

ОГЛЯД СТРУКТУР ІНТЕЛЕКТУАЛЬНИХ АГЕНТІВ ТА ІННОВАЦІЙНИЙ ПІДХІД ДО ІНТЕГРАЦІЇ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ У АРХІТЕКТУРУ BDI

М. В. Крихівський, Т. О. Ваврик*, Л. М. Гобир

ІФНТУНГ; 76019, м. Івано-Франківськ, вул. Карпатська, 15, тел. +38 (042) 727139;
e-mail: tetiana.vavryk@nung.edu.ua

Інтелектуальні агенти стали центральною темою сучасних досліджень у сфері штучного інтелекту. Вони є важливим компонентом сучасних інформаційних технологій, що використовуються для автоматизації складних завдань, таких як обробка даних, прийняття рішень та адаптивне навчання. Розробка інтелектуальних агентів є складним, багатогранним процесом, що включає вибір архітектури, розробку алгоритмів прийняття рішень та інтеграцію з середовищем. Протягом останніх десятиліть з'явилося багато нових технологій та підходів, які дозволяють значно покращити ефективність агентів у реальних застосуваннях. Досліджуються концептуальні засади інтелектуальних агентів, різні підходи до їх проєктування та ключові технології, що дозволяють створювати автономні та адаптивні системи. Авторами розглядається загальний алгоритм функціонування інтелектуального агента, його основні етапи та архітектура. Представлено огляд концепції інтелектуальних агентів, включаючи їх базові властивості, моделі взаємодії з навколишнім середовищем та адаптивні механізми. У статті представлено огляд сучасного стану проєктування інтелектуальних агентів, підкреслюючи значні досягнення та основні виклики у цій галузі. Основна увага зосереджена на архітектурі BDI (Belief-Desire-Intention), яка є однією з найпоширеніших та ефективних підходів до створення інтелектуальних агентів. Переконавання, бажання та наміри агентів моделюються для забезпечення більш адаптивної та гнучкої поведінки у динамічних середовищах. Стаття пропонує новий підхід до інтеграції рекурентних нейронних мереж (RNN) у архітектуру BDI для покращення здатності агентів до обробки послідовних даних та врахування контексту попередніх подій. RNN забезпечують ефективне запам'ятовування та аналіз тимчасових залежностей, що є ключовим для прийняття рішень в умовах невизначеності та змінних обставин. Інтеграція RNN дозволяє агентам не лише оновлювати свої переконання в режимі реального часу, а й покращувати процеси формування бажань та намірів, роблячи їх більш точними та адаптивними. Основні результати дослідження демонструють, що впровадження RNN у BDI архітектуру значно підвищує ефективність та продуктивність агентів, дозволяючи їм швидко адаптувати свої стратегії до нових умов. Це відкриває нові можливості для застосування таких агентів у різних сферах, включаючи кібербезпеку, управління енергією, охорону здоров'я та багато інших. Загалом, ця стаття робить важливий внесок у розвиток теорії та практики проєктування інтелектуальних агентів, пропонує інноваційний підхід до інтеграції рекурентних нейронних мереж у архітектуру BDI, що забезпечує більшу адаптивність та гнучкість у динамічних середовищах.

Ключові слова: інтелектуальні агенти, архітектура інтелектуальних агентів, рекурентні нейронні мережі.

Intelligent agents have become a central theme in modern artificial intelligence research. They are an essential component of contemporary information technologies, used to automate complex tasks such as data processing, decision-making, and adaptive learning. The development of intelligent agents is a complex, multifaceted process that includes the selection of architecture, the development of decision-making algorithms, and integration with the environment. Over the past decades, many new technologies and approaches have emerged that significantly improve the efficiency of agents in real-world applications. Conceptual foundations of intelligent agents, various approaches to their design, and key technologies that enable the creation of autonomous and adaptive systems are explored. The authors examine the general algorithm of the intelligent agent's functioning, its main stages, and architecture. An overview of the concept of intelligent agents is presented, including their basic properties, models of interaction with the environment, and adaptive mechanisms. The article provides an overview of the current state of intelligent agent design, highlighting significant achievements and key challenges in the field. The focus is on the BDI (Belief-Desire-Intention) architecture, which is one of the most prevalent and effective approaches to creating intelligent agents. The beliefs, desires, and intentions of agents are modeled to ensure more adaptive and flexible behavior in dynamic environments. The article presents a novel method for incorporating recurrent neural networks (RNNs) into the BDI architecture to improve agents' ability to comprehend sequential data and evaluate the context of past events. RNNs provide efficient memory and analysis of temporal connections, which is critical for decision-making in uncertain and changing environments. Integrating RNNs enables agents to update their beliefs in real time while also improving the processes of developing desires and intentions, making them more accurate and adap-

tive. The main findings of the research demonstrate that the implementation of RNNs in the BDI architecture significantly enhances the efficiency and productivity of agents, allowing them to quickly adapt their strategies to new conditions. This opens up new opportunities for applying such agents in various fields, including cybersecurity, energy management, healthcare, and many others. Overall, this article makes a significant contribution to the development of the theory and practice of intelligent agent design, offering an innovative approach to integrating recurrent neural networks into the BDI architecture, which ensures greater adaptability and flexibility in dynamic environments.

Keywords: intelligent agents, agent architecture, recurrent neural networks.

Вступ

Поняття «агент» було вперше введено в контексті штучного інтелекту (ШІ) в 1970-х роках, але значне поширення отримало в останні десятиліття з розвитком комп'ютерних технологій та зростанням потужностей обчислювальних систем. На сучасному етапі розвитку науки та технологій інтелектуальних агентів (ІА) стають все більш важливими компонентами складних систем, які взаємодіють з навколишнім середовищем та приймають рішення на основі обробки великої кількості даних. Від архітектури до математичних моделей, проектування таких агентів вимагає глибокого розуміння різних дисциплін, включаючи інформатику, кібернетику, теорію систем та штучний інтелект.

У цій статті розглянуто основні принципи та підходи до проектування ІА, з особливим акцентом на їхню архітектуру та математичне моделювання. Ми прагнемо створити цілісне уявлення про теоретичні основи та практичні аспекти цього процесу, що дозволить розробникам і дослідникам ефективно використовувати ці знання для створення надійних та адаптивних систем.

Актуальність цієї статті полягає у тому, що ІА відіграють ключову роль у багатьох галузях, включаючи робототехніку, автоматизацію, системи підтримки прийняття рішень та, навіть, побутові застосування. Розуміння основ проектування таких агентів допомагає дослідникам та розробникам створювати ефективні та надійні системи, що можуть виконувати складні завдання з мінімальним втручанням людини. Окрім того, стаття також надає важливу теоретичну базу для подальших досліджень і вдосконалення алгоритмів штучного інтелекту (ШІ), що робить її незамінним джерелом знань для науковців і фахівців у цій сфері.

Метою статті є всебічний огляд та аналіз ключових аспектів створення ІА. Стаття спрямована на забезпечення теоретичної та практичної основи для розуміння процесу розробки таких агентів, починаючи від їхньої архітектури та закінчуючи математичними моделями, які описують їх поведінку та функціонування. Це

прагнення узагальнити існуючі знання та досвід, а також запропонувати нові підходи та методи, що сприятимуть підвищенню ефективності та надійності інтелектуальних систем у різних сферах застосування.

Аналіз сучасних закордонних і вітчизняних досліджень та публікацій

Методологія BDI (Belief-Desire-Intention), або переконання-бажання-наміри (ПБН), є одним із підходів до моделювання агентів у галузі штучного інтелекту. Вона базується на психологічних теоріях, що пояснюють, як люди приймають рішення, і переносить ці принципи на штучних агентів, що дозволяє їм діяти раціонально та автономно.

Основні компоненти ПБН моделі [1] – це переконання (Beliefs), бажання (Desires) та наміри (Intentions). Переконання – це знання або переконання агента про світ, які можуть бути правдивими або хибними. Це інформація, яка допомагає агенту оцінювати навколишню ситуацію та приймати рішення. Бажання – це цілі або завдання агента, яких він хоче досягти. Бажання спрямовують його дії та визначають те, що агент намагається здійснити. Наміри – це конкретні плани або дії, які агент обирає для досягнення своїх цілей. Це рішення, які агент приймає, виходячи з своїх переконань та бажань.

ПБН архітектура дозволяє агенту працювати у складних динамічних середовищах, приймаючи розумні та гнучкі рішення. Агент постійно оновлює свої переконання на основі нової інформації, коригує свої бажання, коли змінюються його пріоритети або умови, та обирає нові інтенції, коли це потрібно для досягнення цілей. Це надає агенту здатності адаптуватися до змін та непередбачуваних ситуацій.

Важливою характеристикою ПБН агентів є те, що вони здатні вирішувати конфлікти між різними бажаннями та інтенціями, враховуючи пріоритети та ресурси. Наприклад, якщо агент має кілька суперечливих цілей, він зможе оцінити, яка з них важливіша або більш досяжна у поточний момент, та вибрати відповідну дію. Проте, дослідники [2,3] відмічають відсутність

деяких компонентів гнучкості в моделі BDI та проблемність навчання.

Авторами оглядової статті [3] розглянуто основні підходи до кожного компонента архітектури ПБН, те, як вони були реалізовані в мовах програмування агентів, і показано компроміси, властиві кожному з підходів. Зроблено висновки про те, що для нових областей, які повинні інтегрувати субсимвольну інформацію від датчиків у переконання агентів, знадобляться більш насичені уявлення, ніж в роботі [1]. Це може передбачати використання вкладених та невизначених переконань.

Довгострокова автономія інтелектуальних агентів (IA) стає необхідною для багатьох автономних систем. Це стимулює подальші дослідження підходів до цілей у контексті того, що цілі підтримання та їх взаємодія з цілями досягнення стають більш важливими, ніж властивості та обмеження цілей, такі як пріоритети та дедлайни.

У роботі [3] виділяється проблема справжнього одночасного виконання планів. Цей напрям передбачає дослідження з формалізації взаємодій між одночасно виконуваними гілками та між циклом міркувань IA та тривалими діями. Інша область майбутніх досліджень стосується використання технік штучного інтелекту, таких як машинне навчання, для доповнення заздалегідь визначених правил плану або генерування нових правил. Авторами підкреслено перспективу підтримки невизначених переконань, що є важливим, коли декларативна ціль може вважатися «досягнутою» (або недосяжною), і для підходів до планування намірів на основі симуляцій чергувань.

У статті [4] запропоновано архітектуру виконання плану, яка підтримує різну семантику завдань. Кожна мета, що визначена під час циклу обговорення агента, може бути задоволена через завдання, визначені в різній семантиці. Можливість підтримувати різну семантику завдань забезпечує дві основні переваги. Першою перевагою є повторне використання застарілих артефактів у планах агента. Другою перевагою є спрощення адаптації архітектури агента до різних стандартів організації бізнесу. Щоб інтегрувати різноманітну семантику завдання в архітектуру виконання плану, автори [4] використали плавно переглянута версію попередньо сформульованої моделі робочого процесу, у якій різна семантика завдання була б зменшена перед виконанням. Інтегрування ієрархічної мережі завдань і семантики OWL-S в цю архітектуру виконання плану, дозволило перевірити

її міцність, з точки зору підтримки виконання різних семантик завдань в архітектурі агента.

Авторами роботи [5] створено структуру, яка додає знання для покращення вибору плану в популярній парадигмі програмування агентів ПБН. На відміну від попередніх пропозицій, наведений підхід може добре масштабуватись із складністю бібліотеки планів агента. Технічно розроблено нову міру впевненості, яка дозволяє агенту динамічно регулювати свою залежність від навчання, полегшуючи в принципі нескінченну кількість фаз (повторного) навчання.

Традиційний агент ПБН має три основні обчислювальні компоненти: ті, що генерують переконання, генерують наміри та виконують наміри. Вони виконуються послідовно і циклічно. Це може спричинити кілька проблем. Серед них нездатність постійно спостерігати за навколишнім середовищем у динамічних середовищах може бути катастрофічною. Авторами роботи [6] запропоновано загальну структуру для моделі паралельного агента ПБН. За цією загальною структурою можна створювати паралельні агенти ПБН з різними конфігураціями залежно від доступності фізичних ресурсів. Ці агенти мають низку переваг перед послідовними, а саме, зміни в середовищі агента можуть бути виявлені негайно, надзвичайні ситуації будуть вирішені негайно, підтримка надається на рівні архітектури для перегляду бажань/намірів і розгляду зв'язків між цілями, коли генерується нове переконання/бажання.

Концепції IA представлені в дослідницькій книзі [7], у якій сформована теорія агентів та адаптацію агентів у різних контекстах. Агенти різного порядку складності повинні бути автономними у використовуваних правилах. Агент повинен мати мозок, за допомогою якого він може виявити правила, що містяться в даних. Оскільки правила є інструментами, за допомогою яких агенти змінюють середовище, будь-яку адаптацію правил можна розглядати як еволюцію агентів. Оскільки невизначеність присутня в будь-якому контексті, авторами книги [7] описано, як можна ввести глобальну невизначеність з локального світу в опис правил. У книзі представлено загальний опис агента еволюційної адаптації, дії та мета дії такого агента, абстрактне та формальне представлення дії, зв'язок системи та метасистеми з адаптивним агентом, мозок агента за допомогою теорії морфогенетичних нейронів, логічну структуру адаптивного агента, зворотний зв'язок і гіперзворотний зв'язок в адаптивному агенті, поле адаптації в модальний логічний простір як логічний інструмент в адаптивному агенті, дію

агента у фізичному домені та застосування агентів у роботах та еволюційних обчисленнях.

Автори роботи [8] представляють багато-агентну архітектурну теорію для розробки інтелектуальних асоціативних гібридних систем. Теорія архітектури описана як на рівні структури завдання, так і на рівні обчислень. Ця архітектура призначена для вирішення проблем розробки агентів знань та інформаційних агентів.

Архітектури автономних агентів – це методології проєктування сукупності знань і стратегій, які застосовуються до проблеми створення локального інтелекту. У статті [9] зроблено спробу об'єднати ці знання з кількома архітектурними традиціями, приділяючи особливу увагу особливостям, які, як правило, вибиралися під тиском широкого використання в системах реального світу. Автори визначили, що наступні стратегії надають значну допомогу в проєктуванні автономних інтелектуальних агентів, а саме, модульність, яка спрощує як дизайн, так і контроль, ієрархічно організований вибір дій, який фокусує увагу та забезпечує пріоритетність, коли різні модулі конфліктують; та паралельний моніторинг середовища, який дозволяє системі бути чутливою та оппортуністичною, дозволяючи переключити увагу та переоцінити пріоритети. Ми пропонуємо огляд чотирьох архітектурних парадигм: III на основі поведінки; дво- і тришарові системи; архітектури переконань, бажання та наміру (зокрема PRS); і Soar/ACT-R. Задokumentувавши тенденції в кожній із цих спільнот щодо створення компонентів вище, ми стверджуємо, що ця конвергентна еволюція є переконливим доказом корисності компонентів. Потім ми використовуємо цю інформацію, щоб рекомендувати конкретні стратегії для дослідників, які працюють за кожною парадигмою, для подальшого використання знань і досвіду в галузі загалом.

Механізми емоцій часто використовуються в штучних агентах як метод покращення вибору дій. Порівняння між агентами є складними через відсутність єдності між теоріями емоцій, завданнями агентів і типами вибору дій, що використовуються. Набір архітектурних якостей пропонується у роботі [10] як основа для проведення порівнянь між агентами. Аналіз існуючих архітектур агентів, які включають механізм емоцій, може допомогти триангулювати можливості дизайну в просторі, окресленому цими якостями. З огляду на це, для аналізу обрано дванадцять автономних агентів, які включають механізм емоцій у вибір дій. Кожен агент розбирається за допомогою цих архітектурних якостей (архітектура агента, механізм вибору

дій, механізм емоцій і представлення емоційного стану, а також модель емоцій, на якій він базується). Це допомагає розмістити агентів в архітектурному просторі, підкреслює контрастні методи реалізації схожих теоретичних компонентів і підказує, які архітектурні аспекти важливі для виконання завдань. Представлено початкову структуру, що складається з ряду рекомендацій щодо розробки механізмів емоцій у штучних агентах, заснованих на кореляції між виконуваними ролями емоцій та аспектами механізмів емоцій, які використовуються для виконання цих ролей. У висновку обговорюється, як можна вирішити проблеми з цим типом дослідження та якою мірою розробка основи може допомогти при майбутніх дослідженнях.

Логіка ПБН є важливим і широко використовуваним теоретичним апаратом для представлення та міркування про раціональних агентів. Однак вона є неповною щодо перегляду наміру, перевизначення наміру, процесу обговорення та перегляду переконань. Крім того, деякі раціональні агенти, особливо люди, не мають наближених міркувань, добре представлених логікою ПБН. У статті [11] визначено нечітку семантику для мови ПБН, яка здатна усунути ці обмеження. Запропоновано розширити логіку ПБН до нечіткої логіки ПБН.

У статтях [12-14] проведено аналіз педагогічних та інформаційних засад використання інтелектуальних агентів в адаптивних системах електронного навчання з метою створення структури адаптивної системи електронного навчання. Авторами описано характеристики, функції та взаємодії агентів, які беруть участь у кожному модулі адаптивної архітектури, а також характеристики, функції та взаємодії інтелектуального агента для ухвалення навчальних рішень.

У роботі [15] представлено підхід до розробки ІА, який можна використовувати в мультиагентній системі для вирішення складної проблеми. Запропонований агент має можливість навчатися за допомогою штучних нейронних мереж.

Авторами статті [16] розглядаються проблеми створення надійних децентралізованих програм способом, сумісним із наскрізною моделлю. Створена ними програма Mandrake базується на розумінні того, що можливість агента отримувати повідомлення в будь-якому порядку та обробляти їх відповідно до його власного прийняття рішення є вирішальною для реалізації сенсу програми.

Автори публікації [17] зазначають, що III – ніщо без інтелектуальних агентів, і досліджу-

ють, скільки типів агентів визначено в штучному інтелекті. Вони виділяють п'ять категорій, які згруповані за діапазоном можливостей і ступенем сприйманого інтелекту, і роблять висновки, що кожна з них здатна втілити ідеї ШІ в життя.

У дослідженнях [18-21] розшифровано багатогранну природу інтелектуальних агентів, проаналізовано їх застосування в реальному світі, переваги та потенційні обмеження. Авторі розглянули класифікацію інтелектуальних агентів, що включає різні типи, такі як прості рефлексивні агенти, модельно-рефлексивні агенти, цільові агенти та корисні агенти. Кожен варіант наділяє агента різними можливостями та функціями, що відповідають різноманітним вимогам додатків ШІ.

Невирішені питання

Проектування ІА залишається складним завданням через низку невирішених проблем. Однією з головних проблем є забезпечення агента здатністю до розуміння та інтерпретації контексту, адже обробка мови вимагає тонкого розуміння контекстуальних нюансів. Ще одним важливим аспектом є навчання агентів з урахуванням етичних міркувань [22] та уникнення упередженості, що виникає через перекоси в навчальних даних. Крім того, адаптивність агентів до нових, непередбачуваних ситуацій і їх здатність до самонавчання та покращення в реальному часі також викликають занепокоєння. Безпека та конфіденційність даних користувачів залишаються критичними питаннями, які потрібно розв'язувати. Важливим є також забезпечення природної та ефективної взаємодії з користувачами, що включає виявлення та відповідь на емоції. Усі ці питання потребують інноваційних підходів та подальших досліджень для досягнення значного прогресу.

Створення універсальної архітектури інтелектуальних агентів, що поєднує адаптивність, масштабованість і здатність інтегрувати різні підходи, залишається важливим і складним завданням. Однією з ключових проблем є забезпечення агентів здатністю до гнучкого реагування на зміну середовища, що вимагає ефективних механізмів самонавчання та адаптації до нових умов. Це включає в себе використання методів машинного навчання та штучного інтелекту для постійного покращення та коригування поведінки агентів.

Масштабованість є важливим аспектом, оскільки для роботи з великими системами потрібно забезпечити ефективний розподіл ресурсів і обробку даних. Це вимагає розробки архі-

тектури, здатної до горизонтального масштабування, що дозволяє збільшувати потужність системи шляхом додавання нових вузлів або компонентів без зниження продуктивності.

Інтеграція різних підходів, таких як реактивні, розподілені та гібридні, є важливою для створення універсальної архітектури, здатної вирішувати різноманітні завдання. Це потребує забезпечення узгодженості та взаємодії між компонентами системи, що використовують різні підходи, без втрати ефективності. Важливо також враховувати питання безпеки, конфіденційності та етичних міркувань при розробці таких систем. Усі ці аспекти вимагають інноваційних підходів та ретельного планування для досягнення гармонійної та ефективної взаємодії компонентів інтелектуальних агентів.

Пропозиція для вирішення проблеми:

1. Інтеграція рекурентних нейронних мереж у архітектуру інтелектуальних агентів структури ПБН.
2. Генерація ідей щодо вдосконалення архітектури ПБН агентів на основі рекурентних нейронних мереж.

Очікувані результати:

1. Гнучка та масштабована архітектура, яка адаптується до різноманітних середовищ.
2. Підвищення ефективності агентів у реальному часі.

Виклад основного матеріалу дослідження

Сучасні підходи до проектування інтелектуальних агентів включають низку методологій та технологій, що дозволяють створювати ефективні та гнучкі системи для розв'язання різноманітних завдань. Одним з таких підходів є агентно-орієнтоване програмування, яке базується на концепції автономних агентів, що взаємодіють з оточенням й іншими агентами для досягнення своїх цілей.

У основі сучасного проектування інтелектуальних агентів лежить використання машинного навчання та штучного інтелекту. Одним з ключових аспектів є навчання агентів на основі підкріплення (reinforcement learning), що дозволяє агентам навчатися на власному досвіді через взаємодію з середовищем. Такі агенти використовують алгоритми, що дозволяють їм удосконалювати свої стратегії та приймати оптимальні рішення на основі отриманих винагород.

Ще одним важливим підходом є використання нейронних мереж, зокрема, глибокого навчання (deep learning). Глибокі нейронні ме-

режі дозволяють агентам аналізувати великі обсяги даних і виявляти приховані закономірності, що підвищує їхню здатність до передбачення та прийняття рішень. Такі мережі можуть бути використані для обробки візуальної інформації, розпізнавання мови, а також для розв'язання складних задач, таких як ігри чи стратегічне планування.

Сучасні підходи також передбачають використання мультиагентних систем, де кілька агентів взаємодіють між собою для досягнення спільних цілей. У таких системах агенти можуть співпрацювати, координувати свої дії та обмінюватися інформацією, що дозволяє підвищити ефективність і надійність системи загалом. Для цього використовуються методи розподіленого штучного інтелекту та алгоритми консенсусу.

Крім того, важливим аспектом проектування інтелектуальних агентів є забезпечення їхньої адаптивності та масштабованості. Сучасні агенти повинні бути здатні адаптуватися до змін у середовищі, а також масштабуватися для роботи в різних умовах і з різним рівнем навантаженості. Для цього використовуються методи самоорганізації та автономного управління.

Отже, сучасні підходи до проектування інтелектуальних агентів базуються на використанні машинного навчання, нейронних мереж, мультиагентних систем і методів адаптивного управління. Ці підходи дозволяють створювати гнучкі, ефективні та надійні системи, здатні розв'язувати складні задачі в різноманітних галузях.

Існує кілька інших методологій для моделювання агентів, кожна з яких має свої особливості та відрізняється від ПБН підходу:

1. Реактивні агенти: Ці агенти працюють на основі безпосередніх реакцій на зміни в навколишньому середовищі. Вони не мають внутрішнього уявлення про світ чи довгострокових цілей. Натомість вони реагують на вхідні стимули за допомогою заздалегідь визначених правил або алгоритмів. Це робить їх швидкими та ефективними у простих і стабільних середовищах, однак у складних та динамічних умовах вони можуть бути менш гнучкими.

2. Цільові агенти (Goal-oriented agents): Ці агенти працюють на основі цілей, які вони намагаються досягти. Вони використовують планування та розв'язання задач для досягнення своїх цілей. Цей підхід дозволяє агентам бути більш раціональними та адаптивними, оскільки вони можуть планувати свої дії на основі поточної ситуації та ресурсів.

3. Агенти на основі користувацьких моделей (User model-based agents): Ці агенти використовують моделі користувачів для розуміння та прогнозування їх поведінки. Вони можуть адаптувати свої дії та відповіді на основі індивідуальних характеристик та потреб користувачів. Це корисно у застосунках, де важлива взаємодія з користувачем, наприклад, у рекомендаційних системах або персональних асистентах.

4. Соціальні агенти: Ці агенти враховують соціальні взаємодії та відносини з іншими агентами. Вони можуть співпрацювати, конкурувати або вести переговори з іншими агентами для досягнення своїх цілей. Соціальні агенти зазвичай використовують моделі багатоагентних систем та ігрову теорію для ефективної взаємодії.

5. Гібридні агенти: Цей підхід поєднує елементи кількох методологій для створення більш потужних і гнучких агентів. Наприклад, агент може використовувати реактивні стратегії для швидкої реакції на зміни в середовищі та планування на основі цілей для довгострокових завдань. Гібридні агенти можуть мати складну архітектуру, але вони здатні краще адаптуватися до різноманітних умов.

Інтеграція рекурентних нейронних мереж (РНМ) до ПБН агентів дає можливість підвищити їхню здатність до навчання та адаптації у динамічних середовищах. Рекурентні нейронні мережі, особливо такі їх варіанти, як LSTM (Long Short-Term Memory) і GRU (Gated Recurrent Unit), забезпечують агентів здатністю до обробки та зберігання інформації про попередні стани, що є критично важливим для прийняття рішень в умовах невизначеності.

Модель ПБН на основі РНМ складається з кількох етапів. Спершу агент формує свої переконання, зчитуючи інформацію з середовища та застосовуючи РНМ для обробки історичних даних. Потім, на основі отриманих переконань, агент визначає свої бажання, що можуть бути представлені у вигляді цілей або бажаних станів. Нарешті, агент формує наміри, визначаючи дії, які необхідно виконати для досягнення цих бажань. Рекурентні нейронні мережі відіграють важливу роль на всіх етапах, забезпечуючи гнучкість і здатність до адаптації агентів у змінних умовах.

Комбінація ПБН та РНМ дозволяє створювати агентів, які мають глибоке розуміння контексту та можуть ефективно реагувати на нові виклики, засновані на їхньому минулому досвіді. Це відкриває нові перспективи для застосування таких агентів у різних сферах, включаю-

чи кібербезпеку, управління енергією та багато інших.

Пропонуємо кілька ідей щодо вдосконалення архітектури BDI агентів на основі рекурентних нейронних мереж, які можуть виявитися новаторськими:

1. Гібридна модель із глибинними підсиленнями: Додавання елементів глибинного підсиленого навчання до BDI архітектури дозволить агенту навчатися оптимальним стратегіям у динамічних середовищах. Це дозволить поєднати переваги BDI моделі з потужними можливостями навчання з підкріпленням.

2. Інтеграція з графовими нейронними мережами (GNN): Використання графових нейронних мереж для моделювання зв'язків між переконаннями, бажаннями і намірами агента. Це дозволить більш гнучко представляти складні взаємозв'язки та покращити прийняття рішень.

3. Адаптивне обчислення на основі контексту: Запровадження механізмів адаптивного обчислення, що дозволять агентам динамічно змінювати свою стратегію в залежності від контексту. Використання рекурентних нейронних мереж для обробки контекстної інформації та коригування переконань, бажань і намірів у реальному часі.

4. Багатошарова рекурентна архітектура: Розробка багатошарової рекурентної архітектури, де кожен шар відповідає за окремий аспект процесу прийняття рішень. Наприклад, один шар може бути відповідальний за обробку переконань, інший – за формування бажань, а третій – за генерування намірів. Це дозволить створити більш модульний та масштабований підхід.

5. Інкorporація зовнішніх знань через трансферне навчання: Використання трансферного навчання для інтеграції зовнішніх знань з різних доменів. Це дозволить агенту бути більш універсальним та адаптивним, застосовуючи знання, здобуті в інших контекстах.

Інтеграція рекурентних нейронних мереж в архітектуру інтелектуальних агентів ПБН приносить численні переваги, що роблять їх більш потужними та ефективними. Перш за все, РНМ надають агентам здатність до запам'ятовування послідовних подій та виявлення тимчасових залежностей у даних, що є критично важливим для прийняття рішень в умовах невизначеності. Завдяки їхній структурі, вони можуть обробляти історичну інформацію та зберігати контекст минулих подій, що дозволяє агентам краще розуміти поточні ситуації і адаптувати свої дії відповідно до попередніх досліджень.

Друге, РНМ можуть значно покращити процес формування переконань в BDI агентах. Вони ефективно обробляють динамічні та послідовні дані, що робить їх ідеальними для аналізу інформації в режимі реального часу. Це дозволяє агентам швидко оновлювати свої переконання відповідно до змін у середовищі, забезпечуючи їм більшу гнучкість та точність у прийнятті рішень.

Також, інтеграція РНМ дозволяє агентам більш ефективно прогнозувати майбутні події та наслідки своїх дій. Здатність РНМ до моделювання часових рядів та прогнозування на основі минулих даних дозволяє ПБН агентам краще планувати свої дії та визначати найбільш оптимальні шляхи досягнення своїх цілей.

Крім того, використання РНМ сприяє підвищенню адаптивності агентів у складних та динамічних середовищах. Завдяки здатності до навчання та адаптації, РНМ можуть допомогти ПБН агентам швидко підлаштовувати свої стратегії відповідно до нових умов, що дозволяє їм більш ефективно реагувати на зміни та виклики.

Нарешті, інтеграція РНМ в ПБН архітектуру сприяє підвищенню загальної ефективності та точності агентів. Завдяки їхній здатності до обробки великих обсягів даних та виявлення прихованих патернів, РНМ можуть допомогти BDI агентам приймати більш інформовані та оптимізовані рішення, що покращує їхню продуктивність та здатність до виконання складних завдань. Інтеграція рекурентних нейронних мереж в архітектуру BDI надає агентам значну перевагу у здатності до навчання, адаптації та прийняття рішень у динамічних та невизначених середовищах.

Рекурентні нейронні мережі мають кілька ключових переваг порівняно з іншими архітектурами нейронних мереж, такими як перцептрони зворотного розповсюдження (Feedforward Neural Networks) або згорткові нейронні мережі (Convolutional Neural Networks).

Перш за все, РНМ здатні обробляти послідовності даних різної довжини, що робить їх особливо корисними для задач, пов'язаних з часовими рядами або текстовими послідовностями. Завдяки своїй рекурентній природі, вони можуть зберігати інформацію про попередні стани та враховувати контекст минулих подій при прийнятті рішень.

На відміну від перцептронів зворотного розповсюдження, які не мають механізму для збереження історичної інформації, РНМ мають петлі зворотного зв'язку, що дозволяють їм накопичувати та використовувати знання з попе-

редніх ітерацій. Це забезпечує більшу гнучкість та точність у задачах, де контекст і послідовність подій відіграють важливу роль.

Порівняно із згортковими нейронними мережами, які зазвичай використовуються для обробки статичних зображень або фіксованих вхідних розмірів, РНМ більш підходять для обробки динамічних та змінних даних. Хоча РНМ чудово справляються з розпізнаванням просторових патернів і характеристик, вони не мають тих же можливостей для аналізу тимчасових залежностей, що робить РНМ більш ефективними в таких задачах, як обробка тексту, прогнозування часових рядів або розпізнавання мовлення.

Крім того, розширені варіанти РНМ, такі як LSTM (Long Short-Term Memory) та GRU (Gated Recurrent Unit), долають проблеми затухання градієнтів, що дозволяє їм більш ефективно зберігати довготривалу пам'ять та навчатися на довших послідовностях. Це робить їх особливо потужними для задач, де важливо враховувати тривалі контекстні залежності.

Основними перевагами RNN є їх здатність обробляти послідовності даних різної довжини, зберігати історичну інформацію та враховувати контекст, що робить їх ідеальними для ряду застосувань, включаючи обробку природної мови, прогнозування часових рядів та розпізнавання мовлення.

Висновки

Інтелектуальні агенти є ключовим елементом у багатьох сучасних технологіях, що забезпечує автоматизацію складних завдань. З їх допомогою можна розв'язувати проблеми в таких сферах, як управління ресурсами, робототехніка, аналіз даних і персоналізоване обслуговування.

Науковий огляд сучасного стану проектування інтелектуальних агентів показав значний прогрес у цій галузі, підкреслюючи важливість адаптивних і гнучких підходів до розробки агентів. Аналіз різних архітектур та методів продемонстрував переваги та недоліки кожного з підходів, що сприяло глибшому розумінню механізмів прийняття рішень та здатності до навчання в динамічних середовищах.

Запропонована інтеграція рекурентних нейронних мереж у архітектуру інтелектуальних агентів структури ПБН відкриває нові перспективи у розробці агентів, здатних ефективно реагувати на змінні умови та виконувати складні завдання. Використання РНМ дозволяє агентам враховувати тимчасові залежності та історичну інформацію, що підвищує точність і ада-

птивність їхніх дій. Це сприяє більш ефективному оновленню вірувань та прогнозуванню майбутніх подій, що є критично важливим для досягнення цілей агентів.

Аналіз показав, що використання РНМ у ПБН агентах значно покращує їхню здатність до обробки динамічних даних та збереження контексту, що дозволяє швидко адаптувати стратегії до нових умов. Це відкриває нові можливості для застосування таких агентів у різних сферах, включаючи кібербезпеку, управління енергією та багато інших. Впровадження РНМ дозволяє підвищити загальну ефективність та точність агентів, що сприяє більш оптимізованому прийняттю рішень та виконанню складних завдань.

Інтеграція рекурентних нейронних мереж в архітектуру інтелектуальних агентів BDI має значний потенціал для розвитку та вдосконалення інтелектуальних систем, здатних ефективно вирішувати складні завдання в умовах невизначеності та динамічних змін. Це робить запропонований підхід перспективним напрямком для подальших досліджень і розробок у галузі штучного інтелекту. Запропоновані в статті універсальна архітектура та математична модель відкривають нові можливості для розвитку інтелектуальних агентів, здатних ефективно адаптуватися до змінних умов, масштабуватися та інтегрувати різні підходи. Ці рішення сприяють підвищенню продуктивності та ефективності інтелектуальних систем, що дозволить їх широко застосовувати в різних галузях – від промисловості до обслуговування клієнтів. У кінцевому рахунку, впровадження цих рішень сприятиме створенню більш інтелектуальних, гнучких та надійних систем, здатних задовольнити потреби сучасного світу.

Напрями подальших досліджень:

1. Варто дослідити питання оптимізації архітектури РНМ для специфічних задач ПБН агентів. Вивчення різних варіантів рекурентних нейронних мереж, таких як LSTM (Long Short-Term Memory) та GRU (Gated Recurrent Unit), а також їх параметризації та налаштувань, може дозволити досягти більш високої продуктивності та точності.

2. Цікавою областю досліджень є розробка гібридних моделей, що поєднують РНМ з іншими типами нейронних мереж, такими як згорткові нейронні мережі або графові нейронні мережі. Це може дозволити агентам одночасно обробляти як послідовні, так і просторові або мережеві дані, підвищуючи їхню здатність до аналізу складних ситуацій.

3. Розробка методів глибинного підсиленого навчання для ВПБН агентів з інтеграцією РНМ з акцентом уваги. Дослідження впливу різних стратегій навчання з підкріпленням на поведінку агентів може допомогти створити більш адаптивні та оптимізовані рішення для прийняття рішень у динамічних середовищах.

4. Варто звернути увагу на розробку / методів пояснюваного штучного інтелекту (Explainable AI) для ПБНІ агентів з РНМ. Розробка інструментів та методів, що дозволяють краще розуміти внутрішню логіку та процес прийняття рішень агентами, може підвищити їхню довіру та прийнятність у реальних додатках.

5. Важливим напрямком досліджень є застосування запропонованих моделей у реальних сферах, таких як кібербезпека, управління енергією, охорона здоров'я та інші. Вивчення ефективності та результативності інтеграції РНМ у ПБН архітектуру в різних контекстах допоможе визначити найкращі практики та напрямки для подальших удосконалень.

Література

1. Padgham L., Winikoff M. Developing Intelligent Agent Systems. A Practical Guide. Wiley, 2004. P. 225. DOI: [10.1002/0470861223](https://doi.org/10.1002/0470861223)
2. Saadi A., Maamri R., Sahnoun Z. Behavioral flexibility in Belief-Desire- Intention (BDI) architectures. *Multiagent and Grid Systems*. 2020. No 16(4). P. 343-377. DOI: <https://doi.org/10.3233/MGS-200335>
3. De Silva L. Meneguzzi F., Logan B. BDI Agent Architectures: A Survey, Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence Survey track. 2020, P. 4914–4921. DOI: <https://doi.org/10.24963/ijcai.2020/684>
4. Ekinici E., Halaç T., Erdur C., Çetin Ö., Cakirlar I., Dikenelli O. Satisfying agent goals by executing different task semantics: HTN, OWL-S or plug one yourself. *Autonomous Agents and Multi-Agent Systems*. 2013. No 26(2). DOI: <https://doi.org/10.1007/s10458-011-9185-2>
5. Singh D., Sardina S., Padgham L., James G. Integrating Learning into a BDI Agent for Environments with Changing Dynamics. *International Joint Conference on Artificial Intelligence*. 2011. P. 2525–2530. DOI: <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-420>
6. Zhang H., Huang S. Y. A general framework for parallel BDI agents in dynamic environments. *Web Intelligence and Agent Systems Journal*. 2008. No 6(3). P. 327–351. DOI: <https://doi.org/10.1109/IAT.2006.8>
7. Germano R., Lakhmi C. J. Intelligent Agents. Theory and Applications. Springer Berlin, Heidelberg. 2004. P. 402. DOI: <https://doi.org/10.1007/978-3-540-44401-5>
8. Khosla R., Dillon T. Engineering Intelligent Hybrid Multi-Agent Systems. Springer New York. 1997. P. 410. DOI: <https://doi.org/10.1007/978-1-4615-6223-8>
9. Bryson, J. Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental & Theoretical Artificial Intelligence*. 2000. No 12(2). P. 165–189. DOI: <https://doi.org/10.1080/095281300409829>
10. Rumbell T., Barnden J., Denham S., Wennekers T. Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Auton Agent Multi-Agent Syst*. 2012. No 25. P. 1-45. DOI: <https://doi.org/10.1007/s10458-011-9166-5>
11. Cruz A., dos Santos A. V., Santiago R. H. N., Bedregal B. A Fuzzy Semantic for BDI Logic. *Fuzzy Information and Engineering*. 2021. No 13(2). P. 139-153. DOI: <https://doi.org/10.1080/16168658.2021.1915455>
12. Calegari R., Ciatto G., Mascardi, V., Omicini A. Logic-based technologies for multi-agent systems: a systematic literature review. *Auton Agent Multi-Agent Syst*. 2021. No 35(1). <https://doi.org/10.1007/s10458-020-09478-3>
13. Круглик В. С., Прокоф'єв Є. Г., Маринов, А. В. Аналіз можливостей використання інтелектуальних агентів в адаптивній системі електронного навчання. *Педагогічні науки: теорія та практика*. 2021. № 4. С. 295-302. URL: <https://doi.org/10.26661/2786-5622-2021-4-44>
14. Лопатто І. Ю., Говорущенко Т. О., Капустян М. В. Інтелектуальний агент верифікації врахування інформації предметної галузі в процесі розроблення програмних систем. *Вісник Хмельницького національного університету*. 2022. № 1. С. 116-119. URL: <http://journals.khnu.km.ua/vestnik/?p=12131>
15. Маринов А. В., Круглик В. С. Використання інтелектуальних програмних агентів для створення адаптивного середовища електронного навчання на базі lms moodle. *Міжнародна науково-практична конференція «Цифрова трансформація та диджитал технології для сталого розвитку всіх галузей сучасної освіти, науки і практики»*. Zbiór prac_Tom 2. 2023. С. 306-308. URL: https://repo.btu.kharkov.ua/bitstream/123456789/29446/1/zbior_prac_tom_2_26012023-306-308.pdf
16. Noulamo T., Djimeli-Tsajio A., Kameugne R., Lienou, J. A Generic Intelligent Agent Design Approach Based on Artificial Neural

Networks. *World Journal of Engineering and Technology*. 2023. No 11. P. 682-697. URL: [10.4236/wjet.2023.114046](https://doi.org/10.4236/wjet.2023.114046)

17. Christie S. H., Chopra A. K., Singh M. P. Mandrake: Multiagent Systems as a Basis for Programming Fault-Tolerant Decentralized Applications. *Autonomous Agents and Multi-Agent Systems*. 2022. No 36, A. No 16. DOI: <https://doi.org/10.1007/s10458-021-09540-8>

18. Latham N. Types of intelligent agent, 2024. URL: <https://www.probecx.com/en-au/blog/types-of-intelligent-agent>

19. Intelligent Agent, 2023. URL: <https://www.larksuite.com/en-us/topics/ai-glossary/intelligent-agent>

20. Harjyot K. What are AI Agents: Types, Benefits, Applications, and Examples. 2024. URL: <https://www.signitysolutions.com/blog/ai-agents>

21. Hiren Dhaduk. What is an AI Agent? Characteristics, Advantages, Challenges, Applications. 2023. URL: <https://www.simform.com/blog/ai-agent/>

22. Крихівський М. В., Крихівська С. М. Проблеми людино-машинної взаємодії в контексті штучного інтелекту. *Збірник тез доповідей науково-практичної конференції «Інформаційні технології в освіті, техніці та промисловості»*. 2024. С. 176-177. URL: <https://stlnau.in.ua/samoosvita/item/2024/iit241010.pdf>

References

1. Padgham L., Winikoff M. Developing Intelligent Agent Systems. A Practical Guide. Wiley, 2004. P. 225. DOI: [10.1002/0470861223](https://doi.org/10.1002/0470861223)

2. Saadi A., Maamri R., Sahnoun Z. Behavioral flexibility in Belief-Desire- Intention (BDI) architectures. *Multiagent and Grid Systems*. 2020. No 16(4). P. 343-377. DOI: <https://doi.org/10.3233/MGS-200335>

3. De Silva L. Meneguzzi F., Logan B. BDI Agent Architectures: A Survey, Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence Survey track. 2020, P. 4914–4921. DOI: <https://doi.org/10.24963/ijcai.2020/684>

4. Ekinci E., Halaç T., Erdur C., Çetin Ö., Cakirlar I., Dikenelli O. Satisfying agent goals by executing different task semantics: HTN, OWL-S or plug one yourself. *Autonomous Agents and Multi-Agent Systems*. 2013. No 26(2). DOI: <https://doi.org/10.1007/s10458-011-9185-2>

5. Singh D., Sardina S., Padgham L., James G. Integrating Learning into a BDI Agent for Environments with Changing Dynamics. *International Joint Conference on Artificial Intelligence*. 2011.

P. 2525–2530. DOI: <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-420>

6. Zhang H., Huang S. Y. A general framework for parallel BDI agents in dynamic environments. *Web Intelligence and Agent Systems Journal*. 2008. No 6(3). P. 327–351. DOI: <https://doi.org/10.1109/IAT.2006.8>

7. Germano R., Lakhmi C. J. Intelligent Agents. Theory and Applications. Springer Berlin, Heidelberg. 2004. P. 402. DOI: <https://doi.org/10.1007/978-3-540-44401-5>

8. Khosla R., Dillon T. Engineering Intelligent Hybrid Multi-Agent Systems. Springer New York. 1997. P. 410. DOI: <https://doi.org/10.1007/978-1-4615-6223-8>

9. Bryson, J. Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental & Theoretical Artificial Intelligence*. 2000. No 12(2). P. 165–189. DOI: <https://doi.org/10.1080/095281300409829>

10. Rumbell T., Barnden J., Denham S., Wennekers T. Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Auton Agent Multi-Agent Syst*. 2012. No 25. P. 1-45. DOI: <https://doi.org/10.1007/s10458-011-9166-5>

11. Cruz A., dos Santos A. V., Santiago R. H. N., Bedregal B. A Fuzzy Semantic for BDI Logic. *Fuzzy Information and Engineering*. 2021. No 13(2). P. 139-153. DOI: <https://doi.org/10.1080/16168658.2021.1915455>

12. Calegari R., Ciatto G., Mascardi, V., Omicini A. Logic-based technologies for multi-agent systems: a systematic literature review. *Auton Agent Multi-Agent Syst*. 2021. No 35(1). DOI: <https://doi.org/10.1007/s10458-020-09478-3>

13. Kruhlyk V. S., Prokofiev Ye. H., Marynov A. V. Analiz mozhlyvosti vykorystannia intelektualnykh ahentiv v adaptivnii systemi elektronnoho navchannia. *Pedahohichni nauky: teoriia ta praktyka*. 2021. No 4. P. 295-302. URL: <https://doi.org/10.26661/2786-5622-2021-4-44> [in Ukrainian]

14. Lopatto I. Yu., Hovorushchenko T. O., Kapustian M. V. Intelektualnyi ahent veryfikatsii vrakhuvannia informatsii predmetnoi haluzi v protsesi rozroblennia prohramnykh system. *Visnyk Khmelnytskoho natsionalnoho universytetu*, 2022. No 1. P. 116-119. URL: <http://journals.khnu.km.ua/vestnik/?p=12131> [in Ukrainian]

15. Marynov A.V., Kruhlyk V.S., Vykorystannia intelektualnykh prohramnykh ahentiv dlia stvorennia adaptivnoho seredovyshcha elektronnoho navchannia na bazi lms moodle. *Mizhnarodna nauково-praktychna konferents*

«Tsyfrova transformatsiia ta dydzhytal tekhnolohii dlia staloho rozvytku vsikh haluzei suchasnoi osvity, nauky i praktyky». Zbiór prac_Tom 2. 2023. P. 306-308. URL: https://repo.btu.kharkov.ua/bitstream/123456789/29446/1/zbior_prac_tom_2_26012023-306-308.pdf [in Ukrainian]

16. Noulamo T., Djimeli-Tsajio A., Kameugne R., Lienou, J. A Generic Intelligent Agent Design Approach Based on Artificial Neural Networks. *World Journal of Engineering and Technology*. 2023. No 11. P. 682-697. URL: [10.4236/wjet.2023.114046](https://doi.org/10.4236/wjet.2023.114046)

17. Christie S.H., Chopra A. K., Singh M. P. Mandrake: Multiagent Systems as a Basis for Programming Fault-Tolerant Decentralized Applications. *Autonomous Agents and Multi-Agent Systems*. 2022. No 36, A. No 16. DOI: <https://doi.org/10.1007/s10458-021-09540-8>

18. Latham N. Types of intelligent agent, 2024. URL: <https://www.probecx.com/en-au/blog/types-of-intelligent-agent>

19. Intelligent Agent, 2023. URL: https://www.larksuite.com/en_us/topics/ai-glossary/intelligent-agent

20. Harjyot K. What are AI Agents: Types, Benefits, Applications, and Examples. 2024. URL: <https://www.signitysolutions.com/blog/ai-agents>

21. Hiren Dhaduk. What is an AI Agent? Characteristics, Advantages, Challenges, Applications. 2023. URL: <https://www.simform.com/blog/ai-agent/>

22. Krykhivskyi M. V., Krykhivska S. M. Problemy liudyno-mashynnoi vzaiemodii v konteksti shtuchnoho intelektu. *Zbirnyk tez dopovidei naukovo-praktychnoi konferentsii «Informatsiini tekhnologii v osviti, tekhnitsi ta promyslovosti»*. 2024. P. 176-177. URL: <https://stlnau.in.ua/samoosvita/item/2024/iit241010.pdf> [in Ukrainian]