Inteligenţa Artificială

Raluca Vasilescu

1996

Subiect: Un scurt istoric al IA

Cuvinte cheie: inteligenţă, test Turing, sisteme expert

Cunoştinţe necesare: nici una

Cuprins

1 O definiţie?

Multă lume a încercat să răspundă la întrebarea: ce este inteligenţa artificială? Dificultăţile găsirii unei definiţii pentru acest termen sunt (in principal) două: întîi, că nu se prea ştie ce este însăşi inteligenţa naturală; apoi, că cei ce încearcă sa formuleze o definiţie sunt complexaţi de realizările – departe de a justifica un nume aşa pompos – ale acestui domeniu al ştiinţei calculatoarelor. Aşa că ei se văd siliţi să definească mai degrabă ce ar trebui să fie inteligenţa artificială, decît ceea ce ea este efectiv in prezent.

Un mic istoric al acestei discipline s-ar putea să lumineze mai bine asupra naturii şi ţelurilor ei, decît încercarea (mai mult sau mai puţin şchioapă) de a o defini precis.

Articolul de faţă vorbeşte pe scurt despre cîteva din sistemele celebre, dar nici pe departe singurele, construite înlăuntrul inteligenţei artificiale.

2 Testul Turing

E clar că, dacă n-ar fi existat calculatoarele, atunci nu s-ar fi povestit astăzi despre inteligenţa artificială. Sau s-ar fi povestit in romane SF. Deci, originea inteligenţei artificiale ca ramură a informaticii se ascunde cam prin anii construcţiei primelor calculatoare electronice, adică în momentul în care omul şi-a pus problema: cît de puternice pot fi aceste unelte noi, capabile să efectueze calcule complicate? Pot fi ele făcute să gîndească?

Pornind de la întrebarea aceasta (cum spunea cineva, important e să ştii ce întrebare să pui), şi încercînd să o reformuleze în termeni oarecum obiectivi, direct constatabili, matematicianul englez **Alan Turing** a imaginat testul care îi poartă numele şi care, odată satisfăcut de o maşină, ar dovedi "inteligenţa" ei.

Turing a pornit de la o idee foarte naturală – dacă nu ştim să definim în termeni precişi inteligenţa, însă spunem despre om că este inteligent, atunci am putea să spunem şi despre o altă creatură acelaşi lucru în cazul în care s-ar comporta la fel ca o fiinţa umană. Rămîne de văzut care aspecte ale comportamentului omenesc sunt într-adevăr relevante pentru inteligenţa. De exemplu, hrănirea sau reproducerea s-ar putea să nu fie.

Testul Turing porneşte de la un joc, creat tot de Turing – "jocul imitaţiei" – cu trei jucători: o maşină (A), un om (B) şi un al doilea om (C). A şi B nu se găsesc în aceeaşi cameră cu C. C nu ştie care dintre ceilalţi doi jucători este maşina şi nu poate să îi vadă sau să le vorbească direct. Comunicarea se poate face în scris sau printr-un terminal. Scopul lui C este să deosebească maşina de om, pe baza răspunsurilor la orice fel de întrebări. Dacă C nu reuşeşte, atunci maşina poate fi considerată inteligentă.

În 1950, cînd apărea articolul despre acest test, autorul prezicea că, în cinzeci de ani (deci, aproximativ in anul 2000), va fi posibil să existe un calculator capabil să joace jocul imitaţiei atît de bine încît şansa lui C de a identifica corect omul să fie mai mică de 70% după cinci minute de joc. Informaticienii de astăzi sunt mult mai puţin optimişti. De fapt, există chiar două tabere – unii care cred în posibilitatea de a satisface (cîndva) testul lui Turing, şi alţii care sunt convinşi de contrariu.

Ca să încheiem această secţiune, trebuie să spunem că, deşi a suscitat multe discuţii, testul Turing simbolizează idealul pe termen lung al inteligenţei artificiale ca ramură a informaticii. Turing considera că cel mai bun drum către realizarea unei maşini care să treacă testul său nu este programarea unui calculator dotat cu o mulţime fixă de cunoştinţe, ci, mai degrabă, educarea unei maşini-copil, capabilă să înveţe din experienţă şi să folosească limbajul natural ca să-şi îmbogăţească cunoştinţele. Ea ar putea să-şi rezolve problemele proprii şi să-şi împlinească propriile sale planuri, dînd dovadă de inteligenţă practică în viaţa de zi cu zi. Vom vedea că, de fapt, fiecare dintre aceste trăsături ideale s-au constituit în subdomenii ale inteligenţei artificiale: învăţarea automată, prelucrarea şi înţelegerea limbajului natural, achiziţia de cunoştinţe, construirea şi satisfacerea unor planuri.

3 Prima perioadă – jocurile şi demonstrarea de teoreme

Începuturile inteligenţei artificiale pot fi văzute imediat după al doilea război mondial, în primele programe care rezolvau puzzle-uri sau care jucau anumite jocuri. Au existat două motive (pe lîngă fascinaţia pe care o exercită asupra multora) pentru care jocurile au fost printre primele domenii de aplicare a inteligenţei artificiale – întîi, că performanţa programului este uşor de măsurat (de cele mai multe ori, sau cîştigi sau pierzi un joc); apoi, că regulile sunt ,în general, simple şi puţine la număr, deci pot fi uşor descrise şi folosite.

Jocurile cu care s-a experimentat îndeobşte au fost cele de şah şi de dame. Ideea era foarte simplă – fiind dată o poziţie pe tablă, se încerca să se genereze toate secvenţele posibile de mutări de la acel moment încolo, considerînd că adversarul alege întotdeauna mutarea cea mai bună. Dacă o secvenţă ajungea într-o stare cîştigătoare, atunci ea era cea de urmat. Problema practică de care s-a lovit această idee a fost că numărul combinaţiilor de explorat era foarte mare (de exemplu, la şah, de ordinul 35^{100}). Bineînţeles că oamenii, cînd sunt puşi în situaţia să joace, restrîng numărul combinaţiilor posibil cîştigătoare folosindu-se de experienţa de pînă atunci (de exemplu, nu mai încearcă, de cele mai multe ori, să calculeze cum ar putea cîştiga dacă, fiind la a treia mutare, ar ceda regina pe gratis, ci elimină de la bun început varianta cu pricina). De-aici, concluzia imediată care s-a tras din această perioadă – că pînă şi cînd este vorba de jocuri, e nevoie ca programul aibă cunoştinţe adiţionale (altele decît regulile jocului).

Unul din programele de referinţa din această perioadă a fost programul de jucat dame al lui Samuels. Acest program, pe lîngă faptul că juca cu un adversar, îşi folosea experienţa dobîndită în partidele anterioare ca să-şi îmbunătăţească performanţele. El ţinea minte anumite poziţii ca din start cîstigătoare sau dezastruoase şi nu le mai calcula secvenţele ulterioare de mutări (deci se comporta aproximativ ca omul din exemplul de mai sus, care nu-şi mai pune problema să cedeze regina).

Celălalt domeniu care a suscitat interes în această primă perioadă a fost demonstrarea de teoreme. Acest domeniu se aseamănă cu cel al jocurilor prin faptul că performanţele sunt simplu de evaluat (supui spre demonstrare o teoremă clasică). Ceea ce trebuie să i se descrie calculatorului sunt setul de axiome şi regulile de inferenţă (adică, regulile prin care se obţin noi adevăruri din nişte adevăruri date). În această arie au fost create mai multe programe interesante, printre care "The Logic Theorist" al lui Newell, care demonstra teoreme din primul capitol al cărţii "Principia mathematica" de Whitehead şi Russell, şi un program al lui Gelenter care demonstra teoreme de geometrie.

Trebuie spus că in această primă perioadă (care a ţinut cam pînă în 1965) performanţele pe care le-a obţinut inteligenţa artificială n-au fost ameţitoare, în primul rînd pentru că nici o problemă cu adevarat semnificativă (adică a cărei soluţie să fi ajutat substanţial vreo activitate umană) nu a fost rezolvată. Totuşi, două concluzii s-au impus (care concluzii guvernează pîna astăzi disciplina de care ne ocupăm):

- majoritatea problemelor poate fi redusă la o problemă de căutare. Pentru cine nu ştie, o problemă de căutare seamănă cu următoarea situaţie: X vrea să ajungă în oraşul A, şi se află la o răscruce fără indicatoare; ca să ajungă în A, X o ia pe fiecare din drumurile de la răscruce; dacă oraşul de la capătul unui drum nu e A, atunci X se întoarce înapoi şi porneşte pe un alt drum.

- căutarea trebuie să fie călăuzită de anumite cunoştinţe despre domeniul problemei. Păstrîndu-ne în cadrul aceluiaşi exemplu, am putea spune ca X ştie că nici un drum care urcă nu duce la A. Atunci el ar putea evita de la bun început drumurile ascendente care pornesc din răscrucea cu pricina, limitîndu-şi căutarea numai la drumurile descendente.

4 A doua perioadă – înţelegerea limbajului natural

Anii 1965-1975 formează cea de a doua perioadă a inteligenţei artificiale. Acum lumea se preocupă de "înţelegere", adică vrea să facă maşina să inţeleagă limbajul natural, în special povestiri şi dialoguri.

Un program faimos al timpului a fost ELIZA. ELIZA simula comportamentul unui psiholog, conversînd în engleză cu pacienţii. Cunoştinţele programului despre engleză ca şi despre psihologie erau codificate sub forma unui set de reguli simple. Iată un mic dialog cu ELIZA (tradus în româneşte):

Pacient: Toţi bărbaţii sunt la fel.

ELIZA: In ce fel?

Pacient: Ne plictisesc întotdeauna cu cîte ceva.

ELIZA: Poţi să dai un exemplu?

Pacient: Păi, prietenul meu m-a trimis aici.

ELIZA: Prietenul tău te-a trimis aici.

Pacient: Spune că sunt mai tot timpul deprimată.

ELIZA: Îmi pare rău să aud că eşti deprimată.

Pacient: E adevărat. Sunt nefericită.

ELIZA: Crezi că venind aici ai putea fi ajutată să nu mai fi nefericită?

Pacient: Am nevoie de ajutor, măcar asta e clar.

ELIZA: Ce-ar însemna să primeşti ajutor?

Pacient: Poate aş învăţa să mă împac cu mama.

ELIZA: Povesteşte-mi despre familia ta.

ELIZA ştia un set mic de cuvinte cheie, şi avea una sau mai multe reguli pentru fiecare dintre ele. Cînd întîlnea un astfel de cuvînt în propoziţiile pacientului său, folosea una din regulile corespunzătoare cuvîntului respectiv. De exemplu, oricărei propoziţii care conţinea cuvintele "mamă", "tată" etc., i se răspundea cu "Povesteşte-mi despre familia ta." Procedeul ELIZEI se numeşte "nu-nţelege dar le potriveşte" ("pattern matching" în engleză).

Un amănunt mai special despre ELIZA este impactul pe care ea l-a avut asupra pacienţilor săi umani. Autorul ei, Weizenbaum, a fost uimit să constate "cît de rapid şi de puternic oamenii au ajuns să se implice emoţional în comunicarea cu calculatorul şi cît de mult şi-l imaginau ca pe o fiinţa umană". Însăşi secretara sa, care îl văzuse lucrînd la program, i-a cerut să plece din cameră în timp ce discuta cu maşina.

Un exemplu mai neobişnuit de program a fost PARRY. PARRY simula un bolnav de paranoia, iar motivul pentru care a fost construit era să testeze un model psihologic al paranoicului. Cînd mai mulţi medici au fost invitaţi să-l consulte, în jumătate din cazuri PARRY a fost recunoscut ca un pacient autentic.

Alt program interesant este SHRDLU, considerat de bună seamă una din culmile epocii. El era în stare să priceapă comenzi date in engleză. Aceste comenzi erau folosite pentru a modifica o lume de cubuleţe, asemănătoare celor pe care copii le folosesc la joacă. De asemenea, SHRDLU răspundea unor întrebări legate de configuraţia blocurilor (de tipul "Ce culoare are blocul de sub piramida roşie?"). Mai mult, SHRDLU putea să construiască planuri ca să îndeplinească comenzi de tipul "Pune piramida albastră peste blocul verde". Imaginaţi-vă că peste blocul verde mai erau aşezate alte două blocuri. În cazul acesta, programul le îndepărta, ca să facă loc piramidei albastre!

Din păcate, lumea blocurilor era prea simplă, şi ca atare SHRDLU n-a adus prea mare folos imediat în rezolvarea vreunei probleme concrete.

Un program mai ambiţios s-a numit MYCIN. El îşi propunea să diagnosticheze bolile infecţioase de sînge şi, de asemenea, să recomande tratamente. Cum cei mai mulţi oameni nu pot face acest lucru, neavînd cunoştinţele necesare, un astfel de sistem s-ar fi putut dovedi mult mai util decît celelalte programe de pînă atunci. MYCIN se folosea cunoştinţe prealabile care înglobau cunoştinţele de diagnoză pe care un medic le-ar fi putut poseda. Aceste cunoştinţe erau exprimate sub forma unor reguli de tipul: "Dacă temperatura corpului este 38, atunci boala este meningită" (exemplul este simplificat, şi fără pretenţii de exactitate ştiinţifică).

Într-un test care compară analiza făcută de MYCIN unui număr de cazuri cu cea a unor medici de diferite nivele de calificare şi experienţă, judecătorii au considerat ca preferabilă sau echivalentă soluţia dată de MYCIN faţa de cea a adevăraţilor medici.

Deşi MYCIN n-a fost niciodată folosit efectiv, el arată că probleme care pînă atunci fuseseră lăsate exclusiv în seama experţilor umani pot fi rezolvate de maşină. De altfel, MYCIN a deschis calea unei lungi serii de sisteme expert, adică sisteme care se comportă ca nişte experţi umani în domenii foarte limitate. Ele sunt capabile să treacă, de cele mai multe ori, aşa-numitele teste Turing restrînse, cu întrebări exclusiv din aria pe care se presupune că ar trebui să o stăpînească.

5 A treia perioadă – sisteme expert evoluate

Această perioadă se întinde cam din 1975 pînă în zilele noastre. Se poate spune că inteligenţa artificială a devenit mai lucidă, mai critică cu privire la ea însăşi, şi într-o anumită măsură, mai pragmatică. Entuziasmele cu iz psihologic legate de înţelegere s-au mai temperat, şi, în acelaşi timp, au apărut primele sisteme expert eficiente şi cu folosire rentabilă în industrie. Interesul cade mai mult pe o reprezentare compactă şi uniformă a cunoştinţelor; lumea începe să se îndoiască de oportunitatea unor metode generale de rezolvare a problemelor.

Prin anii 80 a fost creat unul din primele sisteme expert folosite în industrie – R1 (acum numit XCON), construit la Universitatea Carnegie Mellon, S.U.A cu colaborarea DEC (Digital Equipment Corporation). R1 se ocupă de configurarea unor sisteme de calcul (orice calculator VAX fabricat de DEC e configurat cu R1). El a fost pus în producţie, şi de atunci interesul lumii pentru inteligenţa artificială a crescut considerabil. Ca urmare, informaticienii din această sfera s-au împărţit în "implementatori" de tehnici de inteligenţa artificială şi în cercetători.

Tot acum s-au făcut unele progrese în domeniul învăţării automate. De pildă, programul AM (Automated Mathematician – Matematicianul Automatizat), proiectat să descopere legi matematice, reuşeşte să inducă concepte precum cardinalitatea şi aritmetica întregilor, avînd drept cunoştinţe iniţiale conceptele şi axiomele teoriei mulţimilor. AM este selectiv în sensul că reţine numai cunoştinţe "interesante"; de asemenea, el îşi modifică gradual cunoştinţele.

6 Concluzii

Inteligenţa artificială îşi propune la modul ideal să afle cum pot fi maşinile să devină la fel de inteligente ca oamenii. O măsură a inteligenţei maşinii este testul Turing. In imediatul viitor nu se conturează posibilitatea ca vreun calculator să treacă acest test.

Istoria inteligenţei artificiale vorbeşte despre realizările ei efective. Multă vreme, această disciplină s-a păstrat în zona experimentelor. De curînd, cercetările ei au început să fie privite cu interes de către industrie, graţie mai ales sistemelor expert.