

УДК 004.383.8:001.53

О. В. Артьомов, канд. мед. наук,  
Е. С. Бурячківський\*, канд. мед. наук

**ІНТЕЛЕКТ ПОЗА БІОЛОГІЧНИМ НОСІЄМ  
(МЕДИКО-БІОЛОГІЧНИЙ І ФІЛОСОФСЬКИЙ ПОГЛЯД  
НА ПРОБЛЕМУ ШТУЧНОГО РОЗУМУ)**

*ДУ «Інститут очних хвороб і тканинної терапії  
ім. В. П. Філатова НАМН України», Одеса, Україна,*

*\* Одеський національний медичний університет, Одеса, Україна*

---

УДК 004.383.8:001.53

О. В. Артьомов, Е. С. Бурячківський\*

**ІНТЕЛЕКТ ПОЗА БІОЛОГІЧНИМ НОСІЄМ  
(МЕДИКО-БІОЛОГІЧНИЙ І ФІЛОСОФСЬКИЙ ПОГЛЯД  
НА ПРОБЛЕМУ ШТУЧНОГО РОЗУМУ)**

*ДУ «Інститут очних хвороб і тканинної терапії ім. В. П. Філатова НАМН України», Одеса, Україна,*

*\* Одеський національний медичний університет, Одеса, Україна*

Бурхливий розвиток інформаційно-комп'ютерних технологій призвів до завищеної оцінки їх реальних можливостей. Так, необґрунтовані думки про можливості штучного інтелекту нерідко пов'язані зі спробами оцінити наявність розуму за допомогою тесту Тьюрінга. Разом із тим будь-яка розумова діяльність, яку може продемонструвати комп'ютер, потребує не мислення, а алгоритмування — маніпуляцій з синтаксичними конструкціями. Становлення людської індивідуальності починається не з інформації, а з відчуттів — електрофізіологічних процесів, що не створюють якої-небудь інтелектуальної продукції, що може бути об'єктом тестування. Це несвідоме формує усвідомлення власної індивідуальності та власної цінності. Тому мислення і розум назавжди залишаться прерогативою людської особи, оскільки здатність мислити породжена не алгоритмом, а тим, що його ініціює — електрофізіологічним процесом, що відбувається в нейронних мережах проміжного мозку. Цей генератор несвідомого своєю постійною присутністю і дією на неокортекс створює в ньому не тільки розуміння того, що він робить, але і того, що це робить саме він. Таким чином, людський розум — це не тільки алгоритм, але і те не-алгоритмізоване, що його створює.

**Ключові слова:** штучний інтелект, розум, особистість.

UDC 004.383.8:001.53

O. V. Artyomov, E. S. Buryachkivsky\*

**INTELLIGENCE OUT OF THE BIOLOGICAL MEDIA  
(MEDICAL-BIOLOGICAL AND PHILOSOPHICAL VIEW ON  
THE PROBLEM OF ARTIFICIAL MIND)**

*SI "The Filatov Institute of Eye Diseases and Tissue Therapy of  
the National Academy of Medical Sciences of Ukraine", Odesa, Ukraine,*

*\* The Odesa National Medical University, Odesa, Ukraine*

Information and computer technologies have approached a level that allows us to discuss not only the possibility of creating an artificial intelligence, but also the prospects for the existence of a natural mind outside the scope of the carrier, the human being, assigned to it by nature. Unreasonable overestimation of the real possibilities of information technology is due to the fact that their authors rely on the criterion of the computer's intelligence from the position of the Turing test.

At the same time, any thinking activity, the result of which can be demonstrated by a computer, does not require thinking, since it requires manipulation with syntactic constructions. The formation of a human individuality does not begin with the accumulation of information, but with sensations, i. e. perception of external signals and their transfer to neural networks.

Computer intellect requires an algorithm from the very beginning, and at the origins of the formation of human consciousness there is an impulse — an electrophysiological influence that does not carry information and does not create any intellectual products that could be the object of testing. Thus, the human individuality begins not with the processing of information, but with that precedes it — the unconscious. The later the unconscious creates an awareness of one's own individuality and own value. This non-algorithmic awareness can not be reproduced or created as a computer program. The unconscious — “Me-in-itself”, which forms an individuality, appears before us in the form of the “Me-for-us”.

Therefore, the Turing test does not allow to prove the presence of thinking of artificial intelligence as the presence of the “Me-in-myself”, i. e. consciousness of one's own worth. So, under the same circumstances, the functionalists can state that an artificial intelligence has a consciousness, since it answers all conceivable questions in the same way as a person.

However, for the physicalists this argument proves nothing, and their position here is entirely in the spirit of Kant. Thus, the problem of artificial intelligence lies in the features of the formation and organization of human intelligence, and not in the possibilities of information technology. Information technologies based on algorithms can not connect the conscious and the unconscious — what happens in the process of the formation of human individuality. Therefore, thinking and reason will forever remain the prerogative of the human person.

After all, the ability to think is generated not by an algorithm, but by that initiates it — by an electrophysiological process occurring in the neural networks of the diencephalon of the brain. This peculiar generator of the unconscious, by its constant attendance and influence on the neocortex, creates in that not only an understanding of what it is doing, but also of the fact that it is *he* who does it. Thus, the human mind is not only an algorithm, but also the non-algorithmizable, which creates it.

**Key words:** artificial intellect, mind, identity.

Сучасні інформаційно-комп'ютерні технології наблизилися до такого рівня, який дозволяє предметно обговорювати не тільки можливість створення штучного інтелекту (ШІ), а й навіть перспективи існування природного розуму за рамками відведеного йому природою носія — людини. Так, поряд з поширеним поняттям завантаження (uploading, downloading), у світі інформаційних технологій в останні роки набув популярності термін сетлеретика (від англ. *settle* — розміщувати, заселяти) як інноваційний напрям у галузі інформаційних технологій. Цей термін вбирає в себе комплекс ідей так званого трансгуманізму, що припускають можливість перенесення особистості в комп'ютер з наступним необмежено довгим її існуванням у віртуальному світі. Інакше кажучи, проблему безсмертя людського тіла, яку не можуть розв'язати медицина і біологія, пропонують замінити цілком розв'язуваним, з точки зору інформаційних технологій, завданням перенесення особистості на синтетичний носій.

Звичайно, сучасні технології дозволяють створювати інформаційні програми, здатні відтворювати не тільки багато сторін людської діяльності та творчості, а й навіть імітувати почуття шляхом оцінки параметрів зовнішнього або внутрішнього середовища та реагування на їхні зміни. Разом з тим, не торкаючись самих перспектив, варто поставити собі питання: чи є створення штучного розуму чисто технічною проблемою? Для аналізу цього питання необхідне всебічне осмислення таких базових понять, як інтелект і особистість, не тільки з позиції кібернетики, а й медицини та біології. Однак сьогодні в рамках

інформаційних технологій, з позиції медико-біологічних наук і навіть у філософії існують розбіжності в розумінні того, що таке інтелект або особистість. Тим не менш, не вирішивши питання про суть людського розуму, не можна ставити завдання трансгуманістичного значення, як це робить сетлеретика. Потрібно відзначити, що необґрунтовані заяви і переоцінка реальних можливостей інформаційних технологій пов'язані з тим, що їхні автори спираються на критерій розумності комп'ютера з позиції тесту Тьюрінга, запропонованого ще в середині минулого століття. Так, на світанку комп'ютерних технологій А. Тьюрінг запропонував визначати ШІ за здатністю таким чином підтримувати розмову зі звичайною людиною, що остання не зможе зрозуміти, чи розмовляє вона з іншою людиною або з машиною. Такий ШІ буде проявляти інтелектуальну поведінку, що не відрізняється від людської в нормальних ситуаціях. Припускають, що саме на порозі розв'язання цієї проблеми стоять сучасні інформаційні технології, що породжує ілюзію появи в недалекій перспективі штучного розуму, рівноцінного людській особистості. Більше того, у суспільній свідомості вже насаджуються побоювання: чи не може штучна істота, що перевершує людину за інтелектом і володіє свободою волі, витіснити саму людину? Деякі схильні до епатажу відомі фахівці з ШІ, як, наприклад, Рей Курцвейл і Девід Чалмерс, активно обговорюють апокаліптичні наслідки наукового прогресу, нерідко переводячи дискусію в площину наукової фантастики, де з подачі Вернора Вінджа стає популярною тема так званої технологічної сингу-

лярності — драматичних змін у суспільстві, пов'язаних з появою ШІ [3].

Як уже було відзначено вище, подібні думки значною мірою викликані переоцінкою тесту Тьюрінга в оцінці кінцевої мети на шляху до створення штучного розуму. На некоректність такого підходу багато разів указував Джон Серл, який заявляв, що не можна вважати систему мислячою тільки на тій підставі, що вона здатна пройти тест Тьюрінга. Критикуючи концепцію так званого сильного ШІ, Дж. Серл припускав за комп'ютерами лише імітацію ментальних функцій людини, що дозволяє зараховувати їх до слабких ШІ. Будь-яка розумова діяльність, результат якої може продемонструвати ШІ, на думку Дж. Серла, не потребує від комп'ютера мислення, оскільки для цього достатньо маніпуляцій із синтаксичними конструкціями. Тому сильний ШІ, на його думку, неможливий у принципі [1].

Є чимало противників цього підходу (С. Пінкер, Д. Чалмерс, Д. Хофштадтер), проте їхня позиція, зокрема, аргумент «філософський зомбі», багатьма визнається непереконливою. Більш обережні критики, що допускають сильний ШІ, виходять з того, що такий можливий на основі самонавчальної програми і сенсорів для сприйняття світу. Такий ШІ зможе взаємодіяти зі світом, пізнавати його і бути здатним навчитися подібно до дитини, а, отже, повинен мати свідомість як людина. Однак чи не прийдемо ми в цьому випадку до того, що робот виявиться не сильним ШІ, а штучно створеною людиною? Подібні думки висловлював Дж. Серл у своїй концепції біологічного натуралізму, що зводиться до ексклюзивності свідомості, підтримуваного людським мозком і його нейронними процесами [2].

Протилежна, вкрай радикальна, позиція ґрунтується не тільки на тому, що свідомість має виключно інформаційну природу, а й допускає можливість її копіювання і перенесення на інженерний пристрій. Цієї позиції так чи інакше дотримуються багато прихильників так званого функціоналізму, які ототожнюють свідомість з функціями, які вона виконує. Сетлеретика лише проектує цю ідею в площину практичної значущості та технічної здійсненності. Так, прихильники концепції перенесення свідомості допускають сканування людської свідомості без руйнування при цьому мозку і подальше розміщення отриманої кальки на адекватному носії так, щоб структура людського мозку була повністю відтворена. Чи можна вважати це кроком до безсмертя особистості? Адже якщо власник оригіналу постаріє і

помре, то копія людського розуму зможе безмежно довго існувати.

Ця позиція, з точки зору здорового глузду, уявляється абсурдною, проте досить непроста в плані наукового спростування. Дійсно, не можна заперечувати, що розум забезпечений матеріальною структурою, а його діяльність можна алгоритмізувати. Адже здатність нейронних мереж переробляти інформацію не залежить від будь-яких віталістичних, тобто нефізичних, сил, які не можуть бути відтворені руками людини. У такому разі, що, крім знань і технічних можливостей, заважає повторити подібну структуру й алгоритм на іншому носії? Адже показана можливість створення *in vitro* головного носія життя — гена (ДНК) — з органічних і неорганічних елементів. Тому хоча науково-технічна база для створення штучної свідомості далека від вимог заявленої мети, деяким ентузіастам ідея не здається принципово недосяжною; і вони вже готові відрахувати десятиліття, призначаючи дату «технологічної сингулярності».

Так чи інакше суперечка про можливість ШІ повертає нас до визначень особистості, індивідуальності та інтелекту, які в медико-біологічній науці сформувалися ще до комп'ютерної ери. Наприклад, індивідуальність зазвичай визначається як сукупність характерних особливостей і властивостей, що відрізняють одного індивіда від іншого за своєрідністю психіки та інших якостей. Поняття особистості практично рівнозначне, але застосовується виключно до людини для відображення її соціальної природи і як суб'єкта соціокультурного життя, носія індивідуальної основи, що розкривається в рамках соціальних відносин, спілкування і предметної діяльності. Усі ці визначення позиціонують особистість як форму прояву людської індивідуальності. Як можна помітити, при такому підході від ШІ досить демонстрації інтелектуальних здібностей і знань індивідуума, тобто того, на чому концентрують увагу прихильники функціоналізму, що припускають можливість створення сильного ШІ.

Разом із тим потрібно враховувати, що для людського індивідуума найважливішим елементом є самосвідомість, тобто свідомість самого себе, яка створює уявний образ «Я», що не може бути відображений у будь-якій зовнішній інтелектуальній функції. Важливо відзначити, що становлення людської індивідуальності починається не з накопичення інформації, а з відчуттів, тобто сприйняття зовнішніх сигналів і передачі їх у нейронні мережі. При цьому спочатку навіть

такі складні сенсорні сприйняття, як зір і слух, без яких важко уявити формування інтелекту, не пов'язані з інтелектуальними актами. На початковому етапі формування особистості немає розуміння того, що відбувається, — воно приходить поступово, починаючись з безумовних рефлексів і поступово залучаючи нейронні мережі неокортекса. Тому новонароджена дитина, маючи всі анатомічні структури для зору і слуху, протягом багатьох днів не тільки не бачить як доросла людина і не сприймає мову, але навіть далеко не відразу починає відчувати біль. Звичайно, вона його сприймає, бо має для цього і рецептори, і відповідні нейрони для формування безумовних рефлексів. Однак поки не відбудеться усвідомлення власної індивідуальності, нейронні процеси, які відбуваються на периферії, не будуть аналізуватися на рівні більш високих нейронних мереж. У дорослої людини може відбутися повернення до даної ситуації: при деяких тяжких психічних розладах, що ведуть до втрати власного Я, людина, яка втратила поняття власної особистості, може не відчувати біль. При цьому можуть зникати і вегетосоматичні рефлекси, асоційовані з болем, як, наприклад, мідріаз. Це побічно вказує на те, що формування свідомості (вірніше, усвідомлення себе цілісною особистістю) починається вже в структурах проміжного мозку, складовою частиною якого є таламічний фільтр, який взаємодіє з підкірковими вузлами кінцевого мозку. Інакше кажучи, все це відбувається набагато нижче неокортекса, де працюють алгоритми, що забезпечують інтелект.

Таким чином, якщо при створенні комп'ютерного інтелекту вже з самого початку «було слово» — алгоритм, то біля витоків формування людської свідомості лежить імпульс — фізичний (електрофізіологічний) вплив на окремі нейрони, що не несе інформації, тобто не викликає будь-якої інтелектуальної продукції, яка могла б бути об'єктом тестування. Так, людське Я-індивідуальність починається не з обробки інформації, а з того, що передує цьому, — з несвідомого. Присутність несвідомого в подальшому формує усвідомлення власної індивідуальності, в числі іншого виявляється в усвідомленні власної цінності. Це не має алгоритму: усвідомлення не можна відтворити або створити заново як комп'ютерну програму, хоча, безсумнівно, воно існує на основі пам'яті і пов'язаної з цим інформації. Звичайно, зовнішні елементи когнітивного сприйняття любові до себе комп'ютер може відтворити. Однак це буде любов Я-для-нас, тимчасом як любов Я-в-собі не може бути зведена до зовнішніх про-

явів, вона не доступна спостерігачеві в повному обсязі, тому що залишається в несвідомому.

Тут з'являється ще одна особливість людського інтелекту, що відрізняє його від ШІ. Так, пам'ять комп'ютера існує як алгоритм відтворення того, що в ньому закладено, тобто як зв'язок між двома або кількома інформаційними програмами. Пам'ять людини значно ширша: вона допускає створення образу того, що з самого початку не має сенсу, чого не можна уявити інформаційно, і не допускає алгоритмізації. З цим, зокрема, пов'язаний обговорюваний з середини минулого століття феномен так званої *квалія*, деякі дослідники вважають епіфеноменологічними проявами свідомості.

У цьому, мабуть, і полягає феномен Я як власної цінності, не доступної для комп'ютерного інтелекту. Звичайно, слідуючи тесту Тьюрінга, цей феномен також можна «нав'язати» комп'ютеру. Однак це буде лише імітація почуття власної цінності при його реальній відсутності.

Як видно, несвідоме Я (Я-в-собі), яке формує індивідуальність, постає перед нами у формі Я-для-нас. Тому тест Тьюрінга не дозволяє довести наявність у ШІ свідомості, тобто присутність усвідомлення власної цінності — Я-в-собі. У цьому сенсі Я-в-собі можна розглядати як трансцендентну частину свідомості, своєрідну проекцію кантовського ноумена зі світу пізнаваних речей на ту єдину «річ», за допомогою якої здійснюється саме пізнання.

Як відомо, поняття «трансцендентності» у філософії І. Канта використовується для характеристики ноуменів — речей-в-собі, які не проявляються у вигляді доступних спостереженню явищ, проте не можуть бути безпосередньо сприйняті в чуттєвому досвіді, і про їхнє існування ми дізнаємося лише умоглядно. Тому, за одних і тих же обставин, функціоналісти можуть заявити, що якийсь ШІ має свідомість, тому що на всі мислимі запитання відповідає абсолютно так само, як людина. Однак для фізікалістів (як Дж. Серл) цей аргумент нічого не доводить, і їхня позиція тут цілком у дусі Канта.

Повертаючись до сетлеретики, потрібно відзначити, що Я-в-собі не можна повторити не тільки як інформаційну програму (воно, в силу сказаного вище, просто не може бути такою), але і як інший матеріальний носій, хоча те, що зумовлює Я-в-собі, безсумнівно, матеріальне, тобто має структуру. Це зумовлено тим, що свідомість сама створює себе на основі примітивних електрофізіологічних актів, які поступово формують несвідоме — річ-в-собі, і нарешті постає у вигляді



свідомої частини, що розкривається як річ-для-нас. Саме остання частина свідомого, втілена в особистості, і фігурує практично в усіх дефініціях, пропорованих у психології або філософії. Без належної критики це підхоплюється інформаційними технологіями, деякі представники яких починають поспішно заявляти про можливість створення ШІ, хоча в реальності даний напрям наукової думки здатний відтворити лише одну з частин людського інтелекту, яка не представляє всю особистість.

Таким чином, проблема полягає в особливостях становлення й організації людського інтелекту, а не в можливостях інформаційних технологій. Інформаційні технології, засновані на алгоритмах, не здатні зв'язати свідоме і несвідоме — те, що відбувається в процесі формування людської індивідуальності. Намагаючись повторити людський інтелект, ми розриваємо тривалий еволюційний шлях його становлення, біля витоків якого стоять нейрофізіологічні акти, засновані на безумовних рефлексів, тобто які не потребують мислення (свідомості). Саме ці акти забезпечують як фізіологічну діяльність, так і виживання організмів. Наприклад, примітивній гідри не потрібне мислення, щоб піти з небезпечного середовища, — для цього досить безпосередньої трансформації рецепції в рух. У міру ускладнення нервової системи — від гангліїв до багатоступеневої ієрархії підкіркових структур і кори термінального мозку — лавиноподібно збільшується кількість нервових клітин і утворених ними нервових мереж. Ці мережі не обов'язкові для здійснення життєво важливих функцій, що забезпечують метаболізм і пов'язану з ним рухову активність. Однак еволюційний розвиток, усе далі віддаляючи від периферії частину нервових мереж, перетворює їх у наглядача за більш примітивними нервовими процесами. Із кількох рефлексорних дуг неокортекс починає вибирати більш ефективні для організму. Так формується аналізуюча система, що зіставляє сигнали від різних нейросенсорів, щоб перетворити їх у найбільш ефективну результуючу відповідь. Такий аналіз починається вже на рівні довгастого мозку і таламуса, хоча навряд чи їх можна назвати анатомічним притулком свідомості. Проте на цьому рівні нервової діяльності вже формується алгоритм, втім, така робота безглузда з позиції комп'ютерних технологій — вона суперечить їхній суті. Адже аналітична робота в підкіркових і стовбурових структурах спрямована на фільтрацію нейрофізіологічного шуму — безлічі нервових імпульсів. Цей «шум» не має самостійного

значення; поки з безлічі нервових сигналів не будуть виділені ті, що сформуєть алгоритм, який дійде до кори, їх не можна побачити як свідомі акти. Тому всі комп'ютерні програми спрямовані на розпізнання і відтворення цього останнього етапу, тобто для комп'ютера свідомість починається відразу ж з осмисленого сигналу — тут зайві шуми, що не несуть конкретної інформації. Комп'ютеру не потрібне несвідоме, хоча воно необхідне для становлення розуму в живому організмі. У цьому несвідомому ховається особистість і власне Я — не те, яке може бути доступне тесту Тьюрінга у вигляді «розумної» відповіді, а яке назавжди залишається річчю-в-собі. Непізнаність цього Я зумовлюється не лише тим, що воно не має матеріальної структури. Структури, що забезпечують розум цілком матеріальні, а значить, і відтворювані. Однак відтворювані вони тільки в процесі свого створення, коли в тому єдиному, яке повинно стати особистістю, повторюється весь еволюційний шлях — від примітивних рефлексів до інстинктів, потім до несвідомого, а від останнього — до людського інтелекту. Звичайно, в принципі інформаційні технології можуть повторити практично все, що має алгоритм. Однак вони не можуть повторити те, що привело до створення розуму і залишилося прихованим у ньому як матеріальний носій особистості — «Я-в-собі», яке в нашій повсякденній свідомості залишається просто власним Я.

**Ключові слова:** штучний інтелект, розум, особистість.

#### ЛІТЕРАТУРА

1. Серл Дж. Сознание, мозг и программы. *Аналитическая философия: становление и развитие*. Москва, 1998.
2. Серл Дж. Рациональность в действии / пер. с англ. А. Колодия, Е. Румянцевой. Москва: Прогресс-Традиция, 2004. 336 с.
3. [https://ru.wikipedia.org/wiki/%D0%A1%D0%BE%D0%B7%D0%BD%D0%B0%D0%BD%D0%B8%D0%B5\\_\(%D0%BF%D1%81%D0%B8%D1%85%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%8F\)](https://ru.wikipedia.org/wiki/%D0%A1%D0%BE%D0%B7%D0%BD%D0%B0%D0%BD%D0%B8%D0%B5_(%D0%BF%D1%81%D0%B8%D1%85%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%8F)).

*Надійшла до редакції 23.10.2018*

*Рецензент д-р мед. наук, проф. І. В. Савицький,  
дата рецензії 25.10.2018*