

ΦΙΛΟΣΟΦΙΑ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗ

CHARLES McCARTHY

Πολύ φοβούμαι ότι η εικόνα που το πλατύ κοινό έχει σχηματίσει για τον επιστήμονα που ασχολείται με τους υπολογιστές είναι αυτή του πρακτικού επιστήμονα κι όχι του ψυχολόγου ή του φιλοσόφου. Αυτή η εικόνα όμως οφείλεται σε μία παρανόηση· στην πραγματικότητα είναι όντως πολύ συνηθισμένο για τους επιστήμονες των υπολογιστών, ως μέρος της καθημερινής τους ρουτίνας, να καταπιάνονται με τη λύση φιλοσοφικών προβλημάτων. Αυτό αληθεύει ιδιαίτερα σε τομείς όπως είναι η παράσταση της γνώσης, η επεξεργασία της φυσικής γλώσσας και η σημασιολογία του προγράμματος. Έτσι νομίζω δεν θα προκαλέσει παρεξήγηση το να θέσει ένας επιστήμονας των υπολογιστών δύο επιστημολογικά ερωτήματα. Σκοπεύω δε να τα θέσω μ' ένα καντιανό τρόπο, μολονότι το δεύτερο δεν συμπεριλαμβάνεται στα θέματα που κυρίως απασχόλησαν τον Καντ¹. Το πρώτο ερώτημα έχει πολύ παλιές ρίζες· είναι το εξής: «Πώς η γνώση είναι δυνατή;» Το δεύτερο ερώτημα είναι εξ' ίσου παλαιό, είναι το «Πώς το σφάλμα είναι δυνατόν;».

Επειδή ακόμα κι εκείνοι οι επιστήμονες των υπολογιστών που περισσότερο ασχολούνται με τη φιλοσοφία προτιμούν τη διατήρηση ενός ερασιτεχνικού φιλοσοφικού επιπέδου, γι αυτό κι εγώ θα αποφύγω να δώσω οριστικές απαντήσεις στα παραπάνω ερωτήματα. Παραταύτα θα περιοριστώ στο να κάνω μερικές προτάσεις· αυτές θα υπαγορευθούν από τις έννοιες και τα αποτελέσματα που χρησιμοποιούνται και προσφέρουν αξιόλογες λύσεις στη θεωρητική επιστήμη των υπολογιστών. Ίσως ο όρος «θεωρητική επιστήμη των υπολογιστών» να είναι ανοίκειος· κάποιος θα μπορούσε να τον ερμηνεύσει ως την «a priori επιστήμη των υπολογιστών». Καθώς όμως a priori είναι το τμήμα της γνώσης που κερδίζεται ανεξάρτητα από την εμπειρία, η a priori επιστήμη των υπολογιστών συνίσταται σε ό,τι γνωρίζουμε γι' αυτά και τη λειτουργία τους, ανεξάρτητα από την εμπειρία μας με τις πραγματικές μηχανές των υπολογιστών. Θέτοντας την άποψη αυτή μ' ένα άλλο τρόπο, λέμε ότι η επιστήμη των υπολογιστών είναι μία a priori επιστήμη, όταν μεθοδολογικά είναι ανεξάρτητη από το σχεδιασμό, την ανάλυση ή την έρευνα του πρακτικού μηχανικού εξοπλισμού ή των προγραμμάτων².

Μελετώντας το πρόβλημα της αιτίας και του σκοπού της θεωρητικής επιστήμης των υπολογιστών προσπαθώ μόνος, χωρίς βοήθεια, να τροποποιήσω μία μη θεωρητική προκατάληψη που επικρατεί για τις σχέσεις ανάμεσα στην επιστήμη των υπολογιστών και τη φιλοσοφία. Ορισμένα νέα πορίσματα της επιστήμης των υπολογιστών έχουν σημαντικά επιδράσει στο φιλοσοφικό διάλογο. Αρκεί να αναφέρει κανείς το πανέξυπνο πρόγραμμα του Weizenbaum Eliza και τα προγράμματα του Winston που έχουν την ικανότητα στοιχειώδους εννοιολογικής μορφοποίησης. Όμως αυτού του είδους η επίδραση της επιστήμης των υπολογιστών πηγάζει από τη μη θεωρητική πλευρά του θέματος³. Αυτό που η φιλοσοφία οφείλει ακόμα να προαγάγει είναι η επίδραση των πραγματικά θεωρητικών τομέων. Υπάρχουν λόγοι γι' αυτό· αυτό που κυρίως λέγεται είναι ότι η θεωρητική επιστήμη

των υπολογιστών δεν έχει αναπτυχθεί τόσο, ώστε να μπορεί να έχει κάποια επίδραση σε ο,τιδήποτε. Η «θεωρία» στην επιστήμη των υπολογιστών πέρασε μία δύσκολη περίοδο και η ωριμότητα σ' αυτό τον τομέα έχει προχωρήσει σχετικά αργά. Η επίδρασή της μόνο τώρα αρχίζει να φαίνεται ακόμα και σε θέματα που βρίσκονται πιο κοντά της. Αυτό το κείμενο είναι μία προσπάθεια να εξοικονομηθεί ο χαμένος χρόνος.

Επιστρέφοντας στην πρώτη μου ερώτηση, «Πώς η γνώση είναι δυνατή;», πρέπει να τονίσω ότι αυτή, κατά τη γνώμη μου, ανήκει στην ίδια κατηγορία με πολλές άλλες, παρόμοιες μεταξύ τους ερωτήσεις. Σ' αυτή την κατηγορία συμπεριλαμβάνονται ερωτήσεις όπως «Πώς η πίστη είναι δυνατή;», «Πώς η ελπίδα είναι δυνατή;» και «Πώς η επιθυμία είναι δυνατή;». Η πραγματική μου πρόθεση, όταν θέτω την «ερώτηση για τη γνώση», είναι να τις ρωτήσω όλες αυτές μαζί. Αν κάποιος θέλει να βάζει ταμπέλες, θα μπορούσε να πει ότι θέτοντας την ερώτηση για τη γνώση, καθώς και τις άλλες, θέτω την ερώτηση της προθετικότητας (intentionality). Όπως θα προσπαθήσω να δείξω με σαφήνεια, όλες αυτές οι καταστάσεις (γνώση, πίστη, ελπίδα και επιθυμία) είναι προθετικές καταστάσεις και η προθετικότητα μιας κατάστασης μαρτυρείται από ορισμένες εξαιρετικά ιδιόμορφες (ιδιαίτερες) αναφορικές ιδιότητες (referential properties). Έτσι μετά τα παραπάνω η ερώτηση που θέτω είναι: «Πώς είναι δυνατόν να υπάρχουν προθετικές καταστάσεις, καταστάσεις της διάνοιας, οι οποίες εμφανίζουν ορισμένες χαρακτηριστικές αναφορικές συνάφειες;»

Οι μοναδικές αναφορικές ποιότητες, οι χαρακτηριστικές των προθεσιακών καταστάσεων ξεκινούν από μία συναφή ποιότητα που θα μπορούσαμε να ονομάσουμε «αναφορότητα». Όπως οι καλύτερες αγγλικές λέσχες, αυτό το είδος της «αναφορότητας» είναι πολύ περιοριστικό (δρα μ' ένα πολύ περιοριστικό ή αποκλειστικό τρόπο). Για να δείξω τι εννοώ, ας κυττάξουμε την περίπτωση του νεαρού Willard. Ο Willard πιστεύει ότι η Αθήνα βρίσκεται στην Ελλάδα, γνωρίζει όμως πολύ λίγα πράγματα για την Αθήνα ή για την Ελλάδα. Συγκεκριμένα έχει ακούσει για την Ακρόπολη, όμως δεν γνωρίζει που αυτή βρίσκεται. Συνεπώς μολονότι ο Willard πιστεύει ότι η Αθήνα είναι μία πόλη στην Ελλάδα, δεν πιστεύει ότι η πόλη στην οποία βρίσκεται η Ακρόπολη είναι η Ελλάδα. Κάποιος μπορεί να πει ότι η πίστη του Willard «η Αθήνα βρίσκεται στην Ελλάδα» είναι μία πίστη σχετικά με την Αθήνα αλλά μ' ένα περιοριστικό τρόπο· η αναφερόμενη πίστη δε συμπεριλαμβάνει το γεγονός ότι στην Αθήνα βρίσκεται η Ακρόπολη. Με το επιθυμητό νόημα του «περί τινος» η πίστη του Willard είναι για την Αθήνα, ενώ δεν είναι σχετική με την πόλη, όπου βρίσκεται η Ακρόπολη.

Όμως, εκτός από τον περιοριστικό χαρακτήρα, υπάρχουν και άλλα προβλήματα που σχετίζονται με την «αναφορότητα». Οι προθετικές καταστάσεις μπορεί να αναφέρονται, και συχνά αναφέρονται, σε πράγματα ανύπαρκτα. Ο φίλος του Willard, που έχει μεγάλη φαντασία και είναι επαρκώς απληροφόρητος, μπορεί να πιστεύει ότι η Ουτοπία του Thomas More ή η Λιλλιπούπολη του Swift βρίσκονται στην Ελλάδα. Μπορεί ακόμα να πιστεύει σε μυθικά τέρατα ή στον Αη-Βασίλη. Μπορούμε να πούμε ότι αυτές οιπίστεις είναι σχετικά με (γύρω από) την Ουτοπία, τη Λιλλιπούπολη, τα μυθολογικά τέρατα και τον Αη-Βασίλη, πράγματα σεβαστά, μολονότι αυτά δε μπορεί να υπάρξουν. Η αποτυχία ύπαρξης επηρεάζει ολοκληρω-

τικά και με αρνητικό τρόπο τις προθεσιακές καταστάσεις· κάποιος εύκολα μπορεί να κατασκευάσει ανάλογα παραδείγματα για επιθυμίες, ελπίδες, φόβους ή σκέψεις. Ο Willard, για παράδειγμα, θέλει ο Αη-Βασίλης να του φέρει κάτι ή φοβάται ότι δε θα του φέρει τίποτα⁴.

Το ερώτημα για τις προθεσιακές καταστάσεις οδηγεί σε ένα άλλο της φιλοσοφικής γεωγραφίας: Πού πρέπει να τοποθετηθούν αυτές μέσα σ' ένα κατανοητό, λογικό ή επεξηγηματικό πλαίσιο, το οποίο πρέπει να συμπεριλαμβάνει καταστάσεις με λογικά διαφορετικές ποιότητες; Από πού πηγάζουν οι προθεσιακές καταστάσεις και πώς σχετίζονται με τα άλλα είδη καταστάσεων; Στις μέρες μας το προτιμητέο κατανοητό σύστημα αναφοράς είναι το φυσικό· φαίνεται να συμφωνούμε με τον Παρμενίδη ότι «όλα είναι ένα», τουλάχιστον εφ' όσον η φράση αναφέρεται στη φυσική. Πιστεύουμε ότι τα πράγματα θα μπορούσαν να είναι ευχάριστα και τακτικά, εάν ο κόσμος ήταν μία συλλογή από φυσικά γεγονότα, εάν όλα τα γεγονότα ήταν φυσικά γεγονότα, και εάν όλες οι εξηγήσεις περιορίζονταν στην αναφορά φυσικών αλληλεπιδράσεων. Όμως, αν η φυσική μάς προσφέρει τη σύνθεση, η προθετικότητα φαίνεται ότι πρέπει να ανθίσταται στην αφομοίωση. Κι αυτό για ένα λόγο: ότι τα φυσικά γεγονότα (οι φυσικές καταστάσεις) στερούνται την αναφορότητα που έχει περιοριστικό χαρακτήρα: αυτά έχουν εντελώς δημοκρατικές ιδιότητες αναφοράς. Όταν, για παράδειγμα, ο πληθυσμός της Αθήνας αυξάνει, τότε αυξάνει αυτόματα και ο πληθυσμός της πόλης στην οποία βρίσκεται η Ακρόπολη, ανεξάρτητα από το αν κάποιος γνωρίζει ή πιστεύει ή ελπίζει ή εύχεται η Ακρόπολη να βρίσκεται στην Αθήνα. Ακόμα, ένα φυσικό αντικείμενο ή συμβάν δεν μπορεί να σχετίζεται με φυσικό τρόπο ανύπαρκτα αντικείμενα ή συμβάντα. Ο πύργος του Άιφελ είναι ανατολικά της Νέας Υόρκης και νότια του Greenwich, όμως δεν υπάρχει κάποιος τρόπος που να τον ορίζει στο χώρο ως προς την Ουτοπία του More. Τελικά οι προθεσιακές καταστάσεις διακρίνονται από τις φυσικές και δεν είναι δυνατόν να αντικατασταθούν από εκείνες. Φαίνεται να μην υπάρχει χώρος για γνήσια προθετικότητα μέσα σ' ένα σύστημα, όπου κυβερνά η φυσική.

Μολονότι η αναγωγή σ' ένα φυσικό σύστημα δεν είναι ο σκοπός μας, οι δυσκολίες παραμένουν. Οι φυσικές καταστάσεις δρουν με τρόπο ώστε να προκαλούν τις προθεσιακές. Εξετάζοντας τις ποικίλες διαφορές μεταξύ τους, είναι δύσκολο μερικές φορές να δει κανείς πώς συμβαίνει αυτό. Η πίστη μου ότι έξω στο δρόμο υπάρχει κίνηση προέρχεται (δημιουργείται) από το θόρυβο που έρχεται απ' έξω και, τελικά, από το πραγματικό πήγαινε-έλα που γίνεται στο δρόμο. Η πίστη μου είναι το τελευταίο στάδιο μίας αιτιώδους αλυσίδας (μίας αλληλουχίας αιτιωδών καταστάσεων), της οποίας η πρώτη αρχή είναι μία φυσική κατάσταση, η κατάσταση που επικρατεί στο δρόμο. Κάθε κρίκος της αλυσίδας φαίνεται μη-προθεσιακός, κάθε κρίκος, εκτός από τον τελευταίο, στερείται το «περί τινος» της προθεσιακής κατάστασης. Όμως η πίστη μου έχει «αναφορότητα». Όμως από πού προέρχεται η αναφορότητα; Πώς μπορεί το τέλος μίας αλυσίδας από διαφορετικές φυσικές αιτίες ξαφνικά να έχει ένα χαρακτηριστικό το οποίο όλοι οι άλλοι κρίκοι, είτε μεμονωμένα είτε εν συνδέσει, στερούνται;

Υπάρχει μία κλασική φιλοσοφική απάντηση σ' αυτή την ερώτηση καθώς και στις άλλες που σχετίζονται με την προθετικότητα. Είναι απλό το να αρνηθεί κανείς

ότι ο κόσμος είναι μόνο φυσικός και ότι κάθετί επιδέχεται κοινές και φυσικές ερμηνείες. Μάλλον (συνεχίζει η απάντηση) ο κόσμος αποτελείται από δύο βασιλεια, το φυσικό και το πνευματικό· οι ποιότητες των δύο βασιλείων και οι νόμοι που τα κυβερνούν είναι πολύ διαφορετικοί και δεν επιδέχονται καμμία αφομοίωση. Ένας υποστηρικτής αυτής της θέσης θα μπορούσε να ισχυρισθεί ότι στην περιγραφή των αιτιωδών σταδίων της πίστης μου παρέλειψα κάτι στο τέλος της αλυσίδας. Δεν έδωσα την πρέπουσα σημασία στο γεγονός ότι στο τέλος αυτής της αιτιώδους αλυσίδας βρίσκεται ένα ανθρώπινο ον, συγκεκριμένα δε ένα ον με νου. Κατά συνέπεια, αφού υπάρχει ένας νους (συνεχίζει η παραδοσιακή απάντηση), η τελική κατάληξη της αιτιώδους αλυσίδας είναι διαφορετική από αυτή που εγώ έχω περιγράψει. Οι νόες, τελικά, είναι ασυνήθιστες οντότητες: έχουν εκτός από τις φυσικές ιδιότητες και μη-φυσικές. Σπουδαιότερες ανάμεσά τους είναι η προθετικότητα και η «αναφορότητα». Η πηγή της προθετικότητας, η πηγή της σκέψης μου, είναι (ή έτσι η παραδοσιακή άποψη θα μπορούσε να το παρουσιάσει) η φυσικά ανεξιχνίαστη φύση του ίδιου του νου.

Αυτή η απάντηση στο πρόβλημα της προθετικότητας είναι μία έκφραση του φιλοσοφικού δυϊσμού: ο νους είναι νους και η ύλη είναι ύλη και ποτέ αυτά τα δύο δε θα συναντηθούν. Όμως αυτή είναι μία απάντηση που αρνούμαι να δεχθώ, τουλάχιστον όχι χωρίς κάποιο αγώνα. Απέχω από το να μη δεχθώ ότι ο νους είναι ένα ιδιόρρυθμο είδος πράγματος· αντίθετα είναι εφοδιασμένος με όλα τα είδη των μηχανισμών που ακόμα πρέπει να τα καταλάβουμε ή ίσως να τα συλλάβουμε ορθά. Όμως ευχαρίστως θα δεχόμουν ότι μία λύση στο πρόβλημα της προθετικότητας μπορεί να προκύψει από την κατανόηση και τη διαλεύκανση του γεγονότος ότι με κάποιο τρόπο έκανα λάθος στην περιγραφή της αιτιώδους σειράς που οδηγεί στην πίστη μου σχετικά με το θόρυβο στο δρόμο. Παραδέχομαι ότι πιθανόν να αστόχησα στην περιγραφή της πραγματικής κατάστασης με την παράλειψη σχετικών λεπτομερειών. Μπορεί ίσως παρά να έκανα λάθος από την παράλειψη γεγονότων σχετικών με το ανθρώπινο μέρος της αλυσίδας, μάλλον να έχω παραβλέψει και ορισμένους εν δράσει παράγοντες που σχετίζονται με τους κρίκους της αλυσίδας που δεν έχουν να κάνουν με τον άνθρωπο. Ειδικότερα, μπορεί να υπάρχουν προθεσιακές όψεις σ' ένα καθαρά φυσικό σύμπαν, που όμως να έχουν παραμεληθεί. Λαμβάνοντας υπ' όψιν όλα τα παραπάνω, μπορεί να μη λύνονται όλα τα ερωτήματα που απαρτίζουν αυτό που έχω ονομάσει «γεωγραφικό» πρόβλημα της προθετικότητας, όμως τουλάχιστον ξεκαθαρίζεται η έκταση στην οποία αυτά εκτείνονται.

Δεν είναι τόσο δύσκολο να σκεφτούμε ότι μερικές μη-πνευματικές πραγματικότητες μετέχουν σ' ένα τύπο προθετικότητας· στην πραγματικότητα αυτό το κάνουμε αρκετά συχνά. Στην επιστήμη των υπολογιστών ερμηνεύουμε —κι αυτό θεωρείται ολόσωστο— τη λειτουργία των προγραμμάτων, της μνήμης και της κεντρικής μονάδας επεξεργασίας (CPU) με αναφορά στις διαδικασίες που συμβαίνουν ανάμεσα στα μικρότερα στοιχεία πληροφοριών (bits). Αναφερόμαστε στα μικρότερα τμήματα πληροφοριών και λέμε ότι αυτά αποθηκεύονται, ξαναδοκιμάζονται, μεταφέρονται, περνούν, κωδικοποιούνται ή τυπώνονται και, κατά περίπτωση, ότι δημιουργούνται ή καταστρέφονται.

Εν ολίγοις, αντιμετωπίζουμε τα στάδια λειτουργίας ενός υπολογιστή ως

πραγματικούς μεταφορείς πληροφοριών. Και η ομιλία μας δείχνει ότι αποδίδεται στα μικρότερα τμήματα πληροφοριών η περιοριστική αναφορότητα των προθεσιακών καταστάσεων. Ένα κελλί της μνήμης του υπολογιστή μπορεί να διατηρήσει την πληροφορία ότι η Αθήνα βρίσκεται στην Ελλάδα, χωρίς να κρατήσει και την πληροφορία ότι η πόλη στην οποία βρίσκεται η Ακρόπολη είναι στην Ελλάδα. Επίσης ένα κελλί μνήμης μπορεί να κωδικοποιήσει την πληροφορία ότι ο Αη-Βασίλης είναι ένας χαρούμενος φανταστικός γέροντας. Ο υπολογιστής δεν είναι ο μοναδικός μη πνευματικός κομιστής πληροφοριών. Σχεδόν κάθε φυσική πραγματικότητα μπορεί να κάνει κάτι τέτοιο. Οι φυσαλίδες μέσα σ' ένα θάλαμο μάς δίνουν πληροφορίες για τα υποατομικά σωματίδια. Τα σημερινά σύννεφα με πληροφορούν για τον αυριανό καιρό. Ο θόρυβος της κάνουλας από το μπάνιο με πληροφορεί ότι το νερό τρέχει.

Οι μετεωρολόγοι θα πρέπει να είναι πολύ ικανοποιημένοι με την περιγραφή των νεφών ως αιωρούμενων σταγονιδίων νερού. Όσον αφορά την ακουστική, ο θόρυβος του νερού που τρέχει από την κάνουλα είναι μία τυπική περίπτωση κυμάτων στον αέρα. Εντούτοις, από μία ορθή επιστημολογική σκοπιά ίσως πρέπει να αντιμετωπίσουμε όμοια τους υπολογιστές, τα σύννεφα, τις φυσαλίδες και το θόρυβο των σταγόνων που τρέχουν από την κάνουλα, ως κομιστές μίας πληροφορίας και ως καταστάσεις που χαρακτηρίζονται από προθετικότητα. Τουλάχιστον, έτσι νομίζω.

Ακριβέστερα, προτείνω να αντιμετωπίσουμε την όχι τόσο τυπική «πληροφορική συζήτηση» των επιστημών των υπολογιστών με έναν όχι τόσο τυπικό τρόπο και να νοήσουμε την ποσότητα πληροφοριών ως αφηρημένη μη στατιστική ποσότητα, που σχετίζεται φυσικά με καταστάσεις των φυσικών αντικειμένων και διαδικασιών. Οι πιο κοινές αφηρημένες ποσότητες με τις οποίες μπορεί να συγκριθεί είναι η προσανατολισιμότητα στη γεωμετρία του φυσικού κόσμου ή η κατεύθυνση της κίνησης στην κλασική φυσική. Η κατεύθυνση της κίνησης μου παρέχει όντως ένα πετυχημένο παράδειγμα, για να περιγράψω αυτό που έχω στο μυαλό μου, καθώς αυτή καθορίζεται από τη σύγκριση με καθορισμένα κατευθυντικά πρότυπα (σταθερά). Για να δώσουμε κάποιες κατευθύνσεις, ορίζουμε ένα πλαίσιο αναφοράς, ένα σύστημα από κατευθυντικές συντεταγμένες. Οι επί μέρους κατευθύνσεις διακρίνονται σε αναφορά προς το γενικό πλαίσιο, που το ίδιο είναι πρωταρχικό. Δεν γίνεται καμμία προσπάθεια να αναχθεί το ίδιο το πλαίσιο σε άλλες μη-κατευθυντικές ποιότητες. Η ποσότητα των πληροφοριών προσδιορίζεται εν πολλοίς από τον τρόπο που χρησιμοποιούμε ένα σύστημα συντεταγμένων. Το πλαίσιο αναφοράς εδώ καθορίζεται από το λογικό χώρο των προτάσεων. Μέσα στο σύστημα η διάκριση του πληροφοριακού περιεχομένου ενός γεγονότος γίνεται με μία λογική πρόταση, το πληροφοριακό περιεχόμενο της οποίας είναι όμοιο, ανάλογα με τις περιστάσεις, με τη δημιουργημένη από το γεγονός πληροφορία. (Επί τη ευκαιρία, επιλέγουμε προτάσεις και για να προσδιορίσουμε πίστεις — λέμε ότι αυτή πιστεύει ότι το σπίτι της βρίσκεται εκεί που βρίσκεται η καρδιά της ή ότι αυτή είναι η καλύτερη ώρα στην Αγγλία· και στις δύο περιπτώσεις το «ότι» και όσα το ακολουθούν δηλώνουν μία κρίση. Δεν είναι λοιπόν άξιο απορίας ότι η πληροφορία μπορεί να έχει και αυτή τον περιοριστικό χαρακτήρα της πίστης).

Δεν είμαι ο πρώτος, ούτε καν ανάμεσα στους πρώτους, που συνιστά την

αποδοχή μίας σχετικά ανεξάρτητης από την έννοια του νου άποψης για το περιεχόμενο μίας πληροφορίας. Η ιδέα ήδη έχει κερδίσει αξιοσημείωτο χώρο στην αναλυτική φιλοσοφία του νου με τις προσπάθειες ορισμένων φιλοσόφων, κατά κύριο λόγο του Fred Dretske⁵. Η θέση του Dretske, που θεμελιώνεται στη στατιστική θεωρία της πληροφορικής των Shannon και Weaver, βρίσκεται αντιμέτωπη με πολλές δυσκολίες, που πιστεύω ότι μπορεί να ξεπεραστούν με την αντικατάσταση της στατιστικής βάσης με μία αλγεβρική βάση.

Η εναλλακτική βάση που έχω στο μυαλό μου, χρησιμοποιεί ιδέες αρκετά εύχρηστες στον τομέα της μερικώς διατεταγμένης δηλωτικής σημασιολογίας (denotational semantics). Στην προκειμένη περίπτωση αρκεί να γνωρίζουμε μόνο ότι η δηλωτική σημασιολογία ισχυρίζεται ότι λειτουργεί ως θεωρία προτύπων για τις γλώσσες προγραμματισμού. Τα μερικώς διατεταγμένα σύνολα που ονομάζονται «πεδία» (domains) αποτελούν τα πρότυπα. Τα προγράμματα ερμηνεύονται με την απόδοση σ' αυτά «σημασιών» οι οποίες είναι κατάλληλες απεικονίσεις μεταξύ των πεδίων. Όπως στην παραδοσιακή θεωρία προτύπων, σκοπός αυτής της σημασιοδοτικής προσπάθειας είναι να δει τις ιδιότητες των αντικειμένων της τεχνητής γλώσσας (εδώ των προγραμμάτων) όπως αυτές εικονίζονται στις αλγεβρικές σχέσεις μεταξύ των σημασιών των αντικειμένων αυτών⁶.

Οι έννοιες του πεδίου και της απεικόνισης καθώς και η αλγεβρική θεωρία που στηρίζεται επάνω τους, είναι εφαρμόσιμες στην πληροφορική και γενικά στη μεταφορά πληροφοριών. Κατ' αρχάς τα «πεδία» μπορεί να θεωρηθούν σύνολα από τα μικρότερα τμήματα πληροφοριών τα οποία, όπως και οι προτάσεις, είναι στενά συνδεδεμένα δια υπαλληλίας. Δεύτερον, οι απεικονίσεις ανάμεσά τους, όπως μπορεί να καταδειχθεί, αντιστοιχούν σε πετυχημένες υπολογίσιμες μεταφορές τμημάτων πληροφοριών ανάμεσα σε ένα πεδίο και σε ένα άλλο⁷. Λαμβάνοντας υπόψη αυτές τις δύο ιδέες, μπορεί κάποιος να θέσει με αποκλειστικά δηλωτικό (denotational) τρόπο αξιώματα που καθορίζουν τι πρέπει να θεωρηθεί αποδεκτή ανταλλαγή πληροφοριών. Δυστυχώς τα όρια της παρούσης εργασίας αποκλείουν την παρουσίαση αυτών των αξιωμάτων, καθώς και μία συζήτηση για τις συνέπειές τους. Μπορώ μόνο να ισχυρισθώ ότι αυτά προσφέρουν μία γοητευτική εναλλακτική πρόταση στη θεωρία του Dretske· από αυτά τα αξιώματα μπορεί κανείς να αντλήσει τις βασικές αρχές της θεωρίας του Dretske, ενώ ταυτόχρονα αποφεύγει τις προαναφερθείσες δυσκολίες.

Έχοντας ως υπόβαθρο μία αφηρημένη θεωρία της πληροφορικής έρχομαι να προτείνω κάτι πολύ απλό: ο ανθρώπινος νους δεν είναι τόσο παράξενος. Με πολλούς τρόπους συγγενεύει με τους υπολογιστές, τα σύννεφα και τα τηλέφωνα ως προς το ότι όλα είναι φορείς πληροφοριών κι έχουν το χαρακτηριστικό της προθετικότητας. Ο νους διαφέρει μόνο ως προς το βαθμό κι όχι ως προς το είδος από τους πιο «φυσικούς» μεταδότες πληροφοριών. Το να έχει κάποιος πίστεις δεν είναι κάτι εντελώς διαφορετικό στο είδος από τις διάφορες καταστάσεις που σ' αυτά τα δημιουργήματα ισχύουν μ' έναν απόλυτα φυσικό τρόπο⁹. Εν ολίγοις η προθετικότητα έχει ως χαρακτηριστικό της την περιοριστικότητα — όμως δεν είναι μία ιδιότητα που χαρακτηρίζει αποκλειστικά τα ανθρώπινα όντα.

Η δεύτερη ερώτησή μου, «Πώς το λάθος είναι δυνατόν;», είναι επίσης σχετική με τα θέματα των μαθηματικών της πληροφορικής. Εντούτοις μία προσέγγιση

αυτού του τομέα δε θα γίνει από την πλευρά των αλγεβρικών ή σημασιολογικών ιδιοτήτων των πληροφοριών, αλλά από αυτό που μπορούμε να ονομάσουμε «αριθμητική» τους. Θα υποστηρίξω την ιδέα ότι τα θεωρητικά όρια που συνάπτονται με υπολογισμούς και συλλογισμούς που γίνονται με μηχανές, ισχύουν και για τους υπολογισμούς και συλλογισμούς των ανθρώπων. Η θεωρία του λάθους απλά προσφέρει μία ευκαιρία να φανεί καθαρά η φιλοσοφική σπουδαιότητα αυτής της ιδέας.

Ένας φιλόσοφος που λέει ότι προσφέρει μία θεωρία για την ωφέλεια της γνώσης έχει την υποχρέωση να αναπτύξει και μία σχετική θεωρία για τη δυνατότητα να κάνουμε λάθη· κάποιος που θέλει να εξηγήσει πώς συμβαίνει να βλέπουμε ορθά ένα θέμα έχει υποχρέωση να μας πει και πώς καταφέρνουμε να αντιμετωπίζουμε τόσα θέματα λανθασμένα. Μερικά χωρία από τον Descartes μάς προσφέρουν ένα λαμπρό παράδειγμα των δυσκολιών που αυτό το καθήκον επιβάλλει. Στον τέταρτο «Στοχασμό» του ο Descartes θεωρεί ότι έχει αποδείξει πως η ορθότητα της γνωστικής διαδικασίας εξασφαλίζεται από την καλοσύνη του Δημιουργού. Το πρόβλημα του λάθους εμφανίζεται αμέσως: οι ατέλειες στη γνωστική διαδικασία ρίχνουν μία σκιά στην καλοσύνη του Δημιουργού και δημιουργούν αμφιβολίες. Ο Descartes αναγκάζεται να ρωτήσει: «Αν ο Θεός είναι τόσο καλός, ώστε να εγγυάται την ορθότητα της γνώσης μας, γιατί δεν είναι επαρκώς καλός, ώστε να εξασφαλίζει ότι ποτέ δε κάνουμε λάθη;»

Η απάντηση του Descartes είναι ότι ο Θεός δεν έχει σχεδιάσει το λάθος μέσα στο μηχανισμό του νου μας, εμείς κάνουμε λάθη χαλώντας αυτό το μηχανισμό. Ο Descartes προχωρεί στην ερμηνεία των δυσλειτουργιών αυτών συνδυάζοντας μία θεώρηση του νου ως συνόλου ικανοτήτων, με κάτι που θα μπορούσαμε να ονομάσουμε «στερεοφωνική» θεωρία της κρίσης. Το στερεοφωνικό γνωστικό σύστημα συντίθεται από δύο ικανότητες· υπάρχει η ικανότητα της διάνοιας, το λογικό τμήμα του νου, και η ικανότητα της βούλησης, το τμήμα εκείνο που επιλέγει ή αποφασίζει. Η κρίση, κατά τον Descartes, προκύπτει, όταν τα δύο τμήματα εργάζονται μαζί: η διάνοια σκέπτεται τι να ισχυρισθεί και η θέληση ισχυρίζεται. Η κρίση είναι ορθή, όταν τα τμήματα του νου είναι απόλυτα ισορροπημένα, όταν η διάνοια έχει ολοκληρώσει τη θεώρηση των πραγμάτων και συμφωνεί πλήρως με τις αποφάσεις της βούλησης. Κάθε τμήμα, λέει ο Descartes, είναι «τέλειο και ολοκληρωμένο»¹⁰. Το λάθος άρα δεν μπορεί να προέλθει μόνο από τη μία από τις δύο ικανότητες και, όπως με την κρίση, προέρχεται από τη «στερεοφωνική ρήξη». Η ρίζα του λάθους, πιστεύει ο Descartes, βρίσκεται στο ότι η σκέψη τείνει να μείνει πίσω από την απόφαση· έχουμε μία τάση να ισχυριζόμαστε πράγματα στα οποία η διάνοια δεν έχει βάλει τη σφραγίδα της αποδοχής.

Οι σκέψεις του Descartes στο θέμα του λάθους παρέχουν ένα άριστο και ολοκληρωμένο δείγμα των θεωριών που χαρακτηρίζουν το λάθος «κακή λειτουργία». Αυτές οι θεωρίες πιστεύουν ότι η συσκευή της γνώσεως είναι τέλεια, ως έχει. Το λάθος δεν είναι γνώρισμα της σχεδιάσεως του μηχανισμού ούτε της καλής λειτουργίας του. Συμβαίνει, μόνο όταν η συσκευή παθαίνει μία διακοπή, όταν αποτυγχάνει να δουλέψει έτσι όπως έχει σχεδιασθεί να δουλεύει. Κάνουμε λάθος,

όταν κάποια από τις γνωστικές διόδους είναι χαλαρή ή ένα σύρμα της σκέψης σπάει ή (κατά τον Descartes) όταν το κάνουμε να σπάσει.

Τέτοιου τύπου θεωρίες είναι ανίσχυρες, γιατί δεν δίνουν λογικές ερμηνείες για τα διάφορα είδη των κοινών λαθών¹¹. Παραδείγματα ενός από αυτά τα είδη αποτελούν τα λάθη που συσχετίζονται με τα μαθηματικά ή λογικά αινίγματα, αυτά που εμφανίζονται στις εφημερίδες ή στα βιβλία με αινίγματα. Οι μέσοι άνθρωποι, όταν αντιμετωπίζουν πετυχημένα αινίγματα, είναι πιο πιθανό να απαντήσουν λανθασμένα ακόμα και μετά από αρκετή σκέψη. Όσο καλύτερο είναι το αίνιγμα, τόσο πιθανότερο είναι ότι ο μέσος άνθρωπος θα το λύσει λανθασμένα. Σ' ένα πραγματικά έξυπνο αίνιγμα θα ήταν κάτι αξιοσημείωτο, αν κάποιος δεν έκανε σφάλμα. Εντούτοις τα λάθη που γίνονται κατά τη λύση αινιγμάτων δεν είναι συμπτώματα κακής γνωστικής λειτουργίας — το αντίθετο μάλιστα. Μία συνηθισμένη λανθασμένη απάντηση σ' ένα απατηλό ή περίπλοκο πρόβλημα είναι ένα είδος επιβεβαίωσης ότι καθετί στη νοητική μηχανή λειτουργεί. Η διαρκής παρουσία τέτοιων σφαλμάτων αποτελεί σπουδαίο αποδεικτικό γεγονός ενάντια στις θεωρίες περί κακής λειτουργίας¹².

Το ριζικό λάθος των θεωριών που εκλαμβάνουν το σφάλμα σαν κακή λειτουργία μπορεί κανείς να το δει, αν εξετάσει αυτό το θέμα από τη θέση υπεροχής της θεωρίας του υπολογισμού. Εξετάζοντας έτσι το θέμα, μπορούμε να δούμε ότι ο Descartes έκανε λάθος να νομίζει ότι δεν είναι δυνατόν να πέσουμε σε λάθος, όταν η νοητική ικανότητα λειτουργεί όπως πρέπει. Στην πραγματικότητα μπορούμε να διαπιστώσουμε ότι ίσως έσφαλε για μαθηματικούς κυρίως λόγους. Τώρα είμαστε σίγουροι ότι κανένα σκεπτόμενο ον δεν μπορεί να είναι «τέλειο και ολοκληρωμένο»: είναι αναμενόμενο κάθε νοητική ικανότητα, κάθε πρακτική και γενικά εφαρμόσιμη μέθοδος συλλογισμού να διαπράττει «ενδιάθετα» λάθη ως ένα απαραίτητο παρεπόμενο της λειτουργίας της. Για να βγάλουμε ορθά συμπεράσματα, πρέπει να κάνω μία γρήγορη ανασκόπηση των γνωστών αποτελεσμάτων που μας έχει δώσει η θεωρία του υπολογισμού.

Ένας τρόπος να περιγράψουμε τις σημερινές προσπάθειές μας στη θεωρία του υπολογισμού είναι να αντιπαραθέσουμε τις δύο ποικιλίες του από υπολογιστική άποψη δύσκολου προβλήματος, το άλυτο και το δυσεπίλυτο. Τα άλυτα προβλήματα είναι «πανδύσκολα»: αυτά είναι προβλήματα για τα οποία δεν μπορεί να υπάρξει μία γενική αλγοριθμική λύση. Στην περίπτωση προβλημάτων $P(n)$, όπου n είναι ένας ακέραιος αριθμός, το $P(n)$ είναι άλυτο, εάν δεν μπορεί να υπάρξει ένας πεπερασμένος, υπακούων σε κανόνες μηχανισμός που να δέχεται οποιοδήποτε n ως δεδομένο και μετά από πεπερασμένο χρόνο υπολογισμού να δίνει ένα ορθό αποτέλεσμα για το αν η $P(n)$ ισχύει ή όχι. Με άλλα λόγια, τα άλυτα προβλήματα είναι εκείνα για τα οποία δεν μπορεί να υπάρξει απολύτως κανένα πρόγραμμα που να μας δίνει τη λύση τους: οι λύσεις τους δεν μπορούν καν να εκφραστούν σε καμία γλώσσα προγραμματισμού. Υπάρχουν πολλά άλυτα προβλήματα (στην πραγματικότητα γνωρίζουμε ότι η συντριπτική πλειονότητα των δυνατών προβλημάτων είναι άλυτα): το πιο γνωστό από αυτά είναι ίσως το “halting problem”. Η αδυναμία λύσεως στο πρόβλημα αυτό σημαίνει ότι δεν υπάρχει κάποια γενική διαδικασία που να ανακαλύπτει τα λάθη σ' ένα πρόγραμμα κατά το χρόνο ροής του: αυτή είναι μία εργασία που ο υπολογιστής ποτέ δε θα μπορέσει να φέρει σε

πέρας. Το σπουδαίο είναι ότι η αδυναμία αυτή είναι θέμα απλώς μαθηματικών. Δηλαδή, το να βρεθεί μία λύση δεν είναι δύσκολο ή βαρετό ή απαγορευτικά δαπανηρό. Είναι μαθηματικώς αδύνατο.

Σε σύγκριση με την «πανδυσκολία» των άλυτων προβλημάτων, τα δυσεπίλυτα προβλήματα είναι εκείνα που «πραγματικά πολύ δύσκολα» δέχονται μία πρακτική λύση. Σε γενικές γραμμές ένα πρόβλημα είναι δυσεπίλυτο, όταν ο χρόνος ή ο χώρος που απαιτεί για τον υπολογισμό λύσεων, ακόμα και σε περιπτώσεις που θα ήταν δυνατές κάποιες λύσεις, θα ξεπερνούσε τη διάρκεια ζωής ή τη μνήμη κάθε πραγματικού μηχανήματος¹⁴. Τώρα είμαστε λογικά βέβαιοι ότι υπάρχουν στοιχειώδη προβλήματα που είναι δυσεπίλυτα. Ένα από αυτά είναι το test για τους πρώτους αριθμούς· έχουμε κάθε λόγο να πιστεύουμε ότι δεν υπάρχει μία εφαρμόσιμη πορεία μέσω της οποίας να καθορίζουμε αν ένας ακέραιος αριθμός έχει ή όχι πρώτους παράγοντες¹⁵. Μολονότι η δυσκολία επιλύσεως εκτιμάται βάσει παραγόντων που δεν είναι λογικά αναγκαίοι (π.χ. δυνατός χρόνος υπολογισμού), αυτή, όπως και η αδυναμία λύσης, είναι ένα καθαρά μαθηματικό γεγονός. Διαπιστώνουμε ότι ένα πρόβλημα είναι δύσκολο να επιλυθεί όχι επί τη βάσει της εμπειρίας, αλλά επί τη βάσει της μαθηματικής απόδειξης.

Μία σημαντική διαφορά ανάμεσα στα άλυτα και στα δυσεπίλυτα προβλήματα είναι αυτή της σπουδαιότητας. Τελικά τα περισσότερα άλυτα προβλήματα αφορούν μάλλον λίγους μνημένους και δεν προκαλούν αϋπνία στους επιστήμονες των υπολογιστών. Τα δυσεπίλυτα προβλήματα όμως δημιουργούν σίγουρα δυσκολίες και καθυστερήσεις. Εκείνα συχνά είναι τόσο στοιχειώδη και πρακτικά, ώστε η ανάγκη να βρεθεί ικανοποιητική λύση δεν μπορεί λογικά να αγνοηθεί.

Όπως πάντα, «η μαθηματική πενία αλγοριθμικής τέχνας κατεργάζεται», οι επιστήμονες των υπολογιστών έχουν συγκεντρώσει ως εργαλείο τους έναν εξοπλισμό από στρατηγικές για την αντιμετώπιση των δυσεπίλυτων προβλημάτων. Κοινό χαρακτηριστικό σε αρκετές στρατηγικές είναι η «εξαπάτηση με μέτρο». Ο καλύτερος τρόπος για να εξηγήσω αυτή την έννοια είναι με ένα παράδειγμα. Δύο από τις πιο συνηθισμένες «απάτες» τις συναντάμε σε δύο διαφορετικά είδη αλγορίθμων, τον «προσεγγιστικό» και τον «πιθανολογικό». Και στις δύο περιπτώσεις η «απάτη» έγκειται στο ότι ακολουθούμε μία πορεία, όταν έχουμε να λύσουμε ένα πρόβλημα, η οποία αυστηρά μιλώντας δεν μας δίνει μαθηματικώς ορθά μέσα για την εξεύρεση λύσεων σ' αυτό το πρόβλημα. Όμως ακόμα κι έτσι η πορεία αυτή είναι κατάλληλα σχεδιασμένη, ώστε να δίνει ένα αποτέλεσμα «μετρίως» λανθασμένο. Ο προσεγγιστικός αλγόριθμος εγγυάται ότι προσφέρει «λύσεις» που είναι προσεγγίσεις και βρίσκονται μέσα σ' ένα περιθώριο ανέκτου σφάλματος¹⁶. Όπως και το ίδιο το όνομα μάς λέει, ο πιθανολογικός αλγόριθμος χρησιμοποιεί τυχαιοποιητικές (randomizing) μεθόδους, για να ξεπεράσει τις δυσκολίες και να καταλήξει σ' ένα αποτέλεσμα που έχει πολλές πιθανότητες να είναι κοντά στο **ορθό**¹⁷. Φυσικά δεν υπάρχει περίπτωση να ξεπεράσουμε ένα μαθηματικό γεγονός· εάν ένα πρόβλημα είναι πραγματικά δυσεπίλυτο, οποιοσδήποτε ικανοποιητικός αλγόριθμος μπορεί μόνο να φιλοδοξήσει να είναι η γενική λύση του προβλήματος. **Ακόμα** κι αν ένας αλγόριθμος δώσει αποτελέσματα σε «κοινές» ή «τυπικές» περιπτώσεις του προβλήματος, υπάρχουν οπωσδήποτε και άπειρες περιπτώσεις

στις οποίες θα δώσει λανθασμένες απαντήσεις. Εν ολίγοις το τμήμα του πρακτικού υπολογισμού είναι ότι εγκλείει την πιθανότητα δημιουργίας ενός όχι αμελητέου αριθμού λαθών. Η θεωρία του υπολογισμού δείχνει ότι η πρακτικότητα των αλγορίθμων έχει ένα επιστημολογικό τμήμα που μπορούμε να αποφύγουμε να πληρώσουμε, μόνο αν διαπράξουμε μία μικρή «απάτη».

Είναι λογικό να σκεφθούμε ότι τα όρια που ισχύουν για τον πρακτικό υπολογισμό ισχύουν και για την πρακτική λογική. Ο νους, όπως και οτιδήποτε άλλο, κατ' ανάγκη δεσμεύεται από τα μαθηματικά και υπόκειται στους περιορισμούς που αυτά θέτουν. Η ανθρώπινη λογική, εφόσον είναι ευρέως εφαρμοσμένη, πρέπει να διαπράττει «ενδιάθετα» λάθη. Ευτυχώς αρκετά στοιχεία υπέρ αυτής της απόψεως έχουν συγκεντρωθεί από τους πειραματικούς ψυχολόγους. Για παράδειγμα ο Johnson-Laird παραθέτει πειράματα που αποκαλύπτουν ότι η συνάφεια (context) περιορίζει σοβαρά την ικανότητα των ανθρώπων να εκτελέσουν έστω και απλούς συλλογισμούς. Ο Tversky και ο Kahneman έχουν συντάξει ένα σύντομο κατάλογο των φυσικών μας τάσεων να χρησιμοποιούμε κατά την επίλυση προβλημάτων ενστικτώδεις τρόπους που συχνά οδηγούν σε λάθος¹⁸. Μετά από μία εισαγωγή στα δυσεπίλυτα προβλήματα δεν πρέπει να εκπλαγούμε που οι άνθρωποι από τη φύση τους σκέπτονται μάλλον ενστικτωδώς ή πιθανολογικά παρά καθαρώς παραγωγικά. Τελικά οι συγκεκριμένοι «χώροι προβλημάτων» στους οποίους οι άνθρωποι διαρκώς εμπλέκονται είναι τουλάχιστον τόσο περίπλοκοι, όσο κι εκείνοι των αφηρημένων προβλημάτων που τα θεωρούμε δυσεπίλυτα. Αναγνωρίζω ότι η έρευνα σ' αυτές τις σκοτεινές περιοχές συνεχίζεται ακόμα και είναι πολύ νωρίς να προλέγουμε το τελικό αποτέλεσμα με απόλυτη βεβαιότητα. Εντούτοις τα αποτελέσματα από ποικίλους χώρους υπάρχουν και, αν θα έπρεπε να εκφέρω γνώμη, θα έλεγα βέβαια ότι ο Descartes σίγουρα έσφαλε: το λάθος είναι μαθηματικώς αναγκαίο επακόλουθο της σκέψης. Αν κάποιος σχεδίασε τη νοητική μας ικανότητα, το λάθος περιλαμβανόταν στο σχέδιο. Το λάθος δεν είναι κατ' ανάγκη σύμπτωμα γνωστικής δυσλειτουργίας, αλλά μπορεί να είναι ένας μαθηματικά αναπόφευκτος όρος της γνώσης εν γένει.

Θα κλείσω επαναλαμβάνοντας τις ερωτήσεις μου και τις προτεινόμενες απαντήσεις. Πρώτα ρωτήσαμε: «Πώς η γνώση —και κάθε προθετική κατάσταση— είναι δυνατή;» Η απάντησή μου ήταν: «Αυτή είναι μία πολύ παράξενη ερώτηση. Αντιμετωπίζοντας ευρύτερα την προθετικότητα βλέπουμε ότι οι προθεσιακές καταστάσεις δεν είναι απίθανες, αλλά διαρκώς παρούσες. Ο κόσμος από μόνος του είναι ένας χώρος προθέσεων. Ανησυχούμε για τέτοιες ερωτήσεις, μόνο αν αγνοούμε αυτό το γεγονός». Η δεύτερη ερώτηση ήταν: «Πώς το λάθος είναι δυνατόν;» Η προτεινόμενη απάντηση είναι: «Και αυτή η ερώτηση είναι πολύ περίεργη, στηρίζεται σε εντελώς λάθος βάση. Αντιθέτως είναι πιθανό να ανακαλύψουμε ότι είναι μαθηματικώς αδύνατο να είναι το λάθος αδύνατο».

Τέλος πρέπει να ξεκαθαρίσω στον αναγνώστη —αν δεν το έχω πετύχει ως τώρα— ότι οι προτάσεις μου έχουν ανάγκη κάποιων βασικών προϋποθέσεων. Σημαντικότερη είναι η παραδοχή ότι η θεωρία του υπολογισμού θα μας οδηγήσει στην κατασκευή ακριβών μαθηματικών μοντέλων της γνωστικής διαδικασίας. Αυτό με τη σειρά του στηρίζεται σε μία άλλη παραδοχή: ότι, όπως ο χώρος και ο χρόνος

ή όπως οι αλληλεπιδράσεις των στοιχειωδών σωματιδίων, η νόηση έχει χαρακτηριστικές τυπικές ιδιότητες και αυτές είναι οι ιδιότητες με τις οποίες θα εκτιμήσουμε την ορθότητα των υπολογιστικών μοντέλων. Κατά συνέπεια η ψυχολογία και η επιστημολογία δεν μπορούν μόνο να γίνουν φυσικές επιστήμες, μπορούν σε κάποιο βαθμό να γίνουν αφηρημένες. Μαζί αυτές οι προϋποθέσεις αντανakλούν ή και εμπεριέχουν μία τάση στη γνωστική επιστήμη, μία θεωρητική τάση που θα ήθελα να υιοθετήσω. Και κατά σύμπτωση αυτή δεν είναι μία τάση που ο Descartes θα έπρεπε αναγκαστικά να βρει προβληματική — τουλάχιστον από θεολογική πλευρά. Αν όλοι οι νόες υπακούουν ακριβώς στους μαθηματικούς νόμους, τότε ο νους του Θεού θα δεσμεύεται τόσο πολύ από αυτούς, όσο και κάθε άλλος.

ΣΗΜΕΙΩΣΕΙΣ

1. Τουλάχιστον στην πρώτη *Κριτική*. Ο Καντ εισηγείται μία λεπτολογημένη θεωρία για το λάθος στην *Κριτική της δυνάμεως της κρίσεως*.
2. Υπάρχουν φιλόσοφοι που θεωρούν την έννοια του “a priori” ατελή, όμως δεν είμαι ένας από αυτούς. Τα φαινόμενα βεβαιώνουν ότι η χρήση της έκφρασης “a priori” σημαίνει μία γνωσιοθεωρητικά σημαντική διάκριση και πιστεύω ότι τουλάχιστον αυτή την εποχή τα φαινόμενα δεν απατούν. Αν υπάρχει κάποια ατέλεια σχετικά με την απριορικότητα, είναι η επικρατούσα φιλοσοφική ερμηνεία της.
3. Μία πιθανή εξαίρεση στην τάση που εξέθεσα είναι η σχετικά δημοφιλής και υπό όρους επιτυχής θεωρία του λειτουργισμού στη φιλοσοφία του νου. Πρέπει εντούτοις να σημειωθεί ότι η αρχική ιδέα του λειτουργισμού, δηλαδή η αναλογία ανάμεσα στις νοητικές και στις λογικές καταστάσεις, είναι δημιούργημα των αρχών της δεκαετίας του 1960, μίας εποχής κατά την οποία οι θεωρητικές έννοιες πάνω στην οποία αυτή στηρίζεται (η μηχανή του Turing και η κατάσταση υπολογισμού) ήταν ιδέες συναφείς όχι τόσο με την επιστήμη των υπολογιστών, αλλά με ένα κλάδο των μαθηματικών που ασχολείται με το ιδιαίτερο εκείνο είδος συναρτήσεων που είναι γνωστές ως recursive functions.
4. Αυτό απλά αποτελεί το αντίστοιχο των ανησυχιών του Brentano για τη λεγόμενη «προθεσιακή ανυπαρξία».
5. Πβ. F. Dretske, *Knowledge and the Flow of Information*, MIT Press, 1981.
6. Η πιο προσιτή εισαγωγή μάς δίνεται στο έργο του Joseph Stoy *Denotational Semantics: The Scott-Strachey Approach to Programming Language Theory*, MIT Press, 1977.
7. Αυτές οι ιδέες για τα πεδία και την απεικόνιση (μαζί με άλλες σχετικές με το θέμα) οφείλονται στον Dana Scott. Δες την εργασία του “Domains for denotational semantics”, *Proceeding of ICALP '82, Lecture in Computer Science*, Volume 140, Springer-Verlag, 1982.
8. Τα πλεονεκτήματα της προτεινόμενης αλγεβρικής θεωρίας είναι αναρίθμητα. Κάποιος μπορεί από τα δηλωτικά αξιώματα να συναγάγει τα «αξιώματα» του Dretske (π.χ. the Xerox Principle) χωρίς να εμπλακεί στα πολύπλοκα μονοπάτια της στατιστικής. Στη θεωρία μου για τις πληροφορίες δεν υπάρχει καμμία προϋπόθεση ότι αυτός ο μηχανισμός ανταλλαγής πληροφοριών είναι μία φυσική διαδικασία, όπως είναι για τον Dretske. (Η θεώρηση μόνο των πληροφοριών δεν θα έπρεπε να απέκλειε την πιθανότητα οι άγγελοι να είναι μεταφορές πληροφοριών. Στην περιγραφή του Dretske δεν υπάρχει τρόπος για μερική ή ατελή ανταλλαγή πληροφοριών, ενώ η μερικότητα είναι βασικό στοιχείο της αλγεβρικής προσέγγισης).
9. Πρέπει αυτό να εξετασθεί με προσοχή· δεν υποστηρίζω ότι ακόμα και τα βαλτόμετρα ή οι θερμοστάτες έχουν κάποιες πίστεις, έστω και πολύ υποτυπώδεις, διαφορετικά από την ανθρώπινη νοητική δομή, η δομή της πορείας πληροφοριών των βαλτομέτρων δεν είναι αρκετά περίπλοκη γι' αυτό.
10. Βεβαίως και έπρεπε να το πει. Το να αποδεχόταν το αντίθετο θα ήταν σα να έθετε εν αμφιβόλω ένα θεμελιώδες αίτημα του ορθολογισμού.
11. Αυτό δεν πρέπει να θεωρηθεί άρνηση του εμφανούς γεγονότος ότι ορισμένα είδη κακής πνευματικής λειτουργίας είναι υπεύθυνα για ορισμένα είδη λαθών.
12. Ανεξέλεγκτες αντιδράσεις σε οπτικές απάτες, όπως το διάγραμμα Muller-Lyer, δίνουν άλλο ένα παράδειγμα του λάθους που δεν οφείλεται σε δυσλειτουργία.
13. Το περιεκτικό έργο του Hartley Rogers, *Theory of Recursive Functions and Effective computability*, McGraw Hill, 1967, είναι η καλύτερη πηγή για περισσότερες πληροφορίες για την αδυναμία λύσεως.
14. Λέω «Μιλώντας κατά προσέγγιση» γιατί οι επιστήμονες που εξετάζουν την περιοχή των δυσεπίλυτων προβλημάτων και τα μαθηματικά της περίχωρα έχουν διακρίνει διαφορετικά είδη δυσεπίλυτων προβλημάτων και υπάρχει διαφωνία μεταξύ τους ποια από αυτά είναι «πραγματικά δυσεπίλυτα». Εντούτοις όταν ερχόμαστε σε συγκεκριμένες περιπτώσεις η διαφωνία περιορίζεται. Μία τέτοια περίπτωση δίνει το θεώρημα Meyer και

Stockmeyer: υπάρχει μία πρόταση αδυνάμου (weak) δευτέρας τάξεως αριθμητικής εκφραστή με 675 σύμβολα: η εξέτασή της θα απαιτούσε 20 δισεκατομμύρια χρόνια υπολογισμού.

15. Το πρόβλημα των πρώτων αριθμών, μαζί με τον «πλανόδιο πωλητή» και τα προβλήματα του Hamilton, συμπεριλαμβάνονται στην κατηγορία των «NP προβλημάτων». Αυτά τα προβλήματα είναι επιλύσιμα σε χρόνο ο οποίος είναι εκθετική συνάρτηση του μεγέθους των δεδομένων, αλλά όχι προφανώς σε χρόνο ο οποίος είναι πολυώνυμο του μεγέθους των δεδομένων. Το πρόβλημα των Meyer και Stockmeyer (σημ. 14) είναι ακόμα χειρότερο: λύνεται μόνο σε ό,τι ονομάζεται «υπέρ-εκθετικός» χρόνος.
16. Ο αλγόριθμος του Christofides για το πρόβλημα του «πλανόδιου πωλητή» εγγυάται ένα αποτέλεσμα του οποίου το μήκος δεν είναι περισσότερο από το 50% από το μήκος της πραγματικής λύσης. Πβ. Lewis και Papadimitriou, "The efficiency of algorithms", *Scientific American*, Volume 238, Ιανουάριος 1978, σσ. 96-109.
17. Για λεπτομέρειες δες το άρθρο του D.J.A. Welsh, "Randomized algorithms", *Discrete Applied Mathematics*, volume 5, 1983, σσ. 133-145.
18. Γι' αυτά τα αποτελέσματα και την προβληματική, δες P. Wason και P. Johnson-Laird, *Psychology of Reasoning*, Harvard, 1972, και A. Tversky και D. Kahneman, "Judgement under uncertainty: heuristics and biases", *Science*, CLXXXV, 1974, σσ. 1124-31.

* Η μετάφραση στην ελληνική έγινε από τη δεσποινίδα Βασιλική Μπούκη και τον κύριο Γιάννη Στεφάνου.

DR CHARLES McCARTHY
CENTRE FOR COGNITIVE SCIENCE
DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF EDINBURGH