

Müller, Vincent C. (2006), 'Πενήντα χρόνια τεχνητής νοημοσύνης: γιατί δεν επιτύχαμε ακόμα; [Fifty years of artificial intelligence: Why have we not succeeded yet?]', *Cogito*, 4, 48-49. Reprint: (short version) Angeliofóros newspaper, 26.02.2006, p. 38. – <http://www.sophia.de>, <http://orcid.org/0000-0002-4144-4957>

---

Vincent C. Müller

American College of Thessaloniki

## Πενήντα χρόνια τεχνητής νοημοσύνης - γιατί δεν επιτύχαμε ακόμα;

### 1 Οι Αρχές

Έχουμε την αίσθηση ότι ο υπολογιστής είναι κάτι το καινούριο, το πολύ μοντέρνο, ενώ υπάρχουν αριθμομηχανές από τον 16ο αιώνα: Ο Blaise Pascal και ο G. W. Leibniz φτιάχνανε μηχανήματα για αυτόματους υπολογισμούς - και ο πρώτος τα πούλησε κιόλας. Αυτά τα μηχανήματα ακολουθούσαν τους ίδιους κανόνες που ακολουθεί και ένας άνθρωπος ο οποίος θέλει να κάνει υπολογισμούς - πώς προσθέτουμε δυο μεγάλους αριθμούς; Υπάρχει μια απλή συνταγή, και ακολουθώντας την βήμα-βήμα, μηχανικά, θα φτάσουμε στο αποτέλεσμα. Μια τέτοια βήμα-βήμα συνταγή για τον χειρισμό συμβόλων λέγεται «αλγόριθμος». Η αριθμομηχανή παρακολουθεί τον αλγόριθμο για τον οποίο είναι φτιαγμένη.

Αντί λοιπόν να έχουμε μια μηχανή για τον κάθε αλγόριθμο, θα ήτανε καλό να είχαμε μια μηχανή που να μπορεί να αλλάζει αλγόριθμο. Εμείς αλλάζουμε κάποιες ρυθμίσεις και η μηχανή τρέχει έναν άλλον αλγόριθμο - αυτό ονομάζεται «προγραμματισμός» και ο Charles Babbage ήτανε ο πρώτος που σχεδίασε ένα τέτοιο μηχανήμα το 1840 (το «Analytical Engine», που όμως δεν ολοκληρώθηκε ποτέ).

Ο Alan Turing περιέγραφε το 1936 ένα μηχανήμα, το οποίο μπορεί να προγραμματιστεί για οποιονδήποτε αλγόριθμο - αυτό το μηχανήμα σήμερα το ονομάζουμε «υπολογιστή»: ένα μηχανήμα το οποίο μπορεί να τρέχει οποιοδήποτε αλγόριθμο πάνω σε σύμβολα, ένα καθολικό μηχανήμα. Οι πρώτοι υπολογιστές φτιάχτηκαν το 1941 (Konrad Zuse «Z3») και το 1944 («Harvard Mark I»). Από το 1948 οι υπολογιστές έχουν την βασική δομή που έχουνε και τώρα (την αρχιτεκτονική «von-Neumann») και το 1951 άρχισε οι βιομηχανική παραγωγή τους.

Μου φαίνεται σημαντικό να ξέρουμε καλά τι σημαίνει υπολογιστής αν θέλουμε να κρίνουμε αν ένα τέτοιο μηχανήμα (και όχι οποιοδήποτε μηχανήμα) μπορεί να έχει νοημοσύνη. Την ερώτηση «Μπορεί ένας υπολογιστής να σκέφτεται», τη θέτει ξεκάθαρα ο Turing το 1950 και το 1956 διοργανώνεται ένα συνέδριο στο Dartmouth College όπου και χρησιμοποιείται για πρώτη φορά επίσημα ο όρος «τεχνητή νοημοσύνη». Το 2006 συμπληρώνονται 50 χρόνια τεχνητής νοημοσύνης και ήρθε η ώρα να θέσουμε τις βασικές προβληματικές της. Η τεχνητή νοημοσύνη είναι σε κρίση ταυτότητας: από τη μια μεριά είναι πλέον μια καλά εδραιωμένη και πολύ πετυχημένη μηχανική, από την άλλη δεν φαίνεται να έχει καμία ελπίδα να φτιάξει ένα ολοκληρωμένο σύστημα νοημοσύνης. (Γι' αυτό και άρχισα να γράφω ένα βιβλίο για το συγκεκριμένο θέμα.)

### 2 Η δοκιμασία του Turing

Για να δούμε ποιος μπορεί να είναι αυτός ο μακρινός στόχος της τεχνητής νοημοσύνης θα είναι χρήσιμο να δούμε καταρχάς το πώς ξεκίνησε. Είναι εντυπωσιακό ότι ο Turing το 1950,

σε μια στιγμή που υπήρχαν περίπου 6 υπολογιστές στον κόσμο, δημοσιεύει το άρθρο «On Computing Machinery and Intelligence» όπου θέτει την ερώτηση αν μπορεί ένας υπολογιστής να σκέφτεται και, για να αποφύγει την ασάφεια του «σκέφτομαι», προτείνει μία δοκιμή. Αν θέλουμε να ελέγξουμε ένα μηχάνημα για τεχνητή νοημοσύνη, θα βάλουμε έναν ελεγκτή, ο οποίος θα «μιλάει» μέσω ενός πληκτρολογίου (με «chat») και με το μηχάνημα και με έναν άλλον άνθρωπο - χωρίς όμως να ξέρει ποιος είναι ποιος. Αν ο ελεγκτής έχει μόνο μια δυνατότητα 50/50 να κρίνει σωστά ποιο είναι το μηχάνημα και ποιος ο άνθρωπος, τότε το μηχάνημα όντως έχει νοημοσύνη.

Ο Turing προέβλεπε ότι σε 50 χρόνια θα μπορούσε κανείς να μιλάει για σκεπτόμενα μηχανήματα χωρίς αυτό να θεωρείται περίεργο. Όπως επίσης ότι σε 50 χρόνια θα υπήρχαν μηχανήματα το οποία θα είχαν μια πιθανότητα 70% να περάσουν το ελεγκτικό τεστ μέσα σε 5 λεπτά.

Αν και η δοκιμασία του Turing σήμερα δεν έχει πολλούς υποστηρικτές, μου φαίνεται χρήσιμη, παρά τους εμφανείς περιορισμούς της. Φυσικά και είναι μία δοκιμασία μόνο για ανθρώπινη νοημοσύνη και άρα δεν είναι μια επαρκής δοκιμή για γενική νοημοσύνη (ο έξυπνος Λέων - ο γάτος μας - δεν θα είχε καμία δυνατότητα να το περάσει). Ένας γενικότερος χαρακτηρισμός της νοημοσύνης θα ήταν, για παράδειγμα, η δυνατότητα ευέλικτης συμπεριφοράς στην επιδίωξη στόχων. Επιπλέον η δοκιμασία του Turing ελέγχει μόνο την έξυπνη συμπεριφορά και δεν κοιτάει καθόλου πώς δημιουργείται αυτή, δεν ελέγχει το «μαύρο κουτί». Το τελευταίο δεν με ανησυχεί ιδιαίτερα επειδή δεν πιστεύω ότι μπορούμε να φτιάξουμε ένα μηχάνημα που «κάνει τον έξυπνο», αλλά δεν είναι. (Σε περίπτωση βέβαια που κάποιο μηχάνημα περνούσε τη δοκιμασία, τότε φυσικά και θα ήθελα να δω το «μαύρο κουτί».) Το καλό με τη δοκιμασία του Turing είναι ότι είναι τόσο ευέλικτη, που μας επιτρέπει να μιλάμε για τα πάντα. Το μόνο όριο είναι αυτό το οποίο δεν μπορούμε να λέμε (και γι' αυτό μας έλεγε ο Wittgenstein να σιωπούμε). Τι ερώτηση θα θέτατε για να καταλάβετε αν μιλάτε με υπολογιστή; (Εμένα μου φαίνονται πιο δύσκολες οι εκφορές που δεν είναι ερωτήσεις, όπως η πρόκληση «Γράψε μια απάντηση με τρεις λέξεις».)

Τι έγινε λοιπόν με τις προβλέψεις του Turing; Όσον αφορά τα «μηχανήματα που σκέφτονται», θα έλεγα ότι είχε δίκιο - παρ' όλο που ίσως δεν το λέμε στην κυριολεξία. Υπολογιστές όμως που περνούν τη δοκιμασία του, δεν υπάρχουν. Μερικά «chatbots» κυκλοφορούν στο Internet (ο «Alan» φαίνεται να είναι ο καλύτερος), αλλά κανένα δεν απαντάει παραπάνω από μερικές ερωτήσεις και κανένας σοβαρός επιστήμονας της τεχνητής νοημοσύνης δεν ασχολείται με αυτό το θέμα.

### 3 Η κλασική τεχνητή νοημοσύνη

Ο λόγος για τον οποίο δεν ασχολείται κανείς είναι απλός: η σημερινή τεχνολογία δεν έχει καμία ελπίδα να περάσει τη δοκιμασία.

Ο πιο βαθύς λόγος είναι ότι η τεχνητή νοημοσύνη έχει πάρει μια άλλη κατεύθυνση. Να το πούμε πολύ γενικά, η κλασική τεχνητή νοημοσύνη βασιζόταν στην ιδέα ότι ο ανθρώπινος νους είναι ένας βιολογικός μηχανισμός ο οποίος τρέχει αλγόριθμους - και τώρα έχουμε τον υπολογιστή, ο οποίος μπορεί να τρέχει οποιοδήποτε αλγόριθμο! Μας έμεινε μόνο να βρούμε αυτούς τους αλγόριθμους και να προγραμματίσουμε τον υπολογιστή μας (όπως περιγράφει ο Haugeland το 1985 στο *Artificial Intelligence: The Very Idea*). Αυτό το «μόνο» αποδείχτηκε πάρα πολύ μεγάλο.

Φαινόταν ότι μερικά προβλήματα μπορούμε να τα χειριστούμε σχετικά εύκολα, όπως το σκάκι. Ήταν γνωστό από νωρίς ότι κανένας υπολογιστής δεν μπορεί να υπολογίζει όλα τα πιθανά βήματα ενός παιχνιδιού (είναι περίπου  $10^{120}$ , «1» με 120 «0», περισσότερα από τα άτομα στο σύμπαν), βρέθηκαν όμως έξυπνοι αλγόριθμοι και η δύναμη των υπολογιστών όλο

και μεγάλωνε, μέχρι που το 1997 ο «Deep Blue II» νίκησε τον παγκόσμιο πρωταθλητή (των ανθρώπων), Γκάρυ Κασπάροβ (ενώ είχε χάσει την προηγούμενη χρονιά). Παρά τις επιτυχίες αυτές, στην δεκαετία του 1980 έγινε φανερό ότι οι μέθοδοι που λειτούργησαν σε περιορισμένα προβλήματα (όπως το σκάκι, όπου τίποτα άλλο πέρα από τα 64 τετράγωνα δεν έχει σημασία), δεν μεταφέρονται εύκολα σε προβλήματα της ευρύτερης νοημοσύνης, όπου δεν είναι σαφή τα όρια του συστήματος ως προς το τι έχει σημασία και τι δεν έχει. Αποδείχτηκε τελικά ότι η τεχνητή νοημοσύνη πρέπει να φτιάχνει ένα ολόκληρο ρομπότ, με αντίληψη, γνώση, κίνηση κλπ. Ένα βασικό πρόβλημα είναι πώς να προγραμματίσει στην LISP ή την PROLOG για να προσέχει αυτό που μας φαίνεται πολύ απλό, το πλαίσιο της καθημερινής γνώσης, της ανθρώπινης εμπειρίας, της ζωής!

#### **4 Η τεχνητή νοημοσύνη σήμερα**

Ο πιο σημαντικός κλάδος της τεχνητής νοημοσύνης σήμερα είναι αυτός που σιωπηλά εγκατέλειψε την αρχική μεθοδολογία που παρακολουθούσε τον άνθρωπο και την αντικατέστησε περιορίζοντας την τεχνητή νοημοσύνη σε κάποια είδη προβλημάτων. Η τεχνητή νοημοσύνη ασχολείται με γλώσσα, λογική, αντίληψη, γνώση, κίνηση κλπ. και γενικώς με προβλήματα τα οποία δεν είναι σαφώς ορισμένα. Δεν μας ενδιαφέρει αν οι λύσεις εννοούνται σε ένα γενικότερο πλαίσιο νοημοσύνης, ή αν είναι παρόμοιες με τις ανθρώπινες λύσεις. Όπως αναφέραμε παραπάνω, αυτή η κατεύθυνση έχει βασικά προβλήματα, αλλά είναι και εξαιρετικά πετυχημένη σε μερικά τεχνικά ζητήματα. Για παράδειγμα, τον Οκτώβριο του 2005, 4 αυτοκίνητα διέσχισαν πάνω από 200 χλμ. στην έρημο, φτάνοντας στον προορισμό χωρίς οδηγό (το «2005 DARPA challenge»).

Μια διαφορετική προσέγγιση ακολουθούν οι ερευνητές που υποστηρίζουν ότι δεν πρέπει να αρχίζουμε από τον «υψηλό νου», αλλά από τα βασικά πράγματα. Ερευνητές σαν τον R. Brooks πρότειναν να αρχίσουμε με βασικά κομμάτια, όπως η κίνηση ενός ποδιού - και έχουν φτιάξει απροσδόκητα κατασκευάσματα, όπως ένα «έντομο» ρομπότ με έξι πόδια, το οποίο περπατάει χωρίς κανέναν κεντρικό έλεγχο. Αυτή η μέθοδος «από κάτω προς τα πάνω» έχει πολλές δυνατότητες, αλλά μάλλον θα χρειάζεται να συμπληρωθεί κάποια στιγμή με μια πιο κλασική μέθοδο με αναπαραστάσεις.

Τέλος, αρκετοί νευροεπιστήμονες (όπως οι Damasio, LeDoux, Roth και οι Churchlands) πιστεύουν ότι είμαστε σε θέση να κατασκευάσουμε ένα μοντέλο για τις λειτουργίες του εγκεφάλου - και αυτό θα ανοίξει τον δρόμο για τον προγραμματισμό αυτού του μοντέλου. Προς το παρόν δεν ξέρουμε αρκετά για την λειτουργία του εγκεφάλου μας. Σίγουρα είναι απίστευτα περίπλοκο με περίπου 100 δις. νευρώνες με 1000-10000 συνδέσεις ο καθένας. Φέτος ξεκίνησε το μεγάλο «Blue Brain Project», το οποίο στοχεύει στην κατασκευή ενός μοντέλου μιας νευροκορτικής κολόνας με 60.000 νευρώνες - περίπου ένα εκατομμυριοστό του εγκεφάλου.

#### **5 Η τεχνητή νοημοσύνη του μέλλοντος**

Με τις τεχνολογίες του παρόντος μάλλον θα δυσκολευτούμε να φτάσουμε στην κατασκευή μηχανών με τεχνητή νοημοσύνη. Κατά την γνώμη μου, θα δούμε άλλες τεχνικές λύσεις με την κλασική τεχνητή νοημοσύνη και μέθοδο «από κάτω προς τα πάνω», αλλά δεν περιμένω να υπάρξει ριζοσπαστική πρόοδος πριν μάθουμε πολλά παραπάνω για τον εγκέφαλό μας. Υπάρχουν πολύ καλοί λόγοι για να λέμε ότι το μυαλό μας δεν είναι υπολογιστής και δεν περιμένω να μπορούμε να φτιάχνουμε νόηση μόνο με υπολογιστή. Αλλά θα μπορούσαμε με άλλα μηχανήματα, γιατί όχι;

Ούτε επίσης έχει βάση η συνηθισμένη αντίδραση ότι εμείς, οι άνθρωποι, έχουμε ελεύθερη βούληση και φαντασία και κανένα μηχάνημα δεν μπορεί να έχει αυτά τα προσόντα. Με ποια έννοια έχουμε εμείς ελεύθερη βούληση; Με την έννοια ότι μπορούμε να κάνουμε ό,τι επιθυμούμε (μέσα στα όρια που θέτει η λογική και οι νόμοι της φύσης)] και ότι μπορούμε να αλλάξουμε τις επιθυμίες μας με τη σκέψη. Αυτό μπορεί κάλλιστα να είναι απλώς ένας μηχανισμός. (Και αν ο ντετερμινισμός είναι λάθος, τότε είναι λάθος και για τα μηχανήματα.) Και ναι, αν μόνο ο Θεός μπορεί να φτιάχνει νοημοσύνη, τότε η τεχνητή νοημοσύνη είναι αδύνατη.

Ο στόχος να κατασκευαστεί το «τέλειο ρομπότ» θα είναι εξαιρετικά δύσκολο να επιτευχθεί, αφού ένα τέτοιο ρομπότ πρέπει να έχει γλώσσα, αντίληψη, γνώση, κίνηση, λογική, αισθήσεις και βούληση - και όλα αυτά με την μια! (Δεν μπορεί να έχει, π.χ. λογική και να αποκτήσει βούληση αργότερα.) Το ενδιαφέρον είναι ότι η δοκιμασία του Turing εξακολουθεί να ισχύει και μετά από 50 χρόνια: όποιο μηχάνημα περάσει τη δοκιμασία, θα έχει και τεχνητή νοημοσύνη.