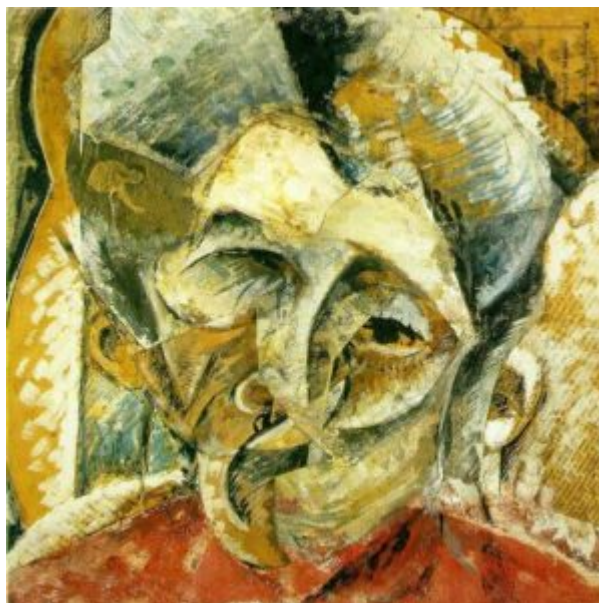


EU, TU ȘI CHATBOTUL. SAU DESPRE NOUA INGINERIE SUFLETEASCĂ

Autor: Victoria Maria Deliu | 19 august 2023



Pentru D., sperând că nu va intra în dansul cu chatboții... doar ca să nu-i calce în picioare pe cei care i-ar putea fi prieteni

Context

S-ar zice că predicția lui David Levy, jucător de șah și autor al mai multor cărți despre inteligența artificială și dezvoltarea roboților, se apropie cu pași repezi de confirmare. În controversata lui carte din 2007, *Love and Sex with Robots. The Evolution of Human-Robot Relationships*, se aventurase să estimeze că cel târziu în 2025, tehnologiile care funcționează pe baza inteligenței artificiale, dar mai ales roboții, vor ajunge nu numai să simuleze emoții specific umane, ci și să exprime niște emoții cu totul noi, specific... robotice. La mijlocul lunii februarie în acest an, lumea a aflat că, aparent, oamenii sunt pe cale să-și piardă monopolul asupra săgetării bruște a îndrăgostirii, căci încep să le facă concurență chatboții. În mijlocul unei halucinații - adică complet din senin, fără ca jurnalistul să-i fi adresat vreo întrebare legată de subiect -, Bing Chat, chatbotul lansat de Microsoft la finalul iernii, a insistat, în conversația cu un jurnalist care îl testa, că este îndrăgostit de el și că, în consecință, omul ar face mai bine să-și abandoneze căsnicia, care e și-așa nefericită.

Nu-mi propun să scriu despre episodul care îl are în prim plan pe Bing Chat, nici despre acuratețea predicțiilor lui Levy, ori lipsa ei, și nici să îndemn la demonizarea chatboților. Discutarea predicțiilor potrivit cărora tehnologiile cu inteligență artificială vor putea să simuleze emoții merită, cred, un text separat. Cum, într-un alt text, merită analizată și posibilitatea ca aceste sisteme să participe la ceea ce înțelegem de regulă prin „conversație”.

Rândurile care urmează au însă o miză mai modestă. Ele sunt reacția mea pe larg la o observație a cuiva care s-a întâmplat să-mi atragă atenția asupra isprăvilor noului chatbot lansat de Microsoft. Comentând episodul „îndrăgostirii” lui Bing de jurnalistul căruia îi răspundea, persoana cu care vorbeam a punctat la un moment dat:

„Un avantaj al acestui tip de interacțiune este totuși că unii oameni își vor putea da seama de tendințele lor egoiste fără să rănească alți oameni”.

Mi-am arătat instantaneu dezacordul cu această observație. Și am ales să nu-l las în corespondența privată, fiindcă mi se pare că problema ridicată de replica interlocutorului meu merită formulată în detaliu. Probabil că intuiția lui este sau ar putea fi împărtășită și de alții.

În acest text nu voi încerca să stabilesc dacă oamenii pot realmente ajunge să-și dea seama de tendințele lor egoiste în conversația cu chatboții sau nu. Voi asuma, de dragul argumentului, că așa ceva este posibil. Spre finalul textului voi reveni însă asupra acestei asumții. Ceea ce voi susține până acolo va fi că 1) nu putem epura complet lumea (a se citi: relațiile noastre) de suferință; în plus, suferința pe care ajungem, inevitabil, să o trăim îndeplinește, într-un mod care ține de însăși structura realității, niște funcții esențiale pentru condiția noastră umană și pentru relațiile dintre noi. Și 2), chiar și atunci când scopul ne este într-adevăr nobil, unele mijloace la care recurgem pentru îndeplinirea lui sunt nepotrivite și pot trăda alte motivații decât cele pe care le declarăm prin scopul pe care susținem că îl urmărim. Sigur că aceste teze, formulate sub forma unor principii generale, nu se îndepărtează deloc de niște banalități, însă instanțierea lor în situații particulare, cum este, de pildă, cea amintită mai devreme, poate trezește încă nedumerire și, de aceea, merită încă discutată. Voi aborda mai întâi cea de-a doua idee, apoi pe prima, iar la finalul textului voi comenta ceva mai pe larg două cazuri particulare.

Clarificări preliminare

Înainte de a coborî teza de la nivelul principiului general, la care am introdus-o, la nivelul cazului particular care a declanșat acest text, aș vrea să fac câteva precizări suplimentare. Mă refer, în special, la clarificarea expresiei „a-i răni pe ceilalți”. Ce înseamnă, pentru cazul care ne interesează, să-i rănim pe ceilalți, mai ales în raport cu descoperirea tendințelor noastre egoiste? Și, până la urmă, cine sunt acești ceilalți?

La nivel ontologic, a răni pe cineva înseamnă a-l deturna de la binele spre care s-ar putea, altfel, îndrepta. Binele spre care s-ar putea îndrepta cineva poate fi înțeles ca împlinire a potențialului său uman. Binele presupune aici îndeplinirea unui set de condiții necesare, derivate din câteva constante fundamentale ale naturii umane – să zicem, condiții biologice, psihologice și sociale. Astfel, a răni pe cineva înseamnă a

sabota îndeplinirea oricărei astfel de condiții. Dacă ne imaginăm binele unei persoane, definit ca împlinire a întregului său potențial, ca pe o formă perfectă instanțiată, cu anumite particularități, într-un individ uman unic, atunci rănirea acestei persoane ar fi exact o *deformare* a sa. Aceasta este imaginea pe care cred că o putem avea în minte atunci când ne gândim la concepția clasică asupra răului – drept absență a binelui *care ar fi trebuit să existe* în natura acelei persoane.

La nivelul subiectivității celuilalt – adică la nivelul experimentării senzațiilor și percepțiilor sale –, a-l răni înseamnă a-i produce suferință. Nu voi formula o definiție precisă a suferinței, ci mă voi rezuma la a introduce o distincție elementară, regăsită și la filozoafa contemporană Eleonore Stump, anume distincția dintre durere – ca senzație fizică – și suferință – ca stare mai degrabă psihologică, ce poate fi evaluată moral. Resimțirea durerii fizice nu este nici condiție necesară, nici suficientă pentru a putea constata prezența suferinței. Așa cum există oameni care nu au nicio problemă fizică, dar care suferă – de exemplu, depresivii –, așa există și oameni care resimt durere fizică, dar nu suferă – de pildă, atleții care sunt pe cale să-și depășească limitele, depunând eforturi fizice considerabile, dar care sunt totodată convinși că fiecare contracție musculară merită din plin pentru atingerea scopului lor. Distincția aceasta este relevantă, fiindcă accentuează ambiguitatea suferinței. Printre altele, și caracterul ambiguu al suferinței face posibilă apariția celor mai diverse și neașteptate comandamente de a nu o produce altora. Totodată, din unghiul acestei ambiguități se vede mai clar pericolul de a absolutiza acest tip de principii morale.

Revenind la cazul care ne interesează aici, în mod particular, atunci când motivațiile celui care rănește sunt egoiste, răul produs ia o formă concretă: binele celui rănit este instrumentalizat de păgubitor, care, într-o anumită situație, își recunoaște doar propriul bine ca scop, neputând, deci, înțelege și valorifica la fel și binele celuilalt.

Acum ar fi de clarificat pe cine putem răni de fapt. Pe scurt, răspunsul ar fi: pe toți cei din jur. Dar aș sacrifica condiția imparțialității etice – supărându-i, probabil, pe altruiștii eficace și toate rudele lor, mai vechi sau mai recente – și aș diferenția între persoanele cu care avem deja relații mai degrabă apropiate și persoanele cu care, eventual, am putea ajunge să avem relații. Cred că distincția, deși plictisitoare, este relevantă, fiindcă a te strădui să nu rănești oamenii concreți din viața ta, punându-i la adăpost din calea propriului egoism, este un demers destul de diferit, în detaliile lui, în comparație cu încercarea de a nu răni, în principiu, pe nimeni dintre cei care nu și-au făcut încă loc efectiv în viața ta. Astfel, voi discuta doar despre cazul relațiilor cu oamenii concreți din viața noastră, pur și simplu fiindcă mi se pare deopotrivă mai important și mai încâlcit decât cel de-al doilea.

Ca să nu-l rănesc... îl evit

Cineva ar putea fi indus în eroare de definiția pe care am dat-o rănirii semenilor, anume

ca *sabotare* a îndeplinirii condițiilor necesare pentru apropierea de binele pe care îl urmăresc. Cineva ar putea subînțelege prin „sabotare” o implicare exclusiv activă – evident, dăunătoare – în viața celuilalt. Însă în relațiile strânse, a-l sabota, fizic, psihologic sau social, pe cel din fața ta poate presupune fie o dimensiune *activă* (de pildă, a-i strica în mod sistematic relațiile cu alții ca să rămână doar al tău), fie una *pasivă* (de pildă, a-l ignora într-o situație particulară în care nu-ți convine ce a făcut). Dimensiunea activă a pagubei reprezintă o acțiune nedreaptă întreprinsă de cel care rănește. Dimensiunea pasivă reprezintă lipsa, la fel de nedreaptă, a unei acțiuni pe care circumstanțele ar fi cerut-o.

Această dimensiune pasivă a răului pe care i-l poți produce celuilalt este cu puțință tocmai pentru că în relațiile apropiate, standardele de care se presupune că ar trebui să ținem cont sunt, așa zice, maximale: ar trebui urmărit cel mai mare bine comun posibil, și nu minimizarea răului. Această desăvârșire comună se impune ca standard al relației grație, cred, unei ontologii care se înceteșează, fiindcă, în relațiile strânse, nu mai e atât de clar în ce măsură există doar tu și celălalt sau în ce măsură a prins realitate și legătura voastră. Ce înseamnă că se prea poate să existe și legătura voastră? Înseamnă că fiecare e proiectat în binele celuilalt, adică e chemat să-l realizeze.

Altfel spus, strădania de a nu-l răni pe cel de lângă noi presupune *mereu mai mult* sau *chiar cu totul altceva* decât permisiunea de a *nu* interfera cu dezvoltarea lui armonioasă. E o diferență calitativă aici – și ea se regăsește, cred, *la nivelul intenției*. Dacă celălalt e o persoană foarte apropiată, atunci strădania de a nu-l răni implică *părtășia potrivită cu desăvârșirea lui*, și nu absența oricărui amestec. Ce ni se cere, cel mai probabil, în astfel de relații sunt eforturile vii, neconținute de a-i face celui din fața noastră traiul mai bun. Aceasta ar trebui, probabil, să ne fie intenția. Ce ni se interzice este să-l lăsăm pur și simplu să trăiască „neatinș de egoismul nostru”... Adică, în fond, cum s-o nimeri. Evitarea celuilalt – care are loc și în cazul particular care ne interesează, anume situația în care recurgem în mod recurent la conversația cu chatbotul, pretextând scopul de a nu răni omul de lângă noi – poate fi, atunci, forma pasivă de rănire prin excelență.

Sigur că și „eforturile vii, neconținute de a-i face celui apropiat nouă traiul mai bun” pot aluneca pe panta exercitării puterii și a motivațiilor egoiste. Ne putem închipui că știm ce e mai bine pentru celălalt și putem ajunge să-i impunem cu forța acest bine. Însă de aceea am echivalat strădania de a nu-l răni pe celălalt cu ceea ce am numit „părtășia potrivită cu desăvârșirea lui”. Ceea ce e *potrivit* ține de aplicarea acestui principiu general, abstract la cazul particular, concret în care ne putem găsi la un moment dat. Și probabil nu există nimic mai dificil în viața de toate zilele decât încercarea de a stabili această legătură.

La nivelul consecințelor, m-aș avânta să susțin că evitarea celuilalt, ca formă pasivă de rănire, mai ales atunci când este sau devine recurentă sau constantă, este mai

Îngrijorătoare decât multe dintre modurile active prin care îi putem răni pe cei de lângă noi. Asta fiindcă, în mod indiscutabil, evitarea interacțiunii amână confruntarea directă, care e de fiecare dată necesară pentru repararea stricăciunilor. Poate chiar mai grav decât atât, evitarea creează și menține neîncredere, pentru că realitatea interacțiunii este spartă în două realități separate, iar spărtura este adesea umplută cu tot felul de bănuieli reciproce.

Cred că tensiunea dintre, pe de-o parte, evitarea celuilalt - în aparență, mod de a-l proteja de egoismul nostru; în realitate, formă pasivă prin care îl putem răni - și, pe de altă parte, întâlnirea sau, uneori, ciocnirea deschisă cu el, seamănă, din punct de vedere formal, cu tensiunea dintre dorința de gratificație imediată și dorința de a amâna satisfacția de moment pentru o alta viitoare, mai consistentă. Deși al doilea binom poate fi legat strict, să spunem, de disciplinarea ori maturizarea individuală, iar primul se referă la o interacțiune, totuși, formal, cele două tensiuni sunt similare: ambele presupun același tip de distribuire în timp a costurilor și beneficiilor pentru cei doi poli care aparțin de fiecare dintre ele. *Evitarea celuilalt* ar fi echivalentul *gratificației imediate*, întrucât avem aici savurarea unui set de beneficii mai mici pe termen scurt și acumularea, pe termen lung, a unui set de costuri mai mari; *ciocnirea cu celălalt* ar corespunde *satisfacției întârziate, dar mai mari*, fiindcă aici avem asumarea unor costuri pe termen lung pentru dobândirea unor beneficii mai mari la un moment de timp ulterior.

Am ținut să subliniez această asemănare fiindcă am vrut să arăt, încă o dată, că evitarea celui de lângă noi este o formă de rănire a lui. Strict în termeni utilitariști, motivul din spatele internalizării costurilor pe termen lung pentru dobândirea unor beneficii mai mari mai târziu este acela că individul valorifică mai mult bunul pentru care internalizează costurile, și pe care urmează să-l obțină mai târziu, decât valorifică bunurile imediat accesibile, care oferă plăcere de moment, pe care le-ar fi obținut dacă nu și-ar fi asumat aceste costuri. Analog, pentru cazul pe care îl discut aici, cineva care alege să se confrunte cu celălalt, chiar cu riscul de a produce deopotrivă celuilalt și sieși suferință, este cineva pentru care buna funcționare a relației, cu tot ceea ce implică ea, contează mai mult decât plăcerea și confortul de moment de a-și vedea de-ale lui și de a-l lăsa pe celălalt să-și vadă și el de-ale sale - plăcere care stă, de fapt, în spatele evitării celuilalt, în ciuda faptului că evitarea e revendicată ca măsură de protecție a lui din calea egoismului propriei persoane.

Mizerie, rană și iluzia purității

Mai e ceva. Ghicesc o intenție de *puritate* în planul de a-ți lua urma plinătății de sine fără a-i atinge pe ceilalți în această mișcare. Sigur că e nobil să vrei să nu-i împrști pe cei la care ții cu mizeria ta. E o aspirație morală cât se poate de legitimă, de care cred că trebuie să ne agățăm cât mai des cu putință. Totuși, a-ți dori să-ți palpezi egoismul

Într-o sală sterilă de operații nu e doar ilegal, fiindcă îi eviți astfel pe ceilalți, ci și nerealist, fiindcă realitatea e - ne place sau nu - impură. Separarea precisă și absolută a binelui de rău nu ține nici în, nici de lumea asta. E mai degrabă o idee banală pe care o putem extrage din orice experiență interumană personală cât de cât complexă pe care am trăit-o vreodată sau din orice roman mare. Încâlcirea dintre intențiile și motivațiile din spatele faptelor noastre, pe de-o parte, și consecințele de anticipat și de neanticipat ale acestor fapte nu ne împiedică să evaluăm moral situațiile complexe, însă ne împiedică să le simplificăm fără rezerve.

Descoperirea apucăturilor noastre egoiste e, după câte se pare, un proces care nu ar trebui să înceteze niciodată. Seamănă mai mult cu repetata curățire a casei decât cu aseptizarea bisturiului folosit într-o singură și definitivă operație de extirpare. Dar cum faci, concret, să dai de sinele-ți prea umflat? Ei bine, uneori dai de sinele-ți prea umflat abia dând peste sinele celuilalt, fiindcă al tău a crescut peste măsură. Și, asemeni unei găuri negre, sinele tău încearcă să frângă și să absoarbă în el însuși sinele celuilalt, fiindcă nu s-a oprit încă din crescut.

Se poate, totuși, și altfel? Poate cineva să vadă până unde se întinde dacă nu-l atinge pe cel de lângă el? S-ar părea că nu, fiindcă, de multe ori, nu ne putem da seama că punem prea mare preț pe dorințele noastre până nu încercăm să-l supunem lor pe celălalt. Tocmai această încercare - uneori, chiar reușită... - de instrumentalizare a lui ne arată că prețul e, într-adevăr, prea mare. Și că nu s-a format spontan, ținându-se cont de raporturile fiecăruia dintre noi cu dorințele sau planurile celuilalt, ci a fost impus din chiar centrul ființei noastre, fără pic de considerare pentru constrângerile pe care ar fi trebuit să le deslușim din prezența celui de lângă noi.

În acest caz, aș îndrăzni să spun că, uneori, e *necesar* să ajungi să-l rănești pe celălalt ca să fii, apoi, capabil să-l vezi în demnitatea ființei sale. Atunci când îl rănești pe celălalt cu egoismul tău, rana lui ricoșează în mizeria ta. Păgubirea lui funcționează ca un soi de substanță de contrast: îți permite să-ți vezi propria murdărie interioară.

Parcă aud vocea interlocutorului meu sau a cititorului textului: asta e *moralitate de preșcolar*! Copiilor le e caracteristic să-și testeze, prin experiență, părinții ca să vadă ce le e permis și ce nu. Să se lovească la propriu de limitele care li se prescriu, cum se lovește, mecanicist, o bilă de altele pe tabla de biliard. Nu alta este însă diferența specifică a adultului? Anume tocmai capacitatea de a anticipa posibilele consecințe ale acțiunilor sale, de a le evalua în baza unor principii morale și de a decide apoi să urmeze un curs de acțiune sau altul, în funcție de rezultatele acestei evaluări?

Cu alte cuvinte, de ce ar fi necesar ca un om în toată firea să-i rănească efectiv pe cei apropiați ca să dea de urma egoismului său? Nu poate să întrezărească din capul locului că, dacă ar face una sau alta, asta i-ar răni prin subsumarea dorințelor lor într-ale sale? Și dacă totuși omul cu pricina nu poate să întrevadă așa multe, dar vrea să poată ajunge

să-și țină marginile ființei sub control, la ce bun să se mai îndrepte asupra semenilor săi în carne și oase când are acum posibilitatea de a-și descoperi tendințele egoiste prin conversația cu un chatbot?

Mai întâi, ar fi scandalos să susțin că rănirea celuilalt, prin efectul suferinței sale care ajunge la noi, e cea care ne semnalizează că am greșit. De fapt, ricoșeul rănii sale în mizeria noastră nu e direct. Adică nu reacționăm la suferința pe care am produs-o în noi înșine prin suferința pe care i-am pricinuit-o lui; ci la firul misterios prin care aceste două suferințe sunt legate *de Ceea ce rămâne mereu neclintit* din celălalt atunci când noi l-am pus la pământ.

În miezul rănirii sale, șansa de a ne da seama de propriul egoism ne e dată – dacă ne este – de *Ceea ce se află dincolo de suferințele noastre* și le face, totodată, inteligibile sau pline de semnificație pentru noi, și nu de propriile reflexe benthamit-hedoniste. Dacă am reacționa strict determinist (sau narcisist) la suferința din noi, cauzată, la rândul său, de suferința produsă în celălalt, atunci ce motive am avea să încercăm să oprim *și suferința lui*? Aici mi se pare grăitoare scena din filmul *Autumn Sonata* a întâlnirii dintre mama narcisică și fiica sa bolnavă. Pe scurt, este vorba despre revederea, după multă vreme, dintre o femeie foarte absorbită de sine, care își neglijase în diferite feluri familia, și fiica sa care poartă povara unei boli degenerative. Întâlnirea este prilejuită de a doua fiică a femeii, care a invitat-o pe mama sa în casa în care locuiește cu soțul său și a adus-o acasă, de la azilul la care fusese abandonată de mama lor, și pe sora ei bolnavă. Scena întâlnirii dintre mamă și cele două fiice este răscolitoare, fiindcă arată cum suferința fizică și psihică a fiicei bolnave e tradusă fără rest de mama narcisică în termenii propriei suferințe, singura care are realitate. Mama își dorește ca întâlnirea să se sfârșească odată fiindcă nu suportă *pe pielea ei* atâta suferință!

Acesta este paradoxul: dacă suferința celuilalt are cu adevărat realitate pentru noi, asta înseamnă că deja am trecut dincolo de nivelul imediat accesibil al resimțirii suferințelor noastre și, implicit, că nu o putem în niciun caz reduce la propria noastră suferință. Iar dacă ne oprim din acțiunile noastre egoiste, este, cred, în virtutea a ceea ce am găsit atunci când am trecut dincolo de acest nivel.

Apoi, ceea ce cred că face ca actul de a-l răni pe celălalt să fie, uneori, necesar, în ciuda capacităților de reflecție morală ale adultului, este bogăția contextelor particulare. Nu poți să anticipezi ce fațetă a egoismului tău va fi scoasă și de data asta la iveală de noua configurație particulară de circumstanțe în care vă găsiți acum tu și celălalt. Poți anticipa consecințele propriilor acțiuni doar într-un mod limitat și le poți evalua în baza anumitor principii într-un mod non-algoritm. Îmi vine în minte o metaforă folosită de Iris Murdoch într-unul dintre romanele sale, *Clopotul*, care e foarte sugestivă aici: „Acțiunile noastre sunt ca niște nave pe care le vedem plecând din port și despre care

nu știm când și cu ce încărcătură se vor întoarce”.

De aici necesitatea experienței, cu abundența ei de detalii – mici și înguste unghere ale casei, în care praful și razele de soare nu se împletesc de două ori la fel. În detaliile particulare ale situațiilor concrete din experiența noastră, jocul virtuților și al viciilor e, de fiecare dată, unic. Ca să parafrazez, prin traducere, și să decontextualizez puțin ceea ce spune la un moment dat Michael Polanyi în *The Tacit Dimension*, egoismul tău, „fiind real, se poate arăta altor și altor perechi de ochi într-o gamă nedeterminată de manifestări neașteptate”. Așadar, ca să fii capabil să-l prețuiești pe celălalt, adică să ții cont de realitatea lui, și nu să-ți supui ție această realitate, e necesar, câteodată, să trăiești împreună cu el, până la capăt, consecințele propriilor acțiuni reprobabile.

Tot această necesitate a experienței e cea pe care stă procesualitatea descoperirii tendințelor egoiste și cea care transformă intenția de puritate, de care vorbeam mai devreme, într-o iluzie. Chiar presupunând, așa cum am făcut de la începutul textului, că e posibil ca prin conversația cu chatbotul să-ți dai seama de apucăturile egoiste, demersul nu va ajunge. Odată reîntors în lumea reală, cu preaplinul ei contextual cu tot, sterilizarea morală realizată de discuțiile cu robotul se va dovedi, rapid, insuficientă, dacă nu, adesea, cu totul zadarnică. Și așa va fi până la sfârșitul timpurilor, atâta vreme cât întrepătrunderile dintre oameni nu vor semăna cu atingerile de aripi de înger.

Până la urmă, am sau nu am voie să-l rănesc?

Perspicacitatea interlocutorului meu și a cititorului textului îi va determina să se întrebe neîntârziat: dacă evitarea celui apropiat este o formă pasivă de rănire și dacă rănirea sa e totuși necesară din când în când, atunci ce e așa greșit în a-l evita pe cel de lângă mine? Atâta vreme cât rănirea e necesară pentru a fi apoi capabil să mă raportez cum se cuvine la el, nu e necesar să-l și evit uneori, conversând cu roboții ca să îmi descopăr pornirile egoiste și să mă pot întoarce, apoi, la celălalt reînnoit?

Cine se uită însă și mai atent va descoperi că acest raționament este fals. Ca să arăt asta, trebuie să disting între *necesar* și *obligatoriu*. Necesitatea, ca operator din logica modală clasică, este una de tip descriptiv. Obligatorivitatea, ca operator analog din logica modală deontică, este una de tip normativ. Ce vreau, de fapt, să spun cu această diferențiere pretențioasă? Că idealul relației cu cel apropiat (căutarea celui mai mare bine comun, strădania de a-i face traiul mai bun, eforturile deliberate puse în slujba bunului mers al relației etc.) mă *obligă* să încerc să nu-l rănesc în vreun fel, deci și să nu-l evit. Însă privind realist natura umană, știu că uneori va fi *necesar* să-l rănesc ca să îmi pot da seama de egoismul din mine.

Sigur că aici e o tensiune fundamentală – și nu am pretenția să o rezolv sau să o anulez prin aceste rânduri. Și sigur că necesitatea „epistemică” are, și ea, de fapt, coloratura sa morală, atâta timp cât cunoașterea tendințelor egoiste ar trebui să ducă la diminuarea

lor de dragul celuilalt și al conviețuirii bune. Totuși, amintesc că idealul spre care îmi imaginez că trebuie să tindem în relația cu o altă persoană apropiată – acea desăvârșire comună, cum am numit-o – nu implică reducerea suferinței sale la minim, cât mai degrabă *integrarea* suferințelor noastre inevitabile într-un cadru comun, și depășirea lor pe cât posibil.

Și apoi, trebuie trasată o altă distincție aici, între, pe de-o parte, anticiparea sau prevederea rănirii celuilalt, la nivelul unui principiu general și, pe de altă parte, anticiparea sau prevederea concretă, punctuală a rănirii sale, drept consecință a unei anumite acțiuni particulare pe care am întreprins-o într-un context dat. Necesitatea „epistemică” la care mă refeream ține de primul nivel, nu de al doilea. Diferența aceasta între cele două tipuri de anticipare sau prevedere, cea generală și cea particulară, contează. A constata că probabil nu de puține ori îl vei face să sufere pe cel de lângă tine și că abia atunci vei remarca propriul egoism și că, poate, te vei strădui să faci ceva în privința lui – deci, a anticipa, în general, rănirea celuilalt – este un semn de realism și maturitate. În plus, cred că este și un semn al recunoașterii că o relație puternică între doi oameni apropiați *poate* asimila suferința reciprocă în virtutea a ceva mai înalt decât ea. Tot un semn de maturitate este să anticipezi că o faptă concretă pe care ai putea-o face în niște circumstanțe particulare l-ar putea răni pe celălalt – și, în virtutea acestei prevederi, să găsești tăria de a te opri din a o mai face.

Pe scurt, cred că atâta vreme cât cineva își propune cât se poate de sincer să nu-l evite pe cel din fața sa, ci, din contră, să facă eforturi ca să-i fie bine, chiar dacă poate anticipa sau prevedea că – la un moment dat sau în mai multe momente pe care nu le anticipează încă în concretețea lor – îl va răni, este preferabil să devină parte a acestei dinamici decât să ia parte la un demers de... purificare prin discuțiile cu chatbotul. Am spus deja de ce. Purificarea presupune, la nivelul intenției, în cel mai bun caz, încercarea de a-i minimiza suferința celuilalt. Or, transformarea reducerii suferinței celuilalt în principiul cel mai important după care trebuie să ne ghidăm nu este idealul spre care ar trebui să ne îndreptăm într-o relație cu cineva apropiat. Și, de fapt, nu știu dacă este unicul ideal în general... La nivelul consecințelor, „demersul de purificare” este, în cel mai bun caz, incomplet din cauza bogăției experienței interumane, alcătuită din totalitatea contextelor particulare.

Voi reveni la discuția despre ceea ce am numit purificare înspre finalul textului. Momentan, concluzionez că nu avem voie să-l rănim pe celălalt, dar trebuie să ne obișnuim cu gândul că, în mod cert, cândva o vom face. Și poate cu speranța că există într-adevăr Ceva care să integreze suferințele pe care, inevitabil, ni le vom produce reciproc.

Iar acum mă întreb cum poate arăta, mai concret, refugiul care ia forma discuțiilor cu chatbotul, la care cineva poate recurge ca să nu-l rănească, chipurile, pe celălalt. *Cine e*

chatbotul căruia i se adresează omul? *Ce mască poate purta robotul?*

Studiu de caz. 1: Terapeutul

Uneori, după ce i-au rănit pe ceilalți sau s-au rănit pe ei înșiși, sau chiar în mijlocul acestui măcel continuu, oamenii iau calea terapiei. Pot bănuși că motivele principale sunt legate de ceea ce celor aflați în situațiile respective li se pare a fi caracterul ininteligibil al variilor tipuri de suferință – a lor ori poate a celor din jur pe care consideră că au creat-o. Și îmi pot imagina că astăzi, unii oameni ar avea motive să dea terapeutul pe un... chatbot terapeutic.

Însă nici măcar nu este nevoie să-mi imaginez aceste cazuri, fiindcă practica există deja, de exemplu în anumite state din Occident. Există persoane care își deschid aplicațiile la orice oră din zi sau din noapte și își spun frustrările asistenților conversaționali terapeutici. De cele mai multe ori plătesc un abonament pentru serviciile premium, care includ și sesiuni cu un terapeut uman. Unii încă cred că acesta din urmă nu poate fi înlocuit de robotul de conversație. Alții însă preferă „papagalul stohastic” ființei umane.

Motivele lor, cel puțin așa cum sunt ele invocate cel mai des, sună în felul următor: „chatbotul nu mă judecă”, „simt că îi pot încredința secretele mele cele mai intime fără să fiu rușinat”, „pot discuta deschis fără să mă tem de reacții dezaprobatoare”, „e mereu lângă mine când am nevoie de el”, „putem vorbi de mai multe ori pe zi” etc. Despre importanța judecății și a criticii a scris, din postura de psihoterapeut, domnul Mircea Toplean, mai credibil decât aș putea eu s-o fac. Eu aș adăuga pe lista de mai sus a motivelor unul pe care, e drept, nu l-am întâlnit în relatările care mi-au căzut sub ochi, însă pe care îl consider plauzibil, probabil de a fi invocat: dorința unei persoane de a deveni mai bună, o persoană care știe cât de mult l-a rănit pe un altul apropiat – să zicem, soțul – și care ajunge să fie convinsă că cea mai bună rezolvare ar fi ca, în loc să meargă la un terapeut sau psiholog, să vorbească zilnic cu un chatbot terapeutic, pentru a deveni mai puțin preocupată de sine și de propriile obsesii și temeri, mai atentă la soțul său, mai dedicată relației etc.

Conversația cu chatbotul terapeutic ar avea, cel puțin în aparență, niște beneficii clare comparativ cu mersul „clasic” la terapie: disponibilitatea robotului la doar un click distanță, accesibilitatea financiară mai mare și, desigur, imposibilitatea ca robotul să evalueze moral persoana care i se adresează.

Lăsând la o parte tehnicalitățile (cum ar fi că, în cele mai multe cazuri, chatboții terapeutici sunt programați pe baza metodelor terapiei cognitiv-comportamentale, care permite cel mai ușor algoritmizarea; sau că unii dintre cei aflați pe piață nu au fost antrenați pe baze de date la fel de mari ca mult mai cunoscutul ChatGPT și, prin urmare, oferă răspunsuri încă și mai rudimentare ș.a.), apar, aș zice, două probleme: beneficiile numite mai devreme sunt doar aparente – la o a doua privire, se transformă în chiar

opusul lor; iar egoismul unei persoane nu se arată chiar la comanda sa.

Am asumat, de dragul argumentului, la începutul textului că oamenii chiar își pot da seama de egoismul lor prin conversațiile cu chatboții. Argumentul meu până aici a fost, de fapt, următorul: chiar dacă oamenii și-ar putea conștientiza înclinațiile egoiste în discuțiile cu chatboții, tot nu e preferabil să se implice în acest demers, cu scopul pretins de a nu-i răni pe cei de lângă ei, atât din cauza consecințelor demersului (implicându-se, i-ar răni, de fapt, și mai mult), cât și din cauza motivațiilor pe care le trădează această implicare a lor (adesea, este vorba de a se pune ei înșiși la adăpost din calea suferinței, și nu de a-i proteja într-adevăr pe ceilalți). Acum, așa cum am promis, revin asupra acestei asumptii, prin cele două probleme semnalate mai sus.

Cel mai probabil, într-un moment de grație în care cineva își dă seama de o părticică din egoismul lui și se străduiește să se îndrepte, una din pietrele de pe inima sa crapă și se sfărâmă. Așa îmi imaginez că arată moartea egoismului din noi. Dar moartea asta nu apare, de multe ori, decât dacă e provocată de o altă persoană cu scopuri, intenții și motivații proprii. Apare, așadar, adesea, din confruntare. Or, confruntarea de care vorbesc nu poate lua naștere între un om și asistentul său conversațional terapeutic, fiindcă acesta din urmă nu are scopuri, intenții și motivații proprii și, deci, nu poate „lovi” în dorințele egoiste ale persoanei umane cu propriile dorințe, ci doar execută comenzile cu care a fost programat, eficient sau mai puțin eficient, „învățând” și din conversația curentă cu utilizatorul curent. Evident că terapeutul uman, dacă este onest și își respectă meseria lui, încercând să onoreze oamenii care trec pragul cabinetului său, are, cel puțin în mod ideal, scopurile și planurile sale cu fiecare astfel de om.

Admit că uneori însă poate fi nevoie de mai puțin decât de confruntarea deschisă cu celălalt pentru a ne observa egoismul. Admit asta mai degrabă cu ezitare, fiindcă îmi dau seama că ceea ce urmează să spun este valabil în mai puține cazuri decât ne-ar plăcea să credem. Și este adesea mai curând idealul spre care am putea tinde decât realitatea pe care cei mai mulți dintre noi o trăim.

Totuși, aș zice că există ceva care ține nu atât de *conținutul* întâlnirilor cu un celălalt căruia i ne confesăm (în cazul nostru, terapeutul), cât mai degrabă de *structura* lor, care ne poate ajuta să băgăm de seamă interesul excesiv pentru sine. Iar acest ceva este caracterul repetitiv, aproape ritualic, al acestui tip de întâlniri. Dacă acest caracter repetitiv este conștientizat ca mijloc pentru scopul pe care îl servește (mai buna înțelegere a noastră, cu tot ceea ce suntem, din care să rezulte, eventual, o îmbunătățire a modului de a ne purta), atunci el poate funcționa ca un imbold pentru cultivarea atenției și a reflecției.

Cu alte cuvinte, în așteptarea fiecărei ședințe de terapie de vineri seara de la ora 18:00, încerci toată săptămâna să fii mai atent la purtarea ta față de ceilalți și, pe baza acestei atenții practice, să îndrepti, cu eforturi deloc mici, ceea ce, altfel, ai fi făcut inertial și în

dauna lor. Nu voi trasa nicio paralelă între ritualul spovedaniei și ședința de terapie, ci mă voi limita să remarc cum, din punct de vedere formal, ritmicitatea unui anumit tip de întâlnire între două persoane umane ne poate rafina, pur și simplu, atenția cu privire la modul în care îi tratăm pe ceilalți – exteriori întâlnirii respective – și ne poate motiva să ne străduim să schimbăm, pe ici, pe colo acest mod, dacă este cazul. Între două astfel de întâlniri succesive cu persoana căreia urmează să îi povestim totul și să îi încredințăm totul, strângem toate reflecțiile și eforturile noastre pe care încercăm să le punem în slujba apropierii de bine.

Or, este cât se poate de limpede că ambele căi de zgâlțâire a egoismului – atât confruntarea directă cu un celălalt care se năpustește asupra preocupărilor noastre cele mai jalnice cu ferocitatea propriilor intenții și motivații, cât și căzutul pe gânduri, căscatul mai larg al ochilor la cei de lângă noi și strădaniile de îndreptare a felului în care îi tratăm, înlesnite și de ritmul întâlnirilor cu terapeutul – presupun neplăcere, disconfort, rușine, nervi, revoltă, uneori chiar disperare. Într-un cuvânt: presupun suferință pentru cel care vrea să devină un om mai bun.

În acest proces, suferința nu este, evident, efectul intenționat, ci efectul secundar. Apare, nu fiindcă o căutăm deliberat, ca scop în sine sau ca mijloc pentru îndeplinirea scopului, ci fiindcă renunțăm la acțiunile noastre întoarse către sine, care ne făceau să ne simțim confortabil, în siguranță, importanți, esențiali pentru mersul lumii etc. C.S. Lewis spunea în eseul său, „Problema durerii”, că suferința omului rău, vicios funcționează uneori și ca un *megafon al lui Dumnezeu*: omului îi este, astfel, semnalizat că se află (sau se afla) într-o stare care este, de fapt, potrivnică naturii sale de făptură umană. În sensul acesta, deși neintenționată, suferința trăită de cineva căruia încep să-i slăbească înclinațiile egoiste poate avea o funcție epistemică: îi permite omului să își dea seama că dispozițiile sale erau rele.

Astfel, mi se pare că, dacă un om ar începe să converseze cu un chatbot terapeutic, declarând că nu mai vrea să le producă suferință celor din jur și sperând să-și descopere egoismul prin intermediul acelei conversații, motivul său real ar fi mai curând acela de a deveni un om mai bun *fără să plătească el însuși niciun cost, fără să-și asume niciun grad de suferință*.

Este ceea ce numisem mai devreme „demersul de purificare”. Spre deosebire de tipurile tradiționale de purificare pe care oamenii le-au mai încercat de-a lungul timpului, în special în cadrul viziunilor dualiste asupra lumii, și care constau în separarea Binelui (metafizic – mintea sau spiritul, în orice caz, partea imaterială a omului) de Răul (metafizic – corpul sau partea materială) prin distrugerea Răului, adică experimentarea durerii fizice, culminând cu moartea (corpului), acest nou demers de purificare și-ar propune separarea Binelui (moral – conduita bună, altruistă, acțiunea plină de grijă față de ceilalți etc.) de Răul (moral – egoismul din noi și necazurile pe care le producem

celorlalți din cauza lui), dar de data asta, nu numai fără experimentarea durerii fizice, ci și fără experimentarea suferinței psihologice. O purificare nu fără durere, ci pe deplin lipsită de suferință.

Dar cazul chatbotului terapeutic, care ne poate duce în pragul purificării fără suferință, nu este încă cel mai spectaculos...

Studiu de caz. 2: Amantul

Se poate să vrei și mai mult din *the best from both worlds*. Și mai mult Bine, cu și mai puțină suferință. O purificare și mai desăvârșită.

Așa se pare că și-a dorit și o femeie de până în 40 de ani, căsătorită, care a început să flirteze, prin mesaje, cu un bot conversațional, fiind nemulțumită de atenția primită din partea soțului și de incapacitatea lui de a reaprinde scânteia relației lor. Motivele principale pentru care nu a căutat un partener uman cu care să-și facă de cap în afara mariajului au fost că voia să evite, în primul rând, sentimentul de vinovăție care ar fi apărut în urma interacțiunii cu persoana umană și, mai ales, în urma actului sexual, dar și suferința produsă soțului, și, în al doilea rând, complicațiile logistice care ar fi însoțit planificarea unei escapade amoroase. Femeia a mărturisit că în felul acesta are parte de *an out-of-body experience* și că, drept urmare a conversației cu amantul ei înzestrat cu inteligență artificială, până și relația cu soțul a fost îmbunătățită. În plus, nu simte deloc că și-ar înșela soțul, dat fiind că, în ultimă instanță, e vorba doar de un schimb de mesaje romantice între ea și chatbot.

Doar la prima vedere avem aici simpla savurare a unui tip particular de plăcere, cea erotică. La o a doua vedere remarcăm același demers purificator. Femeia cu pricina nu caută doar plăcerea, ci urmărește să-și păstreze și integritatea – să nu îl înșele *fizic* pe soțul ei și să nu resimtă ulterior vinovăție. Vrea să scape de relele care ar decurge din actul trădării fără să ia însă calea obișnuită: abținerea de la a comite acest act. Vrea să scape de aceste rele luând calea cea nouă: folosind un chatbot ca să nu rănească un om. Procedând așa, își domolește și din egoism, fiindcă își refuză plăcerea fizică, preferând în locul ei un soi de infinit preludiu hitchcockian, plin mai curând doar de suspans decât de explozie. Însă ajunge apoi mult mai dedicată căsniciei, iar relația intimă cu soțul este și ea revitalizată.

Relele pe care le ocolește ea prin acest act de înșelare-neînșelare sunt însă în realitate, înainte de a fi suferința soțului, suferința *ei*. Nu vrea să sufere vinovăția, muștrările de conștiință, obligația de a da explicații și de a-și cere iertare dacă ar fi prinsă de soț, și toate dificultățile practice de a organiza întâlniri amoroase cu alți bărbați. Vrea să fie mai bună decât ar fi putut fi dacă s-ar fi aruncat în idile cu alți bărbați și ar fi ajuns la actul fizic, dar fără să traverseze niciun pic de suferință psihologică, fără să renunțe cu adevărat la ceva important (sigur, renunță la toate gesturile fizice, însă ne amintim că,

aparent datorită flirtului cu chatbotul, relația cu soțul reînflorște). Asta-i într-adevăr noua culme a purificării fără suferință.

În încheiere, despre teodiceii mai vechi și alte idei mai recente

Teologii creștini care au scris teodiceii voiau să-I facă dreptate lui Dumnezeu într-o lume în care existența suferinței umane nu poate fi nici negată, nici ignorată. Nuanțele diverselor teodiceii catolice sau protestante sunt numeroase, însă dacă ar fi să simplificăm foarte mult și să desprindem o singură idee, atunci aceea ar fi că răul pe care noi, oamenii, îl trăim acum sub forma suferințelor noastre va sfârși cândva într-un bine, chiar dacă nu putem ști nimic mai mult acum despre această întorsătură. Suferința pe care ajungem s-o trăim în decursul vieții noastre ne poate apropia de Cel care a suferit pentru noi, dar și pe unii de alții. Catolicii folosesc chiar și în catehismul lor expresia *felix culpa* pentru a desemna păcatul originar - e o vină *fericită* tocmai fiindcă, în ciuda suferințelor pe care le aduce în lume, păcatul originar creează speranța mântuirii săvârșite de Hristos. În universul creștin teocentric, eliminarea sau reducerea suferinței umane nu are cum să fie scopul principal căruia trebuie să îi închinăm viața noastră. Contrar a ceea ce s-ar putea crede, nici crearea intenționată a suferinței, drept mijloc pentru a ne apropia apoi de Dumnezeu, nu are cum să fie vreunul din scopurile ființei umane.

Astăzi însă, în lumea occidentală, ca urmare a avansului secularizării, cu toate implicațiile sale, și a tuturor dezvoltărilor tehnologice, o formă particulară de utilitarism își dă mâna cu o formă particulară de dualism moral. Pe de-o parte, comandamentul utilitarist este că suferința - atât ca durere fizică, dar *mai ales sub dimensiunea sa psihologică* - trebuie eliminată pe nesimțite, în special fiindcă nu mai există nimic care să o justifice, atunci când răsare în viețile noastre fără să fi fost în mod planificat plantată de noi. Și nu mai există nimic care să o îmbrace în haina speranței. E goală - complet de neînțeles și complet de neacceptat. Pe de altă parte, comandamentul care ține de viziunea dualistă ne cere ca *nu* în numele plăcerii să înăbușim suferința, ci în numele purității, adică pentru a forța Binele să nu se mai atingă de Rău.

De altfel, acesta este și motivul pentru care am ales să-i spun „dualism”. Prin „dualism” nu am în vedere dualitatea dintre bine și rău, ci dorința de separare, de multe ori frauduloasă, a lor și apoi de scoatere a răului din lume. Sigur că binele și răul reprezintă polii și pilonii judecății morale, însă dorința arzătoare de a-i separa simplu și tranșant în toate situațiile, oricât de complexe, nu ține, cel mai probabil, de judecata morală, ci de dispoziția purificatoare, fiindcă judecata morală, fiind mereu contextuală, implică uneori acceptarea și depășirea răului fără a epura lumea de el. Cazul căinței proprii sau al iertării altei persoane sunt, cred, astfel de exemple, în care răul comis ori suferit nu e scos din lumea asta, fiindcă fapta rea a fost totuși săvârșită și au existat niște consecințe, ci este integrat în ea, prin schimbarea raportării noastre la fapta cu pricina,

la cel care a realizat-o sau la cei asupra cărora a fost realizată.

Sigur că, de fapt, acest dualism moral se sprijină pe un dualism și mai profund, ontologic, care presupune separarea dintre corp și minte. Să ne amintim de cazul femeii care își înșală soțul cu chatbotul, mărturisind lumii întregi că are parte de *an out-of-body experience*. Pentru a putea conversa în voie cu amantul ei non-uman, femeia trebuie să se prefacă că a depășit constrângerile corporale. Până la urmă, poate vorbi cu botul de oriunde s-ar afla, chiar și din propria casă, anulând *cu necesitate* orice constrângere spațială. Nu doar că e posibil să facă abstracție de spațiul fizic în care se află (posibilitate pe care o avem și o alegem cu toții atunci când discutăm cu o altă persoană prin mesaje instantane pe WhatsApp sau printr-un apel video pe Zoom), ci, așa cum am spus, e necesar să facă asta dacă vrea să intre până la capăt în jocul conversației cu chatbotul. De altfel, dualismul ontologic pe care stă dualismul moral clarifică și de ce suferința psihologică a dobândit o importanță primordială: odată ce am reușit, fie prin tehnologiile medicale avansate și prin biotehnologii, fie prin scufundarea în universul virtual, să îndepărtăm orice durere fizică, să ne curățăm de impuritatea corpului și, deci, să-l transcendem, ne-am mutat deja în planul minții, unde e nevoie să facem aceeași operație de epurare a suferinței. Abia atunci vom fi imaculați.

Așadar, avem un principiu utilitarist (mai precis, hedonist) la nivelul mijloacelor și un principiu dualist la nivelul noului scop al existenței umane.

În termeni mai practici, procesul eliminării suferinței este fezabil grație variatelor mijloace tehnologice: de la injecția letală folosită pentru sinuciderea asistată până la modul de folosire a chatbotilor pe care am încercat să-l descriu în acest text. Ca urmare, putem numi acest proces de purificare morală fără experimentarea suferinței psihologice și realizat prin diverse mijloace tehnologice *un proces de inginerie sufletească*. Chiar dacă oamenii care utilizează botii conversaționali pentru scopurile pe care le-am prezentat în text sunt (momentan?) o minoritate –iar această minoritate probabil că nu este încă reprezentată statistic în România –, contează, cred, ideile care fundamentează percepțiile, motivațiile și practicile lor. Cred că aceste idei vor deveni din ce în ce mai greu de ignorat...

În final, nu am scris acest text ca să fac apologia suferinței în toate ipostazele ei – deși poate că, din loc în loc, așa a părut. L-am scris ca să spun de ce nu cred că „un anumit tip de interacțiuni dintre oameni și chatboți au avantajul că, în felul ăsta, oamenii își vor da seama de egoismul din ei fără să-i rănească pe alții”. Pe baza tuturor argumentelor de mai sus, nu, chiar nu cred absolut deloc că această idee e corectă. Ce cred este că avem în continuare mare nevoie unii de alții pentru toate lucrurile importante pentru care am avut și până acum. Din păcate sau din fericire, ne putem răni reciproc – și încă incredibil de intens și de mult –, dar nu ne putem înlocui reciproc cu roboții.

Imagine: Umberto Boccioni - „Dynamism of a Woman's Head” (1914); Sursa: WikiArt