

INTELIGENȚA ARTIFICIALĂ

Un pas înainte în evoluție?

JUAN CARLOS NUÑO

Traducere de Anca Dinu

CUPRINS

Capitolul 0. AlphaGo	7
Capitolul 1. Introducere	15
Nașterea unei discipline	17
Ce înseamnă inteligența artificială?	24
Rolul matematicii	31
<i>Inteligență extraterestră</i>	32
Inteligența artificială în secolul XXI	36
Calcul neuronal	38
Capitolul 2. Inteligența artificială operativă	45
<i>Operații binare. Cum calculează computerele?</i>	46
Învățare automată	50
Algoritmi de învățare	56
Rețele neuronale	60
<i>Regula lui Hebb</i>	68
Tehnici probabiliste	70
<i>Formula lui Bayes</i>	72
Algoritmi evolutivi (genetici)	74
Inteligența colectivă	79
Sisteme dinamice inteligente	82
<i>Sisteme dinamice</i>	84
Capitolul 3. Aplicații ale inteligenței artificiale	87
Creativitatea artificială	88
Un Rembrandt al secolului XXI	90

O piesă de Bach	90
Un <i>bestseller</i> ?	92
Mașini cu cinci simțuri	93
Medicină inteligentă	98
Analiza și prevenirea criminalității	100
Analiza tendințelor financiare și prognoza riscurilor	101
Capitolul 4. Problema conștiinței	105
Conviețuirea cu inteligența artificială	112
Persoane electronice	115
Pericolele roboticii	118
Inteligența supraumană?	120
<i>Legile roboticii lui Asimov</i>	122
<i>Educația în secolul inteligenței artificiale</i>	124
Nemurirea robotică	126
Epilog	131
Lecturi recomandate	135

Capitolul 0

AlphaGo

*Lumea începe să își dea seama că pășările
sunt mult mai inteligente decât se credea,*
(Jennifer Ackerman)

Poate fi oare un program informatic ca AlphaGo considerat echivalentul unui personaj? Judecând după definițiile standard ale personajului ca „persoană care se distinge, este calificată sau reprezentativă în viața publică”, sau „ființă reală sau ficțională prezentă într-o operă literară, teatrală sau cinematografică”, răspunsul este mai degrabă negativ, deoarece un program informatic nu poate fi o persoană și nici măcar o ființă reală sau ficțională. Și totuși, așa după cum mulți probabil ne-am gândit, ar putea deveni un personaj.

Iată cum inteligența artificială se consolidează ca un domeniu din ce în ce mai larg, care impune cu necesitate redefinirea unor concepte fundamentale ale cunoașterii. Pe parcursul cărții, ne vom deplasa pe acest teren vast, fără granițe clare și în continuă expansiune. Astfel, din această perspectivă, de ce nu ar merita un rol de personaj principal un program informatic dotat cu niște calități excepționale pentru abordarea unor jocuri dificile, precum jocul de go? Un algoritm de-a dreptul remarcabil, capabil nu numai să joace, dar să și câștige contra campionilor ființelor umane.

Indiferent însă de punctul de vedere pe care îl adoptăm, cazul programului AlphaGo este ilustrativ pentru dezvoltarea inteligenței artificiale. AlphaGo a fost unul dintre primele programe care a înfruntat

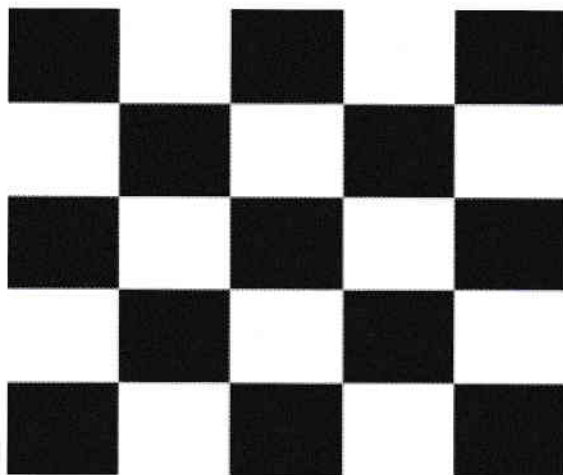


Figura 1. Discretizarea, în special a variabilelor spațiale, constituie baza matematicii algoritmice. Jocul de șah utilizează o rețea discretă ca tablă de joc pe care se mută piesele. Complexitatea acestui joc, ca și a multor alora, derivă tocmai din numărul mare de posibilități de așezare a pieselor pe tablă. Sursa: Yair Neuman. *Mathematical Structures of Natural Intelligence*, Springer International Publishing (2017).

inteligența umană într-una dintre activitățile la care aceasta excellează: jocul.

În mod firesc, jocurile de logică și de strategie, spre deosebire de cele de forță și abilitate (sportive), au făcut mereu parte din societatea umană și, într-un anumit sens, s-au dovedit utile în definirea nivelului de inteligență al celor care le practică – cel puțin într-unele dintre manifestările sale –. Din această categorie fac parte jocul de dame, șahul, go, acesta din urmă larg răspândit în țările asiatice. Din acest motiv, dezvoltarea inteligenței artificiale este strâns legată de obiectivul de a aborda aceste jocuri și de a învinge campionii umani. Mai mult, unii autori susțin chiar că jocurile, și nu numai cele tradiționale de masă,

dar și cele digitale, au jucat un rol decisiv în dezvoltarea inteligenței artificiale¹.

Într-adevăr, jocurile sunt complexe și interesante, îmbogățesc interacțiunea dintre om și mașină, sunt apreciate și stârnesc interesul oamenilor, mai ales pe al tinerilor. În plus, acestea au nevoie de conținuturi, ceea ce fără doar și poate alimentează imaginația și, mai ales în cazul jocurilor interactive, generează un volum mare de date care pot fi utilizate ulterior pentru a crea noi resurse pentru același joc. Așadar, dată fiind importanța jocurilor în dezvoltarea inteligenței artificiale, și analogia evidentă cu insecta care a dus la descoperiri importante în biologie – musca de fructe –, unul dintre aceste jocuri, și anume șahul, a fost etichetat ca „*Drosophila* inteligenței artificiale”.²

În particular, una dintre versiunile programului DeepBlue, aparținând companiei IBM, l-a învins pe campionul de șah Garri Kasparov în anul 1997. La puțin timp după, a fost prezentat Watson, un program capabil să răspundă întrebărilor adresate în limbaj natural, care în 2011, a participat la emisiunea din Statele Unite ale Americii *Jeopardy!*, bătându-i pe campionii edițiilor precedente; astfel, și-a adjudecat un milion de dolari. Acest fapt are o importanță care depășește cu mult cadrul computațional, demonstrând clar potențialul economic al algoritmilor de acest fel.

Fiind recunoscut ca mult mai dificil decât șahul, jocul go a atras atenția imediat cercetătorilor din domeniul inteligenței artificiale, în special celor de la DeepMind, un departament al companiei Google. În pofida regulilor simple, jocul go oferă o cantitate uriașă de alternative, fapt care îl face un joc extrem de complex. Estimarea numărului

¹ Yannakakis, G.N., Togelius, J., 2018, *Artificial Intelligence and Games*, Springer International Publishing.

² După John McCarthy, cercetătorul rus Alexander Kronrod a fost cel care a spus „Șahul este *Drosophila* Inteligenței artificiale”. (A se consulta Marsland, T. A., Schaeffer, J. (eds.), 1990, *Computers, Chess, and Cognition*, Springer.)

minim de poziții posibile pe tabla de go este de ordinul 10^{170} (pe tabla standard, de 19×19 poziții).³

Un astfel de număr de combinații imposibil de conceput depășește vârsta universului măsurată în secunde. În acest joc, învinge cel care reușește să cucerească cel mai mult teritoriu, în timp ce elimină piesele adversarului (numite pietre), „sufocându-l”, adică încercuindu-i piesele.

Programul expert în acest joc se numește AlphaGo, și deja există mai multe versiuni optimizate ale acestuia. Una dintre acestea, cunoscută sub numele de AlphaGo Fan, l-a învins pe campionul european de go Fan Hui în 2015. O versiune ulterioară, AlphaGo Lee, a câștigat patru partide – dintr-un total de cinci – într-un turneu contra unuia dintre cei mai buni jucători ai lumii, Lee Sedol, jucat în Seoul, Coreea de Sud, între 9 și 15 martie 2016. Un an mai târziu, în mai 2017, programul își adjudeca din nou victoria împotriva campionului mondial de la acea dată, Ke Jie. Ultima versiune a acestui program, AlphaGo Zero, lansată la sfârșitul anului 2017, s-a confruntat cu versiunile precedente, iar rezultatele au fost spectaculoase, reușind să își adjudece 100% dintre victorii contra celei mai bune versiuni de până atunci, AlphaGo Lee.

Dar ce anume face această ultimă versiune atât de puternică și net superioară capacităților umane la acest joc? Practic, ultima versiune se distinge de toate celelalte anterioare prin patru aspecte esențiale. Primul, și cel mai important, este faptul că antrenarea acesteia se bazează exclusiv pe un proces de învățare autonomă prin consolidare – asupra căruia vom reveni în cele ce urmează – plecând de la o stare inițială oarecare și fără intervenții din partea omului. În al doilea rând, folosește ca input doar piesele albe și negre prezente pe tablă.

³ În go, numărul de opțiuni atinge niveluri astronomice. Pe o tablă normală de 19×19 , primul jucător are la dispoziție 361 de poziții posibile, iar al doilea, 360 și așa mai departe. În total, 361!, care înseamnă aproximativ 10^{768} , adică 1 urmat de 768 de zerouri. Simetria și alte considerații reduc acest număr uriaș la cel citat în text, 10^{170} (Hendler, J., Mulvehill, A.M., 2016, *Social Machines. The Coming Collision of Artificial Intelligence, Social Networking, and Humanity*, Apress).

Al treilea aspect privește faptul că programul are la bază o singură rețea neuronală. În al patrulea și ultimul rând, programul folosește un arbore de căutare bazat pe rețeaua neuronală pentru a evalua posibile mișcări și poziții, fără a apela la optimizări secvențiale de tip Monte Carlo. În plus, la antrenare, algoritmul prevede o metodă de căutare la distanță care permite o învățare mai eficientă, mai stabilă și mai precisă.⁴ Faptul că programului îi este necesar un antrenament minimal pentru a învăța operații atât de complexe i-a uimit chiar și pe experții din domeniu.⁵

Această victorie reflectă dezvoltarea imprevizibilă cunoscută de inteligența artificială. De fapt, mare parte dintre experți erau conștienți că acest moment avea să vină. Însă lucrul cu adevărat surprinzător este că a venit înainte de al doilea deceniu al secolului XXI.⁶ După cum am subliniat mai devreme, posibilitățile de a muta piesele la jocul go sunt cu mult mai numeroase decât cele de la șah și, în consecință, tehnicile uzuale nu reușeau să le trateze într-un timp rezonabil. Și totuși, apariția noilor tehnici de învățare profundă (*deep learning*, în engleză) a adus un salt calitativ și cantitativ fără precedent în istoria acestei discipline.

§

Asistăm la un moment extrem de prolific pentru cercetarea în domeniul inteligenței artificiale. Nu trece nicio zi fără să apară în mass-media vreo știre despre această disciplină științifică. Numărul de articole publicate în domeniu a explodat. Inteligența artificială este utilă

⁴ A se consulta Silver, D. et al., 2017, „Mastering the Game of Go Without Human Knowledge”, *Nature* 550: 354-359.

⁵ Yang, G.Z. et al., 2018, „The Grand Challenges of Science Robotics”, *Science Robotics* 3.

⁶ A se consulta LazyProgrammer. 2016, *Unsupervised Deep Learning in Python*.

În atâtea probleme interesante și semnificative, că până și neinițiații sunt uimiți în continuu de progresele remarcabile ale domeniului. Mai mult, știrile de acest fel provoacă o reacție de tipul piele de găină, similară reacțiilor la filmele cu suspans, determinată de teama de necunoscut. Este ușor să ne lăsăm influențați de senzația că inteligența artificială este peste tot și că pentru a putea profita de ea este nevoie să începem să ne extindem cunoștințele științifice în multe alte domenii, în afară de știință și inginerie. În particular, matematica nu este străină de o revoluție similară, ceea ce justifică perfect decizia de includere a unui volum dedicat inteligenței artificiale într-o serie de volume pe tema marilor descoperiri ale matematicii.

Nivelul de rigoare matematică acumulat în domeniul inteligenței artificiale este o chestiune controversată, motiv de îngrijorare nu pentru puțini cercetători din domeniu. După cum afirmă Frederick Hoffman, dacă „s-ar intenționa ca matematica să asigure un fundament solid pentru dezvoltarea unei discipline, s-ar putea discuta despre comportamentul unei inteligențe artificiale bazate pe fundamente matematice insuficiente”.⁷ Cu toate acestea, pentru mulți experți care lucrează în acest mediu, în ziua de azi, lucrul cel mai important este să poată aplica algoritmi ca atare pentru a rezolva eficient probleme concrete. În cel mai rău caz, nici nu se pune problema unei justificări matematice pentru anumite metode care sunt utilizate exclusiv în inteligența artificială. După părerea multor cercetători, sistematizarea riguroasă a cunoștințelor, care s-au acumulat în domeniu într-un ritm frenetic, ar trebui să devină încet-încet o linie de cercetare în sine.

Pentru un profesor al unui departament de matematici aplicate, frecventarea domeniului inteligenței artificiale constituie o provocare remarcabilă. Cu toate că recunosc pe deplin relevanța matematicii în dezvoltarea inteligenței artificiale, nu mă voi putea referi exclusiv la rezultate demonstrate matematic; căci astfel nu aș mai putea explora teritorii speculative, care abundă totuși de oportunități în domeniul

⁷ Hoffman, 1998.

inteligenței artificiale. Mai mult, înțelegerea anumitor fundamente matematice care stau la baza unor rezultate poate presupune noțiuni avansate care nu sunt la îndemâna tuturor cititorilor. Faptul că nu dispunem de o justificare riguroasă pentru metodele folosite, în bună parte, îngreunează descrierea acestora. Așadar, profesorul – precum și autorul acestor pagini – este obligat să se miște în permanență în vastul teritoriu dintre adevărat și posibil, ceea ce poate genera anumite temeri.

Pe de altă parte, formatul redus al acestei serii de volume limitează destul de mult amploarea acestei cărți. Cititorii vor înțelege cu siguranță dificultatea de a condensa judicios într-un singur text o varietate imensă de subiecte mai mult sau mai puțin actuale din domeniul inteligenței artificiale. În mod firesc, într-o disciplină cu una dintre cele mai mari rate de producție științifică, este greu de justificat neglijarea actualității imediate. Dar nu se pot ignora nici primii pași care au marcat nașterea inteligenței artificiale, în cea de-a doua jumătate a secolului XX. Analog, importanța majoră a aplicațiilor din acest domeniu nu poate eclipsa importanța principiilor elementare care stau la baza algoritmilor și metodelor utilizate. Această carte, fără a se îndepărta niciun pic de la rigoarea științifică, nu poate evita totuși speculațiile și controversele născute în jurul inteligenței artificiale în societatea noastră. Din acest motiv, ne vom ocupa de tratarea consecințelor socio-politice pe care domeniul le-a provocat pe termen scurt. În ceea ce privește consecințele pe termen lung, vom prezenta eventualele scenarii de luat în seamă, odată trecută furtuna actuală care scutură fundamentele culturale și filosofice ale umanității.

După cum își pot da seama cititorii, acest volum introductiv în inteligență artificială nu poate trata chiar toate temele despre care am vorbit în paragraful anterior. În schimb, invit cititorii să aprofundeze diversele domenii conexe. De altfel, suntem convinși că o abordare din punctul de vedere al inteligenței artificiale a disciplinelor clasice ar produce un nou mod de a înțelege umanitatea, precum și societatea complet schimbată dintr-un viitor nu atât de îndepărtat. Între

timp, această carte vizează cu precădere anumite aspecte formale ale inteligenței artificiale: algoritmi de învățare, implementarea inteligenței la mașini și conviețuirea oamenilor cu mașinile inteligente. Din păcate, chiar și cu o astfel de abordare, nu va fi loc în această carte pentru o mulțime de chestiuni interesante, pe care sperăm că cititorii vor ști să le completeze din numeroasele surse menționate ca referințe sau ca lecturi recomandate.

Tratarea unor aspecte atât de diverse ale aceluiași subiect provoacă o senzație de amețeală, care trebuie ținută tot timpul sub control. Poate călătoria în universul inteligenței artificiale este modul cel mai potrivit de a experimenta această senzație. Sperăm în orice caz că pentru cititori acest volum va fi un stimul pentru a găsi texte cu un conținut mai detaliat și mai riguros. Cel puțin, ne dorim ca această lectură să trezească emoția intensă a învățării prin joc. Dacă așa va fi, va fi meritat osteneala.

Capitolul 1

Introducere

*Calculatorul poate face doar ce este programat să facă,
deci Inteligența artificială nu este posibilă.*
(Ada Lovelace, primul programator)

Calculatoarele? Cu siguranță sunt inutile, sunt capabile doar să dea răspunsuri.
(Picasso)

În acest capitol vom pune bazele pentru înțelegerea conceptului de „inteligență artificială”. Prezentarea va trebui să fie ușor ambiguă, deoarece nu există un consens larg asupra definiției acestei discipline. Mai mult, în acest caz, lipsa definiției se resimte cu atât mai mult, cu cât numele în sine reunește două dintre conceptele cele mai eluzive din istoria științei, cel de „inteligență” și cel de „artificialitate”.

Singurul lucru sigur este că domeniul este într-o expansiune continuă și că își modifică permanent propriile obiective și metodologii de analiză. Multe cercetări în alte ramuri ale științei și ingineriei converg spre inteligența artificială, fapt ce impulsionează extraordinar dezvoltarea acesteia, în special în ceea ce privește calculul și programarea algoritmilor de învățare folosiți la *crearea* inteligenței artificiale. Astfel, matematica oferă contextul adecvat pentru înțelegerea logicii învățării, în timp ce fizica pune la dispoziție cadrul pentru implementarea fizică a mașinilor, fie că este vorba despre calculatoare sau de dispozitive capabile să efectueze sarcini specifice, ca, de pildă, roboții.⁸ Pe de altă parte, biologia constituie cadrul principal de

⁸ Termenul *robot* a fost folosit pentru prima dată de către scriitorul cehoslovac Karel Čapek (1890-1938), într-una dintre operele sale, pentru a descrie niște muncitori artificiali.

referință pentru inteligența artificială. Evoluția a generat sisteme biologice capabile să întreprindă acțiuni inteligente nu numai în cazul oamenilor, dar și în cel al mării majorități a speciilor de animale. Încă de la început, realitatea biologică a fost referința și izvorul de inspirație pentru inteligența artificială. Filosofia este un alt domeniu important pentru inteligența artificială, mai ales în ceea ce privește crearea unor ființe noi, dotate cu autonomie, dar care populează un univers antropocentric și care pot fi supuse unor reguli de conviețuire diferite de regulile uzuale ale eticii. Într-un astfel de context, devine extrem de interesantă poziția teologiei cu privire la pătrunderea inteligenței artificiale în lumea noastră. Sociologia este știința dedicată studiului relațiilor umane și, ca atare, în momentul în care în aceste relații intră și alte inteligențe autonome, ar trebui să studieze în mod aprofundat noile reguli de conviețuire, și să le redefinească adecvat pentru noua societate. În sfârșit, este de datoria politiciii să propună legi care să reglementeze normele de conviețuire civilă în societăți caracterizate de diverse tipuri de inteligență.

Mare parte a aplicațiilor inteligenței artificiale vizează rezolvarea de probleme (în engleză *problem solving*), chiar dacă nu chiar în sensul dat de Pablo Picasso, în fraza citată în mottoul capitolului, ci într-unul mult mai proaspăt și inovator, pe care cu siguranță pictorul l-ar fi susținut. Mare parte dintre autori semnalează aplicații specifice inteligenței artificiale în următoarele domenii: creativitatea artistică, medicina sau securitatea cibernetică, precum și sarcini ca șofatul automat, recunoașterea automată a limbii, traducerea automată sau implementarea roboților de companie.

Stăpânirea tehnicilor pentru rezolvarea problemelor precum cele menționate mai sus presupune ani de studiu și de specializare. Nu ar fi realist să pretindem că vom intra în toate detaliile acestor probleme într-un text precum cel de față. Și nu ar fi corect nici ca acest text să fie citit cu intenția de a învăța printr-o astfel abordare generală să rezolvăm probleme.

Această lectură introductivă ar trebui mai degrabă să folosească la orientarea interesului cititorului spre probleme despre care poate a mai auzit sau spre unele extrem de recente care își fac loc fulgerător pe scena inteligenței artificiale. În secțiunea următoare, vom trece în revistă pe scurt originile acestei discipline, pentru ca apoi să putem pune bazele analizei conceptului de inteligență artificială.

Nașterea unei discipline

Orice disciplină științifică traversează o perioadă de maturizare. În anumite cazuri, este nevoie de mult timp pentru a ajunge la un nivel corespunzător pentru a putea fi numită cu adevărat o știință. Și în aceste cazuri, parcursul normal al domeniului presupune debateri continue asupra temelor tratate în cadrul disciplinei, asupra modalităților de dezvoltare și asupra succeselor obținute. Istoria inteligenței artificiale constituie un bun exemplu de traseu de dezvoltare, traversând acum un moment crucial pentru afirmarea sa ca știință.

Sub anumite aspecte, acest nume, menit să descrie această disciplină științifică – poate, și nu se poate nega acest lucru, un pic pompos –, care studiază modul în care o mașină poate fi înzestrată cu inteligență, implică ambiția oamenilor de a înțelege nu numai inteligența umană, ci și de a-și depăși propria origine și dezvoltare biologică. Încă înainte ca știința modernă să domine în departamentele universitare, au existat numeroase încercări de abordare a inteligenței artificiale, lucru extrem de uman și, în același timp, divin. Conform acestei linii de gândire, noi, oamenii, fiind dotați cu inteligență, suntem capabili să fim conștienți de noi înșine și de existența lui Dumnezeu și să ținem cont de faptul că am fost creați după chipul și asemănarea Sa. Acest mod de a gândi sfida totuși filosofia și știința secolului al XVII-lea, moment în care s-a întrevăzut posibilitatea de a crea inteligență artificială prin intermediul unor ființe artificiale.

Lista numelor care ar trebui citate ca având contribuții la dezvoltarea inteligenței artificiale este lungă. Cititorii interesați pot consulta