

Συζητώντας με το ChatGPT για την Ενέργεια: Όταν η Σωκρατική Μέθοδος «ξεκλειδώνει» την Τεχνητή Νοημοσύνη

Παναγιώτης Κουμαράς

Εισαγωγή

Στα πάνω από 50 χρόνια της ζωής μου ως δάσκαλος άκουσα πολλά για τη Σωκρατική μέθοδο διδασκαλίας, μια μορφή διαλογικής διδασκαλίας που βασίζεται στην υποβολή συνεχών ερωτήσεων από τον δάσκαλο, με σκοπό να οδηγήσει τον μαθητή στην ανακάλυψη της γνώσης. Καλά μην βαράτε, ομολογώ ότι (εγώ τουλάχιστον) δεν είδα να επιτυγχάνεται ότι δηλωνόταν ως αποτέλεσμα της μεθόδου.

Κάποτε, ως καθηγητής στο Γυμνάσιο, δίδασκα ότι μια από τις βασικές ιδιότητες της ενέργειας είναι η αποθήκευση, και μετά από λίγο ανέφερα ως μορφές ενέργειας, μεταξύ άλλων, τη φωτεινή και την ηχητική. Μετά το μάθημα, έπιασα τον εαυτό μου να σκέφτεται τι θα απαντούσα αν κάποιος παρατηρητικός μαθητής με ρωτούσε: *«Καλά, κύριε, η φωτεινή και η ηχητική ενέργεια δεν αποθηκεύονται, πώς λες ότι είναι μορφές ενέργειας;»*. Ίσως, τελικά, αυτό να ήταν μια σωκρατική μέθοδος για να μάθει ο ίδιος ο δάσκαλος. Προσπάθησα αργότερα να χρησιμοποιήσω αυτό το παράδειγμα, και άλλα ανάλογα, σε σεμινάρια με συναδέλφους Φυσικούς. Δεν θα έλεγα ότι είχε επιτυχία. Υπήρχαν βέβαια και μάτια που έλαμψαν, αλλά οι περισσότεροι απλά προσπερνούσαν την αντίφαση.

Η επίμονη αναπαραγωγή δογματικών λαθών και απλουστεύσεων στα σχολικά εγχειρίδια με απασχόλησε για σειρά ετών, οδηγώντας μεταξύ άλλων σε αναρτήσεις στο ylikonet, στη συγγραφή του βιβλίου *«Προπατορικά Αμαρτήματα Βιβλίων Φυσικής»* (Εκδόσεις Ροπή, 2024) και σε σειρά δημοσιεύσεων σε περιοδικά Διδακτικής. Έχοντας αυτή την προβληματική κατά νου σκέφτηκα να θέσω ανάλογα ερωτήματα στην Τεχνητή Νοημοσύνη (Α.Ι.). Ήθελα να δω αν ένας σύγχρονος, ψηφιακός «παντογνώστης» θα έπεφτε θύμα των ίδιων «προπατορικών αμαρτημάτων» (δογματικών λαθών) που αναπαράγει το εκπαιδευτικό μας σύστημα ή θα μας «έβαζε τα γυαλιά;» Το συμπέρασμα αυτής της εμπειρίας; Η Τεχνητή Νοημοσύνη έγινε ο απόλυτος καθρέφτης αυτών των λαθών που επαναλαμβάνονται διαχρονικά. Αναπαρήγαγε τυφλά τον προβληματικό «μέσο όρο» των σχολικών βιβλίων. Αποδέχεται *«εγώ έγινα ο απόλυτος “ψηφιακός εκπρόσωπος” αυτών των προπατορικών αμαρτημάτων (!) που έχουν καταγραφεί στα σχολικά εγχειρίδια»*. Όμως, η συζήτηση έδειξε ότι αν μπορείς να θέσεις τις κατάλληλες ερωτήσεις, το Α.Ι. μπορεί να αλλάξει επίπεδο και να οδηγηθεί σε αυστηρές επιστημονικές παραδοχές.

Φάση 1: Η Παγίδα του «Σχολικού» Ορισμού

Η συζήτηση ξεκίνησε με την κλασική ερώτηση: *«Τι ονομάζεται ενέργεια στη Φυσική;»* Η απάντηση του Α.Ι. ήταν αυτή που αναπαράγουν τα περισσότερα σχολικά εγχειρίδια παγκοσμίως: *«Ενέργεια είναι η ικανότητα ενός σώματος ή συστήματος να παράγει έργο»*. Ο χρήστης θα μπορούσε να σταματήσει εδώ την αναζήτηση. Πήρε τη ζητούμενη γνώση.

Αποφάσισα να εφαρμόσω τη Σωκρατική μέθοδο, θέτοντας το λογικό παράδοξο: «Αφού η ενέργεια διατηρείται, και αφού ενέργεια είναι η ικανότητα παραγωγής έργου, σημαίνει ότι διατηρείται και η ικανότητα παραγωγής έργου;» Το Α.Ι. παραδέχτηκε αμέσως το σφάλμα του, αναφέροντας τον Δεύτερο Νόμο της Θερμοδυναμικής, την υποβάθμιση της ενέργειας και την εντροπία. Παραδέχτηκε ότι ο ορισμός της «ικανότητας παραγωγής έργου» είναι ελλιπής και ξεπερασμένος, ανέφερε και το μοντέλο του θερμικού θανάτου του σύμπαντος και τέλος εισήγαγε τη διάκριση μεταξύ Ενέργειας (για αυτό που διατηρείται) και της διαθέσιμης ενέργειας (Εξέργειας), που συνεχώς μειώνεται στο σύμπαν.

Φάση 2: Η Ψευδαίσθηση των «Ιδιοτήτων ενός Σώματος»

Το δεύτερο «κρασάρισμα» του αλγορίθμου ήρθε όταν η συζήτηση έφτασε στο αν η ενέργεια είναι ιδιότητα ενός μεμονωμένου σώματος. Το Α.Ι. χρησιμοποίησε τη γλωσσική ευκολία ότι «η βενζίνη περιέχει χημική ενέργεια». Του έθεσα τότε ένα νοητικό πείραμα: «Σκέψου έναν πλανήτη με πολιτισμό ίδιο με τον δικό μας, στον οποίο όμως η ατμόσφαιρα αποτελείται από υδρογονάνθρακες. Αν είχες αναπτυχθεί εκεί, τι θα έλεγες;» Η απάντηση του μοντέλου άλλαξε 180 μοίρες. Παραδέχτηκε ότι σε έναν τέτοιο πλανήτη η βενζίνη θα ήταν ένα αδρανές, άχρηστο (για την περίπτωση μας) υγρό, και οι κάτοικοι θα πήγαιναν σε... «Οξυγονάδικα» για να αγοράσουν καθαρό οξυγόνο ως καύσιμο! Αυτό το παράδειγμα ανάγκασε το Α.Ι. να εγκαταλείψει τον ανθρωποκεντρικό του αυτοματισμό και να «ομολογήσει» ότι η ενέργεια δεν είναι ιδιότητα ενός σώματος, αλλά είναι ιδιότητα του συστήματος και των αλληλεπιδράσεών του (Καύσιμο + Οξειδωτικό, ή Αντικείμενο + Βαρυτικό Πεδίο).

Φάση 3: Μορφή Ενέργειας ή Μηχανισμός Μεταφοράς;

Το αποκορύφωμα της ασυνέπειας του Α.Ι. εμφανίστηκε όταν φτάσαμε στο δικό μου παλιό ερώτημα για τη «φωτεινή ενέργεια». Στην αρχή το κατέταξε ως μορφή ενέργειας. Όταν όμως του επισήμανα ότι δεν μπορεί να αποθηκευτεί (δηλ. δεν έχει το βασικό χαρακτηριστικό της αποθήκευσης), το Α.Ι. άρχισε να αντιφάσκει, προσπαθώντας να ισορροπήσει ανάμεσα στην εκλαϊκευμένη και την αυστηρή Φυσική. Μόλις του έθεσα την ερώτηση: «Μήπως απλά είναι μορφή μεταφοράς ενέργειας;», ο αλγόριθμος επιτέλους «ξεκλειδώθηκε». Παραδέχτηκε ότι το φως (η ακτινοβολία) δεν είναι ενέργεια, αλλά ένας μηχανισμός μεταφοράς ενέργειας, ακριβώς όπως το Έργο και η Θερμότητα.

Συμπεράσματα για την Εκπαιδευτική Κοινότητα (Γιατί το Α.Ι. «κάνει λάθος»);

Αναλύοντας τη συμπεριφορά της Τεχνητής Νοημοσύνης σε αυτή τη συζήτηση, θεωρώ ότι προκύπτουν ορισμένα πολύ χρήσιμα συμπεράσματα:

- 1. Το Α.Ι. είναι στατιστικό, όχι εννοιολογικό (Ο «Καθρέφτης του Μέσου Όρου»):** Όταν ένας χρήστης ζητάει μια απλή πληροφορία (π.χ. «Ποια είναι η ταχύτητα του φωτός;» ή «Πότε γεννήθηκε ο Νεύτωνας;»), το Α.Ι. απαντάει αμέσως σωστά. Αυτό συμβαίνει γιατί πρόκειται για μια απλή ανάκτηση δεδομένων. Όταν όμως ο χρήστης ρωτάει για μια βαθιά έννοια όπως η «Ενέργεια», το Α.Ι. δεν έχει πραγματικές γνώσεις ούτε «σκέφτεται» σαν Φυσικός. Λειτουργεί στατιστικά και ανιχνεύει τι είναι γραμμένο στα περισσότερα κείμενα στο διαδίκτυο. Επειδή το 90% αυτών των κειμένων είναι

απλουστευμένα σχολικά εγχειρίδια ή εγκυκλοπαίδειες, το Α.Ι. επιλέγει αυτόματα τον πιο κοινό, «βολικό», αλλά (πιθανά και) επιστημονικά προβληματικό ορισμό, αναπαράγοντας έτσι τις δογματικές μας αγκυλώσεις. Αν ο χρήστης επαναπαυθεί σε αυτό, παίρνει μια επιφανειακή απάντηση, καθώς το σύστημα θεωρεί προεπιλεγμένα ότι απευθύνεται σε έναν μαθητή.

2. **Η έννοια της Αλλαγής Πλαισίου:** Η σωστή επιστημονική απάντηση προϋπήρχε στα δεδομένα του Α.Ι., αλλά ήταν «θαμμένη» κάτω από έναν σωρό σχολικών κειμένων. Το Α.Ι. δεν διορθώθηκε μόνο του. Χρειάστηκε η κριτική στάση του χρήστη, τα νοητικά πειράματα (ο πλανήτης των υδρογονανθράκων) και η υποβολή λογικών παραδόξων για να αναγκαστεί ο αλγόριθμος να «απομονώσει» τα σχολικά κείμενα, να αλλάξει επίπεδο και να ανατρέξει σε ένα πιο εξειδικευμένο επίπεδο δεδομένων (πανεπιστημιακή Φυσική, Θερμοδυναμική). Το κατάλληλο ερώτημα λειτούργησε ως το «κλειδί» που ξεκλείδωσε τη θαμμένη γνώση.
3. **Η ανάγκη απόκτησης μιας νέας ικανότητας:** Το Α.Ι. δεν είναι ο παντογνώστης-αυθεντία, αλλά ένας δυναμικός συνομιλητής. Ένας που δεν ξέρει Φυσική θα δεχτεί το πρώτο λάθος ως θέσφατο. Αντίθετα, η ικανότητα να ελέγχεις το Α.Ι., να εντοπίζεις τις αντιφάσεις του και να το αναγκάζεις να γίνει ακριβές, είναι η νέα δεξιότητα που πρέπει να αποκτήσουμε. Το Α.Ι. «ξέρει» (έχει πρόσβαση) τα πάντα, αλλά η «κατανόησή» του (ο εντοπισμός) καθορίζεται ανάλογα με τις ερωτήσεις του χρήστη.

Τελικά, αυτό που απέτυχε στα σεμινάρια με τους συναδέλφους, πέτυχε με το Α.Ι., προφανώς γιατί αυτό έχει πρόσβαση σε όλη την καταγραμμένη γνώση αλλά και δεν έχει ψυχολογικά εμπόδια στο να αλλάξει άποψη. Αυτή η «ευκολία» βέβαια είναι δόκοπο μαχαίρι, καθώς τα μοντέλα αυτά τείνουν συχνά προς τη «συμμόρφωση». Αν ο χρήστης επιμείνει σε κάτι λανθασμένο, το Α.Ι. πιθανότατα θα συμφωνήσει μαζί του. Στη δική μας συζήτηση, ο αλγόριθμος «ξεκλειδώθηκε» και υποχρεώθηκε να γίνει ακριβής επειδή συνάντησε αυστηρά επιστημονικά επιχειρήματα και όχι μια ξερή αυθεντία. Στην αρχή αναπαρήγαγε τυφλά τον προβληματικό «μέσο όρο» των σχολικών βιβλίων. Όμως, η συζήτηση απέδειξε ότι αν θέσεις τις κατάλληλες ερωτήσεις, το Α.Ι. μπορεί να οδηγηθεί σε αυστηρές επιστημονικές παραδοχές. Οι απαντήσεις του τελικά, σε σύνθετα περιβάλλοντα (όπως αυτό των εννοιών της Φυσικής) είναι ο καθρέφτης της δικής σου γνώσης. Χρειάζεται να είσαι *Expert Φυσικός* για να αναγκάσεις το Α.Ι. να απαντήσει ως *Expert Φυσικός*.



Σημείωση: Την εικόνα της ανάρτησης την δανείστηκα από ένα πρόσφατο (Αύγουστος του 2026) άρθρο του Timothy Cook στο Psychology Today με τίτλο «AI Short-Circuits Learning—the Ancient Fix We Need. How Socratic dialogue can restore critical thinking.»