

# ET CETERA

*Τεχνητή νοημοσύνη vs Ανθρώπινη σοφία:*

*Ποιά θά επικρατήσει;*

(μιά συνομιλία μέ τόν καθηγητή Πληροφορικής καί ειδικό τής Τεχνητής Νοημοσύνης Θεμιστοκλή Παναγιωτόπουλο γιά τό παρόν καί τό μέλλον τών Μηχανών καί τήν επίδρασή τους στήν ποίηση)

ΣΥΝΕΝΤΕΥΞΗ-ΚΕΙΜΕΝΟ: ANNA ΓΡΙΒΑ

**Κ**υριακή πρωί. Ανηφορίζω πρός τό Παγκράτι, μέ προορισμό ένα καφέ πίσω από τόν λόφο του Άρδηττου. Πάντα μέ συγκινοῦσε αὐτή ἡ περιοχή. Ἴσως γιὰτί στόν νοῦ μου ὁ Ἄρδηττος εἶναι συνδεδεμένος μέ τά Μικρά Ἐλευσίνια Μυστήρια τῆς ἀρχαιότητος, ἀλλά καί μέ τόν Σεφέρη, ἀφοῦ κάπου ἐκεῖ, στά πλαϊνά τοῦ λόφου, βρίσκεται τό σπίτι του. Αὐτή τῆ φορά ὅμως δέν ἐπισκέπτομαι τῆ συγκεκριμένη πλευρά τῆς πόλης οὔτε γιά νά νιώσω τήν αὔρα τῶν ἀρχαίων τελετῶν οὔτε γιά νά θυμηθῶ ἕναν ἀπό τούς πιό ἀγαπημένους μου ποιητές. Ἔρχομαι γιά ἕναν λόγο πού σχετίζεται πολύ περισσότερο μέ τό παρόν καί τό μέλλον παρά μέ τό παρελθόν. Ὅδεῶ στό ραντεβού πού ἔδωσα πρό ἡμερῶν μέ ἕναν ἀπό τούς πιό διακεκριμένους διεθνῶς Ἑλληνες ἐπιστήμονες, ὁ ὁποῖος ἐξειδικεύεται στήν Τεχνητή Νοημοσύνη καί ἔχει μεγάλο ἐρευνητικό ἔργο στόν τομέα αὐτό, τόν Θεμιστοκλή Παναγιωτόπουλο, καθηγητή Πληροφορικής στό Πανεπιστήμιο Πειραιῶς. Ἀνυπομονῶ νά συζητήσω μαζί του γιά... ποίηση. Παράδοξο; Καθόλου! Ἡ ἀνάπτυξη τοῦ ChatGPT ἔχει ἀπό καιρό «θορυβήσει» τούς ποιητές. Ἄλλοτε ἀστελευόμαστε γι' αὐτό, ἄλλοτε πειραματιζόμαστε μέ τά ἔργα πού μπορεῖ νά συνθέσει ἡ μηχανή, κι ἄλλοτε συζητᾶμε μεταξύ μας γιά ὅσα διαβάζουμε, προσπαθώντας νά καταλάβουμε τί συμβαίνει τώρα καί τί θά συμβεῖ ἀργότερα. Πάντοτε ὅμως νιώθω πῶς, ἐπειδή δέν κατέχουμε τήν ἐπιστημονική γνώση, ὑπάρχουν πράγματα πού μᾶς διαφεύγουν, πού τά ἀντιμετωπίζουμε διαισθητικά ἢ μέ ἐρμηνεῖες οἱ ὁποῖες ἀγγίζουν περισσότερο τήν ἐπιστημονική φαντασία παρά τήν πραγματικότητα... Αὐτός ἦταν καί ὁ λόγος γιά τόν ὁποῖο, μιά μέρα πού συζητούσαμε μέ τήν ομάδα τῆς Ποιητικῆς σχετικά μέ τήν Τεχνητή Νοημοσύνη, σκέφτηκα πῶς θά ἦταν ὠραία ἰδέα νά πάρουμε μιά συνέντευξη ἀπό τόν πλέον εἰδικό, καί μάλιστα ἕναν εἰδικό πού ὄχι μόνο ἔχει τή γνώση, ἀλλά χαρακτηρίζεται συγχρόνως ἀπό εὐρύτητα παιδείας, ἡ ὁποία τοῦ ἐπιτρέπει νά βλέπει τήν ἐπιστήμη καί τίς προόδους τῆς καί μέσα ἀπό σκοπιά φιλοσοφική.

Ευελπιστώ πώς ή συζήτηση πού καταγράφω παρακάτω θά σᾶς φανεί τόσο χρήσιμη, διαφωτιστική καί ενδιαφέρουσα ὅσο φάνηκε καί σ' ἐμένα...

Ἀγαπητέ Θέμη, κατ' ἀρχάς θά ἤθελα νά σέ εὐχαριστήσω πού δέχτηκες τήν πρόσκληση νά παραχωρήσεις στήν Ποιητική αὐτή τή συνέντευξη. Θά ἤθελα νά ἀρχίσουμε ἀπό τόν βασικό προβληματισμό πού ἐγείρεται τόν τελευταῖο χρόνο στόν ποιητικό, ἀλλά καί στόν εὐρύτερο καλλιτεχνικό, χῶρο σέ σχέση μέ τίς δυνατότητες τῆς Τεχνητῆς Νοημοσύνης (καί κυρίως τοῦ ChatGPT): Μπορεῖ ἡ μηχανή νά υπερκεράσει τήν πρωτότυπη ἀνθρώπινη ποιητική ἐμπνευση; Μπορεῖ δηλαδή, πέρα ἀπό τήν ἀνασύνθεση καί τή μίμηση τῶν ὑπαρκτῶν ποιητικῶν «μοντέλων», πού τήν ἔχουμε δεῖ ὡς ἐπί τό πλεῖστον νά ἐφαρμόζει, νά ἀναπτύξει νέα εἶδη ποίησης, νέες καλλιτεχνικές πρωτοπορίες, νέους δημιουργικούς δρόμους; Καταλαβαίνεις πώς μέσα ἀπό αὐτό τό ἐρώτημα θά ἤθελα νά μᾶς μιλήσεις εὐρύτερα γιά τόν τρόπο μέ τόν ὁποῖο «σκέπτεται» καί λειτουργεῖ ἡ μηχανή σέ σχέση μέ τήν καλλιτεχνική δημιουργία.

Θά ἤθελα κι ἐγώ ἀπό τή μεριά μου νά σᾶς εὐχαριστήσω γιά τήν εὐκαιρία πού μοῦ δίνετε γι' αὐτή τήν πολύ ενδιαφέρουσα συνέντευξη. Γιά νά ἀπαντήσουμε στά ἐρωτήματα περί τῆς σχέσης μεταξύ τεχνητῆς νοημοσύνης καί τέχνης, θά ἦταν ἀπαραίτητο νά σκεφτοῦμε: Τί εἶναι ἄραγε ἡ καλλιτεχνική δημιουργία; Νομίζω πώς ἡ γνήσια καλλιτεχνική δημιουργία, στήν ὑψηλότερη μορφή της, εἶναι μιά ἐσωτερική ἐμπειρία. Τό προκύπτον βίωμα εἶναι τόσο ἰσχυρό ὥστε συνεπαίρνει τόν καλλιτέχνη, ἐπιζητώντας νά ἐκδηλωθεῖ σέ μιά «ἀπτή» δημιουργία. Ὅλο αὐτό τό βίωμα παραπέμπει στήν κυρίαρχη παρουσία τῆς ψυχῆς του, πού πραγματικά πάλλεται ἀπό ἀγάπη, συγκίνηση καί ἐμπνευση γιά τό ἀντικείμενο τῆς καλλιτεχνικῆς του ἔκφρασης καθώς πραγματοποιεῖ τή δημιουργία. Θά λέγαμε πώς πρόκειται γιά μιά ἰδιαίτερου τύπου ἐσωτερική «ἐρωτική σχέση». Καί βέβαια, ἡ νόηση βοηθᾶ στή σύλληψη, πάντα σέ συνεργασία μέ αὐτή τήν ἐσωτερική δραστηριότητα πού περιέγραφα προηγουμένως.

Νά συμπληρώσω ἐδῶ ὅτι μέ τήν παραπάνω περιγραφή ἐννοῶ πώς καλλιτεχνική δημιουργία μπορεῖ νά πραγματοποιήσῃ ὄχι μόνο αὐτός πού ὀρίζουμε ὡς καλλιτέχνη, ἀλλά καί ἕνας οἰκοδόμος, ἕνας ἐπιπλοποιός, ἕνας φιλόσοφος, ἕνας δάσκαλος – καί καταλαβαίνετε ὅτι αὐτό ἀφορᾶ κάθε ἐξωτερίκευση τοῦ ἐαυτοῦ. Καί βέβαια, ἡ καλλιτεχνική δημιουργία ἔχει τίς διαβαθμίσεις της. Στήν πιό «χαμηλή» της μορφή, ἡ παραγόμενη σύνθεση θά ἐμοιαζε περισσότερο σάν μιά σύνδεση ἐξαρτημάτων κατώτερης ποιότητας, τά ὁποῖα σχηματίζουν μιά μέτρια, μιά λίγο πολύ κοινότοπη σύνθεση, μιά καρικατούρα ἢ μιά κακοτεχνία.

Τώρα ἄς μιλήσουμε καί γιά τίς δυνατότητες τῆς Τεχνητῆς Νοημοσύνης γενικότερα, καί εἰδικότερα γιά τό ChatGPT. Ἡ Τεχνητή Νοημοσύνη (TN, Artificial Intelligence, AI) εἶναι ἕνα γνωστικό πεδίο, ὑποπεδίο τῆς ἐπιστήμης τῆς

Πληροφορικής, στο όποιο όμως οι έρευνητές πολλές φορές, εισάγοντας ιδέες από άλλους επιστημονικούς χώρους και πεδία, όπως ή γνωστική ψυχολογία, ή φιλοσοφία, τά μαθηματικά, ή βιολογία, ή γλωσσολογία, αναπτύσσουν κάποιες θεωρίες, μοντέλα και συστήματα με σκοπό τή διαμόρφωση ολοκληρωμένων συστημάτων τά όποία μιμούνται, κατά κάποιο τρόπο, τή ανθρώπινη νοημοσύνη και συμπεριφορά.

Στήν ΤΝ όμιλοϋμε για τήν αναπαράσταση και τήν έπεξεργασία τής γνώσης, όπου ως γνώση προσδιορίζουμε κάτι ευρύτερο από τήν όργάνωση και τήν έπεξεργασία τής πληροφορίας, πού άφορά τήν έπιστήμη τής Πληροφορικής. Όρίζουμε τήν αναπαράσταση του κόσμου μέσω τών ιδιοτήτων διαφόρων αντικειμένων και τών δυναμικά εξέλισσόμενων σχέσεων τους. Πρόκειται για αναπαράσταση άφηρημένων έννοιων και πραγματικοτήτων, όπως ο χώρος, ο χρόνος, οι νόμοι τής φυσικής πραγματικότητας ή τής πραγματικότητας τών ανθρώπων και τών κοινωνιών τους κ.λπ.

Έπίσης, στήν ΤΝ μιλοϋμε για δύο μεγάλες ύποπεριοχές, τή συμβατική και τή μή συμβατική ΤΝ. Στόν χώρο τής συμβατικής ΤΝ, έχουμε τό πεδίο του Λογικού Προγραμματισμού μέσω τής Μαθηματικής Λογικής (Προτασιακή και Κατηγορηματική Λογική) και τής Άσαφούς Λογικής μέσω τής Θεωρίας Συνόλων (Θεωρία Συνόλων + Λογική + Γλωσσολογία), τής Έπεξεργασίας Φυσικής Γλώσσας και τής Άναγνώρισης και Παραγωγής Γλώσσας και Φωνής (μέ στοιχεία από τή γλωσσολογία), τής παραγωγής και έκτέλεσης σχεδίων, πλάνων πού εξέλισσονται μέσα στόν χρόνο, τών Σημαιολογικών Δικτύων και τών Πλαισίων/Σεναρίων (μέ ιδέες από τή γνωστική ψυχολογία). Από τόν χώρο τής μή συμβατικής ΤΝ, θά μπορούσαμε νά αναφερθοϋμε στά Τεχνητά Νευρωνικά Δίκτυα, πού χρησιμοποιούν ιδέες από τή βιολογία, τή δομή και τή λειτουργία βιολογικών νευρικών κυττάρων, στους Γενετικούς Άλγορίθμους, πού χρησιμοποιούν ιδέες από τή γενετική (δομή DNA), και στους Έξελικτικούς Άλγορίθμους, πού χρησιμοποιούν ιδέες από θεωρίες εξέλιξης κ.λπ. Βέβαια, όλα αυτά είναι ένα μικρό μέρος του καταλόγου, αλλά δέν θά έπεκταθώ περαιτέρω.

Η περίπτωση τών προγραμμάτων συνομιλίας (Chat Bots), όπως τό ChatGPT, έντάσσεται στόν χώρο τής Έπεξεργασίας, Κατανόησης και Παραγωγής Φυσικής Γλώσσας. Ειδικά για τό ChatGPT, θά πρέπει νά αναφέρουμε τή χρήση τεχνικών Μηχανικής Μάθησης (Τεχνητά Νευρωνικά Δίκτυα), μέσω έπιβλεπόμενης ή μή έπιβλεπόμενης μάθησης. Τό ChatGPT έχει χρησιμοποιήσει έναν άπίστευτα μεγάλο όγκο βιβλίων και επιστημονικών άρθρων, καθώς και άλλες πηγές, τίς όποιες έχει άποθηκεύσει, όργανώσει και προεπεξεργαστεί, ώστε νά μπορεί νά τίς διαχειριστεί sé πραγματικό χρόνο. Έπίσης, έχει εκπαιδευτεί από συζητητές (άνθρώπους) ώστε νά μπορεί νά πραγματοποιήσει διάλογο εύγενικό, άποδεκτό, κατανοητό και φιλικό. Έτσι, ο διάλογος μέ τό ChatGPT είναι πολύ ευχάριστος, και κατά τή διάρκειά του μάς δίνει τήν ψευδαίσθηση ότι κατανοεί τά έρωτήματά μας και άπαντά σ' έμάς τίς διάφορες έρωτήσεις. Μπο-

ρεί επίσης να του ζητηθεί να «δημιουργήσει» ποιήματα, και τότε κάνει εμφανίζοντας ένα έντυπωσιακό (για μηχανή) αποτέλεσμα.

Όμως, ως καταλάβουμε καλά ότι το ChatGPT, αν αφαιρέσει κανείς τα δεδομένα κειμένων, εικόνων κ.λπ. με τα οποία το έχουν τροφοδοτήσει, δεν διαθέτει εμπειρία ζωής· δεν διαθέτει γενικώς εμπειρίες. Άς μή μας έντυπωσιάξει λοιπόν το γεγονός ότι μπορεί να γράφει ακόμη και λογισμικό, να αναπαράγει εικόνες, να συνθέσει εικόνες με κείμενο ποιητικής ή μη ποιητικής μορφής.

Η «δημιουργία» του ChatGPT δεν είναι επιστημονική ή καλλιτεχνική δημιουργία. Είναι μία παρουσίαση στοιχείων τα οποία έχει «άποστηθείσει» με τρόπο φιλικό και κατανοητό για τον αναγνώστη.

Για να δώσω μία αναλογία, είναι σαν να έχουμε έναν πολυμαθή ο οποίος μπορεί να μας μιλήσει με μεγάλη λεπτομέρεια για την κολύμβηση, τα στοιχεία της, τις τεχνικές κ.λπ., αλλά δεν έχει ποτέ του κολυμπήσει ο ίδιος και δεν μπορεί καν να συλλάβει τα παιχνίδια στο νερό, την προσωπική εμπειρία όταν το νερό κυλά πάνω στο πρόσωπό μας, όταν κολυμπάμε ή κάνουμε ένα μακροβούτι, τούς ήχους, τα κύματα, το βίωμα της απόλυτα γαλήνιας θάλασσας, το πέταγμα των γλαρών, τη μαγική εμπειρία της δύσης του ήλιου στο βάθος του ορίζοντα, τα χρώματα της θάλασσας που αλλάζουν καθώς περνάει η ώρα, όλη αυτή την υπέροχη πολυεμπειρία που σχετίζεται με τη θάλασσα, ή οποία έχει εμπνεύσει τόσους και τόσους καλλιτέχνες.

Σύμφωνα με αυτά, θα μπορούσα να υποστηρίξω επομένως με ασφάλεια ότι η μηχανή δεν μπορεί να υπερκεράσει την πρωτότυπη ανθρώπινη ποιητική εμπνευση. Δεν μπορεί, πέρα από την ανασύνθεση και τη μίμηση των υπαρκτών ποιητικών «μοντέλων», που έχουμε δει ως επί το πλείστον να κάνει, να αναπτύξει νέα είδη ποίησης, νέες καλλιτεχνικές πρωτοπορίες, νέους δημιουργικούς δρόμους.

*Αν θεωρήσουμε πώς ο άνθρωπος σε μεγάλο βαθμό, όπως ορίζει και το εμπειριστικό μοντέλο του John Locke, δημιουργεί ιδέες και θεωρίες ξεκινώντας από τα αισθητηριακά ερεθίσματα του περιβάλλοντος, θα μπορούσε με ανάλογο τρόπο και η μηχανή να μεταβεί από τα ερεθίσματα-δεδομένα που εισάγει σε αυτήν ο άνθρωπος σε δικές της ιδέες και θεωρίες;*

Ο Τζόν Λόκ, νομίζω, επιθυμεί, μεταξύ άλλων, να θεμελιώσει μία εμπειριστική γνωσιοθεωρία, καθώς μάλλον δεν αποδέχεται τη δυνατότητα της ένορατικής αντίληψης για τη γνώση. Βέβαια, θα έλεγα ότι πράγματι, σε έναν συντριπτικό βαθμό, η γνώση που διαθέτει ένας άνθρωπος για τον κόσμο προέρχεται από το σύνολο των εμπειριών του και τη συνακόλουθη συλλογιστική, που τον οδηγεί σε νέα συμπεράσματα. Όμως το ερώτημα παραμένει: Μία εμπειρία που βιώνεται μέσω των αισθήσεων είναι μία αποκλειστικά αισθητική-ύλική

έμπειρία; Ή είναι άραγε μία πολυδιάστατη έμπειρία, μέρος τής όποίας άποτελεϊ ή έμπειρία του αισθητού;

Ή μηχανή δέν θά διαθέτει τέτοιου τύπου πολυδιάστατη έμπειρία, άκόμη και άν κατορθώσουμε, μέσω κάποιων ρομπότ για παράδειγμα, νά αναπτύξουμε συστήματα τά όποία θά διαθέτουν, έστω, όλες τίς ύλικές Ικανότητες τών αισθήσεων (όραση, άκοή, γεύση, όσφρηση, άφή). Θά μπορούσε λοιπόν ίσως μία μελλοντική μηχανή νά έχει τήν Ικανότητα νά αναπτύσσει θεωρίες για τή φυσική πραγματικότητα, άποστραγγισμένη όμως από τήν ευρύτερη έμπειρία πού άποκτᾶ ένας άνθρωπος όταν έρχεται σέ έπαφή μαζί της.

*Οί καλλιτεχνικές συλλήψεις, όπως και ή νόηση του ανθρώπου εν γενει, βασίζονται συχνά σέ μία αισθητηριακή αντίληψη του κόσμου, ενώ ή μηχανή μιλά για τόν κόσμο βασιζόμενη σέ δεδομένα πού τροφοδοτούνται από τούς ανθρώπους προγραμματιστές της και έπομένως περιορίζονται και κατευθύνονται από τή συγκεκριμένη όπτική συγκεκριμένων ανθρώπων. Μίλησέ μας γι' αυτή τή διαφορά, πώς μπορεί νά έπηρεάζει τό εύρος τών δυνατοτήτων τής μηχανής και άν στό μέλλον ύπάρχει ή πιθανότητα άμβλυνσης ή έξάλειψής της.*

Πράγματι, ή όπτική του προγραμματιστή μέ κάποιον τρόπο άποτυπώνεται στό πρόγραμμα. Από τήν άλλη όμως, στήν ΤΝ προσπαθούμε νά άποκρυσταλλώσουμε σέ κώδικα λογισμικού τήν αναπαράσταση του κόσμου όπως φανταζόμαστε ότι τή διαχειρίζεται ό μέσος ανθρώπινος έγκέφαλος. Βέβαια, νομίζω, ούτε οί νευροεπιστήμες δέν έχουν άπαντήσει άκόμη πειστικά και ολοκληρωμένα στό πώς ακριβώς άποθηκεύει, διαχειρίζεται και έπεξεργάζεται έμπειρίες, γεγονότα κ.λπ. ό ανθρώπινος έγκέφαλος. Είμαστε άκόμη σέ έμβρυικό στάδιο σέ αυτό τόν τομέα.

Θεωρώ ότι όποιαδήποτε μηχανή κι άν φτιάξουμε, όσο έξελιγμένη κι άν είναι, όσο γρήγορη στους συλλογισμούς και στήν έπεξεργασία, όσο μεγάλη δυνατότητα άποθήκευσης δεδομένων κι άν έχει —και όλα αυτά έξαιρετικά υπέρτερα του ανθρώπινου έγκεφάλου—, άν και θά μπορεί νά πραγματοποιοει κάποιες λειτουργίες και δραστηριότητες πολύ, μά πολύ, καλύτερα από τόν πλέον ευφυή άνθρωπο, πάντα θά τής λείπει αυτό τό ιδιαίτερο, αυτό τό έξωυλικό —άλλά όχι στήν πραγματικότητα λιγότερο άπτό για όσους τό έρευνούν— στοιχείο, ή ψυχή και τό πνεύμα, και βέβαια κάτι άκόμη βαθύτερο, πού κάποιος τό άποκαλοϋν «Έγώ».

*Οί άνθρωποι χαρακτηρίζονται από έτερότητα και ποικιλία, γεγονός πού συμβάλλει στή μοναδικότητα τής έκφρασής τους, ή όποία δέν είναι άπόλυτα προβλέψιμη. Αντιθέτως, οί μηχανές πού γνωρίζουμε μέχρι τώρα χαρακτηρίζονται από μία όμοιόμορφη, άπόλυτα ντετερμινιστική άνταπόκριση, ανά-*

λογα με τα δεδομένα που έχουν και το πρόβλημα που τους ζητείται να λύσουν. Η τέχνη όμως βασίζεται στη μοναδικότητα, την έμπνευση και το απρόβλεπτο. Άραγε στο μέλλον θα υπάρχουν μηχανές με τέτοιες δυνατότητες και χαρακτηριστικά; Μήπως διαγράφεται ένα τέτοιο μέλλον όταν γίνει εφαρμογή και έκτεταμένη χρήση των κβαντικών υπολογιστών;

Όμολογουμένως, όχι μόνο οι άνθρωποι, αλλά ολόκληρη η Δημιουργία (τό σύμπαν στο σύνολό του) και η φύση (ή φυσική δημιουργία στον πλανήτη μας) χαρακτηρίζονται από αυτό το στοιχείο της ποικιλίας – τόσο μεγάλης που δεν μπορεί κανείς να βρει δύο πανομοιότυπα όντα ή σώματα, ή ακόμη και φαινόμενα, γύρω μας. Έτσι, βλέπουμε μία μοναδικότητα τόσο στη φύση όσο και στις εκδηλώσεις της, παρά τα παρόμοια στοιχειώδη χαρακτηριστικά που μοιράζονται δύο οντότητες. Φαίνεται, ή εύλογα υποθέτουμε, ότι οι φυσικοί νόμοι είναι πανομοιότυποι σε όλη την έκταση του σύμπαντος και σε όλη τη διάρκεια του χρόνου.

Όμως, αν και ο μη ντετερμινισμός είναι άγκιστρωμένος στα βάθη της ουσίας της ύλοενέργειας, αρχίζει να συγκλίνει προς έναν ντετερμινισμό όσο εξετάζουμε όλο και πιο σύνθετες δομές. Παρ' όλα αυτά, και ενώ οι δομές γίνονται, τουλάχιστον σε επίπεδο βασικής οργάνωσης (για παράδειγμα, το ανθρώπινο σώμα με τα συστήματα οργάνων ή τα πλανητικά συστήματα), σε μεγάλο βαθμό ντετερμινιστικές, οι εκδηλώσεις περιλαμβάνουν προβλέψιμες αλλά και απρόβλεπτες πλευρές, που διαφοροποιούνται όχι μόνο στον άνθρωπο, αλλά και σε κάπως εξελιγμένα όντα, όπως τα ζώα. Αυτό δείχνει ότι πρέπει να υπάρχει ένας ιδιαίτερος παράγοντας, τον οποίο ίσως θα μπορούσαμε να αποκαλέσουμε συνειδηση και ιδιαιτερότητα στην προσωπικότητα και τον χαρακτήρα στα εξελιγμένα όντα.

Αυτή η μοναδικότητα στην αντίληψη, την εκδήλωση και την έκφραση έχει, όπως φαίνεται, μεγάλη συγγένεια με τη λεγόμενη «καλλιτεχνική δημιουργία». Όπως ανέφερα και προηγουμένως, ακόμη και αν αναπτυχθούν υπερεξελιγμένοι κβαντικοί υπολογιστές, που θα λειτουργούν σε υπερταχύτητες και πιθανώς θα έχουν και τη δυνατότητα μη ντετερμινιστικής «συμπεριφοράς», το περιεχόμενο της αντίληψής τους θα είναι ψευδοαισθητηριακό, με την έννοια ότι οι υπολογιστές πάντα θα υπολείπονται όσον αφορά τις διαστάσεις της έμπειρίας, ακόμα και αν διαθέτουν εξαιρετικά εξελιγμένο τρόπο έκφρασης. Θα διαφέρουν δηλαδή στην ποιότητα του είναι. Για την ακρίβεια, όχι μόνο θα υστερούν, αλλά δεν θα έχουν καθόλου μερίδιο στο ουσιαστικό είναι, παρόλο που μπορεί να υπερτερούν στην ποιότητα του φαίνεσθαι.

Πολλές φορές οι καλλιτέχνες λένε πώς μέσω της τέχνης τους καλύπτουν έσωτερικές ανάγκες και κίνητρα. Οι μηχανές όμως, όπως τις ξέρουμε μέχρι σήμερα, δεν έχουν κάποιο κίνητρο· εκτελούν εντολές. Ποιά είναι ή επίπτωση

στό άποτέλεσμα τής τέχνης πού παράγει ή μηχανή λόγω τής άνυπαρξίας τέτοιων έσωτερικών διεργασιών έντός τής;

Στήν ΤΝ προσπαθοΰμε νά μοντελοποιήσουμε άκόμη και τίς έννοιες τών άναγκών, τών κινήτρων, τών συναισθημάτων. Μπορούμε δηλαδή νά φτιάξουμε συστήματα πού θά λειτουργοΰν μιμούμενα έξελιγμένα όντα μέ τέτοια χαρακτηριστικά. Τά ίδια όμως τά προγράμματα δέν έχουν άνάγκες, κίνητρα και συναισθήματα.

Όποτε, και πάλι, ή όποια «καλλιτεχνική δημιουργία» τών μηχανών, ένώ μπορεί νά παράγει έντυπωσιακά άποτελέσματα, θά στερείται οΰσιαστικού περιεχομένου. Κατά τή γνώμη μου, τό άποτέλεσμα τών μηχανών θά είναι έντυπωσιακό, και πιθανώς χρήσιμο, αλλά κατά πόσο άραγε θά περιέχει γνήσια καινοτομία και έμπνευση όταν λείπει ή οΰσια;

Η μηχανή, προκειμένου νά όδηγηθεί στους στόχους τής, έξάγει άποτελέσματα άκολουθώντας μιά σειρά άλγοριθμικών βημάτων, πού συνιστοΰν τήν «άποδεικτική» διαδικασία τής. Τό θεώρημα τής μή πληρότητας τοΰ Gödel όμως έδειξε ότι δέν μπορούν νά άποδειχθοΰν όλες οι άληθείς προτάσεις, δηλαδή δέν είναι όλα τά πράγματα άποδείξιμα. Αυτό προφανώς θέτει έναν *de facto* περιορισμό στή λειτουργία τής μηχανής, στον όποιο δέν υπόκειται κατ' άνάγκη ό άνθρωπος, καθώς μέσω τής φαντασίας του μπορεί νά διατυπώνει άληθείες χωρίς νά έχει υποχρεωτικά προηγηθεί κάποια άποδεικτική διαδικασία. Αυτό τό στοιχείο τής φαντασίας δέν διαφοροποιεί ριζικά τόν άνθρωπο από τή μηχανή όσον άφορᾷ τήν καλλιτεχνική δημιουργία;

Μοΰ έχετε θέσει έρωτήματα πού κανονικά θά χρειάζονταν βιβλία όλόκληρα για νά τά άπαντήσω πειστικά και όλοκληρωμένα. Έν πάση περιπτώσει, πράγματι ό διάσημος μαθηματικός Κούρτ Γκαϊντελ απέδειξε ότι υπάρχουν μαθηματικές προτάσεις οι όποιες, ένώ είναι άληθείς όποτε και για όποια δεδομένα τίς έξετάζουμε, δέν μπορούν νά άποδειχτοΰν - ότι δέν υπάρχει συλλογιστική διαδικασία ή όποια νά μπορεί νά όδηγηθεί στήν άπόδειξη τών προτάσεων αυτών. Δηλαδή, θά λέγαμε πώς ή άλήθεια είναι ύπερσύνολο τής άπόδειξης και τής συλλογιστικής πού όδηγεί σέ αυτήν. Αυτό είναι πραγματικά έντυπωσιακό και σοκαριστικό για εκείνους πού πιστεύουν ότι μέ τή συλλογιστική άπόδειξη θά φτάσουμε σέ όλες τίς άλήθειες.

Υπήρξε ένας γνωστός μαθηματικός, ό Σρινιβάσα Ραμανούτζαν, πού έζησε από τό 1887 ως τό 1920, αυτόδίδακτος, μέ σχεδόν καθόλου εκπαίδευση στά καθαρά μαθηματικά, ό όποιος είχε έντυπωσιακή συνεισφορά σέ διάφορα πεδία τής έπιστήμης τών μαθηματικών, όπως ή μαθηματική άνάλυση, ή θεωρία αριθμών, οι άπειροστικές σειρές, τά συνεχή κλάσματα. Όταν τόν κάλεσαν στό Πανεπιστήμιο τοΰ Κέμπριτζ και τούς παρουσίασε τά έξαιρετικά πολύ-

πλοκα μαθηματικά του, τοῦ ζήτησαν τίς ἀποδείξεις γιά νά μπορέσουν νά τά δημοσιεύσουν. Ὁ ἴδιος δέν πολυκαταλάβαινε τί θά πεί ἀπόδειξη καί ὑποστήριξε ὅτι ὁ θεός τοῦ μιλοῦσε στόν ὕπνο του! Τίς προτάσεις του καί τά θεωρήματά του τά ἐπεξεργάζονταν καί προσπαθοῦσαν νά τά ἀποδείξουν ἐπιφανεῖς μαθηματικοί γιά τά ἐπόμενα ἑκατό χρόνια. Καταλαβαίνετε μέ αὐτό τό παράδειγμα τούς περιορισμούς τῶν μηχανῶν καί τήν ὑπεροχή τοῦ πνεύματος.

*Μέ ποιόν τρόπο μποροῦμε νά εἴμαστε ἀσφαλεῖς ἔναντι τῶν μηχανῶν; Θεωρεῖς πώς θά μποροῦσαν κάποια στιγμή νά ἀναπτύξουν ἓνα δικό τους «ἔνστικτο αὐτοσυντήρησης» καί νά θέσουν τόν ἑαυτό τους στό ἐπίκεντρο; Καί σέ μιά τέτοια περίπτωση, τό πρόβλημα θά μποροῦσε νά λυθεῖ «βγάζοντάς τες ἀπό τήν πρίζα», σταματώντας σέ μιά ὥρα ἀνάγκης ὄλο τόν μηχανισμό τοῦ Διαδικτύου, ἢ κάτι τέτοιο εἶναι πλέον ἀδύνατον;*

Σέ καμία περίπτωση δέν πρέπει νά αισθανόμαστε ἀσφαλεῖς ἀπέναντι στίς μηχανές καί στήν πιθανή Τεχνητή Ὑπερνοημοσύνη πού μποροῦν νά ἀναπτύξουν. Δέν θά μποροῦσαν νά ἀναπτύξουν ἀπό μόνες τους ἔνστικτο αὐτοσυντήρησης, γιατί δέν ἔχουν ἀντίληψη οὔτε τῆς γέννησης, οὔτε τῆς ζωῆς, οὔτε τοῦ θανάτου. Θά μποροῦσαν ὅμως νά προγραμματιστοῦν ὥστε νά διαθέτουν προσομοιωμένο ἔνστικτο αὐτοσυντήρησης. Ἄν φροντίσουμε ὡς προγραμματιστές νά ὑπάρχει μηχανισμός τερματισμοῦ, αὐτός θά ὑπάρχει. Ἄν προγραμματιστοῦν νά μὴν ἔχουν τέτοιο μηχανισμό, δέν θά τόν ἔχουν. Ὡστόσο, τό ἀκόμη μεγαλύτερο πρόβλημα μέ τή Μηχανική Μάθηση εἶναι ὅτι, ὅταν ἐκπαιδεύεται ἓνα Τεχνητό Νευρωνικό Δίκτυο, δέν γνωρίζουμε τί γνωρίζει καί τί δέν γνωρίζει.

*Ποιά εἶναι ἡ δική σου πρόβλεψη γιά τό βραχυπρόθεσμο καί τό μακροπρόθεσμο μέλλον τῆς Τεχνητῆς Νοημοσύνης; Τί θά ἔλεγεσ εἰδικότερα στούς καλλιτέχνες τοῦ λόγου; Μήπως θά πάψει τελικά νά εἶναι ἀναγκαῖος ὁ ρόλος τους;*

Τό μέλλον τῆς ΤΝ εἶναι ἀπρόβλεπτο. Ὑπάρχουν ἑκατομμύρια ἐρευνητές πού ἀσχολοῦνται μέ αὐτό τό θέμα. Ἀπρόβλεπτη γιά ἐμένα εἶναι καί ἡ χρήση τῆς. Σίγουρα θά φέρεи μιά ἐπανάσταση. Τό θέμα εἶναι ἂν θά φέρεи καί μιά φυλακή γιά τό ἀνθρώπινο πνεῦμα. Ἡ γνήσια καλλιτεχνική δημιουργία δέν ἔχει τίποτα νά φοβηθεῖ· οἱ μετριότητες ὅμως θά δυσκολευτοῦν πολύ.

*Θέμη, ἐκ μέρους τῆς ομάδας τῆς Ποιητικῆς σέ εὐχαριστοῦμε θερμά γιά ὅσα μάθαμε ἀπό τή συνέντευξη αὐτή.*

Ἐγώ σᾶς εὐχαριστῶ.