

## WHAT IS IT LIKE TO BE A CONSCIOUS MACHINE?

**Florin George Popovici**

**PhD Student, "Ștefan cel Mare" University of Suceava**

*Abstract: It is no longer a novelty to claim that technology evolves in an accelerated and yet confusing manner. Whether we are talking about computer science, robotics, neuroscience, Artificial Intelligence or cognitive science, it is natural for us to ask ourselves if at some point across their struggle humans will be able to design and build machines endowed with consciousness. This means an artificial system able not only to represent and describe the world with all its remarkable features, including its own place and the way it functions in the nature, but being able to imagine alternative and possible worlds as well. Also a machine endowed with consciousness should prove the way it understands human interlocutors, the ability to interact, act and react, to operate rationally but also emotionally, to evaluate alternatives and make decisions in an autonomous way. For skeptic minds this seems to be just an unfeasible science-fiction scenario, for optimistic creators of thought experiments it is the supreme challenge in the extensive field of the philosophy of mind. I try to argue that the main issue we must address before any attempt to build a conscious device or before we even ask ourselves about its theoretical possibility is to clarify what the human consciousness itself means. This is not an easy task, on the contrary it is the supreme challenge that has attracted endless controversy and debates, more than that, it seems to deliver the philosophical discourse into the desert of the paradoxical thinking. We try to understand better the human consciousness by struggling to design and build conscious systems. The present approach is an ambitious endeavour to analyse, beyond any constraints and limitations, some of the most relevant contributions and arguments established to admit or to reject the machine consciousness paradigm, even if it could be nothing more than just a hypothetical possibility. Knowing what it is like to be a conscious machine could be nothing more than just an intractable and hopeless philosophical problem.*

*Keywords: machine, consciousness, neural networks, mental states, computation*

Ritmul de evoluție și dezvoltare al tehnologiilor este accelerat și deopotrivă deconcertant. Indiferent dacă vorbim despre știința computerelor, informatică, robotică, neuroștiințe, Inteligența Artificială sau despre științele cogniției, ajungem să ne adresăm întrebarea firească dacă, la un moment dat în străduințele noastre, vom ajunge să proiectăm și să construim mașinării înzestrate cu conștiință. Dar oare care sunt trăsăturile pe care un astfel de dispozitiv ar trebui să le dețină? Dacă admitem că o astfel de creație artificială nu este nimic mai mult decât un sistem cognitiv capabil să-și reprezinte și să descrie lumea cu toate trăsăturile sale remarcabile, inclusiv locul său propriu și modul în care funcționează în natură, poate că lucrurile nu sunt deloc departe de a se produce. Dacă însă dăruim unui astfel de sistem și privilegiul de a imagina<sup>1</sup> lumi alternative, scenariii ontologice posibile în care să includă propriul statut epistemic, atunci misiunea noastră întâmpină dificultăți serioase. De asemenea, o mașinărie înzestrată cu privilegiul conștiinței ar trebui să dovedească nu doar capacitatea de interacțiune cu

---

<sup>1</sup> Igor Aleksander, în cartea *How to Build a Mind. Toward Machines with Imagination* (Columbia University Press, 2001) stabilește o corelație între *imaginație* și *conștiință*, găsind că forța conștiinței se concretizează în însăși capacitatea de a imagina pe care creierul, ca dispozitiv evoluat de o complexitate remarcabilă, o deține.

interlocutori umani, dar și modul în care înțelege partenerii de dialog umani sau non-umani. Ar trebui să fie capabilă să acționeze și să reacționeze, totodată să opereze în mod rațional, pe baza unor raționamente și algoritmi decizionali, dar și emoțional, totodată să evalueze alternative aflate în ambele registre (rațional – emoțional) după care să adopte cea mai potrivită decizie într-un mod autonom. Pentru mințile sceptice atribuirea tuturor acestor însușiri, precum și multe altele, unei mașinării pare a fi doar un scenariu de ficțiune irealizabil. Însă pentru creatorii optimiști ai experimentelor de gândire din domeniul extins al filosofiei minții, aceasta pare să fie provocarea cognitivă supremă. Îmi propun să argumentez faptul că principala problemă pe care trebuie să o abordăm înainte de orice încercare de a construi un dispozitiv conștient sau înainte de a ne întreba chiar despre posibilitățile sale teoretice este clarificarea prealabilă a ceea ce înseamnă însăși conștiința umană. Aceasta nu este o sarcină ușoară, dimpotrivă, reprezintă un aspect care a atras controverse și dezbateri nesfârșite, mai mult decât atât, a purtat discursul filosofic în deșertul gândirii paradoxale, acolo unde locuiesc „zombi” sau ființe lipsite de conștiință<sup>2</sup>, totodată liliaci a căror conștiință rămâne un mister insolubil<sup>3</sup>. Atunci când abordăm problematica atât de complexă a conștiinței și a condițiilor ei de posibilitate într-un dispozitiv artificial, riscăm să devenim prizonierii unei circularități: pe de-o parte, ne străduim să înțelegem și să descriem adecvat conștiința umană, pentru a proiecta și, eventual, a construi sisteme cognitive conștiente, pe de altă parte, intenționăm producerea unor astfel de artefacte cu scopul de a ne înțelege mecanismele de funcționare ale propriului sistem cognitiv. Abordarea prezentă, deloc lipsită de riscuri, reprezintă rezultatul eforturilor de a analiza, dincolo de orice constrângeri și limitări, unele dintre cele mai relevante contribuții și argumente existente pentru admiterea sau respingerea paradigmei conștiinței mecanice, chiar dacă aceasta nu ar putea fi decât o posibilitate ipotetică. Este posibil ca, dobândind achiziții despre cum este să fii o mașinărie conștientă, să nu ai totuși nici cea mai vagă idee despre existența vreunei experiențe desfășurate dincolo de învelișul de silicon. Un astfel de scenariu, oricât de plauzibil, este susceptibil să plaseze ceonceptul de „conștiință mecanică” în lungul șir de subiecte filosofice fără o rezolvare satisfăcătoare. Rândurile următoare dobândesc, în această lumină, statutul unei invitații la reflecție cu privire la condițiile de posibilitate ale mașinilor capabile de privilegiul remarcabil al conștiinței.

Demersul proiectării și producerii mașinării conștiente, chiar și sub forma exercițiului mental, presupune clarificarea prealabilă a unui aspect extrem de important, acela al stabilirii semnificației termenului de *conștiință*. Este firesc ca întrebarea cu privire la condițiile de posibilitate ale conștiinței într-un dispozitiv mecanic sau electric (sau de orice altă natură) să fie precedată de descoperirea răspunsului relevant la întrebarea originară, *die Urfrage*, după o expresie heideggeriană: „Ce reprezintă sau ce este conștiința umană?” În ceea ce privește formularea răspunsului la o astfel de interogație, constatăm existența unei serii abundente de perspective și încercări explicative, poziționări teoretice, polemici și controverse mai mult sau

<sup>2</sup> concept sau construct teoretic ce și-a făcut loc în universul de discurs filosofic, caracterizând o ființă identică omului din perspectivă fiziologică (scenariu relativ simplu în privința conceptualizării), însă lipsită în totalitate de experiențe conștiente, de o viață interioară, subiectivă privată. Pentru David J. Chalmers existența unor astfel de creaturi nu este deloc o dificultate, chiar dacă ea este nimic mai mult decât concretizarea unui experiment mental. Pentru Daniel C. Dennett sau Paul M. Churchland însă existența unor astfel de creaturi este absurdă, în măsura în care conștiința nu poate fi disociată de acțiuni precum mersul, vorbirea, alegerea etc.

<sup>3</sup> Thomas Nagel lansează celebra sa întrebare „oare cum este să fii un liliac?”, prin care argumentează imposibilitatea accesului la stările conștiente ale alterității, caracterul eminent subiectiv al experiențelor mentale, în ciuda oricărui eforturi cognitive sau scenarii imaginative posibile.

mai puțin fructuoase<sup>4</sup>. Eforturile comprehensive provin nu doar din sfera filosofiei, vulnerabilă în fața ispitei speculative, ci vin mai ales dinspre teritoriul cu pretenții de exactitate al științelor, indiferent ce denumire poartă: științele cogniției, neurobiologie sau neurofiziologie, biologie evoluționistă, psihologie cognitivă ș.a.m.d. Domenii ale cunoașterii precum cele enumerate anterior își propun să contureze un înțeles adecvat pentru conștiință, termen a cărui mănunchi de ambiguități (*a cluster concept*) rezultă, după David J. Chalmers, din faptul că înglobează, pe lângă numeroasele fenomene distincte, reunite sub denumirea de *awareness*, și ceea ce numim *fenomenality, conscious experience*<sup>5</sup>. Chiar și așa, clarificarea empirică, întemeiată științific, a unor astfel de fenomene, oricât de incomplete și susceptibile de revizuire ar fi, nu ridică dificultăți insurmontabile, de unde și plasarea explicațiilor sub spectrul generic de *the easy problem of consciousness*. Cu adevărat problematică rămâne însă „experiența conștientă”, tradusă prin calitățile pe care le folosim în limbajul descriptiv al stărilor noastre interioare<sup>6</sup> atunci când spunem, de pildă, că avem senzația unei culori oarecare, când descriem o suferință sau împărtășim o bucurie, când ne amintim un eveniment din trecut sau navigăm cu ajutorul imaginației în viitor. Înțelegerea și descrierea stărilor ce însoțesc trăirile noastre reprezintă cel de-al doilea aspect problematic al conștiinței, *the hard problem of consciousness*. Întrebarea privitoare la posibilitatea unei mașini conștiente vizează mai curând aspectul problematic al fenomenalității, adică al experiențelor noastre subiective și a unui eventual transfer al acestora într-un sistem cognitiv produs artificial. David J. Chalmers inventariază modalitățile de abordare a unei astfel de probleme dificile. Prima strategie este ignorarea tezei, eschiva, explicarea altui aspect referitor la conștiință. Cei care asumă o astfel de strategie consideră că este mult prea dificil să traduci în limbaj științific o astfel de „realitate”. Cea de-a doua strategie constă în negarea realității conștiinței, problema „experienței” fiind înlocuită cu aspecte observabile, precum accesibilitatea stărilor mentale, tipurile de descriere a acestor stări etc. Se neagă realitatea experienței conștiente pe baza criteriului clasic al verificabilității empirice: există doar ceea ce pot verifica cu ajutorul unor instrumente de observație rigurose concepute. O a treia strategie rezidă în pretenția de a explica deplin sau exhaustiv fenomenul conștiinței, chiar dacă se abandonează perspectiva funcționalistă<sup>7</sup>, totuși și o astfel de modalitate sfârșește în ceea ce pare a fi un scenariu magic. A patra abordare, plauzibilă, dar limitată, constă în explicarea structurii experienței. Operând un paralelism între structurile diferitelor sisteme cognitive, precum sistemul vizual și structura experienței, cercetătorii din acest areal nu spun ce ar trebui să spună de fapt: cum se produce experiența conștientă? Cel de-al cincilea mod prin care se urmărește „soluționarea” problemei dificile a conștiinței este acela care constă în izolarea substratului

<sup>4</sup> Există în spațiul filosofic, cu privire la conștiință, o abundență deconcertantă de „-isme”: dualism, materialism sau fizicalism, interacționism, epifenomenalism, funcționalism ș.a., fiecare orientare având adepții și contestatarii ei.

<sup>5</sup> David J. Chalmers inventariază câteva dintre fenomenele reunite sub denumirea de *conștiință*, în articolul „The Hard Problem of Consciousness”, publicat în *The Blackwell Companion to Consciousness*, volum editat de Max Velmans și Susan L. Schneider, Wiley-Blackwell, 2007, pp.225-235: abilitatea de a discrimina, categoriza și de a reacționa la stimulii mediului extern, intergrarea informației într-un sistem cognitiv, capacitatea de a raporta (descrie) stările mentale, abilitatea sistemului de a accesa propriile stări mentale, concentrarea atenției, controlul deliberat al comportamentului, abilitatea de a diferenția între starea de veghe și aceea de somn.

<sup>6</sup> *qualia* sau *calitățile subiective* inefabile ale experienței noastre, *felul în care lucrurile ne apar nouă*, precum felul în care ni se arată un obiect, ceea ce simțim atunci când vedem o anumită culoare, „proprietățile experienței conștiente” (cf. Daniel C. Dennett, „Quining Qualia”, în *Philosophy of Mind. Classical and Contemporary Readings*, volum editat de David J. Chalmers, Oxford University Press, 2002, p.226).

<sup>7</sup> funcționalismul abordează conștiința nu ca pe un adaos extern, un ingredient suplimentar la înzestrările noastre – percepția, gândirea, memoria, afectivitatea etc., ci mai curând ca pe o funcție cerebrală, un proces intrinsec, inseparabil de creierul nostru ajuns la un anumit nivel de dezvoltare și complexitate.

experienței: știm că experiența apare din anumite procese cerebrale, ne propunem chiar să le identificăm și să le descriem cu ajutorul instrumentelor oferite de știință și tehnologie, însă nu dispunem de o explicație rezonabilă cu privire la *cum* și *de ce* survine experiența. Inventarierea unor astfel de „soluții” la problema dificilă a conștiinței este în măsură să ofere reperi valoroase atunci când ne propunem să investigăm posibilitatea existenței conștiinței într-un artefact. Conștiința poate fi livrată cu ușurință în zona misterelor supreme, considerând-o o enigmă de nedezlegat rezultată din limitările propriului aparat cognitiv, imposibil de explicat prin raportare la dispozitivele electrice sau mecanice ce constituie un computer<sup>8</sup>. Încurajate de progresul cercetărilor din sfera neurobiologiei, există voci care admit conștiința ca fiind nimic altceva decât un nume convențional pentru o suită de procese informaționale desfășurate la nivelul creierului. Astfel de cercetări dăruiesc studiului conștiinței un suflu nou de optimism epistemologic, invitându-ne să admitem că anumite procese din creier – oricât de greu ar fi de identificat – reprezintă sursa experiențelor noastre subiective, de factură non-fizică sau corporală. Totuși rămâne provocatoare probarea modului în care conștiința survine, fie ca funcție corelată a neuronilor asociați în rețele<sup>9</sup>, fie ca proces derulat în structurile informaționale de profunzime ale neuronilor, în microtuburile despre care teoretizează mecanica cuantică<sup>10</sup>. Există și voci care, negând caracterul misterios al conștiinței, dimpotrivă, îi conferă acesteia un rol firesc în scenariul evolutiv, ca particularitate sau fenomen biologic prezent în organismele vii ajunse la un anumit nivel de organizare a activității neuronale. Dintr-un astfel de unghi, conștiința este un fenomen natural explicabil în aceeași manieră în care demersul științific reușește să explice metabolismul, reproducerea, deriva continentală, lumina, gravitația ș.a.<sup>11</sup> Indiferent de suita de poziționări filosofice sau de abordări ce leagă filosofia de științele cogniției, biologia evoluționistă sau neurobiologie, de mecanica cuantică sau de alte domenii epistemice, conștiința rămâne o problemă filosofică extrem de controversată, de nerezolvată de pe pozițiile unui punct de vedere reduționist, fie el materialist sau idealist.

Pentru a evolua în direcția construcției unei mașini conștiente este necesară adoptarea premisei izomorfismului dintre creier și o mașină de calcul inteligentă. Creierul reprezintă fără îndoială un dispozitiv biologic fascinant, pe care cercetătorii din sfera tehnologiei, în special aceia din sfera Inteligenței Artificiale, îl adoptă ca model pentru posibilitatea construcției a ceea ce se dorește a fi o mașinărie conștientă. O astfel de abordare este relativ nouă, dobândind chiar și un nume: *machine model of consciousness* (MMC), un demers multidisciplinar ce înglobează, pe lângă filosofie, contribuții remarcabile din sfera psihologiei, neuroștiinței, informaticii și matematicii<sup>12</sup>. Construcția unei mașini înzestrate cu conștiință are ca scop nu doar să aducă cercetările din sfera Inteligenței Artificiale cu un pas mai departe, ci să aducă o lumină mai clară

<sup>8</sup> John R. Searle, *The Mystery of Consciousness* (The New York Review of Books, 1997), Colin McGinn, *The Mysterious Flame. Conscious Mind in a Material World* (Basic Books, 1999).

<sup>9</sup> articolul „Towards a neurobiological theory of consciousness” (*The Neurosciences*, Vol.2, 1990: pp 263-275), despre *corelatele neurale ale conștiinței*, semnat de Francis Crick și Christof Koch.

<sup>10</sup> modelul mecanicii cuantice (propus de gânditori precum Henry P. Stapp, Roger Penrose) mută discursul explicativ referitor la conștiință la nivel subatomic, în structurile de profunzime ale neuronului, acolo unde se găsesc fibrele de proteine, numite microtuburi, în interiorul cărora au loc vibrații ce pot adopta o „superpoziție cuantică” (*quantum superposition*), fenomene pentru a căror existență nu există însă suficiente dovezi. Dualismul cuantic se traduce prin aceea că evenimentele ce au loc în creierul nostru sunt perechi deopotrivă fizice și mentale, aflate într-o conexiune dinamică, procesuală.

<sup>11</sup> Daniel C. Dennett, *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness* (MIT Press, 2015).

<sup>12</sup> scopul principal al unui astfel de demers este acela ca prin intermediul construcției unei mașini inteligente să poată „captura” și conștiința (Igor Aleksander, „Machine Consciousness”, în *The Blackwell Companion to Consciousness*, ed.cit., p.90.)

în clasică problemă a înțelegerii conștiinței umane. Față de dispozitivele inteligente, o astfel de mașină ar trebui să îndeplinească o serie de criterii<sup>13</sup>, cum ar fi: dețină o mai mare autonomie în raport cu programatorul, să dobândească libertate de mișcare față de programarea prealabilă, să-și reprezinte propriul rol la nivelul mediului înconjurător, să acționeze pe baza unei activități interioare, contemplative, mai mult decât să reacționeze în urma consultării unei liste de opțiuni sau să acționeze contingent, întemeiat pe asocierea de situații preexistente stocate într-o memorie. Dacă astfel de cerințe nu sunt îndeplinite, atunci dispozitivul, oricât de inteligent, nu poate fi etichetat ca fiind unul conștient. Deși atribuim eticheta de inteligență altor specii, găsim de cuvânt să fim sceptici atunci când trebuie să etichetăm ca fiind conștient un artefact sau o mașină pe care o construim. Scenariile și experimentele mentale o apucă în acest sens cu mult înainte realității. Se cuvine să diferențiem, pe baza criteriilor menționate mai sus, între mașina înzestrată cu inteligență și cea căreia îi atribuim privilegiul conștiinței, în măsura în care cele două proprietăți nu se suprapun. O posibilă dificultate în fața proiectării și dezvoltării unei mașinării conștiente ar fi posibilitatea auto-reflexivității sau a monitorizării propriilor procese interne. Ar însemna ca aceste mașini să fie capabile să se scaneze pe ele însele, gândirea care se gândește pe sine ca gândire, cu o expresie consacrată, fapt ce pare paradoxal în cazul unui dispozitiv. Ca sistem biologic cognitiv, creierul este conștient în măsura în care este capabil să se auto-monitorizeze, să se analizeze pe sine. Ca și cum s-ar oglindi, s-ar „privi” din afară și, astfel, ar înțelege propriile mecanisme, operații, structuri, moduri de acțiune și de adoptare a deciziilor etc. Totuși, așa cum admite Drew McDermott, în lucrarea *Mind and Mechanism*, „Dacă un computer se modelează pe sine precum ar modela o piesă de mobilier, modelul ar putea fi destul de detaliat și totuși să nu ducă la conștiință.”<sup>14</sup>. Condițiile necesare unui sistem pentru dobândirea atributelor asociate cu conștiința – precum deținerea de un subsistem senzorial-perceptiv ce-i permite să acceseze mediul înconjurător, capacitatea de a adopta decizii în mod autonom, posesia sentimentului propriei identități etc. – se pare că sunt necesare dar nu și suficiente. Oricât de dificilă ar fi implementarea unor astfel de atribute la nivelul unui dispozitiv tehnic, să admitem că aceasta ar fi posibilă. Dificultatea nu constă neapărat în auto-referențialitate – a face din funcționarea proprie un obiect al analizei și interpretării, ci mai degrabă în abilitatea de a decide dacă merită sau nu o astfel de replicare și de a utiliza rezultatele. „Este ușor să creezi un program de computer care reflectează la propria funcționare, la fel de ușor cu a face o plăcintă. (...) Însă majoritatea programelor sunt concentrate pe o arie limitată de sarcini, fapt care le permite să se ignore pe ele însele.”<sup>15</sup>. Ceea ce este mai provocator decât inserarea unor astfel de funcții – percepție, reprezentare, gândire, memorie, adoptarea de decizii, acțiune și reacție în raport cu condițiile din mediul extern etc. – este faptul de a face dovada prezenței unor stări ce formează registrul complex al conștiinței. Altfel spus, de a face dovada stărilor subiective ale experienței conștiente. Aici, dificultățile înregistrate deja în universul de discurs al umanului împiedică transferul conștiinței către domeniul artefactelor, ca o consecință inevitabilă.

Dacă ne plasăm în universul de discurs a ceea ce poartă denumirea de *computational models of the consciousness*, atunci se cuvine să operăm o distincție între ceea ce înseamnă o mașină înzestrată cu inteligență artificială și o mașină conștientă. Un dispozitiv capabil să imagineze lucruri care la un moment dat nu sunt prezente în câmpul percepției, totodată să își

<sup>13</sup> Igor Aleksander inventariază criteriile de atribuire a conștiinței unei mașini inteligente, în articolul „Machine Consciousness”, din volumul *The Blackwell Companion to Consciousness* (editori: Max Velmans și Susan L. Schneider, Blackwell Publishing, 2002, pp.87-98).

<sup>14</sup> Drew McDermott, *Mind and Mechanism*, MIT Press, 2001, p.143.

<sup>15</sup> *Ibidem*, p.143.

reprezintă lucruri fără a apela la registrul experiențelor anterioare, poate avea statutul de mașinărie conștientă<sup>16</sup>. Dacă este să analizăm criteriile sau, altfel spus, condițiile pe care ar trebui să le îndeplinească un dispozitiv mecanic sau electronic (artificial) pentru a fi considerat conștient, ar trebui ca acesta să dovedească posesia unor înzestrări suplimentare. De pildă, Igor Aleksander admite că un mecanism, pentru a deține proprietatea conștiinței, trebuie să dispună de o înțelegere suficientă a unui interlocutor uman, să dea dovadă de raționare reactivă, contemplativă, să fie capabilă de mecanisme de nivel scăzut, de genul celor existente într-un organism biologic, să fie capabilă să-și reprezinte lumea înconjurătoare, în multitudinea de însușiri pe care aceasta o deține, cu toate celelalte organisme ce o populează, reprezentare a cărei probare rămâne deocamdată problematică. Două aspecte se cuvin semnalate în acest punct: este necesară în ecuația mașinii conștiente existența unui corp? Sau este suficient să gândim conștiința într-un mecanism mai mult sub forma unui *software*, ceea ce exclude condiția corporalității ca suport? Dacă admitem această asumție, atunci putem transfera (*download*) conștiința fără probleme în orice dispozitiv cognitiv, indiferent de natura, structura sau calitățile acestuia. Un astfel de scenariu frizează însă deocamdată teritoriul literaturii științifico-fantastice.

Dacă ar fi să concluzionăm, constatăm că, de-o parte a versantului unui proiect îndrăzneț intitulat „mașinăria conștientă”, se situează optimiștii. Sunt cei care admit că este posibil să existe, la un moment dat, conștiință în interiorul artefactului, chiar dacă, interogându-i cum vor proceda, nu reușesc să ne convingă cu privire la modul în care se produce experiența conștientă la un astfel de nivel. Există autori care se străduiesc să arate că, pentru a exista conștiință, nu este cazul să existe informații care pătrund în aparatul senzorial și acțiunile pe care acestea le provoacă, atenție selectivă, emoții, memorie, auto-reflecție, limbaj, experimentarea lumii înconjurătoare<sup>17</sup>. De altă parte a versantului conștiinței, se găsesc pesimiștii, cei care deconstruiesc o astfel de asumție. Chiar dacă admitem ceea ce adepții optimiști ai singularității cred, anume că în câteva decenii se va construi o mașină care să experimenteze subiectiv ceea ce noi resimțim și plasăm sub semnul conștiinței, totuși problema modului în care vom putea face dovada arhitecturii unei astfel de lumi interioare, încapsulată în artefact, va fi destul de greu dacă nu chiar imposibil de soluționat. Ne găsim într-un impas filosofic similar celui de altădată, atunci când am admis imposibilitatea de a resimți ceea ce simte un liliac sau o altă ființă căreia i-am oferit cu îndrăzneală statutul de ființă conștientă. Impasul se datorează existenței hiatus-ului cognitiv (*the explanatory gap*) între ceea ce resimțim noi și ceea ce experimentează obiectul contemplației noastre. Această neputință ne-a făcut să căutăm tot felul de artificii conceptuale reduționiste prin care fie am negat existența unor astfel de experiențe subiective, fie le-am considerat funcții ale unui dispozitiv biologic, traduceri ale unor algoritmi sau construcții semantice. Tot astfel, este dificil dacă nu chiar imposibil să probăm faptul că acolo, dincolo de carcasa metalică, în structurile de siliciu sau chiar în țesătura biologică pe care sperăm să o reproducem în laborator, poate locui cineva, o entitate menită a proba existența unei vieți interioare. Stările interne ale unei astfel de mașinării riscă să rămână pentru noi private, principial inaccesibile, și chiar dacă am putea să le observăm mișcărilor, reacțiile, comportamentul, tot ne-ar fi dificil să stabilim o legătură cauzală între conduita lor observabilă și stările lor interioare sau activitatea lor mentală. Asemănarea, transferul imaginar pe care îl folosim ca suport al raționamentului nostru – când ne propunem să justificăm prezența stărilor interioare, a fluxului conștiinței la ceilalți oameni – nu funcționează sau, cel mult, își vedește limitele în cazul în care Celălalt reprezintă un artefact sau o mașinărie.

<sup>16</sup> Igor Aleksander, *How to Build a Mind. Toward Machines with Imagination*. Columbia University Press, 2001.

<sup>17</sup> <https://spectrum.ieee.org/biomedical/imaging/can-machines-be-conscious>

Poate că răspunsurile nu ar trebui radicalizate, mai curând ar fi cazul să admitem un spectru menit a ne provoca să concepem lucrurile într-o manieră nuanțată. La nivel principial, ca experiment de gândire valoros prin consecințele pe care le declanșează, conștiința poate fi reprodusă mecanic în interiorul unei mașinării. Aceasta numai în condițiile în care concepem conștiința în manieră biologică, behavioristă sau computațională, ca inventar al unor fenomene, procese sau funcții ce pot fi studiate cu ajutorul instrumentarului generos al științei. Astfel de fenomene sunt cele despre care ne vorbește David J. Chalmers, atunci când disociază între „problema ușoară” și „problema dificilă”, insurmontabilă a conștiinței. Dacă concepem conștiința drept o succesiune de operații, algoritmi, calcule de factură logico-matematică, mai puțin un proces ce funcționează după regularitățile biologiei, ale fizicii și chimiei, lucrurile se simplifică iar posibilitatea unui răspuns afirmativ la condițiile de posibilitate ale unei mașinării conștiente dobândește virtuțile unui scenariu cel puțin plauzibil. Dacă însă așezăm între paranteze conștiința ca proces, funcție sau aspect corelativ neuronal-biologic, totodată ca sumă de algoritmi traductibili în limbajul informatic-computațional și, în paralel cu aceasta, admitem în universul nostru de discurs un termen provocator precum cel de *experiență*, atunci consecințele sunt diferite. Ca sumă de înzestrări achiziționate în cadrul unui scenariu evolutiv – gândire, rezolvarea de probleme etc. – creierul poate să nu fie ultimul „adăpost” al conștiinței și, implicit, să o putem transfera în alt „ambalaj”, fie el și unul din cabluri sau circuite de siliciu. Înțelegerea conștiinței drept *fenomenalitate*, *subiectivitate* ridică serioase dificultăți în fața proiectării și construcției unei mașinării conștiente. În acest caz, se cuvine să ne consolăm cu gândul că un computer conștient, o mașinărie înzestrată cu experiențe și stări interne, rămâne deocamdată o problemă de imaginație a cărei oportunitate și utilitate în plan filosofic sau științific se cuvine stabilită prin dialog prolific și schimb de idei.

## BIBLIOGRAPHY

- Aleksander, Igor. (2001). *How to Build a Mind. Towards Machines with Imagination*, Columbia University Press.
- Aleksander, Igor. (1996). *Impossible Minds. My Neurons, My Consciousness*, Imperial College Press.
- Bijl, Aart. (1995). *Ourselves and Computers. Differences in Minds and Machines*. MacMillan Education.
- Block, Ned; Flanagan, Owen; Guzeldere, Guven. (ed.). (1997). *The Nature of Consciousness*. Massachusetts, MIT Press.
- Crane, Tim. (2003). *The Mechanical Mind. A Philosophical Introduction to Minds, Machines and Mental Representations*. London, New York, Routledge Taylor & Francis Group.
- Fetzer, H. James. (2001). *Computers and Cognition. Why Minds are not Machines*. Dordrecht, Netherlands, Springer-Science+Business Media.
- Husbands, Philip; Holland, Owen; Wheeler, Michael. (ed.). (2008). *The Mechanical Mind in History*. Massachusetts, The MIT Press.
- Nagel, Thomas. 1994. *Oare ce înseamnă toate astea?* București: All Educational.
- Pagel, J.F.; Kirshtein, Philip. (2017). *Machine Dreaming and Consciousness*. Academic Press, Elsevier Inc.
- Taylor, John. (1999). *The Race for Consciousness*, Massachusetts, The MIT Press.

**I.Boldea, C. Sigmirean, D.-M.Buda**

***THE CHALLENGES OF COMMUNICATION. Contexts and Strategies in the World of Globalism***

- Velmans, Max; Schneider, Susan L. (ed.). (2007). *The Blackwell Companion to Consciousness*, Willey-Blackwell.
- Watson, Ian. (2012). *The Universal Machine. From the Dawn of Computing to Digital Consciousness*. Berlin, Heidelberg, Springer Verlag.

**Resursă internet:**

<https://spectrum.ieee.org/biomedical/imaging/can-machines-be-conscious>