

А. Матвійчук

## МОЖЛИВОСТІ ТА ПЕРСПЕКТИВИ СТВОРЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ

---

*Досягнення в теорії обчислень, формальній логіці, поява комп'ютерної техніки в середині ХХ ст. стимулювали дослідження можливості математичного відтворення розумової діяльності, створення штучного інтелекту. З огляду на традицію західної філософської думки, яка розвивалась під впливом платонівського вчення, роботи були орієнтовані переважно на відтворення розумової діяльності шляхом маніпулювання символічними виразами згідно з правилами, що враховують їхню структуру. У пропонованій статті розглянуто передумови зародження наукового напрямку з проблеми штучного інтелекту, сучасні можливості створення штучно-інтелектуальних технологій і ймовірні наслідки появи в майбутньому мозкоподібних систем.*

Ідеї Платона й інших афінських мислителів чітко визначили розвиток філософії, логіки, власне, стали дороговказом для більшості наукових напрямів. І сьогодні надзвичайно важко відійти від тисячолітніх узвичаєнь і спрямувати науковий пошук у русло, яке суперечило б давньогрецькому світогляду. Зокрема, Платон зазначав: усяке знання слід подавати у вигляді точних визначень, якими зможе користуватись будь-хто. Якщо ж цього не відбувається, тобто знання про те, як щось робиться, не можна перетворити у знання про те, що робити, — то це не знання, а віра, упевненість [1]. Отже, за Платоном, роботу художників, поетів і, врешті, науковців, які творять щось принципово нове залежно від натхнення, годі описати точними правилами. Діяльність, де потрібні майстерність, інтуїція чи відчуття традиції, Платон називає «беззмістовною одержимістю».

Погляди Платона передбачають цілковиту відмову від інтуїції та оцінок, що мають характер міркувань. Для цього треба звести

всі семантичні зв'язки (сміслові значення об'єктів і співвідношення між ними) до формальних синтаксичних конструкцій. Переконання щодо можливості такої тотальної формалізації пізнання багато століть домінує у західній думці.

Уперше синтаксичну концепцію мислення як процесу обчислення сформулював у 1651 р. англійський філософ і літератор Томас Гоббс: «Коли людина міркує, вона лише створює в умі підсумкову суму шляхом складання частин ... адже міркування ... це не що інше, як розрахунок» [2; 75, 76]. Німецький філософ і математик ХVІІ ст., винахідник двійкової системи обчислення (покладеної в основу цифрових комп'ютерів) Готфрід Лейбніц працював над створенням універсальної мови міркування, що зведе звичайну аргументацію до обчислень. Її базисом будуть універсальні характеристики — характеристичні числа і правила їх комбінування, за допомогою яких можна описати будь-яку ситуацію чи розв'язати будь-яку проблему: «І якщо хто-небудь

---

© МАТВІЙЧУК Андрій Вікторович. Доктор економічних наук. Директор Інституту моделювання та інформаційних технологій в економіці Київського національного економічного університету імені Вадима Гетьмана. 2011.

мав би сумніви у тому, що я стверджую, я б відповів йому: «Давайте розрахуємо, панове!» — і ми, взявши перо і чорнила, швидко б вийшли із ускладнення» [3].

Переконання Гоббса і Лейбніца сформувалися під впливом платонівської концепції. Ці вчені створили синтаксичну теорію мислення, висунули гіпотезу про можливість формалізації міркувань із застосуванням універсальних символів і операцій над ними, що стало провідною ідеєю у визначенні в 50-х рр. XX ст. програми досліджень зі створення штучного інтелекту.

### ПЕРШІ КРОКИ ШТУЧНОГО ІНТЕЛЕКТУ

**Ш**тучний інтелект як науковий напрям зародився разом з цифровими комп'ютерами. У 1950 р. англійський математик Алан Тьюринг оприлюднив статтю «Обчислювальні машини та інтелект», де зазначив, що «наш інтерес до «мислячих машин» виник завдяки машині особливого роду, яку зазвичай називають «електронним або цифровим комп'ютером» [4; 436]. У цьому ж матеріалі він описав базову структуру цифрового комп'ютера, взаємодію основних його складників, запропонувавши покласти в основу розрахунків двійкову систему обчислення. Тьюринг показав, що обчислювальні машини розв'язують розрахункові задачі будь-якої складності, а оскільки всі цифрові комп'ютери логічно ідентичні (не треба створювати нову машину під кожен нову задачу — їх усі розв'язуватиме лиш один комп'ютер, якщо наділити його відповідною програмою), він назвав їх універсальними машинами. Згодом виникла назва «універсальні машини Тьюринга». Крім того, він запитав: «Чи можуть такі машини мислити?» [4; 434]. Для відповіді вчений запропонував біхевіористичне тестування комп'ютерних систем на інтелектуальність, яке також назвали на його честь.

Стандартна інтерпретація тесту Тьюринга така: суддя взаємодіє з одним комп'ютером

і однією людиною. На підставі відповідей він визначає, з ким розмовляє: з людиною чи комп'ютерною програмою. Завдання програми — ввести суддю в оману. Учасники не бачать один одного. Якщо суддя не може визначитися — машина пройшла тест. Щоб протестувати саме інтелект машини, а не вміння розпізнавати усне мовлення, спілкування ведеться за допомогою клавіатури комп'ютера-посередника. Відповіді даються через певні проміжки часу, щоб суддя не робив висновку, виходячи з їхньої швидкості. За часів Тьюринга комп'ютери надавали відповіді повільніше за людину, а тепер навпаки.

Тест Тьюринга сформульовано у рамках панівної синтаксичної концепції мислення, яка походить ще від Платона і Гоббса. Він унаочнює Тьюрингове розуміння сутності інтелекту і мислення, передбачає, що стандартний цифровий комп'ютер, яким управляють рекурсивно застосовні правила, чутливі до порядку розташування елементарних символів (які відіграють роль вхідних даних), здатен визначити будь-яку систематичну сукупність відповідей на довільні запитання ззовні. Вказівка Тьюринга на взаємодію з машиною через клавіатуру комп'ютера-посередника є несуттєвим обмеженням. Висновок правомірний, навіть якщо комп'ютер взаємодіятиме зі світом складнішими способами: через розпізнавання зображень, мовлення тощо. Зрештою, ця взаємодія являє собою таку саму рекурсивну функцію Тьюринга, тільки складнішої конструкції. Залишається визначити складну функцію, що керує реакціями людини на зовнішні впливи, а потім написати комп'ютерну програму (множину рекурсивно застосовних правил) для її обчислення. Завдання, що їх окреслив Тьюринг, покладено в основу наукової програми класичного штучного інтелекту.

Перші результати були вражаючі для свого часу. Комп'ютери виконали дії, які

можна було сприйняти за прояв розуму. Вони реагували на команди, розв'язували складні арифметичні задачі, грали в шашки й шахи, підтримували простий діалог, навіть доводили теореми. Результати поліпшувалися з розвитком комп'ютерної техніки й ускладненням програм.

Схожі дослідження проводили в СРСР. Зокрема, Інститут кібернетики АН УРСР під керівництвом академіка В.М. Глушкова автоматизував доведення математичних теорем, створив систему перекладу з російської на українську, алгоритм аналізу змісту фраз, нову формальну систему — алгебру алгоритмів, завдяки чому було формалізовано практичні завдання розробки комп'ютерних систем, побудовано математичну теорію їх проектування тощо [5].

### ШТУЧНИЙ ІНТЕЛЕКТ І ЗНАННЯ, НАБУТІ ЦИВІЛІЗАЦІЄЮ

Проривом у створенні інтелектуальних систем за класичною синтаксичною концепцією був універсальний розв'язувач задач (General Problem Solver) Аллена Ньюелла, Кліффа Шоу, Герберта Саймона [6], базований на евристичному пошуку. Ідея програми виникла під час спостереження за студентами під час розв'язання логічних задач. А. Ньюелл і Г. Саймон зауважили, що ті часто застосовували «прямолінійні» прийоми, наприклад, замінювали довгий вираз коротшим.

Ньюелл, Шоу, Саймон відтворювали розумову діяльність у контексті сталих традицій, орієнтуючись на маніпулювання набором символів, перетворюваних відповідно до чітко встановлених правил. Як і фрази звичайної мови, вирази класичних систем штучного інтелекту, до яких належить універсальний розв'язувач задач, складні, утворені шляхом систематизації простих символів. Через обмежену кількість символів нові смислові вирази komponуються з наявних і балансують між синтаксичною структурою

і семантикою. Символи групуються у відомі терміни, що робить символічне подання відносно простим і зрозумілим. Тобто відкритість і зрозумілість символічних систем штучного інтелекту робить їх придатними для людино-машинного спілкування.

У 1976 р. Аллен Ньюелл і Герберт Саймон сформулювали гіпотезу про фізичну символічну систему (гіпотеза Ньюелла–Саймона): необхідні і достатні засоби для реалізації базових інтелектуальних дій у широкому сенсі має фізична символічна система [7–9]. Тобто без символічних обчислень неможливо виконати осмислені дії. Здатності проводити символічні розрахунки цілком достатньо, щоб реалізувати інтелектуальні дії. А оскільки комп'ютер має таку здатність, на його основі можна створити штучний інтелект.

Перші результати універсального розв'язувача задач вселяли певність, що нарешті мрія філософів знайшла для свого втілення технічні засоби і програмну реалізацію набору правил, спроможних перевести міркування в розрахунок. А. Ньюелл і Г. Саймон навіть поспішили проголосити, що «інтуїцію, інсайт і навчання не можна більше вважати винятково прерогативою людини: ними володіє будь-який потужний комп'ютер, запрограмований відповідним чином» [10; 6]. Г. Саймон у 1965 р. прогнозував, що протягом найближчих 20 років машини виконуватимуть усі дії, на які здатна людина [11; 96].

Едвард Фейгенбаум писав: «В евристичних програмах розв'язання проблем передбачено, що пошук рішення у просторі задачі спрямовують і контролюють евристичні правила. Представлення, що окреслює простір задачі, визначається ставленням дослідника до цієї проблеми та його точкою зору, і воно ж зумовлює вид рішення. Обравши для задачі вдалий спосіб представлення, можна істотно підвищити ефективність процесів пошуку рішення. Вибір

способу представлення задачі — справа дослідника, що розробляє програму. Це — творчий акт» [12]. Його називають знаходженням глибинної структури задачі або інсайтом. Без втручання людини на етапі «творчих актів» — постановки задачі, написання алгоритму, встановлення цілей — подібні системи обійтись не можуть.

У такі програми треба закласти множину альтернатив. На основі фіксованих правил обробки визначеного переліку фактів програма вибирає найадекватнішу. Системи штучного інтелекту, ґрунтовані на класичній синтаксичній концепції, здійснюють розрахунки з огляду на встановлені цілі й обмеження, проте не реалізують інтелект у широкому сенсі (творче мислення), а завдання виконують у межах обраної альтернативи.

Один з найавторитетніших дослідників штучного інтелекту Марвін Мінський у 1968 р. писав: «Я не бачу підґрунтя для припущення, що інтелект може існувати поза високоорганізованим масивом знань, без зв'язку з моделями і процесами. Для нашої цивілізації завжди було природним вважати, що інтелект кристалізований у деякому ізольованому утворенні — називайте його свідомістю, розумом, інсайтом, гештальтом або як вам завгодно; але при цьому вирішення проблеми підмінюється її назвою. Здатність розв'язувати задачі, властива людині з високим рівнем інтелекту, пояснюється частково більшою досконалістю її евристик, використовуваних для організації структури її знань, а частково самою структурою; вірогідно, ці два моменти певною мірою неподільні. Немає жодних підстав вважати, що можливі якінебудь інші шляхи до того, що називається інтелектом, крім використання адекватних, цілком конкретних знань чи модельних структур» [13; 27].

За Мінським, зовнішній світ — це маса ізольованих фактів, які розум інтерпретує

в термінах інших ізольованих фактів, збережених у пам'яті й організованих за допомогою своєрідного каталогу чи бази правил. Тож розум сприймає окремі деталі навколишнього світу як миттєві знімки ситуацій, а потім збирає їх у термінах моделі, створеної на основі раніше накопичених фактів. Мінський зазначав, що будь-яка впорядкованість і регулярність у людській поведінці спостерігається за умов підпорядкування детермінованим правилам. В іншому випадку вона абсолютно довільна і непередбачувана. Він навіть не припускав існування регулярної поведінки, яку не визначають правила.

#### КРИТИКА ШТУЧНОГО РОЗУМУ

У 1972 р. позиції Мінського та ін. поборників класичної синтаксичної концепції штучного інтелекту жорстко розкритикував Х'юберт Дрейфус у своїй книзі «Чого не можуть обчислювальні машини: критика штучного розуму» [14, 15]. На його думку, регулярна поведінка, яку не визначають правила, не тільки можлива, але й немінуча, бо конструювання правил для всіх потенційних ситуацій неможливе: «Наша спрямована контекстом діяльність, у термінах якої ми постійно заново оцінюємо релевантність і значення конкретних об'єктів і фактів, цілком регулярна, проте ця регулярність не повинна і не може повністю підпорядковуватися правилам» [15; 240].

Х. Дрейфус зауважує, якщо «розвинути гіпотезу, за якою будь-яку поведінку необхідно обов'язково розуміти як таку, що виникає в результаті застосування деякого набору правил, то ми зіштовхуємося з нескінченним процесом звернень до правил, правил застосування правил і т.д. — з регресом у нескінченність. Цей процес не може завершитися зверненням до звичайних фактів, оскільки, як впливає з вихідного твердження, самі факти завжди впізнаються й інтерпретуються через деякі правила» [15; 258].

Він різко критикує системи штучного інтелекту з формальним описом ситуації у вигляді правил, аргументовано показує їхню принципову неспроможність до розумової діяльності. Доводи Х. Дрейфуса цікаві: «Значення, яке має бути приписано кожному логічному елементу, залежить від інших логічних елементів, і для того, щоб можна було сказати, що ті чи інші вхідні сигнали формують спочатку образи, а врешті-решт — об'єкти й усвідомлені висловлювання, кожен з цих сигналів має бути співвіднесений за деякими правилами з іншими сигналами. Проте відповідно до різних правил елементи можуть мати різну інтерпретацію, вибір же правила визначається контекстом. Однак обчислювальна машина може розпізнавати контексти лише за допомогою певних правил.

Цей спосіб аналізу, який нав'язує обчислювальна машина, суперечить нашому досвіду. Феноменологічний опис нашого досвіду наводить на думку, що ми завжди перебуваємо в деякій ситуації або контексті, що його переносимо з найближчого минулого і підлаштовуємо під сучасність у термінах тих подій, які у світлі попередньої ситуації здаються нам значимими. Кожен раз ми зіштовхуємося не з беззмистовними елементами, у термінах яких маємо ідентифікувати контексти, а тільки з фактами, що вже мають інтерпретацію і своєю чергою визначають ситуацію, у якій ми перебуваємо. Життєвий досвід людини можна зрозуміти лише в тому випадку, коли він організований у термінах ситуацій, у яких уже задано релевантності і значення» [15; 261].

Факти реального світу, на думку Дрейфуса, не мають придатного для всіх контекстів однозначного тлумачення, яке б могли задати розробники систем штучного інтелекту, — інтерпретація залежить від обставин. Можливість встановлення релевантності і значень елементів учень бачив лише в ієрархії контекстів, хоч і критикував

Мінського за спроби створити машину, яка не мислить, а моделює мислення<sup>1</sup>. Розуміючи обмеженість такого підходу, його відмінність від пізнання і прийняття рішень людиною, Дрейфус робить висновок про неможливість реалізації штучного інтелекту шляхом маніпулювання символами за допомогою рекурсивно застосованих правил. Однак він не відкидав можливості створення штучного інтелекту взагалі.

Загалом роботу Дрейфуса зустріли вороже й оцінили як недалекоглядну і необ'єктивну. А недосконалість систем штучного інтелекту пояснювали недостатньою потужністю комп'ютерів. Однак наприкінці 70-х рр. XX ст. швидкість поширення сигналів у них стала приблизно в мільйон разів вищою, ніж у мозку людини, а тактова частота центрального процесорного пристрою майже на стільки ж переросла частоту будь-яких коливань, виявлених у мозку. При цьому комп'ютерні системи практично не наблизилась до біологічних у вирішенні таких когнітивних задач, як розпізнавання образів, переклад текстів тощо.

Погляди Дрейфуса поділяв професор Каліфорнійського університету в Берклі Джон Сірл, який займався філософськими аспектами створення штучного інтелекту. Він ставив під сумнів, що коректне маніпулювання структурованими символами шляхом рекурсивного застосування правил,

<sup>1</sup> Тези М. Мінського «На запитання про ті чи інші об'єкти досліджуваного світу можна відповісти деяким твердженням щодо поведінки відповідних структур у тій моделі світу, яка є в того, до кого поставлено запитання» [15; 426], «Коли людині ставлять запитання загального характеру щодо неї самої, вона намагається дати загальний опис наявної в неї моделі своєї природи» [15; 428] Дрейфус нищівно розкритикував: «М. Мінський ... не усвідомлює, що «регрес у нескінченність» виникає через те, що він наголошує на введенні моделей, і що це ускладнення унаочнює філософську непослідовність його позиції, згідно з якою ми ніколи нічого не знаємо напямую, а лише через призму моделей» [17; 236].

котрі враховують їхню структуру, може становити сутність свідомого розуму. В 1980 р. Сірл продемонстрував можливість проходження тесту Тьюринга побудованою за принципами фізичних символічних систем штучною системою, яка не володіє інтелектуальними властивостями і не розуміє суті запитань. У статті [16] наведено уявний експеримент з «китайською кімнатою» — уособленням штучної системи, що проходить тест на інтелектуальність. У кімнаті людина коректно відповідає на запитання незнайомою їй китайською мовою, спираючись на підказки, що містяться тут же (алгоритм поведінки). При цьому людина зовсім не розуміє, що саме говорить. За Сірлем, така поведінка, як і кімната в цілому, не інтелектуальна<sup>2</sup>.

Цей експеримент ілюструє здатність машини імітувати людську поведінку в розмові, механічно (неінтелектуально) дотримуючись заздалегідь установлених правил. Такі алгоритми не усвідомлюють сутності речей, не узагальнюють досвід, аби застосувати його в майбутньому. Відповідно, Сірл доходить висновку, що мислення не еквівалентне формальним символічним розрахункам. Крім того, друга і третя аксіоми Сірла «людський розум оперує смисловим змістом (семантикою)» і «синтаксис сам по собі не утворює семантику, його недостатньо для існування семантики» [17; 27] свідчать: маніпулювання символами

<sup>2</sup> Тест Тьюринга має також безліч інших зауважень. По-перше, він перевіряє тільки аналіз синтаксичних конструкцій, а не самі високоінтелектуальні дії, наприклад, виконання складних завдань чи висування оригінальних ідей. По-друге, не визначає інтелектуальними такі когнітивні завдання, як розпізнавання образів тощо. Тест провокує обман: навіть найдосконаліша програма повинна прикидатись не надто розумною, щоб пройти перевірку (для цього деякі розробники навчали комп'ютерні системи робити граматичні помилки). Якщо ж машина здатна розв'язати якусь розрахункову задачу, непосильну для людини, вона за визначенням провалить тест.

недостатньо для наявності сенсу. Тож моделювання мислення ще не є самим мисленням. І програмна реалізація комп'ютерних алгоритмів, навіть здатних здійснити синтаксичний, семантичний, морфологічний, графематичний, ін. аналізи лінгвістичних конструкцій і дати обґрунтовані відповіді з бази варіантів, не наділяє систему інтелектом, тобто справжнім розумінням суті запитань і відповідей.

Сірл вказує, що системи його часів здатні відтворити синтаксис, але не реалізують семантику, а це і є головна ознака мислення (інтелекту). Однак він не відкидає можливості створення таких систем у майбутньому.

Розуміння штучного інтелекту за Тьюрингом Дж. Сірл назвав слабким штучним інтелектом. Натомість «сильний» — це той, який зреалізує програма, що буде не просто «моделлю розуму, вона сама буде розумом у буквальному сенсі слова, у тому самому сенсі, у якому людський розум — це розум» [17; 26].

Публікація Сірла спровокувала поміж психологами, філософами, дослідниками штучного інтелекту широку дискусію щодо коректності висновків. Лунали навіть думки, що хоч кімната з її вмістом працює надто повільно, ця система все-таки розуміє китайську мову. Серед критиків варто згадати Пола і Патрицію Черчлендів, які, крім зауважень, запропонували створити систему штучного інтелекту на базі нейронних мереж [18]. У полеміці з ними Дж. Сірл, критикуючи нейронні мережі (як і інші конекціоністські системи) як інструментарій для реалізації штучного інтелекту, наводить новий приклад з китайською мовою, коли в гімнастичному залі багато осіб дають відповіді, базуючись на тих самих правилах, що й людина у «китайській кімнаті». І показує відсутність у такої системи інтелектуальних властивостей. Дж. Сірл зазначає, що для проведення розрахунків

послідовні і паралельні архітектури абсолютно ідентичні: будь-яке обчислення, виконане машиною з паралельним режимом роботи, зrealізує машина з послідовною архітектурою [17; 28].

На це також вказував дослідник штучного інтелекту Дж. Фодор: «будь-яка операція, яку виконує нервова система, ідентична деякій послідовності елементарних операцій» [19; 629]. А якщо «істинна точка зору, згідно з якою причиною сприйняття глибини виступає текстурний градієнт, якщо центральна нервова система відповідає уявленню, яке склалося про неї в найпроникливіших дослідників, то деякі з дій, які виконує нервова система, і деякі фізичні процеси, що відбуваються в ній, коли ми усвідомлюємо глибину, можна описати такими поняттями, як «обчислення текстурних градієнтів», «обробка інформації про текстурні градієнти», «обчислення похідних текстурних градієнтів» тощо» [19; 632].

Х. Дрейфус критикував такі міркування: «Стверджувати, що мозок при розрахунку текстурного градієнта неодмінно здійснює низку операцій, так само абсурдно, як стверджувати, що, обертаючись своїми орбітами навколо Сонця, планети розв'язують диференціальні рівняння або логарифмічна лінійка (аналогова машина), визначаючи корінь квадратний, здійснює ту ж послідовність кроків, що і цифрова машина, яка шукає відповідне значення, записане у двійковій системі обчислення» [15; 118].

#### **БІОЛОГІЧНИЙ КОМПОНЕНТ ПОВЕДІНКИ МАШИН**

Дрейфус передбачав можливість реалізації штучного інтелекту на базі аналогових систем, зазначаючи, що дискретні не здатні до розумової діяльності, схожої на людську [15; 259, 268]. Відомий авторитет у галузі моделювання мислення і психіки український учений Микола Михайлович Амосов, який очолював відділ біокіберне-

тики Інституту кібернетики АН УРСР, вказував на дискретність функціонування будь-якого інтелекту, наголошуючи на його суто мережевій структурі [20; 78]. Зокрема, у книзі «Алгоритми розуму» зазначено: «Будь-який інтелект функціонує дискретно. Якщо говорити точніше, то це сполучення неперервних і дискретних процесів. Утім, чи існують узагалі суто неперервні процеси? У всякому разі, у складних системах будь-яке неперервне є лише статистикою великої кількості окремих подій. У мозку, наприклад, уся діяльність нейронів виражається окремими імпульсами» [20; 85].

Під керівництвом Амосова з 1963 р. проводились роботи зі створення моделей інтелекту за принципом семантичних мереж із системами підсилення-гальмування. Усвідомлюючи технічну обмеженість тогочасних комп'ютерів і неспроможність проведення масштабних паралельних розрахунків [20; 29], Амосов паралельно працює над розробкою систем алгоритмічного інтелекту (за принципами фізичних символічних систем із послідовною архітектурою), що ґрунтуються на біологічних механізмах функціонування мозку, проектує велику кількість евристичних моделей складних систем тощо.

На дотримання біологічної правдоподібності в системах штучного інтелекту вказувало чимало дослідників. Зокрема, Дж. Фодор, намагаючись застосовувати для створення штучного інтелекту експертні системи з послідовним опрацюванням символічної інформації, наполягав, щоб «операції, на яких базується поведінка машини, були того ж типу, що й операції, які лежать в основі поведінки організму» [21; 141].

Прогрес у когнітології — науці про знання, що вивчає методи і прийоми його одержання, обробки, зберігання, використання, — дає змогу пояснити й описати когнітивні процеси, що відбуваються в мозку і відповідають за вищу нервову активність. Їх

осмислення може стати підґрунтям для побудови систем сильного штучного інтелекту, здатних до самостійного навчання, творчості, вільного спілкування з людиною. У ході вивчення структури мозку і процесів, які там відбуваються, сформульовано два загальних підходи до побудови систем штучного інтелекту:

1) семіотичний — створення символічних моделей штучного інтелекту за принципом «згори донизу» (top-down) шляхом побудови експертних систем, баз знань, систем логічного висновку, які імітують високорівневі психічні процеси: мислення, міркування, мову, емоції, творчість тощо;

2) біологічний — створення паралельних розподілених процесорів із природною здатністю до навчання і роботою за принципом «знизу догори» (bottom-up), в основі якого лежить вивчення нейронних мереж і еволюційних обчислень, що моделюють інтелектуальну поведінку на основі дрібніших «неінтелектуальних» елементів<sup>3</sup>.

Відповідно, один із магістральних напрямів досліджень зі створення систем штучного інтелекту — це конструювання мозкоподібних архітектур. Академік Олексій Григорович Івахненко<sup>4</sup> навіть тлумачить штучний інтелект як властивість цифрової обчислювальної машини чи мережі нейроподібних елементів реагувати на інформацію, що надходить на її вхідні пристрої, майже так само, як реагує в тих же інформаційних умовах людина [24; 5].

Процеси видобування, обробки, зберігання, зчитування інформації реалізовані у штучних нейронних мережах подібно до біологічної нервової системи. Концепція

<sup>3</sup> Принципи конструювання інтелектуальних систем згідно з обома підходами, їх застосування для розв'язання різноманітних задач розкрито, зокрема, в авторських працях [22, 23].

<sup>4</sup> Винахідник методу групового урахування аргументів, розробник розпізнавальної самонавчальної системи на базі дворядного персептрона «Альфа» та багато ін.

обробки інформації в нейронних мережах походить із принципу паралелізму, що робить їх гнучкими. Масовий паралелізм (сотні тисяч нейронів і більше) надає нейронним мережам особливу форму робастності (стійкості до грубих зовнішніх впливів із суттєвою невизначеністю). Якщо обчислювальні процеси розподілені між великою кількістю нейронів, тоді майже неважливо, що стан окремих нейронів відрізняється від очікуваного. Викривлений чи неповний вхідний сигнал так чи інакше можна розпізнати; ушкоджена мережа здатна працювати задовільно, а навчання не обов'язково має бути максимально точним. При цьому продуктивність мережі в межах деякого діапазону знижується досить повільно. Крім того, можна додатково підвищити робастність мережі, представляючи кожен властивість окремою групою нейронів [25].

Суттєвих результатів у моделюванні із застосуванням нейронних мереж домоглася школа Воррена Мак-Каллоха [26]. Однак мережі з формальних нейронів виявились нездатними відтворювати складні функції мозку. Неefективними виявились також спроби використовувати такі мережі для керування роботами [27]. Великий інтерес викликали праці Френка Розенблатта, який сформулював низку принципів нейродинаміки мозку і застосовував їх для побудови персептронів [28]. Однак строгий аналіз, який здійснили Марвін Мінський і Сеймур Пейперт [29], продемонстрував принципову обмеженість нейронних мереж, побудованих на порогових елементах, для розв'язання низки простих задач<sup>5</sup>. Таким чином, до поч. 70-х рр. XX ст. загальна криза нейронного підходу стала очевидною.

<sup>5</sup> Ці функціональні обмеження персептронів поборили в 1986 р. Руммельхарт, Хінтон, Вільямс [31], які запропонували використання диференційованих функцій активації та розробили алгоритм оптимізації ваг міжнейронних зв'язків у мережах персептронного типу — Error Back-Propagation Algorithm.



## ВАРІАЦІЇ НА ТЕМУ НЕЙРОМЕРЕЖ

Термін «штучний інтелект» виник у 60-х рр. завдяки експертним системам як науковий напрям, альтернативний нейронним мережам. Поміж його засновників був Марвін Мінський [30], який, хоч і зробив один з найбільших внесків у теорію нейронних мереж<sup>6</sup>, не зараховував цей інструментарій до систем штучного інтелекту.

Мінський вважав недоречною апеляцію до архітектури мозку, його нейронних структур, декларував необхідність моделювання роботи людини зі знаннями. Поставивши в центрі уваги операції з формально-логічними мовними конструкціями, він свідомо орієнтувався на імітацію обробки інформації лівою півкулею мозку. Системи, що її виконували, названо експертними, оскільки вони відтворюють логіку міркувань експерта в певній галузі. Ці міркування реалізуються з використанням правил висновку, отриманих в експерта і формалізованих на підґрунті обраного математичного інструментарію.

Саме структура безпосередньо пов'язаного з архітектурою представлення знань стала ключовою вадою нейронних мереж — вони не передбачають однозначного встановлення правил прийняття рішень. На сьогодні не існує формалізованої теорії оптимізації структури цих мереж або оцінки впливу архітектури на подання знань. Незалежно від вибору останньої, мережа здобуває знання про предметну галузь у процесі навчання. Ці знання подаються у вигляді ваг синаптичних зв'язків мережі і параметрів нейронів, завдяки чому мережа

<sup>6</sup> У 1951 р. разом з Деном Едмондсом він сконструював першу нейромережеву машину SNARC, фактично ставши в наступні 50 років одним з найбільш впливових і авторитетних фахівців у галузі штучного інтелекту. Перша конференція з проблем штучного інтелекту відбулася в США в 1969 р. — саме тоді вийшла книга Мінського і Пейпєрта «Персептроні» з критикою нейромережевих структур [31].

легко адаптується й узагальнює, однак не описує повноцінно розрахункового процесу, застосованого в ухваленні рішення. Інколи це серйозно обмежує використання нейромережевого підходу, особливо в задачах, де необхідно пояснити результат.

У 1988 р. Фодор і Пилишин оприлюднили критичні зауваження про обчислювальну адекватність нейронних мереж у розв'язанні когнітивних і лінгвістичних задач [32]. Мережі не задовольняють два основні критерії пізнання — ментальне уявлення і розумові процеси. За Фодором і Пилишином, їх мають системи штучного інтелекту, проте не нейронні мережі: ментальне уявлення характеризується комбінаторним синтаксисом і семантичною структурою; розумові процеси чутливі до комбінаторної структури уявлення. Науковці наголошують: штучний інтелект передбачає наявність ментального уявлення, де пізнання йде як послідовне опрацювання символічної інформації. Вони стають на позиції класичного синтаксичного підходу до створення систем штучного інтелекту — використання символічної мови для подання загальних знань про предметну галузь і конкретних знань про способи реалізації завдання [32, 33]. Цей підхід можна втілити на сучасних комп'ютерах нейманівської архітектури, що послідовно обробляють інформацію.

Однак деякі анатомічні особливості мозку принципово відрізняють його від архітектури традиційних комп'ютерів. Нервова система — це паралельна машина, у тому сенсі, що сигнали обробляються мільйонами різних елементарних процесорів водночас. Наприклад, сітківка ока передає в мозок складний вхідний сигнал не послідовно, як в електронних обчислювальних машинах, а через сукупність майже мільйона окремих елементів, що надходять одночасно до закінчення зорового нерва, після чого їх також одночасно опрацьовує мозок.

Думка теж не плине послідовно, переходячи від одного нейрона до іншого, а водночас активує велику кількість нейронів, що відповідають деякому образowi в пам'яті. Ці нейрони активують багато інших, які містять образ, асоціативно пов'язаний з попереднім. Поява конкретного образу залежить від того, які нейрони активовано. Послідовна обробка інформації не забезпечує переходу від однієї групи багатьох обчислювальних елементів до іншої в логічній послідовності. Тож Сірлеві аргументи з гімнастичним залом щодо ідентичності паралельних і послідовних розрахунків виглядають надуманими.

### СЛАБКИЙ І СИЛЬНИЙ ШТУЧНИЙ ІНТЕЛЕКТ

Для слабкого штучного інтелекту, який здатний виконувати когнітивні завдання, достатньо послідовної обробки символічної інформації. Що ж до сильного, здатного до мислення, самостійного навчання, творчості, варто уважніше поглянути на сутність інтелекту. У пригоді нам стануть філософія й основи біології, власне, розуміння поведінки простіших мислячих істот.

Вищі тварини постійно приймають розумні рішення. Наприклад, якщо собака побачить більшого за себе, то швидше за все відбіжить, якщо ж більший пес буде за парканом, гавкатиме на нього, підбігши до огорожі. Бачимо елементарну логіку за схемою «якщо..., то..., інакше...». Це не рефлекс, подібний до реакції на вогонь чи біль або звичайного пошуку їжі, а мислення, прийняття рішень.

У зв'язку з цим постають питання про характер думок, організацію процесу мислення. У людей плин думок має вигляд логічних мовних конструкцій: за однією фразою зринає інша, виникають варіанти розвитку подій тощо (саме тут з'являється місце для символічних систем штучного інтелекту). Однак тварини не володіють мо-

вою, не формують речень і синтаксичних конструкцій. У них прийняття рішень має якусь іншу форму. Вищі істоти постійно приймають рішення, спираючись на закладені в мозку моделі поведінки. І людина мислить так, коли думки не набувають вигляду логічних мовних конструкцій, — рішення виникає миттєво, без логічного обґрунтування.

Розмірковуючи, людина оперує не лише словами, а й образами, асоціаціями. Часом вона точно розуміє, що хоче сказати, але не може підібрати влучне слово. Більше того, коли ми згадуємо подумки, то не говоримо: «Я був у великій світлій кімнаті, посередині стояв стіл, два м'яких крісла, під стіною — шафа з книгами. Там були чоловік і жінка. Вони виглядали впевнено, були такого зросту, мали такі фігури, обличчя, одяг. Вони говорили про можливість побудови систем штучного інтелекту». У такому описі багато інформації (можливо, і несуттєвої) втрачено. У пам'яті ситуація зафіксована насиченіше, проте не абсолютно чітко. Зазвичай людина згадує ситуацію миттєво, не в формі речення чи іншої синтаксичної конструкції.

М. Амосов із цього приводу писав: «Головне мислення — образне, а не мовне. Образні картини значно багатші від словесних, у них набагато більше розмаїття, тобто інформації. Якщо уявити кількість моделей-образів у корі, які одночасно «живуть», проявляють активність, то хіба можна порівняти їх з кількістю слів? Передача образів словами повільна, неповоротка й невиразна. Людина мислить переважно образами, а внутрішня мова, яку ми найчастіше ототожнюємо з думками, тільки коментує образне мислення» [20; 128].

Зазначимо, що мова для людини є лише проекцією її думок. Однак самі думки можуть розвиватись і без трансформації в лінгвістичну форму. Відповідно, можна припустити, що процес мислення у людини

проходить, скоріше за все, через деякі збережені у пам'яті образи, але людина усвідомлює це через мову, носієм якої вона є. За М. Амосовим, «взаємодія моделей-образів становить основу граматики й синтаксису мови, які ми звикли зараховувати до якостей, властивих лише їй. Мова тільки позначає, маркує основні образи. Правда, тим самим дії з моделями стають більш раціональними й ефективними» [20; 125].

Відповідно до цього можна припустити, що і тварини можуть мислити подібними категоріями (образами, але у простішій формі), однак вони не трансформують їх через словесні вирази. Біологічним поясненням є елементарна нерозвиненість відповідної ділянки мозку, що надає образам лінгвістичного представлення.

Отже, для втілення штучного інтелекту за принципами живої природи доцільно будувати систему, що оперує саме образами. За необхідності можна додати модуль для лінгвістичної інтерпретації результатів. Це, звісно, суперечить загальноприйнятій гіпотезі Ньюелла–Саймона про реалізацію сильного штучного інтелекту лише на основі фізичних символічних систем, але наведена аргументація показує прийняття логічно обґрунтованих рішень у живій природі без володіння символічною мовою. Відповідно, для створення штучного інтелекту важливо вміти оперувати образами, які реалізують семантику, але їх представлення не обов'язково має бути синтаксичним або взагалі лінгвістичним.

Семантику, як правило, представляють вербально, однак для існування мислення це не обов'язкова умова. Людина здатна продукувати логічні послідовності без лінгвістичних конструкцій, ґрунтуючись лише на образах. Тобто для виникнення синтаксису семантика обов'язкова. Однак для існування семантики синтаксис зовсім не потрібний. І парадоксально те, що хоча синтаксис і вищий за семантику (властивий

лише найбільш інтелектуально розвинутій живій істоті — людині), його ми можемо відтворити, а семантику поки що ні. Адже вона закладена у глибинних пластах нашої свідомості як уміння оперування саме сутностями, а не їхніми лінгвістичними описами. Якщо ми це усвідомимо, а не відтворюватимемо очевидні синтаксичні конструкції, називаючи це інтелектуальними системами, то дійсно наблизимось до реалізації сильного штучного інтелекту.

Проектуючи системи штучного інтелекту за принципом експертного встановлення логічних правил у символічних категоріях, ми свідомо штучно вбудовуємо в систему знання експерта. Ці системи не продукують нових знань — вони здатні лише оптимізувати власні параметри для образних вхідних і вихідних змінних. Природа йшла іншим шляхом.

#### МОЗКОПОДІБНІ СИСТЕМИ

Логіка підказує доцільність конструювання мозкоподібних систем та інтерпретації розрахунків у нейроподібних елементах. На наш погляд, найадекватнішим інструментарієм для відтворення розумових процесів різних живих істот виступають класичні штучні нейронні мережі, зокрема, асоціативної пам'яті. Як і в біологічних нервових системах, там для збереження образів задіяно великі групи нейронів.

Системи паралельної обробки інформації здатні вирішувати надскладні задачі миттєво, незалежно від кількості нейронів, задіяних у розрахунках. Ця перевага стає очевиднішою зі зростанням кількості нейронів на кожному рівні. На швидкість обробки інформації тут не впливає ані число елементів, що беруть участь в обчисленні, ані складність функцій активації. Паралельна система нечутлива до дрібних помилок, функціонально стійка: утрата кількох (і навіть великої кількості) зв'язків не змінює суттєво загальний хід розрахунків.

Паралельна система запам'ятовує чимало інформації в розподіленому вигляді, при цьому доступ до будь-якого фрагмента цієї інформації просто блискавичний. Інформація зберігається у вигляді певних конфігурацій ваг окремих синаптичних зв'язків, сформованих протягом попереднього навчання. Навчання можливе також у фазі звичайної роботи мережі, з огляду на правило Хебба<sup>7</sup>, що ототожнює таку поведінку з біологічним прототипом. Потрібна інформація звільняється в міру того, як вхідний вектор даних проходить і трансформується через цю конфігурацію зв'язків на зразок асоціативної пам'яті живих істот, коли за одним образом спливає інший.

Цими образами можуть бути, зокрема, слова, що утворюють речення<sup>8</sup>. Слова при тому мають смисловий зміст та асоціативні зв'язки зі збереженими в пам'яті візуальними, звуковими, іншими образами, таким чином вони набувають семантики — головної характеристики інтелекту.

Зв'язки в асоціативній мережі можуть бути встановлені як між самими образами, так і між образами і відношеннями між ними<sup>9</sup>. При цьому від останніх через систему зв'язків нейронної мережі можна знов перейти до одного зі збережених у пам'яті образів. Таким чином, можна реалізувати синтаксис у вигляді лінгвістичних структур, що володіють смисловим навантаженням. Але взаємозв'язки тут будуть не строго прописуватись, як в експертних системах, а самостійно виділятися у процесі навчання, утворюючи семантично значимі синтаксичні конструкції.

Наголосимо, що паралельна система не маніпулює символами відповідно до струк-

турних правил. Маніпулювання символами — одна з багатьох «інтелектуальних» навичок, яким мережа може навчитися або не навчитися.

Зважаючи на викладене, можна припустити, що побудована за описаним принципом нейронна мережа виконуватиме інтелектуальні дії і без символічної мови і здатності до створення синтаксичних конструкцій. Але для того, щоб у нас з'явилась можливість розуміння процесу проходження «думок» у такій мережі, можна додати шар нейронів, що інтерпретуватиме образи лінгвістично, подібно до вихідної зірки Гроссберга чи розпізнавального шару в асоціативній нейронній мережі.

Упевнений, що реалізація описаної схеми не викличе значних ускладнень. Але як спонукати її функціонувати і виконувати ці дії? Навіщо їй це робити? Що буде стимулом для нейронної мережі займатися подібними розрахунками? Це питання, гадаю, набагато складніше за вибір конструкції мережі.

Ознайомлення з численною літературою з нейробіології дає нам деякі підстави припустити, що для живого організму такими стимулами можуть бути саме різні нейромедіатори (біологічно активні хімічні речовини, за допомогою яких здійснюється передача електричного імпульсу через синаптичний простір між нейронами). Причому низка нейромедіаторів, таких як дофамін, серотонін, ендорфіни тощо, виділяються як реакція на досягнення якоїсь мети і сприяють піднесеності, радості, одержанню насолоди живим організмом. Метою може бути задоволення як суто біологічних потреб, так і, наприклад, природної допитливості чи розв'язання людиною певного завдання. Є нейромедіатори, такі як епінефрин (адреналін), норепінефрин чи, знову ж таки, дофамін, які сприяють підвищенню активності, викликають занепокоєння тощо. Ацетилхолін, навпаки, сприяє заспокоєнню і бере

<sup>7</sup> Що частіше активізується міжнейронний зв'язок, то сильнішим він стає.

<sup>8</sup> Так реалізується синтаксис — специфічна мозкова діяльність людини.

<sup>9</sup> Відношення — це поєднувальна конструкція, що визначає взаємозалежність між образами.

участь у процесі запам'ятовування інформації, нейром'язовій передачі імпульсів.

Зауважимо, що емоційний прояв впливу нейромедіаторів, очевидно, відрізняється від їхньої участі в процесі обробки інформації. Коли відбувається передача сигналів для прийняття якихось рішень або вживання певних дій, то імпульси проходять у мозку, активізуючи окремі ланцюги нейронів. Тут видається, що можна було б обійтися тільки двома функціями нейромедіаторів — посилення і гальмування. Для прояву ж емоційних реакцій цього недостатньо. Крім того, у цьому процесі задіюються не окремі з'єднання нейронних зв'язків — нейромедіатори впливають на значні ділянки мозку.

Технологічно не можна вирішити завдання передачі емоційного стану шляхом формального призначення якихось нейронів відповідальними за центр задоволення потреб, а далі просто підвищувати чи знижувати активність цих нейронів залежно від ситуації. Штучна нейронна мережа самостійно не визначить, яка ділянка відповідальна за біль, страх, радість, задоволення. І ми теж не можемо формально призначити її такою, як це можливо в алгоритмічних задачах.

Відповідно, на наш погляд, це і буде основним гальмом на шляху створення систем штучного інтелекту — знайти аналог тій розмаїтості хімічних елементів (нейромедіаторів), які беруть участь у процесі обробки сигналів у мозку, а не просто імітувати дві функції — зниження і підвищення активності нейронів. Але, гадаю, і це завдання можна вирішити і втілити у штучній системі. Як варіант, різні нейромедіатори у системах штучного інтелекту можуть бути відтворені різними електричними сигналами, що характеризуватимуться власними амплітудами, частотами тощо, і кожен нейрон у штучній нейронній системі буде мати фільтри, здатні пропускати окремі види сигнала,

що імітуватимуть різні функції нейромедіаторів.

Перші системи штучного інтелекту (якщо вони все-таки з'являться) матимуть обмежений функціонал<sup>10</sup>. Однак стрибок до рівня людського розуму відбудеться надзвичайно швидко<sup>11</sup>, і щабель розвитку штучного інтелекту віддалятиметься від людського дедалі швидше. Математик і письменник Вернор Віндж на симпозіумі, проведеному Центром космічних досліджень NASA ім. Джорджа Льюїса й Огайським аерокосмічним інститутом (1993), наголосив: «Протягом найближчих 30 років у нас з'явиться технічна можливість створити надлюдський інтелект. Невдовзі після цього людська епоха буде завершена» [35]. Він переконаний: розумна машина не буде «інструментом» людей так само, як люди не стали знаряддям кроликів, берестянок чи шимпанзе [35]. З цим згодні багато науковців, що працюють над створенням штучного інтелекту.

Цікавим із цих позицій виглядає тест на інтелектуальність [36], згідно з яким найпереконливішим доказом мислення машин є їхня інтелектуальна мобільність, тобто спроможність розв'язати ту чи іншу реальну задачу з урахуванням трьох законів робототехніки А. Азімова:

- робот не завдає людині шкоди своїми діями і не допускає її через бездіяльність;
- робот підкоряється наказам людини, коли вони не суперечать першому закону;
- робот турбується про свою безпеку, доки це не суперечить першому і другому законам [37].

Стосовно такого тесту на інтелектуальність зауважимо, що насправді далеко не

<sup>10</sup> Багато вчених вказують на некоректність створення повністю сформованого інтелекту дорослої людини, як прагне класична синтаксична концепція.

<sup>11</sup> Еквівалентність потенціалів людського і штучного інтелекту зазвичай називають точкою технологічної сингулярності.

всі люди дотримуються цих законів, що не заважає їм лишатися інтелектуальними істотами. А жодна машина свідомо ще не заподіяла шкоди людям, хоча машини точно не володіють інтелектом. Отже, така перевірка штучно створеної системи на інтелектуальність не витримує ніякої критики.

Однак цей тест, як і самі закони робототехніки, порушує інші, глибші проблеми. Якщо системи сильного штучного інтелекту на засадах біологічної правдоподібності таки з'являться, як встановити в них поведінкові обмеження? Адже чіткі правила припустимі в експертних системах. Натомість системи паралельної обробки інформації містять силу-силенну характеристичних параметрів, які не піддаються змістовій інтерпретації. Тож невідомо, які параметри необхідно змінити, щоб нейронні мережі функціонували належним чином (параметри цих мереж, як і нейронів мозку, визначаються у процесі навчання на реальних даних). Оскільки це розподілені системи, штучні обмеження на визначені параметри мало що дадуть, бо сигнали завжди можуть пройти іншим шляхом.

У зв'язку з цим логічно виникає ідея «виховувати» штучний інтелект, ніби дитину. Цікаві думки щодо «морально-етичного виховання» штучного інтелекту викладено в [38]. Однак якщо людині моральні принципи прищеплюють страхом перед карами, починаючи від смерті і позбавлення волі і закінчуючи осоромленням і релігійними запереннями про неминучу розплату після смерті, то на штучний інтелект це не подіє.

Академік М. Амосов говорив про спеціальне «натренування» соціальних почуттів штучного інтелекту, підвищення активності моделей слів «обов'язок», «совість», «честь», «патріотизм» і пов'язаних з ними образів [20; 72, 139]. Хоча він усе-таки сумнівався в дотриманні штучним інтелектом законів робототехніки, бо той здатен до творчості, перевтілення, врешті, регулювання цих

«вроджених» переконань. М. Амосов скептично ставився до забезпечення дотримання обмежень в поведінці штучного інтелекту, наголошуючи, що це питання найголовніше у програмі його створення [20; 207].

Відповідно, основою для створення штучної інтелектуальної системи могла би слугувати здатна до самовдосконалення структура мозку дорослої психічно зрівноваженої розумної людини.

\* \* \*

Якщо все-таки буде сконструйовано архітектуру штучної системи, що могла би слугувати носієм інтелекту, то, напевно, найперспективніші для людства дослідження у найближчому часі можуть стосуватися сканування людського мозку і зчитування інформації з нейронів з записом на електронні носії (перепрошую за відхід у царину фантастики, але тоді кожна людина в буквальному сенсі дістане шанс на безсмертя).

Одним з можливих наслідків створення штучного інтелекту буде, як би це не звучало, завершення людської епохи, а з іншого боку — новий рівень існування, неймовірно швидкий технологічний розвиток, безсмертя і багато іншого, про що людство давно мріяло.

Чи варта гра свічок? Особисто я не дав би ствердної відповіді. Однак зупинитися вже неможливо. Багато країн вкладає величезні кошти в дослідження зі створення штучного інтелекту. І навіть якби їх заборонили, усвідомивши потенційну загрозу існуванню людства, технологічні, військові, економічні, медичні, інші бонуси так чи інакше спонукали б розвиток штучно-інтелектуальних технологій. Якщо ми і зійдемо з цієї дороги, хтось неминуче попрямує нею далі.

1. Платон. Сочинения. — М.: Мысль, 1971. — Т. 3, Ч. 1. — 610 с.
2. Гоббс Т. Левиафан, или Материя, форма и власть государства церковного и гражданского // Гоббс Т. Сочинения: В 2 т. — М.: Мысль, 1964. — Т. 2. — С. 45–678.

3. *Leibniz G.W.* Selections. — New York: Charles Scribner's Sons, 1951.
4. *Turing A.M.* Computing Machinery and Intelligence // *Mind*. — 1950. — V. 59, № 236. — P. 433–460.
5. *Сергієнко І.В.* Інформатика в Україні: становлення, розвиток, проблеми. — К.: Наукова думка, 1999. — 354 с.
6. *Newell A., Shaw J.C., Simon H.A.* Report on a General Problem-Solving Program // *Proceedings of the International Conference on Information Processing*. — 1959. — P. 256–264.
7. *Newell A., Simon H.A.* GPS: A Program that Simulates Human Thought // *Computers and Thought* (E.A. Feigenbaum, J. Feldman, eds.). — New York: McGraw-Hill, 1963. — P. 101–126.
8. *Newell A., Simon H.A.* Human Problem Solving. — Englewood Cliffs, NJ: Prentice-Hall, 1972. — 920 p.
9. *Newell A.* Physical Symbol Systems // *Cognitive Science*. — 1980. — № 4. — P. 135–183.
10. *Simon H.A., Newell A.* Heuristic Problem Solving: The Next Advance in Operations Research // *Operations Research*. — 1958. — V. 6. — P. 1–10.
11. *Simon H.A.* The Shape of Automation for Men and Management. — New York: Harper & Row, 1965.
12. *Feigenbaum E.A.* Artificial Intelligence: Themes in the Second Decade // *IFIP Congress*. — 1968. — V. 2. — P. 1008–1024.
13. *Minsky M.* (ed.). Semantic Information Processing. — Cambridge Mass.: MIT Press, 1968. — 440 p.
14. *Dreyfus H.L.* What Computers Can't Do: A Critique of Artificial Reason. — New York: Harper & Row, 1972. — 260 p.
15. *Дрейфус Х.* Чего не могут вычислительные машины: Критика искусственного разума / Пер. с англ. — Изд. 2-е. — М.: Книжный дом «ЛИБРОКОМ», 2010. — 336 с.
16. *Searle J.R.* Minds, Brains and Programs // *Behavioral and Brain Sciences*. — 1980. — V. 3, № 3. — P. 417–458.
17. *Searle J.R.* Is the Brain's Mind a Computer Program? // *Scientific American*. — 1990. — № 262(1). — P. 26–31.
18. *Churchland P.M., Churchland P.S.* Could a Machine Think? // *Scientific American*. — 1990. — № 262. — P. 32–37.
19. *Fodor J.A.* The Appeal of Tacit Knowledge in Psychological Explanation // *The Journal of Philosophy*. — 1968. — V. LXV, № 20. — P. 627–640.
20. *Амосов Н.М.* Алгоритмы разума. — К.: Наукова думка, 1979. — 223 с.
21. *Fodor J.A.* Psychological Explanation: An Introduction to the Philosophy of Psychology. — New York: Random House, 1968. — 165 p.
22. *Матвійчук А.В.* Моделивання фінансової стійкості підприємств із застосуванням теорій нечіткої логіки, нейронних мереж і дискримінантного аналізу // *Вісник НАН України*. — 2010. — № 9. — С. 24–46.
23. *Матвійчук А.В.* Штучний інтелект в економіці: нейронні мережі, нечітка логіка: Монографія. — К.: КНЕУ, 2011. — 439 с.
24. *Ивахненко А.Г.* Предисловие // *Амосов Н.М.* Алгоритмы разума. — К.: Наукова думка, 1979. — 223 с.
25. *Hinton G.E.* Shape Representation in Parallel Systems // *Proceedings of the 7<sup>th</sup> International Joint Conference on Artificial Intelligence*. — 1981. — P. 1088–1096.
26. *McCulloch W.S.* The Reliability of Biological Systems // *Self-Organizing Systems* (M.C. Yovits, S. Cameron, eds.). — London: Pergamon Press, 1960. — P. 264–281.
27. *Сутро Л., Киллмер У.* Совокупность решающих устройств для управления роботом // *Интегральные роботы* / Пер. с англ. под ред. Г.Е. Поздняка. Вып. 1. — М.: Мир, 1973. — С. 112–161.
28. *Rosenblatt F.* The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain // *Psychological Review*. — 1958. — № 65. — P. 386–407.
29. *Minsky M.L., Papert S.A.* Perceptrons. — Cambridge, MA: MIT Press, 1969. — 263 p.
30. *Minsky M.L.* Steps Towards Artificial Intelligence // *Proceedings of the Institute of Radio Engineers*. — 1961. — V. 49. — P. 8–30. Передруковано в: *Computers and Thought* / E.A. Feigenbaum, J. Feldman, eds. — New York: McGraw-Hill, 1963. — P. 406–450.
31. *Rummelhart D.E., Hinton G.E., Williams R.J.* Learning Internal Representation by Back-Propagation Errors // *Nature*. — 1986. — № 23. — P. 533–536.
32. *Fodor J.A., Pylyshyn Z.W.* Connectionism and Cognitive Architecture: a Critical Analysis // *Cognition*. — 1988. — V. 28. — P. 3–72.
33. *Pylyshyn Z.W.* Cognition and Computation: Issues in the Foundations of Cognitive Science // *Behavioral and Brain Sciences*. — 1980. — № 3:1. — P. 154–169.
34. *Матвійчук А.В.* Моделивання та аналіз економічних систем на підґрунті теорії нечіткої логіки: Дис. ... д. екон. наук: 08.00.11. — К., 2008. — 470 с.
35. *Vinge V.* The Coming Technological Singularity: How to Survive in the Post-Human Era // *Proceedings of a Symposium Vision-21: Interdisciplinary Science & Engineering in the Era of CyberSpace, held at NASA Lewis Research Center (NASA Conference Publication CP-10129)*. — 1993.
36. *Большаков В.І., Дубров Ю.І.* Інтелектуальна мобільність логічної техніки // *Вісник НАН України*. — 2010. — № 8. — С. 57–64.
37. *Азімов А.* Кінець Вічності: вибрані твори / Пер. з англ. — К.: Дніпро, 1990. — 767 с.
38. *Широчин В.П.* Слово об інтелекте. Концептуальные основы системной психологии. — К.: ТОО «ВЕК», 1999. — 304 с.

*А. Матвійчук*

МОЖЛИВОСТІ ТА ПЕРСПЕКТИВИ  
СТВОРЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ

Резюме

У статті досліджено історичні, філософські, технологічні аспекти створення систем штучного інтелекту, розкрито різні підходи до їх конструювання. Показано місце експертних систем і нейромережових технологій у цьому процесі. Досліджено адекватність тестів штучних систем на інтелектуальність, подано авторські міркування щодо можливості створення штучного інтелекту і ролі нейромедіаторів у процесі мислення. Запропоновано підхід до конструювання систем штучного інтелекту.

*Ключові слова:* тест на інтелектуальність, асоціативна нейронна мережа, нейромедіатори.

*A. Matviychuk*

POSSIBILITY AND PROSPECTS  
IN ARTIFICIAL INTELLIGENCE CREATING

Abstract

The paper covers historical, philosophical, technology aspects of artificial intelligence systems creating. Different approaches to such systems constructing are clarified. The place of expert systems and neural networks in this process is shown. The artificial system tests for intellectuality are examined as to their adequacy. The author gives his own ideas about the possibility of artificial intelligence creating and the role of neurotransmitters in the process of thinking. He also presents approach to artificial intelligence systems construction.

*Keywords:* test for intellectuality, associative neural network, neurotransmitters.