Τεχνητή Νοημοσύνη (ΤΝ) - Artificial Intelligence (AI)

Παρελθόν, Παρόν και Μέλλον

Παναγιώτης Σταματόπουλος

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Ιανουάριος 2018

Τι είναι η ΤΝ;

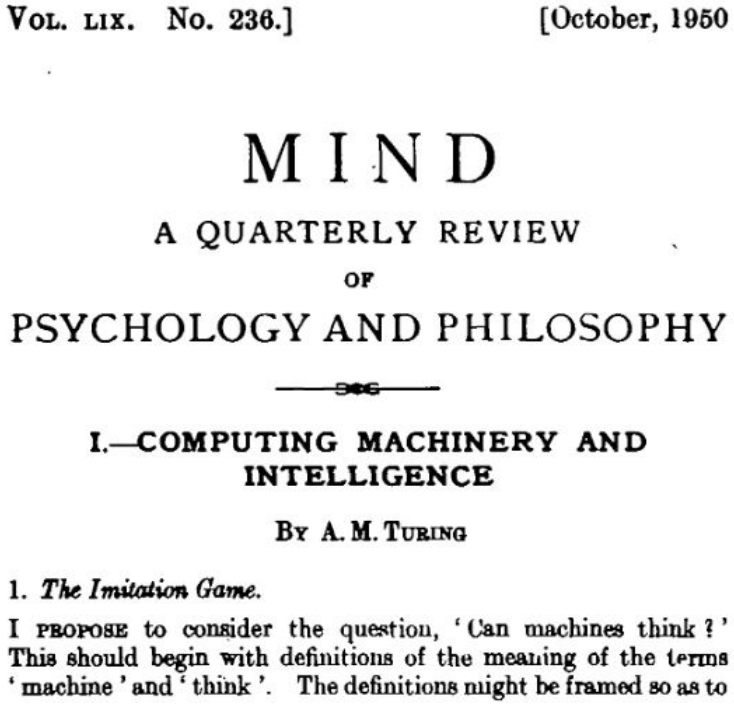
- Είναι η κατασκευή ευφυών μηχανών, όπου ως ευφυΐα για μία οντότητα ορίζεται η ικανότητά της να λειτουργεί κατάλληλα και προβλεπτικά στο περιβάλλον της
- Είναι η μελέτη των διαδικασιών που κάνουν δυνατή την αντίληψη, τον συλλογισμό και τη δράση από μηχανές
- Είναι η κατασκευή μηχανών οι οποίες θα κάνουν πράγματα που σήμερα οι άνθρωποι τα κάνουν καλύτερα

Επιστήμες που αλληλεπιδρούν με την ΤΝ

- Φιλοσοφία
- Μαθηματικά
- Οικονομικά
- Νευροεπιστήμες
- Ψυχολογία
- Γλωσσολογία

Η δοκιμή του Turing

- Alan Turing (1912-1954)
- Δοκιμή για τον χαρακτηρισμό ενός τεχνητού κατασκευάσματος ως «ευφυούς»



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

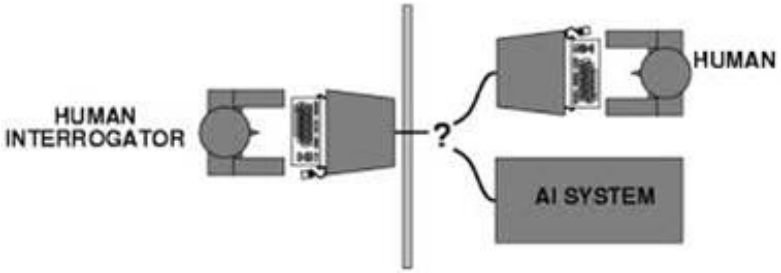
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to



Ομήρου Ιλιάδα - Οι τρίποδες του Ηφαίστου

Ραψωδία Σ, στίχοι 369-381, μετάφραση Καζαντζάκη-Κακριδή

Να πηγαινόρχεται τον πέτυχε στα φυσερά ιδρωμένος
με βιάση, τι είκοσι μαστόρευε μονοφοράς τριπόδια,
στο στέριο γύρω αρχονταρίκι του να στέκουν τοίχο τοίχο.
Μαλαματένιες ρόδες άρμοζε στου καθενού τον πάτο,
μες στων θεών μοναχοσάλευτα[1] τη σύναξη να μπαίνουν
και πίσω να γυρνούν στο σπίτι του
- που να σαστίζει ο νους σου!

[1] «αυτόματοι» στο πρωτότυπο κείμενο

Αριστοτέλεια Λογική

«*Οι Έλληνες είναι άνθρωποι.*» **+** «*Οι άνθρωποι είναι θνητοί.*»
Άρα: «*Οι Έλληνες είναι θνητοί.*»

Από τα
«*Αν κάτι είναι Α, τότε είναι και Β*» και
«*Αν κάτι είναι Β, τότε είναι και Γ*»
συμπεραίνεται το
«*Αν κάτι είναι Α, τότε είναι και Γ*»

Ή, συνολοθεωρητικά:

$$A \subseteq B \ \wedge \ B \subseteq \Gamma \ \Rightarrow \ A \subseteq \Gamma$$

Άλλοι ονειροπόλοι

- Leonardo Da Vinci, 1495, σχεδίαση ανθρωποειδούς ρομπότ
- Jacques de Vaucanson, 1738, κατασκευή μηχανικής πάπιας
- Jacques Offenbach, 1881, μηχανική κούκλα (Ολυμπία) στην όπερα «Τα Παραμύθια του Χόφμαν»
- Karel Čapek, 1921, R.U.R. (Rossum's Universal Robots)
- Isaac Asimov, 1950, «Εγώ, το Ρομπότ»

Βασική ΤΝ, εργαλεία και περιοχές εφαρμογών

- Βασική ΤΝ
  - Αναζήτηση σε χώρους καταστάσεων (Search in state spaces)
  - Αναπαράσταση γνώσης και συλλογιστική (Knowledge representation and reasoning)
- Εργαλεία ΤΝ
  - Νευρωνικά δίκτυα (Neural networks)
  - Εξελικτικοί αλγόριθμοι (Evolutionary algorithms)
- Περιοχές εφαρμογών ΤΝ
  - Κατάστρωση σχεδίου (Planning)
  - Μηχανική μάθηση (Machine learning)
  - Κατανόηση φυσικής γλώσσας (Natural language understanding)
  - Αναγνώριση εικόνων (Image recognition)
  - Ρομποτική (Robotics)

Παραδείγματα προβλημάτων ΤΝ

- Να κατασκευασθεί αυτόματα το ωρολόγιο πρόγραμμα μαθημάτων ενός σχολείου, δεδομένου ενός συνόλου περιορισμών και εκπαιδευτικών στόχων.
- Να επιλυθεί μηχανικά το εξής: «*Στην πόλη έγινε ένας φόνος. Το θύμα ήταν ο Victor. Η αστυνομία συνέλαβε τρεις υπόπτους, τον Abbott, τον Babbitt και τον Cabot. Ο Abbott ισχυρίζεται ότι ο Babbitt ήταν φίλος του Victor και ότι ο Cabot μισούσε τον Victor. Ο Babbitt αρνήθηκε ότι ήταν στην πόλη την ημέρα του φόνου και επίσης είπε ότι δεν γνώριζε τον Victor. Ο Cabot κατέθεσε ότι είδε τον Abbott και τον Babbitt να είναι με τον Victor λίγο πριν το έγκλημα. Η αστυνομία είναι βέβαιη ότι ακριβώς ένας από τους Abbott, Babbitt και Cabot είναι ο ένοχος. Επίσης, υποθέτει ότι οι δύο που είναι αθώοι λένε την αλήθεια. Ποιος είναι ο δολοφόνος;*»

Παραδείγματα προβλημάτων ΤΝ (συν.)

- Να διαχωρίζονται με αυτόματο τρόπο τα e-mails μας σε ανεπιθύμητα (junk) και επιθυμητά.
- Να προσδιορισθούν από ένα ρομπότ οι ενέργειες που πρέπει να εκτελέσει για να επιτύχει ένα στόχο, για παράδειγμα, να τακτοποιήσει με τον βέλτιστο δυνατό τρόπο ένα σύνολο από δέματα μέσα σε ένα δωμάτιο.
- Να μεταφρασθεί αυτόματα ένα αυθαίρετο κείμενο από κάποια φυσική γλώσσα σε μία άλλη, π.χ. από Κινέζικα στα Ελληνικά.
- Να ελεγχθεί αν οι φωτογραφίες των δραστών μίας ένοπλης ληστείας μοιάζουν με κάποιες μέσα από μία βάση δεδομένων ατόμων με βεβαρυμένο ποινικό μητρώο.

Η γέννηση της ΤΝ

Dartmouth Summer Research Project on Artificial Intelligence
(Dartmouth College, New Hampshire, 1956)

- John McCarthy (1927-2011)
- Nathaniel Rochester (1919-2001)
- Claude Shannon (1916-2001)
- Marvin Minsky (1927-2016)
- ....

Σταθμοί στην ιστορία της ΤΝ

- McCulloch/Pitts, 1943: Μοντέλο τεχνητών νευρώνων για την προσομοίωση της εγκεφαλικής λειτουργίας
- Turing, 1950: «Computing Machinery and Intelligence»
- Newell/Simon, 1956: Logic Theorist (αυτόματες αποδείξεις μαθηματικών θεωρημάτων)
- McCarthy, 1958: Γλώσσα προγραμματισμού Lisp
- Gelernter, 1959: Επίλυση γεωμετρικών προβλημάτων
- Samuel, 1959: Πρόγραμμα για το παιγνίδι της ντάμας
- Newell/Simon, 1961: GPS (General Problem Solver)

Σταθμοί στην ιστορία της ΤΝ (συν.)

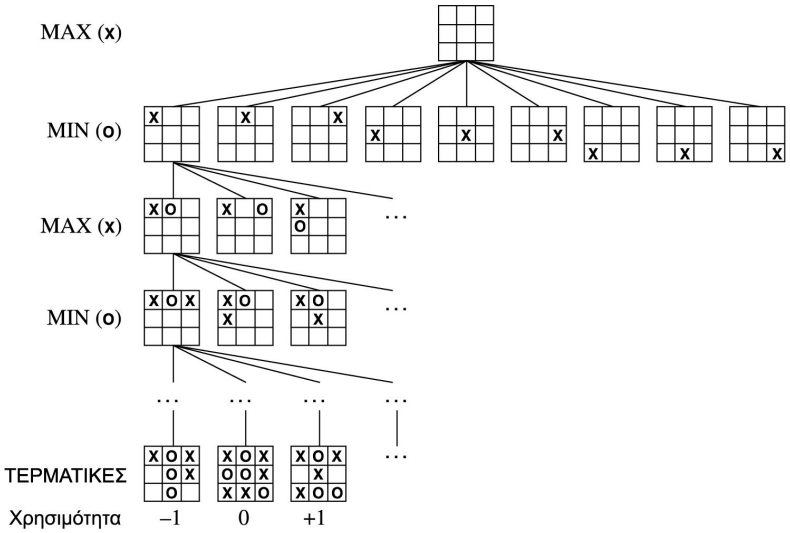
- SRI International, 1966-1972: Shakey the robot
- Stanford, 1969-1973: Έμπειρα συστήματα DENDRAL, MYCIN, PROSPECTOR
- Colmerauer, 1972: Γλώσσα προγραμματισμού Prolog
- 1974-1980: Ο πρώτος «χειμώνας» της ΤΝ
- 1979-1983, Quinlan: Δέντρα απόφασης (ID3, C4.5, C5.0)
- Japan's FGCS, 1982-1990: Fifth Generation Computer Systems project
- 1985-1986: Αλγόριθμος οπισθοδιάδοσης του σφάλματος στα νευρωνικά δίκτυα

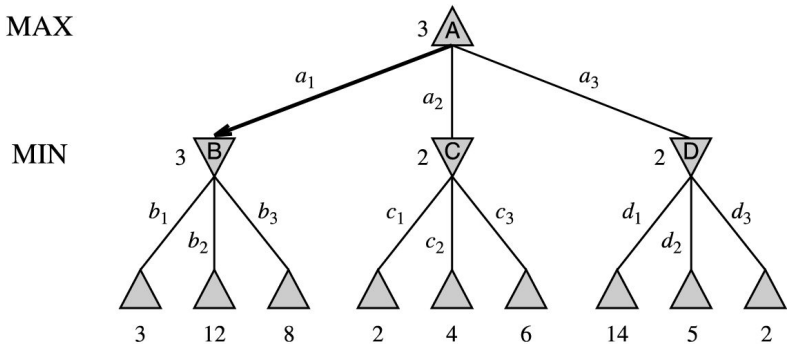
Σταθμοί στην ιστορία της ΤΝ (συν.)

- 1987-1993: Ο δεύτερος «χειμώνας» της ΤΝ
- IBM, Deep Blue (DB), 1996: Garry Kasparov – DB = $4 - 2$
- IBM, Deep Blue (DB), 1997: Garry Kasparov – DB = $2\frac{1}{2} - 3\frac{1}{2}$
- 2000-....: Έμφαση στη μηχανική μάθηση, έκρηξη του «deep learning» και των «big data»
- 2010-....: Αυτο-οδηγούμενα οχήματα
- IBM, Watson, 2011: Νικητής στο παιγνίδι Jeopardy!
- DeepMind, AlphaGo, 2016: Lee Sedol – AlphaGo = $1 - 4$
- DeepMind, AlphaGo, 2017: Ke Jie – AlphaGo = $0 - 3$

Παιγνίδια δύο παικτών

- Παιγνίδια μηδενικού αθροίσματος, τέλειας ή ατελούς πληροφόρησης, με ή χωρίς παράγοντα τύχης
- Αλγόριθμοι
  - Minimax και κλάδεμα άλφα-βήτα
  - Αναζήτηση δέντρου monte carlo





Λογική πρώτης τάξης και αναγωγή

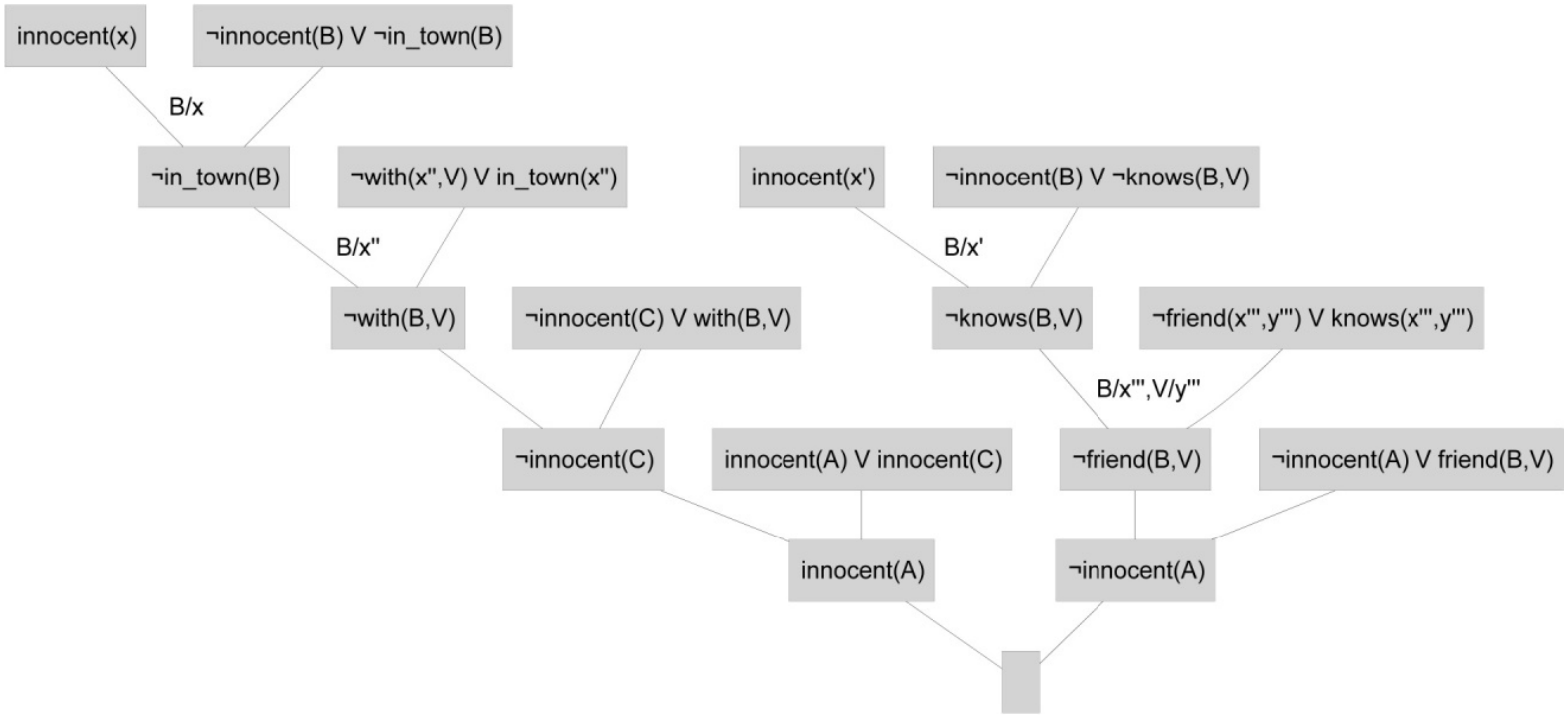
- Διατύπωση της (άμεσης και υπονοούμενης) γνώσης του κόσμου σε λογική πρώτης τάξης
- Προσθήκη της άρνησης του αποδεικτέου και απόδειξη της αντίφασης μέσω αναγωγής

$$innocent(A) \rightarrow friend(B, V)$$
$$innocent(A) \rightarrow hates(C, V)$$
$$innocent(B) \rightarrow \neg in_town(B)$$
$$innocent(B) \rightarrow \neg knows(B, V)$$
$$innocent(C) \rightarrow with(A, V)$$
$$innocent(C) \rightarrow with(B, V)$$
$$\forall x\, [with(x, V) \rightarrow in_town(x)]$$
$$\forall x\, \forall y\, [friend(x, y) \rightarrow knows(x, y)]$$
$$\forall x\, \forall y\, [hates(x, y) \rightarrow knows(x, y)]$$
$$\neg innocent(A) \vee \neg innocent(B) \vee \neg innocent(C)$$
$$(innocent(A) \wedge innocent(B)) \vee (innocent(A) \wedge innocent(C)) \vee (innocent(B) \wedge innocent(C))$$

Αποδεικτέο: $\exists x\, [\neg innocent(x)]$

Άρνηση αποδεικτέου: $\forall x\, [innocent(x)]$

Λογική πρώτης τάξης και αναγωγή (συν.)



Άρα, ο δολοφόνος είναι ο Babbitt!

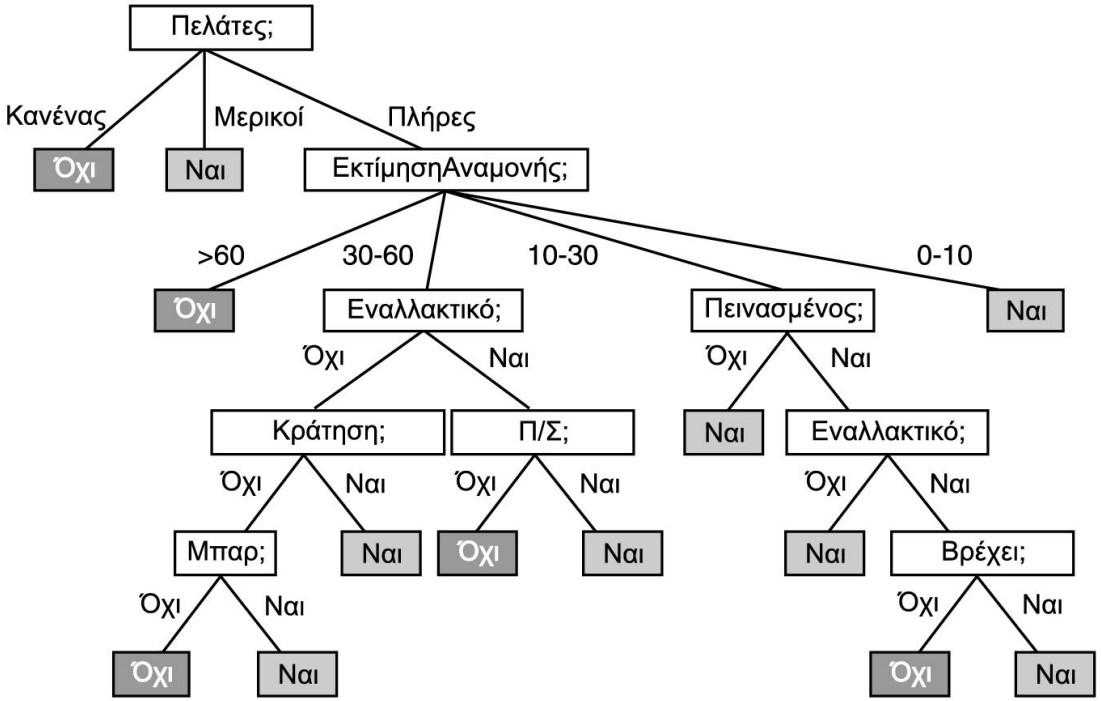
Μηχανική μάθηση (Machine learning)

- Μάθηση με επίβλεψη
- Στόχος είναι η κατασκευή ενός μοντέλου για μία έννοια
- Διαθέτουμε δεδομένα εκπαίδευσης
- Παράδειγμα: Το πρόβλημα του εστιατορίου

Α/Α	Εναλ.	Μπαρ	Π/Σ	Πειν.	Πελ.	Κόσ.	Βρ.	Κρ.	Είδος	Αναμ.	ΘΠ
1	N	O	O	N	Μερ.	$$$	O	N	Γαλλ.	0–10	N
2	N	O	O	N	Πλ.	$	O	O	Ταϋλ.	30–60	O
3	O	N	O	O	Μερ.	$	O	O	Αμερ.	0–10	N
4	N	O	N	N	Πλ.	$	O	O	Ταϋλ.	10–30	N
5	N	O	N	O	Πλ.	$$$	O	N	Γαλλ.	>60	O
6	O	N	O	N	Μερ.	$$	N	N	Ιταλ.	0–10	N
7	O	N	O	O	Καν.	$	N	O	Αμερ.	0–10	O
8	O	O	O	N	Μερ.	$$	N	N	Ταϋλ.	0–10	N
9	O	N	N	O	Πλ.	$	N	O	Αμερ.	>60	O
10	N	N	N	N	Πλ.	$$$	O	N	Ιταλ.	10–30	O
11	O	O	O	O	Καν.	$	O	O	Ταϋλ.	0–10	O
12	N	N	N	N	Πλ.	$	O	O	Αμερ.	30–60	N

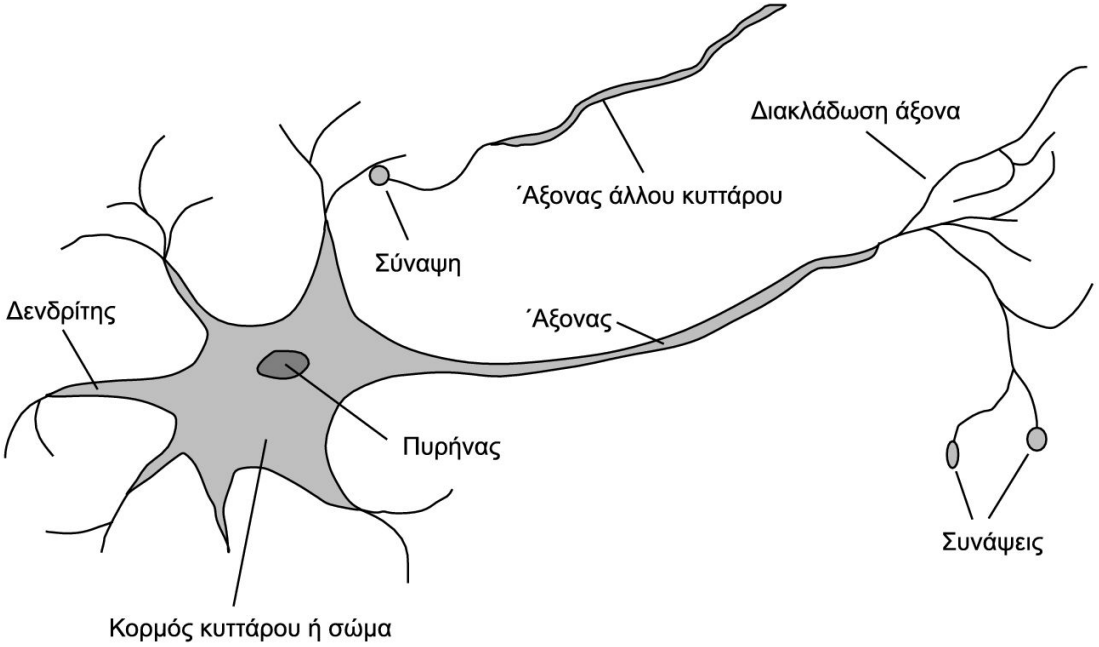
Μηχανική μάθηση (συν.)

- Δέντρα απόφασης
- Αλγόριθμος εκπαίδευσης βάσει του κέρδους πληροφορίας (information gain)



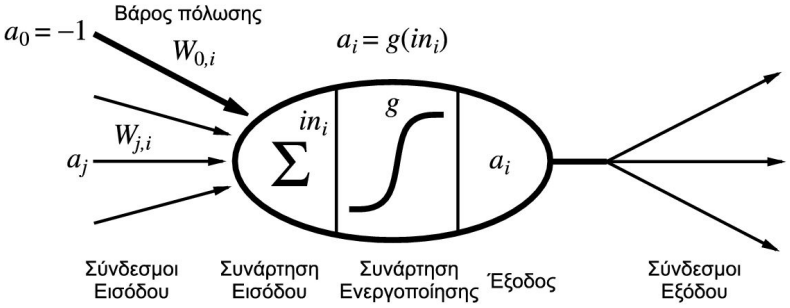
Νευρωνικά δίκτυα (Neural networks)

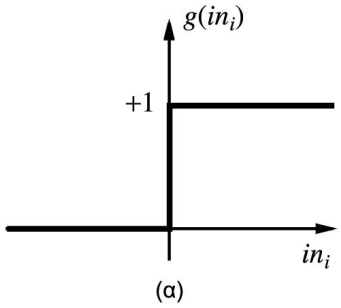
- Βιολογικός νευρώνας
- $\sim 10^{11}$ νευρώνες και $\sim 10^{14}$ συνάψεις στον ανθρώπινο εγκέφαλο

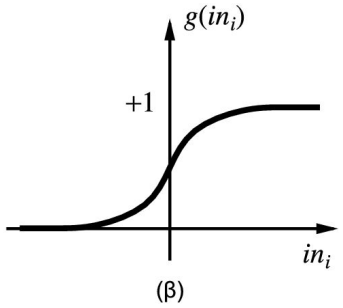


Νευρωνικά δίκτυα (συν.)

- Τεχνητός νευρώνας
- Συναρτήσεις ενεργοποίησης
  - Βηματική ($g(x) = 1$, αν $x \geq 0$, και $g(x) = 0$, αν $x < 0$)
  - Σιγμοειδής ($g(x) = \frac{1}{1+e^{-ax}}$, συνήθως $a = 1$)
  - Γραμμική ($g(x) = x$)
  - Υπερβολική εφαπτομένη ($g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$)

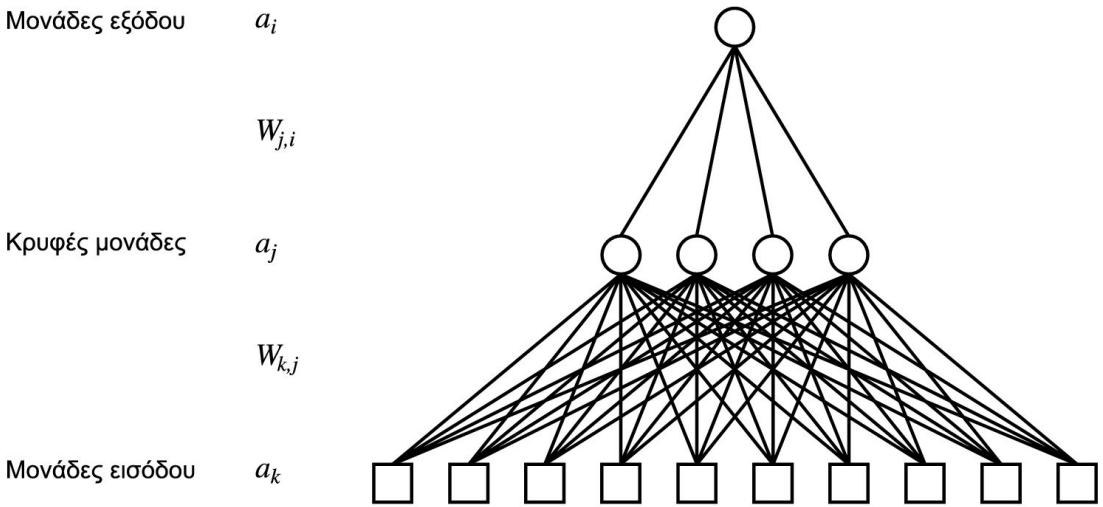






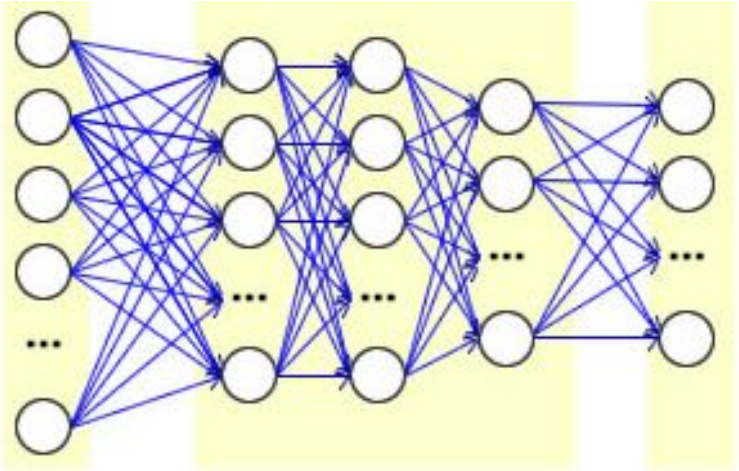
Νευρωνικά δίκτυα (συν.)

- Δίκτυα εμπρόσθιας τροφοδότησης
- Αλγόριθμος εκπαίδευσης των βαρών μέσω οπισθοδιάδοσης του σφάλματος (error backpropagation)

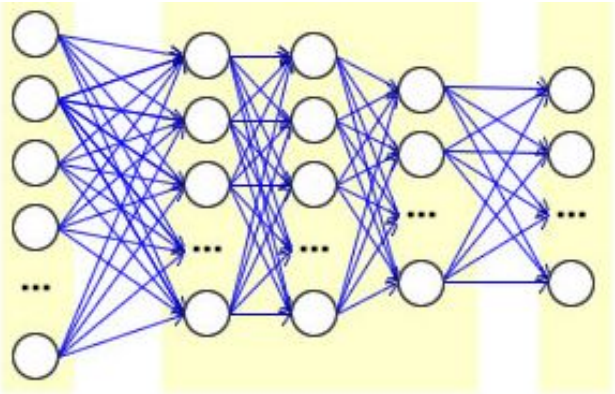


Βαθιά νευρωνικά δίκτυα (Deep neural networks)

- Πολλά κρυμμένα επίπεδα νευρώνων
- Πολλά βάρη προς εκπαίδευση
- Χρειάζονται πολλά παραδείγματα και σημαντική υπολογιστική ισχύς για την εκπαίδευσή τους
- Εφαρμογές σε αναγνώριση εικόνων, ήχου και σε παιγνίδια



Αναγνώριση εικόνας και ήχου



.....

- Amazon Alexa
- Google Assistant
- Apple Siri

Τρέχουσες περιοχές εφαρμογών της ΤΝ

- Μεταφορές (έξυπνα οχήματα, αυτο-οδηγούμενα οχήματα, σχεδιασμός μετακινήσεων)
- Οικιακοί ρομποτικοί βοηθοί (σκούπες)
- Υγεία (κλινική βοήθεια, ανάλυση δεδομένων, ρομποτική χειρουργική, υποστήριξη ηλικιωμένων)
- Εκπαίδευση (έξυπνα μαθησιακά εργαλεία, online συστήματα μάθησης)
- Δημόσια ασφάλεια
- Διασκέδαση

Ερευνητικές τάσεις στην ΤΝ

- Μηχανική μάθηση μέσω μεγάλου όγκου δεδομένων για αναγνώριση αντικειμένων σε εικόνες, χαρακτηρισμό βίντεο, αναγνώριση δραστηριοτήτων, ανάκτηση πληροφορίας από ήχο και φυσική γλώσσα, κλπ.
- Ενισχυτική μάθηση (Reinforcement learning)
- Ρομποτική και μηχανική όραση
- Κατανόηση και επεξεργασία φυσικής γλώσσας
- Πληθοπορισμός (Crowdsourcing)
- Αλγοριθμική θεωρία παιγνίων (Algorithmic game theory)
- Το ίντερνετ των πραγμάτων (Internet of Things - IoT)
- Νευρομορφική υπολογιστική (Neuromorphic computing)

Φιλοσοφικά (και όχι μόνο) ερωτήματα για την ΤΝ

- Ισχυρή ΤΝ: Μπορούν να σκεφτούν οι μηχανές;
- Ασθενής ΤΝ: Ή μήπως απλώς ενεργούν σαν να είναι ευφυείς;
- Μπορεί να επιτευχθεί Τεχνητή Νοημοσύνη παντός σκοπού (Artificial General Intelligence - AGI);
- Αν ναι, πότε; Είμαστε κοντά;

Πηγές

- Satinder P. Singh, Shaul Markovitch. “Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence.” AAAI Press, San Francisco, California, USA, February 4-9, 2017.
- Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller. “Artificial Intelligence and Life in 2030.” One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016. Doc: http://ai100.stanford.edu/2016-report. Accessed: September 6, 2016.
- Stuart Russell, Peter Norvig. “Artificial Intelligence: A Modern Approach.” 3rd edition, Prentice Hall, 2010.
- Nils J. Nilsson. “The Quest for Artificial Intelligence.” Cambridge University Press, 2009.
- Stuart Russell, Peter Norvig. “Τεχνητή Νοημοσύνη: Μία Σύγχρονη Προσέγγιση.” 2η έκδοση, Επιμέλεια Ελληνικής Έκδοσης: Γιάννης Ρεφανίδης, Κλειδάριθμος, 2004.