

## СВІДОМІСТЬ, КОМП'ЮТЕРИ ТА НОВИЙ МОЗОК МЕРІ

Д.П. Сепетий

Чи можуть комп'ютери, що виконують відповідну програму, мислити, мати суб'єктивні переживання, усвідомлювати тощо? Аргументація, що проливає світло на це питання, важлива як з точки зору філософського осмислення сучасних проблем штучного інтелекту, так і з точки зору загальної, вічної філософської проблеми розуміння природи свідомості, її відношення до фізичної реальності (матерії). В академічній філософії аналітичної традиції (що переважає в англomовних країнах), де філософія свідомості в останні десятиліття стала центральною, «комп'ютерний» аспект проблеми широко обговорюється. З одного боку, багато прибічників матеріалістичних напрямків, особливо функціоналізму, підтримують точку зору про можливість «сильного штучного інтелекту» (strong Artificial Intelligence) — реалізації справжнього мислення через виконання комп'ютером відповідної програми. Цю можливість допускають також деякі прибічники «слабких» версій психофізичного дуалізму (наприклад, Девід Чалмерс у своїй головній праці «The Conscious Mind»<sup>1</sup> [3]). Більшість опонентів матеріалізму заперечують таку можливість, як правило, на основі загальних аргументів проти матеріалізму. Існує також дуже впливовий і широко обговорюваний спеціальний аргумент — уявний експеримент Китайської кімнати, запропонований Джоном Сьорлем, — який, на переконання автора, доводить принципову нездоланну відмінність між комп'ютерною алгоритмічною обробкою інформації та мисленням, заснованим на розумінні [4]. Наявний стан дискусії залишає простір для

<sup>1</sup>На мій погляд, найбільш точним перекладом назви цього твору українською буде «Свідома психіка».

нових аргументів, що проливають певне світло на проблему як в загальнотеоретичній, так і в етичній перспективі.

## Історія Мері

Уявіть собі комп'ютерну програму під назвою «Мозок». Вона на вході отримує потоки сигналів, більш-менш аналогічні тим, які людський мозок отримує від нервової системи. Програма обробляє ці потоки за дуже складним алгоритмом, класифікує сигнали на зорові, слухові, нюхові, смакові, дотикові, больові і на виході формує інші потоки сигналів, аналогічні тим, які з людського мозку надходять до м'язів. Програма поділена на багато підпрограм, які імітують різні функціональні блоки мозку. Ці підпрограми частину вхідних даних отримують з входу програми «Мозок», а частину — від інших підпрограм, і, в свою чергу, формують потік даних на виході, частина якого йде на вхід до інших підпрограм, а частина — на вихід програми «Мозок». Серед цих підпрограм функціональних блоків є підпрограми з назвами «Зорові відчуття», «Біль», «Радість», «Сум», «Бажання» тощо.

Різні підпрограми програми «Мозок» час від часу зчитують дані з бази даних «Пам'ять» і записують в неї нові дані. Перед запуском програми початкові дані в базу даних «Пам'ять» може завантажити програміст.

Програма «Мозок» обмінюється даними з програмою «Тіло»: отримує від неї потоки даних, що імітують сигнали від нервової системи, обробляє їх і передає результат — потоки даних, що імітують сигнали, які надходять від мозку до рухових нервів, — до програми «Тіло».

Програма «Тіло» імітує функціонування всього тіла, крім мозку. Вона містить у собі багато підпрограм, таких як «Чуттєві нерви», «Рухові нерви», «М'язи», «Ліва рука», «Голова», «Серце» тощо, між якими відбувається складний обмін даними, що імітує відповідні процеси у тілі людини. (У спрощеному варіанті можна обійтися без більшості внутрішніх органів.)

Зрештою, остання програма — «Зовнішнє середовище» — імітація фізичного світу, в якому ми живемо. Програма формує різноманітні дані, які передає на вхід до програми «Тіло», і, у свою чергу, отримує від неї дані про рухи різних зовнішніх частин тіла людини (або тіл кількох людей). Зрозуміло, що програма «Зовнішнє середовище» буде містити в собі підпрограми «Земля», «Небо», «Трава», «Дерево», «Будинок», «Стіл», «Борщ» тощо для створення програмних моделей

об'єктів відповідних типів, а також з одного чи кількох програмних об'єктів типу «Людина», які міститимуть інформацію про віртуальне розташування у просторі віртуального людського тіла (усієї його зовнішньої поверхні) і будуть обмінюватися даними з програмою «Тіло». Зокрема, програма «Зовнішнє середовище» буде розраховувати (згідно з законами фізики) усі взаємодії між фізичними тілами, які вона моделює, динамічно формувати та змінювати зображення на моніторі людини та усього, що відбувається навколо неї і з нею.

Практично, створити таку модель (програму) так, щоб вона виглядала реалістично (відтворювала достатньо близько до дійсності достатньо складні моделі поведінки людини у різних ситуаціях) дуже важко. Хоча примітивні реалізації такої моделі цілком здійсненні. Напевно, в наукових центрах, які працюють над проблемами штучного інтелекту, вони вже є, — але поки що вони дуже далекі від реалістичного відтворення складної людської поведінки. Проте якими б недосконалыми не були ці реалізації, вони можуть вдосконалюватись і все краще моделювати функціонування людського мозку та людську поведінку. Здається, що для такого вдосконалення немає принципових меж. Якщо матеріалізм або епіфеноменалізм (теорія про те, що хоча свідомість і є чимось нематеріальним, вона є пасивним продуктом мозку і не впливає на жоден фізичний процес) є істинним, то в принципі можливо створити програмну модель, яка імітуватиме людську поведінку настільки добре, що ми не зможемо відрізнити цю імітацію від поведінки реальної людини.

Навряд чи нам, людям, колись вдасться створити таку модель. Але уявіть, що десь на далекій планеті Зет живуть розумні істоти, інтелектуальні здібності яких набагато вищі за наші людські. З якогось спостережного пункту вони вивчають поведінку людей і розробляють відповідні комп'ютерні моделі. Уявіть, що ви — вчений, який працює над проблемами штучного інтелекту. Одного чудового ранку зетіанський космічний корабель забрав вас на свою планету, і ось зетіани показують вам свої видатні наукові досягнення. Особливо вони пишаються своєю новою комп'ютерною моделлю людської психіки. На екрані монітора ви бачите ніби кінофільм, в якому людина в цілком звичайних ситуаціях поводить себе цілком звичайним чином. Але вам пояснюють, що це — не кінофільм, а програма, організована за описаним вище принципом. Вам пояснюють, що ви можете змінювати пам'ять та риси характеру віртуальної людини, завантажуючи різні шаблони. Ви можете маніпулювати оточенням багатьма різними способами, можете ставити людину, яку ви бачите на моніторі, у різноманітні життєві си-

туації і спостерігати, як вона поводитиметься. Наприклад, ви можете поставити цій людині підніжку, і подивитися, як вона відреагує, якщо впаде. Ви побачите вираз болю на її обличчі; можливо, ви почуєте лайку або стогін (до комп'ютера підключений динамік). Ви можете через мікрофон звернутися до людини і побачити, як вона здивується чи злякається, почувши це звертання і не побачивши нікого.

Поведінка людини, яку ви бачите на моніторі, виглядає цілком природною. Але чи Ви вважатимете (знаючи, як це зроблено), що в комп'ютері справді міститься людська особа, яка відчуває біль, здивування, страх, справді розуміє звернені до неї слова? Питання не про те, чи це справжня людина в смислі справжності її тіла (зрозуміло, що комп'ютерне зображення людського тіла і людське тіло — це зовсім різні речі). Питання про те, чи ви маєте справу зі справжньою свідомістю, справжнім людським Я, яке справді відчуває, бажає, мислить. Чи вважатимете ви, що в момент запуску програми починає існувати нове людське Я, і що воно припиняє своє існування (помирає) в момент зупинки програми? Навряд чи. Скоріше, Ви вважатимете, що це неймовірно добра імітація, але не справжнє людське Я.

Між іншим, зетіани розповіли вам, що реально працює не один комп'ютер, а система з трьох комп'ютерів, які обмінюються даними через спеціальний «інтерфейс» (спрощено можете уявити це як велику кількість дротів (волокон), якими сигнали передаються з одного комп'ютера на інший): на одному з них працює модель зовнішнього середовища (програма «Зовнішнє середовище»), що включає в себе і зовнішні обриси тіла, яке ви бачите на моніторі, на другому — модель тіла (крім мозку), на третьому — модель мозку з пам'яттю. Зетіани сказали, що цей поділ має значення для подальшої розробки їхнього проекту.

Минуло п'ять років. Вас знову забрали на планету Зет. Цього разу вам не показували комп'ютерів. Але ви мали змогу познайомитися з новим дивом зетіанської техніки — роботом Робі. Робі за поведінкою дуже схожий на людину. Він спілкується з вами як розумна істота, яка вас цілком розуміє; він каже, що відчуває біль, радість, сум, голод. Правда, оскільки він зроблений з металу й пластмаси, і внутрішні структури його тіла відрізняються від людських, то й зовнішні прояви його «відчуттів» дещо відрізняються. Так, він не «відчуває» болю в деяких із тих ситуацій, в яких його відчуваєте ви, оскільки вони не завдають пошкодження і не несуть загрози його тілу. Але якщо його вдарити кувалдою по руці, він закричить ніби від болю й побіжить «лікуватися» у ремонтну майстерню. «Голод» для нього настає, коли

виснажується електроенергія, і він «харчується», вставивши пальці в розетку.

Як зетіанам вдалося створити таке диво? Вони пояснили, що принцип — той самий, що й з віртуальними людьми, яких ви бачили минулого разу. Тільки зовнішнє середовище і тіло, крім мозку, цього разу реальні. Поки що тіло не людське, а штучне; але його взаємодії з зовнішнім середовищем та «мозком» здійснюються за тими ж принципами, що й у людини (і за тими ж, що й у віртуальній версії минулого разу). А от «мозок» — це комп'ютер, поміщений в голову робота, на якому працює програма «Мозок», дещо модифікована порівняно з минулим разом — пристосована до особливостей тілесної організації робота (зокрема, дещо змінені підпрограми «Біль» та «Голод»).

Розглянемо питання: чи цей робот справді відчуває біль, здивування, страх, справді розуміє звернені до нього слова? Чи є в нього справжня свідомість, справжнє Я, яке справді відчуває, бажає, мислить. Ви знаєте, що його «мозок», це, по-суті, такий самий комп'ютер з такою самою (несуттєво модифікованою) програмою, як і у випадку віртуальної людини, яку ви бачили на моніторі комп'ютера минулого разу. Очевидно, що якщо та віртуальна людина не мала справжнього Я, насправді не відчувала й не розуміла, а лише була неймовірно доброю імітацією, то у випадку робота ситуація та сама. Це — феноменальний зомбі — істота, що поводить ся так, ніби відчуває й розуміє, а проте насправді не має ніяких суб'єктивних переживань і розуміння, а лише дуже добре їх імітує.

Минуло ще п'ять років. Ви знову на планеті Зет. Цього разу вам не показують ані комп'ютерів, ані роботів — ніякої нової техніки. Вас познайомили з жінкою з Землі на ім'я Мері, яка є відомим вченим-нейрофізиком і вже п'ять років перебуває на планеті Зет, де бере участь у зетіанських наукових експериментах. Ви мило з нею поспілкувалися про наші земні справи та зетіанську науку. Але, на жаль, в усьому тому, що вона розповіла про свою участь в зетіанських експериментах, ви не знайшли нічого справді нового, важливого й цікавого. Зрештою, простирчавши на Зет тиждень і так і не побачивши ніяких нових див техніки, ви з розчаруванням звернулися до зетіан: «Нащо ви мене сюди привезли?!». Зетіани самовдоволено похіхікали і розповіли вам історію Мері.

Виявляється, Мері сама не знає головного про ті експерименти, в яких вона брала участь. Головне — це напрочуд вдалий експеримент з пересадки людині мозку — штучного замість органічного. Штучний мозок — це той самий комп'ютер, на якому працює програма «Мозок».

Зетіанам вдалося створити дуже компактну версію такого комп'ютера — за розмірами значно меншу, ніж справжній людський мозок. Їм вдалося створити також мікроскопічні «перехідники», які поєднують нервові закінчення (звичайно поєднані з нейронами мозку) зі штучним мозком і здійснюють потрібні перетворення сигналів. Їм також вдалося створити мікроскопічні трансформатори енергії, які перетворюють ту енергію, яка надходить до людського мозку, на електроенергію, на якій працює комп'ютер. Коротше кажучи, все обладнання, необхідне для заміни органічного мозку на штучний, з такими розмірами, що усе це справді вміщується на місці органічного мозку. Їм вдалося також зчитати всю пам'ять людини і записати її в комп'ютерну пам'ять штучного мозку. І перша ж операція дала чудовий результат. Людина, над якою її здійснили, має такі самі спогади, звички, характер, її розумові здібності навіть трошки покращилися (зетіани зробили великий резерв пам'яті та деякі інші вдосконалення). Минулого року вона літала на Землю, спілкувалася з рідними, друзями. Ніхто, навіть рідна мати, не запідозрили, що з нею щось сталося. Ця людина — Мері. Операцію зробили вночі під час сну; сама Мері про це не знає і ні про що не здогадується.

— Не буду від вас приховувати, — сказав зетіанин, — цього разу ми привезли вас не для того, щоб похвалитися своїми новими науковими досягненнями. Ми вирішили зробити ще одну пересадку комп'ютерного мозку людині, і обрали вас для цієї почесної місії. Не бійтеся, ви навіть нічого не відчуєте, просто заснете і прокинетеся з іншим мозком. Спогади, звички, характер — усе залишиться як і раніше. Інтелектуальні здібності стануть навіть кращими. Гарантія — 100 відсотків.

— Але ж це вбивство! Якщо все, що ви сказали, правда, то Мері вже немає, а є лише її дуже точна зомбі-імітація. Я відмовляюся.

— Ніяких «але». Годі з нас цих дурних земних розмов про суб'єктивність, «я» і т.п. Здавалося б, освічена людина, науковець, а вірить у такі дурниці! Ви б ще душу й Бога сюди приплели. А вашої згоди ніхто не питає. Науковий прогрес понад усе.

## Емерджентистська філософія свідомості Д. Чалмерса та Дж. Сьорля

В цій перспективі розглянемо погляди двох відомих сучасних філософів, яких можна віднести до напрямку емерджентизму — Девіда Чалмерса та Джона Сьорля. Емерджентизм — це теорія про те, що

свідомість виникає (emerges), коли людина народжується (або дещо раніше, ще в утробі) і зникає-знищується (перетворюється на ніщо) за інших. Вона не є тотожною з фізичними станами і процесами, які її породжують і визначають її стани, а є чимось новим, якісно відмінним від них. Загальна слабкість цієї теорії полягає в тому, що оскільки свідомість не утворюється з чогось уже існуючого, а виникає, то вона виявляється якоюсь дуже дивною річчю, яка виникає в буквальному сенсі з нічого і перетворюється на ніщо. Деякі прибічники цієї теорії (зокрема, Дж. Сьорль) вважають, що емерджентні якості — щось звичайне у природі. Натомість, Д. Чалмерс доводить, що, наскільки можливо судити з того, що ми знаємо про природу, свідомість є унікальною: в природі немає нічого іншого справді емерджентного (по відношенню до мікрофізичного базису).

Які ж умови відповідають за виникнення свідомості? Чалмерс і Сьорль представляють дві основні альтернативні версії, які можна назвати функціоналістською та субстратною.

Чалмерс відстоює думку про те, що «свідомість є “організаційний інваріант”»: тобто, що будь-яка система з правильною функціональною організацією буде мати той самий свідомий досвід, незалежно від того, з чого вона зроблена» [3, с. xv].

Це означає, що віртуальні люди з комп'ютерної програми зетіан (якщо вона достатньо точно імітує усі важливі функції мозку людини) мають повноцінне людське Я, свідомість, суб'єктивно переживають та усвідомлюють. Запускаючи програму, ви даєте життя новій людині, припиняючи її роботу — ви вбиваєте людину. Правда, здається, що ви завжди можете її воскресити, запустивши програму.

Але тут виникають проблеми. Ось ви дали життя віртуальній людині на ім'я «Джон». Коли вам набридло з ним бавитися, ви зберегли його пам'ять і закрили програму, тим самим вбивши Джона. Наступного дня ви вирішили воскресити Джона: включили комп'ютер, обрали потрібну базу пам'яті, яка зберігала пам'ять Джона, і запустили програму. Але випадково ви натиснули кнопку запуску програми двічі. І ось на комп'ютері відкрилося дві копії програми, в кожній з яких — свій власний Джон. Є два Джона, з однаковими спогадами й характером. Або кнопку, яка запускає програму, заїло і Джони почали плодитися один за одним. Хто з них — той Джон, якого ви вчора вбили? Чи жоден з них не є тим Джоном, а усі вони є новими людьми? Джон, якого ви вчора вбили, помер назавжди. Ви — вбивця Джона. А якщо ви позакриваєте ці програми, то станете серійним вбивцею.

Якщо припустити, що засаднича ідея програми «сильного штучно-

го інтелекту» — що комп'ютерні моделі можуть не лише імітувати, але й реалізувати свідомість — є істинною, то робота з реалізації цієї програми має бути заборонена!

Інший цікавий варіант (запропонований Недом Блоком): уявімо, що ціла нація чи мешканці цілої планети почнуть імітувати-відтворювати структуру й динаміку взаємодій, що мають місце у людському мозку (кожна особа грає роль одного нейрона) [2]. На думку Чалмерса, і в цій системі виникне нова людська свідомість «Я» [3, с. 252-253].

На цьому фоні більш правдоподібною може видатися позиція Джона Сьорля, який вважає, що свідомість існує завдяки певним властивостям того матеріалу, з якого утворюється мозок, разом з його структурною організацією; що свідомість не лише не можна створити на основі комп'ютерної функціоналістської моделі, але що вона навіть не може утворитися на основі іншого матеріалу, аніж той, з якого утворюється мозок (припустимо, що структурну організацію та функціональну динаміку мозку можливо реалізувати не на вуглеводневій, а на якійсь іншій хімічній основі, — хоча сумнівно, що це є фізично можливим).

Проте, Сьорль надто спрощує проблему і по суті ігнорує те, що власне породжує психофізичну проблему. Сьорль стверджує:

Мозок причинно зумовлює певні «психічні» феномени, такі як свідомі психічні стани, і ці стани просто є властивостями мозку на вищому рівні. Свідомість є властивістю мозку вищого рівня (або емерджентною властивістю) у цілком нешкідливому смислі слів «вищий рівень» та «емерджентний», подібному до того, як твердість є вищого рівня емерджентною властивістю молекул  $H_2O$ , коли вони мають решітчасту структуру (крига), а рідкий стан схожим чином є вищого рівня емерджентною властивістю молекул  $H_2O$ , коли вони, грубо кажучи, катаються одна навколо одної (вода). Свідомість є психічною і тому фізичною властивістю мозку в тому смислі, в якому рідкий стан є властивістю системи молекул [5, с. 14].

Але ж ця аналогія хибна! Справді, мозок причинно зумовлює психічні стани, а от молекули води та утворювана ними структура не зумовлюють причинно твердість чи рідкість. «Твердість» чи «рідкість» — це поняття, які описують на макрорівні ті самі структурно-динамічні властивості, які на мікрорівні описуються фразами «решітчаста структура молекул» та «молекули, які катаються одна навколо одної». В першому випадку маємо причинний зв'язок двох різних речей — фізичних станів мозку та психічних станів свідомості. У другому — два різні описи одного й того ж — сукупності молекул  $H_2O$ .



Джон Белоф так коментує твердження Сьорля, що свідомість є емерджентною властивістю мозку «так само, як твердість та рідкість є емерджентними властивостями систем молекул» [5, с. 112]:

Ця аналогія [...] сама себе зраджує. Ми маємо, зрештою, логічно послідовну теорію, яка пов'язує рух молекул з такими макровластивостями як твердість, рідкість, теплота тощо. [...] не існує жодної мислимій теорії, яка пов'язувала б зміст нашого особистого феноменального досвіду з тими процесами, які можливо спостерігати в нашому мозку з застосуванням відповідної апаратури. Сьорль, тим не менше, наполягає на своєму. Так, він пише: «Той факт, що психічні структури відповідають на вищому рівні<sup>2</sup> нейронним властивостям, жодною мірою не применшує їх причинної дієвості. Твердість клапана є причинно відповідною на вищому рівні щодо його молекулярної структури». Знову ж таки, аналогія розвалюється. Фізична хімія може пояснити чому «твердість клапана є причинно відповідною на вищому рівні щодо його молекулярної структури»; не існує аналогічної теорії, яка пояснювала б, чому досвід болю має відповідати на вищому рівні «мікроподіям в мозку». Ці мікроподії могли б мати місце і в разі, якби досвіду болю не існувало. Коротше кажучи, Сьорль ігнорує відмінність між високорівневими властивостями, з якими може впоратися фізикалістський аналіз, та [...] суб'єктивними переживаннями, чие існування ми мусимо визнати, але які явно не піддаються ніякому подібному поясненню [1, с. 515-516].

У випадку з твердістю клапана у відношенні до його молекулярної структури Сьорль робить ту ж саму помилку, що й у випадку з водою та молекулами  $H_2O$ : він не відрізняє причинний зв'язок від тотожності, зміщує їх під одним поняттям — відповідності на вищому рівні (надвідповідності, супервентності).

Щоб позбутися цієї — досить звичайної у філософії свідомості — плутанини, Девід Чалмерс запропонував розрізняти природну (причинну, каузальну) надвідповідність і логічну надвідповідність.

Логічна надвідповідність події (множини подій) А щодо множини подій В — це таке відношення, коли В (нижчий рівень) утворюють А (вищий рівень); В без А логічно неможливе; немає логічно можливих світів з В, але без А. Логічна надвідповідність означає тотожність, а не причинний зв'язок. Це — випадки води і молекул  $H_2O$ , твердості клапана та його молекулярної структури. Взагалі, усі матеріальні (фізичні в широкому смислі) факти та властивості є логічно надвідповідними щодо мікрофізичних фактів та властивостей. Свідомість (психіка) не

<sup>2</sup> «Відповідають на вищому рівні» — в оригіналі — *supervene* — поняття, якому немає прямого відповідника в українській мові. Детальніше про «відповідність на вищому рівні» (або «супервентність») див. далі.

є такою; саме тому вона не належить до царини фізичного (матеріального); і це означає, що матеріалізм хибний.

Природна (причинна, каузальна) надвідповідність А відносно В — це таке відношення у дійсному світі, що завжди, коли має місце В, то має місце і А; але немає логічної надвідповідності; логічно можливе В без А; логічно можливі світи, у яких при наявності В відсутня А. Природна (причинна) надвідповідність означає, що немає тотожності; є причинне відношення, зумовлене законами природи, що діють у нашому світі: А не тотожне В; А і В — різні речі (події); В не утворює, а спричинює А.

Якби свідомість була логічно надвідповідною щодо мозку, то це означало б, що матеріалізм істинний. Але свідомість не є логічно надвідповідною щодо мозку. Чалмерс припускає, що свідомість є природно (причинно) надвідповідною щодо мозку; це означало б, що слабкий дуалізм (епіфеноменалізм) істинний.

Я вважаю більш правдоподібним, що причинні відношення між свідомістю та мозком більш складні, двонаправлені; немає логічної надвідповідності, а причинна надвідповідність неповна, так що свідомість частково автономна, і є причинно-наслідкові зв'язки не тільки від мозку до свідомості, але й від свідомості до мозку. Якщо це так, то сильний дуалізм (інтераціоналізм) істинний.

А ось ще одне цікаве міркування Сьорля — про інтенціональність: чим би вона не була, вона є «так само каузально залежною від специфічної біохімії свого походження, як лактація, фотосинтез чи будь-які інші біологічні явища» [4, с. 147]. Це дуже нагадує знамениту заяву матеріаліста XIX ст. Карла Vogta: «мозок виробляє думки так само, як печінка — жовч». Не помічено лише однієї «дрібнички»: ми можемо взяти біохімічні продукти лактації, фотосинтезу чи продукування жовчі, піддати їх хімічному аналізу, з'ясувати, з яких молекул та атомів вони складаються. А як щодо хімічного аналізу й складу інтенціональності та думок? Процеси лактації, фотосинтезу, продукування жовчу — як і будь-які фізичні (у широкому сенсі) процеси — це лише велика множина фізичних рухів, зміни взаємного розташування різних фізичних тіл (атомів, іонів, електронів тощо) і утворення ними нових структур. Але як би ми не групували атоми та інші мікрочастинки і як би ми не описували їх структури, ми не отримаємо свідомості й інтенціональності. Ці атоми та інші мікрочастинки могли б рухатися точно так само і утворювати такі самі структури без ніякої свідомості. Таку можливість драматично представляє аргумент зомбі, який захищає Девід Чалмерс.

Ми можемо уявити собі феноменального зомбі — людиноподібне тіло, яке зовні поводить себе (у відповідних ситуаціях здійснює відповідні рухи відповідних частин тіла) точно так само, як і людина; яке утворюється з атомів і молекул з точно такими ж фізичними властивостями, упорядкованими в точно такі ж просторові структури, як і в людському тілі; всередині якого відбуваються точно такі ж рухи атомів і молекул, як і всередині людського тіла, — і яке при цьому абсолютно нічого не відчуває й не усвідомлює, — усі його внутрішні й зовнішні рухи здійснюються чисто автоматично, як результат фізичних взаємодій. Адже в фізичних взаємодіях немає нічого, що передбачало б суб'єктивні переживання й усвідомлення; вони здійснюються автоматично за законами фізики. (Саме це, власне, й розділяє фізичні процеси та свідомість.)

Звичайно, можливість реального існування таких зомбі видається дуже малоправдоподібною, — але лише тому, що кожний з нас знає, з власного досвіду, наскільки його поведінка залежить від його свідомості (суб'єктивних переживань та мислення) і наскільки тілесні процеси впливають на його свідомість. Але з точки зору матеріалізму усе це — лише ілюзія — дуже дивна, нічия ілюзія. Усе, що реально відбувається — це фізичні взаємодії різноманітних часточок нашого тіла, які здійснюються автоматично за законами фізики, — точно так само, як і в уявленому нами випадку зомбі. А це означає, що насправді ми і є такими зомбі.

Але ми знаємо, що це не так. Кожен з нас (навіть ті матеріалісти, які це заперечують) знає, що крім того, що різні частини його тіла здійснюють різноманітні рухи, що атоми й молекули всередині його тіла взаємодіють і рухаються (автоматично) за фізичними законами, відбувається ще дещо: він суб'єктивно відчуває біль, тепло, холод, радість, сум, любов, ненависть, він бажає чогось, він мислить про щось, і він усвідомлює, що він відчуває, бажає й мислить. І, зрештою, він існує не лише як фізичне тіло, але й як істота, що суб'єктивно відчуває, бажає, мислить та усвідомлює.

Загалом, у світлі запропонованого уявного експерименту, припущення про те, що виконання комп'ютером відповідної програми може бути реалізацією справжнього мислення, заснованого на розумінні, справжньої свідомості, виглядає малоправдоподібним. Більш того, наукові дослідження, спрямовані на створення «штучного інтелекту», є етично неприпустимими, оскільки (в разі їх успішності) вони включають регулярні вбивства людських осіб. Хоча філософія свідомості Девіда Чалмерса (близький до функціоналізму епіфеноменалізм) по-

діляє зазначені недоліки, вона містить важливі аргументи проти матеріалізму. Більш правдоподібна в частині заперечення функціоналістського підходу до свідомості філософія свідомості Джона Сьорля, але натомість він тлумачить свідомість на основі хибних аналогій, що не враховують належним чином унікального характеру свідомості як особистого «суб'єктивного виміру», який (як показують, зокрема, аргументи Д. Чалмерса) неможливо зрозуміти в термінах фізичних процесів. Таким чином, врахування переваг і недоліків теорій Чалмерса і Сьорля вказує в напрямку до більш «сильної» — інтеракціоністської — версії дуалізму. Потребує окремого обговорення запропонований Дж. Сьорлем уявний експеримент Китайської кімнати та дискусія, що розгорнулася довкола нього в аналітичній філософії. В філософії Д. Чалмерса більш детального розгляду заслуговують його аргументи на підтримку епіфеноменалізму, а також раселіанського монізму, до якого він схилився у період після «The Conscious Mind». Загалом, в «західній» філософії свідомості кількох останніх десятиліть запропоновано чимало цікавих підходів та нової аргументації, з якими український читач, на жаль, мало знайомий; їх дослідження та спроби розвитку сприятимуть поглибленню й збагаченню розуміння природи свідомості, її відношення до мозку і загалом до фізичної реальності.

## Література

- [1] *Beloff J.* The mind-body problem // The Journal of Scientific Exploration. — 1994. — Vol. 8. — №. 4. — С. 509-522.
- [2] *Block N.* Troubles with functionalism // Perception and Cognition: Issues in the Foundation of Psychology / Ed. by C.W. Savage. — Minneapolis: University of Minnesota Press, 1978. — С. 261-325.
- [3] *Chalmers D.* The Conscious Mind. — New York, Oxford: Oxford University Press, 1996.
- [4] *Searle J.* Minds, brains and programs // The Philosophy of Mind: Classical Problems / Contemporary Issues / Ed. by B. Beakley and P. Ludlow. — Cambridge, London: The MIT Press, 2006. — С. 133-149.
- [5] *Searle J.* The Rediscovery of the Mind. — A Bradford Book, 1992.

*Надійшла до редакції 25 лютого 2014 р.*