

been little studied in the scientific literature, which is associated with the unpredictability of cyberattacks and the absence in many cases of real data, as well as available methods for predicting them.

Mathematical modeling of time series of the intensity of cyberattacks per enterprise is presented to provide comprehensive solutions and predictions of strengthening the enterprise's resistance against current targeted cyber threats. We consider a first-order nonlinear differential equation, the Bernoulli equation, which describes the process of the time series of the intensity of cyberattacks. The analysis of the intensity function of cyberattacks is carried out analytically due to the power-law  $p$ -transformation by the analytical function. Statistical data on the number of cyberattacks at the enterprise are considered, provided that a scheduled audit is carried out once a quarter. The types of cyberattacks to defeat network infrastructure, proprietary applications, the level of patches and server configurations, standard software, and their number at the enterprise for certain time periods are presented. A geometric visualization of the change in the steepness of the logistic curve of the intensity of cyberattacks is presented at various parameter values with a uniform step for the period between scheduled audits when applying  $p$ -conversion.

**Keywords:** cyber security, cyberattack intensity, Bernoulli equation, defeat, logistic curve.

УДК 004.85

к.т.н., доц. **Бойчук В.О.** (ХМНУ)

к.е.н., доц. **Бойчук А.А.** (ТНЕУ)

**Бойчук М.В.** (ХМНУ)

**Бурдюг О.В.** (ВІКНУ)

DOI: <https://doi.org/10.17721/2519-481X/2020/66-07>

## МЕТОД ФОРМУВАННЯ ПОСЛІДОВНОСТІ ДІЙ ІНТЕЛЕКТУАЛЬНИХ АГЕНТІВ

У статті запропоновано підхід, де реалізація формування послідовностей дій інтелектуальних агентів виконується по аналогії з діяльністю біологічних організмів з використанням механізму емоцій для динамічного налаштування організму на виконання дій. Таким чином імітуються функції лімбічної системи в організації рухів на основі мотиваційної поведінки. При плануванні в першу чергу визначається загальний стан агенту. Використовуючи отриманий стан формується послідовність дій. Такий підхід дасть можливість динамічно переналаштувати послідовність і реагувати на небезпечну ситуацію або на зміну внутрішнього стану агенту.

Інтелектуальний агент отримує з сенсорів і рецепторів ознаки початкової умови по ній визначається ціль та формується послідовність дій. Елементами послідовності дій є елементарні дії. Елементарна дія характеризується набором вхідних параметрів для функціонування. Ознаки передумови відповідають першій дії в послідовності, остання дія в послідовності прив'язана до ознаки цілі.

Послідовність дій агенту представляється оргграфом, де вершини визначають елементарні дії, а ребра визначають ступінь сили зв'язку між ними. Початкові умови відповідають першій дії в послідовності, з неї розпочинається реалізація послідовності дій. Ознаки цілі відповідають останній вершині в послідовності дій

Ваги зв'язків змінюються при встановленні змінних загального стану, що дає змогу виконати послідовність дій в реальному масштабі часу з динамічним переналаштуванням і вибрати серед характерних для конкретного стану послідовностей дій. Метод формує послідовність дій, яка ініціюється емоційними станами, і переводить її в послідовність автоматичних дій на основі досягнення цілі і яка в майбутньому буде виконуватись в нормальному стані. Для перевірки функціонування методу реалізований симулятор агенту-роботу в середовищі програми V-REP. Отримані результати можуть бути використані для інтелектуального планування на основі підкріплення при керуванні агентами, роботами на виробничих підприємствах, військовими агентами, потоками міського руху, логістичними системами, соціальними явищами.

**Ключові слова:** інтелектуальний агент, планування, модель,  $Q$ -навчання, емоційні стани, навчання з підкріпленням.

**Вступ.** Планування - це процес генерації уявлень про майбутню поведінку до того, як отримані таким чином плани будуть використані для реалізації цієї поведінки. Результатом планування зазвичай є деяка множина дій, а також накладені на них часові та інші обмеження і передбачається, що ці дії будуть виконуватися будь-яким агентом або агентами.

Більшість підходів, що застосовуються в штучному інтелекті при плануванні, не прагнуть будувати функціональні моделі людського мозку або його відділів. Рішення, пропонувані теоріями штучного інтелекту, будують моделі, поведінка яких схожа з поведінкою людини. Але внутрішній склад цих моделей, як правило, не є моделлю нервової тканини або процесів, що протікають в нервовій тканині живої істоти. Тобто штучний інтелект хоч і займається відтворенням окремих функцій творчої діяльності людини, але його методи кардинально відрізняються від природного, біологічного протікання інтелектуальних процесів.

Як правило інтелектуальне планування засноване на апараті математичної логіки, а міркування, необхідні для формування плану, зводяться до логічного висновку. Таким чином, логічний висновок, в даному випадку, і буде моделлю міркувань. Традиційне планування не може бути безпосередньо застосовано до задач реального світу, оскільки проблемний простір пошуку занадто великий. А між тим біологічні істоти за рахунок навчання з підкріпленням можуть планувати свої дії з достатньою для вживання ефективністю. Запозичення таких ідей планування повинно підвищити ефективність функціонування інтелектуальних агентів.

**Аналіз останніх досліджень і публікацій.** Якщо розглянути планування в природі, то вищі біологічні істоти не навчаються з нуля, але використовують невеликий набір простих моделей поведінки. Ці поведінки низького рівня можуть порівняно легко спроектовані або засвоєні, але задача координації цих поведінок досить складна. Для перемикання і запам'ятовування(підкріплення) поведінок використовується механізм емоції.

Розглянемо приклади реалізації навчання на основі емоцій з підкріпленням при плануванні дій інтелектуальних агентів.

Архітектура EB [1] заснована на емоціях, в якій традиційна адаптивна система навчання доповнюється системою емоцій, яка відповідальна за навчання та поведінку. Агент має деякі вроджені емоції, що визначають його цілі, і він потім засвоює емоційні асоціації середовища-стану, які визначають його рішення. Агент використовує алгоритм Q-навчання для навчання вибору поведінки під час взаємодії зі своїм оточенням.

Архітектура EB II [2] - складається з двох основних систем: цільової системи (GS) та адаптивної системи (AS). Цільова системи оцінює ефективність роботи адаптивної системи з точки зору гомеостатичних змінних і визначає, коли поведінку слід перервати. Обчислюється так зване значення «благополуччя», яке використовується як підкріплення адаптивною системою та визначає кроки запуску. Адаптивна система дізнається, яку поведінку вибрати на кроках запуску, використовуючи методи навчання з підкріпленням, спираючись на нейронні мережі для зберігання значень корисності.

Сукупність перцептивних значень та внутрішніх значень використовувались при обчисленні одного мультимірного емоційного стану. Цей стан, у свою чергу, використовувався для визначення навчання на кожному часовому кроці та значні відмінності у його значенні вважались релевантними подіями, що використовуються для запуску механізм вибору поведінки.

Цілі чітко визначені і пов'язані з гомеостатичними змінними. Ці гомеостатичні змінні асоціюються з трьома різними станами: цільовим (Target), відновлення(Recovery) та небезпечним (Dangerous). Стан кожної змінної залежить від її безперервного значення, яке групується на чотирьох якісних категоріях: оптимальний (Optimal), прийнятний (Acceptable), дефіцитний (Deficient) і небезпечний(Danger). Змінна залишається в цільовому стані до тих пір, поки є її значення оптимальні або прийнятні, але вона повертається до свого цільового стану лише після того, як його значення знову стануть оптимальними.

Архітектура ALEC (Asynchronous Learning by Emotion and Cognition)[3], яка має на меті покращити ефективність навчання шляхом розширення архітектури EB когнітивною системою, що доповнює свої поточні можливості адаптації на основі емоцій правилами, які створюються з взаємодії агент-середовище. Різні можливості навчання двох системи та їх взаємодія можуть створити більш потужну адаптивну систему.

В іншій системі планування дій агентів [4] формується набір поведінкових правил, представлений в MYCIN подібній формі, тобто в формі продукцій з коефіцієнтами впевненості. Вплив емоцій на вчинення дії реалізується як позитивний зворотний зв'язок між вихідним сигналом (поточна дія) і поведінковими правилами.

Основна система управління робота оперує тільки термінами поведінкових дій. Ця система управління не використовує таких дій, як "Повернути ліворуч або праворуч", "Рухатися вперед або назад" і т.п. Всі ці дії перенесені на нижній рівень управління. Основна система управління застосовує складні поведінкові процедури: пошуку їжі, процедури сну (відпочинку) і т.д.

Висновок проходить через блок "Збудження/Гальмування" для активації відповідних про процесів збудження і гальмування. Параметр збудження реалізований через аналог штучного нейрона. Цей елемент бере вхідні сигнали (від давачів і блоку "Потреби"), підсумовує їх і передає на вихід. На відміну від стандартного нейрона тут не використовуються вагові коефіцієнти (зважені входи), а застосовується вага для всіх входів. Значення цієї ваги є параметром збудження.

При використанні конекціоністського підходу в алгоритмі Q-Learning табличне представлення Q-функції замінюється нейронною мережею [5]. На входи мережі подаються стани, а вихідними даними є оцінки Q-значень. Таким чином, ніяких серйозних змін в класичний Q-Learning не вноситься, просто змінюється засіб зберігання оцінок Q-значень. Використовується методика роботи з нейронною мережею, запропонована Ліном, яка полягає в застосуванні окремої нейронної мережі для кожної дії.

На кожній ітерації роботи алгоритму поточний стан системи подається на входи кожної нейронної мережі, однак оновлення ваг здійснюється тільки для тієї нейронної мережі, дія якої була вибрана. Прокручування списку стан-дія в зворотному порядку дозволяє виробляти навчання на більш правильних оцінках.

Прокручування списку стан – дія в зворотному порядку дозволяє розробляти процес навчання на більш правильних оцінках. Однак послідовність кроків, яку виконує система, може виявитися неоптимальною і, отже, оцінки на яких буде проводитися навчання, також виявляються неоптимальними.

Якщо розглянути недоліки цих методів, то можна сказати, що хоча вони й використовують навчання підкріпленням, однак не призначені для динамічного, гнучкого планування і навчання послідовності дій.

**Постановка задачі.** Для розв'язку задачі планування діяльності агенту необхідно здійснити послідовність дій при яких середовище перейде з початкового стану у потрібний цільовий стан.

У природному середовищі організми виконують ідентифікацію та класифікацію початкових умов і цілі керування використовуючи нейронні мережі на основі параметрів зовнішнього середовища і внутрішнього стану організму. Послідовність дій аналогічно реалізується з використанням нейронних мереж головного, спинного мозку і т.д.

Відповідно вищесказаному реалізацію планування дій можна виконувати по аналогії з діяльністю біологічних організмів з використанням механізму емоцій для динамічного налаштування організму на виконання дій.

**Основна частина.** Регулювання емоцій здійснюється лімбічною системою, яка складається з старих відділів переднього мозку.

Таким чином ми будемо імітувати функції лімбічної системи в організації рухів на основі мотиваційної поведінки.

Назвемо аналоги емоцій загальними станам агента. Вони визначаються множиною змінних  $y_1, \dots, y_k$ , від 0 до 1, де значення кожної змінної відображає інтенсивність стану.

При плануванні в першу чергу визначається загальний стан агента. Використовуючи отриманий стан формується послідовність дій. Такий підхід дасть можливість динамічно переналаштовувати послідовність і реагувати на небезпечну ситуацію або на зміну внутрішнього стану агента.

У випадку, якщо інтенсивність деякої змінної загального стану вище деякого порогу, агент може виконувати тільки стандартні стереотипні дії, які характерні для цього стану. Якщо ж інтенсивність змінної загального стану менше цього порога, то може бути вибрана інша послідовність дій. У випадку отримання інших параметрів загального стану системи дії агента можуть динамічно змінюватись.

Дані в агент надходять з зовнішніх сенсорів і з внутрішніх рецепторів. Для планування дій і наступної обробки ці дані різних типів переводяться в символний та кількісний вигляд. Дану функцію реалізують компоненти обробки сигналів з сенсорів і рецепторів. Фізично даними компонентами можуть бути і нейронні мережі різних типів і компоненти з функціями виконання обчислень і перетворень над кількісними даними.

Дані компоненти отримують набір кількісних значень  $x_1, \dots, x_n$  по яких можна обчислити стан агента і використати при формуванні послідовності дій.

Загальна структура агента по запропонованому підходу зображена на рис. 1.

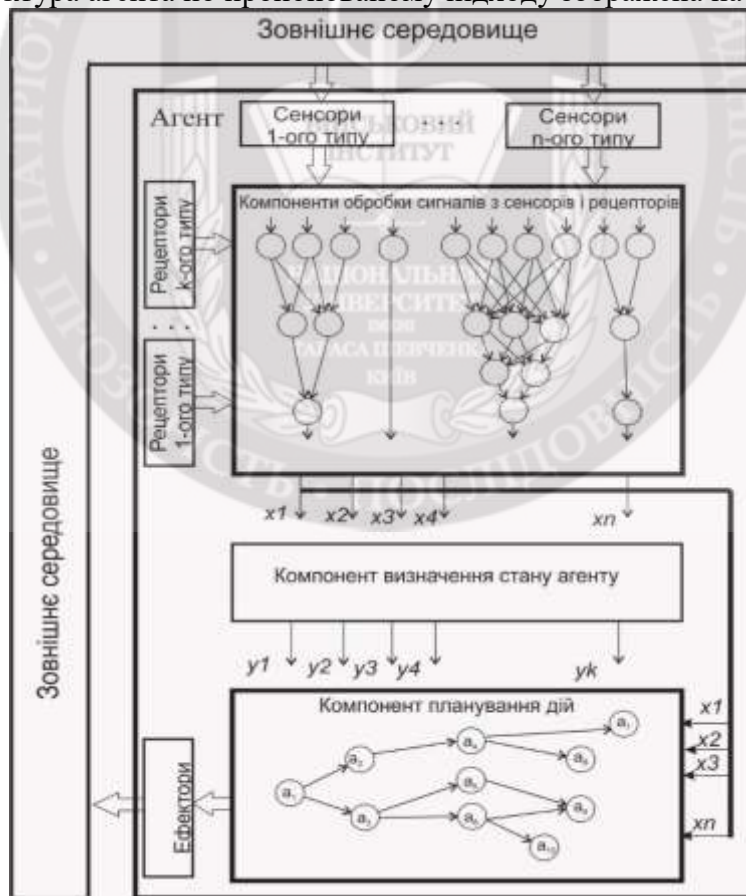


Рисунок 1 – Загальна структура агента

Запропонуємо основні положення моделі представлення послідовності дій на базі представленого вище підходу

Інтелектуальний агент отримує з сенсорів і рецепторів ознаки початкової умови, по ній визначається ціль та формується послідовність дій. Елементами послідовності дій є елементарні дії  $a_1, a_2, a_3, \dots, a_n$  для використання агентом. Елементарна дія характеризується набором вхідних параметрів для функціонування  $x_{11}, x_{12}, \dots, x_{1m}$ .

Ознаки передумови  $q_i$  відповідають першій дії в послідовності, остання дія в послідовності прив'язана до ознаки цілі  $g_i$ .

Відповідно вищесказаному необхідно скомпонувати з елементарних дій їх стабільні послідовності, які б відповідали початковим умовам і приводили до деякої цілі.

Візьмемо в якості прикладу рухи деякого агента-робота, який переміщається по площині. Елементарними діями робота є поворот, рух вперед, назад. Параметрами дій можуть бути визначені швидкість руху, кут повороту та ін. Якщо в напрямку робота швидко переміщається великий об'єкт, то робот-агент має класифікувати ознаки великого об'єкту і швидкого руху в напрямку агента. Робот має перейти в загальний стан страху. На основі цього стану він повинен спочатку зупинитися, а потім розвернутися і максимально швидко рухатись від джерела небезпеки. Такі початкові умови мають генерувати негайну послідовність дій, яка з людської точки зору викликається такою емоцією, як страх. Ціль цієї послідовності дій є відсутність великого об'єкту в полі зору.

За аналогією з біологічними системами з множин впорядкованих елементарних дій складаються стандартні послідовності дій для рефлекторних реакцій на визначені початкові умови.

Модель представлення послідовності дій для навчання з підкріпленням:

$$L=(Q,G,X,W,A),$$

де  $q_i \in Q$  - ознаки початкового стану,  $g_i \in G$  - ознаки цілі дій,  $a_i \in A$  - елементарна дія,  $x_{ij} \in X$  - вхідні параметри елементарних дій,  $w_{ij} \in W$  - вага переходу від одної елементарної дії до іншої.

Представимо послідовність дій агенту направленим графом, де вершини визначають елементарні дії, а ребра визначають ступінь сили зв'язку між ними. Початкові умови  $q$  відповідають першій дії в послідовності, з неї розпочинається реалізація послідовності дій. Ознаки цілі  $g$  відповідають останній вершині в послідовності дій.

На рис. 2.3 показаний приклад графу послідовності дій.

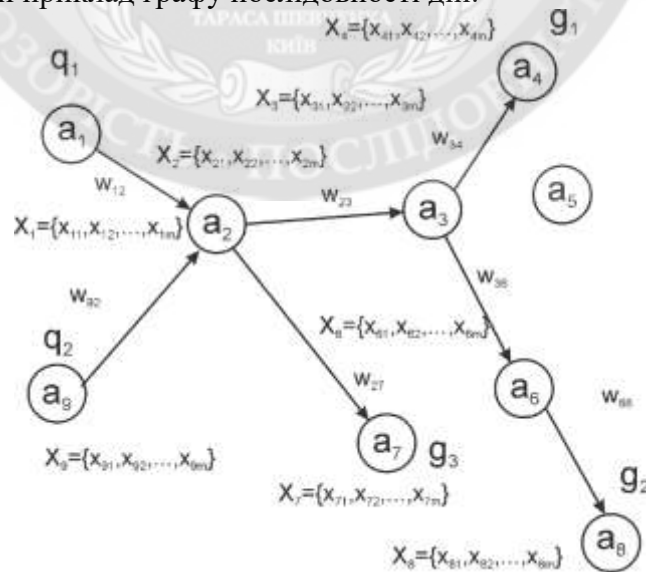


Рисунок 2 – Граф послідовності дій

З вершини в загальному випадку може бути декілька варіантів переходів до інших вершин, ваги всіх ребер, які виходять з однієї вершини в сумі повинні дорівнювати одиниці.

У випадку, коли вага ребра більше деякого порогового значення  $w_{\text{thres\_max}}$ , то вибирається наступна елементарна дія, у яку входить це ребро.

У випадку, коли вага усіх ребер менше цього порогу  $w_{\text{thres\_max}}$ , то агент аналізує поточний стан і реалізує перехід згідно ребер, які більше деякого порога  $w_{\text{thres\_min}}$ .

Якщо існує такий шлях від ідентифікації ознак початкових умов до ознак цілі, в якому всі послідовні ребра мають вагу більше  $w_{\text{thres\_max}}$ , то послідовність дій по цьому шляху виконується автоматично.

Агент початково налаштований на стандартні дії після ідентифікації початкових умов і отримання загального стану.

Ваги зв'язків змінюються при встановленні змінних загального стану, що дає змогу виконати послідовність дій в реальному масштабі часу з динамічним переналаштуванням і вибрати серед характерних для конкретного стану послідовностей дій.

Функція залежності ваги зв'язку від змінних загального стану має наступний вигляд:

$$w_{ij}=f(y_1, \dots, y_n),$$

де  $y_1, \dots, y_n$  - значення змінних загального стану 1..n;  $k_1, \dots, k_n$  - коефіцієнти для позначення значення  $y_1, \dots, y_n$  для відповідного ребра,  $0 < k_i < 1$ .

Для прикладу:

$$w_{ij}=k_1y_1+k_2y_2+\dots+k_ny_n=\sum_{i=1}^n k_i \cdot y_i,$$

лінійний варіант функції залежності ваги зв'язку від змінних загального стану.

Відповідно кожній дузі ставиться у відповідність набір параметрів змінних загального стану  $y_1, \dots, y_n$  зі значення інтенсивностей станів(емоцій) 1..n;  $k_1, \dots, k_n$  - коефіцієнти для позначення важливості  $y_1, \dots, y_n$  для даного ребра

Модель розроблена для імітації виконання дій в біологічних організмах і коефіцієнти  $k_i$  у загальному можуть відповідати різним нейромедіаторам в синапсах між нейронами головного мозку і можуть змінюватись при навчанні. Успішне проходження шляху від початкових умов до цілі по шляху впливає на ваги ребер цього шляху. Тобто змінюються коефіцієнти  $k$  вздовж шляху від початкової умови до цілі при успішному її досягненні.

Відповідно, метод формування послідовності дій виходить з уявлень про формування дій в біологічних організмів, а саме:

1. Існує множина станів агента (емоційних) станів  $y_1, \dots, y_n$ . Кожен стан може мати значення від 0 (повна відсутність) до 1 (найбільш виражений, який пригнічує всі інші емоційні стани).

2. Змінні, які характеризують емоційні стани, мають початкові значення, які характерні для агента.

3. Емоційні стани встановлюються на деякий наперед визначений час, по спливанні якого вони вертаються до початкового значення.

4. Існує початковий нормальний стан з деяким значенням в якому агент перебуває по замовчуванню.

5. Задається множина стандартних дій, які може виконувати агент  $a_1, \dots, a_n$ .

6. Виконання кожної дії обумовлюється набором параметрів на вході дії  $x_1, \dots, x_n$ .

7. Існує початкова множина значень сенсорних і рецепторних входів, які встановлюють визначені емоційні стани. Множина може збільшуватись при навчанні.

8. Для кожного стану агента вище деякого порога існують прості стереотипні дії, які агент обов'язково виконує.

9. Ймовірність вибору дій для виконання може змінюватись в залежності від встановлених станів.

10. Стани успіху і невдачі встановлюються при співпадінні/неспівпадінні поточних значень сенсорних і рецепторних входів з ознаками цілі і використовуються для підкріплення послідовності дій.



Тобто, мета методу сформуванати послідовність дій, яка ініціюється емоційним станом, і перевести в послідовність автоматичних дій на основі досягнення цілі і яка в майбутньому буде виконуватись в нормальному стані .

На рис. 3 показаний можливий граф дій робота-агенту, який рухається по площині, що обмежена стінами. Робот повинен навчитися уникати перешкоду (стіну) і повертати при цьому у відповідну сторону.

Робот має три ультразвукові сенсори, перший який визначає відстань до перешкоди по напрямку рухи і два інших, які визначають перешкоду праворуч або ліворуч під кутом 75 градусів.  $x_1$  - параметр, який задає відстань до перешкоди по напрямку руху,  $x_2$  – параметр на основі показників двох сенсорів, який приймає значення 0 або 1 в залежності де знаходиться перешкода, праворуч або ліворуч.

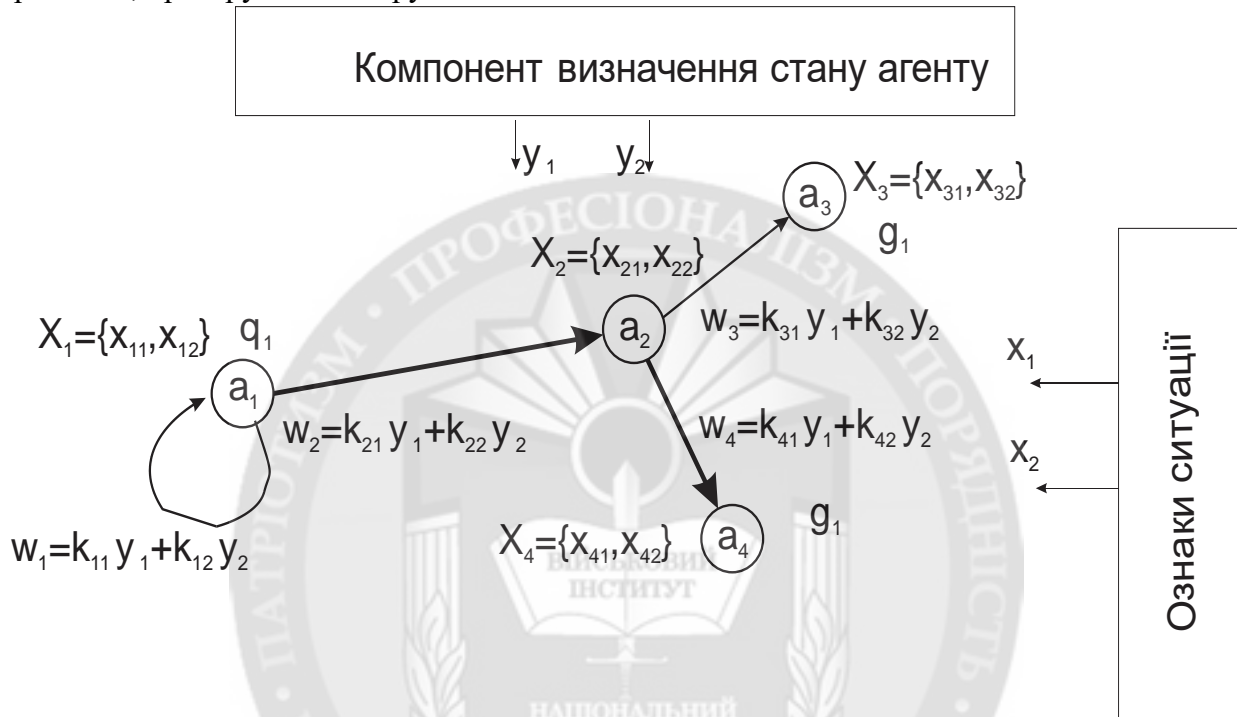


Рисунок 3 – Приклад графу моделі при виконання завдання уникання перешкоди

$y_1$  – показник рівня нормального стану, впливає на рівень функціонування агента при відсутності ідентифікованих емоційних станів

$y_2$  – показник рівня стану страху, росте при наближенні агента до перешкоди

$q_1$  – початкові умови, відстань до перешкоди менше порога

$g_1$  – ціль дій, відсутність перешкоди по ходу руху агента

$k_1$  – коефіцієнт рівня впливу нормального стану на дугу переходу

$k_2$  – коефіцієнт рівня впливу стану страху на дугу переходу

$a_1$  – рух робота по прямій

$a_2$  – зупинка агента

$a_3$  – поворот ліворуч  $90^\circ$

$a_4$  – поворот праворуч  $90^\circ$

$x_1$  – відстань до перешкоди

$x_2$  – напрямок знаходження перешкоди

$\Delta k$  – крок коректування коефіцієнтів  $k$  при навчанні

При наближенні до агента до перешкоди на близьку відстань менше деякого порогу зростає рівень страху  $y_2$ . Це є і початковими умовами для планування дій, цілю яких є відсутність перешкоди по ходу руху агента.

Спочатку вибирається стандартна дія при високому рівні страху - зупинка, потім інші дії з набору дій після виконання дій будується послідовність дій і перевіряється досягнення цілі. При досягненні цілі проводиться коректування коефіцієнтів по ланцюжку  $k_1 + \Delta k$ ,  $k_2 - \Delta k$ ,  $\Delta k$  –

коефіцієнт швидкості навчання. Тобто при частому успішному проходженні послідовності дій уникнення перешкоди буде викликатися не страхом, а нормальним станом. І виконуватись або в автоматичному режимі або з вибором в яку сторону доцільно повертати, щоб уникнути перешкоди, на основі аналізу  $x_1, x_2$ .

Для перевірки функціонування методу згідно задачі на рис. 3 на мові Lua написаний симулятор агента в середовищі програми V-REP для моделі робота dr20. Програма V-REP забезпечує точну симуляцію робота і дає розробнику простий фреймворк, щоб практикуватися в створенні ПЗ для роботів.

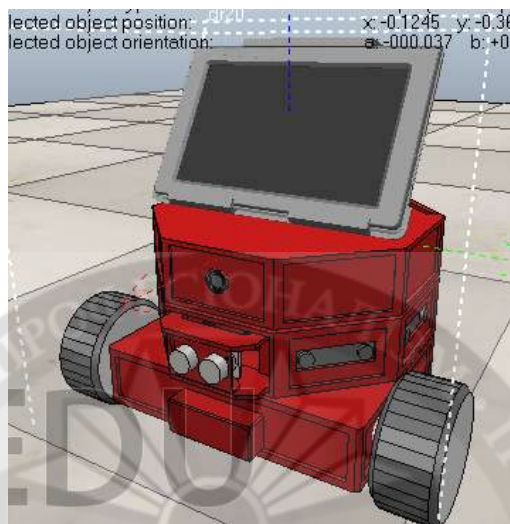


Рисунок 4 – Модель робота dr20 в середовищі V-REP

Початкове значення нормального стану вибрано рівним 0,6. Значення змінної страху для початку вибирається рівним 0. При досягненні успіху значення нормального стану зростає до 1, при невдачі зменшується до 0,2, його мінімального значення. Значення змінної стану страху збільшується до одиниці при зустрічі перешкоди. Стан страху падає з одиниці до нуля за виконання двох дій циклу моделювання.

Моделювання виконувалось на комп'ютері с процесором Intel Atom 570 с 2 Гігабайтами оперативної пам'яті з використанням операційної системи Linux Mint. Час навчання коливався від 5 до 10 хвилин з мінімумом при значенні  $\Delta k$  0,07.

**Висновки.** В статті запропоновано підхід, де реалізація формування послідовностей дій інтелектуальних агентів виконується по аналогії з діяльністю біологічних організмів з використанням механізму емоцій для динамічного налаштування організму на виконання дій. Таким чином імітуються функції лімбічної системи в організації рухів на основі мотиваційної поведінки.

Послідовність дій агенту представляється орграфом, де вершини визначають елементарні дії, а ребра визначають ступінь сили зв'язку між ними. Початкові умови відповідають першій дії в послідовності, з неї розпочинається реалізація послідовності дій. Ознаки цілі відповідають останній вершині в послідовності дій

Метод формує послідовність дій, яка ініціюється емоційними станами, і переводить її в послідовність автоматичних дій на основі досягнення цілі і яка в майбутньому буде виконуватись в нормальному стані.

Для перевірки функціонування методу реалізований симулятор агенту-роботу в середовищі програми V-REP. Отримані результати можуть бути використані для інтелектуального планування на основі підкріплення можуть бути використані при керуванні агентами, роботами на виробничих підприємствах, військовими агентами, потоками міського руху, логістичними системами, соціальними явищами.



#### ЛИТЕРАТУРА:

- 1 Sandra Clara Gadanho. Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions. PhD thesis, University of Edinburgh, 1999.
2. Sandra Clara Gadanho and John Hallam. Emotion-triggered learning in autonomous robot control. *Cybernetics and Systems — Special Issue: Grounding emotions in adaptive systems*, 32(5):531–559, July 2001.
3. Sun R., Peterson T. Learning in Reactive Sequential Decision Tasks: the CLARION Model / Proc. 1996 IEEE ICNN, Washington, DC, USA. Plenary, Panel and Special Sessions Volume. – pp.70–75.
4. Карпов В. Э. Эмоции и темперамент роботов поведенческие аспекты / В. Э. Карпов // Известия РАН. Теория и системы управления. – М.:, 2014, № 5, с. 166–185.
5. Кузьмин В. Использование нейронных сетей в алгоритме Q-learning / В. Кузьмин // Transport and Telecommunication. – Рига: Vol.4, N 1. С 75-86, 2003.

#### REFERENCES:

- 1 Sandra Clara Gadanho. Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions. PhD thesis, University of Edinburgh, 1999.
2. Sandra Clara Gadanho and John Hallam. Emotion-triggered learning in autonomous robot control. *Cybernetics and Systems – Special Issue: Grounding emotions in adaptive systems*, 32(5):531–559, July 2001a.
3. Sun R., Peterson T. Learning in Reactive Sequential Decision Tasks: the CLARION Model / Proc. 1996 IEEE ICNN, Washington, DC, USA. Plenary, Panel and Special Sessions Volume. – pp.70–75.
4. Karpov V. E. Emotsyy u temperament robotov povedencheskiye aspekty / V. E. Karpov // Yzvestiya ran. Teoryia y systemy upravleniya. – М.:, 2014, № 5, s. 166-185.
5. Kuzmyn V. Yspolzovanye neironnykh setei v alhorytme Q-learning / V. Kuzmyn // Transport and Telecommunication. – Ryha: Vol.4, N 1, pp. 75-86, 2003.

к.т.н., доц. Бойчук В.А., к.е.н., доц. Бойчук А.А., Бойчук М.В., Бурдюг О.В.  
**МЕТОД ФОРМИРОВАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ ДЕЙСТВИЙ  
ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ**

*В статье предложен подход, где реализация формирования последовательностей действий интеллектуальных агентов выполняется по аналогии с деятельностью биологических организмов с использованием механизма эмоций для динамической настройки организма на выполнение действий. Таким образом имитируются функции лимбической системы в организации движений на основе мотивационного поведения. При планировании в первую очередь определяется общее состояние агенту. Используя полученное состояние формируется последовательность действий. Такой подход даст возможность динамично перенастраивать последовательность и реагировать на опасную ситуацию или на изменение внутреннего состояния агента.*

*Интеллектуальный агент получает с сенсоров и рецепторов признаки начального условия по ней определяется цель и формируется последовательность действий. Элементами последовательности действий являются элементарные действия. Элементарное действие характеризуется набором входных параметров для функционирования. Признаки предпосылки соответствуют первом действию в последовательности, последнее действие в последовательности привязана к признаку цели.*

*Последовательность действий агенту представляется оргграф, где вершины определяют элементарные действия, а ребра определяют степень силы связи между ними. Начальные условия соответствуют первом действию в последовательности, с нее начинается реализация последовательности действий. Признаки цели соответствуют последней вершине в последовательности действий. Весы связей меняются при установке переменных общего состояния, что позволяет выполнить последовательность действий в реальном масштабе времени с динамической перенастройкой и выбрать среди характерных для конкретного состояния последовательностей действий.*

*Метод формирует последовательность действий, которая иницируется эмоциональными состояниями и переводит ее в последовательность автоматических действий на основе достижения цели и которая в будущем будет выполняться в нормальном состоянии. Для*

*проверки функционирования метода реализован симулятор агенту-работу в среде программы V-REP. Полученные результаты могут быть для интеллектуального планирования на основе подкрепления могут быть использованы при управлении агентами, работами на производственных предприятиях, военными агентами, потоками городского движения, логистическими системами, социальными явлениями.*

*Ключевые слова: интеллектуальный агент, планирование, модель, Q-обучения, эмоциональные состояния, обучение с подкреплением.*

**Ph.D. Boychuk V.O., Ph.D. Boychuk A.A., Boychuk M.V. Burdyug O.V.  
THE ACTION SEQUENCE FORMING METHOD FOR INTELLECTUAL AGENTS**

*The article proposes an approach where the implementation of the formation of sequences of actions of intelligent agents is carried out by analogy with the activities of biological organisms using the mechanism of emotions to dynamically tune the body to perform actions. Thus, the functions of the limbic system are simulated in the organization of movements based on motivational behavior. When planning, first of all, the general condition of the agent is determined. Using the resulting state, a sequence of actions is formed. This approach will make it possible to dynamically reconfigure the sequence and respond to a dangerous situation or to a change in the internal state of the agent.*

*An intelligent agent receives from the sensors and receptors signs of an initial condition, the goal is determined by it, and a sequence of actions is formed. Elements of a sequence of actions are elementary actions. An elementary action is characterized by a set of input parameters for functioning. Signs of the premise correspond to the first action in the sequence, the last action in the sequence is tied to the sign of the goal.*

*The sequence of actions of the agent is represented by a digraph, where the vertices determine the elementary actions, and the edges determine the degree of bond strength between them. The initial conditions correspond to the first action in the sequence, the implementation of the sequence of actions begins with it. Signs of the goal correspond to the last peak in the sequence of actions*

*Link weights change when general state variables are set, which allows you to perform a sequence of actions in real time with dynamic reconfiguration and select sequences of actions that are characteristic of a particular state. The method forms a sequence of actions that is initiated by emotional states and translates it into a sequence of automatic actions based on the achievement of the goal and which in the future will be performed in a normal state.*

*To test the functioning of the method, a agent-work simulator is implemented in the V-REP program environment. The results obtained can be used for intelligent planning based on reinforcements and can be used in the management of agents, work in manufacturing enterprises, military agents, urban traffic flows, logistics systems, and social phenomena.*

*Keywords: intellectual agent, planning, model, Q-learning, emotional states, reinforced learning.*