποία τις καθιστά ακατάλληλες για την περίπτωση της ασφάλειας προσωπικών δεδομένων: το πρόβλημα αυτό αναφέρεται συχνά και ως «λογική παντογνωσία» **[J. Y. Halpern, R. Pucella 2011] [36]** (όπως έχουμε ήδη αναφέρει παραπάνω στην παρούσα εργασία). Πηγάζει από το γεγονός, ότι οι πράκτορες γνωρίζουν όλες τις λογικές συνέπειες της γνώσης που χρησιμοποιούν (διότι αυτές οι συνέπειες μπορεί να ισχύουν σε όλους τους πιθανούς κόσμους). Ένα ανεπιθύμητο αποτέλεσμα της λογικής παντογνωσίας θα μπορούσε να είναι για παράδειγμα, ότι ένας πράκτορας που γνωρίζει τη συνάρτηση – μηχανισμό κωδικοποίησης $C(v)$(ή ένα μέρος) μιας τιμής v, θα μπορούσε επίσης να γνωρίζει και την ίδια την τιμή υ (ή πιθανές τιμές της). Αυτός προφανώς δεν είναι αποδεκτό σε ένα επίσημο μοντέλο παροχής ασφάλειας προσωπικών δεδομένων, όπου οι συναρτήσεις - μηχανισμοί κωδικοποίησης χρησιμοποιούνται ακριβώς, προκειμένου να αποκρύπτονται οι αυθεντικές τιμές από τους παραλήπτες. Το θέμα αυτό σχετίζεται με το γεγονός, ότι οι κανονικές γνωσιακές λογικές, δεν λαμβάνουν υπόψη τους, τους περιορισμούς της υπολογιστικής ισχύος.

Για τον λόγο αυτό, είναι απαραίτητο να καθορίσουμε γνωσιακές λογικές, προκειμένου να αντιμετωπίσουμε τις διάφορες πτυχές της ασφάλειας προσωπικών δεδομένων και να

μοντελοποιήσουμε μία πληθώρα εννοιών και τεχνικών που θα συναντήσουμε. (π.χ. γνώση, γνώση μηδενικών αποδείξεων, έμπιστες γνώσεις κ.τ.λ.). Ας θεωρήσουμε παραδείγματος χάριν, το ακόλουθο απόσπασμα γνωσιακής λογικής:

$$\varphi ::= \varphi_0 \quad (1)$$
$$|\neg\varphi \quad (2)$$
$$|\varphi \wedge \varphi' \quad (3)$$
$$|K_i(\varphi_0) \quad (4)$$
$$|X_i(\varphi_0) \quad (5)$$

$$\varphi_0 ::= receive_{i,j}(x) \quad (6)$$
$$|receive_{i,j}(prim) \quad (7)$$
$$|trust_{i,j} \quad (8)$$
$$|compute_i(x = t) \quad (9)$$
$$|check_i(p) \quad (10)$$
$$|has_i(x) \quad (11)$$
$$|prim \mid p \mid \varphi_0 \wedge \varphi'_0 \quad (12)$$

$$prim ::= proof_{i,j}(p) \mid att \mid prim \wedge prim' \quad (13)$$
$$p ::= att \mid eq \mid p \wedge p' \quad (14)$$
$$att ::= attest_i(eq) \quad (15)$$
$$eq ::= trelt' \quad (16)$$
$$rel ::= = \mid < \mid > \mid \leq \mid \geq \quad (17)$$
$$t ::= c \mid x \mid F(t_1, \ldots, t_n) \quad (18)$$

Η γλώσσα περιλαμβάνει μόνο δύο βοηθητικούς τελεστές: τον συνήθη τελεστή γνώσης K_i και τον τελεστή αλγοριθμικής γνώσης X_i, ο οποίος αναπαριστά τη γνώση,με την οποία ένας πράκτορας iμπορεί ουσιαστικά να δομήσει καταστάσεις. Η ομάδα τελεστών φ_0 περιλαμβάνει και ένα σύνολο από βασικούς όρους των μεταβλητών του συστήματος και τελεστές που χρησιμοποιούνται, για να χαρακτηρίσουν την ίδια την αρχιτεκτονική που θέλουμε να δομήσουμε. Πιο συγκεκριμένα, το $receive_{i,j}(x)$ σημαίνει ότι ο πράκτορας j μπορεί να στείλει τη μεταβλητή x στον πράκτορα i, το $receive_{i,j}(prim)$ σημαίνει ότι ο πράκτορας j μπορεί να στείλει τον όρο $prim$ στον πράκτορα i, όπου ο $prim$ μπορεί να είναι ένας συνδυασμός (μη διαδραστικής) μηδενικών αποδείξεων γνώσης και διαβεβαιώσεων. Μία διαβεβαίωση $attest_i(eq)$ είναι μία απλή δήλωση του πράκτορα i ότι η ιδιότητα eq ισχύει. Η δήλωση αυτή δεν είναι χρήσιμη στον πράκτορα j εκτός αν ο πράκτορας i τον εμπιστεύεται, αλήθεια η οποία εκφράζεται με το $trust_{i,j}$. Η πρωταρχική απόδειξη $proof_{i,j}(p)$ σημαίνει ότι ο πράκτορας i μπορεί να αποδείξει τονόρο p, γεγονός το οποίο μπορεί να ελεγχθεί από τον πράκτορα j. Οι τελεστές $compute_i(x = t)$και $check_i(p)$χρησιμοποιούνται για να εκφράσουν το γεγονός ότι ένας πράκτορας i δύναται να υπολογίσει μία μεταβλητή x, η οποία καθορίζεται από την ισότητα $x = t$ ή να ελέγξει ένα όρο p αντίστοιχα. Το σύμβολο F συμβολίζει τις διαθέσιμες βασικές λειτουργίες όπως (κατατεμαχισμός, ομοιόμορφος κατατεμαχισμός, κρυπτογράφηση, παραγωγή τυχαίας μεταβλητής, αριθμητικές πράξεις, κ.τ.λ.), το c συμβολίζει σταθερές και το x μεταβλητές. Τέλος το $has_i(x)$ είναι ένας τελεστής που συμβολίζει, ότι ο πράκτορας i δύναται να λάβει τη μεταβλητή x (η οποία από μόνη της, δεν μπορεί να πιστοποιηθεί από κάποιον ότι είναι σωστή).

§3.4 Σημασιολογία και Συστήματα Αξιωμάτων

Η σημασιολογία αυτής της λογικής μπορεί να καθορισθεί με τη χρήση ενός ενισχυμένου μοντέλου Kripke ***[8]**(**Παράρτημα «Α»**). $M = (Arch, \pi, D_1, \ldots\ldots, D_n)$, όπου:

- Το Arch είναι ένα σύνολο από πιθανές δομές (που γενικά αποκαλούνται κόσμοι) του υπό εξέταση συστήματος. Στα δικά μας δεδομένα, μία δομή μπορεί να καθορισθεί ως ένα σύνολο τελεστών φ_0, το οποίο χαρακτηρίζει όλες τις διαθέσιμες στους πράκτορες λειτουργίες (ενέργειες).

- Το π είναι μία ερμηνεία όλων των πρωταρχικών όρων $prim$και όλων των συσχετίσεων eq.

- Τα $D_1, \ldots\ldots, D_n$ είναι τα επαγωγικά συστήματα που σχετίζονται με τους πράκτορες $1, \ldots\ldots, n$.

Το επαγωγικό σύστημα που σχετίζεται με έναν πράκτορα i, καθιστά δυνατό το να προσδιορίσουμε τη σημασιολογία του τελεστή X_i (η γνώση που δημιουργήθηκε από τον πράκτορα i). Με άλλα λόγια, κάθε πράκτορας i μπορεί να εφαρμόσει τους κανόνες $\rhd_i$, προκειμένου να εξάγει νέα γνώση. Λογικοί κανόνες του επαγωγικού συστήματος είναι:

- $receive_{i,j}(prim) \rhd_i prim$
- $attest_i(eq), trust_{i,j} \rhd_i eq$
- $proof_{i,j}(p) \rhd_i p$
- $check_i(eq) \rhd_i p$
- $compute_i(x = t) \rhd_i x = t$
- $hash(x_1) = hash(x_2) \rhd_i x_1 = x_2$
- $receive_{i,j}(x) \rhd_i has_i(x)$
- $compute_i(x = t) \rhd_i has_i(x)$

- $has_i(x_1)$,….., $has_i(x_m)$, $dep(x,\{x_1,\dots,x_m\}) \rhd_i has_i(x)$

Με το $dep(x,\{x_1 \dots . x_m\})$ να είναι μία σχέση εξάρτησης γνωστή στον iκαι η οποία ξεκινά με xκαι δύναται να εξαχθεί από τα $x_1,\dots.,x_m$ (είδος κρυπτογράφησης – συλλογισμού).

Προκειμένου να καταστήσουμε λειτουργικές τις αρχιτεκτονικές, μπορούμε να χρησιμοποιήσουμε ένα σύστημα αξιωμάτων, όπως έχουμε κάνει παραπάνω στην εργασία με βάση την επαγωγική αλγοριθμική γνώση **[R. Pucella 2006] [37]**. Τυπικά παραδείγματα αξιωμάτων και συμπερασματικών κανόνων περιλαμβάνουν:

Taut: Όλες οι περιπτώσεις προτασιακών ταυτολογιών. (19)

MP: Από τον φ και το $\varphi \Rightarrow \psi$ συμπεραίνουμε τον ψ (20)

Gen: Από τονφ συμπεραίνουμε το $K_i\varphi$ (21)

K: $K_i(\varphi \Rightarrow \psi) \Rightarrow (K_i(\varphi) \Rightarrow K_i(\psi))$ (22)

T: $K_i(\varphi) \Rightarrow \varphi$ (23)

KC: $K_i(\varphi \wedge \psi) \Rightarrow (K_i(\varphi) \wedge K_i(\psi))$ (24)

XD: Από τα $X_i(\varphi_1),\dots.,\ X_i(\varphi_n)$και $\varphi_1,\dots\dots,\varphi_n \rhd_i \varphi$ συμπεραίνουμε $X_i(\varphi)$ (25)

XT: $X_i(\varphi) \Rightarrow \varphi$ (26)

XC: $X_i(\varphi \wedge \psi) \Rightarrow (X_i(\varphi) \wedge X_i(\psi))$ (27)

Μία ειδοποιός διαφορά μεταξύ των τελεστών K_i και X_i είναι η αδυναμία ισχύος των αξιωμάτων Gen και K για τον X_i, γεγονός το οποίο καθιστά δυνατό για μας να αποφύγουμε το πρόβλημα της λογικής παντογνωσίας (όπως έχει αναφερθεί και παραπάνω στην ίδια εργασία). Αντίθετα, η γνώση που δημιουργείται από τον πράκτορα i εξαρτάται το σχετικό

επαγωγικό σύστημα $\rhd_i$, όπως εκφράζεται και από τον κανόνα XD. Οι κανόνες T και XT εκφράζουν το γεγονός ότι ένας πράκτορας δεν δύναται να εξάγει λανθασμένες ιδιότητες.

Το σύστημα αξιωμάτων που παρατέθηκε σε αυτή την παράγραφο, αποτελεί τη βάση για ένα συμπερασματικό αλγόριθμο και καθιστά δυνατή την απόδειξη των ιδιοτήτων των αρχιτεκτονικών (εκφρασμένες εν προκειμένω ως τελεστές φ_0), όπως θα συζητηθεί παρακάτω.

§3.5 Η Χρησιμότητα και η Χρήση του Μοντέλου Τυπικής Λογικής

Πρωταρχικός σκοπός ενός δομημένου περιβάλλοντος είναι να καταστήσει δυνατό στο σχεδιαστή, να εκφράζει τις απαιτήσεις που εφαρμόζονται στο σύστημα και προαιρετικά, να ορίζει τις επιλογές σχεδίασης που έχουν γίνει (οι οποίες επιβάλλονται έμμεσα από το περιβάλλον ή τον πελάτη). Μιλώντας συγκεκριμένα οι απαιτήσεις αποτελούνται από τρία μέρη:

- Τις λειτουργικές απαιτήσεις: ο σκοπός του συστήματος, ο οποίος εκφράζεται ως ένα σύνολο ισοτήτων $x = t$.
- Γνωσιακές απαιτήσεις και απαιτήσεις ασφαλείας: τιμές που δεν θα έπρεπε (ή θα έπρεπε, ανάλογα την περίπτωση) να είναι γνωστές σε ορισμένους εμπλεκόμενους, γεγονός το οποίο εκφράζεται με τους τελεστές $has_i(x)$ και $\neg has_i(x)$ αντίστοιχα.
- Απαιτήσεις ορθότητας (συνοχής): η πιθανότητα συγκεκριμένοι εμπλεκόμενοι να επιβεβαιώνουν, ότι συγκεκριμένες τιμές είναι ορθές, το οποίο συμβολίζεται με τον τελεστή $X_i(eq)$.

Γενικά, άλλες απαιτήσεις, που δεν έχουν να κάνουν με τη λειτουργικότητα της λογικής, θα μπορούσαν επίσης να ληφθούν υπόψη. Ωστόσο ο συνδυασμός ασφάλειας και

ορθότητας προσδίδει ήδη ένα βαθμό πολυπλοκότητας, η οποία επαρκεί για την παρουσίαση της υπό εξέτασης προσέγγισης.

Οι επιλογές σχεδίασης που προσθέτουν επιπλέον απαιτήσεις στην αρχιτεκτονική μας, μπορούν να προσδιορισθούν ως τελεστές φ_0. Μπορούν να εκφράζουν για παράδειγμα, το γεγονός ότι οι σύνδεσμοι επικοινωνίας είναι (ή δεν είναι) διαθέσιμοι μεταξύ συγκεκριμένων μερών του συστήματος ή ότι συγκεκριμένοι υπολογισμοί (γνώση μηδενικών αποδείξεων, ομοιόμορφη διαίρεση κ.τ.λ.) μπορούν (ή όχι) να αποδοθούν από συγκεκριμένα μέρη του συστήματος. Αρκετές καταστάσεις είναι δυνατές, όταν ο σχεδιαστής εισάγει αυτό το πρώτο πακέτο πληροφοριών:

1. Οι απαιτήσεις ίσως είναι αντικρουόμενες, για παράδειγμα επειδή οι απαιτήσεις ασφαλείας συγκρούονται με τη γνώση ή με δομικές απαιτήσεις ή επειδή οι ίδιες οι δομικές απαιτήσεις, δεν είναι συνεπείς (πράξεις υπολογισμένες από συστατικά, τα οποία δεν έχουν πρόσβαση στις απαραίτητες παραμέτρους ή έλεγχοι που δεν μπορούν να διεκπεραιωθούν, διότι ορισμένες τιμές δεν είναι διαθέσιμες σε μέρη του συστήματος). Το σύστημα επιστρέφει τις εντοπισθείσες αντιφάσεις, το οποίο παρέχει στον σχεδιαστή ορισμένες κατευθύνσεις προς τροποποίηση των δεδομένων που έχει αρχικά εισάγει.

2. Οι απαιτήσεις μπορεί να είναι συνεπείς, αλλά όχι ακριβείς αρκετά για να αποδώσουν μία ολόκληρη αρχιτεκτονική. Στην περίπτωση αυτή, το σύστημα μπορεί να χρησιμοποιήσει μία βιβλιοθήκη από τεχνολογίες ενίσχυσης ασφαλείας (PETs), προκειμένου να τροφοδοτήσει με επιλογές τον χρήση. Ο χρήστης μπορεί να αποφασίσει να εφαρμόσει μία τέτοια τεχνολογία, η οποία εκφράζεται επίσημα με την προσθήκη μιας νέας υπόθεσης. Για παράδειγμα, το $receive_{i,j}(proof_{i,j}(p))$ για μία μηδενικής γνώσης απόδειξη του όρου p, ο οποίος αποστέλλεται από τον πράκτορα jστον πράκτορα i.

3. Οι απαιτήσεις μπορεί να είναι αρκετά ακριβείς, ώστε να συγκεκριμενοποιούν μία μοναδική και ορθή δομή.

Οι πρώτες δύο περιπτώσεις οδηγούν σε νέα επανάληψη της διαδικασίας. Στην τελευταία περίπτωση ο σχεδιαστής έχει καταφέρει μία ικανοποιητικής διάταξης αρχιτεκτονική (η οποία δηλαδή δεν τον εμποδίζει να εκτελέσει μία νέα επανάληψη με διαφορετικές υποθέσεις προς περεταίρω εξερεύνηση της δομής σχεδίασης).

§3.6 Ένα Παράδειγμα Εφαρμογής για Ενίσχυση της Ασφάλειας του Πρωτοκόλλου Επικοινωνίας μεταξύ δύο Συστημάτων

Θα παραθέσουμε εδώ το πώς εφαρμόζεται το πλαίσιο που περιγράφεται στην προηγούμενη παράγραφο, σε μία μικρή περίπτωση, συστημάτων έξυπνων μετρήσεων. Τα έξυπνα δίκτυα που παρέχουν ασφάλεια πληροφοριών και τα συστήματα έξυπνων μετρήσεων έχουν αναλυθεί εκτενώς στο παρελθόν στη σχετική βιβλιογραφία **[M. Jawurek, M. Johns, F. Kerschbaum 2011] [38]**. Σκοπός μας εδώ δεν είναι ούτε να παρουσιάσουμε μία νέα λύση, ούτε να παρέχουμε μία μελέτη κατανόησης των συστημάτων αυτών, αλλά να δείξουμε το πώς η ύπαρξη ενός τυποποιημένου πλαισίου μπορεί να βοηθήσει έναν σχεδιαστή να βρει μία επιθυμητή λύση μεταξύ των διατιθέμενων επιλογών. Εστιάζουμε εδώ στον έλεγχο λειτουργικότητας λογαριασμών και σε πιθανά προβλήματα που μπορεί να δημιουργηθούν μεταξύ απαιτήσεων ασφαλείας των πελατών και ενός υπαλλήλου ελέγχου πληρωμών (ενός πράκτορα δηλαδή), στο κομμάτι της επιβεβαίωσης της ορθότητας ενός ποσού προς πληρωμή.

Ας υποθέσουμε στην πρώτη περίπτωση, ότι το σύστημα αποτελείται από τρία μέρη: το κεντρικό σύστημα του υπαλλήλου o, τον υπολογιστή του χρήστη (πελάτης ή δεύτερος πράκτορας) u και ο μετρητής m. Οι απαιτήσεις για την περίπτωση αυτή είναι οι ακόλουθες:

- Λειτουργικές απαιτήσεις: Σκοπός του συστήματος είναι ο υπολογισμός του ποσού πληρωμής, το οποίο εκφράζεται από την ισότητα $Fee = \sum_{i=1}^{n}(P(C_i))$, όπου το C_i είναι η ακριβής κατανάλωση χρηματικών ποσών στην υπό εξέταση χρονική περίοδο $(i \in [1, n])$και το P είναι το κόστος λειτουργίας.

- Απαιτήσεις ασφαλείας και γνώσεων: τα $\neg has_o(C_i)$ και $has_o(Fee)$αντίστοιχα, εκφράζουν το γεγονός ότι ο υπάλληλος o δεν πρέπει να μπορεί να δει τις τιμές C_iτων χρηματικών ποσών που καταναλώθηκαν, αλλά να μπορεί να δει το ποσό πληρωμής.

- Απαιτήσεις ορθότητας (συνοχής): ο υπάλληλος πρέπει να είναι σίγουρος ότι το ποσό πληρωμής είναι ορθό.

Ας υποθέσουμε ότι στο πρώτο σενάριο, ο σχεδιαστής εξετάζει μία απευθείας σύνδεση μεταξύ του μετρητή και του υπαλλήλου, το οποίο σημαίνει ότι η δομή θα περιλάμβανε τον τελεστή $receive_{o,m}(C_i)$. Αυτή η πιθανότητα θα ερχόταν προφανώς σε σύγκρουση με τις απαιτήσεις ασφαλείας αφού ισχύει: $receive_{o,m}(C_i) \rhd_i has_o(C_i)$. Για τον λόγο αυτό, δύο σύνδεσμοι επικοινωνίας είναι απαραίτητοι: από τον m στον u και από τον u στον o.

Η επόμενη ερώτηση είναι το που θα πρέπει να λάβει χώρα ο υπολογισμός της P: γενικά μιλώντας, αυτό θα μπορούσε να πραγματοποιηθεί στη γραμματεία του υπαλλήλου, στον μετρητή ή στον υπολογιστή του χρήστη. Ανάλογα με τις αρχικές ιδέες της δομής και των περιορισμών που επιβάλλουν οι υπολογιστές, ο σχεδιαστής μπορεί, είτε να εισάγει απευθείας τον κατάλληλο τελεστή $compute_o$, $compute_m$ ή $compute_u$. Διαφορετικά αυτές οι επιλογές θα μπορούσαν να προτείνονται εκ περιτροπής από το σύστημα.

- Η πρώτη επιλογή έρχεται σε αντίθεση με τις απαιτήσεις ασφαλείας προσωπικών δεδομένων (διότι ο υπάλληλος θα χρειαζόταν τις τιμές εισόδου C_i, προκειμένου να υπολογίσει το Fee), εκτός αν ομοιόμορφη κρυπτογράφηση χρησιμοποιηθεί ώστε να επιτρέψει στο υπάλληλο τον υπολογισμό του Fee με κρυπτογραφημένες τιμές C_i.

- Η δεύτερη επιλογή μπορεί να διερευνηθεί περεταίρω, εφόσον δεν δημιουργεί επιπλέον, απαγορευτικά για τη διαδικασία μέτρησης κόστη. Ωστόσο το σύστημα θα εντόπιζε

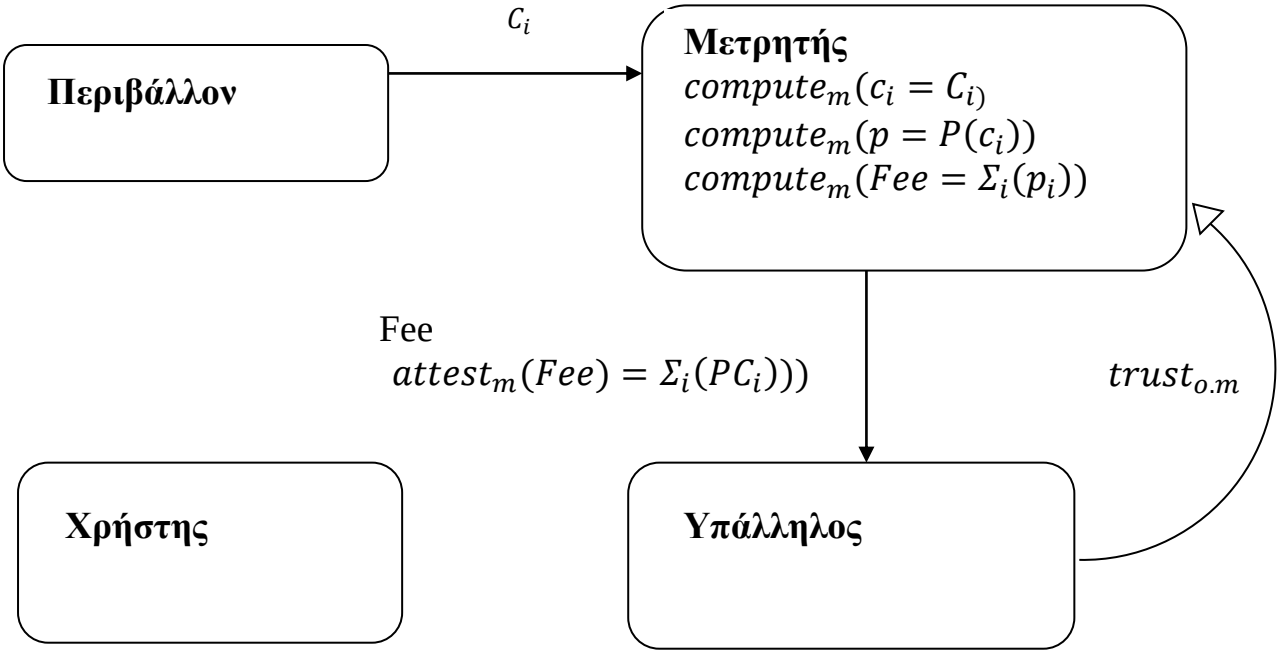
σε αυτή την περίπτωση μία απαίτηση πιστοποίησης (η οποία θα μπορούσε ή όχι να προβλεφθεί από τον σχεδιαστή): ο υπάλληλος πρέπει να θεωρεί έμπιστη τη μέτρηση του μετρητή ($trust_{o,m}$), διότι η μόνη πληροφορία που λαμβάνεται από τον υπάλληλο θα ήταν μία πιστοποίηση σε σχέση με τον μετρητή ($attest_m(Fee = \sum_{i=1}^{n}(P(C_i)))$) και ο μόνος τρόπος επαλήθευσης της πιστοποίησης αυτής, είναι μέσω μίας πιστοποίησης της υπόθεσης (η οποία εκφράζεται από τον επαγωγικό κανόνα $attest_j(eq), trust_{i,j} \rhd_i eq$).

- Η τρίτη επιλογή είναι πιο δελεαστική, αν ο ρόλος του μετρητή περιοριστεί στην διαβίβαση των μετρήσεων καταναλωθέντων ποσών. Όμως αυτή η επιλογή οδηγεί σε μία άλλη απαίτηση: είτε ο υπάλληλος μπορεί να εμπιστευτεί τον χρήστη, (γεγονός το οποίο αποκλείεται από την υπόθεση), είτε ο χρήστης πρέπει να παρέχει στον υπάλληλο απόδειξη της ορθότητας του υπολογισμού του ποσού πληρωμής (το οποίο δίδεται από το $receive_{o,u}(proof_{u,o}(Fee = \sum_{i=1}^{n}(P(C_i))))$).

Αν καμία από αυτές τις επιλογές δεν είναι διαθέσιμη, επιπλέον εμπλεκόμενοι πρέπει να προστεθούν στον υπολογισμό του ποσού πληρωμής. Γενικά, αυτοί οι εμπλεκόμενοι μπορούν να είναι είτε ζεύγη, είτε έμπιστα τρίτα σύνολα ατόμων. Και στις δύο περιπτώσεις, επιπλέον απαιτήσεις θα προέκυπταν: ένας ασφαλές πολλών συνόλων ατόμων σχήμα υπολογισμού θα ήταν απαραίτητο, προκειμένου να διασφαλίσει, ότι τα ζεύγη δεν μαθαίνουν τα καταναλωθέντα ποσά το ένα από το άλλο και πιστοποιήσεις υποθέσεων, προκειμένου να διασφαλίσουν ότι οι υπολογισμοί δεν θα διαδοθούν σε τρίτο σύνολο ατόμων.

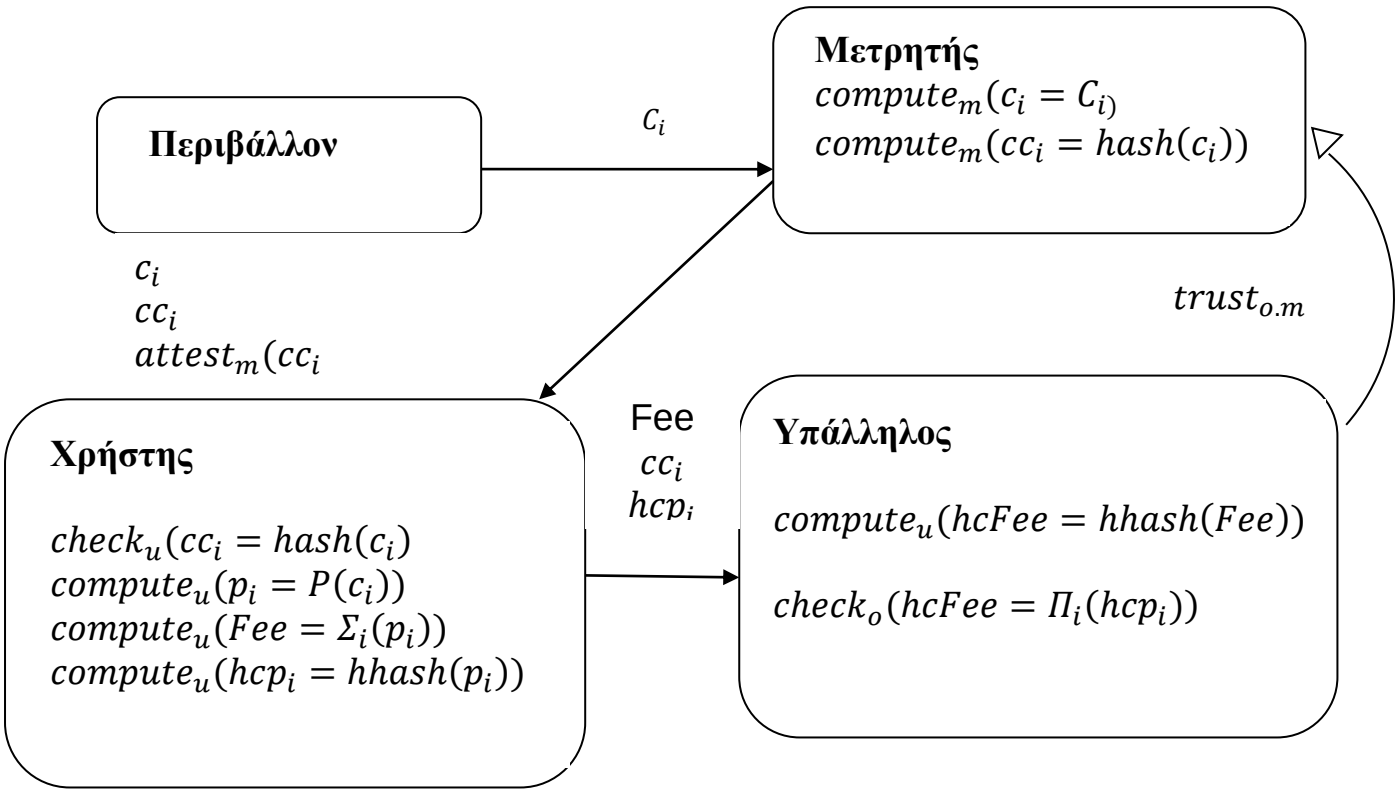
Όπως έχει αναφερθεί και στη δεύτερη παράγραφο, οι σχεδιαστές δεν είναι συνήθως ειδήμονες στη χρήση τέτοιων μεθόδων. Ως αποτέλεσμα, το περιβάλλον σχεδίασης θα έπρεπε να παρέχει εναλλακτικές μορφές διάδρασης, οι οποίες θα απλοποιούν το χρησιμοποιούμενο μοντέλο. Λαμβάνοντας υπόψη ότι οι σχεδιαστές είναι εξοικειωμένοι με το γραφικό περιβάλλον, επιλέξαμε να αναπαραστήσουμε με γραφήματα τις διάφορες περιγραφείσες δομές του συστήματός μας. Αυτές παρατίθενται στα παρακάτω γραφήματα.

Γράφημα 1



Στο παραπάνω γράφημα φαίνεται, ότι ο μετρητής υπολογίζει το ποσό πληρωμής και ο υπάλληλος πρέπει να τον εμπιστεύεται, προκειμένου να το λάβει από αυτόν. Το C_i το ακριβές καταναλωθέν ποσό και τα c_i είναι οι τιμές που χρησιμοποιήθηκαν από τον μετρητή για να υπολογισθεί το συνολικό ποσό πληρωμής.

Γράφημα 2



$attest_m(cc_i = hash(C_i))$
$proof_{u,o}(cc_i = hash(C_i))$
$compute_u(hcp_i = hhash(p_i))$
$proof_{u,o}(p_i = P(c_i))$

Εδώ ο χρήστης υπολογίζει το ποσό πληρωμής και ο υπάλληλος πρέπει να εμπιστεύεται τον μετρητή (για τη χρήση της σωστής τιμής c_i). Το $hhash$ είναι μία ομοιόμορφη συνάρτηση κατατεμαχισμού (hash), η οποία επιτρέπει στον υπάλληλο να ελέγχει, ότι ο κατατεμαχισμός ενός παγκόσμιου ποσού πληρωμής είναι συνεπής με τους κατατεμαχισμούς hcp_i ατομικών ποσών πληρωμής p_i. Η δομή αυτή είναι μία σύνοψη της προτεινόμενης λύσης του προβλήματος της εφαρμογής μας.

ΚΕΦΑΛΑΙΟ 4ο

§4. Επίλογος

Περιγράψαμε μία προσέγγιση, συνδυασμού επαγωγικής γνώσης μεταφρασμένης με βάση τη θεωρία των πιθανών κόσμων, και παραγωγικής γνώσης, η οποία βασίζεται σε ένα επαγωγικό σύστημα, το οποίο επιτρέπει στους πράκτορες να εξάγουν πληροφορίες που γνωρίζουν σαφώς. Αυτή η επιπρόσθετη δομή, το επαγωγικό σύστημα του πράκτορα, μπορεί να χρησιμοποιηθεί για να χαρακτηρίσει πλήρως τις ιδιότητες μιας συνάρτησης άμεσης γνώσης. Πιο συγκεκριμένα, μπορούμε να εξάγουμε έγκυρα και πλήρη συστήματα αξιωμάτων για την χρησιμοποιούμενη λογική με κανονικό τρόπο. Υπάρχουν πολυάριθμες προσεγγίσεις για τη μοντελοποίηση της παραγωγικής γνώσης στη σχετική βιβλιογραφία, με πολλές διαφορετικές φιλοσοφικές θεωρίες και ερμηνείες. Το μοντέλο που παρουσιάζουμε σε αυτή την εργασία, βασίζεται στην έννοια της παραγωγικής γνώσης (η οποία προκύπτει με συμπερασματικούς κανόνες), χρονολογείται τουλάχιστον από τη θεωρία που διατύπωσε ο **[Konolige 1986] [39]** και στοχεύει στην απόδοση μίας συγκεκριμένης υπολογιστικής ερμηνείας της παραγωγικής γνώσης, με βάση γενόμενες παρατηρήσεις. Σημειώνουμε ότι το μοντέλο που παρουσιάσαμε, είναι συνεπές με μία σειρά από γνωσιακές θεωρίες **[Pollock and Cruz 1999] [40],** οι οποίες υποστηρίζουν ότι το σύνολο της γνώσης, εξάγεται ουσιαστικά από παρατηρήσεις.

Ενώ το υπό εξέταση πλαίσιο επεκτείνεται σε περιπτώσεις πολλαπλών πρακτόρων αναλογικά, υπάρχουν πολλά θέματα τα οποία παραμένουν προς επίλυση σε αυτό το επίπεδο. Για παράδειγμα, μία εύλογη ερώτηση είναι το τι συμβαίνει, όταν μεταπέσουμε σε πιο δυναμικά μοντέλα, στα οποία οι παρατηρήσεις πραγματοποιούνται με την πάροδο του χρόνου. Εγείρονται λοιπόν ενδιαφέροντα ερωτήματα, ιδιαίτερα όταν υποθέτουμε ότι οι

πράκτορες δεν έχουν κάποιο μηχανισμό προς συγχρονισμό της λήψης των παρατηρήσεών τους. Το πεδίο αυτό είναι προς μελλοντική διερεύνηση.

Το πλαίσιο που αναλύσαμε είναι χρήσιμο, επειδή η επαγωγική αλγοριθμική γνώση είναι μία ειδική περίπτωση της αλγοριθμικής γνώσης. Έτσι κάθε κατάσταση, η οποία μπορεί να μοντελοποιηθεί στο δικό μας πλαίσιο, μπορεί να μοντελοποιηθεί επίσης και στο ευρύτερο πλαίσιο της αλγοριθμικής γνώσης. Το πλεονέκτημα όμως της λογικής μας, έγκειται στο γεγονός ότι εξάγει έγκυρα και πλήρη συστήματα αξιωμάτων απευθείας από τα επαγωγικά συστήματα που χρησιμοποιεί. Για τον λόγο αυτό, μπορούμε να επινοήσουμε χρήσιμα αποδεικτικά συστήματα για πολλές και ενδιαφέρουσες κατηγορίες εφαρμογών, προσαρμοσμένων σε διαφορετικά κάθε φορά επαγωγικά συστήματα.

Επιπρόσθετα παρουσιάσαμε μία εφαρμογή, στην οποία η επαγωγική αλγοριθμική γνώση υπεισέρχεται, καθιστώντας δυνατή την ενίσχυση του πρωτοκόλλου ασφαλείας ενός συστήματος διαχείρισης τραπεζικών λογαριασμών. Με τον ίδιο τρόπο η εφαρμογή αυτή, μπορεί να επεκταθεί και για ενίσχυση της ασφάλειας πρωτοκόλλων επικοινωνίας μεταξύ συστημάτων επιχειρήσεων του δημοσίου και του ιδιωτικού τομέα.

Το πλαίσιο ανάλυσης σε αυτή την εργασία, ρίχνει φως στο γνωσιακό περιεχόμενο των επαγωγικών συστημάτων, με το να παρέχει μία λογική μέσω της οποίας να μπορούμε να αιτιολογήσουμε την επαγωγική και την παραγωγική γνώση, οι οποίες προκύπτουν από ένα επαγωγικό σύστημα. Ένα ενδιαφέρον ερώτημα προς ανάλυση, είναι το κατά πόσο μία τέτοια προσέγγιση δύναται να ερμηνεύσει το γνωσιακό περιεχόμενο των πιθανοτικών επαγωγικών συστημάτων. Αυτό είναι ένα ερώτημα προς μελλοντική ανάλυση.

ΠΑΡΑΡΤΗΜΑ «Α»

Αστερίσκοι

*[1] Στην υπολογιστική θεωρία ένα πρόβλημα λέμε ότι είναι NP-complete, όταν η λύση του δύναται να διεκπεραιωθεί(σε υπολογιστή) σε πολυωνυμικό χρόνο. Στην απλή διατύπωση του το ερώτημα που θέτει είναι, εάν κάθε πρόβλημα του οποίου η ύπαρξη λύσης μπορεί να ελεγχθεί γρήγορα από έναν υπολογιστή, μπορεί επίσης και να επιλυθεί γρήγορα από αυτόν. Πιο συγκεκριμένα, κάθε εισερχόμενο στοιχείο στο πρόβλημα θα πρέπει να σχετίζεται με ένα σύνολο λύσεων πολυωνυμικού μήκους, των οποίων η ορθότητα δύναται να ελεγχθεί γρήγορα (σε πολυωνυμικό χρόνο), έτσι ώστε το εξερχόμενο στοιχείο για κάθε εισερχόμενο να είναι «ναι» αν υπάρχει έστω μία λύση και «όχι» αν δεν υπάρχει καμία. Εναλλακτικά στο εξής θα χρησιμοποιούμε τον όρο «αποφασιστικό» για το πρόβλημα που είναι NP – complete.

*[2] Ενώ σε αυτή την εργασία εστιάζουμε στη γνώση, τα περισσότερα από αυτά που συμπεραίνουμε, δύνανται να εφαρμοσθούν εξίσου και στην περίπτωση της πεποίθησης (belief). Στην πραγματικότητα, εφόσον δεν υποθέτουμε ότι οι γνωσιακοί αλγόριθμοι απαραίτητα δίδουν ορθές απαντήσεις, κάποιος θα μπορούσε να ισχυρισθεί ότι το είδος της γνώσης που προκύπτει από τους γνωσιακούς αλγόριθμους είναι μία πεποίθηση. Για λόγους συνέπειας με την βιβλιογραφία των γνωσιακών αλγορίθμων ωστόσο, θα συνεχίσουμε να χρησιμοποιούμε την ορολογία «αλγοριθμική γνώση».

*[3] Η χρήση μεταβλητών που γίνεται στα επαγωγικά συστήματα είναι καθαρά για λόγους ευκολίας. Θα μπορούσαμε να αντικαταστήσουμε κάθε επαγωγικό κανόνα με κανόνες, οι οποίοι προκύπτουν από υποκατάστατους βασικούς όρους για όλες τις μεταβλητές και να λάβουμε ένα επαγωγικό σύστημα, το οποίο θα εξάγει τους ίδιους όρους όπως το αυθεντικό.

*4 Για λόγους απλότητας, υποθέτουμε, ότι οι παρατηρήσεις σχηματίζουν ένα σύνολο. Αυτό σημαίνει ότι η επανάληψη των παρατηρήσεων και η σειρά τους εντός του συνόλου δεν μας επηρεάζουν. Θα μπορούσαμε εύκολα να μοντελοποιήσουμε την περίπτωση, όπου οι παρατηρήσεις σχηματίζουν μία αλληλουχία, όμως το πρόβλημα θα γινόταν πιο πολύπλοκο. Επίσης υποθέτουμε, ότι ο αριθμός των παρατηρήσεων είναι πεπερασμένος σε κάθε κατάσταση. Μη πεπερασμένος αριθμός παρατηρήσεων, δεν επηρεάζει κάπως ωστόσο κάτι στην παρούσα παράγραφο. Εικάζουμε ότι τα αποτελέσματα στις ακόλουθες παραγράφους ισχύουν και όταν οι παρατηρήσεις σε μία κατάσταση είναι άπειρες.

*5 Μπορούμε επίσης να επιβάλλουμε περιορισμούς στα επαγωγικά συστήματα και στις υπογραφές. Έχουμε διαχωρίσει στο παρελθόν την έννοια μιας πρωταρχικής υπογραφής, η οποία δεν παρέχει κατασκευαστές για προτασιακούς και βοηθητικούς συνδέσμους. Ένα επαγωγικό σύστημα βασισμένο σε μία πρωταρχική υπογραφή, επιτρέπει μόνο την απόδοση άμεσης γνώσης μέσω πρωταρχικών προτάσεων.

*6 Ο δημιουργός πρέπει να είναι προσεκτικός κατά τον καθορισμό επίσημων συστημάτων αξιωμάτων. Είναι επόμενο, ότι ένα σύστημα αξιωμάτων όπως το παραπάνω με μετά-μεταβλητές, οι οποίες εμφανίζονται εντός όρων, είναι στην ουσία ένα σύνολο αξιωμάτων, όπου κάθε προκείμενη σε ένα αξίωμα είναι περίπτωση βασικού υποκατάστατου κατάλληλου όρου στο σύστημα.

*7 Μία μηχανή Τούρινγκ είναι μια Λογική Υπολογιστική Μηχανή, η οποία αποτελείται από απεριόριστη χωρητικότητα μνήμης, σε μορφή μιας άπειρης ταινίας η οποία είναι χωρισμένη σε τετράγωνα, πάνω στο καθένα από τα οποία, μπορεί να εκτυπωθεί ένα σύμβολο. Η μηχανή Τούρινγκ δεν προορίζεται σαν μια τεχνολογία υπολογιστών, αλλά κυρίως σαν μια υποθετική κατασκευή που αντιπροσωπεύει μια υπολογιστική μηχανή. Οι μηχανές Τούρινγκ βοηθούν τους επιστήμονες να καταλάβουν τα όρια του μηχανικού υπολογισμού.

*[8] Το μοντέλο σημασιολογίας Kripke (επίσης γνωστό και ως σημασιολογία συσχετίσεων ή πλαίσιο σημασιολογίας, συχνά μπερδεύεται με το σημασιολογικό πλαίσιο των πιθανών κόσμων) είναι ένα σημασιολογικό πλαίσιο απόδοσης μαθηματικής λογικής για λογικά συστήματα, το οποίο δημιουργήθηκε το 1950-60 από τον Saul Kripke και τον André Joyal. Αρχικά χρησιμοποιήθηκε για απόδοση βοηθητικών λογικών και αργότερα προσαρμόστηκε στη διαισθητική λογική και σε πληθώρα άλλων συστημάτων. Η εξέλιξη του σημασιολογικού πλαισίου του Kripke ήταν ένα ρηξικέλευθο επίτευγμα για τη θεωρία των μη συμβατικών λογικών, καθότι το θεωρητικό μοντέλο μιας τέτοιας λογικής ήταν ουσιαστικά μέχρι τότε ανύπαρκτο (αλγεβρική σημασιολογία υπήρχε, αλλά θεωρούνταν ως μία «μεταμφιεσμένη σύνταξη»).

ΒΙΒΛΙΟΓΡΑΦΙΑ

Επιστημονικές Αναφορές

[1] Halpern, J. Y., Y. Moses, and M. Y. Vardi (1994). Algorithmic knowledge. In Proc. 5th Conference on Theoretical Aspects of Reasoning about Knowledge (TARK'94), pp. 255–266. Morgan Kaufmann.

[2] Rosenschein, S. J. (1985). Formal theories of knowledge in AI and robotics. New Generation Computing 3(4), 345–357.

[3] Levesque, H. J. (1984). A logic of implicit and explicit belief. In Proc. 4th National Conference on Artificial Intelligence (AAAI'84), pp. 198–202.

[4] Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi (1995). Reasoning about Knowledge. MIT Press.

[5] Konolige, K. (1986). A Deduction Model of Belief. Morgan Kaufmann, Levesque, H. J. (1984). A logic of implicit and explicit belief. In Proc. 4th National Conference on Artificial Intelligence (AAAI'84), pp. 198–202.

[6] Selman, B. and H. Kautz (1996). Knowledge compilation and theory approximation. Journal of the ACM 43(2), 193–224.

[7] Konolige, K. (1986). A Deduction Model of Belief. Morgan Kaufmann.

[8] Halpern, J. Y. and R. Pucella (2002). Modeling adversaries in a logic for reasoning about security protocols. In Proc. Workshop on Formal Aspects of Security (FASec'02), Volume 2629 of Lecture Notes in Computer Science, pp. 115–132.

[9] Dolev, D. and A. C. Yao (1983). On the security of public key protocols. IEEE Transactions on Information Theory 29(2), 198–208.

[10] Giunchiglia, F., L. Serafini, E. Giunchiglia, and M. Frixione (1993). Non-omniscient belief as context-based reasoning. In Proc. 13th International Joint Conference on Artificial Intelligence (IJCAI'93), pp. 548–554.

[11] Ladner, R. E. (1977). The computational complexity of provability in systems of modal propositional logic. SIAM Journal on Computing 6(3), 467–480.

[12] Halpern, J. Y. and Y. Moses (1992). A guide to completeness and complexity for modal logics of knowledge and belief. Artificial Intelligence 54, 319–379.

[13] Higgins, P. J. (1963). Algebras with a scheme of operators. MathematischeNachrichten 27, 115– 132.

[14] Konolige, K. (1986). A Deduction Model of Belief. Morgan Kaufmann.

[15] Hintikka, J. (1962). Knowledge and Belief. Cornell University Press.

[16] Hintikka, J. (1962). Knowledge and Belief. Cornell University Press.

[17] McAllester, D. (1993). Automatic recognition of tractability in inference relations. Journal of the ACM 40(2), 284–303.

[18] Konolige, K. (1986). A Deduction Model of Belief. Morgan Kaufmann.

[19] Halpern, J. Y., Y. Moses, and M. Y. Vardi (1994). Algorithmic knowledge. In Proc. 5th Conference on Theoretical Aspects of Reasoning about Knowledge (TARK'94), pp. 255–266. Morgan Kaufmann.

[20] Enderton, H. B. (1972). A Mathematical Introduction to Logic. Academic Press. Fagin, R., J. Y.

[21] McAllester, D. (1993). Automatic recognition of tractability in inference relations. Journal of the ACM 40(2), 284–303.

[22] Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi (1995). Reasoning about Knowledge. MIT Press.

[23] Kaplan, A. N. and L. K. Schubert (2000). A computational model of belief. Artificial Intelligence 120, 119–160.

[24] Alchourro´n, C. E., P. Gardenfors, and D. Makinson (1985). On the logic of theory change: partial meet functions for contraction and revision. Journal of Symbolic Logic 50, 510–530.

[25] Halpern, J. Y. and Y. Moses (1992). A guide to completeness and complexity for modal logics of knowledge and belief. Artificial Intelligence 54, 319–379.

[26] M. Langheinrich. Privacy by design - principles of privacy aware ubiquitous systems. In Proceedings of the Ubicomp Conference, Springer, LNCS 2201, pages 273–291, 2001.

[27] Y. Poullet. About the e-privacy directive, towards a third generation of data protection legislations. In Data Protection in a Profile World, pages 3–29. Springer, 2010.

[28] E.C. European Commission. Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). inofficial consolidated version after LIBE Commitee vote provided by the rapporteur, 22 October 2013.

[29] S. F. Gurses, C. Troncoso, and C. Diaz. Engineering privacy by design. In Computers, Privacy & Data Protection, (2011).

[30] F. Kerschbaum. Privacy-Preserving Computation (Position Paper). Annual Privacy Forum (APF'12), Cyprus, 2012.

[31] Y. Deswarte and C. A. Melchor. Current and future privacy enhancing technologies for the internet. Annals of Telecommunications 61(3), pages 399–417, 2006.

[32] I. Goldberg. Privacy-enhancing technologies for the internet III: ten years later. In Digital Privacy: Theory, Technologies, and Practices, pages 84–89. TeX Users Group, December 2007.

[33] F. Kerschbaum. Privacy-Preserving Computation (Position Paper). Annual Privacy Forum (APF'12), Cyprus, 2012.

[34] L. Bass, P. Clements, and R. Kazman. Software architecture in practice (3d edition). SEI Series in Software Engineering, Addison Wesley, 2013.

[35] D. Le Metayer. Privacy by design: a matter of choice. In Data Protection in a Profiled World, Springer Verlag, pages 323–334, 2010.

[36] J. Y. Halpern, R. Pucella. Dealing with logical omniscience: Expressiveness and pragmatics. Artif. Intell. 175(1), pages 220–235, 2011.

[37] R. Pucella. Deductive Algorithmic Knowledge. Journal of Logic and Computation 16 (2), pages 287–309, 2006.

[38] M. Jawurek, M. Johns, F. Kerschbaum. Plug-In Privacy for Smart Metering Billing. Privacy Enhancing Technologies Symposium (PETS'11), pages 192–210, 2011.

[39] Konolige, K. (1986). A Deduction Model of Belief.

[40] Pollock, J. L. and J. Cruz (1999). Contemporary Theories of Knowledge (Second ed.). Rowman & Littlefield.

[41] https://el.wikipedia.org/wiki

Επιστημονικά Άρθρα

1. Deductive Algorithmic Knowledge, Riccardo Pucella, Northeastern University Boston, MA 02115 USA (riccardo@ccs.neu.edu)

2. Using Deductive Knowledge To Improve Cryptographic Protocol Verification, Zhiwei Li, Weichao Wang, Department of Software and Information Systems University of North Carolina at Charlotte, NC USA (zli19@uncc.edu and weichaowang@uncc.edu)

3. Privacy by Design: From Technologies to Architectures (Position Paper), Thibaud Antignac and Daniel Le Metayer Inria, Universite de Lyon (thibaud.antignac@inria.fr/daniel.le-metayer@inria.fr)

4. Privacy Architectures: Reasoning About Data Minimisation and Integrity, Thibaud Antignac and Daniel Le Métayer Inria, University of Lyon, France (thibaud.antignac, daniel.le-metayer@inria.fr)