

REFERENCES

1. Adadurov V. (2017). *War of civilizations: social-cultural history of Napoleon Russian march*, vol. 1. Religion-language, Kyev: Laurus, 400 p. (in Russian).
2. Nora P. (2014). *Present, nation, memory*, Kyiv: Klio, 267 p. (in Ukrainian).
3. Ansell R. (2018). Reading and writing travels: Maximillian Mison, Samuel Warring and the afterlives of European voyages. *The English historical review*, vol. 133, issue 565, pp. 1446–1477.
4. Aytov S. Sh. (2018). The modern philosophy of history: historical-anthropological problematics. *Дослідження з історії і філософії науки і техніки*, вип. 26–27, с. 4–9 (in English).
5. Diah N. M. (2014). Overview of the anthropological theories. *Journal of humanities and social sciences*, vol. 4, № 10 (1), pp. 155–164.
6. Field K., Linch D. (2016). The world of Peter Biggs in Civil War era in Los Angeles. *The Western historical quarterly*, vol. 50, issue 4, pp. 43–64.
7. Grossen von M. (2017). *Amsterdam`s Atlantic: Print culture and the making of Dutch Brazil*. Philadelphia: University of Philadelphia, 265 p.
8. Hill M. A. (2018). The myth of Swards Folly. Civil War era. *The Western historical quarterly*, vol. 50, issue 1, pp. 43–64.
9. Hoffman P. (2015). *Why did Europe conquer the world?* Princeton: University of Princeton, 272 p.
10. Maza S. (2017). *Thinking about history*. Chicago: University of Chicago, 255 p.

Received 09.03.2019

Doi: 10.15421/271902

UDK 167.7:004.8

PHILOSOPHICAL ANALYSIS OF ARTIFICIAL INTELLIGENCE EVOLUTION

O. Dobrovolska, V. Shtanko

Kharkiv National University of Radioelectronics, Kharkiv, Ukraine

Tel.: +38-057-702-14-65

E-mail: olena.dobrovolska@nure.ua

ORCID ID 0000-0002-7057-2982

Tel.: +38-057-702-14-65

E-mail: valentyna.shtanko@nure.ua

ORCID ID 0000-0001-5245-3173

Abstract. During the last decade, artificial intelligence (AI) has penetrated into all spheres of human life: science, economy, education, politics, medicine, and others. Due to significant progress of this field, it has become the focus of attention of researchers from different disciplines. In humanities there are works that study the nature of cognition, moral responsibility and legal liability of artificial intelligence, however, philosophical analysis of the current state of the field remains understudied.

Artificial intelligence as a discipline began to develop in the middle of the twentieth century and has experienced several periods of crisis called «AI winter». The latter ones

were caused by technical limitations, on the one hand, and by philosophical problems arising in the field, on the other hand. It has been determined that the main philosophical questions of the field are: inability of artificial intelligence to self-reflection and flexible behavior; it is not able to feel, to create; limits of the computational approach; impossibility of complete formalization of knowledge.

The goal of the research is to find out what factors have led to the success within the field during the last decade and whether any philosophical problems from the list have been solved.

It has been found that evolution in the area is based on the technical achievements of recent years, which helped to implement algorithms that previously required many hours of computations. They are increase of the capacity of processors; distributed computing; development of machine learning, and artificial neural networks; cloud technologies. As a result the goal of the researchers within artificial intelligence has changed. The leading pragmatic approach in the field has caused fragmentation of research, which, in turn, has led to fragmentation of knowledge presentation. The intention to build the comprehensive description of the world has been replaced by need to create task-oriented over-simplified descriptions of small fragments of reality, while the goal to create artificial intelligence has been replaced by developing applications that simulate human activity.

It has been found that successful imitation of human cognitive activity and incapability to solve philosophical problems of the field have led to the evolution of «weak» intelligence, but not of the field itself. However, each area of the field is open to new practical and theoretical achievements, consequently the study of artificial intelligence will remain the focus of attention in the future.

Key words: artificial intelligence, philosophy of artificial intelligence, strong intelligence, weak intelligence, formalization, computational approach.

ФІЛОСОФСЬКИЙ АНАЛІЗ ЕВОЛЮЦІЇ ШТУЧНОГО ІНТЕЛЕКТУ

О. В. Добровольська, В. І. Штанько

Харківський національний університет радіоелектроніки, Харків, Україна

Анотація. Здійснено аналіз сучасного стану розвитку напряму штучного інтелекту з філософської точки зору. Грунтуючись на порівняльно-історичному аналізі, з'ясовано, що спричинило кризи в межах цієї сфери, визначено основні обмеження філософського характеру, які не дозволяють створити «сильну» форму штучного інтелекту: його нездатність до саморефлексії, відчуттів, творчості та гнучкої поведінки; обмеження обчислювального підходу; неможливість повної формалізації знання.

Проаналізовано, чим обумовлені успіхи в межах галузі в останнє десятиліття, і чи були подолані раніше виявлені обмеження. З'ясовано, що еволюція напряму ґрунтується на технічних досягненнях останніх років, що допомогло реалізації алгоритмів, які раніше вимагали багатогодинних обчислень. Визначено, що господарювання прагматичного підходу в галузі призвело до роздрібнення досліджень, фрагментації і спрощення уявлення про світ задля вирішення конкретних задач. З'ясовано, що успішне програмне імітування людської діяльності в різних сферах і неможливість вирішення філософських проблем призвели до розвитку тільки «слабкої» форми штучного інтелекту.

Проведене дослідження показало, що через “однобічний” технічний розвиток і відкритість філософських проблем не можна говорити про еволюцію на пряму в цілому.

Ключові слова: штучний інтелект, філософія штучного інтелекту, сильний інтелект, слабкий інтелект, формалізація, обчислювальний підхід.

1. Вступ

Протягом останнього десятиліття штучний інтелект проникнув до усіх сфер життя людини: науки, економіки, освіти, політики, медицини, побуту, дозвілля та інших. Програми з використанням технологій штучного інтелекту розпізнають зображення, людей, їхні голоси, перекладають тексти на різні мови, спілкуються з людьми та перемагають їх в іграх.

Згідно зі звітом Міжнародного економічного форуму [25], застосування штучного інтелекту є одним із чотирьох факторів, які сприятимуть розвитку бізнесу наступні чотири роки (до 2022 року). Тобто це буде галузь, в яку інвестуватимуть, і яка продовжить інтенсивно розвиватись. У зв'язку з цим напрям штучного інтелекту потрапив до поля зору науковців не тільки технічної та економічної галузей, але й історії, права, психології та філософії.

У межах гуманітарних дисциплін найбільше уваги зосереджено на дослідженні етично-правових питань: чи є системи штучного інтелекту відповідальними за свої дії; якими є моральні, соціальні та правові наслідки дій систем штучного інтелекту; яким є правовий статус роботів; як контролювати інтелектуальні системи, коли їхні навчальні можливості перевищать закладені в них налаштування [7]; відсутність критичного мислення та упереджене ставлення програм до певних верств населення [12]. Також продовжується вивчення питань: «Що таке інтелект, мислення та свідомість?», «За допомогою яких критеріїв можна визначити, що та чи інша система є інтелектуальною?». Крім того, проводиться історичний аналіз розвитку штучного інтелекту як дисципліни, на основі якого створюються прогнози щодо його майбутнього [4; 17]. При цьому питання, які стосуються філософського осмислення дисципліни в цілому в сучасному її стані, залишаються недостатньо висвітленими.

Метою роботи є визначити, що спричинило різкий стрибок у розвитку цього напрямку за останнє десятиліття, і чи є він початком якісно нового етапу його еволюції.

2. Історія розвитку штучного інтелекту

Штучний інтелект як напрям досліджень було виокремлено у 1956 році, коли професор Джон МакКарті визначив його як науку про створення розумних машин та розумних комп'ютерних програм [18]. Як і багато інших сучасних дисциплін, штучний інтелект є міждисциплінарним напрямом, який утворився на межі комп'ютерних наук, психології, фізіології, філософії та лінгвістики. Спричинило таке поєднання те, що поняття інтелекту, здатності до мислення, творчості, самоідентифікації вивчалось і продовжує вивчатись в кожній з цих дисциплін з різних боків. Виникла необхідність поєднати ці знання задля більш цілісного осягнення предмета дослідження.

Наразі штучний інтелект налічує кілька предметних областей, до яких відносять машинне навчання, обробку природної мови, мультиагентні системи,

моделі представлення знань, теорію ігор, робототехніку, комп'ютерне бачення та інші [17]. Причому кожна з цих областей залишається суміжною з іншими науками.

Успіхи в імітуванні деяких процедур людського інтелекту у середині XX століття привели до оптимістичного ставлення дослідників до можливості створення штучного інтелекту у майбутньому. Так, наприклад, американський соціолог та економіст Герберт Саймон, один з авторів програми «General Problem Solver», – універсальної машини для розв'язування задач, передрікав, що до кінця 60-х років комп'ютер стане чемпіоном світу з шахів і буде ефективно використовуватись у психології для вирішення задач, а один із засновників напряму штучного інтелекту Алан Тюрінг вважав, що у кінці XX століття ні в кого не викликатиме сумнівів той факт, що машини можуть мислити [26; 8, с. 29].

Цей оптимізм змінився періодом розчарування через неможливість апаратного втілення їхніх задумів. Виявилось, що створення штучного інтелекту потребує обробки даних великого об'єму (чого не могли забезпечити комп'ютери 50 років тому), що це дуже енергоємний, громіздкий і не вигідний з фінансової точки зору проєкт. Ці фактори призвели до занепаду дисципліни, дослідження в межах якої пошвидко вичерпались тільки на деякий час завдяки успіху експертних систем у 80-х роках. Інтерес до штучного інтелекту відновився у кінці 90-х років – на початку 2000-х через зміни пріоритетів у галузі. Масштабну ідею створення штучного інтелекту, який може виконувати будь-яке завдання як людина або краще за неї, було розділено на більш вузькі, конкретні задачі. В цей час комп'ютер Deep Blue, створений компанією IBM, переміг чинного чемпіона світу з шахів Гаррі Каспарова, було відновлено інтерес до теорії ігор. Також розширилось впровадження інтелектуальних систем для вирішення проблем у галузі транспорту, логістики, медицини, а також для ефективного планування і використання інтелектуальних програм в якості особистого помічника [17].

Якщо зростання обчислювальної потужності та спеціалізація призвели до поновлення розвитку штучного інтелекту на рубежі століть, то експоненціальне збільшення обчислювальної потужності та розширення можливостей зберігання інформації вивели галузь штучного інтелекту за останні роки на новий рівень.

Наразі системи та алгоритми штучного інтелекту спрямовані на вирішення питань, які вважаються буденними для людини (розпізнавання образів, що включає вміння читати, впізнавати людей та інші) та на такі, що вважаються складними (доведення математичних теорем, інтелектуальні ігри – шахи, го та інші, переробка великого масиву даних). В будь-якому разі вони імітують навички людини (такі як сприйняття, асоціювання, передбачення та планування), які допомагають їй досягати певної мети [4, с. 1]. Впливові сучасні компанії все частіше втілюють алгоритми штучного інтелекту, машинного навчання зокрема, у своїй роботі. Так, наприклад, компанії Google, Amazon, Baidu, Microsoft, IBM, Apple, Toyota використовують їх для розуміння і прогнозування поведінки споживача, для обробки природної мови, для пошуку релевантної інформації, для розпізнавання фотографій, для переклада, у безпілотних транспортних засобах. У 2016 році програма AlphaGo, розроблена компанією Google на основі глибинного навчання, виграла у вісімнадцятикратного чемпіону світу з гри го. Технології машинного навчання втілено у багатьох онлайн-сервісах.

3. Філософські проблеми штучного інтелекту

Еволюція наукового напрямку передбачає вирішення технічних і філософських проблем. Останні почали проявлятися у межах штучного інтелекту ще у 50–80-х роках ХХ століття, але їх обговорення продовжується й зараз. Як вже було зазначено вище, напрям штучного інтелекту знаходиться на перетині кількох дисциплін, тому до дискусії долучались філософи, психологи та соціологи, з одного боку, та представники природничих і технічних наук – з іншого. Більшість філософів, такі як Х. Дрейфус, Дж. Серль, Дж. Лукас, Д. Деннет та деякі дослідники (наприклад, Р. Пенроуз) схильні до критики теоретичних засад дисципліни та виділяють принципові обмеження, які накладаються на створення штучного інтелекту (Дрейфус, Серль, Лукас, Деннет). Більш оптимістично налаштовані філософи та дослідники Д. Чалмерс, Д. Льюїс, С. А. Клейн, А. Тюрінг або намагаються спростувати аргументи опонентів, або пропонують розв’язання вузьких задач [5; 24; 26;]. Розглянемо найбільш вагомі проблеми, вирішення яких дозволило б вивести дисципліну на якісно новий рівень.

Першою з них є нездатність штучного інтелекту до саморефлексії та відчуттів. Британський вчений Дж. Джефферсон у своїй промові ще у 1949 зауважив, що машина не може відчувати (а не імітувати) задоволення від своїх успіхів, лестоців, пригнічення від скоєних помилок або невдач [26, с. 444]. Також машина не може усвідомлювати себе. Як доказ на користь цього аргументу американський філософ Джон Серль приводить уявний експеримент під назвою «китайська кімната». Сутність його полягає в тому, що людині, яка не знає китайської, надається текст цієї мовою. Після цього піддослідному також дають інструкції, написані його рідною мовою, завдяки яким він зможе скласти відповідь з ієрогліфів, не розуміючи значення жодного з них. У підсумку здається, що людина володіє китайською мовою через те, що вірно реагує на наданий текст. Діяльність учасника цього експерименту Серль порівнює з комп’ютером, який, незважаючи на вірний результат виконуваних завдань, не усвідомлює значення своїх дій, не мислить і не усвідомлює себе [23, с. 418].

Дослідник також розмежував поняття «сильного» та «слабкого» штучного інтелекту [23, с. 417]. Останній означає інструмент для вивчення людської свідомості, за допомогою якого можна створювати та перевіряти гіпотези, розв’язувати задачі. Сильний штучний інтелект є комп’ютером, який запрограмований таким чином, що може знаходитись у будь-якому когнітивному стані, властивому для людини. Аргумент «китайської кімнати» спрямований саме проти можливості створення «сильного» інтелекту. Те саме стосується і наступного аргументу про неможливість звести інтелект до обчислювальної машини

Аргумент, спрямований проти обчислювального підходу, був сформульований британським філософом Джоном Лукасом у 1961 році [15] та розширений його співвітчизником Роджером Пенроузом через 30 років [20; 21]. Грунтуючись на теоремі Геделя про неповноту, Лукас стверджував, що машина не може бути повною та адекватною моделлю людського розуму [15, с. 113]. Спираючись на ту саму теорему, Пенроуз доводив, що когнітивна діяльність не може бути зведеною до обчислюваної моделі. Це твердження тісно пов’язане з так званим «онтологічним аргументом» Хьюберта Дрейфуса, згідно з яким світ не може бути описаний за допомогою формальної системи [8]. Цей аргумент спрямований проти думки, якої дотримується багато сучасних розробників штучного інтелекту: будь-

який феномен може бути описаний за допомогою об'єктів, відносин, властивостей, класів та інших елементів.

Ще однією проблемою є нездатність машини до творчої або гнучкої поведінки через те, що вона завжди дотримується закладених правил. Іншими словами, результат машинної творчості не буде вважатись чимось новим і оригінальним через те, що машина зробила те, що їй наказали, дотримуючись інструкцій (так зване «заперечення леді Лавлейс») [26; 27].

Проаналізуємо, чи були подолані вищезазначені труднощі та обмеження задля розвитку галузі в цілому.

4. Сучасний стан розвитку штучного інтелекту

Наразі існує пробіл між теоретичними дослідженнями філософів і практичною діяльністю розробників у галузі штучного інтелекту. Ця ситуація пояснюється зосередженням розробників на вирішенні нагальних питань, що гарантує їм фінансове забезпечення. Ще однією причиною є скорочення гуманітарної освіти та господарювання думки, що філософські дослідження є відірваними від реального життя. Недоліками цього становища є повторення помилок, які були зроблені в минулому, перевагами ж є відсутність тиску авторитету, що призводить до активного пошуку шляхів подолання труднощів.

Ще однією особливістю цього напрямку є господарювання в його середовищі обчислювальної парадигми. Це пояснюється тим, що сучасну людину з усіх боків оточує техніка, яка запрограмована на певні дії. Люди звикли, що ці програми, код можна змінювати, переписувати. Такий механістичний погляд переноситься й на людину, в ДНК якої записаний певний код. Через це людей порівнюють з біологічними машинами, а мозок – з комп'ютером.

Змагання мозку людини і комп'ютеру розглядалися у багатьох новітніх дослідженнях, при цьому критерій інтелектуальності в багатьох випадках залишився незмінним – тест Тюрінга – перевірка, в ході якої людина контактує з машиною, не знаючи, машина це чи людина. Ґрунтуючись на своїх особистих враженнях, перевіряючий має визначити, чи спілкувався він з машиною [26].

У 2011 році один з найстаріших у США студентських літературних журналів «Archive» опублікував вірш «For the Bristlecone Snag», написаний комп'ютерною програмою. Її автор надіслав до редакції 26 віршів, не вказавши, що вони є творами машини, але тільки один з них був прийнятим. Таким чином, вважається, що вірш пройшов тест Тюрінга [27].

Ще одним новітнім прикладом відтворення поведінки людини є гра у дебати. У 2019 році шестирічна система «Miss Debater» зіграла з 31-річним Харишем Натарайяном на конференції IBM. Протягом чотирьох хвилин опоненти наводили свої аргументи і відповідали на випадки суперника. Перемога дісталась людині, хоча в пам'яті системи міститься більше 10 мільярдів речень з 300 мільйонів наукових статей [3]. Головним, проте, є той факт, що машина поводитись на рівних з людиною, генеруючи речення промови, синтезуючи людську мову, проявляючи почуття гумору. Тобто машина теж пройшла тест Тюрінга.

Ці два приклади доводять, що завдяки технічним досягненням «слабка» форма штучного інтелекту продовжує розвиватись, освоюючи нові галузі. Перемога в шахах у 1997 довела здатність машин виконувати операції високого рівня складності, участь штучного інтелекту у дебатах стала можливою завдяки відтворенню навичок та вмінь людей (таких як генерування мови, почуття гумору).

Останнє є більш складним завданням, бо, згідно з теорією Поланьї, не може бути описано за допомогою формальної системи. Ця теорія пояснює парадокс Моравека, згідно з яким міркування високого рівня (такі, наприклад, як гра в шахи) вимагає відносно малих обчислень порівняно з відтворенням повсякденних навичок людини. При цьому не йдеться про створення «сильного» інтелекту, бо наведені програми тільки імітують діяльність свідомості.

Проблема розуміння дуже широко обговорюється у середовищі розробників систем штучного інтелекту, бо вона стосується кількох галузей [16]. Розуміння є складним процесом, і ні філософія, ні психологія чи когнітивна наука не дають однозначного визначення цьому поняттю. Ґрунтуючись на прагматичному підході, у штучному інтелекті розуміння було розділено на складові операції, і кожен з розробників працює в своїй області:

- моделювання еволюції мови через взаємодію “машина–людина” та “машина–машина” [16, с. 45; 32];
- відтворення взаємозв’язку мови, тілесних рухів та міміки у роботів [11; 16, с. 46; 6];
- автоматична обробка тексту та розпізнання природної мови [29; 16, с. 47];
- розпізнання емоцій у тексті або усному мовленні [2];
- методи представлення знання [16, с. 52–54] та інші.

Слід зауважити, що все частіше термін «обробка природної мови» замінюється «розумінням природної мови» [29], але насправді це не розуміння, а імітація розуміння. Окрім розуміння, імітуються інтуїція [9], творчість [1], причому науковці визнають, що лише імітують творчі процеси людини і називають її «слабкою творчістю» за аналогією зі «слабким інтелектом» Серля [1].

Окремо треба виділити дослідження у галузі самосвідомості роботів. Одним з підходів до розробок у цьому напрямі є розпізнавання ментальних станів людини через її міміку, іншим є виділення роботом власних внутрішніх станів, не схожих на людські [10]. Знову ж самоконтроль і саморегулювання не є тотожним самосвідомості, а лише імітують її.

Питання, чи може когнітивна діяльність бути зведеною до обчислюваної моделі, досі залишається без відповіді, незважаючи на численні дослідження на перетині когнітивістики, штучного інтелекту та нейронауки [19; 14]. Наприклад, прихильник обчислювального підходу Д. Чалмерс так і не знайшов клас обчислень, якого було б достатньо для відтворення ментальності людини. Філософ тільки припускав, що це можуть бути зв’язки нейронів, обчислення, які відображають каузальну організацію мозку на нижчому рівні, або щось інше [5, с. 310]. Наразі ефективною моделлю в певних областях є нейронні мережі. Завдяки технології глибинного навчання в сучасні системи закладено механізм самонавчання, самовдосконалення та самоналаштування мереж. Вони використовувались для навчання програми AlphaGo, вони ж використовуються у програмі Google Photos. Результати роботи цих програм вражають, проте є в них і недолік: непередбачуваність, бо розробники не можуть контролювати процес навчання. Це призвело, наприклад, до того, що у 2015 році Google Photos розпізнала кількох афроамериканців як приматів [30]. До того ж, як зауважував Норберт Вінер, складні

обчислювальні машини можуть припускати спонтанні функціональні помилки [13, с. 1109].

Проблема формалізації знання наразі залишається нерозв'язною. Її підкріплює теорія про неявне знання Майкла Поланьї, згідно з якою не все знання може бути викладеним формально. До неявного знання належать деякі навички, вміння, наприклад, їзда на велосипеді, гра на музичних інструментах, забивання цвяхів молотком [22]. Через принципову нерозв'язність проблеми у штучному інтелекті до неї був застосований типовий для цієї галузі підхід. Масштабний проект опису світу було замінено на опис вузьких предметних областей, задач або ситуацій. При цьому для кожної задачі виділяють ключові об'єкти, ідеалізуючи їх й абстрагуючись від менш важливих об'єктів, властивостей та зв'язків. Методи представлення знання підбираються окремо для кожного випадку: можуть застосовувати, наприклад, описову логіку чи онтологію [16, с. 52–54]. Іншими словами, уявлення про світ спрощується задля вирішення конкретної задачі.

5. Висновки

Штучний інтелект став продуктом постіндустріального техногенного суспільства з його орієнтацією на креативну діяльність, впровадження нововведень, виробництво послуг, проведення досліджень, підвищення якості життя. За час свого розвитку він зазнав декілька періодів розквіту і занепаду, останні з яких були спричинені обмеженнями як технічного, так і філософського характеру. Серед філософських проблем слід виділити: обмеженість обчислювального підходу, нездатність штучного інтелекту до саморефлексії, відчуттів, творчості, гнучкої поведінки.

Новий спалах досліджень у межах штучного інтелекту був спричинений стрімким розвитком технологій, з одного боку, який допоміг реалізації алгоритмів, які раніше вимагали багатогодинних обчислень. З іншого – він став результатом інтенсифікації процесу інформатизації буття людини і суспільства в цілому. Це призвело до прогресу у сфері «слабкого» інтелекту, до масового створення програм, що імітують людську діяльність, поведінку, відчуття. У межах напряму широко використовується прагматичний підхід, який дозволяє вирішувати вузькі задачі, формалізувати вузькі предметні області. При цьому ймовірність створення “сильного” штучного інтелекту не збільшилась, і наразі ніхто з дослідників не ставить його відтворення за мету. Тобто різкий стрибок у розвитку напряму зобов'язаний тільки технічній складовій, а жодна з філософських проблем не була вирішена, через що не можна говорити про початок якісно нового етапу еволюції штучного інтелекту.

Сучасний стан напряму набуває рис постнекласичної науки через зростання важливості міждисциплінарних досліджень; більшу аксіологічну, економічну і прагматичну спрямованість; технічний прогрес. Кожна окрема сфера в межах цього напряму має потенціал для нових практичних і теоретичних досягнень, тому дослідження штучного інтелекту залишатимуться в центрі уваги і надалі.

REFERENCES

1. *al-Rifaie M. M., Bishop M. (2015). Weak and Strong Computational Creativity. In Computational Creativity Research: Towards Creative Machines. T.R. Besold, M. Schorlemmer, A. Smaill (Eds). Atlantis Press, pp. 37–49.*

2. Appel O. (2017). *A Hybrid Approach to the Sentiment Analysis Problem at the Sentence Level*. De Montfort University, Leicester, Great Britain.
3. Carville O., Kahn J. (2019). A Human Just Triumphed Over IBM's Six-Year-Old AI Debater. *Bloomberg*, February 12. URL: <https://www.bloomberg.com/news/articles/2019-02-12/in-latest-man-vs-machine-human-triumphs-over-ibm-s-ai-debater>
4. Boden M. A. (2016). *AI: Its nature and future*. Oxford University Press.
5. Chalmers D. (1996). *The conscious mind: in search of fundamental theory*. – N.Y.: Oxford University Press. doi: <https://doi.org/10.1007/s11948-017-9975-2>
6. Diamond A., Knight R., Devereux D., Holland O. (2012). Anthropomimetic Robots: Concept, Construction and Modelling. *International Journal of Advanced Robotic Systems*, Vol. 9(5), pp.1–14. doi: <https://doi.org/10.5772/52421>.
7. Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, vol. 20, issue 1, pp. 1–3.
8. Dreyfus H. (1976). What computers can't do. *British Journal for the Philosophy of Science*, vol. 27 (2), pp. 177–185.
9. Dzhabrailov S.V., Rozaliev V.L., Orlova Y.A. (2017). Approaches and implementation of computer simulation intuition. *Internet-journal Naukovedenie*, vol. 9, No 2. URL: <http://naukovedenie.ru/PDF/38TVN217.pdf> (in Russian).
10. Gorbenko A., Popov V. (2012). Robot Self-Awareness: Exploration of Internal States. *Applied Mathematical Sciences*, vol. 6, no. 14, pp. 675–688.
11. Hild M., Siedel T., Benckendorff C., Thiele C., Spranger M. (2014). M. Myon, a New Humanoid. *Language Grounding in Robots*, pp. 25–44. doi: 10.1007/978-1-4614-3064-3_2
12. Howard A., Borenstein J. (2018). The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity. *Science and Engineering Ethics*, no. 24, pp. 1521–1536. doi: 10.1007/s11948-017-9975-2.
13. Jefferson G. (1949). The Mind of Mechanical Man. *British Medical Journal*, pp. 1105–1110.
14. Kriegeskorte N., Douglas PK. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, vol. 21(9), pp. 1148–1160.
15. Lucas, J. R. (1961). "Minds, Machines and Gödel". *Philosophy*, vol. 36, pp. 112–127.
16. Maksimov V. Yu., Klyshinskij E. S., Antonov N. V. (2016). The problem of understanding in artificial intelligence systems. *New Information Technologies in Automated Systems*, pp. 43–60.
17. Martínez-Plumed F., Loe BS., Flach P., O hEigeartaigh S., Vold K., Hernandez-Oralloet J. (2018). The Facets of Artificial Intelligence: A Framework to Track the Evolution of AI. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18) Evolution of the contours of AI*, pp. 5180–5187.
18. Moor J. (2006). The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazine – American Association for Artificial Intelligence*, vol. 27, no 4, pp. 87–91.
19. Naselaris T., Bassett DS., Fletcher AK., Kording K., Kriegeskorte N., Nienborg H. et al. (2018). Cognitive Computational Neuroscience: A New Conference for an Emerging Discipline. *Trends in Cognitive Science*, Vol. 22(5), pp. 365–367. doi: 10.1016/j.tics.2018.02.008.
20. Penrose R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.
21. Penrose R. (1994). *Shadows of the Mind*. Oxford: Oxford University Press.
22. Polanyi M. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy*. University of Chicago Press, Chicago.

23. Searle J. (1980). Minds, brains, and programs. *The behavioral and brain sciences*, vol. 3, pp. 417–457. Sokolov E. N. (2013). *The Psychophysiology of Consciousness* (A. M. Chernorizov, K. A. Chernorizov, D. M. Bowden Trans). – N.Y.: Oxford University Press.
25. The Future of Jobs Report (2018). URL: http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf
26. Turing A.M. (1950). Computing Machinery and Intelligence. *Mind*, vol. 49, pp. 433–460.
27. Walsh T. (2017). *Android Dreams: The Past, Present and Future of Artificial Intelligence*. Hurst & Company, London.
28. Wellens P., van Trijp V., Beuls K., Steels L. (2013). Fluid Construction Grammar for Historical and Evolutionary Linguistics. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. L. Steels, M. Hild. (Ed.). Sofia, Bulgaria: Association for Computational Linguistics, pp.127–132.
29. Young T., Hazarika D., Poria S., Cambria E. (2018). Recent Trends in Deep Learning Based Natural Language Processing. *IEEE Computational Intelligence Magazine*, vol. 13(3), pp. 55–75. doi: [10.1109/MCI.2018.2840738](https://doi.org/10.1109/MCI.2018.2840738)
30. Zhang M. (2015). Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software. *Forbes*, Jul 1 [Електронний ресурс]. – Режим доступа : <https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/#6b7e969e713d>

Received 15.03. 2019

Doi: 10.15421/271903

UDK 161/159.95

MODALITY AND FOLK PSYCHOLOGY

Ja. O. Petik

*G. S. Skovoroda Institute of Philosophy
of the National Academy of Sciences of Ukraine,
Kyiv, Ukraine*

Tel.: +38-098-744-55-45

E-mail: iaroslav.petik@gmail.com

ORCID 0000-0002-6127-5943

Abstract. The connection of the modern psychology and formal systems remains an important direction of research. This paper is centered on philosophical problems surrounding relations between mental and logic. Main attention is given to philosophy of logic but certain ideas are introduced that can be incorporated into the practical philosophical logic.

The definition and properties of basic modal logic and descending ones which are used in study of mental activity are in view. The defining role of philosophical interpretation of modality for the particular formal system used for research in the field of psychological states of agents is postulated. Different semantics of modal logic are studied. The hypothesis about the connection of research in cognitive psychology (semantics of brain activity) and formal systems connected to research of psychological states is stated.