

ΕΚΤΑΚΤΟΣ ΣΥΝΕΔΡΙΑ ΤΗΣ 29ΗΣ ΙΑΝΟΥΑΡΙΟΥ 1991

ΠΡΟΕΔΡΙΑ ΙΩΑΝΝΟΥ ΤΟΥΜΠΑ

ΙΣΧΥΡΗ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ
(Strong Artificial Intelligence)

ΟΜΙΛΙΑ ΤΟΥ ΑΚΑΔΗΜΑΪΚΟΥ Κ. ΝΙΚΟΛΑΟΥ Κ. ΑΡΤΕΜΙΑΔΗ

Κύριε Πρόεδρε,
Κύριοι συνάδελφοι,
Κυρίες και Κύριοι,

Μια περιοχή της Μαθηματικής 'Επιστήμης ή όποια έδω και αρκετά χρόνια προ-
σελκνύει τὸ μεγάλο ενδιαφέρον τῶν ἐπιστημόνων καὶ τοῦ εὐρύτερου κοινοῦ εἶναι ἡ τῆς
Τεχνητῆς Νοημοσύνης (Artificial intelligence) (γιά συντομία ΤΝ).

Ὁ ἀντικειμενικός σκοπὸς τῆς ΤΝ εἶναι νὰ μιμηθεῖ, διὰ μέσου μηχανῶν, ὡς ἐπὶ
τὸ πλεῖστον ἠλεκτρονικῶν, τὴ νοητικὴ δράση τοῦ ἀνθρώπου καὶ πιθανὸν νὰ συντελέσει
στὴ βελτίωση τῶν ἀνθρωπίνων δυνατοτήτων.

Τὰ ἐπιτευχθέντα ἀποτελέσματα στὴν περιοχή τῆς ΤΝ ἐνδιαφέρουν κυρίως τὶς
ἀκόλουθες τέσσερις κατευθύνσεις ἔρευνας.

(α) Ἡ μελέτη κατασκευῆς ΡΟΜΠΟΤ. Ἡ ἔρευνα αὕτὴ ἀσχολεῖται κυρίως μὲ
ἀπαιτήσεις τῆς βιομηχανίας ποὺ ἀναφέρονται σὲ πρακτικούς σκοπούς, ἥτοι στὴν κατα-
σκευὴ μηχανῶν οἱ ὁποῖες θὰ ἐκτελοῦν ἔργα ποὺ ἀπαιτοῦν «νοημοσύνη», ἔργα πολὺπλο-
κα ποὺ ἀπαιτοῦσαν πρῶτα τὴν εὐστροφὴ παρέμβαση ἢ τὸν ἔλεγχο τοῦ ἀνθρώπου.
'Ενδιαφέρει οἱ μηχανὲς αὐτὲς νὰ ἐκτελοῦν τὸ ἔργο τους μὲ ὑπευθυνότητα καὶ πολὺ τα-
χύτερα ἀπὸ τὸν ἄνθρωπο, εἶναι δὲ αὐτὲς ιδιαίτερα χρήσιμες στὶς περιπτώσεις ἐκεῖνες
ποὺ ἡ ἐκτέλεση τοῦ ἔργου εἶναι ἐπικίνδυνη ἂν ἐπιχειρηθεῖ ἀπὸ τὸν ἄνθρωπο.

(β) Ἡ δεύτερη κατεύθυνση ἔρευνας ἀναφέρεται στὸ Ἐμπόριο. Ἐδῶ ἐπιδιώκε-

ται ή ανάπτυξη «Εἰδικῶν Μηχανικῶν Συστημάτων» τὰ ὁποῖα θὰ μποροῦν νὰ κωδικοποιοῦν τήν ὅλη οὐσιαστική γνώση ἑνὸς ἐπαγγέλματος — ἱατρικοῦ, νομικοῦ, κ.λπ. — κατὰ ἐνότητες σέ ἓνα ἡλεκτρονικό ὑπολογιστή. Καί ἀνακύπτει βέβαια τὸ ἐρώτημα: Θὰ εἶναι δυνατόν μιὰ τέτοια ἐνότητα νὰ ἀντικαθιστᾷ τήν πείρα καὶ τήν εἰδικότητα ἐκείνων ποὺ ἀσκοῦν τὰ ἐν λόγῳ ἐπαγγέλματα ἢ θὰ χρησιμεύει αὐτὴ ὡς ἀπλή πηγὴ πληροφοριῶν;

Γενικότερα, τὸ ἐὰν οἱ Ὑπολογιστὲς θὰ μποροῦν νὰ ἐπιδεικνύουν στὸ μέλλον «αὐθεντική νοημοσύνη», ἔχει προφανῶς ἀξιόλογες κοινωνικὲς συνέπειες.

(γ) Μιὰ τρίτη περιοχὴ ἔρευνας, ἡ ὁποία ἔχει σχέση μὲ τὴν TN, εἶναι ἡ Ψυχολογία. Οἱ ἐρευνητὲς ἐλπίζουν ὅτι μὲ τὶς προσπάθειες ποὺ καταβάλλουν νὰ μιμηθοῦν τὴ συμπεριφορὰ τοῦ ἀνθρωπίνου ἐγκεφάλου (ἢ τοῦ ἐγκεφάλου κάποιου ἄλλου ζωικοῦ ὄργανισμοῦ) χρησιμοποιώντας διάφορες ἡλεκτρονικὲς συσκευές — ἢ ἀκόμα ἀποτυγχάνοντας στὶς προσπάθειές τους — εἶναι δυνατόν νὰ προκύψουν χρήσιμες πληροφορίες σχετικὲς μὲ τὴ λειτουργία τοῦ ἐγκεφάλου.

(δ) Τέλος γιὰ λόγους παρόμοιους μὲ ἐκείνους ποὺ ἀνέφερα στὴν προηγούμενη περίπτωσι, ἡ ἔρευνα περὶ τὴν TN μπορεῖ νὰ μᾶς διαφωτίσει ἀναφορικά μὲ βαθιὰ φιλοσοφικά ἐρωτήματα, παρέχοντας νέες ἰδέες σχετικὲς μὲ τὴ σημασία τῶν ἐννοιῶν «πνεῦμα», «διάνοια», «συνείδηση» κ.λπ.

Μέχρι ποιοῦ σημείου ἡ TN ἔχει προχωρήσει στὶς τέσσερις αὐτὲς κατευθύνσεις ποὺ ἀνέφερα; Θὰ μοῦ ἦταν δύσκολο, ὡς μὴ εἰδικός, νὰ δώσω ἔστω καὶ μιὰ συνοπτικὴ εἰκόνα τῆς παρούσας καταστάσεως. Ὑπάρχουν πολλὲς ἐρευνητικὲς ὁμάδες μὲ ἀξιόλογα ἀποτελέσματα ἐπάνω στὰ θέματα αὐτὰ καὶ οἱ ὁποῖες εἶναι διεσπαρμένες σὲ ὅλα τὰ μέρη τοῦ Κόσμου. Ἴσως τὸ μόνο πράγμα ποὺ θὰ μπορούσα νὰ πῶ εἶναι ὅτι, παρὰ τὶς ἐπιτυχίες ποὺ ἔχουν σημειωθεῖ, ἡ μίμηση κάποιου πράγματος ποὺ θὰ μπορούσε νὰ χαρακτηρισθεῖ ὅτι ἀποτελεῖ «αὐθεντικὴ νοημοσύνη» εὐρίσκεται μακριὰ ἀπὸ τοῦ νὰ ἔχει ἐπιτευχθεῖ.

Ὅμως ἡ πρόθεσή μου σήμερα εἶναι νὰ ἀναφερθῶ σὲ μιὰ ἔννοια παραπλήσια τῆς TN, τὴν λεγόμενη Ἰσχυρὴ Τεχνητὴ Νοημοσύνη (ITN).

Κατὰ τὴ θεωρίαν τῆς ITN, τὰ μηχανικὰ κατασκευάσματα ποὺ ἀνέφερα παραπάνω δὲν διακρίνονται μόνο γιὰ τὴ νοημοσύνη τους, ἀλλὰ ἐπιπλέον οἱ λογικὲς αὐτὲς λειτουργίες τῶν μηχανῶν ἀποδίδονται καὶ σὲ πνευματικὲς ἱκανότητες, ἀκόμα καὶ ἂν τὸ μηχανικὸ κατασκεῦασμα εἶναι πολὺ ἀπλό, ὅπως π.χ. εἶναι ἓνας θερμοστάτης.

Ἡ ἰδέα αὐτὴ τῆς κατασκευῆς «σκεπτομένων» ὑπολογιστῶν» γοητεύει πολὺ κόσμον καὶ ἔχει ὡς ἀποτέλεσμα τὴν πώληση μεγάλου ἀριθμοῦ σχετικῶν βιβλίων καὶ περιοδικῶν.

Ἐδῶ καὶ μερικὲς δεκαετίες οἱ ὑποστηρικτὲς τῆς ITN ἰσχυρίζονται ὅτι μετὰ ἑκα-

τὸ ἢ διακόσια χρόνια, ἠλεκτρονικὲς μηχανὲς θὰ εἶναι σὲ θέσῃ νὰ κάνουν κάθε τι ποὺ ὁ ἄνθρωπος μπορεῖ νὰ κάνει. Πιστεύουν ὅτι τὸ ἀνθρώπινο μυαλὸ εἶναι ἓνας Ὑπολογιστὴς κατασκευασμένος ἀπὸ σάρκα καὶ ὑποκείμενος στοὺς φυσικοὺς νόμους, ἢ πλήρης γνώση τῶν ὁποίων θὰ μᾶς ἐπιτρέψει νὰ κατασκευάσουμε ὑπολογιστικὲς μηχανὲς οἱ ὁποῖες θὰ εἶναι σὲ θέσῃ νὰ λύνουν τὶς «διαφορικὲς ἐξισώσεις τοῦ ἐγκεφάλου» σὲ πεπερασμένο χρόνο, καὶ οἱ ὁποῖες θὰ συμπεριφέρονται ἀκριβῶς ὅπως ὁ ἄνθρωπος.

Ἰσχυρίζονται ἀκόμα, οἱ ἀσπαζόμενοι τὴν παραπάνω ἄποψη, ὅτι ὁ ἐγκέφαλος ἀποτελεῖται ἀπὸ ἓνα πεπερασμένο πλῆθος νευρονίων τὰ ὁποῖα ἀποτελοῦν ἓνα πάρα πολὺ πολύπλοκο ἱστὸ καὶ ὅτι κάθε ἓνα ἀπὸ τὰ νευρόνια αὐτὰ ἐκτελεῖ μιὰ ὀρισμένη λειτουργία ἢ ὁποῖα ἀκολουθεῖ τοὺς νόμους τῆς Φυσικῆς.

Ἐνα σύνολο μαθηματικῶν ἐξισώσεων συνδέει ὅλα τὰ σήματα ποὺ ἐμφανίζονται στὸν ἐν λόγῳ ἱστὸ. Ἐνας ἀρκετὰ ἰσχυρὸς ὑπολογιστὴς θὰ εἶναι σὲ θέσῃ, κατὰ τὸ πρότυπο τοῦ ἐγκεφάλου, νὰ λύνει τὶς ἐξισώσεις αὐτές.

Ὑποστηρίζουν οἱ ὁπαδοὶ τῆς ITN ὅτι, σὲ συστήματα τοῦ εἵδους αὐτοῦ «ἡ νοήμων συμπεριφορὰ» δημιουργεῖται μὲ τὴν ὁμαδικὴ λειτουργία τῆς ὅλης ἀρχιτεκτονικῆς δομῆς τοῦ συστήματος, ὁπότε κάθε μεμονωμένο νευρόνιο δὲν μπορεῖ νὰ ἔχει «ἀντίληψη» τῆς ὅλης καταστάσεως.

Ἡ βασικὴ ἰδέα, ὁ ἀκρογωνιαίος λίθος, ἐπὶ τοῦ ὁποίου ἐδράζεται ἡ ITN, εἶναι ὅτι ἡ λεγομένη πνευματικὴ δραστηριότητα δὲν εἶναι τίποτε ἄλλο παρὰ ἡ ἐκτέλεση μιᾶς συγκεκριμένης καὶ πλήρως ὀρισμένης ἀκολουθίας πράξεων, στὴν ὁποῖα ἀκολουθία πράξεων συχνὰ ἀναφερόμαστε μὲ τὸν ὄρο «ἀλγόριθμος».

Οἱ μηχανὲς αὐτές, προσθέτουν οἱ ὑποστηρικτὲς τῆς ITN, θὰ δοκιμάζουν συγκινήσεις, θὰ κρίνουν ποιὸ εἶναι ἀληθινὸ καὶ ποιὸ δὲν εἶναι, θὰ ἐκτιμοῦν τὴν ὁμορφιά, θὰ ἔχουν ἀντίληψη, θὰ ἔχουν αὐτοσυνειδησία καὶ εὐφῶτα.

Μερικοὶ μάλιστα ἀπὸ τοὺς φανατικοὺς φτάνουν μέχρι τοῦ σημείου νὰ πιστεύουν ὅτι οἱ ἐν λόγῳ μηχανὲς θὰ εἶναι καθόλα καλύτερες ἀπὸ ἐμᾶς τοὺς ἴδιους ποὺ τὶς κατασκευάσαμε καὶ ὅτι, ἐνδεχομένως, θὰ διαδεχθοῦν τὸν *Homo Sapiens* στὴν ἐξελικτικὴ κλίμακα τῶν εἰδῶν. Ἐρωτώμενοι, ἂν οἱ Μηχανὲς αὐτὲς θὰ σκέπτονται, ἀπαντοῦν: «Καὶ βέβαια θὰ σκέπτονται. Εἴμαστε καὶ ἐμεῖς μηχανὲς καὶ σκεπτόμαστε. Δὲν νομίζετε λοιπὸν ὅτι καὶ αὐτὲς θὰ μποροῦν νὰ σκέπτονται;»

Μὲ ἄλλα λόγια κατὰ τὴν ITN, ἡ διαφορὰ μεταξὺ τῆς οὐσιαστικῆς λειτουργίας τοῦ ἀνθρώπινου ἐγκεφάλου (συμπεριλαμβανομένων ὅλων τῶν ἐνσυνειδήτων ἐκδηλώσεων) καὶ ἐκείνης ἐνὸς θερμοστάτη, ἔγκειται στὴν πολὺ πρὸ πολύπλοκη δομὴ τοῦ ἐγκεφάλου.

Καὶ κάτι ποὺ εἶναι πολὺ πρὸ σπουδαῖο, εἶναι ὅτι ὅλες οἱ πνευματικὲς ιδιότητες

—σκέπτεσθαι, αισθάνεσθαι, εὐφυνᾶ, ἀντίληψη, συνείδηση — θεωροῦνται ἀπὸ τῆς σκοπιᾶς τῆς ITN ἀπλῶς ὡς διαφορετικές ὕψεις τῆς ὅλης πολύπλοκης λειτουργίας τοῦ ἐγκεφάλου ἦτοι, ὅτι οἱ ιδιότητες αὐτές εἶναι χαρακτηριστικά γνωρίσματα τοῦ ἀλγορίθμου τὸν ὁποῖον ἀκολουθεῖ ὁ ἐγκέφαλος.

Πολλοὶ θετικοὶ ἐπιστήμονες καὶ φιλόσοφοι διαφωνοῦν ριζικὰ καὶ ἔντονα μὲ τοὺς ὑποστηρικτὰς τῶν παραπάνω ἀπόψεων. Ἀντιθέτως, πιστεύουν αὐτοὶ ὅτι οἱ Ὑπολογιστὲς δὲν διαφέρουν ἀπὸ τὶς ὑπόλοιπες μηχανὲς πὺν ἔχει ὁ ἄνθρωπος στὴ διάθεσή του, παρὰ τὸ γεγονὸς ὅτι ἠλεκτρονικὲς μηχανὲς μποροῦν σὲ δεδομένο χρονικὸ διάστημα νὰ ἐκτελέσουν ἔργα πάρα πολὺ πολύπλοκα. Πιστεύουν, οἱ ὡς ἂν ἐπιστήμονες, ὅτι οἱ ἔννοιες «ἀντίληψη» καὶ «σκέψη» δὲν ἔχουν κανένα νόημα γιὰ τὶς μηχανές. Μπορεῖ ὁ Ὑπολογιστὴς νὰ παίξει σκάκι, καὶ μάλιστα τῆς στάθμης ἐνὸς πρωταθλητῆ, ὅμως δὲν ἔχει τὴν ἀντίληψη τῆς ἐννοίας σκάκι, δὲν ἔχει συναίσθηση ὅτι παίξει σκάκι. Εἶναι ἀπλῶς προγραμματισμένος νὰ ἐκτελέσει ἓνα ὁρισμένο ἔργο.

Προσπάθησα νὰ παραθέσω, σὲ πολὺ γενικὲς γραμμὲς καὶ ὅσο ἀπλούστερα μποροῦσα, τὶς ἀπόψεις τῶν δύο αὐτῶν «παρατάξεων», ἡ μεταξὺ τῶν ὁποίων «διαμάχη» ἔλαβε (τὸ 1990) μεγάλη ἔκταση καὶ ἔξαρση, κατὰ πὺν ἀπετέλεσε τὴ βασικὴ αἰτία γιὰ τὴν ἐπιλογὴ τοῦ θέματος τῆς σημερινῆς μου ὁμιλίας, κατὰ τὴν ὁποία θὰ προσπαθῶ νὰ ἐκθέσω καὶ μερικὲς προσωπικὲς μου ἀπόψεις ἐπάνω στοῦ καντὸ αὐτὸ θέμα. Ἐκτὸς ἀπὸ τὰ πολλαπλὰ δημοσιεύματα καὶ συνέδρια πὺν ἔλαβαν χώρα τελευταῖα, τὴν ἔξαρση πὺν ἀνέφερα παραπάνω τὴν ἀποδίδω κυρίως στὰ ἀκόλουθα δύο γεγονότα :

Τὸ πρῶτο γεγονὸς εἶναι ἓνα εἶδος δημόσιας συζήτησης πὺν ἐπιδιώχθηκε καὶ ἔγινε μὲ τὴ δημοσίευση δύο ἄρθρων, τὸ ἓνα δίπλα στοῦ ἄλλο, στοῦ ἴδιο τεύχους (Ἰανουάριος 1990) τοῦ περιοδικοῦ *Scientific American*.

Τὸ πρῶτο ἄρθρο (σ. 20-25) εἶναι τοῦ John R. Searle, καθηγητοῦ τῆς φιλοσοφίας στοῦ Πανεπιστήμιου τῆς Καλιφόρνια στοῦ Berkeley, καὶ ὁ ὁποῖος δὲν εἶναι ὁπαδὸς τῆς ITN. Τὸ δεύτερο ἄρθρο (σ. 26-31) εἶναι τῶν Paul καὶ Patricia Churchland, καθηγητῶν τῆς φιλοσοφίας στοῦ Πανεπιστήμιου τῆς Καλιφόρνια στοῦ San Diego, οἱ ὁποῖοι εἶναι ὁπαδοὶ τῆς ITN.

Μολονότι τὰ ἐπιχειρήματα πὺν παρουσιάζουν οἱ συγγραφεῖς αὐτοὶ εἶναι ὄντως συγκλονιστικά καὶ πειστικά, συμβαίνει κατὰ τὸ παράδοξο: καταλήγουν αὐτὰ σὲ διαμετρικῶς ἀντίθετα συμπεράσματα. Κάθε μία «παρατάξη» εἶναι βεβαία ὅτι ἔχει «δικαίον» καὶ φαίνεται νὰ παραμένει τελείως ἀνεπηρέαστη ἀπὸ τὰ ἐπιχειρήματα τῆς «ἀντιπάλου παρατάξεως».

Τὸ δεύτερο γεγονὸς εἶναι ἡ ἐμφάνιση ἐνὸς βιβλίου πὺν κυκλοφόρησε τὸ 1990 μὲ τίτλο «*The Emperor's new mind*» καὶ τοῦ ὁποῖου συγγραφεὺς εἶναι ὁ Καθηγητὴς

Roger Penrose τοῦ Πανεπιστημίου τῆς Ὁξφόρδης. Ὁ Penrose εἶναι μαθηματικὸς καὶ φυσικὸς. Τὰ εἰς τὸ βιβλίον αὐτὸ ἀναγραφόμενα ἀποτελοῦν τὴ σοβαρότερη ἐπίθεση κατὰ τῆς ITN πὺν ἔχει ἐπιχειρηθεῖ μέχρι σήμερα. Προκειται γιὰ ἓνα βιβλίον πὺν ἂν καὶ δὲν εἶναι ἐντολοδιάβαστο παρουσιάζει ὅμως πολὺ μεγάλο ἐνδιαφέρον.

Ὁ Penrose ἰσχυρίζεται ὅτι τὸ ἀνθρώπινο πνεῦμα, ἡ ἀνθρώπινη διάνοια, ἔχει τέτοια ὀξυδέρκεια, τόση διορατικὴ πὺν ἓνας Ὑπολογιστὴς δὲν θὰ μπορέσει ποτὲ νὰ ἐπιδείξει, γιὰ τὸ λόγο ὅτι ὁ Ὑπολογιστὴς λειτουργεῖ ἀκολουθώντας ἓναν ἀλγόριθμον (ἓνα σύνολον κανόνων βάσει τῶν ὁποίων λύνονται τὰ προβλήματα), ὅμως, ὅπως θὰ ἐξηγήσουμε λεπτομερέστερα παρακάτω, ἔχει ἀποδειχθεῖ ὅτι ὑπάρχουν προβλήματα τῶν ὁποίων ἡ λύση δὲν μπορεῖ νὰ ἐπιτευχθεῖ με κανένα σύστημα κανόνων, με κανένα ἀλγόριθμον.

Ἐπανελημμένα ὁ Penrose τονίζει ὅτι οἱ πνευματικὲς λειτουργίες εἶναι φύσει ἰσχυρότερες ἀπὸ τὶς ὑπολογιστικὲς ἱκανότητες πὺν ἔχει οἰοσδήποτε ὑπολογιστὴς. Παρατηρεῖ ὅτι τὰ περισσότερα ἀπὸ τὰ προβλήματα πὺν παρουσιάζονται στὶς ἐπιστῆμες δὲν εἶναι ἐπιδεκτικὰ ἀλγοριθμικῆς λύσεως. Σταθερὰ ὁ ἄνθρωπος ἐπινοεῖ λύσεις προβλημάτων γιὰ τὶς ὁποῖες δὲν ὑπάρχει ἓνας ἀκολουθητέος γενικὸς ἀλγόριθμος.

Πῶς εἶναι δυνατόν, διερωτᾶται ὁ Penrose, ἓνας ἀλγόριθμος νὰ ἀνακαλύψει θεωρήματα τὰ ὁποῖα βεβαιώνουν τί ἓνας ἀλγόριθμος δὲν μπορεῖ νὰ ἀνακαλύψει; Θὰ δοῦμε παρακάτω ὅτι ὑπάρχουν τέτοια θεωρήματα. Καὶ ὁ Penrose προσθέτει ἐπιγραμματικά: «Τὸ γεγονὸς καὶ μόνο ὅτι ὁ ἐγκέφαλος με ὁδηγεῖ σὲ ἀλήθειες, στὶς ὁποῖες δὲν μπορεῖ νὰ με ὁδηγήσει κανένας ὑπολογιστὴς, με πείθει ὅτι ὁ Ὑπολογιστὴς δὲν θὰ μπορέσει ποτὲ νὰ ὑποκαταστήσει τὸν ἄνθρωπον».

Στὸ ἐκ 466 σελίδων ἀποτελούμενο βιβλίον του ὁ Penrose κάνει μιὰ ἐντύπωση ἀναφορὰ στὶς ἔννοιες: πνεῦμα - συνείδη - σκέψη. Ὁδηγεῖ τὸν ἀναγνώστη σὲ μιὰ Ὀδύσειαν θεμάτων ὅπως εἶναι οἱ Ἀλγόριθμοι — οἱ Μηχανὲς Turing (ἄφηρημένες μαθηματικὲς ἔννοιες χρήσιμες στὴν Πληροφορικὴ) — τὰ Σύνολα τοῦ Mandelbrot [2] — ἡ Μηχανικὴ τοῦ Newton — ἡ Κβαντομηχανικὴ — ὁ Ἐγκέφαλος — ἡ Νευροφυσιολογία — καὶ διάφορα ἄλλα. Στὸ κάθε ἓνα ἀπὸ αὐτὰ τὰ θέματα εὗρισκει ὁ Penrose ἀφρονία ἐνδείξεων ἀνθρώπινων δράσεων οἱ ὁποῖες δὲν εἶναι ἀλγοριθμικὲς, καὶ ὁδηγεῖται στὸ συμπέρασμα ὅτι γιὰ νὰ γίνῃ καλύτερα ἀντιληπτὴ ἡ πνευματικὴ λειτουργία τοῦ ἀνθρώπου πρέπει νὰ περιμένουμε νὰ ἀναπτυχθεῖ περισσότερο ἡ Θεωρία τῶν Κβάντα στὴ Φυσική.

Ὅμως ὥς προσπαθήσουμε νὰ βγάλουμε καὶ μερικὰ συμπεράσματα ὥς τρίτοι παρατηρητές.

Ἡ μελέτη τῶν ἐπιχειρημάτων τῶν δύο ἀντιμαχομένων «παρατάξεων» πὺν ἀνέ-

φερα παραπάνω οδηγεί στην άποψη ότι το όλο πρόβλημα, σε τελευταία ανάλυση, συνοψίζεται σε ένα και μόνο ερώτημα:

«Θα είναι ποτέ ο ύπολογιστής σε θέση να σκέπτεται;»

Οι εν λόγω παρατάξεις απαντούν βέβαια στο ερώτημα αυτό, ή μία θετικά και ή άλλη αρνητικά. Όμως, φαίνεται ότι, και οι δύο πλευρές αντιπαρόχονται και δεν ασχολούνται με το ερώτημα: «ΤΙ ΣΗΜΑΙΝΕΙ ΣΚΕΠΤΟΜΑΙ;», και τούτο διότι θεωρούν το ερώτημα αυτό ως στερούμενο νοήματος.

Όμως, όπως και ο Penrose τονίζει, το ερώτημα αυτό εδρύσκεται στην καρδιά, είναι το νευραλγικό σημείο της συζήτησης των δύο «μονομάχων» του περιοδικού *Scientific American* που ανέφερα προηγουμένως.

Το τί πιστεύουμε ότι σημαίνει ή λέξη «σκέπτομαι» αποτελεί ένα κινούμενο στόχο δια μέσου της Ιστορίας. Έπὶ δύο εκατονταετίες περίπου, από τις αρχές του 1700 και υπό την επιρροήν της Μηχανικής του Newton, όλοι πίστευαν ότι το Σύμπαν ήταν ένα υπέροχο ωρολογιακό σύστημα διεπόμενο από μερικούς απλούς νόμους. Τὴν ἐποχὴ ἐκείνη ὁ ἄνθρωπος προσπαθοῦσε νὰ διερευνήσει τὸ Σύμπαν ἀκολουθώντας τὴ Συμπερασματικὴ Λογικὴ. Ἡ ἀναζήτησις, ἡ ἀνάγκη νὰ διαμορφωθεῖ μιὰ πλήρης ἀντίληψις τῆς ἔννοιᾳς «σκέψῃ» ὡδήγησε κατὰ τὸ 1800 μ.Χ. τοὺς ἐπιστήμονες στὸ νὰ ἐπιχειρήσουν νὰ διατυπώσουν ἕνα Γενικὸ Σύστημα Λογικῆς, στὸ ὁποῖο ὁποιαδήποτε πρότασις νὰ εἶναι δυνατὸν νὰ ἐλέγχεται ὡς ἀληθὴς ἢ ψευδής.

Ἡ προσπάθεια αὐτὴ διατυπώσεως ἑνὸς τέτοιου λογικοῦ συστήματος κατέρρευσε τὸ 1930 μετὰ τὴ δημοσίευσιν ἑνὸς θεωρήματος ἀπὸ τὸν 25ετὴ τότε αὐστριακὸ μαθηματικὸ τοῦ Πανεπιστημίου τῆς Βιέννης Kurt Gödel. Τὸ θεώρημα αὐτό, σὲ γενικὲς γραμμές, βεβαιώνει ὅτι: «Σὲ ὁποιοδήποτε μαθηματικὸ σύστημα ὑπάρχουν προτάσεις οἱ ὁποῖες εἶναι μὲν ἀληθεῖς, ὅμως ἡ ἀλήθεια αὐτῶν δὲν μπορεῖ νὰ ἀποδειχθεῖ ἐντὸς τοῦ συστήματος, μετὰ τὴ χρῆσιν δηλαδὴ μόνο τῶν κανόνων πὺν ἰσχύουν στὸ ἐν λόγω σύστημα». Εἶχα τὴν εὐκαιρίαν νὰ μιλήσω διεξοδικότερα γιὰ τὸ Θεώρημα τοῦ Gödel στὴν ὁμιλίαν μου «Κορυφαῖες στιγμὲς στὰ Μαθηματικά» [3]. Τὸ θεώρημα τοῦ Gödel ἀποδυναμώνει τὴν ἄποψιν ὅτι «σκέψῃ» εἶναι μιὰ μηχανικὴ, μιὰ ἀλγοριθμικὴ διαδικασία.

Σήμερα ἐπικρατεῖ μιὰ διαφορετικὴ ἐρμηνεία τῆς ἔννοιᾳς «σκέπτομαι» μετὰ τὴν ὁποία καὶ συντάσσομαι. Πολλοὶ πιστεύουν ὅτι «σκέψῃ» ΔΕΝ εἶναι ἡ διαδικασία ἐφαρμογῆς τῆς Συμπερασματικῆς Λογικῆς, ἀλλὰ ἡ δημιουργία νέων ἰδεῶν. Ἡ Συμπερασματικὴ Λογικὴ μοιάζει μετὰ μηχανικὴ λειτουργία.

Ἄν ἀναπολήσομε τίς στιγμὲς κατὰ τίς ὁποῖες βρισκόμαστε σὲ κατάστασις ἐνοράσεως, ἀποφαινόμεστε ὅτι διατελούσαμε, κατὰ τίς στιγμὲς αὐτές, ὑπὸ τὸ κράτος μιᾶς συγκινησιακῆς καταστάσεως, καὶ ὅτι πρὶν ἀκόμα ἐκφράσομε μετὰ λέξεις αὐτὸ πὺν

δραματισθήκαμε, αυτό που δημιουργήσαμε κατά τις στιγμές εκείνες, δοκιμάσαμε μιὰ ιδιαίτερη σωματική αίσθηση. Μὲ ἄλλα λόγια θεωροῦμε τὴ «σκέψη» ὡς ἓνα φαινόμενο τὸ ὁποῖο λαμβάνει χώραν προτοῦ τὸ προῖδ' τῆς ἐνοράσεώς μας ἐκφρασθεῖ διὰ τοῦ λόγου, διὰ τῆς γλώσσας, ἐνῶ ἀντιθέτως οἱ Ὑπολογιστὲς οἱ ὁποῖοι εἶναι προγραμματισμένοι μέσα στὴ γλώσσα δὲν εἶναι σὲ θέση νὰ δράσουν ἐκτὸς τῆς γλώσσας, ἐκτὸς δηλαδὴ τῶν δυνατοτήτων γλωσσικῆς ἐκφράσεως.

Ἀναφερόμενος πάντοτε στὸν κινούμενο στόχο πὺν ἀποτελεῖ ἡ ἔννοια «σκέπτομαι», θὰ ἔλεγα ὅτι δὲν ὑπάρχει ἀμφιβολία ὅτι μετὰ πενήντα ἢ ἑκατὸ χρόνια θὰ ὑπάρξουν μηχανεὲς οἱ ὁποῖες θὰ ἐπιτελοῦν ἔργα τὰ ὁποῖα σήμερα θὰ ἀποδίδαμε στὸ «σκέπτεσθαι», παρὰ ταῦτα ὅμως οἱ ἄνθρωποι καὶ τότε θὰ τὶς θεωροῦν ἀπλῶς ὡς μηχανεὲς. Αὐτὸ σημαίνει ὅτι ἡ ἐρμηνεία τῆς ἔννοιας «σκέψη» θὰ ἔχει «μετατοπισθεῖ» τότε καὶ θὰ ἐδρίσκεται «μακρότερα» ἀπὸ ἐκεῖ πὺν ἐδρίσκεται τώρα, ἔτσι ὥστε νὰ διατηρεῖται πάντα μιὰ σαφὴς διάκριση μεταξὺ τοῦ ἀνθρωπίνου ὄντος καὶ τῆς μηχανῆς. Παρατηροῦμε ἐδῶ ὅτι βρισκόμαστε ἐνώπιον μιᾶς «ὀριακῆς» διαδικασίας, μιᾶς διαδικασίας ἡ ὁποία ὁδηγεῖ «ὀριακὰ» ἂν θέλετε τὴ μηχανὴ στὴν ἀνθρώπινη φύση, κάτι πὺν δημιουργεῖ τὸ αἶσθημα ὅτι ὁ Ὑπολογιστὴς δὲν θὰ ὑποκαταστήσει ποτὲ τὸν ἄνθρωπο.

Ἐπανέρχομαι πάλι στὸ Penrose τοῦ ὁποῖου τὰ ἐπιχειρήματα ἐνισχύουν τὴν παραπάνω ἄποψη. Ἡ ITN πρесеβεύει ὅτι, ὅταν κάποτε θὰ ἀποκτήσουμε μιὰ καλύτερη ἀντίληψη τῶν νόμων τῆς Φυσικῆς καθὼς καὶ τῆς δομῆς τοῦ ἐγκέφαλου, τότε θὰ μπορέσουμε νὰ «ἀναπαραγάγομε» τὸν ἐγκέφαλο διὰ μέσον ἑνὸς ὑπολογιστῆ. Ὅμως τί εἶναι ἡ Φυσική; Εἶναι ἡ Φυσικὴ ἱκανὴ νὰ ἀντιληφθεῖ πλήρως τὰ φαινόμενα; Τί σημαίνει ἀκριβῶς «ἀπομίμησης»;

Ἡ ἀνεπάρκεια τῆς Μηχανικῆς τοῦ Newton ὁδήγησε στὴ Θεωρίαν τῆς Σχετικότητος καὶ στὴν Κβαντομηχανικὴ, δὲν ἀποκλείεται δὲ οἱ νόμοι τῆς Φυσικῆς πὺν ἀφοροῦν τὰ Κβάντα νὰ μὴν εἶναι ἀλγοριθμικοί. Ἐξάλλου ἐπειδὴ μερικὰ νοητικὰ φαινόμενα λειτουργοῦν σὲ κλίμακα ὅπου κβαντικὰ φαινόμενα μπορεῖ νὰ ἐπιδροῦν ἐπ' αὐτῶν, οἱ λειτουργίαι πὺν «παριστάνουν» τὴ νοημοσύνη μπορεῖ νὰ μὴν εἶναι ἀλγοριθμικέ, ὁπότε ἡ μηχανικὴ ἀπομίμησή τους νὰ μὴν εἶναι δυνατή. Ἄς γίνουμε κάπως σαφέστεροι.

Ἡ Θεωρία τῆς Σχετικότητος τοῦ Einstein ἡ ὁποία ἀναφέρεται στὴ Βαρύτητα καὶ ἡ Κβαντομηχανικὴ ἡ ὁποία ρυθμίζει τὸ μικρόκοσμο, τὸν κόσμο τῶν σωματιδίων, εἶναι δύο θεωρίαι οἱ ὁποῖαι πρὸς τὸ παρὸν παραμένουν μαθηματικῶς ἀσυμβίβαστες. Γιὰ τὸ λόγο αὐτὸ οἱ ἐπιστήμονες-ἐρευνητὲς προσπαθοῦν νὰ συμβιβάσουν τὶς θεωρίαι αὐτὲς μὲ τὴ δημιουργία μιᾶς Κβαντικῆς Θεωρίας τῆς Βαρύτητας. Μιὰ συνέπεια τοῦ συμβιβασμοῦ αὐτοῦ θὰ ἦταν ὁ προσδιορισμὸς τῶν ὁρίων τῆς Κβαντομηχανικῆς ἡ ὁποία μεταξὺ ἄλλων μᾶς βεβαιώνει ὅτι σωματίδια μποροῦν ἀπότομα νὰ μετακινηθοῦν ἀπὸ ἓνα σημεῖο τοῦ χώρου σὲ κάποιο ἄλλο χωρὶς νὰ διασχίσουν τὸ μεταξὺ αὐτῶν διάστη-

μα! Μήπως τὸ τελευταῖο αὐτὸ φαινόμενο πὸν λαμβάνει χώραν μεταξὺ τῶν ἀτόμων τῆς ὕλης λαμβάνει ἐπίσης χώραν μεταξὺ μεγαλυτέρων σωματιδίων ὅπως εἶναι τὰ κύτταρα τοῦ ἐγκεφάλου; Μήπως ἡ μελλοντική, ἡ ἀναμενόμενη αὐτὴ Κβαντικὴ Θεωρία τῆς Βαρύτητας, εὐρίσκεται πίσω ἀπὸ τὴν ἐρμηνεία τῶν ἐννοιῶν «Συνείδηση», «Σκέψη», οἱ ὁποῖες τόσο μεγάλο καὶ βασικὸ ρόλο παίζουν στὸ θέμα πὸν μᾶς ἀπασχολεῖ; Ὡς περιμένουμε λοιπὸν τὴν περαιτέρω ἀνάπτυξη τῶν Κβαντικῶν Θεωριῶν τῆς Φυσικῆς.

Τελειώνοντας θὰ ἤθελα νὰ προσθέσω καὶ ἓνα ἀκόμη ἐπιχείρημα τὸ ὁποῖο δὲν συνηγορεῖ ὑπὲρ τῆς ITN. Τὸ ἐπιχείρημα βασίζεται σὲ ὀρισμένες ἀπόψεις καὶ σκέψεις πὸν εἶχα ἀναπτύξει σὲ προγενέστερή μου ὁμιλία κατὰ τὴν ἡμέρα τῆς ἐπισήμου ὑποδοχῆς μου στὴν Ἀκαδημία [1]. Οἱ σκέψεις ἐκεῖνες ὀδηγοῦν, κατ' ἐμέ, στὸ «συμπέρασμα» ὅτι ἡ νόηση κάθε ἀτόμου λειτουργεῖ στὰ πλαίσια ἐνὸς συστήματος ἐρμηνείας, τὸ ὁποῖο σύστημα βασίζεται καὶ περιλαμβάνει προκαταλήψεις, προλήψεις, πολιτισμὸ, ἱστορία καθὼς καὶ διάφορες ἀξίες. Τὸ σύστημα αὐτὸ ἐπηρεάζει ἐκεῖνο πὸν μπορεῖ νὰ δεῖ ἢ νὰ μὴ δεῖ ἓνα ἄτομο. Ἐπηρεάζει τὸ τί θὰ θεωρήσει τὸ ἄτομο σπονδαῖο ἢ ὄχι, τὸ τί εἶναι ἀληθὲς ἢ ψευδές. Ἐξάλλου, ὅπως εἶχα παρατηρήσει τότε στὴν ὁμιλία μου, στὴν πράξη, μιὰ πρόταση γίνεται δεκτὴ ἀπὸ μιὰ κοινότητα ἀτόμων ὥς ἀληθὴς ὅταν κανεὶς δὲν εἶναι σὲ θέση νὰ παρουσιάσει ἓνα πειστήριον ἢ κάποιο λογικὸ ἐπιχείρημα τὸ ὁποῖο θὰ πείσει κάποιον νὰ διαφωνήσει.

Ὡς ἐνσυνείδητα ὄντα ἀντιμετωπίζουμε διαρκῶς διαφορετικὰς ἐρμηνεῖς τοῦ ἰδίου φαινομένου, ὅπως συμβαίνει στὴν περίπτωση τοῦ Searle καὶ τῶν Churchlands, οἱ ὁποῖοι εἶναι διαφορετικοὶ παρατηρητὲς τῆς ITN καὶ καταλήγουν σὲ διαφορετικὰ συμπεράσματα γιὰ τὸ ἴδιο φαινόμενο.

Ἡ ὑπαρξὴ αὐτῆς τῶν πολλαπλῶν ἐρμηνειῶν ἐνισχύει τὶς θέσεις τῶν ἀντιτιθεμένων στὴν ITN. Ὅπως συμβαίνει καὶ μὲ ἓνα λογικὸ σύστημα, δὲν μπορεῖ νὰ δοθεῖ μιὰ μοναδικὴ ἐρμηνεία γιὰ ὅλα τὰ φαινόμενα. Οἱ δυνάμεις τῆς ἐνσυνείδητης παρατήρησης πὸν διαθέτει ὁ ἄνθρωπος τοῦ δύνανται τὴν ἰκανότητα νὰ μὴν περιορίζεται σὲ μιὰ συγκεκριμένη ἐρμηνεία ἀλλὰ νὰ δημιουργεῖ καὶ ἄλλες ἐναλλακτικὰς ἐρμηνεῖς. Κατὰ συνέπειαν, ἀφοῦ ἡ Συνείδηση δὲν μπορεῖ νὰ παρουσιασθεῖ σὲ μιὰ μόνο σταθερὴ περιγραφή ἢ ἐρμηνεία, πῶς εἶναι δυνατὸν αὐτὴ νὰ περιορισθεῖ σὲ ἓνα ἀλγόριθμο; Τὸ τελευταῖο αὐτὸ ἐπιχείρημα ἰσχύει ἐπίσης καὶ γιὰ τοὺς ἀλγορίθμους οἱ ὁποῖοι ἔχουν τὴ δυνατότητα νὰ μεταβάλλουν τὶς ἐρμηνεῖς πὸν παρέχουν καὶ τοῦτο διότι οἱ ἐν λόγῳ μεταβολὰς ἀποτελοῦν καὶ αὐτὲς οἱ ἴδιες ἐπίσης κάποια ἐρμηνεία.

Μολονότι τὰ παραπάνω ἐπιχειρήματα ἀποτελοῦν πρὸς τὸ παρὸν ἐνδείξεις μόνον καὶ ὄχι ἀποδείξεις τῶν ἀπόψεων τὶς ὁποῖες ὑποστηρίζουν, ὅμως ἀποτελοῦν ἀναμφισβήτητα μιὰ ὀξεῖα μεταφυσικὴ πρόκληση πρὸς τοὺς ὑποστηρικτὰς τῆς ITN, πρόκληση πὸν ἐπεβάλλετο νὰ γίνῃ δεδομένης τῆς ἀπόψεως πὸν ἐπικρατεῖ, καὶ πὸν κατὰ πολλοὺς

υπερισχύει στο φιλοσοφικό ρεύμα της εποχής μας, ότι ή σκέψη είναι, βασικά, τὸ ἴδιο πράγμα μὲ τὴ δράση ἑνὸς πολὺπλοκου ὑπολογιστῆ. Ἡ ἄποψη αὐτὴ δὲν μπορεῖ νὰ ἀνατραπεῖ μὲ ἀοριστολογίες καὶ ἀθεμελίωτα ἐπιχειρήματα.

Πολλοὶ πιστεύουν καὶ ἐλπίζουν ὅτι διὰ τῆς Φυσικῆς καὶ τῶν Μαθηματικῶν θὰ ἐπιτευχθεῖ θεμελιώδης πρόοδος στὸ νὰ γίνουν πιὸ κατανοητὲς οἱ ἔννοιες «πνεῦμα», «συνείδηση» κ.ἄ. Ἡ Συνείδηση ἀποτελεῖ ἓνα τόσο σπουδαῖο φαινόμενο ὥστε δὲν μπορεῖ νὰ πιστέψει κανεὶς ὅτι εἶναι ἀπλῶς κάτι πὸν «τυχαία» καὶ ὡς διὰ μαγείας παρουσιάζεται νὰ εἶναι τὸ ἀποτέλεσμα ἑνὸς πολυπλόκου ὑπολογισμοῦ.

Ἡ Συνείδηση εἶναι ἓνα φαινόμενο διὰ μέσου τοῦ ὁποῖου ἡ ὑπαρξη αὐτοῦ τούτου τοῦ Σύμπαντος γίνεται γνωστή. Μπορεῖ μάλιστα κανεὶς νὰ ἰσχυρισθεῖ ὅτι ἓνα Σύμπαν πὸν διέπεται ἀπὸ νόμους οἱ ὁποῖοι δὲν ἐπιτρέπουν τὴν ὑπαρξη Συνειδήσεως, δὲν εἶναι κὰν Σύμπαν.

Ἡ Συνείδηση εἶναι τὸ μόνο φαινόμενο διὰ τοῦ ὁποῖου ἓνα «θεωρητικὰ» ὑπαρκτὸ Σύμπαν μπορεῖ νὰ μετατραπεῖ σὲ πραγματικό.

Ὅμως ἄς προσγειωθοῦμε.

Τόσο ὁ Penrose ὅσο καὶ οἱ ὑπόλοιποι ἐρευνητὲς πὸν ἀναφέραμε, θίγουν θέματα τὰ ὁποῖα εἶναι βέβαιο ὅτι θὰ ἀπασχολήσουν τόσο τοὺς φιλοσόφους ὅσο καὶ τοὺς ὑπολοίπους ἐπιστήμονες γιὰ πολλὰ ἀκόμα χρόνια. Ἐκ παραλλήλου εἶναι ἐπίσης βέβαιο ὅτι ὁ ἄνθρωπος θὰ ἐξακολουθήσει νὰ κατασκευάζει μηχανικὰ συστήματα ὁλοένα πιὸ βελτιωμένα μὲ σκοπὸ τὴν πρακτικὴ ἐξυπηρέτησή του.

Ἀναφορικὰ μὲ τὴν ITN θὰ ἤθελα νὰ κλείσω τὴν ὁμιλία αὐτὴ μὲ τὶς ἀκόλουθες σκέψεις:

Ἐμεῖς οἱ ἄνθρωποι θέτομε τοὺς ἑαυτοὺς μας στὴν κορυφὴ τῆς παρούσας ἐξελικτικῆς κλίμακας. Εἴμαστε ἀρκετὰ ἀλαζόνες ὥστε νὰ πιστεῦομε ὅτι ἔχομε τὴ δύναμη νὰ δημιουργήσομε ὄντα προηγμένα ὅσο καὶ ἐμεῖς, ἐνῶ συγχρόνως μᾶς φοβίζει ἡ σκέψη ὅτι αὐτὸ πὸν προσπαθοῦμε νὰ δημιουργήσομε μπορεῖ νὰ γίνῃ πραγματικότητα!

Εὐτυχῶς ὑπάρχει, ὅπως ἀνέφερα παραπάνω, καὶ ἡ ἄλλη ὄψη τοῦ νομίσματος, ὅτι δηλαδὴ εἴμαστε κάτι παραπάνω ἀπὸ μηχανές.

Ἐχοντας μπροστά μας τὴ δεύτερη αὐτὴ, ἀνθρώπινη, ὄψη τοῦ νομίσματος, ἄς μὴ ξεχνᾶμε ΕΚΕΙΝΟΝ πὸν μᾶς ἔδωσε τὴ δύναμη νὰ τὴν βλέπομε.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Ν. Κ. Ἀρτεμιάδης, «Ἡ Ἀρμονία στὴ Φύση: Ὁ ρόλος τῶν Μαθηματικῶν στὴν κατανόησή της», Πρακτικὰ Ἀκαδημίας Ἀθηνῶν, Τομ. 62, 1987.
2. Ν. Κ. Ἀρτεμιάδης, «Ἡ Γεωμετρία τῶν Fractals», Πρακτικὰ Ἀκαδημίας Ἀθηνῶν, Τομ. 63, 1988.
3. Ν. Κ. Ἀρτεμιάδης, «Κορυφαῖες στιγμὲς στὰ Μαθηματικά», Πρακτικὰ Ἀκαδημίας Ἀθηνῶν, Τομ. 63, 1988.
4. P. J. D e n n i n g, «Blindness in the design of intelligent systems», Amer. Sc. 76, 1988.
5. J. R. S e a r l e, «Is the brain's mind a computer program?» Sci. Am. 262, 1990, p. 20-25.
6. P. M and P. R. C h u r c h l a n d : «Could a machine think?», Sci. Am. 262, 1990, p. 26-31.
7. R. P e n r o s e, «The Emperor's New Mind», Oxford University Press, 1990.