
ETICA INFORMAȚIONALĂ ȘI INTELIGENȚA MORALĂ**Laura PANĂ**

Abstract. The setting up of an artificial technical environment and its transition to an intellectual technical environment by the transformation of the information machine into the intelligent machine generates a new theoretical field of research, that of the information ethics, which includes at least four new ethical domains. A specific form of human and artificial intelligence is needed for successful practice of a moral code. The integrated analysis of the intelligence as of the morality, considered as systems with internal structure, is undertaken. An efficient moral conduct, common for humans and machines, is proposed by philosophical foundation, scientific description and practical implementation of an artificial ethics, which is not an application of human morality, but an outcome of moral invention. The future of human and artificial morality is studied.

Semnificațiile etice ale conceperii și folosirii agenților artificiali inteligenți sau problemele morale implicate de consecințele funcționării sistemelor artificiale inteligente în general erau neglijate până recent: se considera că responsabilitatea pentru efectele acțiunii unor astfel de sisteme revine în mod automat utilizatorului, creatorului sau proprietarului. În condițiile actuale problematica etică se pune cu o nouă acuitate, dar nu numai în raport cu mașina considerată inteligentă, așa cum ne-am aștepta, ci și în raport cu omul.

Proiectarea și producerea mașinilor dotate cu programe de inteligență artificială, ca și întreg ansamblul de activități presupuse și permise de aceste produse tehnice generează dezvoltarea unui nou mediu tehnic al cunoașterii și acțiunii, un mediu tehnic intelectual artificial. Apare, în același context, tendința asistării sau înlocuirii factorului uman cu agenți artificiali în efectuarea și conducerea proceselor de producție, servicii, comunicații, cercetare, educație, afaceri, ca și de comerț sau divertisment.

Entitățile dotate cu inteligență artificială îndeplinesc, astfel, tot mai multe funcții în societate și, implicit, în raport cu omul. Aceste entități prezintă un comportament complex și vor fi tot mai autonome. Mai mult, ele vor avea un comportament din ce în ce mai imprevizibil, întrucât

acțiunile unei astfel de mașini vor depinde nu numai de decizii umane, ci și de decizii proprii, ca și de decizii luate de alte mașini.

Nu numai mașina, ci și omul, ca și organizarea sa socială devin mai complexe în cultura actuală. Specia umană evoluează în toate dimensiunile sale: biotică, psihică, socială și culturală; acum ea însăși evoluează către artificialitate.

Chiar ca ființă biologică omul evoluează spre artificialitate, pentru că în procesul natural de îmbătrânire și de involuție ca urmare a unor boli sau accidente noi consumăm și primim din ce în ce mai multe componente artificiale și devenim tot mai mult roboți, prin combinația dintre țesutul biotic și tehnic, prin inventarea și acceptarea mușchilor artificiali sau a sângelui artificial; se adaugă utilizarea în diverse scopuri a unor rudimente de sistem nervos artificial sau proiectarea sistemelor inteligente mixte.

Pe de altă parte, sistemele inteligente artificiale primesc proprietăți suprasenzoriale (ca vederea în infraroșu sau detectarea ultrasunetelor) și o formă eficientă de inteligență abstractă. Așa cum apreciază unii specialiști în știința calculatoarelor, aceste sisteme se vor putea dezvolta prin evoluție „naturală” simulată, ca și prin procese inspirate din cultivarea grădinii sau coacerea pâinii; inginerii vor crea condițiile posibilității unor astfel de sisteme, care se vor autostructura în funcție de condițiile specificate și cerințele formulate.

Omul și mașina interacționează inclusiv în medii cognitive, urmărind scopuri partajate și folosind tehnici comune, în activități tot mai complexe de căutare și de structurare a cunoștințelor, de generare colaborativă a cunoștințelor, de comunicare și diseminare a acestora, în activități de e-learning, de e-work, e-trade (inclusiv knowledge trade) ș. a. m. d. Așa-numiții *webbots*, care pot fi și *knowbots*, fac parte acum din comunitatea virtuală care este, astfel, o comunitate mixtă.

Prin astfel de activități se conturează o posibilă evoluție comună a omului și a mașinii inteligente. În acest proces, dezvoltarea inteligenței artificiale poate fi considerată ca o etapă în evoluția omului. Putem anticipa, în același proces, conturarea a două tendințe convergente: pe de o parte, adecvarea instrumentelor intelectuale artificiale pentru cerințele specifice activităților umane de cunoaștere și, pe de altă parte, adaptarea omului pentru folosirea noilor tehnici de cunoaștere în echipe mixte de inteligență umană și artificială. Se poate prevedea chiar adoptarea de către om a unor noi tehnici vitale în cadrul unui mediu tehnic tot mai artificial. Omul și mașina se vor întâlni, astfel, la mijlocul drumului dintre natural și artificial.

Interacțiunea dintre om și mașină într-o lume tot mai artificială, în condițiile creșterii complexității omului și a mașinii, asociate cu sporirea

accelerată a eficienței mașinii, determină atât necesitatea dotării mașinilor considerate inteligente cu un cod moral adecvat cooperării și coexistenței cu omul, cât și modificarea eticii umane.

1. PREMISELE PENTRU FORMULAREA ETICII INFORMAȚIONALE

În cadrul orientării generate de actuala ”turnură computerială” a filosofiei s-a conturat și o arie de construcție teoretică și practică dedicată eticii. În acest cadru se constituie, astăzi, chiar mai multe domenii de preocupare. Ne vom referi la câteva dintre cele a căror apariție și evoluție este observabilă în volumele de specialitate și în desfășurarea manifestărilor științifice internaționale.

Etica folosirii calculatorului este destinată utilizatorului mașinii informaționale*. Acest domeniu al eticii privește și activitatea tuturor celor care folosesc sau generează informație în mediul virtual al webului. În domeniul menționat al eticii sunt tratate probleme caracteristice, de la cele vizând proprietatea software, la cele referitoare la asigurarea identității și intimității utilizatorului și până la cele ce privesc transmiterea și păstrarea așa-numitei *netiquette*, care constituie adaptarea etichetei la statutul de „netizen”, adică cel de cetățean al spațiului și timpului virtual al internetului.

Etica pentru calculator („machine ethics”) privește calculatorul însuși, ca urmare a faptului că se are în vedere capacitatea mașinii de a induce schimbări în lume, în mod asemănător omului. Orice activitate umană are o componentă morală, iar mașina cu posibilități asemănătoare trebuie asigurată cu funcții de tip moral. Problematika menționată constituie obiectul unei noi direcții actuale de cercetare, în care conlucrează generaliști din domeniul filosofiei și specialiști din domeniul inteligenței artificiale. Considerațiile noastre se vor referi, în mare măsură, la acest nou domeniu al eticii.

Etica de calculator se referă la preocupările filosofice de ordin etic care se bazează pe folosirea calculatorului. Calculatorul este utilizat, în acest caz, atât în formularea și soluționarea unor problemele teoretice generale, cât și în tratarea unor aspectele specifice privind utilizarea calculatorului în domenii filosofice determinate, cum este și domeniul eticii. În acest din urmă caz, se aplică achizițiile teoriei și metateoriei etice, cu ajutorul tehnicii informatice.

Etica informațională globală este dezvoltată ca urmare a posibilităților și problemelor generate de globalizarea cunoașterii,

comunicării și prelucrării cunoștințelor și este abordată în cărțile, revistele și reuniunile filosofice actuale. Una dintre concluziile diferitelor tipuri de lucrări amintite este că etica informațională globală reprezintă chiar ansamblul domeniilor până acum prezentate, dezvoltate pe fondul interesului general trezit de posibilitățile oferite de folosirea tehnicii informatice.

Există și opinii conform cărora etica folosirii calculatorului și a mijloacelor informatice și de comunicare asociate ar fi numai o ramură a eticii profesionale, care se ocupă de codurile de conduită ale profesioniștilor

* Acest termen a fost propus de noi în anul 2000, pentru a desemna calculatorul, oportunitatea folosirii termenului fiind argumentată într-un subcapitol special al cărții „Filosofia culturii tehnice” (15, 430-438).

în calculatoare. Dar acestea din urmă sunt folosite astăzi în toate profesiile. De asemenea, așa cum evidențiază unii dintre filosofii consacrați prin reflecțiile asupra acestui domeniu (2, 285), în etica informațională globală, dezvoltată în continuarea celei anterioare, valorile morale și chiar cele umane în general, ca și standardele de conduită derivate, devin obiectul dezbaterii, consensului și reglementării, dincolo de constrângeri geografice, sociale, politice și chiar de diferențe culturale.

În cadrul a ceea ce am numit etica de calculator se adoptă și se dezvoltă noi metode și mijloace de predare – asimilare ale eticii, alături de alte discipline socio-umane, folosindu-se avantajele apărute prin dezvoltarea învățământului telematic, distribuit, flexibil și bazat pe multimedia, dirijat și asistat prin tehnici ale inteligenței artificiale bazate pe cunoaștere. Se folosesc, în acest domeniu, în continuare, metodele abstracte consacrate, acestea devenind însă mai puternice prin conceperea și folosirea unor programe adecvate. Se modelează și se simulează, astfel, consecințele aplicării teoriilor etice consacrate, descoperindu-se inclusiv slăbiciunile lor și facilitându-se înțelegerea lor. Sunt anticipate efectele morale ale unor decizii sociale și devine posibilă întemeierea analizei deciziilor în probleme morale dificile. Se elaborează proceduri valide de decizie morală pentru diverse domenii de activitate umană.

Ne vom opri, în continuare, asupra unuia dintre domeniile de reflecție filosofică constituite prin tendințele teoretice și practice actuale menționate, și anume, asupra a ceea ce, în literatura de specialitate de limbă engleză se denuște prin *etica mașinii*. În ceea ce ne privește, considerăm că o denumire poate chiar mai adecvată pentru acest domeniu de cercetare

ar fi aceea de *etică artificială*. Acest termen ar avea două grupuri de avantaje.

Primul avantaj este chiar de ordin terminologic și este dat de faptul că

- această denumire trimite automat la domeniul de cercetare de care studiul este în mod efectiv legat, acela al inteligenței artificiale
- este astfel sugerat caracterul interdisciplinar al studiului, ca și identitatea științifică a cercetătorilor implicați, realitatea că acest domeniu este comun filosofilor și specialiștilor în știința calculatoarelor: tipul de pregătire al participanților la reuniunile științifice internaționale dedicate domeniului confirmă și prin observație directă cele afirmate;
- se evită, în mai mare parte, reacțiile subiective subliminale care ar putea apărea pornind de la anumite îngrijorări existente cu privire nu numai la tendințe reale, ci și la presupuse intenții de artificializare a ființării umane, iată, acum, și sub aspectele sale morale.

Cel de-al doilea avantaj este de natură teoretică și constă în creșterea coerenței viziunii dezvoltate de noi (16, 249-323) cu privire la evoluțiile actuale din cadrul filosofiei, viziune generată de faptul că se conturează, astăzi, în cadrul culturii spirituale, *filosofia artificială*. În acest context pare firesc să vorbim și de domenii sau discipline filosofice constitutive noii problematici și noilor forme de expresie filosofică, cum ar fi *etica artificială*, *estetica informațională* sau *filosofia politicii digitale*. Aceste domenii de reflecție filosofică au fost deja tratate de noi în lucrări recente (11) și (14), dar au fost precedate de inițierea de către o serie de autori a unor discipline cum ar fi *axiologia formală* sau *estetica abstractă*, fiind legate și de evoluția de la etica filosofică la *etica științifică*.

În lucrarea noastră vom folosi pentru domeniul nemijlocit abordat denumirea de *etica mașinii*, în pofida precizărilor terminologice care aduc și clarificări teoretice, având în vedere faptul că este vorba despre un domeniu de cercetare în curs de constituire, care are nevoie de un timp de impunere, convergența opiniilor și consonanța termenilor putând reprezenta condiții favorizante.

Ca denumire generică pentru ansamblul preocupărilor etice care folosesc tehnici și tehnologii informatice, întrupate în primul rând în calculator (pentru care vom folosi, cum am arătat, termenul de *mașină informațională*), vom apela, în continuare, la denumirea de *etică informațională*.

2. Asimetria posibilități – dificultăți

în elaborarea eticii mașinii

Posibilitățile practice din domeniul eticii mașinii sunt date de facilitățile oferite de folosirea tehnicii de calcul, iar posibilitățile teoretice sunt întemeiate pe existența unui mare număr de teorii etice; ansamblul acestor posibilități este condiționat de capacitatea folosirii specifice a tehnicii de calcul, dar mai ales de abilitatea depășirii dificultăților aplicării teoriilor etice, a normelor morale și a regulilor concrete, construite în jurul valorilor morale. Dificultățile par să depășească, deocamdată, posibilitățile, deși au fost indicate câteva direcții de studiu fertile.

Dificultățile întâmpinate în formularea eticii mașinii ne apar, în urma studierii domeniului, ca fiind dificultăți conceptuale, culturale și tehnice.

Dificultățile conceptuale, la rândul lor, pot fi date de

- a. complexitatea realității și a descrierii sale conceptuale sub raport etic
- b. caracteristicile procesului cunoașterii teoriilor etice, ca teme atât al practicii morale, cât și al elaborării codului moral al mașinii
- c. problemele aplicării uneia dintre teoriile etice.

Fără a mai discuta problemele teoretice legate de descrierea și explicarea complexității, nici măcar la nivelul societății, vom menționa numai că apariția și dezvoltarea mașinii inteligente este tocmai unul dintre răspunsurile posibile la problema complexității. Vom insista însă asupra unor moduri adecvate de înțelegere și folosire a tehnicii ca activitate umană complexă, ca și asupra complexității valorilor și a omului, ca rezultat al folosirii și al inventării valorilor.

Complexitatea tehnicii este dată, între altele, de faptul că dezvoltarea tehnicii este rezultatul evoluției tuturor domeniilor culturii. Realizările actuale ale tehnicii pot fi neașteptate, dar la fel ca orice alt domeniu al culturii, tehnica este produsul unei sinteze culturale. Nivelul de dezvoltare și tendințele de evoluție ale fiecărui tip de fenomen cultural sunt condiționate de ansamblul dezvoltării culturii. Ceea ce variază sensibil și face diferența între domeniile culturii ține, probabil, de ritmul propriu de dezvoltare și de impactul social al schimbărilor din cadrul lor, de formele de manifestare ale socializării lor.

Ca ilustrare a dependenței reciproce a progreselor din domenii diferite ale culturii putem evidenția condițiile dezvoltării uneia dintre direcțiile de evoluție ale tehnicii actuale, și anume, dezvoltarea ingineriei genetice: descifrarea codului genetic a necesitat cooperarea biologilor cu matematicieni, fizicieni, chimiști și folosirea în echipă a calculatorului. În

același proces complex au fost necesare descoperiri teoretice, experimente practice, instrumente și aparate special create, invenții tehnice.

Invenția însăși nu trebuie înțeleasă numai ca invenție tehnică, iar dacă vorbim de această din urmă formă a invenției, putem să o caracterizăm ca activitate complexă. Din acest punct de vedere, invenția tehnică însăși poate fi definită ca ansamblu al studiilor, activităților constructive, al reflecțiilor și al produselor caracterizate prin noutate și utilitate. În acest context, prin studii înțelegem cercetări nu numai practice, ci și teoretice, în procesele de înnoire tehnică fiind valabilă aprecierea că tehnica se organizează astăzi în jurul științei, chiar în jurul nivelului teoretic al acesteia.

În ceea ce privește una dintre cele mai importante direcții actuale de dezvoltare ale tehnicii, se consideră că mașinile cele mai complexe, calculatoarele, vor evolua pe căi multiple; se au în vedere, în principal, pe lângă direcțiile mai des frecventate în prezent, calculatorul optic, cuantic și biotic. În aceste condiții devine clar că, pentru realizarea unor astfel de variante de evoluție, va fi necesară cercetarea interdisciplinară consacrată corelării teoretice, metodologice și valorificării practice a progreselor din domeniile implicate.

Constructorii de calculatoare vizionari afirmă chiar că viitorul calculatoarelor nu va fi decis pe cale tehnică: tehnologul va crea mai degrabă condițiile autostructurării sistemelor informatice evolute, după modelul evoluției naturale (prin selecție automată a programelor performante) sau după cel al activităților umane primordiale, cum sunt cultivarea grădinii ori fabricarea pâinii. În primul caz, sunt folosite metode de programare care simulează evoluția naturală prin evenimente în cadrul unor populații de programe, care suportă mutații, se integrează în ecosisteme tot mai complexe, păstrându-se variantele cele mai adecvate (1, 148). În cel de-al doilea caz se are în vedere o evoluție culturală a mașinilor, prin reproducerea proceselor de cultivare a aptitudinilor în educația copiilor (7, 187).

Complexitatea și procesualitatea valorilor morale constituie alte aspecte ale caracterului complex al existenței și se încadrează în procesul mai larg al schimbării modelelor culturale, proces care se întemeiază, la rândul său, pe schimbarea modurilor de practicare a valorilor tradiționale, pe apariția unor valori noi și pe relaționarea în forme noi a diferitelor tipuri de valori în comportamentul oamenilor.

Îmbinarea actuală a valorilor morale cu valori științifice, tehnice sau artistice ar putea constitui, de pildă, o bază pentru conceperea și implementarea unui cod moral al mașinii. Ar putea fi simplificate

problemele apărute în acest proces considerându-se că agenții artificiali folosiți în cercetarea științifică vor împrumuta morala omului de știință?

Nu numai complexitatea, ci și complicarea continuă a omului, inclusiv sub raport moral, sunt avute în vedere în studiul nostru. Suntem deci departe de a adera la acele aprecieri conform cărora omul ar evolua din punct de vedere biologic, dar prin dezafectarea aptitudinilor sale naturale și nu împărtășim nici acele evaluări potrivit cărora omul ar involua din perspectivă culturală, prin simplificarea sau chiar dispariția vieții sale spirituale.

Dimpotrivă, adaptarea omului ca ființă culturală este tot mai dificilă, ca urmare a diversificării și intercondiționării modelelor culturale. Chiar dacă eforturile făcute de o mare parte a “ființelor culturale” nu sunt suficiente pentru orientarea sau integrarea culturală la nivelul posibilităților fiecăreia dintre ele, această situație este cauzată tocmai de integrarea lor într-un model cultural dominant, care nu solicită aceste posibilități sau nu permite actualizarea lor.

Totuși, cultura profundă influențează atitudinile culturale, răzbate prin noianul subproduselor culturale și incită cultura comună prin expresii ale sale concentrate și percutante, cărora nu le poate rezista nici un strat al publicului. Mari voci readuc în atenție muzica clasică, pictori de succes reimpun reprezentarea plastică artistică, atracția graficii de calculator stimulează învățarea matematicii, eseistica moral-politică reînvie interesul pentru filosofie, marile viziuni contraintuitive stârnesc resursele cunoașterii comune pentru înțelegerea științelor naturii.

Ca rezultat al sintezei proceselor reale prezentate, în evoluția omului ca ființă biotică, socială, culturală, istorică, astăzi se accentuează inclusiv trăsăturile sale tehnice și chiar cele artificiale. Ca ființă culturală, de pildă, omul evidențiază astăzi trebuințe noi, inclusiv virtuale, satisfăcute în modalități virtuale, într-un mediu virtual. Mai mult, trebuințe umane permanente și importante se manifestă în zilele noastre în forme virtuale. Este vorba de trebuințe atât materiale, cât și spirituale care, deși la fel de adânc resimțite ca altădată, sunt satisfăcute în condiții și cu mijloace virtuale, dintre care unele, de maximă eficiență.

Accentuarea complexității omului ca ființă culturală înseamnă și complicarea moralei umane, a cărei esență este de domeniul vieții spirituale. Spiritualitatea morală, la rândul ei, este rezultatul manifestării conștiinței morale și expresia practicării valorilor morale. Moralitatea devenind mai diversă, ca urmare a unor evoluții ale fenomenului moral în cadrul diferitelor sisteme morale, dar și ca efect al intersectării diferitelor sisteme

morale care aparțin unor modele culturale distincte, este previzibil ca omul și calculatoarele să aibă dificultăți chiar mai mari în a

- înțelege,
- interpreta
- practica valorile morale.

Persistența unor dificultăți de acest fel este însoțită, astăzi, și de existența unor premise favorabile, de apariția unor contexte încurajatoare, ca și de formularea unor previziuni cu grad crescut de probabilitate ce vin dinspre tehnică.

Printre premisele favorabile putem nota faptul că, la fel ca în cazul tuturor fenomenelor sociale, unele aspecte ale comportamentului moral al omului sunt și au fost întotdeauna artificiale, au reprezentat instituire umane, în limitele istorice ale posibilului uman, stabilite preponderent prin extinderea posibilului tehnic (11, 161-235). Morala artificială ar putea constitui deci, și din acest punct de vedere, extinderea caracterului artificial al moralei.

O altă bază preexistentă pentru elaborarea și implementarea eticii mașinii este dată de faptul că diferite aspecte ale comportamentului uman au fost deja modelate și au fost distinse structuri, funcții și procese specifice diverselor tipuri de activități umane. Este adevărat că s-au dovedit mai ușor de modelat și de simulat activitățile abstracte ale omului, care necesită numai respectarea legilor formale ale gândirii, fără implicarea explicită a legilor sociale și naturale, a căror respectare este impusă în funcționarea componentelor materiale a mașinii.

Unele dintre cele mai complexe tipuri de comportament sunt, de pildă, cele afective, puternic implicate în deciziile morale. Aceste comportamente sunt însă, și ele, în parte învățate, în parte inventate. Modelele comportamentului afectiv sunt însușite, se pare, până la vârsta de trei ani. Ele determină, cu forța unor structuri permanent consolidate și cu stabilitatea unor deprinderi niciodată conștientizate, tot parcursul vieții noastre afective.

În știința calculatoarelor s-a studiat, pe lângă recunoașterea și reproducerea formelor în scopuri tehnice sau îndeplinirea altor obiective de cunoaștere perceptivă, cum ar fi rezolvarea problemei culorilor, și reprezentarea nonsentențială, diagramatică a cunoștințelor abstracte, ca și stabilirea ariilor diferitelor domenii ale cunoașterii (cum este și cunoașterea socială) pe cortex și chiar învățarea recunoașterii emoțiilor umane în imagini.

Învățarea culturii este una dintre opțiunile educaționale în condițiile în care însăși cultura este concepută din perspectivă informațională sau este

studiată ca limbaj, deci ca sistem de semne, pe baza căruia sunt generate texte, configurate în funcție de reguli. În consens cu aceste viziuni, unii tehnologi (6, 59) consideră că, pentru a contracara lipsa experienței concrete și dificultatea dobândirii habitusurilor culturale, sisteme informatice similare celor actuale, cum sunt calculatoare neurale, vor fi dezvoltate prin învățarea din experiență și prin asimilarea culturii după modelul educării copiilor.

3. Dificultăți ale cunoașterii și aplicării teoriilor etice

Dificultățile conceptuale în formularea eticii mașinii sunt legate și de complexitatea cunoașterii și aplicării teoriei morale (a eticii). Etica, deși cunoaște și variante științifice, reflectă unele caracteristici ale sistemelor morale reale.

Sistemele morale tradiționale sunt sisteme preferențiale, care nu pot fi satisfăcător corectate prin sistemele etice acreditate. Subiectul uman și grupul de persoane preferate tind să devină

- sistemul de referință al stabilirii criteriilor deciziei,
- sursa conținutului acceptat al valorilor pentru comunități și epoci distincte
- unitatea de măsură a valorii acțiunii
- temeiul evaluării gradului de realizare a valorilor morale.

Etica interesului de grup, a comunității și chiar a majorității nu elimină caracterul preferențial al comportamentului moral. Ar putea rezulta de aici inclusiv că un calculator dotat cu un program care descrie adecvat funcții morale derivate corect dintr-un sistem etic construit pe un model cultural nou ar putea avea un comportament moral mai adecvat decât o persoană.

Chiar cunoașterea și aplicarea teoriilor morale actuale prezintă însă dificultăți începând cu problemele generate de diferențele dintre cunoașterea naturii, respectiv a societății, care înseamnă cunoașterea structurilor și proceselor sociale, a tendințelor și perspectivelor evoluției unui sistem social în corelație cu alte sisteme, dar și cunoașterea valorii și posibilităților persoanelor, a relațiilor interpersonale.

Reprezentarea specificității și complexității cunoașterii sociale se poate întregi prin mențiunea că, în acest domeniu, cunoașterea științifică presupune și chiar generează inclusiv cunoaștere tehnică. Numeroase discipline sociale adoptă și folosesc tehnici de cunoaștere generale și inventează metode și tehnici specifice, ca în cunoașterea sociologică sau psihologică.

Cunoașterea artistică, morală și politologică pot include, de asemenea, aspecte științifice și se pot servi de cunoștințe și mijloace tehnice. Precizăm că, în acest context, *prin tehnică înțelegem*, în primul rând, *proiectul intelectual care orientează orice tip de acțiune* și, în al doilea rând, *ansamblul de procedee ce asigură eficiența activității în diferitele domenii ale cunoașterii*.

Pot fi adăugate două alte mențiuni importante.

1. Fiecare dintre aceste forme de cunoaștere poate avea un anumit grad de obiectivitate, în funcție de obiectul, metoda și instrumentele cercetării, interpretării și valorificării.

2. Numai împreună, ca aspecte ale culturii spirituale, rezultatele acestor demersuri cognitive pot duce la explicarea corectă și la orientarea adecvată a domeniilor acțiunii, dintre care face parte și morală.

Suntem de părere că este necesar să subliniem că însăși *cunoașterea morală* poate fi de mai multe feluri: filosofică, științifică și tehnică.

La rândul ei, *cunoașterea morală științifică* s-a dovedit a fi empirică, teoretică și metateoretică. Formele cunoașterii empirice în etică sunt augmentate astăzi prin posibilitatea căutării, prelucrării și transmiterii informației prin multimedia.

Putem remarca, în acest context, că însăși etica de calculator, în cadrul căreia sunt studiate inclusiv istoria eticii și principalele teorii etice elaborate în cursul ei, este, în mare parte, o etică empirică, folosind inclusiv cunoașterea subiecților, a documentelor, a personalității deschizătorilor de drumuri în domeniu și încurajând aplicațiile care au ca obiect modelarea situațiilor morale problematice sau construirea unor proceduri de decizie în probleme morale care pot influența formularea unor politici publice.

Știința calculatoarelor însăși este expresia îmbinării cunoașterii științifice și tehnice, iar în cadrul ei, mai precis în domeniul cercetării inteligenței artificiale, se conturează și o cunoaștere etică. *Acest domeniu este, de altfel, un nod de redirejare a cercetărilor din științe aparținând biologiei și psihologiei, după ce disciplinele menționate au fost deja transformate, nu numai prin centrarea lor asupra problemelor cunoașterii și acțiunii, ci și ca urmare a integrării unor reflecții filosofice*. S-a consolidat, astfel, atât psihologia cunoașterii (numită de către autorii specializați psihologia mecanismelor cognitive, iar mai nou, psihologie cognitivă), cât și psihologia filosofică, distinctă de lucrările de metapsihologie, dar dezvoltată în conexiune cu necesitățile promovării inteligenței artificiale. Relevarea complexității domeniilor de cunoaștere menționate devine realistă abia după ce reamintim și faptul că, legat de ele

se dezvoltă filosofia psihologiei sau filosofia biologiei, alături de alte specializări ale filosofiei științei.

Studierea și folosirea nivelului spiritual al activității umane, din care face parte și cunoașterea morală, este mai dificilă. O strategie aplicată astăzi este și punerea pe net a rezultatelor acesteia. Acest lucru înseamnă însă numai facilitarea accesului la cunoașterea de tip moral, nu neapărat și valorificarea acestei posibilități. De asemenea, în acest fel este facilitat mai ales procesul didactic și mai puțin cercetarea științifică și tehnică, cu atât mai puțin cercetarea interdisciplinară, ca cea necesară pentru dezvoltarea inteligenței artificiale.

Cunoașterea morală și cunoașterea tehnică sunt sau vor fi îmbinate pentru formularea eticii mașinii. Forma de cunoaștere sintetică devenită astfel posibilă se poate baza inclusiv pe studiul comparativ privind morala umană și etica mașinii. Un astfel de studiu poate evidenția atât diferențele dintre ele, cât și elementele lor comune, dar și perspectivele coevoluției lor, respectiv alternativa posibilă: evoluția moralei umane către o morală artificială.

4. Morala umană și etica mașinii

Anticiparea unei evoluții ca cea menționată ar putea fi inhibată de o serie de diferențe, astăzi evidente, între reala morală umană și posibila morală artificială a mașinii. Dar, așa cum se poate deduce chiar din unele precizări anterioare, aceste deosebiri ne apar în principal ca rezultate ale caracterului practic și eterogen al formării și practicării conduitei morale la om, în opoziție cu coerența aspectului teoretic și eficiența componentei tehnice implicate în modelarea și posibila implementare a acesteia în funcționarea mașinii. Implementarea unui cod etic presupune însă o mașină dotată nu numai cu inteligență și capacitate de decizie situațională, ci și cu discernământ evaluativ și responsabilitate comportamentală.

Morala umană

are o structură socio-culturală complexă
- conținutul valorilor și
- forma normelor

reconstituie
sau cadrele organizării umane
anticipează

Etica mașinii

poate fi derivată din științe ca:
- știința conduitei morale - etica
- alte științe socio-umane
- științele tehnice
- științe ale spiritului cum sunt
și matematicile

este diferențiată în funcție de

- arie culturală
- comunități umane
- categorii sociale
- grupuri profesionale

va fi omogenă, globală, cu

- diferențieri numai pentru
- domenii de activitate
- tipuri de sarcini
- grade de complexitate

În morala umană s-au succedat multiple dihotomii istorice. Categoriile sociale distincte li s-au impus, în aceeași epocă, atât prin sistemul conexiunilor inverse dintre condițiile sociale și conduitele conformiste respectiv inventive determinate, cât și prin norme exterioare, consfințite religios sau prescrise prin lege, coduri morale diferite.

Mai mult, dincolo de aceste diferențe sociale, s-au perpetuat deosebiri elementare, traduse în sisteme de obligații și permisiuni distincte în cadrul aceluiași grup social și chiar familial. Până în zilele noastre persistă, astfel, dihotomii determinate de factori diferiți ca pondere dar, se pare, permanenți: biotici, psihici, sociali, istorici, culturali și profesionali.

Aceste tipuri de factori sunt prezenți în ponderi diferite în geneza și structurarea comportamentului uman individual și colectiv. Ansamblul lor variabil determină, în opinia noastră, două tipuri distincte de comportament uman (individual și social), perpetuate și transmise prin două coduri culturale diferite.

Cultura creației și competenței

Comportament preponderent creativ caracterizat prin:

- capacitate analitică/meticulozitate
- conștiinciozitate
- orientare echilibrată
 - spre $\left\{ \begin{array}{l} \text{viitor} \\ \text{păstrarea} \end{array} \right.$
 - regulilor
 - tradițiilor
 - valorilor
- actualizarea aptitudinilor personale
- folosirea resurselor proprii

Cultura competiției și confruntării

Comportament parțial distructiv caracterizat prin:

- perspectivă sintetică/operativitate
- abilitate
- orientare spre prezent
- apelul la autoritate, mai ales în forma ei de putere și chiar tendințe de folosire a violenței
- capacitatea de $\left\{ \begin{array}{l} \text{identificare} \\ \text{valorificare} \\ \text{dezvoltare} \end{array} \right.$ a resurselor externe.

Sunt practicate, în consecință, și două etici.

ETICA	ETICA
<ul style="list-style-type: none"> • identității, integrității, continuității • grijii 	<ul style="list-style-type: none"> ▪ cuceririi, progresului, riscului ▪ stăpânirii
<ul style="list-style-type: none"> • acumulării rezultatelor benefice tendențelor 	<ul style="list-style-type: none"> -formelor disensului ▪ sporirii -mijloacelor distrugerii
<ul style="list-style-type: none"> • etica comunicării și în plan profesional competenței 	<ul style="list-style-type: none"> ▪ etica competiției.

Ambele etici sunt necesare și justificate. Necesare întrucât reprezintă modalități de orientare ale comunităților umane care le-au permis să „traverseze” istoria până la noi și justificate prin teorii sau chiar orientări teoretice în cadrul reflecției morale.

Dificultățile par, în mare măsură, similare în însușirea valorilor morale și în întemeierea comportamentului moral la om și mașină. Aceste dificultăți vor fi tratate ca dificultăți practice, între care distingem dificultăți generale (culturale), altele particulare (cele științifice), ca și unele specifice (considerate aici a fi cele tehnice). Inclusiv dificultățile culturale pot fi considerate practice dacă prin cultură înțelegem tocmai felul în care sunt trăite și practicate valorile.

5. Dificultăți în conceperea, practicarea și înnoirea valorilor morale

Atât în cazul omului ca ființă socială, cât și în cazul mașinii, ca entitate inventată de om, descrierea, modelarea și implementarea conduitei morale prezintă dificultăți, generate de nivelul dezvoltării științelor socio-umane și de ritmul progreselor tehnice. În ambele cazuri, conduita morală

pare să fie impusă din afară, cel puțin în perioada inițierii vieții spirituale, respectiv, în momentul implementării codului moral.

În ceea ce privește omul însuși, pentru care nu numai calitatea de ființă culturală, ci și aceea de ființă socială este condiționată de practicarea unui sistem de valori, cu un obligatoriu subsistem moral,

- valorile morale sunt induse prin sistemul sancțiunilor pozitive/negative și, ca urmare,
- comportamentul moral se înscrie într-un sistem de obligații care provoacă emoții negative, iar

- viața morală este experimentată subiectiv deseori ca trăire a constrângerii

- opțiunile morale se schimbă greu, întrucât conținutul, procesualitatea și finalitatea valorilor morale sunt conștientizate în grade diferite

- valorile morale sunt vagi, iar normele morale sunt relative la caracteristici ale comunității și aspecte ale vieții individualității

- legătura dintre valori și norme nu este univocă

- discontinuitatea valori – norme morale este prelungită în

- imprecizia normelor morale de tipul „fii corect”, „fii cinstit”, „fii bun”; acestea

- constituie imperative al căror conținut nu este determinat decât în procesul vieții morale reale.

Viața morală autentică, în sensul de viață spirituală, condiționată de accesul la valorile morale și desfășurată prin dezbateră omului cu sine, pe baza acestora, este îngreunată deci de caracterul general al valorilor morale și de lipsa normelor strict corespunzătoare. Astfel, valorile și normele morale nu sunt adecvat înțelese și aplicate nici în practica umană.

Comportamentul moral implică formarea *conștiinței morale*, prezentă însă, chiar la om, în diverse grade și forme, începând cu

- deprinderile morale,
- emoțiile și sentimentele morale și continuând cu
- convingerile morale (care îmbină aspecte cognitive și afective) și
- reflecțiile morale.

Cu privire la convingeri, care sunt componente ale conștiinței morale, este necesar să facem unele precizări, pentru a putea evidenția complexitatea lor și dificultățile pe care le prezintă formarea lor, chiar în cazul omului. Ansamblul convingerilor este

- parțial coerent, ele se aplică alternativ, în contexte adecvate,
- convingerile sunt stabile (inclusiv pentru că adesea nu sunt

conștientizate), dar sunt

- ușor transpozabile la situații diferite (avantaj); ele se dovedesc însă și
- foarte maleabile în procesul trecerii de la intenții la realizări (dezavantaj).

În cazul mașinii,

- valorile orientative sunt de asemenea impuse din afară, dar
- comportamentul mașinii este strict normat prin instrucțiuni explicite; cu toate că, și în acest domeniu,
- valorile sunt implicite, normele/regulile sunt explicite.

Sunt prezente, în domeniul studierii eticii mașinii, și aspecte cognitive comune cu omul:

- codul moral folosit poate fi aplicarea unei teorii etice
- presupune descrierea domeniilor /situațiilor concrete
- necesită învățarea prin antrenare.

Conștiința presupune însă, pe lângă cunoaștere, afectivitate și voință. Ea se obiectivează, între altele, în forme de spiritualitate.

Ca în orice domeniu de activitate spirituală, și în morală, conștiința este explorată și folosită ca resursă pentru creativitate. În acest punct, ne delimităm deci de opiniile conform cărora inconștientul ar fi principala sursă, permanentă și infinită, a creativității.

Ce înseamnă creativitatea în morală? În ce măsură putem fi creatori în raport cu valorile morale? Astfel de probleme pot fi încadrate într-o altă categorie de dificultăți întâmpinate în formularea eticii mașinii, dificultăți pe care le putem numi, generic, dificultăți culturale.

Problemele de tipul menționat se pun, în legătură cu domeniul discutat, în două registre diferite. Întrebarea cu privire la gradul de creativitate acceptabil cu privire la valori în particular și la valorile morale în special se pune în raport cu omul, în timp ce posibilitatea manifestării oricărui tip sau grad de creativitate este pusă global sub semnul întrebării în cazul mașinii.

Atitudinile formulate în privința gradului de creativitate al omului în raport cu valorile morale sunt multiple, așa cum reiese și din lucrările unor autori preocupați de „moralitatea artificială”(4, 292). Am corelat analiza rezultatelor unor astfel de studii cu aplicarea ideii existenței unor particularități ale creativității în diferite domenii ale culturii la sfera culturii morale. Având în vedere și rezultatele unor cercetări concrete proprii privind modul de percepere al unor valori specifice de către categorii socioprofesionale reprezentative pentru viitor, distingem mai multe atitudini

culturale posibile în această privință, fiecare cu anumite grade de probabilitate a valabilității în prezent și în viitoarele decenii.

O atitudine răspândită este aceea potrivit căreia valorile sunt parțial învățate, parțial inventate; ele sunt achiziții intelectuale la fel ca automatismele de vorbire sau de gândire, dar nu constituie cadre imuabile, ci funcționează mai degrabă ca baze ale unor orientări axiologice caracteristice diverselor domenii ale culturii, diferențiat asimilate și aplicate de către indivizi.

În conformitate cu o altă atitudine, recunoașterea și trăirea valorilor țin de conduita practică și sunt marcate de tensiunea raporturilor individ – societate, prin care se produce atât integrarea persoanelor, prin comportamente conformiste, cât și evoluția lor distinctă, prin manifestări creatoare, care au drept rezultat înnoirea valorilor acceptate.

O altă atitudine ar putea fi caracterizată ca „fidelitate creatoare” față de valori, prin care sunt permanent reafirmate valorile acreditate și formele de creație subsumate, fiind aduse, în același timp, inovații personale, cu condiția recunoașterii continue a filiației lor comune.

Regenerarea permanentă (nu simpla persistență) a valorilor funcționale poate fi afirmată ca o altă posibilitate. Valorile preexistă fiecărui individ, reprezentând cadrul cultural în care acesta se integrează și pe care îl reconstruiește permanent în cooperare cu alte personalități creatoare. Nu se produce, deci, simpla continuitate a credințelor sau permanentizarea aspectelor rituale (exterioare) ale tradițiilor; la fel cu diferitele feluri de celule ale corpului nostru, diversele tipuri de valori se înnoiesc la intervale distincte, dar se păstrează integritatea sistemelor din care ele fac parte. Se păstrează astfel cadrul axiologic de bază al culturii, dar se înnoiește conținutul valorilor componente.

Generarea unor valori noi, impuse sau permise de condiții culturale în schimbare este accentuată într-o ultimă atitudine culturală identificată dintre cele posibile. Personalitățile creatoare aparținând unui sistem cultural au tins să se desprindă de influențele acestuia, dar au reușit, până acum, să depășească doar anumite limite personale, sociale și istorice. Numai astăzi, ideile, oricât de noi, pot fi nu numai admise, ci și întrupate, prin intermediul valorilor tehnice, uneori chiar dincolo de previziunile autorilor. În același timp, realizarea valorilor științifice și tehnice excede, deseori, limitele conturate de valorile morale tradiționale.

Valorile tehnice însele nu mai sunt considerate doar valori instrumentale care ar fi mijloace de aplicare, în forme utile, a unor valori intelectuale (ca cele științifice sau filosofice), a unor valori spirituale (cum ar fi cele artistice sau religioase) cu respectarea unor valori santinelă (cele

morale sau juridice). Ele pot funcționa chiar ca valori eliberatoare în raport cu valori dominante (valorile economice și politice), promovând, prin tehnica informatică, nevoile spirituale și chiar virtuale, transformând conflictul, negoțul și compromisul politic în organizare eficientă.

Se manifestă însă, deocamdată, până la dovedirea eficienței acestor tehnici în toate domeniile de activitate, dificultăți în argumentarea necesității continuității unor valori spirituale, iar valorile economice, tehnice sau politice tind nu numai să se autonomizeze în raport cu cele spirituale, dar, uneori, chiar să le dizloce sau să le elimine.

Devine necesară, de aceea, inclusiv o tehnologie a culturii, care să includă managementul cultural. Acest tip de management se impune să se practice și ca un management strategic, care privește orientarea culturii în ansamblul ei.

Etica mașinii reprezintă, în aceste condiții, nu numai un domeniu de cercetare teoretică și un prilej de demonstrare a unor performanțe tehnice, ci o reorientare categorică și semnificativă în gândirea științifică actuală, de reintegrare a reflecției filosofice în știință și tehnică, de redirecționare a inventivității tehnice către problemele umane esențiale.

6. Dificultăți în implementarea unui cod moral în funcționarea mașinii

Se constată că acest tip de dificultăți își are sursa tot în probleme teoretice, iar rezolvarea provine din regândirea modului nostru de concepere a fundamentelor moralei umane, a statutului valorilor, normelor și regulilor morale. Unele dintre aceste probleme conceptuale par insolubile, nu atât ca urmare a complexității, cât din cauza ambiguității și procesualității, respectiv conjuncturalității situațiilor morale. Aspectele menționate sunt puternic influențate de subiectivitatea factorilor implicați, inclusiv de caracteristicile și stările subiective agenților care ar urma să ia decizii în probleme teoretice sau în situații practice.

Astfel de probleme tehnice, legate de dificultăți conceptuale, sunt persistente și interdependente.

Apare, în acest context, mai întâi, o problemă logică: pentru a putea fi considerate "morale" în sensul moralei umane, preferențiale, mașinile ar trebui să funcționeze nu numai după o *logică tehnică*, ci și în conformitate cu o *logică umană*, operaționalizată ca logică a preferinței sau ca logică a convingerii.

Problema care urmează este de natură psihologică: elaborarea și aplicarea eticii mașinii presupune nu numai implementarea *intelenței abstracte*, matematice sau lingvistice, așa cum se întâmplă astăzi, ci și cunoașterea și folosirea *intelenței emoționale*. Vom arăta, în ultimul paragraf, că în conduita morală este necesară combinarea mai multor tipuri de inteligență.

Vor trebui luate în considerare, de asemenea, o serie de probleme tehnice nu doar de ordin conceptual, ci și de ordin practic.

O astfel de problemă practică este cea a găsirii unei căi de completare a deciziei obiective, mașinale, cu trăsături ale deciziei umane, secundată de cea a identificării și reconstituirii abilităților de folosire a avantajelor acestora și de evitare a dezavantajelor caracteristice.

Un alt grup de probleme rezultă din necesitatea ca sistemele informatice dotate cu programe de inteligență artificială avansate, care vor permite implementarea eticii mașinii, *să fie tratate ca entități individuale*, ca urmare a:

- complexității
- funcționalității specifice
- gradelor de libertate atribuite.

Poate cele mai dificile și importante probleme vor decurge din consecințele îmbinării, în cadrul eticii mașinii, a unui comandament teoretic și a unui imperativ practic.

Morala înseamnă și libertate morală. În cazul eticii umane este considerată axiomatică ideea că numai în condițiile asumării și aplicării libere a normelor morale comportamentul uman are valoare culturală.

Și în cazul mașinii, corelația libertate – responsabilitate trebuie să funcționeze în ambele sensuri. Gradul de responsabilitate depinde de gradul de libertate.

Putem constata însă că relația libertate – responsabilitate nu este deplin funcțională nici în ceea ce privește omul. Se acceptă că responsabilitatea decurge din libertate, dar se trece sub tăcere faptul că există, mai mult, o dependență exact determinabilă a gradului de responsabilitate de gradele și de natura libertăților practicate. De asemenea, atât sistemele, cât și grupurile de interese accentuează asupra gradelor de libertate acordate, dar uită că singura libertate cu valoare individuală este cea luată de individul însuși în condiții de responsabilitate. Astfel, responsabilitatea este încurajată drept condiție a libertății, dar libertatea nu este socotită drept cauză a responsabilității.

Din premisele anterioare și discutarea lor rezultă și importante concluzii pentru problematica implementării codului moral al mașinii.

Etica mașinii presupune și libertatea alegerii în aplicarea normelor morale în domenii de acțiune și în situații determinate.

Omul trebuie să-și asume riscurile ce decurg din acordarea libertății. Mai mult, umanitatea, ca multitudine de indivizi și comunități, urmează să accepte și posibilitatea mașinii de a-și înmulți gradele de libertate.

În elaborarea și implementarea eticii mașinii apar *dificultăți conceptuale generate de rezolvarea unor probleme tehnice*, nu numai probleme tehnice legate de persistența unor dificultăți conceptuale. Observăm, în acest context, de asemenea, că atât problemele conceptuale, cât și cele numite tehnice, pot fi interpretate și tratate, cele mai multe, și ca probleme de cunoaștere și acțiune în perimetrul ființării umane, atât individuale, cât și sociale.

Problemele conceptuale identificate în urma cercetării posibilității de concepere și implementare a unui cod moral pentru mașini țin de capacitatea noastră de a înțelege însuși fenomenul moral în completitudinea componentelor sale: spiritualitate, normativitate și activitate. Din perspectiva problematicii urmărite, accentul cade pe norme și punerea lor în acțiune.

Normele morale au anumite particularități: sunt valabile pentru orice tip de activitate umană; toate activitățile umane au o dimensiune morală și pot fi evaluate din această perspectivă.

Moralitatea este una dintre dimensiunile tuturor activităților umane, dar aplicarea normelor morale în domenii diferite cere

- abilități specifice
- achiziția unor instrumente culturale.

Normele morale generale, perene, transpozabile în diverse domenii pot fi corect interpretate și efectiv folosite, în cazul omului, numai prin cultivarea și *folosirea inteligenței, conștiinței și creativității*.

Pot fi asimilate valorile morale cu criteriile de decizie, iar normele morale, care sunt reguli de conduită complexe, cu proceduri standard, pentru a fi implementate în programe de calculator?

Devine necesară formularea unor modalități explicite de transpunere a normelor morale în prevederi operaționale valabile în diverse domenii și de adecvare a lor la diferite împrejurări posibile.

7. Ipoteze de lucru pentru implementarea unui cod moral în conduita mașinii

În vederea elaborării și folosirii unui cod moral al mașinii putem formula mai multe ipoteze de lucru: (1) Ipoteza structurală, (2) Ipoteza funcțională și (3) Ipoteza acțională. În toate aceste ipoteze, trebuie să renunțăm la avantajele existenței conștiinței și a conștiinței morale, să contăm pe folosirea inteligenței, în formele sale artificiale actuale, (care pot constitui însă puncte de plecare pentru studierea unei forme de inteligență adecvate activității morale), ca și pe elemente de conduită creativă, permise de modalitățile de programare euristică.

7. 1. În ipoteza structurală, codul moral ar putea fi implementat prin programe care ar putea conține seturi de norme morale adecvate pentru

- funcții diferite
- domenii determinate de activitate
- tipuri de activitate
- clase de probleme teoretice definite sau situații practice

specificitate cu care ar putea fi confruntate mașinile sau, mai bine zis, toate sistemele artificiale inteligente.

Codul moral ar putea fi însă implementat și prin analogie cu structura inteligenței umane. Astfel, de pildă, pentru implementarea unui cod moral în structura și funcționarea agenților inteligenți nu este suficientă modelarea și implementarea inteligenței abstracte, ci se impune formarea și activarea unor structuri și procese caracteristice judecății, deciziei și acțiunii morale.

Aspectul comportamental al conținutului activității inteligente pare a fi esențial, în acest context, iar asupra acestui aspect vom reveni în cadrul discuției ipotezei acționale.

În ipoteza noastră structurală, între formele de inteligență diferențiate în funcție de domenii de activitate determinate putem adăuga și *intelența morală*, care este necesară pentru respectarea normelor specifice care derivă din sistemul valorilor ce orientează practica morală, dar care ar putea întemeia și componenta intelectuală a creativității morale.

Structura inteligenței morale, odată identificată, de pildă prin raportarea componentelor inteligenței generale la structura fenomenului moral, ar putea fi modelată și folosită în scopul implementării unui cod moral în funcționarea sistemelor inteligente, ceea ce presupune însă, mai întâi, studierea ei de către filosofi, psihologi și tehnologi reuniți în echipe interdisciplinare.

7. 2. Ipoteza funcțională poate fi luată în considerare dacă:
- nu apreciem ca utilă sau nu putem realiza stabilirea locului unui cod moral

- în structura conștiinței umane, respectiv a psihicului artificial;
- identificând într-un fel acest cod în conștiință, nu-l putem accesa, modela și transfera, în forme operaționale;
 - inteligența însăși, chiar diferențiată în diferite forme și după diverse criterii, cum ar fi acela al domeniului de activitate (în cazul nostru, cel al moralității), nu poate fi identificată sub forma unor structuri formale, măsurabile, stabile, corelate și unitar eficiente;
 - se constată că nivelul de inteligență nu determină univoc calitatea comportamentului moral;
 - pentru formarea conduitei morale se impune mai ales cultivarea unor trăsături de personalitate, deci obținerea unui nivel superior de realizare a individului ca entitate psihică, deocamdată inaccesibil abordării tehnice.

Ipoteza de lucru care poate deveni operațională în aceste condiții ar fi aceea a identității funcționale a moralității.

În acest caz se pun însă o serie de alte probleme: este moralitatea o însușire funcțională?

În caz afirmativ, este aceasta o funcție de sine stătătoare? Este un grup de funcții? Este realizată de psihic în ansamblul său? Se bazează pe alte funcții? Este efectuată de structuri și funcții psihice nespecializate dar cooperante?

Opinia noastră, bazată pe o viziune sintetică, este că, dacă putem interpreta moralitatea ca realizare a unei funcții psihice de sine stătătoare, aceasta este mai degrabă o funcție supradeterminată, o funcție spirituală.

Toate activitățile spirituale presupun și subsumează mai multe niveluri distincte de activitate, manifestate în același timp, atât în cursul acțiunilor individuale, cât și în planul activităților cooperante:

- activitate intelectuală,
- activitate automată (deprinderi)
- activitate instinctuală.

Implementarea moralității ca o însușire funcțională necesită, ca urmare, în continuare, o abordare acțională.

7.3. Ipoteza acțională

Eventualele răspunsuri certe sau cel puțin utile obținute pornind de la ipoteza precedentă pot fi folosite ca premise în cazul în care ne propunem să construim o a treia ipoteză, ipoteza acțională.

Comportamentul moral este activitate complexă, care include - componente spirituale,
- componente intelectuale, dar care se bazează și pe

- procese automate (deprinderi) și are drept condiții
- activități instinctuale.

Morala este o activitate spirituală, la fel ca știința, filosofia, arta sau tehnica. Dezvoltarea acestor domenii ale culturii atestă că nivelul superior al comportamentului uman este cel mai eficient, pentru că omul lucrează, în acest caz, la nivelul său specific de existență și de activitate (15, 438, 448-452).

La polul opus, în lumea animală, dar și în lumea creațiilor tehnice neanimate, activitățile automate par mai eficiente.

Sistemul social se poate compara, de asemenea, cu succesiunea modelelor de eficiență inventate de natură pentru creșterea eficienței funcționării sistemelor. Acest scop este realizat însă, pe ansamblu, în mod automat, prin faptul că sistemele, cu cât sunt mai complexe, cu atât sunt mai condiționate.

În cazul sistemului social, se poate produce autocondiționarea, dar corelată cu intercondiționarea și supracondiționarea (intercondiționarea globală). Aceasta are însă, de regulă, nu numai efectele propuse și scontate, ci și consecințe cu atât mai neașteptate și incontrolabile, cu cât sunt mai îndepărtate, indirecte și intricate. Astfel, sistemele sociale evoluează prin sacrificii individuale, ale mediului și chiar ale propriilor perspective, dacă abandonează tendința „naturală” de autostructurare și autodezvoltare caracteristică sistemelor cibernetice

Comportamentul uman individual este, de asemenea, complex și este condiționat de acțiunea tuturor tipurilor de legi - fizice, biotice și sociale -, dar cunoașterea tuturor acestor tipuri de legi nu este obligatorie pentru succesul acțiunii umane însăși, care poate fi eficientă dacă se poate adapta acțiunii concrete a acestor legi sau dacă poate crea condiții care permit realizarea dorințelor în spațiul tendințelor legitime.

Tendința creșterii eficienței prin reinstalarea automatismelor, care s-au dovedit eficiente la niveluri inferioare ale existenței, se manifestă și în plan intelectual. Chiar pe parcursul dezvoltării omului ca individ, avantajele inteligenței fluide, ale memoriei permeabile și maleabile, cele ale deciziei rapide și ferme, ca și ale absenței reflecției ezitante, caracteristice vârstei tinere, facilitează formarea automatismelor intelectuale.

Calitățile menționate sunt suplinite, dar niciodată egalate, la maturitate, prin organizare, experiență și prin folosirea altor oameni, ceea ce definitivează complexul deprinderilor eficiente: omul și instituțiile umane devin automate mai mult sau mai puțin perfecte sub raportul îndeplinirii

obiectivelor instituționale, nu și în privința realizării unei eficiențe umane, care înseamnă eficiența globală a acțiunii, raportată la valorile umane.

Mașinile inteligente și oamenii nu se deosebesc mult în zilele noastre în ceea ce privește activitățile specializate / profesionalizate.

Implementarea codului moral ar putea folosi capacitatea și disponibilitatea omului și a mașinii de a învăța inclusiv prin mecanisme și proceduri automate pentru efectuarea unor funcții și pentru optimizarea propriei activități.

Așa cum oamenii folosesc astăzi mașina informațională de pildă pentru analiza fundamentării deciziei, în viitor oamenii ar putea excela ca psihologi ai mașinilor, ca sociologi ai comunităților virtuale, ca lingviști și matematicieni, moraliști și filosofi, specialiști cu calificări deosebite sau rare, experți, consilieri.

8. Inteligența morală

Având în vedere structura complexă a conduitei morale, ca formă a vieții practice și spirituale, ca și rolul ei în organizarea și desfășurarea vieții individului respectiv comunității, semnalăm existența unui tip corespunzător de inteligență, *intelența morală*, și necesitatea cultivării acesteia în cazul omului, ca și obligativitatea implementării sale, în cazul mașinii inteligente; interacțiunea acesteia cu omul poate fi vitală.

Observăm că în domeniul specializat în studierea inteligenței umane nu se acordă întotdeauna o importanță egală diferitelor tipuri de inteligență, fiind evidențiate și asociate formele mai bine studiate, iar clasificările nefiind făcute după criterii de aceeași importanță. De aici rezultă un număr mare de clasificări, unele tipuri de inteligență apărând în cadrul mai multor modalități de clasificare.

Din punct de vedere structural, inteligența umană a fost analizată progresiv, prin identificarea unui factor general și apoi a unor factori semigenerali, pentru ca să se studieze, în final, și unii factori specifici, care au fost corelați, inclusiv într-un model spațial.

Unele studii au evidențiat particularizări ale inteligenței umane, cum este și *intelența tehnică*. În funcție de aptitudini, de tipul de educație și de activitate, ca și în dependență de experiența în activitate și de mediul specific se conturează și diverse alte tipuri de inteligență, ca inteligența lingvistică, matematică sau muzicală.

Propunem evidențierea posibilităților de diferențiere ale inteligenței în funcție de *un criteriu unitar, acela al domeniului culturii prin care se formează și în care se manifestă formele principale de inteligență*,

acestea putând subsuma diferite tipuri de inteligență care sunt acum incluse în mai multe clasificări diferite, cum este și cazul inteligenței matematice, cel al inteligenței literare, al inteligenței interpersonale sau cel al inteligenței sociale.

După criteriul propus putem distinge *inteligența științifică, inteligența artistică, inteligența tehnică, inteligența politică și inteligența morală etc.*

În această clasificare apar o serie de înnoiri, dintre care unele sunt de ordin metodologic și terminologic, iar două, de ordin teoretic.

Astfel, deși au fost studiate inteligența matematică, inteligența lingvistică, inteligența practică și teoretică, după știința noastră, nu a fost aprofundată *inteligența științifică* și nu a fost folosit termenul corespunzător. De asemenea, deși au fost studiate inteligența muzicală, plastică sau kinestezică, nu a fost caracterizată *inteligența artistică*. La fel, cu toate că numeroase clasificări includ inteligența interpersonală sau inteligența socială, nu a fost stăruitor analizată *inteligența organizațională*, cea care le-ar putea subsuma. Nu a fost studiată nici inteligența politică, care ar putea fi considerată una dintre cele mai influente forme de manifestare ale acesteia. Propunem, ca obiect de cercetare, din perspectiva prezentului studiu, *inteligența morală*, legat de care vom face și o serie de precizări.

Denumirile propuse de noi rezultă dintr-o perspectivă metodologică sistemică, asociată cu o viziune integrativă, în cadrul căreia putem vorbi și de inteligența morală în tipologia inteligenței umane.

Formele particularizate de inteligență pot include forme specializate de inteligență, care la rândul lor vor fi cuprinse, fiecare, o singură dată, într-o singură categorie. Astfel de forme specifice de inteligență se constituie potrivit cerințelor unor domenii de activitate, caracteristicilor unei discipline științifice, unui gen artistic sau unei direcții de dezvoltare tehnică, respectiv în funcție de normele derivate din exigențele unui anumit sistem de valori, ca cele morale.

Analiza și nominalizarea unei forme până acum neidentificate dar funcționale de inteligență, *inteligența morală*, presupun argumentarea legitimității studierii acesteia și a folosirii denumirii menționate. Avem în acest sens două categorii de argumente.

Primul grup de argumente se referă la statutul moralei ca fenomen uman, acela de aspect al vieții spirituale și, deci, al culturii.

Conștiința morală se constituie, așa cum am arătat, în trepte, dar odată dezvoltate nivelurile sale superioare, omul nu mai acționează în

virtutea deprinderilor sau constrângerilor morale, ci ca urmare a unei dezbateri interioare, spirituale.

Valorile morale se constituie într-un sistem particular de valori în cadrul sistemului general de valori caracteristice unei societăți. Realizarea acestui sistem particular de valori necesită și constituirea, cultivarea și manifestarea unui ansamblu de aptitudini umane consacrate, între care și un tip determinat de inteligență, inteligența morală.

Conduita morală presupune chiar combinarea, într-o formă distinctă de inteligență, pe care o numim *inteligență morală*, a mai multor tipuri anterior studiate de inteligență. Este vorba, pe lângă inteligența abstractă și inteligența emoțională, deja menționate, de altele ca inteligența concretă, descriptivă, interpretativă, imitativă, creativă, teoretică și practică.

Denumirea de *inteligență morală*, dată de noi acestei forme complexe de inteligență, care condiționează conduita morală eficientă, este întemeiată atât de *complexitatea cunoașterii morale* (la care ne-am referit în paragraful 3), cât și de *ierarhia structurală a conștiinței morale* (pe care am evidențiat-o în paragraful 5). Se adaugă *specificul relațiilor și practicilor morale*, care impun dezvoltarea și folosirea acestei forme distincte de inteligență.

Am argumentat, într-un studiu anterior, că analiza aptitudinilor umane implicate în conduita morală poate fi dusă până la identificarea aptitudinilor necesare îndeplinirii fiecărei valori morale, iar realizarea valorii morale fundamentale implică practicarea tuturor celorlalte valori morale.

Cel de-al doilea grup de argumente decurge tocmai din practica morală, ca formă fundamentală a conduitei practice. Or, tipurile distincte de inteligență se formează, se cultivă și se manifestă tocmai în funcție de particularitățile și cerințele diferitelor tipuri de activitate.

Situația studierii formei morale a inteligență este similară cu cea a formei tehnice a culturii. Deși a apărut în același timp cu alte forme demult recunoscute ale culturii, iar după unii gânditori, chiar înaintea aceluia, cultura tehnică a intrat mult mai recent în atenția filosofiei culturii, odată cu recunoașterea sau cel puțin cu impunerea tehnicii ca domeniu hotărâtor al dezvoltării sociale.

Inteligența morală nu se poate confunda cu inteligența generală, deși se manifestă în toate tipurile de activități umane. Acest tip de inteligență este activ și eficient în contexte specifice (morale), dar astfel de contexte se întâlnesc în toate domeniile de activitate și fiecare persoană trăiește experiențe morale, indiferent de natura preocupărilor sale profesionale.

În același timp, nivelul inteligenței morale poate fi considerat dependent de nivelul inteligenței generale, dar gradul manifestării inteligenței morale poate fi mult diferit de posibilitățile oferite de nivelul dezvoltării inteligenței generale și chiar de nivelul unor forme de inteligență particulară. Astfel, nivelul inteligenței morale, evaluat în funcție de gradul și eficiența manifestării sale, pare a fi dependent mai ales de conștiința morală, de măsura în care sunt prezente, active și interfuncționale diversele componente deja prezentate ale acesteia.

Exersarea conștiinței morale este relativă și la condițiile concrete ale vieții și activităților individului. Se poate afirma că ea există în măsura în care acționează, iar această afirmație este valabilă și pentru inteligența morală. Inteligența morală depinde nu atât de alte aptitudini mai mult sau mai puțin complexe, și deci de factori psihici, cât de factori spirituali, educaționali și culturali.

Totuși, în scopul implementării unui cod moral în conduita mașinii putem folosi rezultatele studierii unor dimensiuni ale inteligenței umane, care se descompune în aptitudini (aptitudinea verbală, numerică, raționarea, memoria etc.). Formele inteligenței umane, inclusiv inteligența morală, ar putea fi reconstruite în variante tehnice pornind de la astfel de aptitudini, exprimate în funcții adecvat descrise, care ar putea simula judecata morală și raționamentele complexe specifice acestui domeniu al activității spirituale.

Inteligența morală nu poate fi înțeleasă, în opinia noastră, ca fiind o particularizare a inteligenței generale, întrucât determinarea ei în acest fel produce dificultăți. Morala este un domeniu al culturii, moralitatea constituie un complex de relații umane, activitatea morală generează un subsistem al sistemului social, dar omul, pentru a se dovedi moral și pentru a manifesta inteligență morală cumulează și folosește însușiri proprii mai multor forme de inteligență umană.

Putem spune chiar că inteligența morală implică folosirea mai multor tipuri de inteligență umană. Astfel, la nivelul relațiilor și activităților morale este nevoie de inteligență concretă și de inteligență practică, de inteligență imitativă, dar și de inteligență creativă. Viața în comunitatea morală implică inteligență interpersonală și inteligență emoțională. Conștiința morală are drept condiție a existenței și funcționării inteligența intrapersonală. Nivelul științific al cunoașterii morale (etica științifică) presupune inteligență teoretică.

Dacă avem în vedere chiar specializări actuale ale eticii, ca bioetica și tehnocetica sau domenii emergente ale sale, cum este cea abordată de noi, etica mașinii, constatăm că dezvoltarea acestora presupune

intelență matematică și chiar intelență tehnică, iar pe lângă cunoștințe de specialitate, și disponibilitatea către reflecția filosofică. Situaarea la nivelul filosofiei morale implică intelență abstractă și intelență creativă.

Înțelegerea și folosirea valorilor morale, ca rezultat al activității spirituale, solicită, pe de o parte, intelență descriptivă și intelență cristalizată, dar se bazează și pe intelență interpretativă, iar în măsura arătată în paragraful 5, chiar pe intelență creativă. Eficiența activităților sociale, care (toate) implică practicarea valorilor morale, este condiționată de folosirea intelenței sociale.

Din analiza formelor de intelență presupuse, implicate și integrate în existența și folosirea intelenței morale observăm că intelența morală se bazează pe intelența generală, combină, în alcătuirea și funcționarea sa, forme particulare de intelență, iar pentru acoperirea unor noi domenii de activitate, apelează chiar la forme speciale de intelență. *Rezultă că intelența morală reprezintă o formă sintetică de intelență.*

Una dintre căile care vor face posibilă integrarea acestei forme complexe de intelență cu intelența artificială este chiar studiarea cât mai analitică a intelenței morale.

Dacă am supune însă unei analize atente și alte forme la fel de importante ale intelenței umane, am constata, probabil, că *toate formele fundamentale ale intelenței umane sunt forme sintetice de intelență.*

În același timp, putem aprecia că forța intelenței omului, luată în ansamblul ei, constă în *reunirea acestei disponibilități intelectuale generale pentru rezolvarea oricărei probleme semnificative.*

Discutând problema naturii și eficienței activității intelectuale a omului, putem afirma că aceasta din urmă poate fi realizată mai ales prin *concentrarea complexității intelectuale asupra unei singularități.*

Intelența artificială a fost dezvoltată, până acum, ca intelență abstractă, manifestată mai ales ca intelență matematică, respectiv lingvistică. Se pare că, mergând pe această linie, intelența artificială ar putea obține *câștiguri* numai în viteză, maleabilitate, manevrabilitate și universalitate. Toate acestea ar putea fi interpretate, în același timp, ca *pierderi* în profunzime, consecvență, autodeterminare (determinare din interior) și autodiferențiere.

Pentru a se dovedi utilă în activități relevante pentru natura și esența omului, intelența artificială ar putea deveni tot mai specializată. Diferențierea formelor de intelență implementate și folosite nu ar apropia neapărat mașina de om, iar acesta nu ar fi neapărat concurat de mașină, ci mediul artificial al vieții umane ar deveni mai divers, mai funcțional și mai stimulativ pentru creativitate, fără a ajunge obligatoriu mai complex.

La rândul său, omul evoluează spre artificialitate, iar efortul de a concepe și de a construi, respectiv de a forma cultural (spiritual) roboții și mașinile în general nu ar fi decât un alt mod, mai accelerat și mai simplu, de a realiza convergența a două tendințe „naturale” actuale și de a dezvolta societatea în întâmpinarea destinului ei tehnologic.

Moralitatea artificială ar fi o nouă formă de manifestare spirituală, atât umană, cât și tehnică, a activității umane, alături de arta computerizată, de filosofia artificială sau de politica digitală. De observat însă că, inclusiv atunci când vorbim de tehnică, ne referim la o manifestare umană, iar artificialul este, dintotdeauna și în toate formele sale, un produs în întregime uman.

Bibliografie

1. Bedau, M. A., „Philosophical Content and Method of Artificial Life”, Bynum, T. W. and Moor, J. H., *The Digital Phoenix: How Computers are Changing Philosophy*, Blackwell Publishers, Oxford, 1998.
2. Bynum, T. W., „Global Information Ethics and the Information Revolution”, Bynum, T. W. and Moor, J. H., *The Digital Phoenix: How Computers are Changing Philosophy*, (...).
3. P. M. Churchland, „The Neural Representation of the Social World”, Bynum, T. W. and Moor, J. H., *The Digital Phoenix: How Computers* (...).
4. Danielson, P., „How Computers Extend Artificial Morality”, în vol. ant. menționat.
5. Darwall, S. *Philosophical Ethics*, Westview Press, Colorado, Oxford, 1998.
6. Gregory, R., *Viitorul creatorilor de inteligență*, Editura Științifică, București, 2000.
7. Hillis, W. D., *Mașina care gândește*, Editura Humanitas, București, 2001
8. Londen, R. B., „Vices of the Virtue Ethics”, în vol. Crisp, R.; Slote, M., *Virtue Ethics*, Oxford University Press, 1997.
9. Pană, Laura, „Artificial Intelligence and Moral Intelligence”, Paper for the 3rd European Computing and Philosophy Conference, E-CAP 2005, Mälardalen University, Västerås, Sweden, 2-4 June, 2005.

10. Pană, Laura, „Co-evolution of Human and Artificial Cognitive Agents”, *i-C&P 2006, Computers and Philosophy, an International Conference*, Laval, France, May 2-3 2006.
11. Pană, Laura, *Cultura tehnică și industria culturală*, Editura Tehnică, București, 2002.
12. Pană, Laura, „De la etica virtuții la etica virtuală”, Prezentare în cadrul Grupului de cercetări interdisciplinare al CRIFST, Ianuarie 2005.
13. Pană, Laura, „Etica artificială”, în *Filosofia informației și a tehnicii informaționale*, Editura Politehnica Press, București, 2004.
14. Pană, Laura, „Filosofia artificialului și filosofia artificială”, în *Academica*, nr. 34, ianuarie 2005, Anul XV, 171.
15. Pană, Laura, *Filosofia culturii tehnice*, București, Editura Tehnică, 2000.
16. Pană, Laura, *Filosofia informației și a tehnicii informaționale*, (...).
17. Pană, Laura, „Moral Intelligence for Human and Artificial Agents”, în *Machine Ethics, Papers from the AAAI Fall Symposium*, 4-6 November, 2005, Arlington, Virginia USA, AAAI Press, American Association for Artificial Intelligence, Menlo Park, California, 2005.