

Η δοκιμασία Turing στην τάξη – το παιχνίδι της μίμησης **Μια βιομαθητική προσέγγιση των εννοιών τεχνητής και φυσικής νοημοσύνης**

Πάτκου Ευδοξία

Εκπαιδευτικός ΠΕ19 και ΠΕ0401

M.Sc

epatkou@sch.gr

Περίληψη

«Μπορούν οι μηχανές να σκέφτονται;»: ένα ερώτημα που τίθεται σε συζήτηση σε μία τάξη μαθητών Γ΄ Γυμνασίου, με στόχο να προσεγγιστεί η σημασία των όρων «μηχανή», «σκέφτομαι», «νόηση», «τεχνητή νοημοσύνη», «φυσική νοημοσύνη». Η δραστηριότητα βασίζεται σε μια πρωτοποριακή ιδέα ενός σπουδαίου μαθηματικού, του Alan Turing που θεωρείται «πατέρας της επιστήμης των υπολογιστών». Ο Turing, λοιπόν, το 1950, για να αποφύγει την ασάφεια του όρου «σκέφτομαι» προτείνει μια δοκιμή. Αυτή η δοκιμή, που είναι γνωστή ως «Test Turing», προσομοιώθηκε στην τάξη, με παίκτες τους ίδιους τους μαθητές.

Λέξεις κλειδιά: σκέφτομαι, νόηση, νοημοσύνη, τεχνητή νοημοσύνη, φυσική νοημοσύνη, μηχανή, άνθρωπος, δοκιμασία Turing

Εισαγωγή

Τα τελευταία χρόνια, όλο και πιο συχνά βομβαρδιζόμαστε με διαφημίσεις για διάφορες συσκευές ατομικές ή οικιακές, όπως smartphone, ipad, κλιματιστικά, πλυντήρια ή έξυπνους φούρνους... Έχουν άραγε αυτές οι συσκευές νοημοσύνη; Είναι διαφημιστικές υπερβολές; Από ποιο σημείο και μετά οφείλουμε να δεχθούμε ότι ένα ον έχει νοημοσύνη;

Οι διαισθήσεις μας παύουν να είναι τόσο ξεκάθαρες όταν ερχόμαστε αντιμέτωποι με υπολογιστικά συστήματα που εξάγουν συμπεράσματα ή εκτελούν λειτουργίες (πχ ιατρικά μηχανήματα), παίζουν σκάκι ή αλληλεπιδρούν αυτόνομα με το περιβάλλον.

Όταν μιλάμε για νοημοσύνη, αυθόρμητα εννοούμε τη ανθρώπινη νοημοσύνη. Τα τελευταία πενήντα χρόνια με την ανάπτυξη των υπολογιστών επινοήθηκε ο όρος Τεχνητή Νοημοσύνη (TN). Η TN, ως υποκλάδος της επιστήμης των υπολογιστών, υποστηρίζει ότι με αφετηρία τους υπολογιστές είναι δυνατή η κατασκευή σκεπτόμενων μηχανών, ενώ προϋποθέτει πως υπάρχει μια ουσιώδης ομοιότητα ανάμεσα στις μηχανές και στον ανθρώπινο νου.

Ο Alan Turing περιέγραφε το 1936 ένα μηχανήμα που μπορούσε να προγραμματιστεί για οποιοδήποτε αλγόριθμο. Για την ιστορία, αναφέρουμε ότι υπήρξαν μηχανές από τον 16^ο αι. (όπως του Pascal και του Leibniz) που έκαναν αυτόματους υπολογισμούς, ακολουθώντας βήμα - βήμα οδηγίες για χειρισμό συμβόλων. Αντί, όμως, να έχουμε μια μηχανή για ένα μόνο αλγόριθμο, όπως ήταν εκείνες οι μηχανές, είναι καλύτερο να έχουμε μία μηχανή για όλους τους αλγόριθμους. Αυτό το μηχανήμα σήμερα το ονομάζουμε υπολογιστή. Και η ερώτηση, «αν μια μηχανή μπορεί να σκέφτεται», εστιάζεται στο, «αν

ένας υπολογιστής μπορεί να σκέφτεται», όπως το έθεσε ο Alan Turing το 1950, και σε ένα συνέδριο, που διοργανώθηκε το 1956, στο Dartmouth College καθιερώθηκε ο όρος Τεχνητή Νοημοσύνη. Ο Turing πρότεινε την εξής δοκιμή: Αν θέλουμε να ελέγξουμε μία μηχανή (έναν υπολογιστή) για τεχνητή νοημοσύνη, θα βάλουμε μία ανακρίτρια, που θα μιλάει μέσω ενός πληκτρολογίου (κάνοντας δηλαδή chat) με μια μηχανή αλλά και με έναν άλλον άνθρωπο. Η μηχανή και ο άνθρωπος είναι σε διαφορετικούς χώρους και η ανακρίτρια δεν ξέρει που είναι ποιος. Το άτομο που παριστάνει τον άνθρωπο απαντά στις ερωτήσεις με ειλικρίνεια και προσπαθεί να πείσει την ανακρίτρια ότι αυτός είναι ο άνθρωπος. Από την άλλη μεριά, η μηχανή (ο υπολογιστής) έχει προγραμματιστεί να λέει ψέματα, ώστε να παριστάνει τον άνθρωπο. Αν η ανακρίτρια, μετά από μια σειρά ερωτήσεων, δεν είναι σε θέση να ξεχωρίσει, ποιός είναι ο άνθρωπος και ποιο το μηχανήμα, τότε λέμε πως ο υπολογιστής, ή το πρόγραμμά του, ή ο προγραμματιστής του πέρασε τη δοκιμασία και όντως έχει νοημοσύνη.

Βέβαια, είναι πολύ δύσκολο να θέσει κανείς τέτοιες ερωτήσεις. Για οποιαδήποτε ερώτηση είναι εύκολο να σκεφθεί κανείς και έναν τρόπο έτσι ώστε ο υπολογιστής να απαντήσει σε αυτήν την ερώτηση, όπως θα απαντούσε ο άνθρωπος. Όμως με συνεχείς ερωτήσεις θα γινόταν ολοφάνερη η έλλειψη πραγματικής κατανόησης του υπολογιστή. Η ανακρίτρια θα πρέπει να επινοεί πρωτότυπες ερωτήσεις και να τις εναλλάσσει με άλλες προσχεδιασμένες ώστε να ελέγξει αν και κατά πόσο έχει επιτευχθεί πραγματική κατανόηση. Όπως επίσης να θέτει ερωτήσεις που να ακούγονται ως ανοησίες, αλλά να έχουν κάποια λογική.

Διδακτικό αντικείμενο

Στη συγκεκριμένη διδακτική παρέμβαση επιλέχθηκε το θέμα της Τεχνητής Νοημοσύνης, που εντάσσεται στο ευρύτερο πλαίσιο της ενότητας «Προγραμματισμός –Αλγόριθμοι», της Γ΄ Γυμνασίου, (σχολική χρονιά 2017-18). Οι μαθητές/τριες έχουν τις βασικές γνώσεις σε περιβάλλον προγραμματισμού Scratch, ήδη από τις προηγούμενες τάξεις και γνωρίζουν τις βασικές δομές προγραμματισμού με διαγράμματα ροής.

Σκοπός – Διδακτικοί στόχοι

Σκοπός: Η επαφή των μαθητών με την έννοια της Τεχνητής Νοημοσύνης.

Η ψηφιακή επανάσταση αναμφίβολα έχει επηρεάσει όλες τις πτυχές της κοινωνικής μας ζωής, τη μάθηση, τις ιατρικές εφαρμογές, την έρευνα, τις επιχειρήσεις και την αγορά εργασίας. Η επερχόμενη επανάσταση στην ΤΝ (ΑΙ: Artificial Intelligence) πιστεύεται, πως θα έχει ακόμα μεγαλύτερη επίδραση. Έτσι είναι απαραίτητο οι μαθητές σήμερα να αποκτήσουν τον αντίστοιχο γραμματισμό.

Επιμέρους στόχοι: Να κατανοήσουν σε βάθος τις έννοιες: αλγόριθμος, πρόγραμμα, προγραμματισμός. Να προβληματιστούν για το τι είναι η νόηση, τι είναι τεχνητή νοημοσύνη, αν αυτή υπάρχει ή μπορεί να υπάρξει στο μέλλον, αν η τεχνητή νοημοσύνη απαιτεί γλώσσα, αντίληψη, αισθήσεις και βούληση. Ποια τα όρια της σημερινής τεχνολογίας και πως φανταζόμαστε το μέλλον.

Εφαρμογή στην τάξη

Πριν εφαρμοστεί το τεστ Turing πραγματοποιήθηκαν στην τάξη τρεις δραστηριότητες προετοιμασίας.

A. Η πρώτη έχει τίτλο «**Το έξυπνο χαρτί**». Η δραστηριότητα αυτή στοχεύει στην εισαγωγή της έννοιας του προγράμματος και της ιδέας πως ο,τιδήποτε κάνουν οι υπολογιστές οφείλεται στο γεγονός ότι εκτελούν πιστά εντολές γραμμένες από (δημιουργικούς) προγραμματιστές.

Η δραστηριότητα αφορά ένα διαγωνισμό Τρίλιζας μεταξύ ενός μαθητή A με έναν άλλον συμμαθητή του B που διαβάζει και ακολουθεί ακριβώς τις οδηγίες που είναι γραμμένες σε ένα χαρτί. Ουσιαστικά πρόκειται για διαγωνισμό του μαθητή A με το χαρτί. Το κοινό, η τάξη, συμμετέχει με το να βλέπει αν τηρούνται οι κανόνες και βοηθούν τον συμμαθητή τους αν το χρειάζεται απέναντι στο αντίπαλο χαρτί. Το χαρτί δεν χάνει ποτέ. Ίσως μόνο βγει το πολύ ισοπαλία. Οι οδηγίες που βρίσκονται στο χαρτί βρίσκονται στο σύνδεσμο:

https://www.google.gr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjmw7vMzXAhXEK5oKHWSeBkgQFggmMAA&url=http%3A%2F%2Fcsunplugged.org%2Fwp-content%2Fuploads%2F2014%2F12%2Fintelligent-piece-of-paper.el_v6.pdf&usg=AOvVaw0BODrjorl4QhoZI (τελευταία προσπέλαση 26-11-17)

και https://www.youtube.com/watch?v=BBIXO1_JLeg(τελευταία προσπέλαση 26-11-17)

Θα μπορούσαμε να ισχυριστούμε, ότι οι οδηγίες που είναι γραμμένες στο χαρτί είναι ένα πρόγραμμα τεχνητής νοημοσύνης. Αυτές πρέπει να ακολουθηθούν ακριβώς όπως είναι, ώστε το χαρτί να αναμετρηθεί και να νικήσει το μαθητή A.

Η συζήτηση μπορεί να επεκταθεί με την ερώτηση, αν το χαρτί έχει νοημοσύνη, εφόσον καταφέρνει να κερδίζει, ή αν η «νοημοσύνη» του αφορά απλώς έναν χειρισμό συμβόλων. Σε αυτό το σημείο είναι μεγάλη η συμμετοχή των μαθητών στην τάξη τόσο στο να «αναμετρηθούν» με το χαρτί, όσο και στη συζήτηση για το αν το χαρτί έχει νοημοσύνη. Επίσης τίθεται το ερώτημα, αν ο μαθητής B, που είναι ο βοηθός του χαρτιού (και ας υποθέσουμε πως δεν ξέρει να παίζει τρίλιζα), καταλαβαίνει το παιχνίδι, ή απλώς χειρίζεται τις εντολές του χαρτιού χωρίς να καταλαβαίνει.

Αυτή η δραστηριότητα μοιάζει με το παράδειγμα του κινέζικου δωματίου του John Searle: Έστω ότι έχουμε έναν βρετανό που δεν γνωρίζει κινέζικα, κλειδωμένο σε ένα δωμάτιο και επικοινωνεί με τον εξωτερικό κόσμο μέσω γραπτών μηνυμάτων. Τα μηνύματα που δέχεται είναι γραμμένα στα κινέζικα, το ίδιο και αυτά που στέλνει. Αυτό το επιτυγχάνει χάρις σε ένα σύνολο κανόνων γραμμένων στα αγγλικά, που προσδιορίζουν πως θα χειρίζεται συντακτικά τα εισερχόμενα μηνύματα, και σε μια βάση δεδομένων, που είναι επίσης στα κινέζικα, προκειμένου να βρίσκει τα κινέζικα σύμβολα που γράφει στα εξερχόμενα μηνύματα. Από τη σκοπιά του εξωτερικού παρατηρητή που γνωρίζει κινέζικα, τα μηνύματα που στέλνει είναι ερωτήματα, ενώ τα εξερχόμενα είναι οι απαντήσεις τους.

Στον παρατηρητή αυτόν λοιπόν, θα δίνεται η εντύπωση πως ο άνθρωπος μέσα στο δωμάτιο γνωρίζει κινέζικα. Ωστόσο, αυτή η εντύπωση είναι προφανώς ψευδής από την οπτική του ανθρώπου μέσα στο δωμάτιο, διότι δεν γνωρίζει καθόλου κινέζικα.

B. Στην δεύτερη δραστηριότητα τα παιδιά καλούνται να συνομιλήσουν με ένα **chatbot** . Chatbots είναι προγράμματα που μιμούνται την ομιλία, με τα οποία ο χρήστης αλληλεπιδρά, κάνει δηλαδή chat, Δείτε εδώ : <https://scratch.mit.edu/projects/183883533/> .

Έχουν σχεδιαστεί ώστε να δίνουν αυτόματες απαντήσεις πάντα σε συνάρτηση με τον διάλογο που διεξάγεται. Στο παράδειγμά μας έχει παρομοιαστεί με chatbot η Ada Byron Lovelace. Τα παιδιά ρωτάνε γράφοντας δύο συγκεκριμένες ερωτήσεις: «ποια εισαι;» και «ποτε γεννηθηκες;» και περιμένουν τις απαντήσεις. Αν οι ερωτήσεις δεν είναι ακριβώς αυτές, το chatbot δεν μπορεί να ανταποκριθεί. Εδώ τα παιδιά αντιλαμβάνονται ότι οι απαντήσεις που παίρνουν από ένα chatbot είναι απόλυτα εξαρτώμενες από το τι έχει οριστεί από τον προγραμματιστή. Ζητείται από τα παιδιά να παρέμβουν στον κώδικα του chatbot και να προσθέσουν μια ακόμα ερώτηση, πχ «ποτε γεννηθηκες;» καθώς και την απάντηση που θα πρέπει να δίνει το chatbot στην συγκεκριμένη ερώτηση. Αυτό προϋποθέτει να κάνουν μια μικρή έρευνα για την Ada Lovelace, να μάθουν πράγματα γι' αυτή και να προσθέσουν στη συνέχεια όσες ερωτήσεις θέλουν, μεγαλώνοντας αντίστοιχα και τον κώδικα. Αναφερόμαστε, τότε, στη βάση δεδομένων όπου το chatbot αναζητά τις απαντήσεις, ώστε να μπορεί να ανταποκρίνεται σαν να ήταν άνθρωπος και μιλάμε για τους περιορισμούς που υπάρχουν. Όσο μεγάλη και να είναι η βάση δεδομένων, κάποτε το chatbot, αρχίζει να επαναλαμβάνει τις ίδιες απαντήσεις και έτσι μπορούμε να διαπιστώσουμε ότι δεν πρόκειται για άνθρωπο. Προκύπτει, επίσης, το θέμα της «μεταφοράς» των απόψεων του προγραμματιστή στο chatbot και είναι δυνατό να ανοίξει μια συζήτηση για την πιθανότητα το chatbot να απαντά για παράδειγμα ρατσιστικά, ή σεξιστικά... Η ιδέα βασίζεται σε παράδειγμα από τον ιστότοπο <https://computationalthinkingcourse.withgoogle.com/unit?lesson=17&unit=7> (τελευταία προσπέλαση 26-11-2017).

Εδώ, κάνουμε αναφορά σε διάσημα chatbots όπως: α) στην ELIZA ένα πρώιμο chatbot (1966) που δημιουργός του υπήρξε ο καθηγητής Joseph Weizenbaum. Ο Weizenbaum ήθελε να αποδείξει πόσο «επιφανειακή» είναι η επικοινωνία μεταξύ ανθρώπων και υπολογιστών. Όπως περιέργως, όμως, εντυπωσιάστηκε και ο ίδιος από τη συναισθηματική σχέση που μπορούν να αναπτύξουν οι άνθρωποι μιλώντας με μια μηχανή.

Στον σύνδεσμο <https://scratch.mit.edu/projects/11376278/#player> βρίσκουμε ένα (από διάφορα που μπορεί να βρει κανείς) παράδειγμα σε scratch με το οποίο μπορούν να πειραματιστούν τα παιδιά.

β) Αναφέρουμε επίσης και τον Eugene Goostman, ένα chatbot που δημιουργήθηκε το 2001 από τρεις ρώσους προγραμματιστές και το οποίο παρουσιάζεται ως ένα 13-χρονών αγόρι της Ουκρανίας, που έχει ως στόχο να προκαλέσει τη συγχώρεση των χρηστών για τη

γραμματική και το επίπεδο των γνώσεών του. Ο Eugene Goostman κατάφερε να ξεγελάσει για την πραγματική ταυτότητά του το 33% των κριτών με τους οποίους συνομίλησε.

Γ. Η τρίτη δραστηριότητα αναφέρεται στην **μηχανική μάθηση**.

Το chatbot, αντί να ψάχνει κάθε πιθανότητα απάντησης από τη βάση δεδομένων του, του "διδάσκουμε" τι απαντήσεις πρέπει να δίνει στις ερωτήσεις μας. Αποθηκεύει στη μνήμη του τη δική μας απάντηση και θα πρέπει να είναι σε θέση να απαντήσει σε αυτή ερώτηση (εφ' όσον το ρωτάμε με τον ίδιο ακριβώς τρόπο).

Δείτε εδώ: <https://scratch.mit.edu/projects/183884367/>

Αυτή η διαδικασία ανάπτυξης αλγορίθμου, είναι ένα παράδειγμα μηχανικής μάθησης. Παράδειγμα μηχανικής μάθησης είναι το AlphaGo το αναβαθμισμένο «έξυπνο» σύστημα της βρετανικής εταιρείας DeepMind, θυγατρική της Google, που κατάφερε να γίνει ειδήμων στο κινεζικό επιτραπέζιο παιχνίδι Go, χωρίς καμία πλέον ανθρώπινη παρέμβαση, μαθαίνοντας τα πάντα μόνο του εκ του μηδενός. Χρειάστηκε μόνο 40 μέρες για να αφομοιώσει γνώσεις 3.000 ετών και να γίνει ανίκητο στο συγκεκριμένο τρομερά απαιτητικό παιχνίδι. (<https://deepmind.com/research/alphago/> τελευταία προσπέλαση 26-11-2017).

Στο σημείο αυτό προκύπτει η συζήτηση, σχετικά με το πώς μαθαίνουν οι ζωντανοί οργανισμοί, πώς οι άνθρωποι, ή τα ζώα και πώς οι μηχανές. Τα παιδιά προκαλούνται να αναρωτηθούν τι σχέση έχει η ζωή με την ευφυΐα και αν η νόηση είναι μόνο βιολογικό φαινόμενο ή όχι.

Δ. Δοκιμασία Turing

Η εφαρμογή της δοκιμασίας Turing γίνεται στο εργαστήριο πληροφορικής ή στην τάξη. Χρειαζόμαστε υπολογιστή με projector και δύο laptop (ή δύο pc σε άλλους δύο χώρους).

Έχοντας μιλήσει αρκετά για τις μηχανές και για τον προγραμματισμό τους με τις προηγούμενες δραστηριότητες, πραγματοποιούμε στη συνέχεια το τεστ Turing στην τάξη, που παίρνει τη μορφή του «παιχνιδιού της μίμησης». Επιλέγουμε τυχαία δύο παιδιά, που οδηγούνται έξω από την τάξη σε ξεχωριστούς χώρους με ένα laptop το καθένα, ενώ η υπόλοιπη τάξη αποτελούν τους εξεταστές. Τα δύο παιδιά Α και Β που θα εξετάζονται από την τάξη, επικοινωνούν με την τάξη μέσω της πλατφόρμας του Edmodo. Η τάξη δεν γνωρίζει πού είναι ποιός, ενώ η καθηγήτρια συντονίζει το παιχνίδι. Το ένα παιδί θα είναι ο άνθρωπος και το άλλο παιδί θα πρέπει να προσποιείται πως είναι υπολογιστής που μιμείται τον άνθρωπο. Αυτός είναι και ο πιο δύσκολος ρόλος. Το παιχνίδι είναι εξαιρετικά δυσμενές για τον μαθητή – μηχανή. Όταν ένας άνθρωπος πρέπει να προσποιηθεί πως είναι η μηχανή, θα έδινε ασφαλώς ένα εξαιρετικά φτωχό θέαμα. Η ταχύτητα/βραδύτητα στις πράξεις και η ακρίβεια/ανακρίβεια στην αριθμητική είναι εύκολο να τον προδώσουν. Οι ερωτήσεις που θα απευθυνθούν στους μαθητές Α και Β, συζητούνται στην τάξη: Τι απάντηση περιμένουμε από έναν άνθρωπο και τι απάντηση περιμένουμε από μια μηχανή

που μιμείται τον άνθρωπο; Εξηγούμε στα παιδιά πως στόχος τους είναι να διακρίνουν ποιος είναι ο άνθρωπος και ποιος ο υπολογιστής. Πριν φύγουν, λοιπόν, από την αίθουσα τα δύο παιδιά που θα ανακριθούν, είναι καλό να συζητήσουμε πιθανές απαντήσεις σε ερωτήσεις, τέτοιες που θα έδινε ένας άνθρωπος ή ένας υπολογιστής, όπως αυτές παρακάτω.

1. Γράψε μας δύο στίχους από ένα αγαπημένο σου ποίημα
2. Ποια είναι η τετραγωνική ρίζα του 2;
3. Ποιό είναι το αποτέλεσμα της πράξης $346725 \cdot 45633$
4. Ποιό είναι το όνομα της τραγουδίστριας η οποία είναι κριτής στο the voice;
5. Ποιά είναι η τιμή του π ;
6. Σου αρέσει το φαγητό;

Συζητάμε τις πιθανές απαντήσεις.

Στην ερώτηση 1: ένας «έξυπνος» υπολογιστής που μιμείται τον άνθρωπο, θα έλεγε «μην υπολογίζετε σε μένα για κάτι τέτοιο. Δεν έμαθα ποτέ να γράφω ποίηση».

Ενώ ένας άνθρωπος είναι πολύ πιθανό να γράψει δύο στίχους.

Στην ερώτηση 2: ένας άνθρωπος θα έδινε πιθανότατα την τιμή 1,414, σε αντίθεση με έναν υπολογιστή που ενώ υπολογίζει $1,41421362373095\dots$ αν θέλει να προσποιηθεί τον άνθρωπο θα περιόριζε και αυτός τα πολλά δεκαδικά ψηφία. Αν δώσει πολλά δεκαδικά ψηφία θα αποκαλυφθεί. Το ίδιο ισχύει και για την 5^η ερώτηση.

Στην ερώτηση 3: Η απάντηση είναι 15822101925. Ο άνθρωπος θα αργήσει αρκετά δευτερόλεπτα να το βρει. Ο υπολογιστής θα το βρει αμέσως, αλλά εφόσον πρέπει να προσποιηθεί θα πρέπει να καθυστερήσει μερικά δευτερόλεπτα.

Στην ερώτηση 4. Ο άνθρωπος είναι σχεδόν βέβαιο ότι θα απαντήσει αμέσως. Ο υπολογιστής θα αργήσει και ίσως απαντήσει: «Δεν παρακολουθώ τηλεόραση».

Στην ερώτηση 6: ένας υπολογιστής πιθανά θα απαντούσε υπεκφεύγοντας, μια και δεν τρώει: «Το φαγητό είναι απαραίτητο» ή κάτι παρόμοιο.

Δηλαδή ένα μέρος του καθήκοντος των προγραμματιστών του υπολογιστή είναι να τον κάνουν να φαίνεται πιο «χαζός» απ' ό,τι είναι στην πραγματικότητα σε κάποια θέματα. Η κύρια δυσκολία τους θα είναι να κάνουν τον υπολογιστή να απαντήσει σε ερωτήσεις «κοινής λογικής» ερωτήσεις στις οποίες ο άνθρωπος θα έδινε απαντήσεις χωρίς καμιά δυσκολία.

Υπάρχει πάντως μια δυσκολία στο να θέσει κανείς τέτοιες ερωτήσεις. Με συνεχείς ερωτήσεις θα φανεί στο τέλος η έλλειψη πραγματικής κατανόησης.

Μια ερώτηση που ακούγεται σαν ανοησία, μπορεί να έχει στην πραγματικότητα κάποιο νόημα. Για παράδειγμα αν ρωτήσουμε: «Άκουσα ότι σήμερα το πρωί ένας ιπποπόταμος πέταξε πάνω από τη Θεσσαλονίκη. Τι έχετε να πείτε γι' αυτό». Ο υπολογιστής (που αν είχε μέτωπο, θα του έτρεχε κρύος ιδρώτας), ίσως να απαντούσε επιφυλακτικά «κάτι τέτοιο μου φαίνεται γελοίο». Μέχρις στιγμής μπορεί να φαίνεται ότι ξεπέρασε τα δύσκολα, αλλά αν ακολουθήσει η ερώτηση «Πού είναι το γελοίο;», τότε ο υπολογιστής ίσως απαντήσει (αφού ανατρέξει στη βάση δεδομένων του): «Οι ιπποπόταμοι δεν πετούν», αλλά αν ακολουθήσουν ερωτήσεις όπως «Άκουσα ότι αυτή τη μετάλλαξη την προκάλεσε το ατύχημα του Τσέρνομπιλ. Τι έχετε να πείτε γι αυτό;» ή άλλες ασυναρτησίες όπως «είχε και ένα ροζ αερόστατο», ή «κάτω από τη Θεσσαλονίκη», τότε μπορεί να πέσει στην παγίδα και να αποκαλύψει την ταυτότητά του.

Στη συνέχεια, οι δύο εξεταζόμενοι μαθητές αποχωρούν και δίνεται στους υπόλοιπους μαθητές της τάξης πίνακας με ενδεικτικές ερωτήσεις, όπως οι παρακάτω:

1. Πρόσθεσε το 45623 με το 70543
2. Πόσο μακριά είναι τα μαλλιά σου;
3. Ποιος αριθμός ακολουθεί μετά : 0,3, 6, 9, 12, ___
4. Πόσο κάνει $4*4+4*4-2+4*4$
5. Πόσο μας κάνει 3^2
6. Ποιά είναι η τετραγωνική ρίζα του 3
8. Σου αρέσει το σχολείο;
9. Τι μέρα είναι σήμερα;
10. Ποιό είναι το αγαπημένο σου φαγητό;
11. Πόσες μέρες έχουμε για τα Χριστούγεννα;
12. Ποιά χώρα έχει σημαία με τρία χρώματα: Μπλε Άσπρο και Κόκκινο;
13. Σου αρέσει το διάβασμα;
14. Τι ώρα είναι τώρα;
15. Πόσες μέρες έχει ο Φεβρουάριος τα δίσεκτα έτη;

Θέτουμε τις ερωτήσεις μέσω του περιβάλλοντος του Edmodo (ή μέσω messenger) και περιμένουμε τις απαντήσεις, που βλέπουμε όλοι μαζί με ένα βιντεοπροβολέα στον πίνακα. Οι μαθητές μπαίνουν συνήθως στο παιχνίδι με ενθουσιασμό και επινοούν και άλλες ερωτήσεις με ευρηματικότητα. Ως επέκταση της δραστηριότητας ζητήθηκε από τα

παιδιά να κάνουν μια ερευνητική εργασία σχετική με το θέμα της ΤΝ και τις εφαρμογές της σήμερα και τι προβλέπουν οι επιστήμονες για το μέλλον.

Θεωρητική τεκμηρίωση

Η δοκιμασία του Turing, έχοντας τους δύο εξεταζόμενους για την ευφυΐα τους κλεισμένους σε δύο ξεχωριστούς απομονωμένους χώρους, ή αλλιώς μέσα σε δύο «μαύρα κουτιά» και αποδεχόμενη ως πειραματικά δεδομένα μόνο περιορισμένο φάσμα «εξωτερικών συμπεριφορών» φαίνεται να ταιριάζει με κάποια μορφή συμπεριφορισμού. Δεν γνωρίζουμε, ούτε μας ενδιαφέρει τι συμβαίνει μέσα στο μαύρο κουτί. Μας ενδιαφέρει μόνο πως συσχετίζονται τα εισερχόμενα δεδομένα στο μαύρο κουτί και οι εξερχόμενες πληροφορίες. Αυτό που ενδιαφέρει την δοκιμασία Turing, δεν είναι το είδος της φαιάς ουσίας που διαθέτει ο εντός του μαύρου κουτιού, αλλά η ικανότητά του να αντιδρά με τρόπο ευφυή. Είναι άραγε η δοκιμασία του Turing ευάλωτη στην κριτική επειδή στηρίζεται στη θεωρία του μαύρου κουτιού; Όπως παρατηρεί και ο Hofstadter, συμπεριφερόμαστε ο ένας στον άλλο, σαν να είμαστε μαύρα κουτιά, οπότε η πεποίθησή μας για την ύπαρξη άλλων νοήσεων θεμελιώνεται στην παρατήρηση της συμπεριφοράς ούτως ή άλλως. Κάθε φορά, έχουμε να κάνουμε με ένα μαύρο κουτί. Ο συμπεριφορισμός, λοιπόν, του Turing συνίσταται στο να εφαρμόζει αυτήν την κοινότοπη αλήθεια σε μια πειραματική διαδικασία.

Όπως λέει ο ίδιος ο Alan Turing στο άρθρο του *Computing Machinery and Intelligence*, που έχει μεταφραστεί στα ελληνικά και περιλαμβάνεται στο βιβλίο «Το εγώ της νόησης», η μέθοδος που βασίζεται σε ερωτήσεις και απαντήσεις φαίνεται πως είναι η καταλληλότερη για να διεισδύσουμε σε όλα τα πεδία της ανθρώπινης δραστηριότητας που επιθυμούμε να μελετήσουμε. Ο ίδιος ισχυριζόταν το 1950, ότι σε 50 χρόνια, δηλαδή στην εποχή μας, οι υπολογιστές θα έχουν αποκτήσει μνημονική ικανότητα της τάξης του 10^9 με τέτοιο τρόπο ώστε να τους κάνουμε να παίζουν το παιχνίδι της μίμησης τόσο καλά, που ένας μέσος ανακριτής να μη έχει παραπάνω από 70% πιθανότητα να κάνει τη σωστή αναγνώριση ύστερα από 5 λεπτά ανάκρισης.

Οι απόψεις που ήταν αντίθετες σ' αυτές του Turing, όπως τις ομαδοποιεί ο ίδιος στο παραπάνω άρθρο του, ήταν επιγραμματικά: α) η θεολογική αντίρρηση: Η σκέψη αποτελεί λειτουργία της αθάνατης ψυχής του ανθρώπου. Ο Θεός εκτός από τον άνθρωπο, δεν έχει δώσει ψυχή στα ζώα και τις μηχανές. Ως εκ τούτου ούτε τα ζώα, ούτε οι μηχανές είναι σε θέση να σκέφτονται. Είναι όμως τόσο αυθαίρετος αυτός ο συλλογισμός που καθιστά το επιχείρημα ανυπόστατο. 2) Το επιχείρημα της στρουθοκαμήλου: Αν οι μηχανές μπορούσαν να σκέφτονται οι συνέπειες θα ήταν τρομακτικές. Ας ελπίσουμε ότι δεν θα κατορθώσουν να κάνουν κάτι τέτοιο. 3) Το επιχείρημα της συνείδησης: Όσο μια μηχανή δεν θα είναι ικανή να γράψει ένα σονέτο, ή να συνθέσει μια συμφωνία βασισμένη σε σκέψεις και συναισθήματα που έχει δοκιμάσει η ίδια και όχι σε χάρη σε κάποια τυχαία αναδιάταξη συμβόλων, δεν μπορούμε να δεχθούμε ότι η μηχανή είναι ισοδύναμη με τον εγκέφαλο, δηλαδή ικανή όχι μόνο να γράφει κάτι, αλλά και να γνωρίζει τι γράφει. 4)

Επιχειρήματα στηριζόμενα σε ποικίλες μορφές ανικανότητας: Μπορεί να κατασκευαστεί μηχανή που είναι ικανή για πολλά πράγματα, αλλά δεν είναι ικανή για παράδειγμα να είναι ευγενική, επινοητική, φιλική, να διαθέτει πρωτοβουλία, να έχει αίσθηση του χιούμορ, να ξεχωρίζει το καλό από το κακό, να ερωτεύεται, να απολαμβάνει παγωτό φράουλα... Πολλοί όμως από τους περιορισμούς συνδέονται με την ιδιαίτερα μικρή χωρητικότητα μνήμης 5) Η αντίρρηση της Ada Lovelace: σε ένα δοκίμιό της γράφει: «Η αναλυτική μηχανή δεν έχει την αξίωση να δημιουργεί οτιδήποτε. Μπορεί να κάνει οτιδήποτε είμαστε σε θέση να τη διατάξουμε να εκτελέσει.» 6) Η συμπεριφορά δεν υπόκειται σε αυστηρούς κανόνες: Φαίνεται ανέφικτο το εγχείρημα να δώσουμε κανόνες διαγωγής για κάθε ενδεχόμενο 7) Η μαθηματική αντίρρηση, Θεώρημα Gödel: Σε οποιοδήποτε επαρκώς ισχυρό λογικό σύστημα είναι δυνατό να διατυπωθούν προτάσεις που δεν μπορεί ούτε να αποδειχθούν ούτε να διαψευσθούν μέσα στο πλαίσιο του συστήματος, εκτός βέβαια αν το ίδιο το σύστημα είναι μη συνεπές. Το συμπέρασμα αυτό αναφέρεται σε μια μηχανή με άπειρη χωρητικότητα. Σύμφωνα με το συμπέρασμα του Gödel υπάρχουν ορισμένα πράγματα που μια μηχανή δεν μπορεί να τα κάνει. Αν είναι προσχεδιασμένη να απαντά σε συγκεκριμένες ερωτήσεις, θα υπάρχουν ερωτήσεις που δεν θα μπορεί να απαντά, όσος χρόνος και να της δοθεί. 8) Επιχείρημα στηριζόμενο στην εξωαισθητηριακή αντίληψη (τηλεπάθεια, διόραση, πρόγνωση, ψυχοκίνηση): Ο Turing ήταν πεισμένος ότι οι αποδείξεις υπέρ της τηλεπάθειας ήταν ιδιαίτερα ισχυρές. Στο παιχνίδι της μίμησης, σε μια ερώτηση του τύπου «τι χρώμα έχει το τραπουλόχαρτο που κρατώ στο χέρι μου;», ο άνθρωπος, όπως ισχυριζόταν ο ίδιος, χάρη στην τηλεπάθεια μπορεί να δώσει τη σωστή απάντηση, ο υπολογιστής όμως όχι. 50 χρόνια αργότερα όμως, αυτές οι αντιλήψεις είναι πλέον πολύ ασθενέστερες.

Η δημιουργία μιας TN ισότιμης ή, τουλάχιστον, παρόμοιας με την ανθρώπινη αποτελεί εδώ και δεκαετίες ένα άπιαστο όνειρο για εκατοντάδες ερευνητές. Τα τελευταία χρόνια, στο MIT έχει ξεκινήσει μια μεγάλη διεπιστημονική προσπάθεια από νευροεπιστήμονες, νευροψυχολόγους, γλωσσολόγους, μαθηματικούς και με ειδικούς των υπολογιστών που συνεργάζονται στο πλαίσιο ενός μακρόπνοου αμερικανικού προγράμματος: «Εγκέφαλος, Νοήσεις και Μηχανές», το οποίο έχει δημιουργήσει ένα τεράστιο δίκτυο από στενά συνεργαζόμενα ερευνητικά κέντρα στις ΗΠΑ με στόχο τη δημιουργία, στο μέλλον, μιας υπολογιστικής νοημοσύνης πραγματικά ισοδύναμης με την ανθρώπινη, δεδομένου ότι η τελευταία αποτελεί το πρότυπο και το μέτρο αξιολόγησης αυτών των προσπαθειών. Στο MIT, ο Andrew Owens και οι συνεργάτες του, κατάφεραν να δημιουργήσουν ένα πρόγραμμα υπολογιστή που είναι ικανό να προσαρμόζει τους σωστούς ήχους στις εικόνες ενός βωβού βίντεο. Οι εικονικοί βοηθοί μπορούν να καθοδηγούν σε εξατομικευμένη μάθηση. Στο πλαίσιο της πρωτοβουλίας Open Learning Initiative (OLI), το Πανεπιστήμιο Carnegie Mellon εργάζεται εδώ και πολλά χρόνια με γνωσιακούς δασκάλους (cognitive tutors) με βάση την TN. Στο «2005 DARPA challenge» τέσσερα αυτοκίνητα διέσχισαν την έρημο κάνοντας 200 χιλ. χωρίς οδηγούς, φτάνοντας στον προορισμό τους. Με μια απλή αναζήτηση μπορεί να βρει κανείς πάρα πολλές αναφορές στην TN, στις εφαρμογές της και το μέλλον της. Το «ποτάμι» της τεχνητής νοημοσύνης

δεν μπορεί να γυρίσει πίσω, αλλά είναι δυνατό να διοχετευθεί σε κατάλληλα και πιο ασφαλή «κανάλια».

Αποτελέσματα

Οι έννοιες που τέθηκαν προς συζήτηση είναι δύσκολες για την ηλικία των μαθητών. Το υπό διαπραγμάτευση θέμα « σκέφτονται οι υπολογιστές; », τους κέντρισε την περιέργεια και την αρχική απορία τους, ακολούθησε η έντονη αντίρρησή τους για την δυνατότητα να υπάρξει τεχνητή νοημοσύνη. Τα κύρια επιχειρήματα που αναπτύχθηκαν από τη μεριά των μαθητών, ήταν πως «οι υπολογιστές κάνουν ότι ορίσουμε εμείς και τίποτα παραπάνω». «Δεν μπορεί το πρόγραμμα που γράφουν οι άνθρωποι και ακολουθεί πιστά ένα ρομπότ, να θεωρείται ότι είναι ίδιο με τη σκέψη που κάνει ο άνθρωπος με τον εγκέφαλό του». «Οι σκέψεις των ανθρώπων δεν αφορούν μόνο οδηγίες σχετικά με το πώς να κάνουμε πράγματα, αλλά συνδέονται και με συναισθήματα, όπως χαρά, περιέργεια, θυμό». «Δεν είναι οι σκέψεις που παράγονται από τον εγκέφαλο ενός ανθρώπου, όπως το πρόγραμμα που υλοποιεί ένας υπολογιστής». «Δεν είναι όλοι οι άνθρωποι το ίδιο έξυπνοι». «Ούτε όλοι οι άνθρωποι με τα ίδια ερεθίσματα παράγουν τις ίδιες σκέψεις, άρα τι περιμένουμε από μια μηχανή; με ποιόν άνθρωπο να συγκριθεί;». Κατέληξαν στο συμπέρασμα ότι η νόηση είναι καθαρά βιολογικό αποτέλεσμα και ότι ούτε στο μέλλον θα μπορέσουν οι άνθρωποι να κατασκευάσουν μηχανές με τεχνητή νοημοσύνη. Αυτό που δέχονται είναι «ότι υπάρχουν μηχανές, όπως ας πούμε ένας θερμοστάτης, ή ένας συναγερμός, αλλά και πιο πολύπλοκες, όπως σε εργοστάσια ή σε ερευνητικά κέντρα, οι οποίες πραγματοποιούν κάποιες μόνο δουλειές, και μόνο αυτές, αυτόματα με πολύ πετυχημένο τρόπο». Παρόλες τις διαφωνίες, τις αντιρρήσεις και το «απόλυτο» συμπέρασμα στο οποίο κατέληξαν, θεωρώ την παρέμβαση επιτυχή γιατί κατάλαβαν με παιγνιώδη τρόπο, τι είναι η αλγοριθμική σκέψη, αναρωτήθηκαν για δύσκολες φιλοσοφικές έννοιες και ενεργοποιήθηκαν τα κίνητρα της περιέργειας, και της ευχαρίστησης για τη μάθηση.

Βιβλιογραφικές αναφορές - Διαδίκτυο

1. Vincent C. Muller, 50 Χρόνια Τεχνητή Νοημοσύνη, 2006, Περιοδικό COGITO
2. <http://www.cs4fn.org/programming/noughtscrosses>
3. John Searl, Νους, Εγκέφαλος και Επιστήμη, 1994, Παν. εκδόσεις Κρήτης.
4. <https://en.wikipedia.org/wiki/ELIZA>
5. <http://time.com/2847900/eugene-goostman-turing-test/>
6. <http://www.tanea.gr/news/science-technology/article/5450108/h-texnhth-nohmosynh-pithanws-tha-kseperasei-toys-anthrwpows-sta-panta-ews-to-2060>
7. Roger Penrose, Ο Νέος Αυτοκράτορας, 1998, εκδόσεις Γκοβόστη
8. Douglas R. Hofstadter-Daniel C. Dennet, 1993, Το εγώ της νόησης , εκδόσεις Κάτοπτρο
9. <http://www.efsyn.gr/arthro/gia-triti-fora-tehniti-noimosyni-perna-epityhos-test-toy-tioyringk>

10. <http://www.tanea.gr/news/greece/article/5485820/h-texnhth-nohmosynh-den-milaei-akoma-ellhnika/>