

DOI: 10.34185/1991-7848.itmm.2023.01.046

**ОПЕРАЦІОНАЛІЗАЦІЯ КОНЦЕПТІВ «АГЕНТНІСТЬ» ТА «МИСЛЕДІЇ»
НА ЗАСАДАХ ПАРАДИГМИ ГРАНИЧНИХ УЗАГАЛЬНЕНЬ**

Прокопчук Ю.О.

Інститут технічної механіки НАН України

Придніпровська державна академія будівництва та архітектури, Україна

Незалежно від рівня автономії агента, успішна поведінка ґрунтується на взаємодії з навколишнім середовищем та можливості спілкуватися з іншими агентами чи користувачами. Агент може використати ці навички, щоб вивчити правдиву модель знань про навколишнє середовище та, таким чином, передбачити наслідки своїх власних дій [1, 2, 3, 4]. Як виникає почуття агентності? Відповідно до поширеної точки зору [1], агентність стає сильнішою у міру підвищення відповідності між прогнозованим та фактичним ефектом дії. Таким чином, почуття агентності (Sense of Agency - SoA) виникає, коли зовнішні події, які йдуть за дією, узгоджуються з прогнозами ефектів дії. Найважливішим проявом почуття агентності є інтуїція [2]. Фундаментальною основою почуття агентності у людини є Суб'єктивність, яка формується, зокрема, «Світом/Простором фантазій» (Fantasy-Space).

Наявність Світу фантазій дозволяє агенту адаптивно регулювати свій зв'язок із навколишнім середовищем, що сприяє самоідентифікації. Рухаюча сила розвитку Світу фантазій (напрямок суб'єктивної оптимізації) – це породження арт-начерків та арт-композицій арт-начерків (ментально-моторних дій), що викликають максимальні емоції (реакції мозку-тіла). Світ фантазій дозволяє переживати-проживати арт-начерки ситуацій, сцен з максимальними емоціями, повністю відключаючись від "зовнішньої" реальності (метафора «уход в себе»). Іноді у деяких рисах фантазії можна реалізувати «наяву» (таке бажання постає як мета поведінки, уособлюючи Свободу Розуму; Spirituality Intelligence and Creativity). Мікс фантазій та реальності дозволяє досягти особливо яскравих переживань, відчуттів (емоційна нагорода). Це найважливіший аспект суб'єктивності, що багато в чому визначає цілеспрямовану поведінку (емоції людини частково конститууються її діями). Свідомість дозволяє відібрати цільовий гіпотетичний сценарій, найбільш наближений до поточних умов. По суті, будь-які неформальні плани (особливо віддалені) – це певною мірою фантазії-ілюзії. Мікс фантазій та реальності є основою «темних рішень» [2].

Часто люди створюють арт-начеркі (включаючи атрефакти мистецтва),

тому що це дає їм можливість випробувати особливі почуття, наприклад, зануритися у творчу гру та доторкнутися до чогось життєво важливого. Інакше кажучи, метою є «заселення» емпіричного простору фантазій, у якому почуття, що високо цінуються, активізуються і підтримуються в інтерактивному режимі (важливий аспект духовності). Саме так розвивається естетичне почуття краси та досконалості (Sketching to Feel).

Таким чином, одна з ключових особливостей суб'єктивної динамічної логіки - це незбагнений мікс фантазій і реальності що істотно впливає на всю поведінку. Необмежений Світ фантазій багато в чому формує простір емоційних цілей-винагород, будучи однією з причин incomputability (Reward-seeking motivation and behaviour). Основний внутрішній конфлікт і водночас мотивація – це розбіжність між фантазіями та реальністю (незадоволені бажання). Фантазії – це, перш за все, «свобода Розуму» (Conception of Freedom of Mind).

Ментальна свобода вимагає здатності слідувати за своїми думками, куди б вони не вели - краса і цінність у нашій здатності думати про це (навіть найогидніші установки, найупередженіші переконання та наймерзенніші фантазії не є морально забороненими). "Свобода Розуму" є важливою складовою світо подібності ментальної сфери. Психо-соціальним проявом свободи Розуму є «темний фактор особистості» (The Dark Factor of Personality) [5]. Глибока суперечність між внутрішньою «свободою Розуму» та зовнішньою «забороною на неетичну поведінку» (Moral Uncertainty) частково вирішується за допомогою «темних рішень / dark solutions» [2]. Вибір «ступеня темряви» феноменологічно визначає «свободу Волі». Без «креативного невігластва», «D-фактора», «свободи Розуму» та «свободи Волі», «контрольованих галюцинацій», «Світу фантазій», «обману», «темних рішень» сильний ШІ (AGI) не створити (Cognitive Design для Artificial Minds; Origins of Mind). Всі ці питання враховує Парадигма граничних узагальнень (ПГУ: пошук максимально узагальнених, але змістовних начерків концепцій та сутностей) [2].

Значною мірою зі Світом фантазій пов'язане контрфактуальне мислення. Слід також пам'ятати, що людина не тільки «створює себе», а й створює і власну екологічну нішу (соціум, предметне оточення), яка мінімізує невизначеність.

Морфологічні / морфогенетичні втілені обчислення *in vivo* (Morphological / Morphogenetic Embodied Computation) обумовлюють «почуття власності» на наші власні тіла – найважливішої компоненти агентності (The bodily self: "Sense of Ownership" for our own bodies; a rich foundation of primitive, nonconceptual self-consciousness; Primitive Forms of Self-awareness). ПГУ

розкриває суть і відмінність морфологічних та когнітивних обчислень (мережі духовних мереж начерків образів; узагальнене заплутування і т.д.; Spiritual Networks of Sketches) [2]

Миследії – це системо-патерни високого рівня, виконання яких супроводжується «почуттям агентності», «уявною дією» та формуванням «локальних стріл часу» – специфічного потоку спостережень (думку можна розглядати як результат розвитку дії у певному напрямку; ‘Action-thoughts’ as a foundation of mental work; Purposefulness of Thought; «Language of Thought»).

Відомий американський психолог Barbara Tversky вважає, що в основі людського пізнання та мислення лежить рух (Mind in Motion). В рамках діяльнісного підходу в психології миследія (Action-Thoughts [4]) вводиться як холістична одиниця аналізу індивідуальної діяльності, включаючи весь її спектр від елементарних дій-рухів до складної розумової роботи (The fundamentals of cognitive action theory; Cognitive and Affective Aspects of Agency; Beyond Automaticity: The Psychological Complexity of Skill/Action; Knowledge-How: the rich interplay between automatic and cognitive control processes; detailed and well-structured action representations and internal models). Можливість існування думок пов'язана з вродженою здатністю розуму абстрагувати дії від рухів [1, 2, 4]. Здатність справлятися з довгими послідовностями миследій пов'язана зі здатністю людини мислити укрупненими, об'єднаними миследіями, тобто. пов'язувати миследії в ієрархічні структури (мережі начерків патернів, мережі трансформацій). У рамках ПГУ миследії відповідають, зокрема, за формування «Суб'єктивного простору-часу-дій» (Subjective Space-Time-Action) за допомогою базового концепту «локальні стріли часу» (Local ‘Arrow of Time’ - LAoT) - це розгортання «схеми миследії» по досягненню мети (упорядковані множини емоційних подій) [2].

Деякі дії для Розуму, висловлюючись терміном французького філософа Мерло-Понті (Merleau-Ponty), формують «тілесну суб'єктивність» (bodily subjectivity). Мерло-Понті стверджує [6], що існує тип розумної взаємодії зі світом, не пов'язаний з мисленням. У деяких випадках перцептивний досвід людини може негайно ініціювати та спрямовувати дію без необхідності будь-яких проміжних думок. Безумовно, ми свідомо не думаємо про багато речей, які робимо щодня. Це все шаблонні або звичні дії, але це гнучка, розумна взаємодія зі світом, яка потребує певної адаптації до конкретної ситуації. Такі дії ми не відноситимемо до «миследій», що супроводжуються почуттям агентності. Також до миследій не відносяться радикали.

Можна вважати, що поряд із миследіями існує «лінійний» режим роботи

системи управління (The “lazy” control mechanism або режим “автопілота”), який не підтримує жодних внутрішніх моделей світу (основний режим у природі). За такого режиму роботи організму, агента «стріли часу» не породжуються. Відповідно, не виникає суб'єктивний простір-час-дії з усіма наслідками, зокрема, відсутністю агентності.

Суб'єктивний простір-час-дії дозволяє визначити відлуння подій-дій і, отже, розподілені в часі та в соціумі позитивні-негативні наслідки, включаючи «нагородження-покарання» (важливий аспект досвіду та агентності). Ця здатність наразі відсутня в системах зі штучним інтелектом, але обов'язкова для AGI.

З мислєдїями, як правило, пов'язаний тимчасовий та ресурсний синтез, прогноз. Тимчасовий синтез сенсомоторної дії сприймається як процес встановлення узгодженості тривалості дії (очікуваної тривалості), часової послідовності необхідних операцій. Ресурсний синтез пов'язаний з формуванням безлічі доступних (в даний момент) механізмів виконання патерну з опорою на соціум або багатоагентне середовище (це можуть бути радикали; приклад – мережі потреб та можливостей).

Тимчасовий і ресурсний синтез, безперервний прогноз у процесі розгортання дії є основою «уявної дії» - сукупності уявних начерків планів-прогнозів (Imaginal action: mental sketches of plans-forecasts). «Уявна дія» разом з емпіричними спостереженнями в процесі розгортання дії формує унікальну сутність суб'єктивного простору-часу-дій «Уявна дія – Дія / Спостереження» (Imaginal Action – Action / Observations). Унікальність даної сутності забезпечує єдність суб'єктивного та емпірично-тілесного начал. Віддаленою метафорою цієї сутності може бути «подвійна спіраль» ("Double Helix of Action-Thoughts": одну «спіраль» утворюють уявні начерки уявної дії, іншу – емпіричні начерки реальної дії; базовий елемент концепції «стріла часу»).

Творчість, інтуїція, фантазії, почуття агентності займають центральне місце у процесі бачення можливого майбутнього та формування довгострокових намірів (Long-Term Intentions; Sense of Control over Life). Одна з відмінностей людини від інших високоорганізованих тварин полягає у побудові «Образу Майбутнього» («The Image of the Future» - поняття, введене в науковий обіг голландським соціологом Фредом Полаком в 1973; Seeing the Future). Паттерн мислєдїї визначає дію мети (майбутнього стану) на сьогоднішній день (Принцип цільової причинності, Final Cause: «майбутнє тимчасове сьогоднішнього»; Future Making Machine).

З необмеженого репертуару мислєдїй у доповіді більш детально

розглядаються два типу мисленій [2]:

- виконання патерну розрізнення-управління на основі уявної «схеми дії» $V(\{\underline{a}/A\}, \underline{z}/Z)$ або $V = (f/\mu: \{\underline{a}/A\} \rightarrow_e \underline{z}/Z)$ для будь якої Z-задачі розрізнення;
- комунікативні дії або передача-прийом повідомлень з урахуванням семантичних аспектів (смислів).

Задача про «Мінімізацію кодів, знаків у повідомленні для передачі заданого сенсу». Таке завдання виконують критичні нариси, зокрема, «тонкий зріз» евристик у задачах розрізнення [2]. Проілюструємо схему передачі змісту повідомлення та дії культурного реплікатора в меметиці з урахуванням позначень [2]. Нехай агент Ag1 – адресант (комунікатор, ініціатор комунікації), а Ag2 – адресат (реципієнт, комунікант), тобто. одержувач інформації/повідомлення. Агент Ag1 передає повідомлення $Mes = \langle \underline{a}_1, \underline{a}_2, \dots, \underline{a}_k \rangle$ (\underline{a} – значення тесту a , знак/слово), при цьому в його ментальній сфері це повідомлення викликає, зокрема, такі когерентні асоціації та образи:

$$\langle \underline{a}_1^{\uparrow\downarrow}: Cog(\underline{a}_1, P_{a_1}) \{g/\mu\}_{a_1} \{LAoT\}_{a_1}; \underline{a}_2^{\uparrow\downarrow}: Cog(\underline{a}_2, P_{a_2}) \{g/\mu\}_{a_2} \{LAoT\}_{a_2}; \dots; \underline{a}_k^{\uparrow\downarrow}: Cog(\underline{a}_k, P_{a_k}) \{g/\mu\}_{a_k} \{LAoT\}_{a_k} \rangle_{Ag1}, Gs_{Ag1}(Mes),$$

де P_{a_1}, \dots, P_{a_k} - нариси деяких образів; $\underline{a}^{\uparrow\downarrow}$ - конуси узагальнення та деталізації значень тестів; $Cog(\underline{a}, P_a)$ – когі (когнітивні групи); $\{g/\mu\}_{a_i}$ – індуктори заплутування; $\{LAoT\}_{a_i}$ – локальні "стріли часу", в подіях яких зустрічається \underline{a}_i (важливий елемент заплутування); $Gs_{Ag1}(Mes)$ - мережа начерків Mes. У агента Ag2 повідомлення Mes викличе свої когерентні асоціації та образи (суть «культурного реплікатора»):

$$\langle \underline{a}_1^{\uparrow\downarrow}: Cog(\underline{a}_1, P_{a_1}) \{g/\mu\}_{a_1} \{LAoT\}_{a_1}; \underline{a}_2^{\uparrow\downarrow}: Cog(\underline{a}_2, P_{a_2}) \{g/\mu\}_{a_2} \{LAoT\}_{a_2}; \dots; \underline{a}_k^{\uparrow\downarrow}: Cog(\underline{a}_k, P_{a_k}) \{g/\mu\}_{a_k} \{LAoT\}_{a_k} \rangle_{Ag2}, Gs_{Ag2}(Mes).$$

При цьому Банки тестів комунікантів найімовірніше не збігаються: $\{a\}_{Ag1} \neq \{a\}_{Ag2}$. "Стріли часу" дозволяють і автору і читачеві ввести слова своїми умами та емоціями, а не просто осягати повідомлення (історію). По суті, це фігуративна мова для створення верств сенсу, до яких комунікант або читач (Ag2) звертається через почуття, символіку. Можливо, агент Ag1 і сподівався викликати саме таке розгортання смислів у Ag2, якщо Ag1 добре знає зміст ментальної сфери Ag2. Аналізовану ситуацію добре описує крилата фраза Оскара Уальда: «Я відповідаю лише за те, що говорю, але не відповідаю за те, ЩО ВИ чуєте». Класична інформація по Шеннону не враховує подібне розгортання-згортання смислів у комунікантів.

Висновки. Застосування запропонованої технології в системах штучного інтелекту дозволить впровадити когнітивний підхід, зокрема морфологічні та когнітивні обчислення. Крім того, підхід дозволяє накопичувати досвід

вирішення тих чи інших питань з автоматичним формуванням «стріл причинності», глибоких каузальних моделей та «схем образів/дій».

OPERATIONALIZATION OF THE CONCEPTS "AGENCY" AND "ACTION-THOUGHTS" BASED ON THE PARADIGM OF LIMITING GENERALIZATIONS

Prokopchuk Yurii

Abstract. The paper discusses ways to formalize and realize the basic concepts of cognitive science. Formalization is based on the paradigm of limiting generalizations. The foundations of the proposed models are the concepts of "Subjectivity," "Language of Thought," "Freedom of Mind," and "Fantasy Space". As an example, the model of unfolding the meaning of the message is considered. The results of the study can be used to create cognitive agents and robots.

Keywords: Creative Agency, Action-Thoughts, Fantasy-Space, Paradigm of Limiting Generalizations, Subjectivity, Double Helix of Action-Thoughts

Література / Reference

1. Chambon V., Sidarus N., Haggard P. (2014). From action intentions to action effects: How does the sense of agency come about?. *Frontiers in human neuroscience*. 8. 320. 10.3389/fnhum.2014.00320.
2. Prokopchuk Y. (2022). *Intuition: The Experience of Formal Research*. Dnepr, Ukraine : PSACEA Press. 724 p. (in RU)
3. Nosov P., Zinchenko S., Plokhikh V., Popovych I., Prokopchuk Y., Makarchuk D., Mamenko, P., Moiseienko, V., Ben, A. (2021). Development and experimental study of analyzer to enhance maritime safety. *Eastern-European Journal of Enterprise Technologies*, 4 (3 (112)), 27–35. doi: <https://doi.org/10.15587/1729-4061.2021.239093>
4. Simsky A. (2018). Action-thoughts: concept and conception. 10.13140/RG.2.2.34219.26405.
5. Moshagen M., Hilbig B. E., Zettler I. (2018). The dark core of personality. *Psychological Review*, 125(5), 656–688. <https://doi.org/10.1037/rev0000111>
6. Merleau-Ponty, M. (2012). *Phenomenology of perception*. Oxon: Routledge.