

ESTE MINTEA UMANĂ COMPLET ALGORITMICĂ? REMARCI ASUPRA TEOREMELOR DE INCOMPLETITUDINE ALE LUI GÖDEL¹

MIRCEA DUMITRU

Filosofia minții înflorește astăzi. Domeniul s-a dezvoltat extensiv și intensiv, beneficiind de multe genuri de input-uri venite din domenii conexe, în mod special din informatica teoretică și din științele cognitive. Complexitatea și vitalitatea domeniului sunt reflectate de către literatura vastă, care se ramifică în diferite sub-domenii și direcții de cercetare, în care sunt abordate mai multe teme corelate: problema ontologică (așa-numita problemă tradițională a relației dintre minte și corp), problema semantică, problema epistemologică, problema metodologică, inteligența artificială și problemele neuroștiințelor.

În studiul meu, mă voi ocupa de o chestiune din filosofia minții care se află în relație cu filosofia matematicii și, mai specific, cu natura cunoașterii și a raționamentelor matematice. Chestiunea este dacă mintea umană și conștiința inteligentă sunt în mod complet algoritmice. Voi elabora un răspuns, având drept fundal celebrele teoreme de incompletitudine ale lui Gödel. În ceea ce urmează: (i) voi schița, mai întâi, principalele programe și răspunsuri elaborate în legătură cu problema minte-corp din filosofia minții; (ii) apoi, voi furniza o prezentare semi-formală teoremelor de incompletitudine ale lui Gödel; (iii) finalmente, voi prezenta și comenta unele dintre opiniile și concepțiile lui Gödel despre minți și mașini, minte și materie, și contrastul dintre mașinile Turing și așa-numitele minți Gödel. Pe parcursul acestor comentarii se vor configura în mod pregnant concepțiile foarte neortodoxe ale lui Gödel împotriva computabilismului, neuralismului, fizicalismului, paralelismului psiho-neuronal, și chiar împotriva presupuzițiilor filosofice subiacente mașinilor Turing. Indiferent cât de șocante și de ciudate ar putea să ne apară azi, în mod rezonabil, concepțiile extrem de originale ale lui Gödel, prin comparație cu concepțiile ortodoxe standard psihologice și filosofice, care sunt subiacente concepțiilor computaționiste standard despre minte acceptate astăzi în mod curent, ele merită să fie explorate cu atenție și să se bucure pe deplin de întreaga noastră atenție filosofică. Gödel este, la urma urmei, părintele fondator și una dintre cele mai motivante și inspiratoare figuri pentru întregul domeniu de teme și probleme cu care mă ocup în această lucrare. Și chiar dacă, așa cum par să stea

¹ Studiul de față este o traducere, adaptată, în limba română a studiului nostru original, apărut inițial în limba engleză, "Is Human Mind Fully Algorithmic? Remarks on Kurt Gödel's Incompleteness Theorems", în *Evolving Computability. 11th Conference on Computability in Europe*, editori Arnold Beckmann, Victor Mitrana, Mariya Soskova, Springer, 2015, p. 23–33.

lucrurile, nu vom accepta, și probabil nici nu vom *putea* accepta tot ceea ce spune Gödel despre aceste chestiuni, tot avem destul de învățat de la modul în care formulează el întrebările și de la modul în care s-a pronunțat cu privire la toate aceste chestiuni fascinante privitoare la natura și la funcționarea minții noastre matematice.

1. SCHIȚĂ A PRINCIPALELOR PROGRAME ȘI RĂSPUNSURI LA PROBLEMA MINTE-CORP DIN FILOSOFIA MINȚII

Să schițăm acum principalele poziții din metafizica minții. O resursă foarte utilă pentru această temă este (Churchland, 2013), pe care o urmez aici pentru a sistematiza principalele răspunsuri filosofice la problema minte-corp. Ceea ce ne dorim este să clarificăm problema naturii stărilor și proceselor mentale. Și, mai specific, chestiunile pe care le ridicăm sunt: de unde provin stările și procesele mentale și cum sunt acestea corelate cu lumea fizică? Vor supraviețui stările mele conștiente descompunerii mele fizice, după ce voi muri? Este posibil ca un sistem pur fizic (un computer) să fie construit în așa fel încât să aibă experiențe conștiente cu stări calitative distincte (*qualia*)? De unde provin mințile? Ce sunt acestea?

Răspunsurile argumentate la aceste întrebări dificile sunt condiționate de teoria și metodologia împărtășite. Ele sunt dependente de teoria minții pe care o îmbrățișăm, care depinde, la rândul ei, de puterea explicativă și predictivă a teoriei și, totodată, de coerența și simplitatea acesteia. Principalele teorii care au fost susținute în filosofia minții sunt:

Dualismul. Esența tuturor formelor de dualism, așa cum sunt dualismul substanțelor și dualismul proprietăților, constă în ideea că natura stărilor mentale rezidă într-o entitate non-fizică mentală sau spirituală, care se sustrage domeniului fizicii, neurofiziologiei sau al informaticii. Azi, dualismul trece prin momente paradoxale. Indubitabil, cea mai populară și tradițională perspectivă asupra minții timp de milenii, similară poziției argumentate de către diferitele religii majore cu privire la relația dintre minte, suflet și corp, dualismul este aproape cu desăvârșire respins de către filozofii profesioniști contemporani.

Behaviorismul filosofic. Aceasta a fost o concepție extrem de influentă în metafizica minții timp de mai multe decenii în sec. al XX-lea. Ascensiunea cognitivismului în lingvistică și în psihologie a condus la abandonarea treptată a acestei poziții, care a fost extrem de influentă, nu în ultimul rând ca instrument de analiză critică a diverselor programe metafizice speculative. Ca atare, behaviorismul filosofic nu este o teorie despre esența stărilor mentale *per se*; mai degrabă, este un gen de analiză a limbajului în care raportăm și vorbim despre stări mentale. Astfel, propozițiile limbajelor noastre care vizează diferite episoade mentale, așa cum ar fi emoțiile, senzațiile, credințele, dorințele, voințele etc. nu vizează prezumtive ocurențe interne ale unor evenimente sau procese mentale, ci, mai degrabă, ele sunt modalități abreviate sau prescurtate de a vorbi despre comportamente ocurențe în mod real sau actual sau despre comportamente doar posibile. Prin urmare, orice propoziție despre o stare sau un proces mental poate fi reformulată sau parafrazată ca o propoziție mai lungă și mai complexă despre comportament.

Materialismul reducționist (teoria identității). Principala aserțiune a acestei forme de teorie materialistă este aceea că stările mentale sunt (identice cu) stări fizice ale creierului. Mai specific, fiecare tip (sau instanță, în versiuni mai slabe ale acestei teorii) de stare sau proces mental este identic numeric cu un anumit tip (sau instanță) de stare sau proces fizic care se produce în creier sau în sistemul nervos central.

Funcționalismul. Această doctrină, care este concepția dominantă azi asupra minții, spune că trăsătura caracteristică a oricărui tip de stare mentală este setul de relații cauzale pe care acel tip de stare mentală le susține cu (i) inputul care provine din mediul înconjurător, (ii) cu alte tipuri de stări mentale și (iii) cu outputul comportamentului nostru. O stare mentală joacă un rol cauzal și acea stare mentală este definită prin intermediul rețelei sale de roluri cauzale. Potrivit funcționalismului, spre deosebire de behaviorismul filosofic, nu se poate elimina total orice referire la stările mentale; și, pentru a defini un astfel de tip de stare mentală, trebuie să ne referim la un număr de alte stări mentale cu care acea stare în cauză este conectată cauzal. Funcționalismul recunoaște realitatea stărilor mentale care trebuie studiate în mod sistematic. Decurge de aici că psihologia trebuie să fie o știință autonomă față de și ireductibilă la științele fizice (fizică, biologie, neurofiziologie). Psihologia este o știință de sine stătătoare, cu propriile sale legi ireductibile și cu propriul său domeniu de studiu.

Materialismul eliminativist. Aceasta este o concepție profund sceptică cu privire la minte. Se îndoiește radical de conceptele și explicațiile psihologiei simțului comun (care explică acțiunile inteligente ale ființelor umane în termenii puterilor cauzale ale atribuirilor de atitudini propoziționale, cum ar fi credințele, dorințele, speranțele etc.). Materialismul eliminativist se opune și materialismului reducționist, din moment ce o parte definitorie a programului reducționist este aceea de a obține o corespondență unu-la-unu între stările și procesele mentale recunoscute de către psihologia simțului comun și anumite procese neurofiziologice care apar în creier. Această reducere intenționată nu poate fi efectuată. Iar motivul nu este lipsa de ingeniozitate din partea teoreticienilor. Temeiul adânc este non-existența unor astfel de entități precum stările, procesele sau atitudinile mentale care sunt postulate de către cadrul conceptual al psihologiei simțului comun și care, la rândul său, este un gen de ipoteză la care psihologia simțului comun ajunge printr-o inferență în acord cu cea mai bună explicație. Prin urmare, unul dintre motivele cruciale pentru imposibilitatea acestei reducții rezidă în faptul că psihologia simțului comun este confruntată cu unele probleme fatale: ea este literalmente falsă și, în consecință, ea este și o concepție înșelătoare asupra a ce anume determină din punct de vedere cauzal comportamentul nostru și activitatea noastră mentală. Prin intermediul educației științifice, ne așteptăm în mod rațional și putem spera ca treptat să ne putem debarasa de această falsă reprezentare cu privire la propria noastră psihologie, bazată pe conceptul de atribuire de atitudini propoziționale. Acest cadru explicativ va fi eliminat de către descoperirile viitoare din domeniul neuroștiințelor cognitive. Avem aici, totodată, motivația pentru denumirea acestei concepții, anume „materialismul eliminativist”.

O idee generală, care emerge din diferitele răspunsuri la problema minte-corp și care este în mod esențial un laitmotiv pentru dialectica în desfășurare a acestor întrebări și răspunsuri, este aceea că ceea ce explică cel mai bine stările și procesele mentale este o

paradigmă computațională a minții. Mulți teoreticieni argumentează că mintea este un gen de calcul cu simboluri și cu conținuturi mentale reprezentationale. Această paradigmă computațională ne va furniza un răspuns coerent la dificila problemă a integrării a două viziuni distincte despre natura ființelor umane: (i) concepția determinist-cauzalistă, care insistă asupra explicației structurilor complexe bio-chimice din care constau ființele umane în calitatea lor de entități biologice, și (ii) concepția intenționalității reprezentărilor mentale, care insistă asupra ființelor umane în calitatea lor de agenți decidenți raționali și socio-culturali. Speranța actuală este aceea că funcționalismul va rezolva această problemă a integrării celor două concepții despre natura ființelor umane. Versiunea cea mai influentă a funcționalismului contemporan consideră că o teorie computațională a minții este cea mai bună explicație disponibilă a comportamentului uman prin intermediul rolului cauzal al stărilor mentale de a media, în termeni computaționali explițiți, între inputul mediului înconjurător și outputul comportamental. Vom vedea imediat că Gödel respinge atât această viziune asupra minții cât și presuposițiile ei.

2. O SUCCINTĂ PREZENTARE SEMI-FORMALĂ A TEOREMELOR DE INCOMPLETITUDINE ALE LUI GÖDEL

În linii mari și într-o redare semi-formală, ceea ce ne arată Prima Teoremă de Incompletitudine a lui Gödel este că orice sistem axiomatic sau sistem deductiv oarecare consistent T , care este corect (i.e. demonstrează numai propoziții adevărate) și care este suficient de puternic pentru a exprima aritmetica elementară, este cu necesitate incomplet, deoarece o propoziție, pe care o vom denumi G_T , poate fi adevărată corespunzător interpretării aceluia sistem formal T , dar nu poate fi derivată ca teoremă în acel sistem.

Astfel, Gödel arată că ideea comună, potrivit căreia adevărul aritmetic rezidă într-o demonstrație în cadrul unui sistem deductiv formal, este greșită. Gödel a demonstrat această idee revoluționară prin intermediul unei serii de pași ingenioși. Mai întâi, el a construit o propoziție G_T , în limbajul aritmeticii (grație tehnicii concepute de el și pe care azi o numim numărătoare-Gödel), care reprezintă propoziția meta-matematică pe care o decodificăm ca având semnificația: „Propoziția G_T nu este demonstrabilă în sistemul T ”. Adică, G_T spune despre ea însăși că este nedemonstrabilă în T . Decurge că G_T este adevărată dacă și numai dacă (ddacă) G_T nu poate fi demonstrată în T . Să presupunem, mai departe, că T este corect. Dacă G_T ar fi demonstrabilă în T , atunci G_T ar fi falsă și, așadar, nedemonstrabilă în T , deoarece T este un sistem corect și poate demonstra numai propoziții adevărate. Deci, până în acest punct, dacă G_T ar fi demonstrabilă, atunci nu poate fi demonstrată. Prin urmare, G_T nu este demonstrabilă în T , la urma urmei, ceea ce înseamnă că G_T este adevărată. Să presupunem, acum, că G_T nu ar fi demonstrabilă. Atunci G_T este adevărată și, desigur, negația ei, $\sim G_T$, este falsă. Dar T este corect și demonstrează numai propoziții adevărate. Astfel, T nu poate să o demonstreze nici pe $\sim G_T$. Așadar, există o propoziție adevărată, G_T , care spune despre ea însăși că nu este demonstrabilă într-un sistem T și nici acea propoziție, G_T , nici negația ei, $\sim G_T$, nu este demonstrabilă în T . Deci, propoziția G_T este indecidabilă, prin mijloacele sistemului T și, presupunând că T este corect, sistemul T este incomplet.

Adăugarea propoziției G_T la sistemul T nu rezolvă această problemă, deoarece, potrivit aceleiași metode, o nouă propoziție G'_T poate fi construită, în așa fel încât să poată spune despre ea însăși că nu este demonstrabilă în sistemul $T + G_T$, în timp ce propoziția G'_T este adevărată și nici propoziția în cauză, G'_T , nici negația ei, $\sim G'_T$, nu este demonstrabilă în sistemul $T + G_T$.

Până aici, Gödel a arătat că, deoarece G_T este adevărată, dar nedemonstrabilă în T , axiomele sistemului T sunt incomplete. Rezumând această parte a demonstrației, care are drept punct culminant Prima Teoremă de Incompletitudine a lui Gödel, Nagel & Newman în (Nagel și Newman, 1958, p. 67) argumentează convingător că “nu putem deduce toate adevărurile aritmetice din axiome. Mai mult, Gödel a stabilit că aritmetica este în mod esențial incompletă: chiar dacă am asuma axiome suplimentare, în așa fel încât formula adevărată G_T să poată fi derivată formal din mulțimea augmentată de axiome, o altă formulă adevărată, dar nedecidabilă poate fi construită”.

În cea de-a Doua Teoremă de Incompletitudine, Gödel arată cum să construim o propoziție aritmetică A , care are conținutul meta-matematic: „Aritmetica este consistentă”. El demonstrează, apoi, că propoziția „ $A \rightarrow G_T$ ” este demonstrabilă; totuși, din moment ce G_T însăși nu este ea însăși demonstrabilă, Gödel arată că nici A nu este demonstrabilă. Ceea ce decurge din toate acestea este A Doua Teoremă de Incompletitudine care stabilește faptul că consistența aritmeticii nu poate fi stabilită printr-un argument care să poată fi reprezentat în chiar același calcul aritmetic formal (cf. Nagel și Newman, 1958, p. 67).

Dar au toate aceste rezultate meta-matematic fundamentale vreo semnificație filosofică? Și dacă au, care ar fi acea semnificație? Din uriașa literatură care conține reacții și comentarii declanșate de Teoremele de Incompletitudine ale lui Gödel, voi examina punctele de vedere exprimate chiar de Gödel cu privire la propriile sale rezultate și voi face comentarii asupra a trei chestiuni care vizează (a) concepția lui Gödel cu privire la minți, mașini și computabilism, (b) concepția lui Gödel cu privire la minte, materie, fizicalism și paralelismul psiho-fizic, și (c) Gödel vs. Turing, adică concepția lui Gödel despre mașinile Turing și despre mințile Gödel.

3 COMENTARIU ASUPRA UNORA DINTRE IDEILE FILOSOFICE PRINCIPALE SUSȚINUTE DE CĂTRE GÖDEL

Această secțiune a studiului meu se bazează pe (Wang, 1996), care este o resursă extrem de bogată și de fertilă cu privire la concepțiile filosofice ale lui Gödel.

3.1. Gödel despre Minți și Mașini

Gödel nutrește convingerea fermă că nici computabilism, adică ideea potrivit căreia creierul și mintea funcționează, în mod esențial, ca un computer, și nici neuralismul, adică ideea potrivit căreia creierul este un *explanans* suficient pentru fenomenele mentale nu sunt corecte și, drept urmare, el a argumentat viguros împotriva ambelor concepții și le-a respins. Hao Wang, în (Wang, 1996), ne spune că Gödel a fost preocupat de problema dacă computabilismul oferă o explicație completă a proceselor men-

tale, „adică, chestiunea dacă întreaga gândire este computațională – cu un accent special asupra gândirii matematice. Principala preocupare a lui Gödel era aceea de a demonstra că nu toată gândirea matematică este computațională” (Wang, 1996, p. 183).

În una dintre diferitele sale formulări, A Doua Teoremă de Incompletitudine a lui Gödel spune ceva care este relevant pentru capacitatea matematică a minții umane, anume că dacă un computer sau un program care demonstrează teoreme, și care este rezonabil de puternic, este corect (*sound*) și consistent, atunci acel program nu poate să demonstreze adevărul care exprimă propria sa consistență (a programului). Potrivit lui (Wang, 1996), Gödel a extras o concluzie relevantă privitoare la minte din acest rezultat: „6.1.1 Minte umană este incapabilă să formuleze (sau să mecanizeze) toate intuițiile sale matematice. Adică, dacă ea a reușit să formuleze pe unele dintre acestea, chiar acest fapt produce cunoaștere intuitivă nouă, de pildă consistența acestui formalism. Faptul acesta poate fi numit „incompletabilitatea” matematicii. Pe de altă parte, pe baza a ceea ce s-a demonstrat până acum, rămâne posibil să poată exista (și chiar să fie descoperibilă empiric) o mașină de demonstrat teoreme care, în fapt, să fie echivalentă cu intuiția matematică, dar nu poate fi demonstrat că ea este echivalentă cu intuiția matematică și nici nu poate nici măcar să fie demonstrat că mașina va produce numai teoreme corecte ale teoriei numerelor finite” (Wang, 1996, p. 184-185).

Hao Wang, iarăși, ne spune că Gödel era foarte atașat de unele idei cu privire la creația matematică și la caracterul limitat al naturii algoritmice a minții umane și a gândirii matematice (Wang, 1996, p. 186). Aceste idei sunt relevante pentru implicațiile teoremei sale, așa cum se poate lesne observa din următoarea remarcă a lui Gödel: „6.1.8 Teorema mea de incompletitudine face destul de probabil ca mintea să nu fie mecanică sau, altfel, ca mintea să nu poată înțelege propriul ei mecanism. Dacă rezultatul meu este luat împreună cu atitudinea raționalistă pe care a adoptat-o Hilbert și care nu a fost repudiată de către rezultatele mele, atunci [putem infera] rezultatul precis că mintea nu este mecanică. Lucrurile stau așa pentru că, dacă mintea ar fi o mașină, atunci ar exista, contrar acestei atitudini raționaliste, chestiuni teoretice cu privire la numere nedecidabile pentru mintea umană” (Wang, 1996, p. 186-187).

Punctul culminant al tuturor acestor remarci este convingerea puternică a lui Gödel că mintea umană, prin puterile sale intuitive și prin creativitate, este superioară computerelor și că participarea minților individuale la experiența colectivă a speciei umane produce un întreg nou domeniu de posibilități, ceea ce îi permite minții și spiritului uman să depășească puterea mașinilor de calculat. Iată câteva dintre gândurile lui Gödel în această direcție:

„6.1.19 Creierul este o mașină de calcul conectată cu un spirit”.

„6.1.21 Conștiința este conectată cu o unitate. O mașină este compusă din părți”.

„6.1.23 Prin *minte* înțeleg o minte individuală cu un timp de viață nelimitat.”²

Aceasta este, totuși, ceva diferit de mintea colectivă a speciei. Imaginați-vă o persoană angajată în rezolvarea unei mulțimi întregi de probleme: aceasta este ceva apropiat de realitate; oamenii introduc în mod constant noi axiome” (Wang, 1996, p. 189).

² Gödel credea că sufletul uman este nemuritor și că știința va dovedi acest fapt într-o bună zi. Eroul său filosofic era Leibniz.

Fragmentul acesta ne arată cât se poate de grăitor încrederea lui Gödel că, în cele din urmă, putem demonstra superioritatea minții față de computere, grație creativității celei dintâi și a puterii ei de a produce noi idei și intuiții:

„6.1.24 Ar fi un rezultat de mare interes să se demonstreze că cea mai scurtă procedură de decizie cere un interval de timp lung pentru a decide propoziții comparative scurte. Mai specific, poate fi posibil să se demonstreze: Pentru fiecare sistem decidabil și fiecare procedură de decizie pentru acel sistem, există unele propoziții de lungime mai mică decât 200 ale căror cele mai scurte demonstrații sunt mai lungi decât 10^{20} . Un astfel de rezultat ar însemna, de fapt, că computerele nu pot înlocui mintea umană, care poate da demonstrații scurte, furnizând o nouă idee.” (Wang, 1996, p. 189)

3.2. Gödel despre minte și materie

Trasarea unei distincții între minte și materie aduce cu sine ideea metafizică potrivit căreia cele două entități sunt distincte, o doctrină care ne angajează față de o anumită formă de dualism (vezi mai sus). Dificultatea de a le corela (într-o modalitate cauzală sau altfel), din moment ce le-am separat în mod esențial, este ceva de notorietate începând cu opera lui Descartes. Gödel are propria sa modalitate de a formula această celebră chestiune metafizică, anume punând întrebarea dacă “creierul este suficient pentru explicarea tuturor fenomenelor mentale”. Gödel reformulează problema într-un mod mai precis, cantitativ, ridicând problema dacă există suficiente operații cerebrale care să reprezinte operațiile mentale în așa fel încât corespondența dintre creierul fizic și operațiile mentale să fie unu-la-unu sau chiar mai multe-la-unu.

Ortodoxia științifică și filosofică argumentează că există o astfel de corelație, concepție care este cunoscută drept *paralelism psiho-neuronal*. Dacă, mai departe, facem asumția fizicalistă, care este foarte comună astăzi, că toate operațiile neuronale sunt operații fizice de un gen aparte, concepția se transformă într-un *paralelism psiho-fizic*.

Propriul argument al lui Gödel este că mintea este separată de creier (materie). Gödel, contrar întregului establishment științific și ortodoxiei curente, respinge atât paralelismul psiho-neurologic cât și paralelismul psiho-fizic. Observațiile sale cu privire la această temă sunt foarte curajoase și, desigur, șocante pentru mulți dintre noi. Astfel, Gödel spune: „6.2.1 Paralelismul este o prejudecată a timpurilor noastre. 6.2.2 Paralelismul va fi respins (*disproved*) în mod științific (probabil de către faptul că nu există suficiente celule nervoase pentru a realiza operațiile observabile ale minții)” (Wang, 1996, p. 190).

Gödel recunoaște de îndată că nu toate prejudecățile sunt cu necesitate false. O prejudecată este o opinie sau o credință amplu împărtășită a cărei tărie nu este susținută de dovezi solide. De ce susținem cu atâta putere prejudecata paralelismului? O facem, pentru că suntem impresionați de puterea științei și a tehnologiei, ceea ce ne conduce adesea la acceptarea necritică a scientismului. Gödel face următoarea remarcă extraordinară că miezul filosofic al paralelismului în 6.2.2, menționat mai înainte nu este numai o prejudecată filosofică, ci și o pronunțare științifică și empirică ce va fi respinsă. Gödel subliniază această idee ori de câte ori el simte că este important să se facă mai mult loc pentru această idee în spațiul conceptual. El se referă la această noțiune în

următoarele pasaje: „6.2.3 Este o posibilitate logică ca existența minții [separată de materie] să fie o chestiune decidabilă în mod empiric. Această posibilitate nu este o conjectură. [...] în spatele acestei probleme se află o chestiune empirică. 6.2.4 Logica se ocupă cu concepte mai generale; monadologia, care conține legile generale ale biologiei, este mai specifică. Limitele științei: Este posibil ca toate activitățile mentale [...] să fie activități cerebrale? Se poate să existe un răspuns factual la această întrebare. Respingerea ideii că gândirea este o proprietate de o anumită natură cere, de asemenea, respingerea unei anumite concepții despre particulele elementare. Materia și mintea sunt două lucruri diferite. 6.2.5 Simpla posibilitate că se poate să nu existe suficiente celule nervoase pentru a executa funcția minții introduce o componentă empirică în problema minții și a materiei” (Wang, 1996, p. 191).

Gödel accentuează în mod aparte chestiunea metafizică importantă și dificilă a relației dintre minte și materie, considerând-o centrală pentru cercetarea filosofică și critică pentru înțelegerea importanței filosofiei pentru știință. Putem, astfel, să citim următoarea remarcă pe care Gödel o face în conversație cu Hao Wang: “6.2.6 Multe dintre așa-numitele probleme filosofice sunt probleme științifice, atâta doar că nu sunt tratate de către oameni de știință. Un exemplu este dacă mintea este separată de materie. Astfel de probleme ar trebui discutate de către filosofi înainte ca oamenii de știință să fie pregătiți să le discute, așa încât filosofia are drept una dintre funcțiile sale să orienteze cercetarea științifică. O altă funcție a filosofiei este să studieze care este semnificația lumii” (Wang, 1996, p. 191).

Gödel este foarte interesat de clarificarea chestiunii paralelismului dintre minte și materie și exprimă cât se poate de clar poziția pe care o adoptă: “6.2.9 Mintea este separată de materie: este un obiect separat...” Mai mult, Gödel face conjectura foarte curajoasă, mergând complet împotriva spiritului actual al establishmentului științific, că știința însăși va respinge, în cele din urmă, această prejudecată a paralelismului psiho-neuronal: “6.2.11 Cred că mecanismul din biologie³ este o prejudecată a timpului nostru, care va fi respinsă. În acest caz, o respingere bazată pe demonstrație va consta, în opinia mea, într-o teoremă matematică care va avea conținutul că formarea, într-o perioadă de timp geologică, a corpului uman, prin legile fizicii (sau oricare alte legi de un gen similar), pornind de la o distribuție aleatorie a particulelor elementare și a câmpului, este tot atât de puțin probabilă precum separarea din pură întâmplare a atmosferei în elementele sale componente” (Wang, 1996, p. 192).

Și o ultimă remarcă în aceeași direcție. Remarca ne arată convingerea lui Gödel că: (a) creierul este un obiect fizic, (b) mintea (sau spiritul) este o entitate separată de creier și (c) creierul, ca obiect fizic normal, funcționează în modul în care o face pentru că este conectat cu o minte: “6.2.14 Chiar dacă creierul finit nu poate înmagazina o cantitate infinită de informații, spiritul este capabil să facă aceasta. Creierul este o mașină de calcul (*computing machine*) conectată cu un spirit. Dacă creierul se consi-

³ Prin ‘mecanism în biologie’, Hao Wang ne spune că Gödel avea în vedere darwinismul, “pe care el (Gödel, *nota MD*) se pare că îl considera ca pe un set de legi algoritmice (ale evoluției). Chiar dacă pare să creadă că creierul – și putem prezuma, de asemenea, că corpul uman – funcționează ca un computer [...], el pare să ne spună aici că corpul uman este atât de complex încât legile fizicii și ale evoluției sunt insuficiente pentru a explica formarea sa în intervalul de timp estimat în mod obișnuit” (Wang, 1996, p. 192).

deră că este ceva fizic și ca fiind un computer digital, din mecanica cuantică [decurge că] există, atunci, numai un număr finit de stări. Numai conectându-l [creierul] cu un spirit îl poate face să funcționeze într-un alt fel” (Wang, 1996, p. 193).

3.3. Mașini Turing vs. așa-numitele minți Gödel

Gödel a gândit în mod profund asupra naturii algoritmilor și formalizării sistemelor logice. În consecință, el era foarte interesat de opera fundațională de pionerat a lui Alan Turing, dând o apreciere cu totul aparte originalității și fertilității ideilor acestuia. Gödel a ajuns să creadă că propriile sale teoreme de incompletitudine au atins un aspect important al limitelor formalizării, numai după ce Turing a dezvoltat analiza sa, pe care Gödel a susținut-o în totalitate, cu privire la conceptul de procedură mecanică (sau computațională), așa-numita mașină Turing. Mai mult, Gödel era satisfăcut de faptul că mașinile Turing oferă o dovadă pentru teza că există realmente concepte precise și că mintea umană le poate percepe în mod clar. Cu toate acestea, Gödel a detectat o problemă în argumentul lui Turing al adecvării analizei sale a algoritmilor, anume o demonstrație sofisticată a concluziei că mințile și mașinile sunt echivalente (cf. Wang, 1996, p. 194).

Poziția lui Gödel este clar delimitată de către următoarea remarcă: „6.3.5 Demonstrațiile tentative pentru echivalența minților și a mașinilor sunt sofisticate. Un exemplu este pretinsa demonstrație a lui Turing că fiecare procedură mentală pentru producerea unui șir infinit de numere întregi este echivalentă cu o procedură mecanică” (Wang, 1996, p. 197). Gödel explică de ce consideră că demonstrația tentativă a lui Turing este sofisticată: “6.3.6 Turing dă un argument care ar trebui să arate că procedurile mentale nu pot să conducă mai departe decât procedurile mecanice. Totuși, acest argument este inconcluziv, pentru că depinde de supoziția că o minte finită este capabilă [să aibă] numai un număr finit de stări care pot fi distinse între ele” (Wang, 1996, p. 197).

Gödel respinge supoziția că mintea (spiritul) este materie; el spune: „6.3.7 Este o prejudecată a vremurilor noastre că (1) nu există minte separată de materie; într-adevăr, (1) va fi respinsă în mod științific” (Wang, 1996, p. 198). El, apoi, continuă să interpreteze și să reconstruiască argumentul lui Turing și îl găsește valid, numai după ce anumite presupoziii sunt garantate și acceptate: “6.3.8 Este foarte probabil că (2) creierul funcționează, în mod fundamental, precum un computer digital. 6.3.9 Este practic sigur că (2') legile fizice, în consecințele lor observabile, au o limită finită de precizie. 6.3.10 Dacă acceptăm (1), împreună sau cu (2) sau cu (2'), atunci argumentul lui Turing devine valid” (Wang, 1996, p. 198).

Pentru noi, astăzi, este greu să nu acceptăm toate acele presupoziii tocmai pentru că suntem obișnuiți, sau poate că, într-adevăr, avem prejudecata de a gândi cu privire la creier și la minte ca fiind două aspecte ale aceluiași lucru. Totuși, Gödel nu socotea că lucrurile stau așa: “6.3.11 Dacă (i) o minte finită este capabilă numai de un număr finit de stări care pot fi distinse, atunci (ii) procedurile mentale nu ne pot conduce deloc mai departe decât procedurile mecanice. 6.3.12 Argumentul lui Turing (iii) pentru condiția (i) este ideea sa care gravitează în jurul următoarei propoziții: *Vom presupune, de asemenea, că numărul de stări ale minții care trebuie luat în considerație este finit. Temeiurile pentru aceasta sunt de același caracter ca acelea care limi-*

tează numărul de simboluri. Dacă admitem o infinitate de stări ale minții, unele dintre ele vor fi 'arbitrar de apropiate' și vor fi confuze" (Wang, 1996, p. 198).

Gödel este mulțumit cu inferența de la (i) la (ii). El crede, totuși, că (i) poate fi inferat din (iii), numai dacă avem la dispoziție unele asumptii suplimentare. Și deoarece Gödel nu acceptă că creierul este echivalent cu mintea, el continuă prin a respinge atât (i) cât și (ii).

În final, respingerea computabilismului/computerismului mental de către Gödel, convingerea sa adânc susținută că mintea ne poate conduce mai departe decât mașinile, sunt bazate pe ideea că „6.3.13 *Mintea, în utilizarea ei, nu este statică, ci într-o constantă dezvoltare*”. Atunci când ne focalizăm, în mod introspectiv, asupra fluxului conștiinței noastre, suntem frapați de faptul că stările mentale și succesiunea lor nu se bucură de precizia și claritatea stărilor mașinilor Turing. Wang comentează astfel ideea lui Gödel: "... ne dezvoltăm, de-a lungul timpului, atât în mod individual cât și în mod colectiv; și astfel, de pildă, ceea ce apare a fi complex devine simplu și înțelegem lucruri pe care nu le înțelegeam înainte. Și aici, iarăși, simțim că procesul de dezvoltare este oarecum indefinit și nu mecanic" (Wang, 1996, p. 200).

Disponem, noi, la urma urmei, de o demonstrație că mințile ne pot duce mai departe decât computerele și că ele nu sunt, pe de-a-ntregul, mecanice? Nu, nu dispunem de o astfel de demonstrație! Totuși, Gödel promovează un gen de viziune asupra minții care este dinamică și în dezvoltare, viziune care este atât grăitoare cât credibilă: „6.3.14 Deși la fiecare stadiu al dezvoltării minții numărul stărilor sale posibile este finit, nu există nici un temei de ce acest număr să nu convergă la infinit în cursul dezvoltării sale" (Wang, 1996, p. 200).

Un creier mecanic conectat la o minte/spirit creator, non-mecanic, într-o continuă evoluție și dezvoltare, este un „dispozitiv” care trece dincolo de individualismul minților atomice și izolate. Astfel, creierul și mințile pot să creeze gânduri, într-o manieră care reflectă faptul că mințile pot să ne ducă mai departe decât creierul și computerele, indicând, poate, în cele din urmă, superioritatea minții asupra computerelor.

Recunoaștere și mulțumiri. Doresc să mulțumesc d-lui Dr. Victor Mitrana, Universitatea din București, pentru comentariile asupra unei versiuni anterioare a acestui studiu, comentarii care au contribuit la îmbunătățirea formulării argumentelor. Sunt recunoscător d-nei Dr. Daniela Dumitru, ASE București, pentru discuții stimulatoare despre cognitivism și computaționism/computerism. Vreau, de asemenea, să mulțumesc domnișoarei Ioana Andrada Dumitru, studentă doctorandă la Johns Hopkins University, pentru comentariile și sugestiile stilistice pe care le-am acceptat bucuros și care au îmbunătățit claritatea și lizibilitatea studiului.

REFERINȚE BIBLIOGRAFICE

1. Churchland, P.M.: *Matter and Consciousness*, The MIT Press (2013).
2. Hintikka, J.: *On Gödel*, Wadsworth (2000).
3. Nagel, E., Newman, J.R.: *Gödel's Proof*, Routledge (1958).
4. Smith, P.: *An Introduction to Gödel's Theorems*, Second Edition, Cambridge (2013).
5. Tieszen, R.: *After Gödel. Platonism and Rationalism in Mathematics and Logic*, Oxford (2011).
6. Wang, H.: *A Logical Journey. From Gödel to Philosophy*, The MIT Press (1996).