



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Τμήμα Επιστήμης και Τεχνολογίας Υπολογιστών

Μεταπτυχιακή Διπλωματική Εργασία

ΕΠΙΣΚΟΠΗΣΗ ΤΗΣ ΘΕΩΡΙΑΣ ΑΝΑΘΕΩΡΗΣΗΣ ΠΕΠΟΙΘΗΣΕΩΝ



ΠΥΡΠΑΡΑ ΣΠΥΡΙΔΟΥΛΑ

A.M. 2009026

pcst0926@uop.gr

Επιβλέπων: Κούτρας Κωνσταντίνος

Επίκουρος Καθηγητής του Τμήματος Επιστήμης και Τεχνολογίας Υπολογιστών
Πανεπιστήμιο Πελοποννήσου

Τρίπολη 2011

ΑΝΑΘΕΩΡΗΣΗ ΠΕΠΟΙΘΗΣΕΩΝ

ΕΚΤΕΤΑΜΕΝΗ ΠΕΡΙΛΗΨΗ

Στα μέσα της δεκαετίας του 1980, εφευρέθηκαν λογικά εργαλεία με τα οποία ήταν δυνατή η μοντελοποίηση της αναθεώρησης πεποιθήσεων και γνώσεων με εντελώς καινούριες διαδικασίες. Αυτά τα λογικά εργαλεία αποδείχτηκαν ότι είναι εφαρμόσιμα και στα ανθρώπινα πιστεύω και στα περιεχόμενα των βάσεων δεδομένων των ηλεκτρονικών υπολογιστών.

Η ανάγκη για ισχυρότερους τρόπους αναπαράστασης πληροφορίας οδήγησε στην ανάπτυξη των λογικών βάσεων γνώσεων, δηλαδή βάσεων στις οποίες η γνώση εκφράζεται με την χρήση λογικών προτάσεων. Τέτοιου είδους βάσεις γνώσης δίνουν την δυνατότητα αποθήκευσης πληροφορίας η οποία δεν είναι ρητώς καταχωρημένη στην βάση. Η παραγόμενη γνώση όμως προκάλεσε δυσκολίες στην ανάπτυξη μεθόδων ενημέρωσης της γνώσης, αφού η νέα πληροφορία μπορεί να αντιφάσκει είτε με την ίδια την βάση ή με την παραγόμενη γνώση της βάσης. Το πρόβλημα της ενημέρωσης μίας βάσης γνώσης με μία νέα, πιθανά αντιφατική, πληροφορία ονομάζεται αναθεώρηση πεποιθήσεων και είναι ένα πολύ σημαντικό πρόβλημα για διάφορες εφαρμογές.

Η παρούσα εργασία αντιμετωπίζει το πρόβλημα αυτό για προτασιακές βάσεις γνώσης. Εισάγεται ένας νέος, ισχυρότερος τρόπος αναπαράστασης εκφράσεων του προτασιακού λογισμού, με την βοήθεια του AGM μοντέλου, το οποίο μας βοηθά να κάνουμε : α) πρόσθεση πεποιθήσεων, β) αναθεώρηση πεποιθήσεων και γ) αφαίρεση πεποιθήσεων χρησιμοποιώντας κάποιες βασικές αρχές.

Η πρόσθεση αναφέρεται στη συλλογή νέων πληροφοριών (επέκταση πεποιθήσεων), η αφαίρεση στην απώλεια πληροφορίας, ενώ η αναθεώρηση ερμηνεύει τη μερική ή ολική αλλαγή στο σύνολο των πεποιθήσεών μας, εξαιτίας της εμφάνισης μίας νέας πεποίθησης.

Κάθε διαδικασία αναθεώρησης συνοδεύεται από ένα σύνολο ορθολογικών αξιωμάτων.

Τέλος, μελετώνται κάποια νέα προβλήματα που ανοίγονται από την εισαγωγή της αναθεώρησης πεποιθήσεων και προτείνονται τρόποι περαιτέρω βελτίωσης της τεχνικής αυτής στην αναπαράσταση και στην ενημέρωση της γνώσης.

Belief Revision

Extended Abstract

In the middle of the 1980's, logical tools were discovered that make it possible to model changes in belief and knowledge in entirely new ways. These logical tools turned out to be applicable both to human beliefs and to the contents of databases of computers.

The need for more powerful knowledge representation techniques led to the development of logical knowledge bases, where knowledge is expressed using logical propositions. Such knowledge bases allow the existence of information not explicitly stored in the base. However, implicit knowledge introduced several difficulties when it comes to updating the knowledge, as the new data may contradict either the knowledge base itself, or the implicit knowledge of the base. The problem of updating a knowledge base in the face of new, possibly contradictory, information is known as belief revision and is a crucial problem for several applications.

This work addresses this problem for propositional knowledge bases. A new and more powerful representation of propositional expressions is introduced, the AGM model, which helps us to: a) expand one's beliefs, b) revise one's beliefs and c) contract one's beliefs, by using some of the most basic principles.

The expansion refers to gathering new information (extension beliefs), removing the loss of information while reviewing and interpreting the partial or total change in all of our beliefs, because of the entrance of a new belief.

Any change process is accompanied by a set of rational axioms.

Finally, some new problems opened by the introduction of belief revision and methods of further improvement of this technique in knowledge representation and updating are proposed.

Πίνακας Περιεχομένων

Κεφάλαιο 1	
1.1 Εισαγωγή	7
1.2 Ποιο είναι το Πρόβλημα	8
1.3 Ιστορική αναδρομή	12
1.4 Η Προσέγγιση των Alchourron, Gärdenfors και Makinson	14
Κεφάλαιο 2	
2.1 Προτασιακή Λογική	16
2.2 Αναπαράσταση Γνώσης	19
2.3 Ιδιότητες των Προτασιακών Βάσεων δεδομένων	23
2.4 Αποδοχή και Απόρριψη Πεποιθήσεων	23
2.5 Γνωσιακά Σύνολα	25
Κεφάλαιο 3	
3.1 Το AGM Παράδειγμα	27
3.2 Πρόσθεση Πεποιθήσεων	28
3.3 Αναθεώρηση Πεποιθήσεων	32
3.4 Αφαίρεση Πεποιθήσεων	38
3.5 Από την Αφαίρεση στην Αναθεώρηση Πεποιθήσεων και Αντίστροφα	43
Κεφάλαιο 4	
4.1 Το Πρόβλημα της Κατασκευής Συναρτήσεων	48
4.2 Συναρτήσεις Αφαίρεσης Μέγιστης Επιλογής	49
4.3 Συναρτήσεις Αφαίρεσης Πλήρους Επιλογής	51
4.4 Συναρτήσεις Αφαίρεσης Μερικής Επιλογής	53
4.5 Γνωσιακή Κατοχύρωση	56
Κεφάλαιο 5	
5.1 Συμπεράσματα	64
5.2 Μελλοντικές Κατευθύνσεις	67
Παράρτημα Αποδείξεων	70
Βιβλιογραφία	73

Κατάλογος Πινάκων

Πίνακας 1. Αξιώματα Πρόσθεσης Πεποιθήσεων	32
Πίνακας 2. Αξιώματα Αναθεώρησης Πεποιθήσεων	33
Πίνακας 3. Αξιώματα Αφαίρεσης Πεποιθήσεων	39
Πίνακας 4. Αξιώματα Γνωσιακής Κατοχύρωσης	61

Κεφάλαιο 1 : Εισαγωγή

1.1 Εισαγωγή

Μία από τις βασικότερες εφαρμογές των υπολογιστικών συστημάτων είναι η αποθήκευση, ανάκτηση, σύνθεση και αξιολόγηση γνώσης και πληροφορίας για τον κόσμο που μας περιβάλλει. Οι τεράστιες αποθηκευτικές ικανότητες των ηλεκτρονικών υπολογιστών σε συνδυασμό με την δυνατότητα για γρήγορη ανάκτηση των αποθηκευμένων πληροφοριών οδήγησαν από πολύ νωρίς στην χρήση των υπολογιστών σαν αποθήκες πληροφορίας.

Χρησιμοποιώντας τις σύγχρονες τεχνολογίες, όπως τα δίκτυα υπολογιστών και το διαδίκτυο, τα κατανεμημένα συστήματα και τις νέες τεχνολογίες αποθήκευσης, οι υπολογιστές έγιναν ακόμη πιο αποδοτικοί στην διαχείριση μεγάλων ποσοτήτων πληροφορίας. Έτσι, έχουμε την δυνατότητα αποθήκευσης ακριβέστερων αναπαραστάσεων του κόσμου.

Στην προσπάθεια μας για την όλο και ακριβέστερη αναπαράσταση της γνώσης μας όμως, δεν αρκεί μόνο η αποθήκευση όλο και μεγαλύτερου όγκου πληροφορίας. Νέα μοντέλα και τρόποι αποθήκευσης πληροφορίας έχουν αναπτυχθεί, με στόχο την ακριβέστερη αναπαράσταση της πραγματικότητας. Το κάθε μοντέλο έχει διαφορετικές δυνατότητες και χρησιμοποιείται για διαφορετικές εφαρμογές. Βασικά κριτήρια αξιολόγησης του κάθε μοντέλου αναπαράστασης είναι η εκφραστική του δύναμη, δηλαδή το είδος των πληροφοριών που μπορούμε να αποθηκεύσουμε, η ταχύτητα διαχείρισης, ενημέρωσης και συντήρησης της πληροφορίας, καθώς και η δυνατότητα ανάκτησης πληροφορίας που παρέχεται στον χρήστη. Επιπλέον χαρακτηριστικά που θα μπορούσαμε να λάβουμε υπόψιν μας για την αξιολόγηση τέτοιων τεχνικών είναι οι δυνατότητες ενημέρωσης που παρέχονται στον χρήστη, οι δυνατότητες σύνθεσης πληροφορίας για την εξαγωγή συμπερασμάτων ή άλλου είδους επαγόμενης γνώσης, ή ακόμα και ο χώρος που απαιτείται για την αποθήκευση της πληροφορίας.

1.2 Ποιο είναι το Πρόβλημα

Για να είναι χρήσιμο ένα σύστημα αποθήκευσης πληροφορίας, πρέπει ανά πάσα στιγμή να περιέχει την κατά το δυνατόν πληρέστερη, σωστότερη και πιο πρόσφατη πληροφορία για τον κόσμο τον οποίο μοντελοποιεί. Για να γίνει αυτό, πρέπει η αρχική γνώση να αναθεωρείται συνεχώς, με κάθε νέα παρατήρηση που κάνουμε για τον κόσμο. Με άλλα λόγια, πρέπει να οριστεί μία πράξη (την οποία θα συμβολίζουμε με “ * “ , η οποία εφαρμοζόμενη μεταξύ μίας βάσης γνώσης και μίας ενημέρωσης (update), θα μας δίνει μία νέα βάση γνώσης. Ο ορισμός της πράξης αυτής αποτελεί ένα μόνο από τα κομμάτια του προβλήματος της αναθεώρησης πληροφορίας (belief revision). Ο τρόπος αναπαράστασης της γνώσης και της ενημέρωσης αποτελεί ένα υποπρόβλημα άμεσα σχετιζόμενο με το πρώτο. Στην παρούσα εργασία μας ενδιαφέρουν βάσεις γνώσης και ενημερώσεις εκφρασμένες σαν λογικές εκφράσεις του Προτασιακού Λογισμού. Το αντίστοιχο πρακτικό κομμάτι του προβλήματος είναι η αναπαράσταση της γνώσης σε μορφή κατανοητή από ηλεκτρονικό υπολογιστή, καθώς και η ανάπτυξη ενός κατά το δυνατόν αποδοτικότερου αλγορίθμου ο οποίος θα υλοποιεί την πράξη της ενημέρωσης (“ * “) όπως αυτή ορίστηκε παραπάνω.

Το πρόβλημα λοιπόν της ενημέρωσης της πληροφορίας (belief revision), συνίσταται στην εύρεση ενός αλγορίθμου μετατροπής της γνωστής πληροφορίας σε νέα, με βάση κάποια νέα παρατήρηση σχετική με τον υπό μοντελοποίηση κόσμο. Η νέα αυτή παρατήρηση (πληροφορία) μπορεί σε πολλές περιπτώσεις να αντιφάσκει με τις τρέχουσες πεποιθήσεις μας για τον κόσμο. Η περίπτωση αυτή είναι η δυσκολότερη, όπως θα δούμε παρακάτω, αλλά ταυτόχρονα και η πιο ενδιαφέρουσα, δεδομένου ότι όταν δεν υπάρχουν αντιφάσεις αρκεί η απλή συγχώνευση της νέας πληροφορίας με την παλιά.

Η περίπτωση αντιφατικής πληροφορίας δεν θα πρέπει να αντιμετωπίζεται σαν σπάνια, ειδική περίπτωση ενημέρωσης. Υπάρχουν πολλοί λόγοι για τους οποίους μπορεί η νέα (ή αντίστοιχα η παλιά) πληροφορία να είναι λανθασμένη. Επίσης, η γνώση μας είναι κατά κανόνα ημιτελής. Νέα δεδομένα ή γνώση

μπορούν να αναιρέσουν την τρέχουσα οπτική μας για τον κόσμο. Είναι πιθανό να υπάρχει πληροφορία η οποία να είναι δύσκολο, αδύνατο ή πολύ ακριβό να αποκτηθεί. Κάποιο κομμάτι πληροφορίας μπορεί να αφορά προσωπικά δεδομένα και έτσι να μην είναι διαθέσιμο. Σε μία κατανεμημένη βάση, κάποιος χρήστης μπορεί να βλέπει μόνο ένα κομμάτι της γνωστής πληροφορίας και όχι όλη την βάση.

Τέλος, κανείς δεν μπορεί να μας εγγυηθεί ότι ο κόσμος μας είναι στατικός. Κάτι το οποίο ισχύει σε μία δεδομένη χρονική στιγμή μπορεί να μην ισχύει σε κάποια επόμενη, αφού ενδέχεται ο κόσμος μας να αλλάζει δυναμικά. Τα παραπάνω κάνουν σαφές ότι καμία γνώση δεν μπορεί να θεωρείται αυθεντική και απόλυτα σωστή. Έτσι οφείλουμε να είμαστε προσεκτικοί τόσο κατά την απόρριψη όσο και κατά την αποδοχή ενημερώσεων-παρατηρήσεων. Αυτό αποτελεί και το πρώτο μεγάλο πρόβλημα της ενημέρωσης πληροφορίας.

Όπως μπορεί εύκολα να γίνει αντιληπτό, η αξία του προβλήματος της αναθεώρησης γνώσης είναι πολύ μεγάλη. Οι γνώσεις μας για τον υπό μοντελοποίηση κόσμο δεν είναι όπως είδαμε πλήρεις ούτε σωστές εκ των προτέρων. Ακόμα και αν είναι όμως, πρέπει να είμαστε προετοιμασμένοι για το ενδεχόμενο της αλλαγής του δυναμικού κόσμου που μας περιβάλλει. Άρα λοιπόν η ύπαρξη μίας μεθόδου αναθεωρήσεων είναι απαραίτητη, και μάλιστα η μέθοδος αυτή οφείλει να αναθεωρεί την γνώση με τον καλύτερο δυνατό τρόπο. Ένας αλγόριθμος αναθεωρήσεων ο οποίος επιλέγει λάθος πληροφορία προς απόρριψη-αποδοχή θα μας οδηγήσει σύντομα στην μοντελοποίηση ενός κόσμου πολύ διαφορετικού από τον πραγματικό, κάτι σίγουρα μη επιθυμητό – και αυτό αποτελεί μόνο το ένα (το πιο εύκολο) σκέλος του προβλήματος. Το δεύτερο σκέλος έχει να κάνει με την σωστή διαισθητικά μετάβαση από την αρχική βάση στην ενημερωμένη. Το βασικό μας κίνητρο κατά την σχεδίαση μίας μεθόδου αναθεώρησης πληροφορίας οφείλει να είναι η ανθρώπινη διαίσθηση. Δεν αρκεί μόνο η σωστή αποδοχή-απόρριψη γνώσης. Σε μερικές περιπτώσεις μπορούμε με μικρές αλλαγές στις πεποιθήσεις μας να αποδεχτούμε μία νέα αντιφατική παρατήρηση. Σε άλλες περιπτώσεις πάλι, είναι απαραίτητη η πλήρης απόρριψη κάποιας ή κάποιων παρατηρήσεων (ή της ενημέρωσης) προκειμένου να καταλήξουμε σε ένα συνεπές σύστημα γνώσης. Όπως θα δούμε, στην πλειοψηφία των περιπτώσεων, υπάρχουν πολλοί διαφορετικοί τρόποι

ενημέρωσης της γνώσης μας. Η σωστότερη προσέγγιση (θα πρέπει να) υποδεικνύεται από την ανθρώπινη σκέψη.

Στην καθημερινή ζωή, είμαστε υποχρεωμένοι να αλλάζουμε συνεχώς τις πεποιθήσεις μας, προσθέτοντας νέες και αφαιρώντας ή αλλάζοντας παλιές. Νέες γνώσεις για τον κόσμο που μας περιβάλλει καταφθάνουν συνεχώς στον ανθρώπινο εγκέφαλο μέσω των αισθητήριων οργάνων. Όταν μάθουμε κάτι σχετικό με τον κόσμο γύρω μας, τότε η νέα πληροφορία αποτελεί πλέον μέρος των πεποιθήσεων μας. Στην συνέχεια ξεκινάει μία διαδικασία αποδοχής ή απόρριψης της νέας και των παλιών πεποιθήσεων, η αναθεώρηση πραγμάτων που θεωρούσαμε δεδομένα, η μελέτη των συνεπειών της νέας γνώσης σε συνδυασμό με την παλιά. Αν η διαδικασία αυτή προκαλέσει την δημιουργία νέας (επαγόμενης) γνώσης, τότε εισάγεται και αυτή ως νέα γνώση, οπότε η όλη διαδικασία επαναλαμβάνεται από την αρχή.

Οι ακριβείς διανοητικές διαδικασίες ενημέρωσης γνώσης δεν είναι απόλυτα γνωστές και αποτελούν ενεργό πεδίο μελέτης για διάφορες επιστήμες όπως ψυχολογία και φιλοσοφία. Στην πράξη πάντως, οι διαδικασίες αυτές καθαυτές δεν είναι τόσο σημαντικές όσο το αποτέλεσμά τους, δηλαδή το αποτέλεσμα της ενημέρωσης των πεποιθήσεών μας.

Ξεφεύγοντας λίγο από την παραπάνω αφηρημένη μελέτη, ας δούμε τι συμβαίνει στις προτασιακές βάσεις γνώσης. Εδώ έχουμε γνώση η οποία εκφράζεται με ένα σύνολο από λογικές εκφράσεις. Οι προτάσεις αυτές εκφράζουν την τρέχουσα γνώση μας για τον κόσμο. Μία νέα πληροφορία αποτελεί μία νέα λογική έκφραση, η οποία μπορεί απλά να προστεθεί στο σύνολο των λογικών εκφράσεων, οπότε και πετυχαίνουμε την επιθυμητή αναθεώρηση.

Τα προβλήματα αρχίζουν όταν αυτή η νέα γνώση αναιρεί κάποια από τις παλιές. Αυτό μπορεί να συμβαίνει άμεσα (δηλαδή να αντιφάσκει άμεσα με μία παλιά παρατήρηση) ή έμμεσα (δηλαδή να αντιφάσκει με μία από τις συνέπειες κάποιας ή κάποιων παλιών παρατηρήσεων). Αν συμβαίνει αυτό, με την απλή εισαγωγή της νέας γνώσης στο σύνολο πεποιθήσεών μας, θα καταλήξουμε σε ένα σύστημα γνώσης αντιφατικό, στο οποίο θα πιστεύουμε ταυτόχρονα δύο αλληλοαναιρούμενες πληροφορίες. Έτσι, αυτό που πρέπει να γίνει κατά την εισαγωγή κάθε νέας πληροφορίας είναι σε πρώτη φάση η αναζήτηση των

αντιφάσεων ,αν υπάρχουν και σε δεύτερη η μελέτη των διαφόρων τρόπων άρσης των αντιφάσεων ώστε να επιλεχθεί ο βέλτιστος.

Η εύρεση των αντιφάσεων είναι ένα πρόβλημα του οποίου η θεωρητική λύση είναι όπως θα δούμε σχετικά απλή, αν και οι σχετικοί αλγόριθμοι δεν είναι αποδοτικοί στην πράξη. Το δεύτερο πρόβλημα είναι αρκετά δυσκολότερο θεωρητικά αφού δεν είναι απόλυτα σαφές ακόμα και σήμερα ποιος είναι ο «βέλτιστος» τρόπος άρσης μίας αντίφασης. Για να κατανοήσουμε ένα μέρος των δυσκολιών του προβλήματος, ας δούμε ένα σχετικά απλό παράδειγμα [1], σελ. 1.

Η Φιλίππα, μία δεκαεννιάχρονη φοιτήτρια στο Πανεπιστήμιο Πάτρας, μόλις ανακάλυψε ότι ο Νίκος και η Angela δεν είναι οι αληθινοί γονείς της. Υιοθετήθηκε όταν ήταν έξι μηνών από ένα ορφανοτροφείο στο Σάο Πάολο. Τα νέα τάραξαν πραγματικά την Φιλίππα. Ένα μεγάλο μέρος αυτών που πίστευε για όλη τη ζωή της και για την οικογένειά της ήτανε λάθος. Μετά τον αρχικό κλονισμό άρχισε να βάζει τις σκέψεις της σε μια σειρά, δηλαδή αυτό σήμαινε ότι ο Αλέξανδρος δεν είναι πραγματικά ο ξάδελφός της, ότι δεν πήρε τα καφετιά μάτια της από την γιαγιά της και πλέον δεν χρειάζεται να ανησυχεί σχετικά με το ιστορικό της οικογένειας για υψηλή πίεση αίματος από την πλευρά του Νίκου και της Angela. Επιπλέον, έχει πιθανώς συγγενείς κάπου στη Βραζιλία, και εάν το εξέταζε πιο προσεχτικά, μπορεί να έχει το δικαίωμα της βραζιλιανής υπηκοότητας, που θα μπορούσε να είναι βολικό σε εκείνο το μακροχρόνιο ταξίδι που πάντα ήθελε να κάνει στη Λατινική Αμερική.

Αυτό είναι μία χαρακτηριστική περίπτωση αναθεώρησης πεπιοθήσεων, ένας λογικός πράκτορας λαμβάνει νέες πληροφορίες που αλλάζουν τις παλαιές πεπιοθήσεις του. Στην αρχική περίπτωση όπου οι νέες πληροφορίες έρχονται σε αντίθεση με τις αρχικές πεπιοθήσεις του, ο πράκτορας πρέπει να αποσύρει μερικές από τις παλαιές πεπιοθήσεις προτού να μπορέσει να προσαρμόσει τις νέες πληροφορίες και επίσης πρέπει να δεχτεί τις συνέπειες που να προκύψουν από την αλληλεπίδραση των νέων πληροφοριών με τις παλαιές πεπιοθήσεις. Η μελέτη της διαδικασίας της αναθεώρησης πεπιοθήσεων, που έδωσε αφορμή για έναν συναρπαστικό ερευνητικό τομέα με το ίδιο όνομα (Αναθεώρηση Πεπιοθήσεων), μπορεί να επισημανθεί στις αρχές της δεκαετίας του '80. Το άρθρο που θεωρείται ευρέως γνωστό για να χαρακτηρίσει τη γέννηση του τομέα

είναι η δημιουργική εργασία των Alchourron, Gardenfors και Makinson [1]. Στην πραγματικότητα, το πλαίσιο που εξελίχθηκε από το άρθρο [1], το οποίο είναι γνωστό τώρα ως AGM παράδειγμα (ή απλά AGM) από τα αρχικά των τριών ιδρυτών του, είναι το κυρίαρχο άρθρο στην Αναθεώρηση Πεποιθήσεων.

Φυσικά πολλά έχουν συμβεί από το 1985. Οι αρχικές προσεγγίσεις που αναπτύχθηκαν στο άρθρο [1], έχουν επλουτηθεί και μελετηθεί λεπτομερώς, νέες ερευνητικές κατευθύνσεις έχουν προκύψει, συνδέσεις με γειτονικούς τομείς έχουν καθιερωθεί, και πολλά περισσότερα βρίσκονται εν εξελίξει.

Οι παραπάνω παρατηρήσεις αποτελούν μία πρώτη προσέγγιση τόσο της σπουδαιότητας όσο και της δυσκολίας του προβλήματος της αναθεώρησης πεποιθήσεων σε λογικές βάσεις γνώσης. Άλλωστε, η ανάμιξη τόσων διαφορετικών επιστημονικών κλάδων (ψυχολογία, φιλοσοφία, λογική, μαθηματικά, επιστήμη υπολογιστών κτ) στην μελέτη του προβλήματος αυτού αποτελεί μία ακόμα ένδειξη της σημασίας του.

1.3 Ιστορική Αναδρομή

Η αναθεώρηση πεποιθήσεων (αλλαγή πεποιθήσεων, δυναμική πεποιθήσεων) είναι ένα νέο πεδίο στον τομέα της έρευνας και έχει αναγνωριστεί ως αντικείμενο έρευνας από τα μέσα της δεκαετίας του 1980. Το νέο αυτό πεδίο αναπτύχθηκε μέσα από δύο συγκλίνοντες ερευνητικές ερμηνείες.

Μία από αυτές είναι η επιστήμη των υπολογιστών. Από τις αρχές της πληροφορικής, οι προγραμματιστές έχουν αναπτύξει διαδικασίες με τις οποίες οι βάσεις δεδομένων μπορούν να ενημερώνονται. Η ανάπτυξη της Τεχνητής Νοημοσύνης ενέπνευσε τους επιστήμονες να κατασκευάσουν πιο εξελιγμένα μοντέλα αναθεώρησης βάσεων δεδομένων. Τα συστήματα συντήρησης που αναπτύχθηκαν από τον Jon Doyle (1979) ήταν σημαντικά για την εξέλιξη αυτή. Μία από τις πιο σημαντικές θεωρητικές συνεισφορές ήταν η δημοσίευση το 1983 από τον Ronald Fagin, Jeffrey Ullman και Moshe Βέρντι, με την οποία εισήχθη η έννοια των προτεραιοτήτων της βάσης δεδομένων.

Η δεύτερη από αυτές τις δύο ερευνητικές ερμηνείες είναι φιλοσοφικού περιεχομένου. Με μια ευρεία έννοια, η αλλαγή πεποιθήσεων έχει αποτελέσει αντικείμενο φιλοσοφικού στοχασμού από την αρχαιότητα. Στον εικοστό αιώνα,

οι φιλόσοφοι έχουν μελετήσει τους μηχανισμούς, με τους οποίους οι επιστημονικές θεωρίες αναπτύχθηκαν, και έχουν προτείνει τα κριτήρια του ορθολογισμού για τις αναθεωρήσεις των αναθέσεων. Ξεκινώντας τη δεκαετία του 1970, μια πιο εστιασμένη μελέτη έχει γίνει σχετικά με τις απαιτήσεις της ορθολογικής αλλαγής πεποιθήσεων. Δύο κύρια σημεία μπορούν να τονιστούν. Το πρώτο ήταν μια σειρά από μελέτες που διεξήχθησαν από τον Ισαάκ Levi στη δεκαετία του 1970 (Levi 1977, 1980). Ο Levi θέτει πολλά από τα προβλήματα σε αυτόν τον τομέα της έρευνας. Παρείχε, επίσης, μεγάλο μέρος του βασικού επίσημου πλαισίου (Framework). Το έργο William Harper (1977) την ίδια περίοδο είχε επίσης μια διαρκή επιρροή.

Επόμενο ορόσημο ήταν το AGM μοντέλο, που ονομάζεται έτσι μετά από τους τρεις δημιουργούς του, Carlos Alchourrón, Peter Gärdenfors, και David Makinson (1985). Οι Alchourrón και Makinson είχαν προηγουμένως συνεργαστεί σε μελέτες αναθεώρησης του νομικού συστήματος (Alchourrón και Makinson 1981, 1982). Το προηγούμενο έργο του Gärdenfors ήταν πάνω στις συνδέσεις μεταξύ της αναθεώρησης πεποιθήσεων και των υποθετικών προτάσεων (Gärdenfors 1978, 1981). Με τον συνδυασμό των τριών έργων μια δημοσίευση, που παρείχε ένα νέο, πολύ πιο γενικό και ευέλικτο επίσημο πλαίσιο για τις μελέτες της αναθεώρησης πεποιθήσεων. Η παραπάνω δημοσίευση, δημοσιεύθηκε στο περιοδικό *Journal of Symbolic Logic*, το 1985, και οι ιδέες της έχουν αποτελέσει αντικείμενο επεξεργασίας και ανάπτυξης. Το AGM παράδειγμα και οι εξελίξεις που έχουν αναπτυχθεί από αυτό εξακολουθούν να αποτελούν τον πυρήνα της θεωρίας της αναθεώρησης πεποιθήσεων.

Ένας φυσικός προβληματισμός της επιστημολογίας είναι να προτείνει μια θεώρηση δικαιολογημένης αναθεώρησης πεποιθήσεων. Υπάρχουν δύο δρόμοι που μπορεί να ακολουθήσει κανείς σε μια τέτοια περίπτωση:

- Η αξιωματική θεώρηση, η οποία προτείνει αφηρημένες γενικές αρχές που νοηματικά επενεργούν πάνω στη διαδικασία και αξιωματικοποιούν την αναθεώρηση πεποιθήσεων.

- Η παραγωγική θεώρηση, η οποία προσπαθεί να αντλήσει μια θεωρία αναθεώρησης πεποιθήσεων από μια πιο εποικοδομητική επιστημολογική θεωρία[3].

Η πιο δημοφιλής και ευρέως διαδεδομένη αξιωματική θεώρηση είναι η AGM (από τα αρχικά των συγγραφέων) θεώρηση, η οποία είναι κατασκευάσμα των Alchourrón, C. E., P. Gärdenfors, και D. Makinson (1985) και αναπτύχθηκε περαιτέρω από τον P. Gärdenfors (1988, 1992) και τους Alchourrón και Makinson (1982, 1985). Τέτοιου είδους θεωρήσεις προχωρούν σε ένα πολύ μεγάλο επίπεδο αφαίρεσης, αγνοώντας όψεις γνωσιακής ορθολογικότητας, χωρίς την οποία είναι αδύνατο να δημιουργήσουμε πραγματικές αρχές ορθολογικής αναθεώρησης πεποιθήσεων. Μια θεωρία αναθεώρησης πεποιθήσεων οφείλει να καθοδηγείται από μια πιο εποικοδομητική επιστημολογία.

Η θεωρία αναθεώρησης πεποιθήσεων δημιουργήθηκε κατ' αρχήν για να απαντήσει στο ερώτημα πως οι άνθρωποι αλλάζουν γνώμη, ακολουθώντας μια μακρά παράδοση φιλοσοφικού έργου στη θεωρία της αλλαγής πεποιθήσεων. Κατά τις δύο τελευταίες δεκαετίες η πρωταρχική αξιωματική θεώρηση έχει συζητηθεί, ανασκευαστεί και δεχτεί εκτενέστατη κριτική. Άλλα πιο ρεαλιστικά, και πιο εύκολα υπολογιστικά μοντέλα αναθεώρησης πεποιθήσεων έχουν προταθεί. Κανονιστικές θεωρίες αναθεώρησης, οι οποίες υποστηρίζονται από φιλοσόφους και Λογικούς, που μελετάνε το ερώτημα του πως θα έπρεπε οι άνθρωποι να αλλάζουν γνώμες, βαδίζουν παράλληλα με πιο περιγραφικές θεωρήσεις που δημιουργήθηκαν από γνωσιακούς ψυχολόγους και κοινωνικούς επιστήμονες, οι οποίοι θέλουν να ξέρουν πως οι άνθρωποι στη πραγματικότητα λειτουργούν. Ως αποτέλεσμα, η αναθεώρηση πεποιθήσεων έχει γίνει μια από τις πιο ενεργές περιοχές έρευνας, με τη συμβολή μεταξύ γνωσιακής επιστήμης, λογικής, φιλοσοφίας και της τεχνητής νοημοσύνης.

1.4 Η Προσέγγιση των Alchourron, Gärdenfors και Makinson

Στην παραπάνω προσέγγιση, που αναπτύχθηκε από τους Alchourron, Gärdenfors και Makinson η έμπνευση των συγγραφέων ήταν η προσωρινή

παραίτηση από την προσπάθεια εύρεσης συγκεκριμένου αλγορίθμου για το πρόβλημα αναθεώρησης πεποιθήσεων και η αναζήτηση συγκεκριμένων επιθυμητών ιδιοτήτων ενός τέτοιου αλγορίθμου. Οι ιδιότητες αυτές εκφράστηκαν υπό την μορφή λογικών αξιωμάτων, τα οποία ονομάστηκαν αξιώματα AGM (AGM postulates) από τα αρχικά των ονομάτων των συγγραφέων.

Τα αξιώματα αυτά αναφέρονται στην πράξη της αναθεώρησης. Κατά την ανάπτυξη των αξιωμάτων υποτίθεται ότι η γνώση αποτελείται από ένα σύνολο K από λογικές προτάσεις. Το K είναι κλειστό ως προς την σχέση της παραγωγής, δηλαδή αν υπάρχουν λογικές προτάσεις $P_i \in K$, τέτοιες ώστε η πρόταση: $\bigwedge_i P_i$ να έπεται την P , τότε $P \in K$. Αν συμβολίσουμε με $Cn(K)$ το σύνολο των επαγόμενων προτάσεων ενός συνόλου K , έχουμε ότι: $Cn(K) = \{p \mid \exists I, i \in I, p_i \in K : \bigwedge_i p_i \rightarrow p\}$. Με τον συμβολισμό αυτό, θα λέμε ότι ένα σύνολο K είναι κλειστό ως προς την πράξη της παραγωγής αν και μόνο αν $K = Cn(K)$. Τέτοιου είδους σύνολα ονομάζονται και σύνολα πεποιθήσεων (belief sets). Αν ισχύει ότι $Cn(K) = L^*$, τότε λέμε ότι το K είναι ασυνεπές. Εύκολα παρατηρεί κανείς ότι τελικά υπάρχει μόνο ένα ασυνεπές σύνολο πεποιθήσεων, το οποίο συμβολίζουμε με K_{\perp} . Το σύνολο αυτό είναι το μοναδικό που περιέχει την λογική πρόταση F .

Για την αντιμετώπιση του προβλήματος της αναθεώρησης, αρχικά ορίζεται η πράξη της επέκτασης (expansion), η οποία συμβολίζεται με το σύμβολο '+' και είναι η απλή προσθήκη μίας πρότασης p (και των συνεπειών της) σε ένα σύνολο πεποιθήσεων K , χωρίς έλεγχο συνέπειας του αποτελέσματος. Η επέκταση ορίζεται με την βοήθεια του τύπου: $K+p = Cn(K \cup \{p\})$.

Η επέκταση δεν είναι χρήσιμη πράξη διότι δίνει σαν αποτέλεσμα ένα ασυνεπές σύνολο (το K_{\perp}) όταν η νέα πληροφορία αντιφάσκει με την παλιά γνώση. Έτσι ορίζεται και η (πιο ενδιαφέρουσα) πράξη της αναθεώρησης, την οποία θα συμβολίσουμε με " * ".

Οι παραπάνω ταυτότητες έχουν το νόημα ότι για κάθε μέθοδο αναθεωρήσεων η οποία ικανοποιεί τα αξιώματα AGM για τις αναθεωρήσεις, μπορούμε να ορίσουμε μία αντίστοιχη πράξη αφαίρεσης, μέσω της ταυτότητας Harper. Η πράξη αφαίρεσης που θα οριστεί με τον τρόπο αυτό θα ικανοποιεί τα αξιώματα

AGM για τις αφαιρέσεις. Αντίστοιχη ιδιότητα ισχύει κατά την μετατροπή από αφαίρεση σε αναθεώρηση μέσω της ταυτότητας Levi.

Αφού ορίστηκαν τα αξιώματα και μελετήθηκαν οι συνέπειές τους, έγινε μία προσπάθεια να οριστεί κάποια συγκεκριμένη πράξη η οποία να ικανοποιεί τα αξιώματα και διαπιστώθηκε ότι υπάρχουν πολλές πράξεις με αυτές τις ιδιότητες. Επειδή θεωρήθηκε ότι η αφαίρεση είναι πιο εύκολη πράξη διαισθητικά, μελετήθηκαν κλάσεις συναρτήσεων αφαίρεσης και όχι αναθεώρησης. Οι αντίστοιχες συναρτήσεις αναθεώρησης μπορούν να υπολογιστούν μέσω της ταυτότητας Levi σε κάθε περίπτωση.

Οι πράξεις αφαίρεσης χωρίστηκαν σε δύο μεγάλες κλάσεις, τις συναρτήσεις πλήρους τομής (full-meet contraction functions) και μερικής τομής (partial-meet contraction functions).

Κεφάλαιο 2

2.1 Προτασιακή Λογική

Η παρούσα εργασία θα επικεντρωθεί στο μοντέλο των λογικών βάσεων πεποιθήσεων με χρήση προτασιακής λογικής. Η προτασιακή λογική βασίζεται σε μία γλώσσα L , η οποία είναι ένα σύνολο συμβόλων. Η L περιέχει μεταβλητές τις οποίες ονομάζουμε άτομα και συνδέσμους που ονομάζουμε πράξεις ή τελεστές. Τα άτομα μπορούν να είναι πεπερασμένου ή απείρως αριθμήσιμου πλήθους και οι τελεστές που συνήθως χρησιμοποιούνται είναι αυτοί της σύζευξης (\wedge), διάζευξης (\vee), άρνησης (\neg), συνεπαγωγής (\rightarrow) και ισοδυναμίας (\leftrightarrow). Επίσης, η L περιέχει δύο σταθερές τιμές, που συμβολίζουμε με T (αληθές) και F (ψευδές) και ονομάζουμε αληθοτιμές. Για να εξασφαλιστεί η μονοσήμαντη αναγνωσιμότητα των λογικών εκφράσεων που θα προκύψουν από την L , προσθέτουμε στο σύνολο L και τις παρενθέσεις ($(,)$). Έτσι μία τυπική γλώσσα L είναι το σύνολο: $L = \{\wedge, \vee, \neg, \leftrightarrow, \rightarrow, T, F, (,), a_1, a_2, a_3, \dots\}$.

Τα στοιχεία του συνόλου L μπορούν να συνδυαστούν για να σχηματίσουν συμβολοσειρές. Κάθε συμβολοσειρά που σχηματίζεται με συνδυασμό

πεπερασμένου πλήθους στοιχείων του συνόλου L λέγεται έκφραση. Όμως δεν είναι όλες οι εκφράσεις αποδεκτές. Το σύνολο των αποδεκτών εκφράσεων (ή καλά σχηματισμένων τύπων, όπως λέγονται) συμβολίζεται με L^* . Τους καλά σχηματισμένους τύπους θα τις λέμε και προτάσεις. Υπάρχουν ακριβείς κανόνες οι οποίοι καθορίζουν αν μία έκφραση είναι πρόταση ή όχι (δηλαδή αν ανήκει στο L^*). Έτσι, για παράδειγμα, είναι γνωστό ότι για την γλώσσα $L = \{a_1, a_2, a_3, \wedge, \vee, \neg, \leftrightarrow, \rightarrow, T, F, (,)\}$ η συμβολοσειρά " $(\wedge \neg a_1 \vee a_2$ " δεν είναι αποδεκτός τύπος. Οι κανόνες σχηματισμού καλά σχηματισμένων τύπων είναι οι εξής:

- 1) Κάθε λογική έκφραση της μορφής x όπου x : άτομο ή σταθερά (T, F) της L είναι καλά σχηματισμένη (δηλαδή $x \in L^*$).
- 2) Αν η λογική έκφραση x είναι καλά σχηματισμένη, τότε και η λογική έκφραση $\neg(x)$ είναι καλά σχηματισμένη (δηλαδή $x \in L^* \Rightarrow \neg(x) \in L^*$).
- 3) Αν οι λογικές εκφράσεις x, y είναι καλά σχηματισμένες, τότε και οι τύποι $(x) \vee (y)$, $(x) \wedge (y)$, $(x) \rightarrow (y)$, $(x) \leftrightarrow (y)$ είναι καλά σχηματισμένοι (δηλαδή $x, y \in L^* \Rightarrow (x) \vee (y) \in L^*$, $x, y \in L^* \Rightarrow (x) \wedge (y) \in L^*$, $x, y \in L^* \Rightarrow (x) \rightarrow (y) \in L^*$, $x, y \in L^* \Rightarrow (x) \leftrightarrow (y) \in L^*$).
- 4) Όλες οι καλά σχηματισμένοι τύποι ακολουθούν τους κανόνες 1, 2, 3.

Σε πολλές περιπτώσεις, οι κανόνες αυτοί δεν τηρούνται κατά γράμμα. Έτσι θα θεωρούμε πολλές φορές και τον τύπο $\neg a_1 \vee a_2 \in L^*$ σαν καλά σχηματισμένος τύπο, παρόλο που η σωστή μορφή του είναι η $(\neg(a_1)) \vee (a_2) \in L^*$. Πιο συγκεκριμένα, για λόγους αναγνωσιμότητας, θα αποφεύγουμε τις μη απαραίτητες παρενθέσεις, ορίζοντας την προτεραιότητα μεταξύ των τελεστών της L ως εξής:

- 1) Άρνηση: \neg
- 2) Σύζευξη, διάζευξη: \wedge, \vee
- 3) Συνεπαγωγή, ισοδυναμία: $\rightarrow, \leftrightarrow$

Με δεδομένη την παραπάνω προτεραιότητα, μπορούμε να δημιουργήσουμε πιο αναγνώσιμες εκφράσεις με λιγότερες παρενθέσεις, χωρίς να χάνουμε σε εκφραστικότητα.

Κάθε άτομο (μεταβλητή) της γλώσσας δέχεται δύο μόνο δυνατές τιμές, τις T και F (αληθές και ψευδές αντίστοιχα). Κάθε (αυθαίρετη) ανάθεση αληθοτιμών στα άτομα της L ονομάζεται ερμηνεία. Μία έκφραση του συνόλου των προτάσεων L^* δεν είναι από μόνη της αληθής ή ψευδής. Για κάθε δεδομένη ερμηνεία όμως, μπορούμε να βρούμε την αληθοτιμή μίας οποιασδήποτε πρότασης του L^* . Αυτό γίνεται με αναδρομικό τρόπο χρησιμοποιώντας τους παρακάτω κανόνες, οι οποίοι δίνουν το αποτέλεσμα της εφαρμογής κάθε τελεστή:

- 1) $\neg x = T \Leftrightarrow x = F$
- 2) $x \wedge y = T \Leftrightarrow x = T$ και $y = T$
- 3) $x \vee y = T \Leftrightarrow x = T$ ή $y = T$
- 4) $x \rightarrow y = T \Leftrightarrow x = F$ ή $y = T$
- 5) $x \leftrightarrow y = T \Leftrightarrow (x = T$ και $y = T)$ ή $(x = F$ και $y = F)$

Μία πρόταση λέμε ότι βρίσκεται σε κανονική διαζευκτική μορφή (disjunctive normal form – DNF) αν και μόνο αν είναι της μορφής:

$$P = (P_{11} \wedge P_{12} \wedge \dots \wedge P_{1k_1}) \vee \dots \vee (P_{n1} \wedge P_{n2} \wedge \dots \wedge P_{nk_n})$$

όπου τα P_{ij} είναι άτομα ή αρνήσεις ατόμων. Στην τετριμμένη περίπτωση που $n=0$, η παραπάνω έκφραση γίνεται $P=F$. Από σύμβαση θα λέμε ότι και αυτή η έκφραση είναι σε DNF. Επίσης, κάθε ένας από τους όρους της παραπάνω διάζευξης, για παράδειγμα ο τύπος $P_{11} \wedge P_{12} \wedge \dots \wedge P_{1k_1}$, θα ονομάζεται ελάχιστος όρος. Έχει αποδειχθεί ότι για κάθε έκφραση $P \in L^*$ του Προτασιακού Λογισμού υπάρχει μία τουλάχιστον έκφραση $Q \in L^*$ σε DNF, η οποία να είναι ισοδύναμή της, δηλαδή $P \equiv Q$. Αντίστοιχα ορίζεται η κανονική συζευκτική μορφή (conjunctive normal form – CNF). Λέμε ότι μία έκφραση είναι σε CNF αν και μόνο αν είναι της μορφής:

$$P = (P_{11} \vee P_{12} \vee \dots \vee P_{1k_1}) \wedge \dots \wedge (P_{n1} \vee P_{n2} \vee \dots \vee P_{nk_n})$$

όπου τα P_{ij} είναι όπως παραπάνω.

Το πρόβλημα αυτό έχει μεγάλη σχέση με την **αναθεώρηση πεποιθήσεων**, διότι ο έλεγχος ικανοποιησιμότητας της σύζευξης ενός συνόλου εκφράσεων αποτελεί υποπρόβλημα της αναθεώρησης πεποιθήσεων

2.2 Αναπαράσταση της γνώσης

Μία λογική έκφραση είναι αρκετή για να εκφράσει την γνώση που περιέχεται στην βάση, ενώ η προσέγγιση των AGM αξιωμάτων χρησιμοποιεί ένα σύνολο πεποιθήσεων, δηλαδή ένα σύνολο λογικών εκφράσεων, οι οποίες είναι κλειστές ως προς την πράξη της παραγωγής. Η διαφορά αυτή αντικατοπτρίζει μία γενικότερη διαφωνία που υπάρχει στον χώρο και σχετίζεται με την αναπαράσταση της γνώσης. Η αναπαράσταση που θα χρησιμοποιηθεί προφανώς επηρεάζει και τον τρόπο με τον οποίο θα γίνει η ενημέρωση, άρα είναι πολύ σημαντική για το πρόβλημά μας.

Σε αντίθεση με την προσέγγιση των συνόλων πεποιθήσεων – δηλαδή συνόλων από λογικές εκφράσεις τα οποία είναι κλειστά ως προς την παραγωγή – πολλοί ερευνητές προτείνουν την χρήση βάσεων πεποιθήσεων (belief base). Οι προσεγγίσεις που χρησιμοποιούν βάσεις πεποιθήσεων βασίζονται στην παρατήρηση ότι, στην γενική περίπτωση, τα σύνολα πεποιθήσεων αποτελούνται από μεγάλο (ενδεχομένως και άπειρο) πλήθος λογικών προτάσεων. Έτσι, πρακτικά είναι δύσκολο ή αδύνατο να υλοποιήσουμε μεθόδους οι οποίες ενημερώνουν σύνολα πεποιθήσεων. Εντούτοις, για κάθε κλειστό ως προς την πράξη της παραγωγής σύνολο λογικών προτάσεων K , υπάρχει σύνολο B με πολύ μικρότερο πλήθος λογικών προτάσεων για το οποίο ισχύει ότι $Cn(B)=K$. Τέτοια σύνολα ονομάζονται βάσεις πεποιθήσεων και είναι σαφές ότι μία μέθοδος ενημέρωσης που αναφέρεται σε βάσεις πεποιθήσεων θα είναι πολύ ευκολότερα υλοποιήσιμη και αποδοτικότερη σε σχέση με τις μεθόδους που δρουν πάνω σε σύνολα πεποιθήσεων. Μία σημαντική παρατήρηση εδώ είναι ότι εν γένει υπάρχουν περισσότερες της μίας βάσεις για ένα δεδομένο σύνολο πεποιθήσεων και η επιλογή της κατάλληλης βάσης αποτελεί έναν επιπλέον αντικείμενο άξιο μελέτης.

Είναι κοινά αποδεκτό ότι αν θέλουμε να δημιουργήσουμε ένα σύστημα το οποίο θα διαχειρίζεται βάσεις γνώσης σε αποδεκτό χρόνο, είμαστε υποχρεωμένοι να χρησιμοποιήσουμε βάσεις πεποιθήσεων και όχι σύνολα πεποιθήσεων. Τα σύνολα πεποιθήσεων είναι ενδιαφέροντα θεωρητικά κατασκευάσματα, χρήσιμα σε ορισμένες περιπτώσεις σαν ενδεικτικά της ιδανικής κατάστασης, αλλά μικρής αξίας από πρακτικής απόψεως.

Έτσι, τοποθετώντας τέτοιες προτάσεις στη βάση πεποιθήσεων τις διαχωρίζουμε αυτόματα από όλες τις υπόλοιπες, παραγόμενες πεποιθήσεις, κάτι που ενδεχομένως και να είναι επιθυμητό. Στην ίδια εργασία προτείνεται ένας αλγόριθμος ενημέρωσης βάσεων πεποιθήσεων και μελετάται η σχέση του με τις συναρτήσεις πλήρους και μερικής τομής των AGM καθώς και τα αξιώματα τα οποία ικανοποιεί.

Ο Gärdenfors, παρόλο που υποστηρίζει τα σύνολα πεποιθήσεων, παρατηρεί ότι οι βάσεις πεποιθήσεων προσφέρουν μεγαλύτερη εκφραστική ικανότητα, αφού η ίδια γνώση (σύνολο πεποιθήσεων) μπορεί να προέλθει από περισσότερες της μίας βάσεις πεποιθήσεων. Παρόλα αυτά, χρησιμοποιεί την έννοια της γνωσιακής στήριξης (epistemic entrenchment), με την βοήθεια της οποίας τα σύνολα πεποιθήσεων αποκτούν την ίδια εκφραστική δύναμη με τις βάσεις πεποιθήσεων. Η γνωσιακή στήριξη είναι ένα είδος καθοριζόμενης από τον χρήστη διάταξης μεταξύ πεποιθήσεων. Η διάταξη αυτή χρησιμοποιείται όταν πρέπει να επιλέξουμε κάποια πεποίθηση προς απόρριψη (σε αντιφατικές ενημερώσεις), και όπως αποδεικνύει ο Gärdenfors, μας επιτρέπει να εκφράσουμε εμμέσως σχέσεις αιτιότητας.

Όπως έχουμε ήδη τονίσει, στο πρόβλημα της αναθεώρησης γνώσης μας ενδιαφέρει σε μεγάλο βαθμό να ενεργούμε σε συμφωνία με την ανθρώπινη διαίσθηση. Επομένως, πολύ μεγάλο ρόλο στην επιλογή του βέλτιστου τρόπου αποθήκευσης γνώσης παίζει και η αντίστοιχη (αλλά όχι ισοδύναμη) φιλοσοφική αντιπαράθεση ως προς τον τρόπο αποθήκευσης γνώσης στον ανθρώπινο εγκέφαλο. Στην φιλοσοφία υπάρχουν δύο θεωρίες ως προς αυτό, τις οποίες αναλύει με εξαιρετικό τρόπο ο Gärdenfors.

Η θεωρία θεμελιωδών πεποιθήσεων (foundations theory) βασίζεται στην αρχή ότι μαζί με την γνώση πρέπει να αποθηκεύονται και οι αιτίες της γνώσης. Υπάρχουν κάποιες πεποιθήσεις οι οποίες ονομάζονται θεμελιώδεις

(foundational). Οι θεμελιώδεις πεποιθήσεις είναι αυτόνομες και ισχύουν χωρίς να υπάρχει ανάγκη αιτιολόγησής τους, αλλά όλες οι υπόλοιπες οφείλουν να υποστηρίζονται (δηλαδή να έπονται) άμεσα ή έμμεσα από αυτές.

Σε αντίθεση με την θεωρία θεμελιωδών πεποιθήσεων, η θεωρία συναφών πεποιθήσεων (coherence theory), βασίζεται στην αρχή ότι κάθε πεποίθηση είναι αυτόνομη, ανεξάρτητη και αδιαμφισβήτητη. Δεν υπάρχει ανάγκη αιτιολόγησης μίας πεποίθησης και κάθε κομμάτι γνώσης είναι δεκτό από μόνο του, αρκεί να μην αντιφάσκει με την υπόλοιπη γνώση.

Δυστυχώς, καμία από τις δύο θεωρίες δεν έχει αποδειχθεί. Έγιναν πειράματα τα οποία δείχνουν ότι από την στιγμή που μία πεποίθηση γίνει αποδεκτή από κάποιον, αυτός συνεχίζει να την πιστεύει, ακόμα και αν η αιτία η οποία τον ώθησε αρχικά να την κάνει αποδεκτή παύει να υφίσταται. Αυτό υπονοεί ότι η θεωρία συναφών πεποιθήσεων είναι σωστότερη. Οι οπαδοί της θεωρίας θεμελιωδών πεποιθήσεων όμως σχολιάζουν ότι τα πειράματα αποδεικνύουν απλά την πρακτική δυσκολία που υπάρχει στην αποθήκευση τόσων πεποιθήσεων και των αιτιών τους (κάτι το οποίο ισχύει και στους ηλεκτρονικούς υπολογιστές). Έτσι – κατά τους υποστηρικτές της θεωρίας θεμελιωδών πεποιθήσεων – αν μία πεποίθηση A γίνει δεκτή αποκλειστικά λόγω της ύπαρξης μίας άλλης πεποίθησης B, και εκ των υστέρων πάψουμε να πιστεύουμε την B, πρέπει να απορρίψουμε και την A. Το ότι διατηρείται και η A είναι λάθος, γι' αυτό άλλωστε και αν υποδειχθεί στο άτομο υπό εξέταση ότι ο μοναδικός λόγος για τον οποίο πίστευε το A ήταν το B, τότε το άτομο αυτό είναι πρόθυμο να παραιτηθεί και από την πεποίθηση A. Αυτό άλλωστε συμφωνεί και με την κοινή λογική. Επομένως, το αποτέλεσμα των πειραμάτων αιτιολογείται με το επιχείρημα ότι τα εξεταζόμενα άτομα δεν αγνοούν αλλά ξεχνούν τις σχέσεις αιτιότητας.

Με την πρώτη ματιά, θα λέγαμε ότι η θεωρία θεμελιωδών πεποιθήσεων μας παροτρύνει να χρησιμοποιήσουμε βάσεις πεποιθήσεων, ενώ η θεωρία συναφών πεποιθήσεων υπονοεί ότι τα σύνολα πεποιθήσεων είναι προτιμότερα. Αυτό όμως δεν είναι απόλυτα ακριβές, μιας και οι δύο διαχωρισμοί είναι σε μεγάλο βαθμό ανεξάρτητοι.

Η θεωρία θεμελιωδών πεποιθήσεων μας υποχρεώνει να κρατάμε τις αιτίες για τις πεποιθήσεις μας, ωστόσο δεν μας περιορίζει ως προς το ποιές πεποιθήσεις

θα αποθηκεύονται ρητώς και ποιές θα προκύπτουν από τις αποθηκευμένες. Τόσο οι παραγόμενες πεποιθήσεις, όσο και οι αιτίες τους μπορούν να προκύπτουν από τις αποθηκευμένες λογικές προτάσεις. Έτσι, έχουμε την ελευθερία να χρησιμοποιήσουμε είτε βάσεις είτε σύνολα πεποιθήσεων, αρκεί να διατηρούνται οι σχέσεις αιτιότητας. Από την άλλη, η θεωρία συναφών πεποιθήσεων δεν απαιτεί την αποθήκευση σχέσεων αιτιότητας μεταξύ κομματιών γνώσης. Και πάλι όμως μπορούμε να αποθηκεύουμε ρητώς είτε όλες, είτε ένα μέρος των πεποιθήσεών μας, αρκεί να υπάρχει τρόπος το μη αποθηκευμένο κομμάτι γνώσης να προέλθει από το αποθηκευμένο. Οι σχέσεις εξάρτησης που έμμεσα δημιουργούνται κατά την διαδικασία ανάκτησης της πλήρους γνώσης μπορούν απλά να αγνοηθούν κατά τις ενημερώσεις.

Για παράδειγμα, ο Dalal χρησιμοποιεί μία μόνο λογική πρόταση σαν βάση γνώσης η οποία στην πράξη είναι η σύζευξη όλων των προτάσεων του συνόλου πεποιθήσεων που αναπαριστά την γνώση μας. Μπορεί να αποδειχθεί ότι η σύζευξη όλων των προτάσεων ενός συνόλου πεποιθήσεων K ανήκει στο K και έχει ως συνέπειά της όλες τις προτάσεις του K . Επομένως, ο Dalal επέλεξε μία από τις βάσεις πεποιθήσεων της γνώσης, πάνω στην οποία εφαρμόζει τον αλγόριθμό του. Όμως δεν χρησιμοποιείται πουθενά καμία σχέση αιτιότητας, άρα λοιπόν ο Dalal χρησιμοποιεί την θεωρία συναφών πεποιθήσεων.

Πάντως, σε ένα σύστημα το οποίο χρησιμοποιεί βάσεις γνώσεων, όπως το παραπάνω, η επιλογή της θεωρίας παίζει σημαντικό ρόλο ως προς την τρόπο επιλογής της βάσης πεποιθήσεων. Δύο διαφορετικές βάσεις του ίδιου συνόλου πεποιθήσεων θα δώσουν διαφορετικές αιτίες για τις ίδιες πεποιθήσεις, άρα λοιπόν χρειάζεται προσοχή ως προς την επιλογή της βάσης αν δεχτούμε την θεωρία θεμελιωδών πεποιθήσεων. Στην θεωρία συναφών πεποιθήσεων αυτό δεν συμβαίνει, διότι οι σχέσεις αιτιότητας αγνοούνται. Έτσι ο Dalal προτίμησε την απλούστερη δυνατή βάση πεποιθήσεων για τον αλγόριθμό του, ενώ οι υποστηρικτές της θεωρίας θεμελιωδών πεποιθήσεων θα προτιμούσαν σαν βάση πεποιθήσεων ένα σύνολο από στοιχειώδεις πεποιθήσεις, οι οποίες θα μπορούσαν να αντιπροσωπεύουν παρατηρήσεις σχετικές με τον υπό μοντελοποίηση κόσμο και θα έδιναν τις «σωστές» σχέσεις αιτιότητας μεταξύ των πεποιθήσεων της βάσης.

2.3 Αποδοχή και Απόρριψη Πεποιθήσεων

Είδαμε ότι κατά την ενημέρωση μίας βάσης με κάποια αντιφατική ενημέρωση, είμαστε υποχρεωμένοι να απορρίψουμε κάποιο κομμάτι από την γνώση μας. Η επιλογή της γνώσης προς απόρριψη μπορεί να γίνει με διάφορους τρόπους και έχει δύο σκέλη που αφορούν πρώτον την αποδοχή (ή μη) της ενημέρωσης και δεύτερον, εφόσον η ενημέρωση γίνει αποδεκτή, την επιλογή της πληροφορίας από την βάση που θα επιλεγεί προς απόρριψη.

Ένας από τους τρόπους, όπου το σύστημα ενημέρωσης γνώσης θεωρείται ένας αυτόνομος πράκτορας (agent), ο οποίος κρίνει σε κάθε ενημέρωση αν θα την δεχτεί ή όχι, τι πληροφορία θα απορρίψει από την βάση κοκ. Τα κριτήρια επιλογής είναι η αξιοπιστία της πηγής πληροφορίας, οι πεποιθήσεις και οι στόχοι του πράκτορα, οι σχέσεις μεταξύ των πεποιθήσεων αυτών και η απώλεια πληροφορίας ανά περίπτωση.

Ένας συνηθέστερος τρόπος επιλογής είναι με την χρήση της αρχής προτεραιότητας ενημέρωσης του Dalal, η οποία χρησιμοποιείται και στις μεθόδους που ικανοποιούν τα αξιώματα AGM. Η αρχή αυτή λέει ότι η ενημέρωση είναι αδιαμφισβήτητη άρα λοιπόν η γνώση που θα απορριφθεί πρέπει με κάποιο τρόπο να επιλεγεί από την βάση. Η επιλογή της γνώσης που θα απορριφθεί από την βάση γίνεται με γνώμονα την αρχή διατήρησης παλιάς πληροφορίας του Dalal ή το ισοδύναμο κριτήριο οικονομίας πληροφορίας (criterion of informational economy) του Gärdenfors . Η αρχή αυτή ορίζει ότι η απώλεια πληροφορίας οφείλει να είναι η ελάχιστη δυνατή.

Η δυσκολία που προκύπτει με την παραπάνω προσέγγιση είναι η μέτρηση της απώλειας πληροφορίας, πρόβλημα που αντιμετωπίζεται με διαφορετικούς τρόπους στην βιβλιογραφία. Ο Dalal προτείνει μία μέθοδο βασισμένη σε λογικές προτάσεις οι οποίες «εκφράζουν» την χαμένη πληροφορία. Ο Gärdenfors προτείνει την πιο διαδεδομένη τεχνική της γνωσιακής στήριξης.

2.4 Ιδιότητες των Προτασιακών Βάσεων Γνώσεων

Πάνω στις βασικές αυτές αρχές του Προτασιακού Λογισμού έχει χτιστεί μια ολόκληρη θεωρία της οποίας η ανάπτυξη είναι εκτός των στόχων της εργασίας αυτής. Το σημαντικό στοιχείο εδώ είναι οι δυνατότητες των λογικών εκφράσεων

οποιοδήποτε είδους λογικής στην αναπαράσταση δεδομένων (facts) και γνώσης. Είναι γνωστό ότι η λογική, ως τρόπος αναπαράστασης γνώσης, έχει πολύ μεγάλη εκφραστική δύναμη. Η πρωτοβάθμια λογική μπορεί πρακτικά να εκφράσει οποιαδήποτε μορφή γνώσης. Αυτό ισχύει – σε μικρότερο βαθμό όμως – και για την προτασιακή λογική. Στην προτασιακή λογική μπορούμε να εκφράσουμε απλή γνώση (facts, για παράδειγμα: A , $A \wedge B$, $\neg A \vee \neg B$), περιορισμούς ακεραιότητας (integrity constraints, για παράδειγμα: $\neg(A \wedge B)$), ενεργούς κανόνες (triggers, για παράδειγμα: $A \rightarrow B$), σχέσεις συνέπειας ή ισοδυναμίας (για παράδειγμα: $A \rightarrow B$, $A \leftrightarrow B$ κοκ), όπως επίσης και ακολουθίες συλλογισμών προς εξαγωγή συμπερασμάτων (συστήματα αποδείξεων του Προτασιακού Λογισμού). Από την άλλη, υπάρχουν σχέσεις που δεν μπορούν να εκφραστούν με τον Προτασιακό Λογισμό, αλλά χρειάζονται ισχυρότερα εκφραστικά μοντέλα όπως είναι η πρωτοβάθμια λογική. Τέτοιες σχέσεις είναι για παράδειγμα οι σχέσεις στιγμιοτύπου (instance of), τα γνωρίσματα οντοτήτων (entity attributes) και οι σχέσεις υποσυνόλου (IsA). Η μικρότερη εκφραστικότητα των προτασιακών βάσεων γνώσεων αποζημιώνεται από το γεγονός ότι σαν μορφή λογικής ο προτασιακός λογισμός είναι πολύ απλούστερος του πρωτοβάθμιου. Αυτό ισχύει τόσο σε θεωρητικό επίπεδο (θεωρήματα, αποδείξεις, απλότητα έκφρασης), όσο και σε πρακτικό (απλότητα διαχείρισης συμβόλων, πολυπλοκότητα αλγορίθμων).

Σε κάποια από τα γνωστά μοντέλα, όπως στο μοντέλο οντοτήτων-συσχετίσεων, κάθε πληροφορία απαιτείται να αποθηκεύεται ρητώς στην βάση. Έτσι, ότι υπάρχει στην βάση γνώσης μας ισχύει (θεωρείται αληθές), ενώ ότι δεν υπάρχει δεν ισχύει (θεωρείται ψευδές). Αυτό σημαίνει ότι πρέπει να αποθηκεύεται ρητώς κάθε γνωστή πληροφορία, καθώς και οι ενδιαφέρουσες συνέπειες αυτών. Σε άλλα, για παράδειγμα στη γλώσσα Telos, δίνεται η δυνατότητα εξαγωγής συμπερασμάτων (παραγόμενη γνώση).

Στις προτασιακές βάσεις γνώσης (και γενικότερα στις λογικές βάσεις γνώσης) υπάρχει η έννοια της επαγόμενης πληροφορίας. Για παράδειγμα οι εκφράσεις « A » και « $A \rightarrow B$ » είναι γνωστό ότι μας δίνουν ως συνέπεια την έκφραση « B » (modus ponens). Έτσι, αν με βάση κάποια παρατήρηση ισχύει η λογική πρόταση « A », ενώ με βάση μία άλλη παρατήρηση όποτε ισχύει το A ισχύει και το B (« $A \rightarrow B$ »), τότε ο συνδυασμός των δύο αυτών παρατηρήσεων μας δίνει την

έκφραση «B». Αν εκ των υστέρων μία νέα παρατήρηση μας αναιρέσει την γνώση «A», τότε αυτόματα αναιρείται και η γνώση «B», αφού πλέον δεν έχουμε κανένα λόγο να πιστεύουμε το «B».

Η παραπάνω παρατήρηση είναι πολύ σημαντική για δύο λόγους. Πρώτον, διότι δείχνει την συμπάγεια της εκφραστικότητας των προτασιακών βάσεων γνώσεων, αφού μας αρκεί η αποθήκευση μόνο ενός μέρους της γνωστής πληροφορίας και όχι της πλήρους γνώσης μας για τον κόσμο, και δεύτερον, διότι μας δείχνει την έννοια της εξάρτησης προτάσεων από άλλες προτάσεις, ή ισοδυνάμως, της εξάρτησης συγκεκριμένων κομματιών της γνώσης μας από άλλη γνώση.

Η πιο σημαντική ιδιότητα των Προτασιακών βάσεων γνώσεων όμως είναι η δυνατότητα εξαγωγής συμπερασμάτων. Αποτελεί μία πολύ χρήσιμη διαδικασία για πολλές εφαρμογές όπως είναι η εξόρυξη πληροφορίας (data mining), τα συστήματα λήψης αποφάσεων, τα συστήματα μηχανικής μάθησης, στην επεξεργασία επερωτήσεων, στα συστήματα λογισμού και αποδείξεων κα. Σε άλλου τύπου βάσεις πληροφορίας, είναι αναπόφευκτο ένα ενδιάμεσο βήμα μετατροπής της γνώσης σε λογικές εκφράσεις, προκειμένου η πληροφορία να χρησιμοποιηθεί για τις εφαρμογές αυτές. Εξάλλου, η θεωρητική θεμελίωση όλων των μοντέλων αναπαράστασης πληροφορίας βασίζεται με κάποιο τρόπο σε κάποιο είδος λογικής, κάτι το οποίο αποδεικνύει την ισχύ των λογικών βάσεων γνώσεων.

Οι παραπάνω παρατηρήσεις δείχνουν ότι η αναπαράσταση γνώσης με λογικές προτάσεις παρέχει αδιαμφισβήτητα πλεονεκτήματα στον τομέα της εκφραστικότητας. Η εισαγωγή των λογικών προτάσεων όμως δημιούργησε νέα προβλήματα στον τομέα της αποθήκευσης, ανάκτησης και διαχείρισης γνώσης. Ένα από τα βασικότερα τέτοια προβλήματα ήταν η ενημέρωση και συντήρηση της υπάρχουσας γνώσης, πρόβλημα με το οποίο θα ασχοληθούμε στην παρούσα εργασία.

2.5 Γνωσιακά σύνολα

1) Γνωσιολογικές καταστάσεις και σύνολα πεπιοιθήσεων

Οι γνωσιολογικές καταστάσεις είναι οι κεντρικές οντότητες των επιστημολογικών θεωριών. Χρησιμοποιούνται για να αναπαραστήσουν μία πραγματική ή μία πιθανή γνωσιακή κατάσταση ενός ατόμου για συγκεκριμένη χρονική στιγμή. Η ιδέα μιας γνωσιολογικής κατάστασης μιας πεποίθησης δεν είναι η έκφραση της αναπαράστασης των πεποιθήσεων στο μυαλό μας αλλά μία επιστημονολογική ιδέα, η οποία εξαδινεικεύει μία ψυχολογική ιδέα. Όταν μιλάμε για αλλαγή υποθέτουμε ότι υπάρχει κάτι, το οποίο αλλάζει. Αυτό το κάτι, το οποίο αλλάζει είναι η γνωσιολογική κατάσταση. Μπορούμε να αναπτύξουμε πολλές κατασκευές, οι οποίες θα λειτουργήσουν ως μοντέλα των γνωσιολογικών καταστάσεων ή απλά μπορούμε να τις χειριστούμε σαν «μαύρα κουτιά». Δηλαδή όταν δεχόμαστε συγκεκριμένες εισόδους να παίρνουμε συγκεκριμένες εξόδους, χωρίς να μας απασχολεί ο εσωτερικός μηχανισμός, με τον οποίο γίνεται. Το πιο χαρακτηριστικό μιας γνωσιολογικής κατάστασης είναι οι λογικές προτάσεις από τις οποίες αποτελείται. Το σύνολο όλων αυτών των λογικών προτάσεων ονομάζεται σύνολο πεποιθήσεων (belief set), που είναι συνεπείς και κλειστές κάτω από λογικές συνέπειες. Για να είμαστε πιο ακριβείς, μία πρόταση ϕ ανήκει στο σύνολο πεποιθήσεων εάν και μόνο αν η απάντηση από το σύστημα στην ερώτηση εάν ισχύει η ϕ είναι ναι. Κάθε γνωσιολογική κατάσταση έχει ένα σύνολο πεποιθήσεων. Με άλλα λόγια υπάρχει μία λειτουργία, που ονομάζεται λειτουργία υποστήριξης, η οποία αναθέτει σε κάθε γνωσιολογική κατάσταση ένα σύνολο πεποιθήσεων.

Τέλος, βάση μιας πεποίθησης ή ενός συνόλου πεποιθήσεων είναι το σύνολο όλων των λογικών προτάσεων των πεποιθήσεων. Επομένως, θα ισχύει ότι κάθε σύνολο προτάσεων A αποτελεί μία βάση πεποιθήσεων (belief base). Έστω K είναι ένα σύνολο πεποιθήσεων, τότε το σύνολο των προτάσεων A είναι μία βάση πεποιθήσεων για το K εάν και μόνο αν ισχύει: $K = Cn(A)$.

Κεφάλαιο 3

3.1 Το AGM παράδειγμα

Στο AGM παράδειγμα, οι πεποιθήσεις αναπαριστούνται ως προτάσεις της γλώσσας L και τα σύνολα των πεποιθήσεων ως θεωρίες της γλώσσας L . Εδώ πρέπει να σημειωθεί ότι η αναπαράσταση ενός συνόλου πεποιθήσεων ως μία θεωρία, υπονοεί ότι οι πράκτορες είναι παντογνώστες, δηλαδή γνωρίζουν τα πάντα, μια πολύ δυνατή υπόθεση και μία από τις αδυναμίες του AGM παραδείγματος. Η διαδικασία της αναθεώρησης πεποιθήσεων σαν μία συνάρτηση $*$ χαρτογραφώντας μία θεωρία K και μία πρόταση ϕ σε μία καινούρια θεωρία $K*\phi$. Φυσικά πολλοί περιορισμοί πρέπει να γίνουν προκειμένου να αποτυπωθεί σωστά η ιδέα της λογικής αναθεώρησης πεποιθήσεων. Ένας τέτοιος περιορισμός είναι και η αρχή της ελάχιστης αλλαγής (principle of minimal change), σύμφωνα με την οποία ένας λογικός πράκτορας οφείλει να αλλάζει τις πεποιθήσεις του κάνοντας όσο το δυνατό λιγότερες αλλαγές προκειμένου να αφομοιώσει την νέα πληροφορία. Φυσικά με μια πρώτη ματιά, δεν είναι ξεκάθαρο πως κανείς μπορεί να μετρήσει την αλλαγή μεταξύ του συνόλου των πεποιθήσεων ή ακόμα πως η ιδέα της ελάχιστης αλλαγής μπορεί να εκφραστεί με καθαρά λογικό τρόπο.

Ένα σύνολο πεποιθήσεων είναι ομολογουμένως ένας αυστηρός τρόπος αναπαράστασης της κατάστασης των πεποιθήσεών μας. Ωστόσο, ακόμα και σε αυτό το επίπεδο πολυπλοκότητας έχουν εμφανιστεί αρκετά και σημαντικά επιστημολογικά προβλήματα, και για αυτό θα χρησιμοποιήσουμε όσο το δυνατόν πιο απλές μεθόδους αναπαράστασης αυτών των πεποιθήσεων.

Έστω λοιπόν ότι το K είναι ένα συνεπές σύνολο πεποιθήσεων. Για για κάθε πρόταση ϕ υπάρχουν μόνο τρεις διαφορετικές γνωστικές καταστάσεις, οι οποίες εκφράζονται ως εξής:

1. $\phi \in K$: η πρόταση ϕ γίνεται αποδεκτή
2. $\neg\phi \in K$: η πρόταση ϕ απορρίπτεται
3. $\phi \notin K$ και $\neg\phi \notin K$: η πρόταση ϕ είναι απροσδιόριστη

Συνεπώς τα σύνολα πεποιθήσεων χρησιμοποιούνται ως αναπαράσταση των γνωστικών καταστάσεων, ενώ οι τρεις τύποι των γνωστικών αλλαγών ως ομαδοποίηση και έκφραση τέτοιων καταστάσεων. Προκειμένου να καθορίσουμε την επιστημονική διατύπωση για ένα σύστημα πεποιθήσεων βασισμένο σε σύνολα πεποιθήσεων, χρειαζόμαστε ειδικούς κανόνες που θα υποδεικνύουν την ορθότητα των διαδικασιών της πρόσθεσης, της αφαίρεσης και της αναθεώρησης. Στην πρόσθεση πεποιθήσεων προσθέτουμε απλά μία καινούρια πεποίθηση χωρίς να αφαιρούμε τίποτα από το αρχικό σύνολο πεποιθήσεων. Στην αφαίρεση πεποιθήσεων απλά αφαιρούμε μια πεποίθηση από το αρχικό σύνολο πεποιθήσεων. Εδώ πρέπει να σημειώσουμε, ότι όσο η πρόσθεση τόσο και η αφαίρεση πεποιθήσεων είναι συνεπής αλλαγές της κατάστασης των πεποιθήσεων, γιατί και στην πρόσθεση και στην αφαίρεση καμία πεποίθηση δεν έρχεται σε αντίθεση με την αρχική κατάσταση των πεποιθήσεων ενός πράκτορα. Αντίθετα, στην αναθεώρηση των πεποιθήσεων μπορούμε να δεχτούμε την άρνηση μιας πεποίθησης, την οποία είχαμε αποδεχτεί αρχικά. Επομένως η **αναθεώρηση πεποιθήσεων** δεν είναι συνεπής αλλαγή. Ένα χαρακτηριστικό παράδειγμα [2] σελ. 48, για να κατανοήσουμε την **αναθεώρηση πεποιθήσεων** είναι όταν π.χ. είσαι έτοιμος να πληρώσεις το λογαριασμό σε ένα εστιατόριο και προς μεγάλη σου έκπληξη δεν υπάρχουν καθόλου χρήματα στο πορτοφόλι σου, ενώ εσύ ήσουν απόλυτα σίγουρος ότι είχες αρκετά χρήματα για να πληρώσεις τον λογαριασμό.

Ο κύριος στόχος σε αυτό το κεφάλαιο είναι να παρουσιαστεί η θεμελίωση και ανάλυση ποικίλων ορθολογικών αξιωμάτων για κάθε διαδικασία αλλαγής πεποιθήσεων και στο επόμενο κεφάλαιο θα γίνει η παρουσίαση κατασκευαστικών μοντέλων για αυτούς τους τύπους των αλλαγών.

3.2 Πρόσθεση Πεποιθήσεων

Η πρόσθεση είναι ο πιο απλός τρόπος αναπαράστασης της επιστημολογικής αλλαγής όπου “μαθαίνουμε κάτι καινούργιο”. Η πιο συνηθισμένη αιτία τέτοιων αλλαγών είναι οι παρατηρήσεις καθώς και η πληροφόρηση που παρέχεται από άλλους ανθρώπους μέσω ποικίλων πληροφοριακών πηγών.

Εάν το K είναι το αρχικό σύνολο πεποιθήσεων και ϕ μία νέα πληροφορία, τότε η

πρόσθεση του K με το ϕ συμβολίζεται ως $K+\phi$. Ο τελεστής $+$ αναπαριστά τη συνάρτηση που αντιστοιχεί ζεύγη συνόλων πεποιθήσεων και προτάσεων, δηλαδή το $KL \times L \rightarrow KL$.

Η πρόταση αυτή εκφράζεται από το παρακάτω αίτημα:

(K+1) $K+\phi$ είναι ένα σύνολο πεποιθήσεων

Επειδή τα σύνολα πεποιθήσεων χρησιμοποιούνται ως αναπαραστάσεις των καταστάσεων των πεποιθήσεων, πολλές φορές κρίνεται σκόπιμο να μην ασχολούμαστε με επιπλέον λεπτομέρειες που αφορούν τις νέες πληροφορίες. Αυτό γιατί οι νέες πληροφορίες εντοπίζονται κυρίως από τις αλλαγές που οι ίδιες προκαλούν σε ένα σύνολο πεποιθήσεων.

Παραδείγματος χάριν μία αλλαγή μπορεί να περιγραφεί από έναν περιορισμό του τύπου: η πρόταση ϕ έχει ήδη γίνει δεκτή στο νέο σύνολο πεποιθήσεων K . Ο περιορισμός αυτός τυποποιείται σε αίτημα ως εξής:

(K+2) $\phi \in K+\phi$

Τα επόμενα αξιώματα που αφορούν τη διαδικασία της πρόσθεσης θεμελιώνονται από μία διαφορετική ορθολογική πλευρά. Η βασική ιδέα κλειδί είναι ότι, όταν αλλάζουμε τις πεποιθήσεις μας, επιθυμούμε να διατηρήσουμε όσο το δυνατόν περισσότερες από τις παλιές μας πεποιθήσεις. Η πληροφορία γενικά δεν περισσεύει, και η άσκοπη απώλειά της θα πρέπει πάση θυσία να αποφεύγεται. Αυτή η ιδέα αποτελεί ίσως το πιο σημαντικό κριτήριο στην αλλαγή των πεποιθήσεων, το οποίο είναι γνωστό και ως κριτήριο 'άρχή της ελάχιστης αλλαγής' ή 'πληροφοριακής οικονομίας'.

Εάν το ϕ είναι απροσδιόριστο στο K (ή εάν το ϕ γίνεται δεκτό), τότε το $\neg\phi$ απορρίπτεται. Επομένως το ϕ δεν προκαλεί αντίφαση με καμία από τις πεποιθήσεις που βρίσκονται ήδη στο K . Έτσι στη διαδικασία της πρόσθεσης (expansion) διατηρούνται πάντα όλες οι παλιές πεποιθήσεις. Τέλος, το κριτήριο της πληροφοριακής οικονομίας θεμελιώνεται από την παρακάτω ιδιότητα:

1 Εάν το $\neg\phi \notin K$ τότε $K \subseteq K+\phi$

Ωστόσο, στην περίπτωση όπου το $\neg\phi \in K$, η πρόσθεση της πρότασης ϕ οδηγεί σε ανεπάρκεια. Στο ενδεχόμενο αυτό δηλώνουμε το $K+\phi$ ως το άτοπο σύνολο πεποιθήσεων, το οποίο θα συμβολίζουμε ως $K\perp$. Επειδή λοιπόν το $K\perp$ είναι το υπερσύνολο όλων των συνόλων πεποιθήσεων, η ιδιότητα 1 μπορεί να γενικευτεί στο παρακάτω αίτημα:

(K+3) $K \subseteq K+\phi$

Στην περίπτωση όπου το ϕ ανήκει ήδη στο αρχικό σύνολο πεποιθήσεων K η “επαναπρόσθεσή” του δεν θα πρέπει να έχει καμία επίδραση στην κατάσταση των πεποιθήσεών μας. Η παραπάνω ιδέα διατυπώνεται από το τέταρτο στη σειρά αίτημα:

(K+4) Εάν το $\phi \in K$ τότε $K+\phi = K$

Για την ανάλυση του επόμενου αξιώματος υποθέτουμε ότι η γνώση είναι πληρέστερη στο σύνολο πεποιθήσεων H παρά στο σύνολο K . Στην ορολογία των συνόλων πεποιθήσεων οδηγούμαστε ευθύς αμέσως στη σχέση $K \subseteq H$. Τώρα εάν προσθέσουμε την ίδια πρόταση στα σύνολα K και H αντίστοιχα, συμπεραίνουμε ότι το $K+\phi$ θα πρέπει να περιέχει λιγότερες πεποιθήσεις από τις πεποιθήσεις του $H+\phi$. Η παραπάνω σκέψη μπορεί να διατυπωθεί από το αίτημα της μονοτονίας:

(K+5) Εάν το $K \subseteq H$ τότε $K+\phi \subseteq H+\phi$

Από αυτά τα πέντε αξιώματα για τη διαδικασία της πρόσθεσης που έχουν ήδη παρουσιαστεί μπορούμε να εξάγουμε πολύτιμες ιδιότητες, όπως φαίνεται παρακάτω:

2 Εάν $\psi \in K+\phi$ τότε το $K+\psi \subseteq K+\phi$

Ενδιαφέρον έχει η περίπτωση όπου το ψ δεν ανήκει στο K αλλά ισχύει ότι $\psi \in K+\phi$, είναι δηλαδή λογική συνέπεια του ϕ στο K . Επομένως εάν στο K προσθέσουμε το ψ και όχι το ϕ , τότε δεν δεχόμαστε τίποτα παραπάνω στο $K+\psi$ που δεν έχουμε δεχτεί στο $K+\phi$. Άρα ισχύει ότι:

3 $K+\phi = K+\psi$ εάν και μόνο αν $\psi \in K+\phi$ και $\phi \in K+\psi$

Εάν λάβουμε υπόψη ότι τα σύνολα πεποιθήσεων είναι κλειστά ως προς τη λογική τους συνεπαγωγή, δηλαδή ότι ικανοποιούν τη συνθήκη (Cn), τότε μπορούμε να εξάγουμε κάποιες επιπλέον ιδιότητες:

4 Εάν $\vdash \phi \leftrightarrow \psi$ τότε $K+\phi = K+\psi$

Με άλλα λόγια το περιεχόμενο του συνόλου $K+\phi$ καθορίζεται από τη λογική ισχύ του ϕ και όχι από τη λογική του διατύπωση. Η επόμενη ιδιότητα είναι γνωστή και ως “αρχή της σύνθεσης”:

5 $(K+\phi)+\psi = K+(\phi \wedge \psi)$

Η ερμηνεία αυτής της αρχής είναι ότι η πρόσθεση της πληροφορίας ϕ και έπειτα της ψ οδηγεί στο ίδιο ακριβώς αποτέλεσμα με την πρόσθεση της $\phi \wedge \psi$. Επειδή το $K+\phi \wedge \psi$ είναι το ίδιο με το $K+(\psi \wedge \phi)$, η εξιδανίκευση της διαδικασίας της πρόσθεσης είναι ότι η σειρά που καταφθάνουν οι πληροφορίες στο σύνολο των

πεποιθήσεών μας είναι αδιάφορη, δηλαδή η διαδικασία της πρόσθεσης μπορεί να χαρακτηριστεί και ως επιμεριστική:

$$6 \quad (K+\phi)+\psi = (K+\psi)+\phi$$

Μία άμεση συνέπεια της ιδιότητας 5 είναι η εξής:

$$7 \quad \text{Εάν } \neg\phi \in K, \text{ τότε } K+\phi = K \perp$$

Σε αυτή τη εξιδεικευμένη περίπτωση πρόσθεσης, το σύνολο $K \perp$ αποτελεί το σύνολο όλων των προτάσεων, δηλαδή ολόκληρη τη γλώσσα, η οποία θα πρέπει να αποφεύγεται πάση θυσία. Βέβαια είναι απίθανο να αποφύγουμε το $K \perp$ με μία απλή πρόσθεση οποιασδήποτε πεποίθησης ϕ , επειδή $(K \perp)+\phi = K \perp$. Από την ιδιότητα 6 προκύπτει ότι, εάν το $\neg\phi$ γίνει αποδεκτό στο K και χρειαστεί να αλλάξουμε το K προκειμένου να δεχτούμε την αντίθετη πεποίθηση, δηλαδή την ϕ , τότε λόγω αντιφάσεων η διαδικασία της πρόσθεσης δεν θα μπορούσε να χρησιμοποιηθεί.

Τα αξιώματα $(K+1)$ - $(K+5)$ δεν αποκλείουν την πιθανότητα το $K+\phi$ να περιλαμβάνει κάποιες πεποιθήσεις, οι οποίες δεν περιλαμβάνονται στο K και δεν έχουν καμία απολύτως σχέση με το ϕ . Επειδή όμως επιθυμούμε να αποφύγουμε την ύπαρξη αδικαιολόγητων πεποιθήσεων, απαιτούμε το $K+\phi$ να μην περιέχει πεποιθήσεις οι οποίες δεν επιβάλλονται από επιπρόσθετα αξιώματα. Τεχνικά η ιδέα αυτή διατυπώνεται σε αίτημα ως εξής:

$(K+6)$ Για όλα τα σύνολα πεποιθήσεων K και για όλες τις προτάσεις ϕ , το $K+\phi$ είναι το μικρότερο σύνολο πεποιθήσεων που ικανοποιεί τα αξιώματα από $(K+1)$ - $(K+5)$.

Με βάση το νέο αίτημα, ενδιαφέρουσα περίπτωση αποτελεί το σύνολο $Cn(K \cup \{\phi\})$, δηλαδή, το σύνολο όλων των λογικών συνεπαγωγών των παλιών πεποιθήσεων μαζί με τη νέα πεποίθηση ϕ . Στην πραγματικότητα, παρουσία των αξιωμάτων $(K+1)$ - $(K+5)$, αποδεικνύεται ότι το $(K+6)$ είναι ισοδύναμο με τη σχέση $K+\phi = Cn(K \cup \{\phi\})$. Εάν όμως

δεν λάβουμε υπόψη το περιορισμό που εισάγει το $(K+6)$, τα $(K+1)$ - $(K+5)$ υπαγορεύουν ακριβώς το αντίστροφο δηλαδή:

$$8 \quad Cn(K \cup \{\phi\}) \wedge K+\phi.$$

Θεώρημα 3.1 (Garnefors, P. [2]): Η συνάρτηση της επέκτασης $+$ ικανοποιεί τα αξιώματα $(K+1)$ - $(K+6)$ εάν και μόνο εάν ισχύει $K+\phi = Cn(K \cup \{\phi\})$.

Το αποτέλεσμα αυτό εκλαμβάνεται και ως θεώρημα αναπαράστασης για την πρόσθεση, επειδή μας παρουσιάζει μία πιο σαφή εικόνα αυτής της διαδικασίας. Ανακεφαλαιώνοντας, το σύνολο των αξιωμάτων για την πρόσθεση των συνόλων των πεποιθήσεων ορίζουν μονοσήμαντα την πρόσθεση του K με το ϕ , ως το σύνολο όλων των λογικών συνεπαγωγών του K μαζί με το ϕ .

Πίνακας 1: Αξιώματα πρόσθεσης πεποιθήσεων

- (K+1): $K+\phi$ είναι ένα σύνολο πεποιθήσεων
 (K+2): $\phi \in K+\phi$
 (K+3): $K \subseteq K+\phi$
 (K+4): Εάν το $\phi \in K$ τότε $K+\phi = K$
 (K+5): Εάν το $K \subseteq H$ τότε $K+\phi \subseteq H+\phi$
 (K+6): Για όλα τα σύνολα πεποιθήσεων K και για όλες τις προτάσεις ϕ , το $K+\phi$ είναι το μικρότερο σύνολο πεποιθήσεων που ικανοποιείται αξιώματα (K+1)-(K+5).

3.3 Αναθεώρηση Πεποιθήσεων

Αναθεώρηση πεποιθήσεων είναι η διαδικασία κατά την οποία μία νέα πληροφορία - πεποίθηση ϕ , έρχεται σε αντίφαση με τις πεποιθήσεις που ήδη βρίσκονται στο σύνολο πεποιθήσεων K . Σε αυτήν την περίπτωση κρίνεται αναγκαία η αναθεώρηση του K προκειμένου να διατηρηθεί η συνέπεια μεταξύ των παλαιών και των νέων πεποιθήσεων. Υπάρχει βέβαια και μία ειδική κατηγορία αναθεωρήσεων, όπου η πρόταση ϕ είναι εξαρχής συνεπής με το σύνολο πεποιθήσεων K . Στο σημείο αυτό η διαδικασία αναθεώρησης μεταπίπτει αυτόματα σε πρόσθεση πεποιθήσεων. Η αναθεώρηση των πεποιθήσεων συγκαταλέγεται στους "μη-μονότονους" τύπους αλλαγών πεποιθήσεων, επειδή, όταν προσθέτουμε μία νέα πεποίθηση σε ένα σύνολο πεποιθήσεων δεν διατηρούνται όλες οι παλιές πεποιθήσεις του συγκεκριμένου συνόλου.

Ενώ οι παραδοσιακές θεωρίες αλλαγών πεποιθήσεων περιγράφουν τύπους ‘‘συνεπών’’ αλλαγών, όπως ο τύπος της πρόσθεσης, η αναθεώρηση είναι ο τύπος της αλλαγής που χρησιμοποιείται πιο συχνά. Ένας από τους πιο βασικούς λόγους είναι ότι αρκετά θέματα και έννοιες που αντιλαμβανόμαστε ως δεδομένα, τελικά δεν είναι. Και αυτό εξαιτίας της τάσης που έχουμε στο να προδικάζουμε καταστάσεις, να εμπιστευόμαστε πολλές πηγές πληροφοριών ή ακόμη και να εξάγουμε λάθος συμπεράσματα. Προκειμένου λοιπόν να κατανοήσουμε ανθρώπινες συλλογιστικές, η μοντελοποίηση της αναθεώρησης των πεποιθήσεων φαίνεται να είναι πολύ πιο σημαντική από αυτή της πρόσθεσης.

Ένα σημαντικό πρόβλημα στην αναθεώρηση των πεποιθήσεων είναι ο προσδιορισμός των πεποιθήσεων που διατηρούνται στο σύνολο καθώς και αυτών που απομακρύνονται. Μία μερική αλλά ταυτόχρονα ορθολογική λύση σε αυτό το πρόβλημα τη δίνουν τα οχτώ αξιώματα, που παρουσιάζονται παρακάτω.

Πίνακας 2: Αξιώματα αναθεώρησης πεποιθήσεων

- | | |
|--------|---|
| (K* 1) | Για κάθε ϕ και κάθε K το $K * \phi$ είναι σύνολο πεποιθήσεων |
| (K* 2) | $\phi \in K * \phi$ (success) |
| (K* 3) | $K * \phi \subseteq K + \phi$ (inclusion) |
| (K* 4) | Εάν $\neg \phi \notin K$, τότε $K + \phi \subseteq K * \phi$ |
| (K* 5) | $K * \phi = K \perp$ εάν και μόνο εάν $\neg \phi$ (consistency) |
| (K* 6) | Εάν $\phi \leftrightarrow \psi$ τότε $K * \phi = K * \psi$ |
| (K* 7) | $K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$ |
| (K* 8) | Εάν $\neg \psi \notin K * \phi$ τότε $(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$ |

Έστω ότι για κάθε σύνολο πεποιθήσεων K και κάθε πρόταση ϕ υπάρχει μία μοναδική αναθεώρηση που αναπαριστά την ελάχιστη αλλαγή. Σε μαθηματικούς όρους αυτό ισοδυναμεί με την υπόθεση ότι υπάρχει μία συνάρτηση αναθεώρησης $*$ από το $K \times L$ στο K , η οποία αντιστοιχεί ένα σύνολο πεποιθήσεων K και μία πρόταση ϕ σε ένα νέο σύνολο πληροφοριών, το $K * \phi$.

Το πρώτο αίτημα απαιτεί ότι το αποτέλεσμα της διαδικασίας της αναθεώρησης είναι ένα σύνολο πεποιθήσεων και μας λέει ότι ο πράκτορας παραμένει παντογνώστης έχοντας αναθεωρήσει τις πεποιθήσεις του.

(K*1) Για κάθε πρόταση ϕ και κάθε σύνολο πεποιθήσεων K το $K*\phi$ είναι ένα σύνολο πεποιθήσεων

Το δεύτερο αίτημα γνωστό ως αίτημα της επιτυχίας (postulate of success) εγγυάται ότι η νέα πληροφορία περιλαμβάνεται στο $K*\phi$:

(K*2) $\phi \in K*\phi$

Το τρίτο και τέταρτο αίτημα δηλώνουν ότι όταν η νέα πληροφορία δεν έρχεται σε αντίθεση με το αρχικό σύνολο πεποιθήσεων K, τότε δεν υπάρχει λόγος να απομακρύνουμε καμία από τις αρχικές μας πεποιθήσεις. Η καινούρια κατάσταση πεποιθήσεων $K*\phi$ θα περιέχει ολόκληρο το σύνολο K, την νέα πληροφορία ϕ και οποιαδήποτε συνέπεια είναι κλειστή ως προς το $K*\phi$ και τίποτα παραπάνω. Ακόμα εκφράζουν την ιδέα της ελάχιστης αλλαγής στην περίπτωση που η νέα πληροφορία είναι συνεπής με το αρχικό σύνολο K.

(K*3) $K*\phi \subseteq K+\phi$

(K*4) Εάν $\neg\phi \notin K$, τότε $K+\phi \subseteq K*\phi$

Το πέμπτο αίτημα λέει ότι ο σκοπός του πράκτορα πρέπει να είναι η συνέπεια με οποιαδήποτε κόστος. Η μόνη περίπτωση που είναι αποδεκτό ο πράκτορας να κάνει λάθος είναι όταν η νέα πληροφορία είναι από μόνη της ασυνεπής.

(K*5) $K*\phi = K\perp$ εάν και μόνο εάν $\neg\phi$

Τα έκτο αίτημα είναι γνωστό και ως αίτημα άσχετο με την σύνταξη (irrelevance of syntax postulate). Δηλώνει ότι η σύνταξη της νέας πληροφορίας δεν έχει κανένα αντίκτυπο στην διαδικασία αναθεώρησης πεποιθήσεων, αλλά αυτό που έχει σημασία είναι το περιεχόμενό της. Επειδή τώρα τα στοιχεία των συνόλων των πεποιθήσεων καθώς και οι εισροές γνώσης που δημιουργούν την αλλαγή των πεποιθήσεων διαμορφώνονται από προτάσεις, η αλλαγή καθορίζεται από το περιεχόμενο της εισροής και όχι από μία ειδικότερη γλωσσική τροποποίηση. Αυτό σημαίνει ότι λογικές ισοδύναμες προτάσεις πρέπει να οδηγούν σε ίδιες ή πανομοιότυπες αλλαγές. Έτσι προκύπτει το έκτο αίτημα για τη διαδικασία της αναθεώρησης:

(K*6) Εάν $\phi \leftrightarrow \psi$ τότε $K*\phi = K*\psi$

Τέλος, το έβδομο και όγδοο αίτημα δηλώνουν ότι :

Έστω ότι έχουμε δύο ξεχωριστές προτάσεις ϕ και ψ , εάν αναθεωρήσουμε το σύνολο πεποιθήσεων K με την ϕ τότε ο πράκτορας θα είναι τυχερός εάν φτάσει στο σημείο που το νέο σύνολο πεποιθήσεων $K * \phi$ είναι συνεπές με την πρόταση ψ , και μπορούμε να παράγουμε την $K * (\phi \wedge \psi)$, αφού το μόνο που χρειάζεται είναι να προσθέσουμε την ψ στην $K * \phi$.

$$(K * 7) \quad K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$$

$$(K * 8) \quad \text{Εάν } \neg\psi \notin K * \phi \text{ τότε } (K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$$

Από τα παραπάνω αξιώματα που παρουσιάστηκαν για τη διαδικασία της αναθεώρησης και τις πρόσθεσης προκύπτουν οι παρακάτω ιδιότητες:

$$9 \quad \text{Εάν } \phi \in K, \text{ τότε } K * \phi = K$$

Εάν δηλαδή το ϕ βρίσκεται ήδη στο K , τότε η ελάχιστη αναγκαία αλλαγή για να περιληφθεί εκ νέου το ϕ στο K , είναι μηδενική.

$$10 \quad K * \phi = K * \psi \text{ εάν και μόνο εάν } \psi \in K * \phi \text{ και } \phi \in K * \psi \text{ (identity criterion)}$$

Αυτό το κριτήριο δίνει ένα ανεξάρτητο κίνητρο για την ερμηνεία του $K * \phi$ ως την ελάχιστη αλλαγή του K : Εάν το ψ γίνει αποδεκτό στο $K * \phi$, τότε η αναγκαία αλλαγή στο K για να συμπεριλάβει το ψ δεν μπορεί να είναι μεγαλύτερη από την αναγκαία αλλαγή στο K για να συμπεριλάβει το ϕ . Ομοίως, εάν το ϕ βρίσκεται στο $K * \psi$ τότε η αναγκαία αλλαγή του K ώστε να περιλάβει το ϕ , δεν μπορεί να είναι μεγαλύτερη από την αναγκαία αλλαγή στο K για να συμπεριλάβει το ψ . Βασική υπόθεση είναι ότι τα $K * \phi$ και $K * \psi$ πρέπει να είναι ισοδύναμα. Μία επιπλέον ιδιότητα του $(K * 7)$ είναι η ακόλουθη:

$$11 \quad K * \phi \cap K * \psi \subseteq K * (\phi \vee \psi)$$

Επίσης για το $(K * 8)$ μπορούμε να δώσουμε μία ισοδύναμη διατύπωση που να περιλαμβάνει διάζευξη αντί για σύζευξη:

$$12 \quad \text{Εάν } \neg\psi \notin K * (\phi \vee \psi), \text{ τότε } K * (\phi \vee \psi) \subseteq K * \psi$$

Μία χρήσιμη αρχή με ονομασία "factoring condition" είναι η επόμενη:

$$13 \quad \text{Εάν } K * (\phi \vee \psi) = K * \psi \text{ ή } K * (\phi \vee \psi) = K * \phi \text{ ή}$$

$$K * (\phi \vee \psi) = (K * \phi) \cap (K * \psi)$$

Αποδεικνύεται μέσω των βασικών αξιωμάτων $(K * 1)$ - $(K * 6)$ ότι το κριτήριο 13 είναι ισοδύναμο με τη σύζευξη των αξιωμάτων $(K * 7)$ και $(K * 8)$. Τέλος εάν προσθέσουμε τα αξιώματα $(K * 7)$ και $(K * 8)$ στη λίστα των βασικών αξιωμάτων $(K * 1)$ - $(K * 6)$ μερικά από αυτά θα είναι περιττά. Τα αξιώματα που παρουσιάσαμε έχουν κίνητρο την ερμηνεία του $K * \phi$ ως την ελάχιστη αλλαγή

στο K προκειμένου να γίνει αποδεκτή η πρόταση ϕ . Οι ιδιότητες αυτών των αξιωμάτων υποστηρίζουν αυτήν την ερμηνεία. Στην ερώτηση για το αν υπάρχουν επιπλέον αξιώματα για την αναθεώρηση που να θεμελιώνονται πάνω στην ερμηνεία της ελάχιστης αλλαγής, η απάντηση είναι ότι τα $(K * 1)$ - $(K * 8)$ είναι τα απολύτως αναγκαία για να καλύψουν τις λογικές ιδιότητες της αναθεώρησης.

Συνοψίζοντας όλα τα παραπάνω, τα δύο βασικά καθήκοντα που εκτελεί μία συνάρτηση αναθεώρησης πεποιθήσεων είναι α) να προσθέτει μία καινούρια πεποίθηση ϕ σε μία βάση πεποιθήσεων και β) να δικαιολογεί ότι τη νέα αναθεωρημένη βάση πεποιθήσεων είναι συνεπής, (εκτός εάν ϕ είναι ασυνεπές). Το πρώτο μπορούμε να το πετύχουμε με την πρόσθεση της πεποίθησης ϕ . Το δεύτερο μπορεί να επιτευχθεί με την απομάκρυνση της άρνησης της πεποίθησης ϕ . Η συνάρτηση της αναθεώρησης των πεποιθήσεων μπορεί λοιπόν να κατασκευαστεί με δύο υπολειτουργίες: προσθήκη του ϕ και απομάκρυνση του $\neg\phi$. Η παραδοχή που συνιθίζεται να γίνεται είναι :

1. Απομάκρυνση του $\neg\phi$ και
2. Προσθήκη του ϕ

Αυτό εκφράζει και η Levi Identity , που θα αναλύσουμε παρακάτω. Όμως η Levi Identity δεν είναι ο μόνος τρόπος που μπορεί να εκφραστούν οι δύο υπολειτουργίες. Αντίστροφα μπορούμε :

1. Προσθήκη του ϕ και
2. Απομάκρυνση του $\neg\phi$

Οι χαρακτηρισμοί «εσωτερική αναθεώρηση» και «εξωτερική αναθεώρηση», δείχνουν ότι στην εσωτερική αναθεώρηση οι δύο υπολειτουργίες πραγματοποιούνται εντός της βάσης πεποιθήσεων, ενώ στην εξωτερική αναθεώρηση πραγματοποιούνται έξω από το αρχικό σύνολο πεποιθήσεων.

Στην εσωτερική αναθεώρηση, όταν μία βάση πεποιθήσεων K αναθεωρηθεί από δύο διαφορετικές πεποιθήσεις ϕ , ψ τότε η υπολειτουργία της αφαίρεσης εφαρμόζεται στο ίδιο σύνολο πεποιθήσεων. Επομένως για να εφαρμόσουμε εσωτερική αναθεώρηση σε ένα σύνολο πεποιθήσεων K , χρειαζόμαστε μόνο μία τοπική λειτουργία της αφαίρεσης πεποιθήσεων μερικής επιλογής. Στην εξωτερική αναθεώρηση, όμως, η υπολειτουργία της αφαίρεσης εφαρμόζεται σε διαφορετικές βάσεις πεποιθήσεων. Οπότε, για την εξωτερική αναθεώρηση χρειαζόμαστε μία συνάρτηση σφαιρικής αφαίρεσης.

Ποια όμως από τα δύο είδη είναι περισσότερο εφαρμόσιμο στην πράξη; Οι γνώμες διίστανται. Για παράδειγμα [3] σελ. 204:

Ο Αντώνης και η Βεατρίκη είναι ένα παντρεμένο ζευγάρι. Είχα την εντύπωση ότι είναι και οι δύο Καθολικοί. Τότε άκουσα την Βεατρίκη να λέει : «Στον γάμο μας, δεν ήταν ποτέ πρόβλημα το ότι ανήκουμε σε διαφορετικές θρησκείες». Όταν άκουσα αυτό, έπαψα να πιστεύω ότι η Βεατρίκη είναι Καθολική, αλλά διατήρησα την πεποίθησή μου ότι ο Αντώνης είναι Καθολικός (αφού τον είχα δει σε Καθολική εκκλησία αρκετές φορές).

Σύμφωνα με την συνάρτηση της εσωτερικής αναθεώρησης: Πριν δεχτώ αυτό που είπε η Βεατρίκη, θα αναθεωρούσα τις πεπιοιθήσεις μου για αυτούς τους δύο, και θα ανακάλυπτα ότι θα ήταν προτιμότερο να εγκαταλείψω την ιδέα ότι η Βεατρίκη είναι καθολική. Αφού έπαψα να πιστεύω σε αυτήν την πεπιοιθηση, θα δεχόμουν τη νέα πληροφορία ότι ο Αντώνης και η Βεατρίκη ανήκουν σε διαφορετικές θρησκείες.

Σύμφωνα με την συνάρτηση της εξωτερικής αναθεώρησης αμέσως θα δεχόμουν αυτό που Βεατρίκη είπε. Κατανοώντας ότι αυτό είναι συνεπές με τις αρχικές μου πεπιοιθήσεις, θα τις αναθεωρούσα με σκόπο η νέα πληροφορία να είναι συνεπής. Αυτό θα το έκανα απομακρύνοντας την πεπιοιθηση ότι η Βεατρίκη είναι Καθολική.

Σε αυτό το παράδειγμα η συνάρτηση της εξωτερικής αναθεώρησης είναι η πιο κατάλληλη. Γενικά όταν είναι φανερό ότι η νέα πληροφορία πρέπει να γίνει αποδεκτή, αλλά λιγότερο φανερό ποιες από τις αρχικές μου πεπιοιθήσεις πρέπει να απομακρύνθουν τότε η εξωτερική αναθεώρηση είναι πιο κόντα στην πραγματική φυσιολογική διαδικασία.

Το επόμενο παράδειγμα είναι διαφορετικής φύσεως [3] σελ. 205:

Έμεινα έκπληκτος όταν ο Πίτερ είπε ότι ο πατέρας του καθηγητή Inglefield είναι ένας κηπουρός. Πάντα πίστευα ότι είχε αριστοκρατικό υπόβαθρο. Ομως, μετά από έναν δυσταγμό, δέχτηκα αυτό που μου είχαν πει.

Στο παραπάνω παράδειγμα, φαίνεται να υπάρχει μία φράση στην οποία ούτε η καινούργια πεπιοιθηση ούτε η άρνηση της γίνεται αποδεκτή. Η εσωτερική αναθεώρηση είναι πιο κόντα από την εξωτερική για αυτού του είδους την διαδικασία, και για αυτό στις περισσότερες περιπτώσεις είναι πολύ δύσκολο να καθοριστεί το είδος της αναθεώρησης που είναι πιο κατάλληλο.

3.4 Αφαίρεση Πεποιθήσεων

Η διαδικασία της αφαίρεσης των πεποιθήσεων εξελίσσεται, όταν απομακρύνουμε από ένα σύνολο πεποιθήσεων μερικές πεποιθήσεις χωρίς όμως να προσθέτουμε νέες. Διαισθητικά η διαδικασία της αφαίρεσης μπορεί να πραγματοποιηθεί σε μία δημόσια συζήτηση όπου η μία πλευρά υποστηρίζει την πεποίθηση ϕ , ενώ η αντίπαλη της, το $\neg\phi$. Προκειμένου να αποφύγουν και οι δύο πλευρές την ερώτηση την οποία θα πρέπει υποστηρίξουν το ϕ ή το $\neg\phi$ αντιστοίχως καθώς και τις υπόλοιπες πεποιθήσεις που οδηγούν-υπονοούν τα ϕ , $\neg\phi$, παύουν να πιστεύουν προσωρινά σε αυτές τις πεποιθήσεις, ενώ η δημόσια συζήτηση είναι δυνατόν να διεξαχθεί με τις εναπομείναντες πεποιθήσεις. Τελικά η διαδικασία της αφαίρεσης δεν έχει άμεση εφαρμογή στην πραγματικότητα αλλά αποτελεί μία τεχνική ή αλλιώς ένα ενδιάμεσο κρίκο για την πραγματοποίηση της διαδικασίας της αναθεώρησης των πεποιθήσεων.

Το κύριο πρόβλημα στην αφαίρεση είναι ότι, απομακρύνουμε μία πεποίθηση ϕ από ένα σύνολο πεποιθήσεων K , μπορεί να υπάρχουν επιπρόσθετες πεποιθήσεις στο K που επιφέρουν το ϕ (ή άλλες πεποιθήσεις που απο κοινού επιφέρουν το ϕ). Συνεπώς εάν θέλουμε να συμμορφωθούμε με το κριτήριο (C_n), δηλαδή να διαφυλάξουμε τα σύνολα των πεποιθήσεων κλειστά ως προς τη λογική συνεπαγωγή, είναι αναγκαίο να απορρίψουμε και τις υπόλοιπες πεποιθήσεις. Για παράδειγμα, εάν το ϕ γίνει δεκτό στο K εξαιτίας της λογικής συνεπαγωγής του από το ψ και το γ , τα οποία ανήκουν και τα δύο στο K , τότε είτε το ψ είτε το γ (ή ακόμη και τα δύο) θα πρέπει να απορριφθούν. Το πρόβλημα είναι στον προσδιορισμό των πεποιθήσεων που θα πρέπει να απορριφθούν και αυτών που θα πρέπει να διατηρηθούν. Επειδή το κριτήριο της πληροφοριακής οικονομίας εφαρμόζεται και στην αφαίρεση, καμία από τις πεποιθήσεις δεν θα πρέπει να απορρίπτεται άσκοπα.

Η αφαίρεση μίας πρότασης ϕ από ένα σύνολο πεποιθήσεων K παριστάνεται από το συμβολισμό $K-\phi$. Υποθέτουμε ότι για κάθε σύνολο πεποιθήσεων και για κάθε πρόταση ϕ υπάρχει μία "μοναδική" συνάρτηση αφαίρεσης που ορίζει το $K-\phi$. Όπως και στην πρόσθεση έτσι και στην αφαίρεση η ακριβής μαθηματική έκφραση είναι η εξής: Υπάρχει συνάρτηση αφαίρεσης $- : KL \times L \rightarrow KL$, που αντιστοιχεί ένα ζεύγος από σύνολα πεποιθήσεων και προτάσεων σε ένα νέο σύνολο πεποιθήσεων $K-\phi$.

Όπως στη διαδικασία της πρόσθεσης έτσι και εδώ, υπάρχουν οχτώ αξιώματα έχουν ως κίνητρο το κριτήριο της πληροφοριακής οικονομίας. Αυτά τα αξιώματα είναι γνωστά και σαν AGM αξιώματα για την αφαίρεση πεποιθήσεων και φαίνονται στον παρακάτω πίνακα 3.

Πίνακας 3: Αξιώματα αφαίρεσης πεποιθήσεων

- (K ÷ 1) Για κάθε ϕ και κάθε K το $K \div \phi$ είναι σύνολο πεποιθήσεων
 (K ÷ 2) $K \div \phi \subseteq K$
 (K ÷ 3) Εάν $\phi \notin K$, τότε $K \div \phi = K$
 (K ÷ 4) Εάν δεν ισχύει το ϕ , τότε $\phi \notin K \div \phi$
 (K ÷ 5) Εάν $\phi \in K$ τότε $K \subseteq (K \div \phi) + \phi$
 (K ÷ 6) Εάν $\vdash \phi \leftrightarrow \psi$ τότε $K \div \phi = K \div \psi$
 (K ÷ 7) $K \div \phi \cap K \div \psi \subseteq K \div (\phi \wedge \psi)$
 (K ÷ 8) Εάν $\phi \notin K \div (\phi \wedge \psi)$, τότε $K(\div \phi \wedge \psi) \subseteq K \div \phi$

Κάθε συνάρτηση που ικανοποιεί τα αξιώματα από το (K ÷ 1) - (K ÷ 8) ονομάζεται AGM συνάρτηση αφαίρεσης (AGM contraction function).

Το πρώτο αίτημα για την αφαίρεση των πεποιθήσεων αποδεικνύεται από μόνο του.

(K ÷ 1) Για κάθε πρόταση ϕ και κάθε σύνολο πεποιθήσεων K, το $K \div \phi$ είναι ένα σύνολο πεποιθήσεων

Επειδή το $K \div \phi$ προέρχεται από το K μετά την απόρριψη κάποιων πεποιθήσεων, απαιτείται να μην προκύπτουν νέες πεποιθήσεις στο νέο μας σύνολο $K \div \phi$, το οποίο θα είναι μικρότερο του αρχικού συνόλου K. Άρα το δεύτερο αίτημα μπορεί να διατυπωθεί ως εξής:

(K ÷ 2) $K \div \phi \subseteq K$

Στην περίπτωση που το ϕ δεν ανήκει στο K, και έχοντας υπόψη το κριτήριο της πληροφοριακής οικονομίας, το αρχικό σύνολο K δεν αλλάζει καθόλου:

(K ÷ 3) Εάν $\phi \notin K$, τότε $K \div \phi = K$

Το τέταρτο αίτημα προσδιορίζει την επιτυχία της αφαίρεσης. Με λίγα λόγια η πρόταση που θα απορριφθεί από το αρχικό σύνολο πεποιθήσεων K δεν θα

πρέπει να αποτελεί λογική συνεπαγωγή των πεποιθήσεων που διατηρούνται στο $K \div \phi$ (εκτός και αν το ϕ είναι λογικά έγκυρο, δηλαδή είναι ταυτολογία):

(K \div 4) Εάν δεν ισχύει το ϕ , τότε $\phi \notin K \div \phi$

Το πέμπτο αίτημα εγγυάται ότι, εάν αφαιρέσουμε μία πρόταση ϕ από ένα σύνολο πεποιθήσεων K και ύστερα την επαναπροσθέσουμε, το τελικό σύνολο πεποιθήσεων θα περιέχει πεποιθήσεις που βρίσκονταν ήδη στο K . Εάν συνυπολογίσουμε και το κριτήριο πληροφοριακής οικονομίας, τότε το $K \div \phi$ θα πρέπει να είναι ένα μεγάλο υποσύνολο του K . Σε αναλογία με το (K+6) θεωρούμε ότι όλες οι πεποιθήσεις στο K ανακτούνται, αφού πρώτα αφαιρεθεί και επαναπροσθεθεί η ίδια πεποίθηση. Η παραπάνω έκφραση διατυπώνεται σε αίτημα ως εξής:

(K \div 5) Εάν $\phi \in K$ τότε $K \subseteq (K \div \phi) + \phi$

Το αίτημα αυτό είναι γνωστό ως *recovery postulate*.

Έχει αποτελέσει στο παρελθόν αντικείμενο συζήτησης και περαιτέρω μελέτης κυρίως όσον αφορά της εγκυρότητά του. Η εγκυρότητα του *recovery postulate* φαίνεται να αμφισβητείται ακόμη πιο έντονα κυρίως στις πιθανοτικές περιπτώσεις. Έστω για παράδειγμα [2] σελ. 62, ότι το σύνολο K αποτελεί τις πεποιθήσεις μας για μία συγκεκριμένη κατάσταση. Έστω ότι η πρόταση ϕ περιγράφει το γεγονός ότι στο πρωινό σήμερα μου έπεσε το τoστ και ψ το γεγονός ότι το τoστ έπεσε με την πλευρά με το βούτυρο προς τα κάτω. (Η πρόταση ψ αποτελεί μία τυχαία υπόθεση). Επειδή το ψ είναι συνέπεια του ϕ (όχι λογική) προκύπτει ότι στο $K \div \phi$ δεν ισχύει η πεποίθηση ψ . Εάν όμως ξεκινήσουμε από το $K \div \phi$ και

φανταστούμε ότι μου έπεσε το τoστ σήμερα στο πρωινό, δηλαδή εάν προσθέσουμε την πρόταση ϕ , τότε η πρόταση ψ δεν αποτελεί σε καμία περίπτωση λογική συνέπεια του ϕ , επειδή και το $\neg\psi$ είναι εξίσου πιθανόν.

Άλλο ένα παράδειγμα που πρώτα απομακρύνεις και μετά το προσθέτεις πάλι χωρίς να χάσεις τίποτα από την αρχική σου πεποίθηση είναι:

Πίστευα πως είχα τα κλειδιά του σπιτιού μου πάνω μου (ϕ). Τότε έβαλα το χέρι μου στην αριστερή μου τσέπη, εκεί που συνήθως τα βάζω και δεν τα βρήκα. Έτσι έπψα να πιστεύω στη πεποίθηση ϕ , αλλά δεν άρχισα να πιστεύω στην πεποίθηση $\neg\phi$. Μισό δευτερόλεπτο αργότερα, βρήκα τα κλειδιά και πίστεψα πάλι στην πεποίθηση ϕ . Με άλλα λόγια προσθέτοντας μια πεποίθηση ϕ κερδίζουμε όσα είχαμε χάσει όταν την αφαιρέσαμε. Για αυτό το αίτημα πέντε

είναι γνωστό ως *recovery postulate*, το οποίο αποτελεί το αμφιλεγόμενο αξίωμα από τα AGM αξιώματα. Ένα παράδειγμα από το οποίο φαίνεται ότι το *recovery postulate* δεν είναι ισχυρό είναι το παρακάτω [3] σελ. 73:

Έστω ότι πιστεύω ότι η Κλεοπάτρα έχει έναν γιο (ϕ) και ότι η Κλεοπάτρα έχει μία κόρη (ψ), και ακόμα πιστεύω ότι η Κλεοπάτρα έχει ένα παιδί ($\phi \vee \psi$). Τότε λαμβάνω την πληροφορία που με κάνει να πιστεύω ότι η Κλεοπάτρα δεν έχει παιδιά. Αμέσως μετά μαθαίνω από αξιόπιστη πηγή ότι η Κλεοπάτρα έχει ένα παιδί. Φαίνεται απόλυτα φυσιολογικό σε εμένα να προσθέσω την πεποίθηση ($\phi \vee \psi$) στο σύνολο των πεποιθήσεών μου χωρίς να επαναεισάγω το ϕ και το ψ . Οφειλόμενος στον προβληματικό χαρακτήρα του πέμπτου αξιώματος, ο Makinson πρότεινε μια πιο ευρύ κατηγορία συναρτήσεων στα σύνολα πεποιθήσεων. Όρισε την συνάρτηση απόσυρσης (*withdrawal*) σαν μια συνάρτηση που ικανοποιεί τα πέντε μόνο AGM αξιώματα και όχι απαραίτητα το *recovery postulate*.

Ας φανταστούμε τώρα τι θα μπορούσε να συμβεί, εάν πρώτα προσθέταμε και μετά αφαιρούσαμε το ϕ από ένα σύνολο πεποιθήσεων K . Θα καταλήγαμε πάλι στο K ; Αυτό δεν θα μπορούσε να ισχύει σε καμία περίπτωση, διότι, εάν το $\neg\phi$ ανήκει στο K , τότε $K+\phi = K \perp$.

Άρα το $(K+\phi) \div \phi$ δεν είναι απαραίτητα ίδιο με το K .

Μία πρώτη ιδιότητα των αξιωμάτων της πρόσθεσης και των αξιωμάτων $(K \div 1)$ - $(K \div 5)$ είναι η:

$$16 \quad K \div \phi = K \cap (K \div \phi) + \neg\phi$$

Παράλληλα με την ιδιότητα 16 εισάγουμε το επόμενο αίτημα για την αφαίρεση των πεποιθήσεων:

$$(K \div 6) \quad \text{Εάν } \vdash \phi \leftrightarrow \psi \text{ τότε } K \div \phi = K \div \psi$$

Τα αξιώματα $(K \div 1)$ - $(K \div 6)$ αποτελούν το βασικό σύνολο των αξιωμάτων για τη διαδικασία της αφαίρεσης των πεποιθήσεων. Εκτός από το βασικό σύνολο, θα παρουσιαστούν δύο επιπλέον αξιώματα για τη διαδικασία της αφαίρεσης τα οποία ευθύνονται και για τη σχέση μεταξύ των συνόλων $K \div \phi$ και $K \div \phi \wedge \psi$.

Το κίνητρο για το πρώτο από τα δύο αξιώματα είναι ότι, όταν απορρίπτεται μία πεποίθηση της μορφής $\phi \wedge \psi$, υπάρχει η δυνατότητα επιλογής μεταξύ της απόρριψης του ϕ και του ψ . Σε γενικές γραμμές θα μπορούσαμε να απορρίψουμε λιγότερες πεποιθήσεις, όταν αφαιρούμε το $\phi \wedge \psi$ από ένα

σύνολο πεποιθήσεων K από την περίπτωση που αφαιρούμε μόνο το ϕ . Η παραπάνω ανάλυση διατυπώνεται ως εξής:

$$17 \quad \text{Εάν } \vdash \psi \rightarrow \phi \text{ τότε } K-\phi \subseteq \subseteq K-\psi$$

Παρόλα αυτά όμως η 17 μπορεί να ισχυροποιηθεί από το ακόλουθο αίτημα :

$$(K \div 7) \quad K \div \phi \cap K \div \psi \subseteq K \div (\phi \wedge \psi)$$

Με άλλα λόγια, οι πεποιθήσεις που ανήκουν και στο $K \div \phi$ και στο $K \div \psi$ ανήκουν και στο $K \div \phi \wedge \psi$. Σε σχέση με το προηγούμενο παράδειγμα, αν η πεποίθηση γ άνηκε και αυτή στο $K \div \phi$, τότε δεν θα άνηκε στο $K \div \psi$ (για τον ίδιο λόγο που δεν θα άνηκε στο $K \div \phi \wedge \psi$). Οπότε αυτό το παράδειγμα δίνει έναν επιπλέον λόγο για την ισχύ του $(K \div 7)$.

Το όγδοο αίτημα για τη διαδικασία της αφαίρεσης των πεποιθήσεων απαιτεί το εξής: Όταν αφαιρούμε μία έκφραση της μορφής $\phi \wedge \psi$ από ένα σύνολο πεποιθήσεων K , είτε το ϕ είτε το ψ είτε και τα δύο θα πρέπει να απορριφθούν. Εάν απορριφθεί το ϕ , αυτό θα γίνει επειδή είναι λιγότερο κατοχυρωμένο από το ψ . Σε αυτή την περίπτωση η ελάχιστη αλλαγή που γίνεται στο K προκειμένου να απορριφθούν τα $\phi \wedge \psi$ είναι στενά συνδεδεμένη με την αναγκαία ελάχιστη αλλαγή που απαιτείται για να απορριφθεί μόνο το ϕ . Αυτό αποτελεί και το κίνητρο του όγδου αξιώματος:

$$(K \div 8) \quad \text{Εάν } \phi \notin K \div (\phi \wedge \psi), \text{ τότε } K(\div \phi \wedge \psi) \subseteq K \div \phi$$

Άμεση συνέπεια του $(K \div 8)$ και του $(K \div 4)$ είναι η επόμενη αρχή, η οποία εξασφαλίζει ότι το $K \div \phi \wedge \psi$ "καλύπτεται" είτε από το $K \div \phi$ είτε από το $K \div \psi$:

$$18 \quad \text{Είτε } K \div \phi \wedge \psi \subseteq K \div \phi \text{ είτε } K \div \phi \wedge \psi \subseteq K \div \psi$$

Στις περισσότερες περιπτώσεις όπου $\phi \notin K \div \phi \wedge \psi$, το $K \div \phi \wedge \psi$ είναι ίσο με το $K \div \phi$. Ωστόσο, η παρακάτω έκφραση δεν θα μπορούσε να αποτελέσει μέρος αυτών των περιπτώσεων: Εάν οι προτάσεις ϕ και ψ είναι ισοδύναμες ως προς την επιστημονική τους κατοχύρωση (όχι απαραίτητα λογικά ισοδύναμες), τότε και το ϕ και το ψ απορρίπτεται από το $K \div \phi \wedge \psi$, λαμβάνοντας υπόψη ότι μόνο το ϕ απορρίπτεται στο $K \div \phi$. Ωστόσο, μπορεί να αποδειχθεί ότι τα $(K \div 7)$ και $(K \div 8)$ μαζί με το βασικό σύνολο των αξιωμάτων για την αφαίρεση πεποιθήσεων επιφέρουν την επόμενη ιδιότητα (Factoring Condition) :

$$19 \quad \text{Είτε } K \div (\phi \wedge \psi) = K \div \phi \text{ είτε } K \div (\phi \wedge \psi) = K \div \psi \text{ είτε } K \div (\phi \wedge \psi) = K \div \phi \cap K \div \psi$$

Το αντίθετο είναι εφικτό. Υπό την σκέπη των βασικών αξιωμάτων η ιδιότητα 19 επιφέρει και το $(K \div 7)$ και το $(K \div 8)$. Μία πιο ενδυναμωμένη αρχή σε σχέση με την είναι η επόμενη:

$$20 \quad \text{Είτε } K \div (\phi \wedge \psi) = K \div \phi \text{ είτε } K \div (\phi \wedge \psi) = K \div \psi$$

Τέλος όλες οι προηγούμενες παρατηρήσεις - ιδιότητες παρέχουν μία τεράστια υποστήριξη στον ισχυρισμό ότι τα αξιώματα $(K - 1)$ - $(K - 8)$ όντως αποτελούν το καταλληλότερο σύνολο αξιωμάτων για τη διαδικασία της αφαίρεσης των πεποιθήσεων.

3.5 Από την αφαίρεση στην αναθεώρηση πεποιθήσεων και αντίστροφα

Στα προηγούμενα κεφάλαια, η αφαίρεση και η **αναθεώρηση πεποιθήσεων** παρουσιάστηκαν με δύο σύνολα αξιωμάτων. Αυτά τα σύνολα αξιωμάτων είναι ανεξάρτητα μεταξύ τους, με την έννοια ότι τα αξιώματα που αφορούν στην αφαίρεση πεποιθήσεων δεν σχετίζονται με την αναθεώρηση και το αντίστροφο. Μία λογική απορία λοιπόν, θα ήταν πως από την αφαίρεση θα μπορούσαμε να προσδιορίσουμε την αναθεώρηση και το αντίστροφο; Εάν αυτό είναι εφικτό, τότε το σύνολο των αρχικών ιδεών θα μπορεί να αναχθεί και οι επιστημονολογικές θεωρίες που εξετάζουμε θα είναι πολύ πιο εύκολα κατανοητές. Μια τέτοια σύνδεση μεταξύ αυτών των δύο τύπων αλλαγής των πεποιθήσεων προτάθηκε από τον Isaac Levi πριν από τους Alchourron, Gardenfors και Makinson.

Σε αυτήν την ενότητα θα παρουσιαστούν δύο επιχειρήματα που βοηθούν αυτό τον ισχυρισμό: ότι η αναθεώρηση μπορεί να οριστεί από την αφαίρεση και η αφαίρεση από την αναθεώρηση.

Ο Isaac Levi υποστηρίζει ότι υπάρχουν μόνο δύο μορφές αλλαγών πεποιθήσεων: η πρόσθεση και η αφαίρεση. Ο Isaac Levi ενισχύει τη θέση του αποδεικνύοντας ότι οι αναθεωρήσεις μπορούν να αναλυθούν ως μία ακολουθία από αφαιρέσεις και προσθέσεις. Στην ορολογία των συνόλων των πεποιθήσεων αυτή η ιδέα μπορεί να ερμηνευθεί ως εξής: Προκειμένου να κατασκευάσουμε την αναθεώρηση $K * \phi$ για ένα σύνολο πεποιθήσεων K και μία πρόταση ϕ , θα πρέπει πρώτα να αφαιρέσουμε την πρόταση $\neg \phi$ από το σύνολο K και μετά να προσθέσουμε την πρόταση ϕ στο νέο σύνολο $K \div \neg \phi$. Διαισθητικά, προκειμένου

να προσθέσουμε μία αντιφατική πληροφορία ϕ στο K προετοιμάζουμε το έδαφος με την απομάκρυνση της πεποίθησης $\neg\phi$ καθώς και όλες τις πεποιθήσεις που επιφέρουν το $\neg\phi$. Τυπικά η θέση του Levi μπορεί να εκφραστεί από τον ακόλουθο ορισμό:

$$(Def^*) \quad K * \phi = (K \div \neg\phi) + \phi \quad (\text{Levi identity})$$

Η παραπάνω έκφραση παρέχει τον ορισμό της αναθεώρησης μέσα από τους όρους της πρόσθεσης και της αφαίρεσης των πεποιθήσεων. Η έκφραση αυτή καλείται και "Levi identity". Οι Alchourron, Gardenfors και Makinson απέδειξαν ότι οι συναρτήσεις που προέρχονται από τα αξιώματά τους ικανοποιούν και την Levi identity

Εάν το σύνολο $K * \phi$ οριστεί κατ' αυτόν τον τρόπο, τότε υπάρχει η ελάχιστη αναγκαία αλλαγή για την ενσωμάτωση της πεποίθησης ϕ στο σύνολο K . Κάτι τέτοιο άλλωστε είναι αναμενόμενο σε κάθε ορισμό αναθεώρησης ενός συνόλου. Μία πεποίθηση ϕ για να είναι συνεπής με το νέο σύνολο θα πρέπει να έχει συμπεριληφθεί στο αρχικό η άρνησή της.

Επίσης το σύνολο $K \div \neg\phi$ ορίζεται ως η αναγκαία ελάχιστη αλλαγή του K προκειμένου να απορριφθεί η πεποίθηση $\neg\phi$. Τέλος, η πεποίθηση ϕ που θα πρέπει να προστεθεί υπαγορεύει ότι το $K \div \neg\phi$ θα είναι η ελάχιστη αναγκαία αλλαγή που χρειαζόμαστε για να ολοκληρωθεί όλη αυτή η διαδικασία.

Από το αίτημα $(K \div 1)$ και τον ορισμό της πρόσθεσης πεποιθήσεων προκύπτει ότι το σύνολο $K * \phi$, που ορίζεται από την "Levi identity" είναι ένα σύνολο πεποιθήσεων. Μία πολύ προφανής ερώτηση θα ήταν εάν τα εναπομείναντα αξιώματα για την αναθεώρηση των πεποιθήσεων ικανοποιούνται. Η απάντηση δίνεται μέσω των επόμενων δύο θεωρημάτων:

Θεώρημα 3.2 (Alchourron, C.E., P. Gärdenfors and D. Makinson. [2]) Εάν η συνάρτηση αφαίρεσης $-$ ικανοποιεί τα $(K \div 1)$ - $(K \div 4)$ και $(K \div 6)$ και η συνάρτηση πρόσθεσης $+$ ικανοποιεί τα $(K+1)$ - $(K+6)$ τότε η συνάρτηση αναθεώρησης που ορίζεται από την (Def^*) ικανοποιεί τα $(K * 1)$ - $(K * 6)$.

Ας σημειώσουμε ότι το αίτημα $(K \div 5)$ δεν χρησιμοποιείται στην παραγωγή της $*$ συνάρτησης αναθεώρησης.

Θεώρημα 3.3 (Alchourron, C.E., P. Gärdenfors and D. Makinson. [2]) Υποθέτουμε ότι οι προϋποθέσεις του θεωρήματος 3.2 ικανοποιούνται. Τότε (α) εάν το $(K \div 7)$ ικανοποιείται, το $(K * 7)$ ικανοποιείται για την ορισθείσα συνάρτηση αναθεώρησης και (β) εάν το $(K \div 8)$ ικανοποιείται, το $(K * 8)$ ικανοποιείται και αυτό για την ορισθείσα συνάρτηση αναθεώρησης.

Τα αποτελέσματα αυτά ενισχύουν εμφανώς την καταλληλότητα της $(Def *)$ ως ορισμό της συνάρτησης αναθεώρησης.

Έχοντας ολοκληρώσει την παρουσίαση και τον ορισμό συναρτήσεων αναθεώρησης σειρά έχει η μελέτη του αντίστροφου προβλήματος, δηλαδή η κατασκευή συναρτήσεων αφαίρεσης μέσω της αναθεώρησης των πεποιθήσεων. Η ιδέα η οποία παρουσιάστηκε από τον William Harper είναι η εξής: Μία πρόταση ψ γίνεται αποδεκτή στην αφαίρεση της πρότασης ϕ από το σύνολο K , δηλαδή στο σύνολο $K \div \phi$, εάν και μόνο εάν η πρόταση ψ γίνει αποδεκτή και στο K αλλά και στο $K * \neg \phi$. Πιο αναλυτικά, επειδή στο σύνολο $K * \neg \phi$ είναι η ελάχιστη αλλαγή του K , προκειμένου να γίνει αποδεκτή η πρόταση $\neg \phi$, περιέχει όσο το δυνατόν περισσότερη πληροφορία από το K η οποία δεν επιφέρει την πληροφορία ϕ . Φυσικά το σύνολο $K * \neg \phi$ περιέχει το $\neg \phi$ και όλες τις συνεπαγωγές του, ενώ οι πεποιθήσεις που ανήκουν από κοινού στα K και στα $K * \neg \phi$ δεν επιφέρουν την πεποίθηση ϕ . Τυπικά η παραπάνω ανάλυση διατυπώνεται από τον επόμενο ορισμό:

$$(Def \div) \quad K \div \phi = K \cap K * \neg \phi \quad (\text{Harper identity})$$

Η έκφραση αυτή ορίζει την αφαίρεση μέσα από τους όρους της αναθεώρησης των πεποιθήσεων. Επιπλέον η έκφραση αυτή καλείται και "Harper identity".

Επειδή η τομή δύο συνόλων πεποιθήσεων αποτελεί σύνολο πεποιθήσεων, προκύπτει ότι το σύνολο $K \div \phi$ (το οποίο ορίζεται από τη Harper identity) αποτελεί και αυτό σύνολο πεποιθήσεων. Συνεπώς αυτός ο ορισμός ικανοποιεί το αίτημα $(K \div 1)$. Ακολουθούν δύο θεωρήματα, οι αποδείξεις των οποίων ικανοποιούν και τα υπόλοιπα αξιώματα:

Θεώρημα 3.4 (Alchourron, C.E., P. Gärdenfors and D. Makinson. [2]) Εάν η συνάρτηση αναθεώρησης $*$ ικανοποιεί τα $(K * 1) - (K * 6)$, τότε η συνάρτηση αφαίρεσης που ορίζεται από την $(Def \div)$ ικανοποιεί τα $(K \div 1) - (K \div 6)$.

Θεώρημα 3.5 (Alchourron, C.E., P. Gärdenfors and D. Makinson. [2])
Υποθέτουμε ότι η συνάρτηση αναθεώρησης $*$ ικανοποιεί τα $(K*1) - (K*6)$.
Τότε (α) εάν το $(K*7)$ ικανοποιείται, το $(K \div 7)$ θα ικανοποιείται για την ορισθείσα συνάρτηση αφαίρεσης και (β) εάν το $(K*8)$ ικανοποιείται, το $(K \div 8)$ θα ικανοποιείται και αυτό για την ορισθείσα συνάρτηση αφαίρεσης.

Με τον ίδιο ακριβώς τρόπο, όπως και με τη "Levi identity" τα παραπάνω αποτελέσματα ενισχύουν ακόμη περισσότερο την καταλληλότητα των αξιωμάτων. Τα θεωρήματα 3.2 - 3.5 δείχνουν ότι οι συναρτήσεις αφαίρεσης και αναθεώρησης έχουν τις σωστές ιδιότητες. Παρόλα αυτά επιδιώκουμε οι δύο ορισμοί για τις συναρτήσεις να είναι ανταλλάξιμοι υπό την έννοια ότι, εάν ξεκινήσουμε με τον ορισμό και την κατασκευή μιας συνάρτησης αφαίρεσης και έπειτα χρησιμοποιήσουμε τον δεύτερο ορισμό για να αποκτήσουμε τη συνάρτηση αναθεώρησης, θα πρέπει να καταλήξουμε στην αυθεντική-πρώτη μας συνάρτηση. Για να αποδειχθεί αυτό, δηλαδή για να επιτύχουμε την ανταλλαξιμότητα μεταξύ των συναρτήσεων αφαίρεσης και αναθεώρησης πεποιθήσεων θα πρέπει να εφαρμόσουμε το $(Def \div)$ στο σύνολο $K \div \neg\phi$ και ύστερα να το αντικαταστήσουμε στον ορισμό $(Def*)$. Από τη Harper identity έχουμε: $K \div \neg\phi = K \cap K*\phi$ (1) και από τη Levi identity : $K*\phi = (K \div \neg\phi)+\phi$ (2). Άρα από τις σχέσεις (1), (2), από $K*\phi = K* \neg\neg\phi$ και απότο $(K*6)$ έχουμε:

$$21 \quad K*\phi = (K \cap K*\phi)+\phi$$

Για να ελέγξουμε και την αντίστροφη κατεύθυνση θα πρέπει να εφαρμόσουμε το $(Def*)$ στο σύνολο $K* \neg\phi$ και ύστερα να το αντικαταστήσουμε στον ορισμό του $(Def \div)$. Από τη Levi identity έχουμε: $K* \neg\phi = (K \div \phi)+\neg\phi$ (3) και από τη Harper identity : $K \div \phi = K \cap K* \neg\phi$

(4). Άρα από τις σχέσεις (3),(4), από $K \div \phi = K \div \neg\neg\phi$ και $(K \div 6)$ έχουμε:

$$22 \quad K \div \phi = K \cap (K \div \phi)+\neg\phi$$

Τα παραπάνω αποτελέσματα δείχνουν ότι, χρησιμοποιώντας μόνο τα βασικά αξιώματα για την αναθεώρηση και την αφαίρεση και με την βοήθεια των Levi και Harper identities, οι δύο αυτές διαδικασίες είναι όντως ανταλλάξιμες. Προηγουμένως είχε αναφερθεί ότι το "recovery postulate" μπορεί να μην είναι έγκυρο σε όλα τα πλαίσια εφαρμογής του. Αυτό αποτελεί ενδιαφέρουσα περίπτωση για το τι μπορούσε να συμβεί εάν αυτό το αίτημα έπαυε να υφίσταται. Ο Makinson απέδειξε ότι, εάν απορρίψουμε το αίτημα $(K \div 5)$, τότε θα

οδηγηθούμε σε μία σημαντική απώλεια. Αυτό διότι το $(K \div 5)$ δεν μπορεί να παραχθεί από τα αξιώματα $(K \div 1)$ - $(K \div 4)$ και $(K \div 6)$.

Σύμφωνα με τον Makinson , για να γίνουμε πιο ακριβείς , υποθέτουμε ότι υπάρχει μία συνάρτηση \div η οποία ικανοποιεί τα $(K \div 1)$ - $(K \div 4)$ και $(K \div 6)$ και η οποία καλείται συνάρτηση ανάκλησης (withdrawal function). Επειδή το $(K \div 5)$ δεν χρησιμοποιείται στο θεώρημα 3.2 γνωρίζουμε ότι κάθε συνάρτηση ανάκλησης βασισμένη στη Levi identity παράγει μία συνάρτηση αναθεώρησης που ικανοποιεί τα $(K * 1)$ - $(K * 6)$. Εάν τώρα $-$ και $-'$ είναι δύο συναρτήσεις ανάκλησης οι οποίες παράγουν την "ίδια" συνάρτηση αναθεώρησης τότε λέμε ότι οι $-$ και $-'$ είναι ισοδύναμες και πιο συγκεκριμένα "revision equivalent". Επιπλέον σημειώνουμε με $[-]$ την κλάση όλων των συναρτήσεων ανάκλησης που είναι "revision equivalent" με τη $-$. Επίσης λέμε ότι η συνάρτηση ανάκλησης $-'$ είναι μεγαλύτερη από τη $-$ στο σύνολο πεποιθήσεων K εάν $K \div \phi \subseteq K -' \phi$ για όλες τις πεποιθήσεις ϕ .

Ακόμη, ο Makinson απέδειξε το ακόλουθο λήμμα:

Λήμμα 3.1 Έστω K ένα σύνολο πεποιθήσεων. Τότε για κάθε συνάρτηση ανάκλησης $-$ στο σύνολο K , υπάρχει μία "μοναδική" συνάρτηση αφαίρεσης $-'$ στο K η οποία είναι "revision equivalent" με την $-$ και επίσης η $-'$ είναι το μεγαλύτερο στοιχείο της κλάσης $[-]$.

Καταλήγουμε στο παρακάτω συμπέρασμα: Εάν η προσοχή μας στραφεί στην αναθεώρηση των πεποιθήσεων, τότε, αν και η αναθεώρηση ενός συνόλου K μπορεί να παραχθεί από αρκετές συναρτήσεις ανάκλησης, υπάρχει μία μοναδική συνάρτηση εξ αυτών που θα ικανοποιεί και το recovery postulate $(K-5)$. Αυτή η συνάρτηση είναι επίσης και η μοναδική συνάρτηση ανάκλησης που απομακρύνει από το K όσο το δυνατόν λιγότερη πληροφορία. Το παραπάνω επιχείρημα καθιστά ξεκάθαρο τον ειδικό ρόλο του recovery postulate και ενισχύει τη διαισθητική του αποδοχή στα σύνολα των αξιωμάτων που μελετήθηκαν προηγουμένως.

Τέλος, ο Gardenfors παραδέχεται ότι η αφαίρεση πεποιθήσεων πολύ σπάνια είναι μία συνάρτηση η οποία γίνεται μεμονωμένα. Ουσιαστικά η αφαίρεση είναι το πρώτο βήμα της αναθεώρησης πεποιθήσεων και αυτό που συμβαίνει είναι ένας πράκτορας να χάνει την πίστη του σε μια πεποίθηση ϕ αφού πρώτα έχει δεχτεί στοιχεία ότι ισχύει η $\neg \phi$.

Ωστόσο, υπάρχει τουλάχιστον μία περίπτωση κατά την οποία η λειτουργία της αφαίρεσης συμβαίνει ανεξάρτητα της λειτουργίας της αναθεώρησης και αυτό είναι κατά τη διάρκεια της επιχειρηματολογίας. Έστω δύο πράκτορες A και B, που υποστηρίζουν ένα συγκεκριμένο ζήτημα για το οποίο έχουν αντίθετες απόψεις. Είναι πολύ πιθανό ότι για χάρη του επιχειρήματος, δύο πράκτορες (προσωρινά) θα αφαιρέσουν τις πεποιθήσεις τους για την επίτευξη κάποιου κοινού εδάφους, από το οποίο στη συνέχεια θα αρχίσουν να φτιάχνουν υπόθεσή τους.

ΚΕΦΑΛΑΙΟ 4

4.1 Το πρόβλημα της κατασκευής συναρτήσεων

Έχοντας προσδιορίσει τα αξιώματα για τις συναρτήσεις της πρόσθεσης, της αφαίρεσης και της αναθεώρησης πεποιθήσεων, το επόμενο στάδιο που πρέπει να μελετήσουμε είναι η ανάπτυξη κατασκευαστικών μοντέλων για αυτές τις συναρτήσεις. Σημειώνεται ότι στο προηγούμενο κεφάλαιο στην ενότητα 3.5 μας δόθηκε η ευκαιρία να παρουσιάσουμε μία ακριβή κατασκευή συναρτήσεων αφαίρεσης που οδηγεί σε συναρτήσεις αναθεώρησης μέσω της εξίσωσης του Levi (η εξίσωση του Harper ορίζει ακριβώς το αντίστροφο). Στόχος αυτής της ενότητας είναι η παρουσίαση μίας σειράς κατασκευαστικών διαδικασιών που αφορούν συναρτήσεις αφαίρεσης και έμμεσα τις συναρτήσεις αναθεώρησης.

Όταν πραγματοποιούμε την αφαίρεση μίας πρότασης ϕ από μία θεωρία K , το κριτήριο της πληροφοριακής οικονομίας απαιτεί ότι το $K-\phi$ θα περιέχει όσο το δυνατόν περισσότερη πληροφορία από το K χωρίς όμως να παράγει την πεποίθηση ϕ . Επομένως τα μέγιστα υποσύνολα του K που δεν περιέχουν το ϕ έχουν ιδιαίτερο ενδιαφέρον στην περίπτωση που προσπαθούμε να ορίσουμε μία κατασκευαστική μέθοδο για τις συναρτήσεις αφαίρεσης. Οι κατασκευαστικές διαδικασίες που βασίζονται στα μέγιστα υποσύνολα ενός αρχικού συνόλου K μελετώνται λεπτομερώς στις ενότητες 4.2 - 4.4.

Μία διαφορετική προσέγγιση για την κατασκευή συναρτήσεων αφαίρεσης και αναθεώρησης είναι η γνωσιακή κατοχύρωση (epistemic entrenchment) των προτάσεων (ενότητα 4.5). Με λίγα λόγια οι προτάσεις που γίνονται αποδεκτές από κάποιον πράκτορα σε μία δεδομένη κατάσταση πεποιθήσεων έχουν

διαφορετική ισχύ, δηλαδή είναι περισσότερο χρήσιμες από κάποιες άλλες κατά τη διαδικασία λήψης αποφάσεων. Εάν μπορούσαμε να ορίσουμε τη σχετική επιστημονική ισχύ των προτάσεων που ανήκουν σε ένα σύνολο πεποιθήσεων K , τότε ο βαθμός αυτής της ισχύς μπορεί να χρησιμοποιηθεί ως οδηγός για τη διαδικασία της αφαίρεσης των πεποιθήσεων. Συμπερασματικά οι προτάσεις που θα πρέπει να απομακρυνθούν είναι αυτές με το μικρότερο επιστημολογικό ενδιαφέρον. Το σενάριο της γνωσιακής κατοχύρωσης παρουσιάζεται λεπτομερώς στην ενότητα 4.5. Τέλος τα AGM αξιώματα εγγυώνται ότι κάθε συνάρτηση αφαίρεσης (ή αναθεώρησης) μπορεί να ορίσει μία συγκεκριμένη διάταξη προτάσεων, η οποία βασίζεται στην έννοια της γνωσιακής κατοχύρωσης.

4.2 Συναρτήσεις Αφαίρεσης Μέγιστης Επιλογής

Η ενότητα αυτή εστιάζεται στην κατασκευή συναρτήσεων αφαίρεσης για ένα σύνολο πεποιθήσεων K . Η γενική ιδέα είναι να ξεκινήσουμε από το σύνολο K και βάσει κάποιας μεθοδολογίας να επιλέξουμε ποιες προτάσεις (έστω ϕ) θα πρέπει να διαγραφούν από το K , ώστε το νέο σύνολο $K \div \phi$ να μην περιλαμβάνει την πρόταση ϕ ως λογική συνέπεια. Εάν επικεντρώσουμε στο κριτήριο της πληροφοριακής οικονομίας, η απαίτηση, όπου το $K \div \phi$ αποτελεί το μεγαλύτερο υποσύνολο του K , φαίνεται φυσιολογική. Δηλαδή το νέο σύνολο πεποιθήσεων $K \div \phi$ θα πρέπει να περιλαμβάνει όσο το δυνατόν περισσότερη πληροφορία από το K . Για το λόγο αυτό επικεντρώνουμε στα μέγιστα υποσύνολα του K που δεν περιέχουν την πρόταση ϕ και έπειτα προσπαθούμε να βρούμε το καλύτερο από αυτά.

Πριν προχωρήσουμε όμως πρέπει να εισάγουμε κάποιους επιπλέον περιορισμούς:

Ένα σύνολο πεποιθήσεων K' που δεν περιλαμβάνει την πρόταση ϕ ορίζεται ως υπόλοιπο του K ως προς το ϕ εάν και μόνο εάν:

- $K' \subseteq K$
- $\phi \notin K'$
- για κάθε $\psi \in K$ και $\psi \notin K'$ ισχύει $\psi \rightarrow \phi \in K'$

Ο τελευταίος όρος υπαγορεύει ότι, εάν προστεθεί η πρόταση ψ στο σύνολο K' , τότε το K' θα δημιουργήσει την πρόταση ϕ , δηλαδή: $\psi \rightarrow \phi \in K' \Rightarrow \phi \in \text{Cn}(K' \cup \{\psi\})$.

Επειδή το K είναι σύνολο πεποιθήσεων, μπορούμε εύκολα να επιβεβαιώσουμε ότι υπάρχουν πάντα υποσύνολα στο K τα οποία είναι όντως μέγιστα αλλά αποτυγχάνουν στη συνεπαγωγή της πρότασης ϕ (εκτός αν η ϕ είναι ταυτολογία). Το σύνολο όλων των υποσυνόλων K' συμβολίζεται ως $K \perp \phi$. Το σύνολο αυτό αποτελεί τη βάση για όλες τις διαδικασίες αφαίρεσης που θα μελετηθούν σ' αυτήν και στις επόμενες ενότητες. Η ιδέα της κατασκευής μπορεί τώρα να αποτυπωθεί με μεγαλύτερη ακρίβεια:

Η κατασκευή του $K \div \phi$ από το K βάσει της πρότασης ϕ ορίζεται από μία συνάρτηση επιλογής S η οποία διαλέγει ένα μόνο στοιχείο $S(K \perp \phi)$ του συνόλου $K \perp \phi$ όταν $K \perp \phi \neq \emptyset$. Έτσι το $K \div \phi$ ορίζεται ως εξής:

(Def Max)

$$K \div \phi = \begin{cases} S(K \perp \phi), & \text{όταν } K \perp \phi \neq \emptyset \\ \kappa, & \text{διαφορετικά} \end{cases}$$

Επομένως το $K \div \phi$ είναι κάποιο μέγιστο υποσύνολο του K που αποτυγχάνει στη συνεπαγωγή του ϕ . Οι συναρτήσεις αφαίρεσης που ορίζονται μέσω των συναρτήσεων επιλογής ονομάζονται συναρτήσεις αφαίρεσης μέγιστης επιλογής (maxichoice contraction functions). Το παρακάτω λήμμα αποτελεί ένα σημαντικό αποτέλεσμα για τις συναρτήσεις αφαίρεσης μέγιστης επιλογής:

Λήμμα 4.1 (Alchourron, Makinson [2]) Κάθε συνάρτηση αφαίρεσης μέγιστης επιλογής ικανοποιεί τα αξιώματα $(K \div 1)$ - $(K \div 6)$.

Οι Alchourron και Makinson στην [2] επισήμαναν ότι τα αξιώματα $(K \div 7)$ και $(K \div 8)$ δεν ισχύουν στη γενική περίπτωση των συναρτήσεων μέγιστης επιλογής. Για να ισχύσουν και αυτά τα αξιώματα τότε θα πρέπει να εισαχθούν επιπρόσθετοι περιορισμοί. Ένας πρώτος περιορισμός θα μπορούσε να είναι ο εξής:

Μία συνάρτηση μέγιστης επιλογής είναι διατεταγμένη εάν και μόνο εάν υπάρχει μία μερική διάταξη \leq στο σύνολο των υπολοίπων του K ως προς οποιαδήποτε

πρόταση τέτοιο ώστε $K' \leq K \div \phi$ για όλες τις προτάσεις ϕ και όλα τα $K' \in K \perp \phi$. Με άλλα λόγια η συνάρτηση μέγιστης επιλογής είναι διατεταγμένη, εάν το $K \div \phi$ είναι ένα από τα πιο επιθυμητά στοιχεία του $K \perp \phi$ σύμφωνα με την \leq .

Λήμμα 4.2 Κάθε διατεταγμένη συνάρτηση αφαίρεσης μέγιστης επιλογής ικανοποιεί τα αξιώματα $(K \div 1) - (K \div 8)$ και την αρχή 3.10.

Επειδή η αρχή 3.10 είναι αρκετά ισχυρή και μεταξύ άλλων συνεπάγεται τα αξιώματα $(K \div 7)$ και $(K \div 8)$ το παρακάτω θεώρημα θεμελιώνει τη σύνδεση των βασικών αξιωμάτων με τις συναρτήσεις μέγιστης επιλογής:

Θεώρημα 4.1 (Alchourron, Makinson [2]) Κάθε συνάρτηση αφαίρεσης που ικανοποιεί τα $(K \div 1) - (K \div 6)$ και την 3.10 μπορεί να γενικευθεί από μία διατεταγμένη συνάρτηση αφαίρεσης μέγιστης επιλογής

Συμπερασματικά οι συναρτήσεις μέγιστης επιλογής δημιουργούν μέγιστα σύνολα πεποιθήσεων, κάτι το οποίο είναι ανεπιθύμητο. Υπενθυμίζουμε ότι ένα μέγιστο σύνολο πεποιθήσεων K είναι ένα σύνολο πεποιθήσεων τέτοιο ώστε για κάθε πρόταση $\psi \in L$, είτε $\psi \in K$ είτε $\neg\psi \in K$.

Λήμμα 4.3 Εάν $\phi \in K$ και το σύνολο $K \div \phi$ ορίζεται από μία συνάρτηση μέγιστης επιλογής, τότε για κάθε πρόταση ψ είτε $\phi \vee \psi \in K \div \phi$ είτε $\phi \vee \neg\psi \in K \div \phi$.

Πόρισμα 4.1 Εάν μία συνάρτηση αναθεώρησης $*$ ορίζεται από μία συνάρτηση μέγιστης επιλογής – μέσω της ιδιότητας του Levi, τότε για κάθε ϕ τέτοιο ώστε $\neg\phi \in K$, το $K * \phi$ θα είναι μέγιστο.

Από τα παραπάνω αποτελέσματα οι συναρτήσεις μέγιστης επιλογής πράγματι δημιουργούν μέγιστα σύνολα πεποιθήσεων, κάτι το οποίο είναι ανεπιθύμητο. Έτσι το σύνολο $K * \phi$ είναι αρκετά μεγάλο ώστε να συμπεριληφθεί στη διαδικασία αναθεώρησης πεποιθήσεων. Αυτά τα αποτελέσματα επιβάλλουν τη διερεύνηση διαφορετικών τρόπων κατασκευής συναρτήσεων αφαίρεσης πεποιθήσεων.

4.3 Συναρτήσεις Αφαίρεσης Πλήρους Επιλογής

Στην προηγούμενη ενότητα παρατηρήσαμε ότι ένα μέγιστο υποσύνολο του K που αποτυγχάνει στη συνεπαγωγή του ϕ περιέχει αρκετή πληροφορία ώστε να χρησιμοποιηθεί ως $K \div \phi$. Ωστόσο τα στοιχεία του $K \perp \phi$ μπορούν να χρησιμοποιηθούν με διαφορετικούς τρόπους. Μία δεύτερη ιδέα είναι να υποθέσουμε ότι το $K \div \phi$ περιέχει μόνο προτάσεις που είναι κοινές με όλες τις προτάσεις όλων των μέγιστων υποσυνόλων του $K \perp \phi$.

Μαθηματικά, για κάθε σύνολο K και κάθε πρόταση ϕ το $K \div \phi$ μπορεί να οριστεί ως εξής:

$$(Def Meet) \quad K \div \phi = \begin{cases} \bigcap S(K \perp \phi), & \text{όταν } K \perp \phi \neq \emptyset \\ K, & \text{διαφορετικά} \end{cases}$$

Έτσι μία πρόταση ψ ανήκει στο $K \div \phi$ εάν και μόνο εάν εμπεριέχεται σε όλα τα μέγιστα υποσύνολα του K που αποτυγχάνουν στη συνεπαγωγή του ϕ . Μία τέτοια συνάρτηση ονομάζεται συνάρτηση αφαίρεσης πλήρους επιλογής (full meet contraction function).

Για άλλη μία φορά η ερώτηση που προκύπτει είναι εάν τέτοιες συναρτήσεις ικανοποιούν τα αξιώματα των AGM. Το επόμενο λήμμα μας παρέχει μία πρώτη εγγύηση:

Λήμμα 4.4 Κάθε συνάρτηση αφαίρεσης πλήρους επιλογής ικανοποιεί τα αξιώματα $(K \div 1)$ - $(K \div 6)$.

Επιπρόσθετα μία ιδιότητα των συναρτήσεων πλήρους επιλογής είναι η ιδιότητα τομής (intersection condition):

$(K \div 1)$ Για κάθε πρόταση ϕ και ψ ισχύει $K \div \phi \wedge \psi = K \div \phi \cap K \div \psi$

Θεώρημα 4.2 (Alchourron, Makinson [2]) Κάθε συνάρτηση αφαίρεσης που ικανοποιεί τα $(K \div 1)$ - $(K \div 6)$ και την $(K \div 1)$ μπορεί να γενικευθεί από μία συνάρτηση αφαίρεσης πλήρους επιλογής.

Τα παραπάνω αποτελέσματα φαίνεται να θεμελιώνουν τη χρήση των συναρτήσεων πλήρους επιλογής ως ένα τυπικό μοντέλο της διαδικασίας της αφαίρεσης πεποιθήσεων. Ωστόσο, το μειονέκτημα των συναρτήσεων πλήρους επιλογής είναι το αντίστροφο αυτού των συναρτήσεων μέγιστης επιλογής. Οι

συναρτήσεις πλήρους επιλογής οδηγούν σε σύνολα πεπιοιθήσεων που είναι υπερβολικά μικρά.

Λήμμα 4.5 Εάν το $K \div \phi$ ορίζεται από τις συναρτήσεις πλήρους επιλογής και η πρόταση $\phi \in K$ τότε $\psi \in K \div \phi$ εάν και μόνο εάν $\psi \in K$ και $\neg \phi \vdash \psi$.

Με άλλα λόγια εάν αφαιρέσουμε μία πρόταση ϕ από ένα σύνολο πεπιοιθήσεων K , τότε καταλήγουμε με προτάσεις στο K που αποτελούν λογικές συνεπαγωγές του $\neg \phi$. Αυτό οδηγεί στο επόμενο συμπέρασμα:

Πόρισμα 4.2 Εάν μία συνάρτηση αναθεώρησης $*$ ορίζεται από μία συνάρτηση πλήρους επιλογής – μέσω της ιδιότητας του Levi, τότε για κάθε ϕ τέτοιο ώστε $\neg \phi \in K$, το $K * \phi$ θα περιέχει μόνο το ϕ με όλες τις λογικές του συνεπαγωγές. Ανεξάρτητα από τα αρνητικά αυτά αποτελέσματα η ισχύς των συναρτήσεων πλήρους επιλογής παρουσιάζει εξαιρετικό ενδιαφέρον, διότι παρέχει ένα σαφές κάτω όριο για κάθε διαδικασία αφαίρεσης, η οποία ικανοποιεί τα αξιώματα $(K \div 1) - (K \div 6)$.

4.4 Συναρτήσεις Αφαίρεσης Μερικής Επιλογής

Το χαρακτηριστικό, όπως αναλύσαμε σε προηγούμενη ενότητα, είναι ότι οι συναρτήσεις αφαίρεσης πλήρους επιλογής παράγουν πάντα την πιο ελάχιστη πληροφορία. Αντίθετα, οι συναρτήσεις αφαίρεσης μέγιστης επιλογής με την βοήθεια των συναρτήσεων επιλογής επιστρέφουν πάντα ένα μεγάλο σύνολο πεπιοιθήσεων το οποίο είναι μοναδικό. Πρέπει να σημειωθεί ότι αυτές οι συναρτήσεις δεν είναι μεταβατικά ορθολογικές και δεν ικανοποιούν πάντα τα AGM αξιώματα για την συνάρτηση της αφαίρεσης πεπιοιθήσεων.

Δεδομένου της αρχής της ελάχιστης αλλαγής, το πρώτο πράγμα που μας έρχεται στο μυαλό είναι προσδιορίσουμε ένα μέγιστο υποσύνολο του K , το οποίο δεν περιέχει την πληροφορία ϕ , και να παράγουμε μία συνάρτηση

αφαίρεσης. Δυστυχώς όμως υπάρχουν περισσότερα από ένα τέτοια μεγιστικά υποσύνολα, από τα οποία είναι δύσκολο να ξεχωρίσουμε κάποιο. Παρόλα αυτά, τα μεγιστικά αυτά υποσύνολα είναι μια καλή αρχή. Για αυτό κάθε μεγιστικό υποσύνολο που δεν περιέχει το ϕ , το ονομάζουμε « ϕ – remainder». Το σύνολο όλων των « ϕ – remainders» το συμβολίζουμε $K \perp \phi$.

Για να είμαστε πιο σωστοί, χρησιμοποιούμε μια συνάρτηση επιλογής S , η οποία επιλέγει από το $K \perp \phi$, ένα μη κενό υποσύνολο $S(K \perp \phi)$ εφόσον το $K \perp \phi$ είναι και αυτό διαφορετικό του κενού συνόλου. Η συνάρτηση αφαίρεσης μπορεί να οριστεί ως ακολούθως:

$$(\text{Def Part}) \quad K \div \phi = \cap S(K \perp \phi)$$

Κάθε συνάρτηση αφαίρεσης αυτής της μορφής ονομάζεται συνάρτηση αφαίρεσης μερικής επιλογής (partial meet contraction function). Η ιδέα είναι η εξής: η συνάρτηση επιλογής S επιλέγει όλα τα υπόλοιπα του $K \perp \phi$ τα οποία έχουν μεγαλύτερη γνωσιακή ισχύ.

Έτσι μία πρόταση ψ ανήκει στο $K \div \phi$ εάν και μόνο εάν είναι στοιχείο όλων των βέλτιστων υπολοίπων (πιο κατοχυρωμένα) του K ως προς το ϕ . Να σημειωθεί ότι το σενάριο των συναρτήσεων αφαίρεσης μερικής επιλογής περιλαμβάνει ως ειδικές περιπτώσεις τις συναρτήσεις αφαίρεσης μέγιστης και πλήρους επιλογής. Η συνάρτηση αφαίρεσης μέγιστης επιλογής είναι και συνάρτηση αφαίρεσης μερικής επιλογής με το $S(K \perp \phi)$ ως μοναδικό στοιχείο. Η συνάρτηση αφαίρεσης πλήρους επιλογής είναι και συνάρτηση αφαίρεσης μερικής επιλογής με το $S(K \perp \phi)$ ως ολόκληρο το σύνολο $K \perp \phi$.

Λήμμα 4.6 Κάθε συνάρτηση αφαίρεσης μερικής επιλογής ικανοποιεί τα αξιώματα $(K \div 1)$ - $(K \div 6)$.

Θεώρημα 4.3 (Alchourron, C.E., P. Gärdenfors and D. Makinson. [1]) Έστω – μία συνάρτηση αφαίρεσης. Για κάθε σύνολο πεποιθήσεων K η \div είναι η συνάρτηση αφαίρεσης μερικής επιλογής εάν και μόνο εάν ικανοποιεί τα αξιώματα $(K \div 1)$ - $(K \div 6)$ για τη διαδικασία της αφαίρεσης στο σύνολο πεποιθήσεων K .

Υποθέτουμε ότι τα αξιώματα $(K \div 1) - (K \div 6)$ είναι πράγματι θεμελιώδη για κάθε ορθολογική συνάρτηση αφαίρεσης. Από το θεώρημα 4.3 προκύπτει ότι, όταν ψάχνουμε την πιο λογική μέθοδο για την αναπαράσταση συναρτήσεων αφαίρεσης, θα πρέπει να επικεντρωνόμαστε σε συναρτήσεις αφαίρεσης μερικής επιλογής.

Στον ορισμό των συναρτήσεων αφαίρεσης μερικής επιλογής η συνάρτηση επιλογής S θεωρείται δεδομένη. Όμως εάν χρειαστεί να χρησιμοποιήσουμε μία πιο ειδική μέθοδο αφαίρεσης, θα πρέπει να γνωρίζουμε τον τρόπο που θα καθορίσουμε τα πιο ισχυρά στοιχεία (βάσει γνωσιακής κατοχύρωσης) του $K \perp \phi$. Ο πιο απλός τρόπος είναι να επιλέξουμε τα κορυφαία στοιχεία του $K \perp \phi$. Η διάταξη των υπολοίπων θα πρέπει να είναι η ίδια για διαφορετικές επιλογές της πρότασης ϕ . Τεχνικά αυτό επιτυγχάνεται με την εισαγωγή του $M(K)$. Το $M(K)$ αναπαριστά την ένωση όλων των συνόλων της οικογένειας $K \perp \phi$, όπου η ϕ είναι οποιαδήποτε πρόταση στο σύνολο πεποιθήσεων K και δεν είναι λογικά έγκυρη. Τότε υποθέτουμε ότι υπάρχει σχέση \leq στο $M(K)$. Όταν το $K \perp \phi \neq 0$ δηλαδή όταν το ϕ δεν είναι ταυτολογία, αυτή η σχέση χρησιμοποιείται για να ορίσει τη συνάρτηση επιλογής S :

(Def S): $S(K \perp \phi) = \{K' \in K \perp \phi : K'' \leq K' \text{ για όλα τα } K'' \in K \perp \phi\}$

Αυτή η ταυτότητα καλείται και "marking-off identity", διότι η σχέση \leq σημειώνει τα $S(K \perp \phi)$ ως τα πιο ισχυρά μέγιστα υποσύνολα του K που αποτυγχάνουν στη συνεπαγωγή της πρότασης ϕ . Εάν η συνάρτηση επιλογής S ορισθεί μέσω της "marking-off identity", τότε για μερικές \leq θα λέμε ότι η S είναι σχεσιακή ως προς το K και ακόμα η συνάρτηση αφαίρεσης που παράγεται από μερικές τέτοιες συναρτήσεις επιλογής S λέμε ότι είναι σχεσιακή συνάρτηση αφαίρεσης μερικής επιλογής.

Μία συνάρτηση επιλογής S είναι μεταβατικά ορθολογική και μπορεί να παραχθεί από μία δυαδική σχέση $\ll 2^K$.

Μέχρι στιγμής οι συναρτήσεις αφαίρεσης μερικής επιλογής ικανοποιούν τα αξιώματα $(K \div 1) - (K \div 6)$. Για τα αξιώματα $(K \div 7) - (K \div 8)$ θα πρέπει να εισαχθούν επιπλέον ιδιότητες και ορισμοί στις διατάξεις των μέγιστων υποσυνόλων του $K \perp \phi$.

4.4 Γνωσιακή Κατοχύρωση

Οι συναρτήσεις επιλογής αποτελούν ένα επίσημο τρόπο αποκρυπτογράφησης των έχτρα λογικών παραγόντων μιας πληροφορίας φ , οι οποίοι απομακρύνονται όταν απομακρύνεται η πληροφορία φ από την θεωρία K . Αυτοί οι έχτρα λογικοί παράγοντες σχετίζονται με το περιεχόμενο της γνωσιακής αξίας των πεποιθήσεων ενός πράκτορα. Για παράδειγμα, η πεποίθηση ψ «όλοι οι κύκνοι είναι λευκοί» είναι ποιο σημαντική για τον πράκτορα από την πεποίθηση φ «η Λούσυ είναι ένας κύκνος». Συνεπώς, εάν ο πράκτορας χρειαστεί να διαλέξει μεταξύ των δύο πεποιθήσεων, θα πάψει να πιστεύει την δεύτερη γιατί έχει μικρότερη επεξηγηματική αξία.

Εκτιμήσεις σαν τις παραπάνω, οδήγησαν τους Gardenfors και Markison να εισάγουν την ιδέα της γνωσιακής κατοχύρωσης (epistemic entrenchment) . Επομένως, η γνωσιακή κατοχύρωση μιας πεποίθησης ψ είναι ο βαθμός αντίστασης της πεποίθησης ψ να αλλάξει. Όσο πιο κατοχυρωμένη η πεποίθηση ψ είναι τόσο πιο δύσκολα θα αφαιρεθεί κατά την διάρκεια της συνάρτησης αφαίρεσης, από μία άλλη πεποίθηση φ .

Αυτή είναι η βασική ιδέα του Gardenfors, ο οποίος πρότεινε ότι στην αφαίρεση πεποιθήσεων πρέπει να υπάρχει μία σειρά γνωσιακής κατοχύρωσης. Τα ίδια του τα λόγια ήταν: «έστω ότι όλες οι πεποιθήσεις σε ένα σύνολο πεποιθήσεων θεωρούνται γεγονότα (έχοντας έτσι την μέγιστη πιθανότητα) , αυτό όμως δεν σημαίνει ότι όλες οι πεποιθήσεις έχουν ισοδύναμη αξία στον σχεδιασμό ή την εύρεση λύσεων σε προβλήματα. Συγκεκριμένα κομμάτια της γνώσης και των πεποιθήσεων μας σχετικά με τον κόσμο, είναι πιο σημαντικά από άλλα , κατά την διάρκεια σχεδιασμού μελλοντικών μας ενεργειών, σύνδεσης επιστημονικών ερευνών ή επεξήγησης γενικότερα. Θα λέγαμε ότι μερικές πεποιθήσεις σε ένα σύνολο πεποιθήσεων έχουν υψηλότερο βαθμό από κάποιες άλλες. Η ιδέα για την κατασκευή είναι όταν σε ένα σύνολο πεποιθήσεων K κάνουμε αναθεώρηση ή αφαίρεση , οι πεποιθήσεις του K που θα απομακρυνθούν είναι αυτές που έχουν τους χαμηλότερους βαθμούς γνωσιακής κατοχύρωσης».

Η εφαρμογή αυτής της ιδέας γίνεται σε απλές προτάσεις που ανήκουν στη γλώσσα L . Με λίγα λόγια, κατά τη διαδικασία της αλλαγής η γνωσιακή κατοχύρωση της πρότασης είναι αυτή που καθορίζει την ίδια της την πορεία.

Οι βασικές ιδέες για αυτό το μοντέλο είναι οι εξής:

1. Η δυνατότητα ορισμού της γνωσιακής κατοχύρωσης των προτάσεων σε ένα σύνολο πεπιοθήσεων K (ανεξάρτητα με το τι συμβαίνει μέσα σε αυτό) κατά τη διαδικασία της αναθεώρησης ή της αφαίρεσης των πεπιοθήσεων.
2. Όταν ένα σύνολο πεπιοθήσεων αφαιρείται ή αναθεωρείται, οι προτάσεις στο K που θα απομακρυνθούν είναι αυτές με την ελάχιστη γνωσιακή κατοχύρωση.

Θεμελιώδες κριτήριο για τον καθορισμό της γνωσιακής κατοχύρωσης μίας πρότασης είναι πόσο χρήσιμη μπορεί να είναι αυτή η πρόταση κατά τη διάρκεια μίας συζήτησης. Κάποια τμήματα της γνώσης και των πεπιοθήσεών μας είναι πολύ πιο σημαντικά από άλλα.

Πολλές φορές η γνωσιακή κατοχύρωση μίας πρότασης εξαρτάται από την κατάσταση των πεπιοθήσεων στην οποία επέρχεται. Διαφορετικά σύνολα πεπιοθήσεων συνδέονται με διαφορετικές διατάξεις γνωσιακής κατοχύρωσης. Ακόμη και αν τα σύνολα των προτάσεων που εμπεριέχονται στα σύνολα των πεπιοθήσεων έχουν κοινά χαρακτηριστικά, οι διατάξεις γνωσιακής κατοχύρωσης μπορεί να μη συμφωνούν σε όλα αυτά τα χαρακτηριστικά.

Τα ακόλουθα σύμβολα θα χρησιμοποιηθούν για την γνωσιακή κατοχύρωση:

$\Phi \leq \psi$: το Φ είναι τόσο κατοχυρωμένο όσο το ψ

$\Phi < \psi$: το Φ είναι λιγότερο κατοχυρωμένο από το ψ

$\Phi \equiv \psi$: το Φ και το ψ είναι το ίδιο κατοχυρωμένα

Ακόμα το $<$ και το \equiv μπορούν να προσδιοριστούν από το \leq :

$\Phi < \psi$: εάν και μόνο εάν $(\Phi < \psi) \& \neg(\psi \leq \Phi)$

$\Phi \equiv \psi$: εάν και μόνο εάν $(\Phi \leq \psi) \& (\psi \leq \Phi)$

Ο Gardenfors παρουσιάζει ένα σύνολο από πέντε αξιώματα για τη γνωσιακή κατοχύρωση τα οποία θα αποτελέσουν τη βάση για την κατασκευή συναρτήσεων αφαίρεσης και αναθεώρησης. Έτσι βάσει αυτών των αξιωμάτων μπορούμε πλέον να ξεκινήσουμε από μία συνάρτηση αφαίρεσης και να καταλήξουμε σε μία διάταξη γνωσιακής κατοχύρωσης. Όπως αναφέραμε

νωρίτερα, η ιδέα κατασκευής αυτού του μοντέλου είναι ότι οι προτάσεις που δύσκολα απομακρύνονται από ένα σύνολο πεποιθήσεων είναι αυτές με τη μεγαλύτερη επιστημολογική κατοχύρωση. Έστω για παράδειγμα το σύνολο πεποιθήσεων $K \div \phi \wedge \psi$. Στην περίπτωση όπου και το ϕ αλλά και το ψ δεν είναι ταυτολογίες τότε τουλάχιστον μία από τις προτάσεις αυτές θα πρέπει να απομακρυνθεί. Εάν απομακρυνθεί η πρόταση ψ , τότε είναι ξεκάθαρο ότι η ϕ είναι πιο κατοχυρωμένη από την ψ . Σε μαθηματικούς όρους αυτό μπορεί να διατυπωθεί ως εξής:

$$\psi \leq \phi \text{ εάν και μόνο εάν } \psi \notin K \div \phi \wedge \psi$$

Επιστρέφουμε τώρα στο ζήτημα της διατύπωσης των αξιωμάτων για τη γνωσιακή κατοχύρωση. Σε καμία περίπτωση η γνωσιακή κατοχύρωση δεν θα πρέπει να μετριέται ποσοτικά. Για το λόγο αυτό το σύνολο των αξιωμάτων αφορά μόνο την ποιοτική της πλευρά.

Εάν οι ϕ , ψ είναι προτάσεις της γλώσσας L , ο συμβολισμός $\phi \leq \psi$ χρησιμοποιείται ως συντόμευση "το ψ είναι τουλάχιστον κατοχυρωμένο όσο είναι και το ϕ ". Η αυστηρή σχέση $\phi < \psi$ αναπαριστά ότι "το ψ είναι πιο κατοχυρωμένο από το ϕ ". Μπορεί επίσης να οριστεί και ως " $\phi \leq \psi$ και όχι $\psi \leq \phi$ ". Αξίζει να σημειώσουμε ότι η σχέση \leq ορίζεται μόνο σε εξάρτηση με το σύνολο πεποιθήσεων K . Αυτό γιατί, διαφορετικά σύνολα πεποιθήσεων μπορούν να σχετίζονται με διαφορετικές διατάξεις από γνωσιακές κατοχυρώσεις. Το αίτημα της **Μεταβατικότητας** ορίζει την ελάχιστη απαίτηση σε μία σχέση διατάξεων:

(EE1) Για κάθε ϕ , ψ , γ εάν $\phi \leq \psi$ και $\psi \leq \gamma$ τότε $\phi \leq \gamma$ (*transitivity*)

Όταν αφαιρούμε ή αναθεωρούμε ένα σύνολο πεποιθήσεων, μερικές από τις προηγούμενες αποδεκτές προτάσεις θα πρέπει να απομακρυνθούν. Η διάταξη της γνωσιακής κατοχύρωσης μπορεί να χρησιμοποιηθεί για να καθορίσει ποιες προτάσεις θα πρέπει να απομακρυνθούν βάσει του κριτηρίου της πληροφοριακής οικονομίας. Οι λιγότερο κατοχυρωμένες προτάσεις απομακρύνονται πρώτες, έτσι ώστε η έλλειψη πληροφορίας να είναι η ελάχιστη δυνατή. Μία ειδική περίπτωση γνωσιακής κατοχύρωσης διατυπώνεται από το αίτημα της **Κυριαρχίας**, το οποίο απαιτεί ότι, εάν μία πρόταση είναι λογικά ισχυρότερη από μία άλλη, τότε θα είναι λιγότερο κατοχυρωμένη:

(EE2) Για κάθε ϕ , ψ εάν $\phi \vdash \psi$ τότε $\phi \leq \psi$ (*Dominance*)

Διαισθητικά, εάν η πρόταση ϕ παράγει τη ψ και είτε η ϕ είτε η ψ πρέπει να απορριφθούν από το σύνολο πεποιθήσεων K , τότε μικρότερη αλλαγή θα έχουμε εάν απορρίψουμε το ϕ και διατηρήσουμε το ψ παρά να απορρίψουμε το ψ , (Επειδή η θεωρία μας είναι κλειστή ως προς λογική συνεπαγωγή εάν απορρίψουμε το ψ θα πρέπει να απορρίψουμε και το ϕ). Εκ πρώτης άποψης φαίνεται ότι δημιουργείται μία αντίφαση με τα όσα υπαγορεύει το κριτήριο της πληροφοριακής οικονομίας. Όταν μία πρόταση ψ παράγεται από μία πρόταση ϕ , η πληροφοριακή τιμή της ϕ είναι μεγαλύτερη από αυτή της ψ . Ωστόσο, όταν απορρίπτουμε την ϕ , στη νέα κατάσταση πεποιθήσεων δεν απουσιάζει ολόκληρη η πληροφορία που σχετίζεται με την ϕ και μερικές φορές το ψ διατηρείται. Εάν όμως απορρίπταμε το ψ , τότε λόγω της κλειστής συνεπαγωγής του συνόλου των πεποιθήσεων και μέσω της πρότασης ϕ θα είχαμε την επανεμφάνιση της πρότασης ψ . Άρα θα έπρεπε να απορριφθεί και η πρόταση ϕ . Η παραπάνω σκέψη οδηγεί σε μεγαλύτερη έλλειψη πληροφορίας στο σύνολο πεποιθήσεων K . Παρόλο που, ο εν λόγω ισχυρισμός αποδεικνύει ότι το αίτημα (EE2) συμμορφώνεται με το κριτήριο της πληροφοριακής οικονομίας. Γενικά το αίτημα (EE2) φαίνεται να περιλαμβάνει όλα όσα μπορούν να αξιωματικοποιηθούν ‘‘ποιοτικά’’ για τη γνωσιακή κατοχύρωση. Ένα ακόμη πιο ισχυρό αίτημα είναι το αίτημα της **Συνδεδεσιμότητας**:

(EE3) Για όλες τις προτάσεις $\phi, \psi \in K$, $\phi \leq \phi \wedge \psi$ ή $\psi \leq \phi \wedge \psi$
(Conjunctiveness)

Το παραπάνω αίτημα υπαγορεύει ότι, εάν επιθυμούμε να απορρίψουμε το $\phi \wedge \psi$ από ένα σύνολο πεποιθήσεων K , αυτό επιτυγχάνεται μόνο εάν απορρίψουμε είτε το ϕ είτε το ψ . Με λίγα λόγια η έλλειψη της πληροφορίας προκαλείται είτε από την απόρριψη της πρότασης ϕ είτε της ψ . Επίσης από το (EE2) επειδή $\phi \wedge \psi \rightarrow \phi$ και $\phi \wedge \psi \rightarrow \psi$ προκύπτει ακριβώς το αντίθετο, δηλαδή, $\phi \wedge \psi \leq \phi$ και $\phi \wedge \psi \leq \psi$.

Ένας διαφορετικός τρόπος που αποτελεί και κίνητρο για τα αξιώματα γνωσιακής κατοχύρωσης και συγκεκριμένα για το (EE3) είναι η εκτίμηση της \leq μέσω των βαθμών σταθερότητας του Spohn. Πιο συγκεκριμένα ο Spohn θεωρεί ότι μία πρόταση ϕ είναι λιγότερο κατοχυρωμένη από μία άλλη πρόταση ψ , εάν

η ϕ έχει μικρότερο βαθμό σταθερότητας από τη ψ . Η πρόταση του Spohn διατυπώνεται από την επόμενη συνθήκη:

$$4.1 \quad \phi \leq \psi \text{ εάν και μόνο εάν } k(\neg\phi) \leq k(\neg\psi)$$

Εδώ το K ορίζεται ως συνάρτηση υποθετικής διάταξης (ordinal conditional function). Εάν λοιπόν $\eta \leq$ οριστεί κατ' αυτό τον τρόπο τότε θα ικανοποιεί και τα αξιώματα (EE1) - (EE3).

Λήμμα 4.8 Εάν $\eta \leq$ ικανοποιεί τα αξιώματα (EE1)-(EE3), τότε για κάθε ϕ και ψ προκύπτει ότι είτε $\phi \leq \psi$ είτε $\psi \leq \phi$

Κατά την ερμηνεία των αξιωμάτων υποθέσαμε ότι όλες οι προτάσεις ανήκουν στο υπ' εξέταση σύνολο πεποιθήσεων K . Ωστόσο, θα μπορούσαμε να ορίσουμε το \leq για όλες τις προτάσεις που ανήκουν στη γλώσσα L . Εάν μία πρόταση δεν ανήκει στο K , τότε δεν είναι κατοχυρωμένη και συνεπώς αποτελεί το μικρότερο στοιχείο της διάταξης \leq . Όλα τα παραπάνω διατυπώνονται μέσω του αξιώματος της **Ελαχιστότητας**:

(EE4) Όταν $K \neq K_{\perp}$ τότε $\phi \notin K$ εάν και μόνο εάν $\phi \leq \psi$ για όλες τις προτάσεις $\psi \in K$. (Minimality)

Το αίτημα αυτό προκύπτει επίσης και από το Λήμμα 4.8 διότι εάν $\phi \notin K$ τότε $K(\neg\phi) = 0$

Αντιστρόφως οι πιο 'κατοχυρωμένες' προτάσεις είναι αυτές που είναι λογικά έγκυρες (ταυτολογίες) και αποτελούν το μέγιστο στοιχείο στη διάταξη \leq :

(EE5) Εάν $\psi \leq \phi$ για όλες τις προτάσεις ψ τότε $\vdash \phi$. (Maximality)

Το αντίστροφο του αξιώματος (EE5) προκύπτει από το (EE2) διότι, εάν ϕ τότε $\psi \rightarrow \phi$ για όλες τις προτάσεις ψ . Τέλος, μερικές συνέπειες για τα αξιώματα (EE1) - (EE5) είναι οι εξής:

Λήμμα 4.9 Έστω ότι η διάταξη \leq ικανοποιεί το (EE1) και τη συνθήκη 4.8. Επιπρόσθετα ικανοποιεί το (EE3) εάν και μόνο εάν για κάθε πρόταση ϕ , ψ και $\gamma \in K$ εάν $\gamma \leq \phi$ και $\gamma \leq \psi$ τότε $\gamma \leq \phi \wedge \psi$

Λήμμα 4.10 Έστω ότι η διάταξη \leq ικανοποιεί τα (EE1)-(EE3). Τότε για όλες τις προτάσεις ϕ , ψ και $\gamma \in K$ εάν $\psi \wedge \gamma \leq \phi$ τότε $\psi \leq \phi$ ή $\gamma \leq \phi$.

Λήμμα 4.11 Έστω ότι η διάταξη \leq ικανοποιεί τα (EE1)-(EE3). Τότε $\phi < \psi$ εάν και μόνο εάν $\phi \wedge \psi < \psi$

Πίνακας 4: Αξιώματα γνωσιακής κατοχύρωσης

(EE1) Για κάθε ϕ, ψ, γ εάν $\phi \leq \psi$ και $\psi \leq \gamma$ τότε $\phi \leq \gamma$
(EE2) Για κάθε ϕ, ψ εάν $\phi \vdash \psi$ τότε $\phi \leq \psi$
(EE3) Για όλες τις προτάσεις $\phi, \psi \in K$, $\phi \leq \phi \wedge \psi$ ή $\psi \leq \phi \wedge \psi$
(EE4) Όταν $K \neq K_{\perp}$ τότε $\phi \notin K$ εάν και μόνο εάν $\phi \leq \psi$ για όλες τις προτάσεις ψ
(EE5) Εάν $\psi \leq \phi$ για όλες τις προτάσεις ψ τότε ϕ είναι ταυτολογία

Αφού παρουσιάσαμε τα πέντε αξιώματα, το επόμενο βήμα είναι να συνδέσουμε τη σειρά γνωσιακής κατοχύρωσης με την συνάρτηση αφαίρεσης πεποιθήσεων. Υπάρχουν δύο τρόποι να γίνει αυτό μπορούμε να : α) δεδομένου της συνάρτησης αφαίρεσης πεποιθήσεων να προσδιορίσουμε την γνωσιακή κατοχύρωση και β) δεδομένου της γνωσιακής κατοχύρωσης να προσδιορίσουμε την συνάρτηση αφαίρεσης πεποιθήσεων.

Για να ξεκινήσουμε κατασκευάζουμε μία σειρά γνωσιακής κατοχύρωσης από μία συνάρτηση αφαίρεσης. Η οδηγός αρχή αυτής της κατασκευής είναι : $\phi \leq \psi$ θα ισχύει αν και μόνο αν, σε περίπτωση που θέλουμε να αφαιρέσουμε το ϕ ή το ψ , θα αφαιρέσουμε το ϕ . Η περίπτωση να αφαιρέσουμε το ϕ και το ψ , είναι όταν θέλουμε να αφαιρέσουμε την σύνδεση $\phi \& \psi$.

Όμως υπάρχει μια εξαίρεση στον κανόνα, όταν το $\phi \& \psi$ είναι ταυτολογία , τότε και το ϕ και το ψ είναι ταυτολογίες. Επομένως είναι και τα δύο στοιχεία του συνόλου $K \div (\phi \& \psi)$, άρα το $\phi \notin K \div (\phi \& \psi)$ και $\psi \notin K \div (\phi \& \psi)$. Έτσι κάνουμε την παραδοχή ότι $\alpha \equiv \beta$, και ότι θα ισχύει $\phi \leq \psi$, όταν $\phi \& \psi$ είναι ταυτολογία. Επομένως φτάνουμε στον παρακάτω ορισμό:

(C \leq) $\phi \leq \psi$ αν και μόνο αν $\phi \notin K \div (\phi \& \psi)$ ή $\phi \& \psi \in C_n(0)$.

Τώρα, ας προσδιορίσουμε την συνάρτηση αφαίρεσης δεδομένου της σειράς γνωσιακής κατοχύρωσης, το οποίο έχει πιο πολύ ενδιαφέρον. Ο Gardenfors προτείνει τον παρακάτω ορισμό:

(G ÷) $\psi \in K \div \phi$ αν και μόνο αν $\psi \in K$ και ή $\phi < (\phi \vee \psi)$ ή $\phi \in Cn(0)$.

(Gardenfors' s entrenchment – based contraction)

Η σχέση αυτή θα ισχύει μόνο εάν η συνάρτηση αφαίρεσης ικανοποιεί το αίτημα (K ÷ 5) Εάν $\phi \in K$ τότε $K \subseteq (K \div \phi) + \phi$ (recovery postulate).

Απο την άλλη πλευρά ο Hans Rott , μετά από έρευνα έδειξε το παρακάτω ορισμό:

(R ÷) $\psi \in K \div \phi$ αν και μόνο αν $\psi \in K$ και ή $\phi < \psi$ ή $\phi \in Cn(0)$.

(Rott' s entrenchment – based contraction)

Όταν αφαιρούμε μία πεποίθηση που δεν είναι ταυτολογία από ένα σύνολο πεποιθήσεων, τότε θα διατηρούμε ακριβώς τις πεποιθήσεις αυτές που είναι πιο κατοχυρωμένες από το ϕ .

Επομένως , παρατηρούμε ότι η σχέση που προτείνει ο Gardenfors ικανοποιεί και τα οχτώ AGM αξιώματα (K ÷ 1) - (K ÷ 8) , ενώ η σχέση που προτείνει ο Rott ικανοποιεί όλα τα αξιώματα εκτό από το (K ÷ 5) recovery postulate.

Παρατηρούμε λοιπόν ότι το πέμπτο αίτημα και και το πιο αμφιλεγόμενο από τα αξιώματα του Gardenfors. Με μια πρώτη ματιά, ο ορισμός του Rott φαίνεται πιο αρεστός σε σχέση με του Gardenfors, αλλά δυστυχώς δεν ικανοποιεί την αρχή της Αποβολής (Explusiveness) :

Εάν το ϕ , ψ δεν είναι ταυτολογίες τότε $\phi \notin K \div \psi$ και $\psi \notin K \div \phi$.

Σύμφωνα με την παραπάνω αρχή της αποβολής, για οποιανδήποτε δύο μη ταυτολογικές πεποιθήσεις ϕ , ψ σε ένα σύνολο πεποιθήσεων, είτε το ϕ διώχνεται όταν το ψ αποβάλλεται ή τα αντίστροφο. Αυτό είναι ένα μεγάλο μειονέκτημα της συνάρτησης αφαίρεσης γιατί δεν επιτρέπει άσχετες μεταξύ τους πεποιθήσεις να διατηρούνται ανενόχλητες από την αποβολή του ενός ή του άλλου.

Για παράδειγμα [3] σελ. 102, έστω ότι πιστεύω ότι έχω παρκάρει το αυτοκίνητό μου μπροστά από το σπίτι (ϕ) και επίσης πιστεύω ότι ο Σαιξπηρ έγραψε το Tempest (ψ). Εάν πάψω να πιστεύω στην πρώτη πεποίθηση θα

εξακολουθώ να πιστεύω την δεύτερη; Ή εάν πάψω να πιστεύω στην δεύτερη πεποίθηση θα εξακολουθώ να πιστεύω την πρώτη;

Ο Lindstrom και ο Rabinowicz [3] σελ. 103, πρότειναν μια πιο ρεαλιστική κατασκευή της συνάρτησης αφαίρεσης βάσει την γνωσιακής κατοχύρωσης, η οποία τοποθετείται μεταξύ των Gardenfors και Rott, για όλες τις πεποιθήσεις $\varphi \div$ και παρουσιάζεται παρακάτω:

$$K \div_R \Phi \subseteq K \div \varphi \subseteq K \div_G \Phi$$

Παρατήρηση : Έστω \div μία συνάρτηση αφαίρεσης για ένα σύνολο πεποιθήσεων K και \leq η σχέση που προσδιορίζεται από την \div μέσω της $(C \leq)$. Ακόμα έστω ότι \div_G και \div_R οι συναρτήσεις αφαίρεσης των Gardenfors και Rott τότε :

1. Εάν η \div ικανοποιεί τα αξιώματα inclusion, closure, success και extensionality τότε $K \div \varphi \subseteq K \div_G \varphi$, για όλες τις πεποιθήσεις φ .
2. Εάν η \div ικανοποιεί τα αξιώματα closure, failure και conjunctive inclusion τότε $K \div_R \Phi \subseteq K \div \varphi$, για όλες τις πεποιθήσεις φ .

Η γνωσιακή κατοχύρωση προσδιορίστηκε για σύνολα πεποιθήσεων και όχι για βάσεις πεποιθήσεων. Δεν μπορούμε να εφαρμόσουμε τα ίδια και για τις βάσεις πεποιθήσεων. Ένας από τους λόγους το φιλοσοφικό πρόβλημα που εμφανίζεται όχι μόνο στην γνωσιακή κατοχύρωση αλλά σε όλους τους μηχανισμούς επιλογής, οι οποίοι βασίζονται σε μία σχέση ανάμεσα στα στοιχεία της βάσης. Θεωρούμε μία βάση πεποιθήσεων $K = \{ \alpha, \beta, \delta \}$. Μία επιλογή μεταξύ του β και του δ θα εξαρτηθεί από το εάν ή όχι το α θα παραμείνει. Για παράδειγμα $\{ \alpha, \beta \}$ μπορεί να είναι επιστημονικά προτιμότερα από τα $\{ \alpha, \delta \}$, αν και το $\{ \delta \}$ είναι επιστημονικά προτιμότερο από το $\{ \beta \}$. Αυτό μπορεί εύκολα να εκφραστεί με μία σχέση μεταξύ των υποσυνόλων του K , αλλά δεν μπορεί να εκφραστεί με μία σχέση των στοιχείων του K .

Αυτό το πρόβλημα δεν εμφανίζεται στους μηχανισμούς επιλογής, οι οποίοι λειτουργούν σε σύνολα πεποιθήσεων κλειστά ως λογική συνεπαγωγή. Θεωρούμε τέτοια σύνολα $C_n(K)$, τα οποία περιέχουν συνδέσεις όπως $\varphi \& \psi$ και $\varphi \& \gamma$ που μπορούν να εκφράσουν ότι δεν μπορεί η σχέση μεταξύ των στοιχείων του K . Μπορούμε λοιπόν να έχουμε μία σειρά γνωσιακής κατοχύρωσης για το $C_n(K)$, όπως $\psi < \gamma$ και $\varphi \& \gamma < \varphi \& \psi$.

Η πιο επιτυχημένη εφαρμογή για τις βάσεις πεποιθήσεων σχετικά με την γνωσιακή κατοχύρωση είναι η θεωρία των σχέσεων που θέτουμε (ensconcement relations), η οποία αναπτύχθηκε από την Mary – Anne Williams.

Προσδιορίζει την παραπάνω σχέση ως (\leq) σε ένα σύνολο πεποιθήσεων K , ως μεταβατική και συνδετική σχέση που ικανοποιεί τις παρακάτω τρεις ιδιότητες:

1. Εάν $\beta \in K \setminus Cn(0)$, τότε $\{ \varphi \in K \mid \psi \ll \varphi \}$ και το ψ δεν είναι ταυτολογία
2. Εάν το φ δεν είναι ταυτολογία και το ψ είναι ταυτολογία τότε $\varphi \ll \psi$
3. εάν το φ και το ψ είναι ταυτολογίες τότε $\varphi \leq \psi$

Σύμφωνα με την πρώτη υπόθεση οι θεσμοί οι οποίοι θέτονται είναι πιο ισχυροί από το φ , όταν δεν περιέχουν το φ . Σύμφωνα με τις δύο τελευταίες υποθέσεις, εάν υπάρχουν ταυτολογίες σε ένα σύνολο πεποιθήσεων K τότε είναι οι πιο ισχυρές πεποιθήσεις.

Η Mary – Anne Williams έδειξε ότι για κάθε ensconcement – based contraction σε ένα σύνολο πεποιθήσεων K , υπάρχει μία σειρά γνωσιακής κατοχύρωσης αφαίρεσης \div_G που αντιστοιχεί σε ένα σύνολο πεποιθήσεων $Cn(K)$ για όλες τις πεποιθήσεις φ :

$$K \div_W \varphi = (Cn(K) \div_G \varphi) \cap K$$

Με άλλα λόγια, μία φόρμουλα παραμένει μετά από μια ensconcement – based contraction εάν και μόνο εάν είναι μέλος της βάσης πεποιθήσεων και θα έχει παραμείνει μετά από την corresponding ensconcement – based contraction.

ΚΕΦΑΛΑΙΟ 5

5.1 Συμπεράσματα

Η αλλαγή πεποιθήσεων είναι το πεδίο έρευνας που ασχολείται με τη διατύπωση και τυποποίηση των διαδικασιών ορθολογικής διαχείρισης νέων πληροφοριών. Η τυποποίηση αυτή ενέταξε την αλλαγή πεποιθήσεων στον ευρύτερο χώρο της

αναπαράστασης γνώσης, προσελκύνοντας στους κόλπους της όλο και περισσότερους ερευνητές.

Οι ρίζες της αλλαγής των πεπιοιθήσεων εντοπίζονται από διάφορα φιλοσοφικά-γνωστικά πεδία, όμως η εξέλιξή της επικουρείται από ένα ευρύτερο πλαίσιο επιστημών. Εκ πρώτης άποψης, οι επιστήμες που φαίνεται να συνεισφέρουν στην αλλαγή πεπιοιθήσεων είναι αυτές των μαθηματικών, των υπολογιστών και της φιλοσοφίας. Από την άλλη πλευρά η αλλαγή των πεπιοιθήσεων δεν έχει μοναδικό κίνητρο την περαιτέρω επέκτασή της ως ανεξάρτητο πεδίο έρευνας και προβληματισμού. Αντιθέτως θα μπορούσε να ερμηνευθεί ως ένα μεταβατικό και συνάμα συνδετικό στάδιο για τη συνεισφορά σε διαφορετικές επιστημονικές περιοχές.

Επειδή λοιπόν η αλλαγή των πεπιοιθήσεων κατατάσσει, ομαδοποιεί και συστηματοποιεί τη ροή των πληροφοριών που μπορεί να δεχθεί ένας ορθολογικός νους, ο κυριότερος όγκος της βιβλιογραφίας προσανατολίζεται στην αντιμετώπιση διάφορων δυσεπίλυτων προβλημάτων. Τα προβλήματα αυτά εμφανίζονται κυρίως σε τομείς υπολογιστικών εφαρμογών, όπως οι βάσεις δεδομένων, η ρομποτική, η διάγνωση σφαλμάτων, αλλά και σε αρκετές οικονομικές επιστήμες, όπως η θεωρία παιγνίων, το μάρκετινγκ, οικονομικές διαπραγματεύσεις κ.ο.κ.

Κατά την έρευνά μας χρειάστηκε η συνδρομή της επιστήμης των μαθηματικών και πιο συγκεκριμένα της τυπικής λογικής για την ανάπτυξη και θεμελίωση της αναθεώρησης πεπιοιθήσεων.

Τι έχει πραγματοποιηθεί μέχρι τώρα? Οι ερευνητές στη δυναμική των πεπιοιθήσεων έχουν αναπτύξει αρκετούς τύπους διαδικασιών, οι οποίοι χρησιμοποιούνται για να εκτελεστούν αλλαγές στα σύνολα πεπιοιθήσεων ή στις βάσεις πεπιοιθήσεων. Αρκετά αξιώματα έχουν προταθεί, τα οποία χρησιμοποιούνται για να χαρακτηρίσουν αυτές τις διαδικασίες. Παρόλα αυτά δεν είμαστε σε θέση να απαντήσουμε στην ερώτηση τι σημαίνει η λογική αναθεώρηση των πεπιοιθήσεων ενός ατόμου. Όμως είμαστε σε θέση να ξαναθέσουμε την ερώτηση με περισσότερη ακρίβεια. Ειδικότερα, μπορούμε να ξεχωρίσουμε πολλές παραμέτρους, τις οποίες μπορούμε να μελετήσουμε ξεχωριστά τη κάθε μία. Αυτό δεν είναι μικρό επίτευμα αλλά χρίζει περαιτέρω μελέτη και έρευνα.

Από φιλοσοφική άποψη, τα επιστημονολογικά ζητήματα μπορούν πλέον να μελετηθούν από νέα οπτική γωνία. Ο διαχωρισμός ανάμεσα στην συνέπεια και στην ιδρυματική επιστημονολογία μπορεί τώρα να εκφραστεί με έναν πιο ακριβή τρόπο. Οι σχέσεις μεταξύ των τριών ιδεών: της επιστημονικής σημασίας (αξίας), της εθελοντικής αναθεώρησης και των πιθανοτήτων είναι αναγκαίο να ξεκαθαριστούν.

Οι τυπικές μελέτες στη δυναμική των πεποιθήσεων ανέπτυξαν ζητήματα φιλοσοφικής σημασίας: πρέπει η λογική αφαίρεση πεποιθήσεων να ικανοποιεί το αξίωμα ανάκτησης? (postulate of recovery). Πρέπει σε μια πεποίθηση που αντιφάσκει, να προστίθεται πρώτα η νέα πληροφορία και μετά να αφαιρούνται οι παλαιές πληροφορίες, οι οποίες αντιφάσκουν? Ή οι δύο παραπάνω υπολειπουργίες να πραγματοποιούνται ανάποδα? Ποια είναι η σχέση μεταξύ της πληροφορίας και της γνωσιακής κατοχύρωσης της πληροφορίας?

Από την πλευρά των επιστημόνων των υπολογιστών το πιο φλέγον ζήτημα θα πρέπει να είναι: Μέχρι πιο σημείο τα λογικά αξιώματα ικανοποιούνται από διαφορετικά μοντέλα αναθεώρησης βάσεων δεδομένων. Τα αξιώματα που έχουν προταθεί τα τελευταία χρόνια, μας παρέχουν αρκετές πληροφορίες και εργαλεία για να χαρακτηρίσουμε τα γνωσιακά ενημερωμένα συστήματα. Μελέτες σε νευρωνικά δίκτυα έδειξαν ότι μερικά από αυτά τα δίκτυα ικανοποιούν τις βασικές αρχές της αναθεώρησης πεποιθήσεων. Γενικότερα μέσω αυτών των νέων δυνατοτήτων ο χαρακτηρισμός των ενημερωμένων συστημάτων παραμένει ανεξερεύνητος.

Από ψυχολογική άποψη, το πιο σημαντικό ζήτημα θα πρέπει να είναι: κάτω από ποιες συνθήκες μπορεί η συμπεριφορά των ανθρώπων να συμβαδίζει με τα αξιώματα αναθεώρησης πεποιθήσεων? Είναι αποδεκτό ότι οι νόμοι της ανθρώπινης λογικής μπορούν να τοποθετηθούν ανάμεσα στα αξιώματα της δυναμικής πεποιθήσεων. Παρόλα αυτά είναι δυνατό να βρούμε συνθήκες στις οποίες τα αξιώματα ικανοποιούνται και συνθήκες κάτω από τις οποίες τα αξιώματα δεν ικανοποιούνται. Ο διαχωρισμός μεταξύ εσωτερικής και εξωτερικής αναθεώρησης είναι επίσης χρήσιμος στην ψυχολογία.

Δεν έχουμε κανέναν λόγο να περιμένουμε συνεχιζόμενες διαφοροποιήσεις στην δυναμική πεποιθήσεων για αναθεώρηση. Δεδομένου ότι όλα τα κατάλληλα μοντέλα αναθεώρησης πεποιθήσεων φαίνονται στον ορίζοντα, αυτό το πεδίο

έρευνας πρέπει να σημειώσει πρόοδο, αναπτύσσοντας διαφορετικά μοντέλα, το καθένα από αυτά να αντικατροπίζει τα σημαντικότερα χαρακτηριστικά του αντικειμένου της μελέτης.

5.2 Μελλοντικές Κατευθύνσεις

Οι πεπιοθήσεις ενός ορθολογικού υποκειμένου δεν είναι στατικές, αλλά αλλάζουν με το πέρασμα του χρόνου, αφού νέα πληροφορία προστίθεται. Αναθεώρηση πεπιοθήσεων είναι η διαδικασία αλλαγής πεπιοθήσεων με στόχο να κάνουμε αισθητή την παρουσία των νέων πληροφοριών. Ως ορθολογικά υποκείμενα έχουμε ένα πλέγμα πεπιοθήσεων, μέσω του οποίου δημιουργούμε μια ιεράρχηση των αξιών μας. Η αλλαγή πεπιοθήσεων επέρχεται από τη στιγμή που θα αλλάξουμε αυτή την ιεραρχία πεπιοθήσεων, από τη στιγμή δηλαδή που θα επέμβουμε στην ιεράρχηση των επιθυμιών. Παρατηρείται τότε μια αλλαγή συμπεριφοράς, μια ορθολογική ανασυγκρότηση του συνόλου των πεπιοθήσεων. Λέγοντας ορθολογική εννοούμε αλλάζοντας τις επιθυμίες έτσι ώστε να ταιριάζουν με τις αλλαγές των πεπιοθήσεων.

Ένας φυσικός προβληματισμός της επιστημολογίας είναι να προτείνει μια θεώρηση δικαιολογημένης αναθεώρησης πεπιοθήσεων. Υπάρχουν δύο δρόμοι που μπορεί να ακολουθήσει κανείς σε μια τέτοια περίπτωση. Η "αξιωματική" θεώρηση, η οποία προτείνει αφηρημένες γενικές αρχές που νοηματικά επενεργούν πάνω στη διαδικασία και αξιωματικοποιούν την αναθεώρηση πεπιοθήσεων. Ο δεύτερος δρόμος είναι η "παραγωγική" θεώρηση, η οποία προσπαθεί να αντλήσει μια θεωρία αναθεώρησης πεπιοθήσεων από μια πιο εποικοδομητική επιστημολογική θεωρία.

Η πιο δημοφιλής και ευρέως διαδεδομένη αξιωματική θεώρηση είναι η AGM θεώρηση, η οποία, όπως είδαμε είναι κατασκευάσμο των Alchourrón, C. E., P. Gärdenfors, και D. Makinson (1985) και αναπτύχθηκε περαιτέρω από τον P. Gärdenfors (1988, 1992) και τους Alchourrón και Makinson (1982, 1985).

Τέτοιου είδους θεωρήσεις προχωρούν σε ένα πολύ μεγάλο επίπεδο αφαίρεσης, αγνοώντας όψεις γνωσιακής ορθολογικότητας, χωρίς την οποία είναι αδύνατο

να δημιουργήσουμε πραγματικές αρχές ορθολογικής αναθεώρησης πεπιοθήσεων. Μια θεωρία αναθεώρησης πεπιοθήσεων οφείλει να καθοδηγείται από μια πιο εποικοδομητική επιστημολογία.

Η θεωρία αναθεώρησης πεπιοθήσεων δημιουργήθηκε κατ' αρχήν για να απαντήσει στο ερώτημα πως οι άνθρωποι αλλάζουν γνώμη, ακολουθώντας μια μακρά παράδοση φιλοσοφικού έργου στη θεωρία της αλλαγής πεπιοθήσεων. Κατά τις δύο τελευταίες δεκαετίες η πρωταρχική αξιωματική θεώρηση έχει συζητηθεί, ανασκευαστεί και δεχτεί εκτενέστατη κριτική. Άλλα πιο ρεαλιστικά, και πιο εύκολα υπολογιστικά μοντέλα αναθεώρησης πεπιοθήσεων έχουν προταθεί. Κανονιστικές θεωρίες αναθεώρησης, οι οποίες υποστηρίζονται από φιλοσόφους και Λογικούς (που μελετάνε το ερώτημα του πως θα έπρεπε οι άνθρωποι να αλλάζουν γνώμες), βαδίζουν παράλληλα με πιο περιγραφικές θεωρήσεις που δημιουργήθηκαν από γνωσιακούς ψυχολόγους και κοινωνικούς επιστήμονες (οι οποίοι θέλουν να ξέρουν πως οι άνθρωποι στη πραγματικότητα λειτουργούν). Ως αποτέλεσμα, η αναθεώρηση πεπιοθήσεων έχει γίνει μια από τις πιο ενεργές περιοχές έρευνας, με τη συμβολή μεταξύ γνωσιακής επιστήμης, λογικής, φιλοσοφίας και της τεχνητής νοημοσύνης.

Μια κανονιστική θεώρηση μας λει για το πώς οι άνθρωποι θα έπρεπε να σκέπτονται, ενώ αντιθέτως μια περιγραφική θεώρηση μας λει πως πραγματικά σκέπτονται. Υπάρχουν προσεγγίσεις που είτε προσπαθούν να πετύχουν το ένα από τα δύο είτε και ακόμα να συνδυάσουν αυτές τις δυο διαφορετικές όψεις του ίδιου νομίσματος. Στην πραγματικότητα περιγραφικές και κανονιστικές θεωρήσεις συνδέονται στενά. Για ένα πράγμα είναι δύσκολο να καταλήξουμε σε πειστικές κανονιστικές αρχές εκτός αν λάβουμε υπόψιν μας πως στην πραγματικότητα οι άνθρωποι σκέπτονται, κάτι το οποίο ανήκει στη δικαιοδοσία μιας περιγραφικής θεώρησης. Από την άλλη φαίνεται ότι κάθε περιγραφική θεώρηση πρέπει να συνεπάγεται ένα συγκεκριμένο είδος εξιδανίκευσης, και η εξιδανίκευση είναι κανονιστική ως ένα βαθμό.

Εντούτοις, τα βασικά προβλήματα της αναθεώρησης πεπιοθήσεων παραμένουν άλυτα- και ο βασικός λόγος που γίνεται αυτό είναι η έλλειψη ενότητας (ενοποίησης) μεταξύ της "τυπικής" πλευράς και της "γνωσιακής" πλευράς του προβλήματος. Ως αποτέλεσμα πολλά ερωτήματα είναι ακόμα αναπάντητα. Ένα τέτοιο ερώτημα για παράδειγμα είναι εάν η αξιωματική θεώρηση είναι αρκετά

εκφραστική για να συλλάβει όλα τα σχετικά γνωσιακά θέματα της αναθεώρησης πεπιοθήσεων. Επίσης ποια είναι η ακριβής σχέση μεταξύ αναθεώρησης πεπιοθήσεων και πληροφοριακής αναβάθμισης; Πως επηρεάζεται η αναθεώρηση πεπιοθήσεων από την επικοινωνία, και πιο γενικά από την κοινωνική αλληλεπίδραση; Για να απαντηθούν αυτά τα ερωτήματα, η μελέτη της αναθεώρησης πεπιοθήσεων πρέπει να προσανατολιστεί σε έναν πιο ενοποιημένο σκελετό, φέροντας μαζί διαφορετικές προσεγγίσεις από διαφορετικούς τομείς γνώσης

Παράρτημα : Απόδειξη Θεωρήματος 3.3

Θεώρημα 3.3 (Alchourron, C.E., P. Gärdenfors and D. Makinson. [2])
Υποθέτουμε ότι οι προϋποθέσεις του θεωρήματος 3.2 ικανοποιούνται. Τότε (α) εάν το $(K \div 7)$ ικανοποιείται, το $(K*7)$ ικανοποιείται για την ορισθείσα συνάρτηση αναθεώρησης και (β) εάν το $(K \div 8)$ ικανοποιείται, το $(K*8)$ ικανοποιείται και αυτό για την ορισθείσα συνάρτηση αναθεώρησης.

Απόδειξη:

A) Ισχύουν τα αξιώματα $(K \div 1) - (K \div 4)$, $(K \div 6)$ και $(K*1) - (K*6)$. Επίσης ισχύει το $(K \div 7) : (K \div \phi) \cap (K \div \psi) \subseteq K \div (\phi \wedge \psi)$

Θα αποδείξουμε ότι $(K*7) : K * (\phi \wedge \psi) \subseteq (K*\phi) + \psi$.

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \subseteq [(K \div \neg \phi) + \phi] + \psi \rightarrow$$

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \subseteq [(K \div \neg \phi) + (\phi \wedge \psi)] \rightarrow$$

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \subseteq [(K \div (\neg(\phi \wedge \psi) \wedge (\phi \rightarrow \psi))) + (\phi \wedge \psi)] \rightarrow$$

Χρησιμοποιώ το $(K \div 7) : (K \div \phi) \cap (K \div \psi) \subseteq K \div (\phi \wedge \psi)$ και αρκεί να αποδείξω :

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \subseteq [(K \div \neg(\phi \wedge \psi) \cap K \div (\phi \rightarrow \psi)) + (\phi \wedge \psi)]$$

επειδή

$$(K \div \neg(\phi \wedge \psi) \cap K \div (\phi \rightarrow \psi)) + (\phi \wedge \psi) \subseteq (K \div (\neg(\phi \wedge \psi) \wedge (\phi \rightarrow \psi))) + (\phi \wedge \psi)$$

από $(K \div 7)$. Από 3.9 μπορώ να γράψω:

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \subseteq$$

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \cap [(K \div (\phi \rightarrow \psi)) + (\phi \wedge \psi)] \rightarrow$$

$$[K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi) \subseteq (K*(\phi \wedge \psi)) \cap [(K \div (\phi \rightarrow \psi)) + (\phi \wedge \psi)]$$

Θέτω $RHS (1) = (K*(\phi \wedge \psi))$ και $RHS (2) = [(K \div (\phi \rightarrow \psi)) + (\phi \wedge \psi)]$

Έστω $r \in [K \div \neg(\phi \wedge \psi)] + (\phi \wedge \psi)$, επόμενως το r ανήκει και στο RHS (1) γιατί είναι το ίδιο.

Αρκεί να δείξω ότι το r ανήκει και στο RHS (2) ή ότι

$$[K \div \neg(\varphi \wedge \psi)] + (\varphi \wedge \psi) \subseteq [K \div (\varphi \rightarrow \psi)] + (\varphi \wedge \psi)$$

όμως $(\varphi \rightarrow \psi) \equiv \neg\varphi \vee \psi$ και ισχύει ότι

$(\varphi \wedge \psi) \vdash \psi$ και το $\psi \vdash \neg\varphi \vee \psi$, άρα

$(\varphi \wedge \psi) \vdash \neg\varphi \vee \psi$ και $(\varphi \wedge \psi) \vdash (\varphi \rightarrow \psi)$

Προκύπτει πως

$r \in \text{Cn}(K \cup (\varphi \wedge \psi))$, αρκεί να δείξω ότι

$$K \cup (\varphi \wedge \psi) \subseteq \text{Cn}(K \div (\neg\varphi \vee \psi) \cup (\varphi \wedge \psi))$$

Επομένως, $t \in K \Rightarrow t \in \text{RHS} (2)$.

B) Το K είναι μία θεωρία και ισχύει το $(K \div 8)$:

Εάν $\phi \notin K \div (\phi \wedge \psi)$, τότε $K \div (\phi \wedge \psi) \subseteq K \div \phi$ για όλα τα ϕ και ψ .

Θα δείξουμε ότι ισχύει το **$(K * 8)$** : **Εάν $\neg\psi \notin K * \phi$ τότε $(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$**

Παρατηρούμε ότι $\neg\phi \equiv (\neg\phi \vee \neg\psi) \wedge \neg\phi$

Άρα $K \div \phi = K \div ((\neg\phi \vee \neg\psi) \wedge \neg\phi)$

Πάμε στο $(K * 8)$ έστω $\neg\psi \notin K * \phi = (K \div \neg\phi) + \phi = \text{Cn}(K \div \neg\phi) \cup \{\phi\}$

Προκύπτει πως $(\neg\phi \vee \neg\psi) \notin K \div \neg\phi$

Χρησιμοποιούμε το $(K \div 8)$ και έχουμε:

$$K \div \neg\phi = K \div ((\neg\phi \vee \neg\psi) \wedge \neg\phi) \subseteq K \div ((\neg\phi \vee \neg\psi)) = K \div \neg(\varphi \wedge \psi)$$

Και έχουμε $\text{Cn}((K * \phi) \cup \{\psi\}) = \text{Cn}(\text{Cn}[K \div \neg\phi] \cup \{\phi\}) \cup \{\psi\} =$

$$\text{Cn}[K \div \neg\phi] \cup \{\phi \wedge \psi\} \subseteq \text{Cn}((K \div \neg(\varphi \wedge \psi)) \cup (\varphi \wedge \psi)) = K * (\varphi \wedge \psi)$$

Γ) Υποθέτουμε ότι ισχύει το $(K * 7)$: $K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$ για κάθε ϕ και ψ .

Θέλουμε να δείξουμε ότι ισχύει το **$(K \div 7)$** : **$K \div \phi \cap K \div \psi \subseteq K \div (\phi \wedge \psi)$**

Έστω ότι $\alpha \in K \div \phi \cap K \div \psi$

Επειδή $\phi \equiv \neg(\neg\phi \vee \neg\psi) \wedge \neg\phi$ τότε

$$\alpha \in K \div \neg(\neg\phi \vee \neg\psi) \wedge \neg\phi \subseteq K * (\neg\phi \vee \neg\psi) \wedge \neg\phi \subseteq$$

$$\text{Cn} ((K * (\neg \phi \vee \neg \psi) \cup \{\neg \phi\})$$

$$\text{Επομένως} \quad \alpha \in \text{Cn} ((K * (\neg \phi \vee \neg \psi) \cup \{\neg \phi\})$$

$$\text{Οπότε} \quad \alpha \in \text{Cn} ((K * (\neg \phi \vee \neg \psi) = K * (\neg \phi \vee \neg \psi))$$

$$\text{Cn} ((K \div (\phi \wedge \psi)) + \neg (\phi \wedge \psi)$$

$$\text{Από} (K \div 6) \quad \alpha \in \Psi \vee ((K \div (\phi \wedge \psi)) \cup (\phi \wedge \psi)$$

$$\text{Επομένως} \quad \alpha \in \text{Cn} ((K \div (\phi \wedge \psi)) = K \div (\phi \wedge \psi)$$

(Disjunction of premises)

$$\Delta) \text{ Έστω ότι ισχύει } (K * 8) \text{ και } \neg \psi \notin (K * \phi) \Rightarrow (K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$$

Θα δείξουμε ότι ισχύει $(K \div 8)$:

$$\text{Εάν } \phi \notin K \div (\phi \wedge \psi), \text{ τότε } K \div (\phi \wedge \psi) \subseteq K \div \phi$$

$$\text{Έστω } \phi \notin \text{Cn} (K \div (\phi \wedge \psi))$$

$$\text{Σαφώς} \quad \phi \notin \text{Cn} ((K \div (\phi \wedge \psi)) \cup (\neg \phi \vee \neg \psi)) = K * \neg (\phi \wedge \psi)$$

Εφαρμόζουμε $(K * 8)$

$$\text{Cn} ((K \div \neg (\phi \wedge \psi)) \cup \neg \phi \subseteq K * \neg ((\phi \wedge \psi) + \neg \phi) = K * \neg \phi =$$

$$\text{Cn} ((K \div \neg \phi) \cup \{\neg \phi\}$$

$$\text{Άρα} \quad K \div (\phi \wedge \psi) \subseteq \text{Cn} ((K \div \neg \phi) \cup \{\neg \phi\}$$

$$\text{Επίσης} \quad K \div (\phi \wedge \psi) \subseteq K \subseteq \text{Cn} ((K \div \phi) \cup \{\phi\} \Rightarrow$$

$$K \div (\phi \wedge \psi) \subseteq \text{Cn} ((K \div \phi)) = K \div \phi$$

BIBΛΙΟΓΡΑΦΙΑ

- [1] Peppas, P. 2007. " Belief Revision" , Handbook of Knowledge Representation, F.van Harmelen, V. Lifschitz, and B. Porter (eds), Elsevier.
- [2] Garnefors, P. 1988. " Knowledge in Flux - Modeling the Dynamics of Epistemic States," MIT Press, Cambridge, England.
- [3] Hansson, O.S. 1999. " A Textbook of Belief Dynamics. Theory Change and Database Updating" , Kluwer Academic Publishers.
- [4] Peppas, P. 2004. " The Limit Assumption and Multiple Revision" , Journal of Logic and Computation, pp. 355-371, vol. 14(3), Oxford University Press.
- [5] Alchourron, C.E., P. Gärnefors and D. Makinson. 1985 . " On The logic of theory change: Partial meet functions for contraction and revision." Journal of Symbolic Logic.
- [6] Gärnefors, P. 1992. " Belief Revision" , Cambridge University Press
- [7] Logic of Belief Revision (Stanford Encyclopedia of Philosophy) [2006]
- [8] Φλουρής Γεώργιος, Μεταπτυχιακή Εργασία: Αναθεώρηση Πεποιθήσεων σε Βάσεις Γνώσεων [Σεπτέμβριος 2001] Πανεπιστήμιο Κρήτης.
- [9] Αναστάσιος – Μιχάλης Φωτεινόπουλος, Διδακτορική Διατριβή: Αναπαράσταση Γνώσης – Επεκτάσεις στην Αλλαγή Πεποιθήσεων [Πάτρα 2010]
- [10] Wikipedia, The Free Encyclopedia, Αναθεώρηση Πεποιθήσεων
- [11] Φλουρής Γεώργιος [2006] «Αναθεώρηση Πεποιθήσεων και Εξέλιξη Οντολογιών», Πανεπιστήμιο Κρήτης.