

DOI: 10.46340/ephd.2020.6.3.16

Anna OdynetsORCID ID: <https://orcid.org/0000-0002-5841-5722>*Taras Shevchenko National University of Kyiv, Ukraine*

HOW EMOTIONS ARE POSSIBLE IN TECHNOLOGICAL FORM OF LIFE?

Анна Одинець*Київський національний університет ім. Тараса Шевченка, Україна*

ЯК ЕМОЦІЇ Є МОЖЛИВИМИ У ТЕХНОЛОГІЧНІЙ ФОРМІ ЖИТТЯ?

This article is dedicated to the question: how emotions are possible in a technological form of life. To understand in more detail how it is possible to model emotions artificially, it is necessary to find out how they occur in living beings, in particular humans. For this purpose, this article offers a hypothesis which proposes a distinction of two components of consciousness and the model of the process of the emergence of emotions in living beings, as well as a consistent historical-theoretical analysis of the ideas which have influenced the modeling of emotional robots today and can influence in the future. The third part provides examples of successful implementations of models of emotional robots that are able to respond to the tone of voice, human behavior, and are able to learn. As a result, this article concludes that for the development of a full-fledged personality in a technological form of life, emotionality plays an extremely important role, and the latter is impossible without corporality as a source of personal experience.

Keywords: emotions, model of the process of the emergence of emotions, consciousness, artificial intelligence, robot, technological form of life.

Основні побоювання, пов'язані зі штучним інтелектом, роботами, технологічними формами життя, врешті-решт, зводяться до того, що істоти, які будуть здатні обчислювати величезні масиви інформації, не будуть здатні співпереживати і тоді забудуться футуристичні жахи щодо холодних бездушних машин. Бо співпереживання – є саме те, що дозволяє людям існувати в одному культурному та соціальному просторі. Не сухі закони та правила, а саме те, що нас об'єднує, один контекст. А співпереживання базується на емоціях.

Тому, мета цієї статті зрозуміти, що таке емоція, як утворюються емоції та чи будуть здатні технологічні форми життя перейняти таку здатність. Перші два питання тісно пов'язані, але без відповіді на них, ми не зможемо перейти до третього, найціннішого для мого дослідження. Тому, для початку пропоную розглянути наступну гіпотезу:

Свідомість, в сенсі усвідомлення навколишнього середовища та реакції на нього, має два основні складники: когнітивний (як здатність розуміти, пізнавати, вивчати, усвідомлювати, сприймати і переробляти зовнішню інформацію), який відповідає за моделювання, обчислення варіантів, etc., та емоційний, який у свою чергу коригує когнітивний. Тобто емоції – це механізм, що коригує когнітивні процеси. Для мого дослідження важливо зробити акцент на розрізненні цих двох складників та їх поступовій появі у нашій свідомості. Тому що для того, щоб перейти до модулювання емоцій штучної форми життя, необхідно з'ясувати, що саме і для чого ми моделюємо.

Все, що істоти (бо наділені свідомістю не тільки людські особини) сприймають, вони сприймають когнітивно. Емоції примішуються пізніше, в процесі обробки та «усвідомлення» ситуації. Наприклад, людина, що страждає арахнофобією, побачила павука. На перший погляд здається, що страх виник миттєво. Але насправді, для того, щоб виникло почуття страху, необхідно спочатку сприйняти фізично (побачити, почути, відчути), усвідомити, що це (когнітивний складник), а тільки потім з'являється емоція, як внутрішня реакція на сприйняте.

Чому ми не вважаємо, що поведінка інфузорії-туфельки, яка наштовхнулася на об'єкт та оминає його, означає, що вона відчуває емоцію страху? А равлик, який втягує голову в панцир після дотику –

огиду? Бо в них немає міміки. Люди схильні приписувати емоції тільки істотам, які наділені рухомими м'язами обличчя (або взагалі тим, що мають обличчя). Навіть відповідна поведінка не є тим, що провокує нас проводити асоціації з емоційними сигналами.

Отже, на якому етапі з'являються емоції? Відповідаючи на це питання, пропоную розглянути другу частину гіпотези, яка описує процес появи емоцій. Для зручності візуалізації представляю таку схему (Рис. 1):



Рис. 1. Процес виникнення емоцій

Джерело: авторська розробка

Отже, що інформація з навколишнього середовища потрапляє на органи чуттів суб'єкта, якими він може їх сприйняти. Процес сприйняття ділиться на свідомі та несвідомі процеси. Свідомі – це когнітивні, ті, які ми актуально маємо у мисленні, пам'яті, можемо артикулювати, відслідкувати причино-наслідкові зв'язки. Несвідомі, відповідно, ті, які перебігають без залучення свідомого досвіду, це миттєві реакції тіла, «забуті» асоціації. Наприклад, коли ми вперше бачимо людину, вона може викликати яскраву антипатію, але якщо запитати, ми не можемо сказати чому. Тому що, свідомо, артикульовано, немає ніяких доказів на користь такого ставлення. «Я так відчуваю». І саме це відчуття йде з несвідомого, з цілого спектру асоціацій рухів, запахів, виразу очей, які свідомо дуже важко відслідкувати і якщо спробувати, це займатиме багато часу і зусиль. Коли ми не можемо сказати, чому ми щось робимо чи відчуваємо – це не значить, що немає реальної причини, це значить, що нам ще не вдалося її віднайти.

Саме ці несвідомі реакції на те, що було сприйнято з навколишнього середовища, запускають реакції тіла, найважливішою серед яких є викид гормонів, що в свою чергу впливає на все інше: на темп серцебиття, на глибину дихання і ширину зіниць та інше. І ці реакції тіла, власне і є основою емоцій.

Далі, емоції впливають на когнітивний процес, який відбувається набагато повільніше за несвідомий, і вже з цієї сукупності формується відповідь, яка і виходить знову у навколишнє середовище.

Я вважаю, що реакції тіла є основою, оскільки щоб емоціям стати власне тим, що ми маємо на увазі, коли в повсякденній мові кажемо про емоції – має відбутися їх мінімальне усвідомлення.

На користь моєї гіпотези слугують дослідження такого складного явища як алекситемія. Алекситемія – це розрад аутичного спектру, при якому люди не відчувають емоцій, або їм важко їх розрізнити. При цьому вони здатні відчувати фізичні прояви, які традиційно пов'язують з переживанням емоцій. Такі як прискорене серцебиття чи нудота при небезпеці, або тиск при хвилюванні. «Вчені припускали, що цей феномен спричинений порушенням зв'язку між двома півкулями мозку, в результаті чого сигнали від емоційних відділів мозку, переважно в правій півкулі не досягають мовних відділів – у лівій. Це часто трапляється під час хірургічного лікування епілепсії, розповідає Катаріна Герліх-Добре з Аахенського університету. Розріз волокон між півкулями мозку полегшує епілептичні напади, але пацієнти втрачають здатність відчувати емоції. Сканування мозку людей з алекситемією показали, що вони мають аномально щільні зв'язки в цьому відділі, що створює перешкоди передачі емоцій. Андре Алеман з Медичного центру Університету Гронінгена в Нідерландах виявив брак сірої речовини у відділах головного мозку, пов'язаних з увагою. Коли людина з алекситемією дивиться на зображення людей, які відчувають емоції, її мозок наче не реєструє ці

*почуття*¹». Тобто, реакції тіла відбуваються, але головний мозок не може адекватно опрацювати отриману інформацію, не може донести її в мовні центри, артикулювати, а тільки артикульована інформація може бути усвідомлена. Тобто перебіг переривається саме у момент, коли емоції мали б скомпонуватися з когнітивними процесами і як наслідок, відповідь формується тільки з останніх. Але детальний розгляд такого відхилення слугує тільки на користь моєї гіпотези.

Отже, емоція – це внутрішня несвідома реакція організму на сприйняте з навколишнього середовища, яка виникає паралельно з когнітивними процесами та впливає на формування відповіді.

Для того, щоб зрозуміти, як емоції є можливими у нелюдській формі життя, пропоную для початку зробити невеликий історико-теоретичний екскурс, адже практика завжди ґрунтується на теорії. Отримавши уявлення про те, які теорії вплинули, або мають потенціал вплинути на розвиток емоційних робіт, ми зможемо зрозуміти, як їх можливо моделювати. Щоб не відходити надто далеко, пропоную тільки ті теорії, які, з мого погляду, можуть бути релевантними на наш час.

Основна теорія емоцій. Емоції як розвинені програми впливу. Основна теорія емоцій виникла в 1970-х роках із піонерської роботи Сільвана Томкінса, суть якої полягала в тому, що *«первинною мотиваційною системою є афективна система»*². Він запропонував існування дев'яти основних або вроджених афектів, що контролюються програмами, які успадковуються: інтерес, задоволення, здивування, страх, гнів, горе, ганьба, презирство та огида. Саме від розрізнення приємного або болісного бере коріння їх мотиваційна сила, а такі гедоністичні відчуття виникають у результаті сприйняття змін виразів обличчя, що забезпечує у свою чергу *«зворотній мотиваційний зв'язок»*³.

Теорія Томкінса стала основою для розвитку ще двох теорій на своїй базі: (1) перша – це зародження сучасної теорії основних емоцій, яка звертає увагу на універсальність міміки, Пола Екмана та Керрола Ізард; друга – поява еволюційного психологічного підходу, де емоції розуміються як механізм для розв'язання повторюваних еволюційних проблем, з помітними внесками Роберта Плутчіка. В 90-х рр. XX століття, два підходи почали поступово зливатися. Пол Екман зазначає, що *«емоції розвинулися для адаптації при вирішенні основних життєвих завдань»*⁴, таких як *«переживання веселощів, кохання, порятунок від хижаків, протистояння сексуальній невірності, переживання втрати статусу, спричинену невдачею, реакція на смерть члена сім'ї»*. Вони мають адаптивне значення тому, що швидко мобілізують та координують ресурси, необхідні для успішного вирішення життєвих завдань, і тому, що вони передають суспільно важливу інформацію за допомогою тілесних виразів. Як тільки активована основна програма емоцій, а *«каскад змін (без нашого вибору або негайного усвідомлення) відбувається за секунди: емоційні сигнали в обличчі і голосі; задані дії; вивчені дії; діяльність вегетативної нервової системи, яка регулює наш організм; регуляторні структури, які постійно змінюють нашу поведінку; пошук відповідних спогадів та очікувань; і як ми інтерпретуємо те, що відбувається всередині нас і у світі»*⁵.

З цієї думки випливає центральна емпірична гіпотеза основної теорії емоцій: для кожної основної емоції повинні бути тілесні маркери, що складаються з корельованих, специфічних змін на рівні міміки, вегетативної системи, заданих та вивчених дій. Експлікуючи, Екман окреслив межі основних емоцій з точки зору їх: (а) характерних універсальних сигналів, (б) характерної фізіології, (в) особливості думок, спогадів і образів, (г) характерних суб'єктивних переживань, (д) передбачуваного розвитку зовнішності, (е) гомологічної присутності у інших приматів; (є) автоматичної оцінки, налаштованості на відмінні універсали в попередніх подіях, (ж) швидкості їх настання, короткочасність та неприхованість появи⁶. Озброївшись цим визначенням, Екман продовжував стверджувати, що ми маємо емпіричні докази для шести основних програм впливу (щастя, смуток, гнів, страх, огида та здивування), пізніше розширивши список, щоб включити стани, чий основний статус, ймовірно, буде доведений у майбутньому, такі як розвага, зневага, збентеження,

¹ Robson, D. (2015). What is like to have never felt an emotion? *BBC Future*.

<<https://www.bbc.com/future/article/20150818-what-is-it-like-to-have-never-felt-an-emotion>> (2020, серпень, 13).

² Tomkins, S. S. (2008). *Affect Imagery Consciousness: The Complete Edition*. New York: Springer Publishing Company, I, II, III, IV, 4.

³ Там само, 623.

⁴ Ekman, P. (1999). Basic Emotions. in Tim Dalgleish and Mick J. Power (eds.). *Handbook of Cognition and Emotion*, Chichester. UK: Wiley and Sons. doi:10.1002/0470013494.ch3, 46.

⁵ Ekman, P., Cordaro D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, 3(4). doi:10.1177/1754073911410740, 366.

⁶ Ekman, P. (1999). Basic Emotions. in Tim Dalgleish and Mick J. Power (eds.). *Handbook of Cognition and Emotion*, Chichester. UK: Wiley and Sons. doi:10.1002/0470013494.ch3, 46.

хвилювання, провина, гордість за досягнення, полегшення, задоволення, здивування, екстаз, чуттєве задоволення та сором¹. Науково налаштовані філософи часто обмежують свої дискусії щодо емоцій основними афективними програмами, оскільки вони, як стверджується, є єдиними на сьогодні природними видами, виявленими в афективній галузі².

Центральною доказовою базою для основних теорій емоцій є крос-культурні дослідження міміки, започатковані Чарльзом Дарвіном, в яких використовуються техніка розпізнавання виразів обличчя. Суть дослідження полягало у експерименті, де людям демонстрували зображення з різними емоційними виразами обличчя та просили визначити, яку емоцію переживає зображений, з шести-десяти запропонованих термінів, мовою людей, яких запитували. Як зауважує Пол Екман, досі подібні експерименти проводилися з спостерігачами з десятків країн, що виявило високий рівень узгодженості щодо того, яка емоція зображена (показники розпізнавання найбільш співпадають для щастя, смутку та огиди). Говорячи про це, приналежність до даної культури збільшує розпізнавання виразів із цієї культури, що спонукало деяких до думки про те, що емоційні мімічні вирази є універсальною мовою з різними діалектами³.

Здобутки основної теорії емоцій для робототехніки полягають у виділенні основних емоцій, які слугують базовою складовою при моделюванні. Адже весь спектр емоцій надзвичайно широкий і дрібних відтінків є безліч, а виділення вузлових елементів допомагає досягнути весь принцип дії. Другий важливий аспект – це, власне, ствердження універсальності міміки. Адже, моделюючи можливості відображення необхідно робити це так, щоб емоції читалися людьми будь-якої культури.

Емоції як конструкції. Психологічний конструкціонізм.

Впливовою альтернативою для основної теорії емоцій є психологічний конструкціонізм. Дослідники висунули думку про існування специфічних біологічних підписів для емоцій, особливо на рівні вегетативних та нейронних змін. Іншим варіантом є перехід до нової базової теорії емоцій, яка замінює припущення щодо каскадних відповідей, тенденціями до дії з контрольним пріоритетом, що враховувало б певну мінливість відповідей, зберігаючи основну думку про те, що основні емоції спрямовані до вирішення еволюційних проблем. Важливою для нас думкою в цій традиції є створення циклічної схеми емоцій Джеймсом Расселом. Так Рассел в своїй роботі «Циклічна модель афектів» (Рис. 2) :



Рис. 2. Циклічна схема емоцій Дж. Рассела

Джерело: авторський переклад⁴

¹ Ekman, P., Cordaro D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, 3(4), 364-370. doi:10.1177/1754073911410740.

² DeLancey, C. (2002). *Passionate Engines: What Emotions Reveal about Mind and Artificial Intelligence*. Oxford: Oxford University Press.

³ Ekman, P. (1999). Facial Expressions. in Tim Dalgleish and Mick J. Power (eds.), *Handbook of Cognition and Emotion*, Chichester. UK: Wiley and Sons, 301-320. doi:10.1002/0470013494.ch3.

⁴ Russell, J. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. doi:10.1037/h0077714 1161–1178. doi:10.1037/h0077714.

Він вважає, що одним з компонентів, на яких будується емоція є центральний афект («*core affect*»), який є «нейрофізіологічним станом, що доступний як просте неререфлексивне відчуття і є інтегрованою сумішшю гедонічних (задоволення – незадоволення) і збуджуючих (активація – сон) значень»¹. Психологічний конструкціонізм стверджує, що ми завжди знаходимося в певному стані по відношенню до центрального афекта, який є нібито барометр, що інформує нас про наші відношення до подій, які з нами відбуваються. Ці показання можуть бути представлені у вигляді точок вздовж «структури кругообігу», при цьому вертикальна вісь представляє ступінь активації-деактивації, а горизонтальна вісь представляє ступінь задоволення-незадоволення.

Ця схема має важливе значення для майбутніх розробок у сфері робототехніки, оскільки пізніше буде використана як база для відтворення емоцій у роботів, хоча і змінена, а з двомірної перетвориться у тривімірну модель. Але принцип активації та поступового перетікання однієї емоції в іншу залишиться.

Енактивістські теорії емоцій: емоції як енації

Енактивізм – це міждисциплінарна дослідницька програма, представлена групою теорій, спільним для яких є незадоволення вивченням когнітивних процесів у класичних когнітивних науках. Ця програма розвивається на базі натуралістичної епістемології, свідомість розглядається як результат розвитку живої природи, з огляду на складність її форм, як втілена та ситуативна, активна й інтерактивна, її когнітивна здатність здійснюється через активацію у навколишньому середовищі, тобто енактивування середовища.

Енактивізм вплинув на теорію емоцій та вніс дві важливі ідеї:

1. Ідея активної ролі того, хто пізнає зовнішній світ, який для енактивістів, не даний пасивно, а активно проявляється, формується «смыслотворчими силами» суб'єкта пізнання. Ця діяльність «осмислення сенсу» лежить в основі пізнання, як її розуміють енактивісти, і вона доступна всім живим істотам, якими б простими вони не були, оскільки вони є автономними та адаптивними системами².

2. Фокус на втіленому, вбудованому та розширеному характері когнітивних процесів. В той час, коли нейро- та когнітивні науки зосереджуються на дослідженні діяльності мозку, відриваючись від решти тіла та навколишнього середовища, енактивісти стверджують, що ми не зможемо зрозуміти пізнання, якщо нехтувати взаємними причинно-наслідковими взаємодіями між мозком, тілом та навколишнім середовищем, які динамічно розгортаються у часі.

Ці ідеї про те, що складні когнітивні процеси є чутливими до змін у навколишньому середовищі призвело до звернення особливої уваги дослідників на роль міжособистісного спілкування в соціальних середовищах³. Таке розуміння тілесності наближує енактивізм до феноменологічної традиції, яка стверджує, що тіло є досвідченою структурою, а не просто фізичним об'єктом, і те, що ми здатні відчувати, обмежує наш світ.

Енактивісти, на противагу «безтілесній теорії» більшості когнітивних теорій, розглядають емоції як тілесні процеси, а не інтелектуальні. Деякі з дослідників запропонували розглядати емоції, роблячи наголос на їх дієво-орієнтовану природу. Така думка логічно витікає за ствердження визначальної ролі дії в пізнанні. Пізнання, на їх думку, енактивується (вводиться в дію) по-суті телеологічними живими системами у відповідності до цілей їх дії. Не зважаючи на це, серед енактивістів не існує єдиної позиції на природу емоцій, але є низка чітких пропозицій.

Наприклад, показовою є робота Джованні Коломбетті, для якої ключову роль відіграє поняття самоорганізації. На думку дослідниці, «емоційні епізоди – це самоорганізуючі закономірності організму, найкраще описані концептуальними інструментами теорії динамічних систем»⁴, галуззю математики, покликаної враховувати тимчасову еволюцію систем, що змінюються з плином часу. Самоорганізація – це «активний процес формування, відтворення, збереження або удосконалення організації складної динамічної системи, що забезпечує її нормальне існування і функціонування»

¹ Russell, J. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110(1), 147. doi:10.1037/0033-295X.110.1.145.

² Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

³ Griffiths, P., Scarantino, A. (2009). Emotions in the Wild: the Situated Perspective on Emotion. in Philip Robbins and Murat Aydede (eds.), *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press, 437-453. doi:10.1017/CBO9780511816826.023.

⁴ Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press, 53.

як цілісного утворення; це іманентна здатність динамічної системи самостійно підтримувати, відтворювати чи удосконалювати рівень своєї організації при зміні зовнішніх чи внутрішніх умов її існування та функціонування, спрямована на підвищення її стійкості, збереження цілісності, забезпечення ефективних дій чи розвитку»¹. Застосовуючи це поняття, стосовно теорії емоцій, мається на увазі, що емоції, які з цієї точки зору також є складними динамічними структурами, самоорганізуються, що домагає пояснити мінливість емоційних епізодів у часі, оскільки самоорганізуючі системи можуть опинитися в декількох кінцевих станах залежно від взаємодії їх компонентів.

Хоча існує дещо спільне між поглядами психологічних конструкціоністів, щодо емоцій як таких, що виникають у відповідь на зміну навколишнього середовища, але Коломбетті заперечує, що емоції виникають у відповідь на концептуальні дії, а стверджує, що осмислення сенсу є набагато примітивнішим явищем, доступним від бактерій до людини. Істоти беруть участь у «осмисленні сенсу», коли вони оцінюють навколишнє середовище з точки зору того, чи сприяє це їх самозбереженню, і діють так, щоб поліпшити свою життєздатність у навколишньому середовищі, як це робить бактерія, коли відпливає від шкідливої речовини.

Таке розуміння емоційних епізодів, як здатних до самоорганізації, протиставляється уявленню про те, що вони спричинені афективними програмами. Джованні Коломбетті вважає, що поняття базової емоції введено доволіно і не варте збереження, оскільки нівелює необхідність дослідження нейронних, тілесних та поведінкових особливостей нібито не основних емоцій. Ребека Хуфендієк у праці «Втілені емоції: натуралістичний підхід до нормативного явища» робить додатковий висновок про те, що «неосновні емоції проявляють велику кількість особливостей, що не є характерними для нібито основних емоцій»².

Ще одна важлива ідея енактивізму – його антирепрезентативна позиція. Наприклад, Деніел Хутто запропонував «відпустити ідею про те, що емоції представляють ситуації, визначені правдою»³, припускаючи, що емоції не є основними реляційними темами. Наприклад, страх не означає, що реально існує небезпека, а гнів не означає, що мною знехтували. Суть антирепрезентативної позиції Хутто полягає в тому, що якщо приписувати емоціям репрезентативні здатності, це не завжди буде мати пояснювальну цінність. Він припускає, що психологічний стан може вважатися репрезентативним лише у тому випадку, якщо інші системи сприймають те, на що він вказує, або про що говорить⁴.

Альтернативне застосування нерепрезентативної теорії емоцій пропонують Деніел Шаргель і Джесс Принз. Вони роблять акцент на змісті емоцій. На їх думку, емоції створюють можливості дій, з огляду на силу підготовленості тіла, в яке вони включені. Ці можливості «залежать від стану (виникають, як правило, лише після того, як емоція ініціюється), і є імперативними (вони мотивують дію)»⁵. З цієї точки зору, сам страх не становить небезпеки, але створює можливості для втечі, є атрактором, який схиляє агента до певної дії.

Ще одну важливу нейробіологічну гіпотезу «соматичного маркера» запропонував Антоніо Дамасіо. Вона полягає в тому, що емоції спрощують процес прийняття рішення шляхом швидкого маркування можливих варіантів у передній корі головного мозку як позитивних чи негативних з огляду на очікувані емоційні наслідки. Дамасіо стверджує, що пацієнти з вентромедіальним префронтальним пошкодженням стають нераціонально схожими на Гамлета, коли стикаються з тривіальними рішеннями, такими як вибір дати для чергового призначення лікаря, ірраціонально схильні до ризику, коли стикаються з рішеннями з азартними іграми, і ірраціонально нетерплячі, коли стикаються з рішеннями, що вимагають відкладеного задоволення⁶.

Ідеї енактивізму значною мірою впливають на розробки в сфері штучного інтелекту і робототехніки. Висновки, щодо того як організми взаємодіють з навколишнім середовищем,

¹ Шинкарук, В. І. (ред.) (2002). *Філософський енциклопедичний словник*. Київ: Інститут філософії імені Григорія Сковороди НАН України: Абрис, 563.

² Hufendiek, R. (2016). *Embodied Emotions: A Naturalist Approach to a Normative Phenomenon*. London: Routledge.

³ Hutto, D. (2012). Truly Enactive Emotion. *Emotion Review*, 4(2), 176-181. doi:10.1177/1754073911430134.

⁴ Hutto, D. (2012). Truly Enactive Emotion. *Emotion Review*, 4(2), 176-181. doi:10.1177/1754073911430134.

⁵ Shargel, D., Prinz J. (2018). An Enactivist Theory of Emotional Content. in *Hichem Naar and Fabrice Teroni*

(eds.), *The Ontology of Emotions*. Cambridge: Cambridge University Press, 110-129. doi:10.1017/9781316275221.007

⁶ Damasio, A. (2003). Looking for Spinoza: Joy, Sorrow, and the Feeling Brain. *Orlando, FL: Harcourt*.

застосовуються до роботів, робиться припущення, що вони мають бути спроектовані таким чином, щоб, аналогічно організмам, мати змогу взаємодіяти з середовищем в якому перебувають, отримувати досвід, навчатися і розвиватися з його допомогою. А людина, в свою чергу, має змогу взаємодіяти з роботами в одному полі, застосовуючи всі візуальні, аудіальні і тактильні можливості, взаємно досліджуючи і збагачуючись досвідом.

З іншого боку, діяльність в спільноті штучного інтелекту, взаємно, впливає на енактивізм. Посилаючись на методи моделювання еволюційної робототехніки Бером, моделювання поведінки в навчанні Кельсо, та моделювання сенсомоторної активності Сольцмана, МакГанн, Де Ягер та Ді Паоло обговорюють, як ця робота робить динаміку зв'язку між агентом та його середовищем, фундаментом енактивізму, «*емпірично спостережуване явище*». Вони зазначають, що «*такі методи моделювання дозволяють досліджувати простір параметрів зв'язку між агентом і середовищем ... до того моменту, доки їх основні принципи (універсали, якщо такі є, дієвої психології) зможуть бути чітко простежені*»¹.

Отже, ідеї енактивізму дозволяють від теоретичних доробок підійти впритул до практичного втілення модельованих емоцій у технологічній формі життя.

Практична реалізація ментальної моделі робота.

Яким чином людина розуміє, що відбувається процес спілкування з іншим суб'єктом? Спілкування – процес, що потребує щонайменше двох учасників. Кожен з яких має здаватися іншому свідомим ситуації, адекватно реагувати, тобто перебувати в одному культурному і соціальному контексті. А це досягається не тільки доцільним обміном вербальною інформацією (як відомо, вербальна складова при спілкуванні коливається на рівні 20%), а й усім складним комплексом невербальних сигналів, які в свою чергу, контролюються несвідомим процесом, багато в чому покладаючись на емоційну складову. Тобто, для того аби зробити спілкування ж технологічною формою життя більш коректною для них мають бути створені можливості «відчувати» емоції. Як це можливо практично?

На сучасному етапі розвитку, роботи мають здатність відображати певний набір емоцій. Це робиться здебільшого для гуманоїдних роботів, які створюються з метою спілкування та допомоги людям, це роботи-помічники. Робиться це для того, щоб зробити їх взаємодію більш природньою. Тому зростає науковий інтерес до розвитку роботів, які можуть висловлювати емоції за допомогою міміки, імітуючи зовнішності людини, наприклад SAYA² або Geminoid³. Хіройоші Міва з колегами розробляють серію WE-4 з 2001 року: ці роботи можуть висловлювати свої емоції у верхній частині свого тіла (за допомогою рук та обличчя). Але вони не вмюють ходити. Роботи, які мають на меті природне спілкування з людьми, повинні вміти не тільки виражати статичні емоції, але й динамічно змінювати свої емоції у відповідь на зовнішні подразники та змінювати свою поведінку залежно від їх емоційного стану⁴. Тому, дослідники вважають, що роботи повинні мати «особистість», щоб було, що, власне, виражати.

Модель особистості робота була запропонована дослідниками з токійського університету Вседа, зокрема, Хіройоші Міва, в роботі «Особистість робота, заснована на рівняннях емоцій у тривимірному ментальному просторі»⁵. Ця модель була розроблена та застосована для робота WE-4RП. У цій моделі робот має два рівня особистості. Чутливу (ту, яка відчуває) та виразну (ту, яка виражає). («*Sensing personality and expression personality*».) (Рис. 3). Зовнішні подразники інтерпретуються чутливим рівнем особистості робота і це викликає зміну емоційного стану, тоді емоційний стан модифікує поведінку робота через виразний рівень особистості. Одним з перших таких роботів був «Кісмет»⁶ з рухомими повіками, очима та губами, побудований Сінтією Брейзеал та її колегами з Массачусетського технологічного інституту (MIT). Діапазон емоційних виразів,

¹ McGann, M., De Jaegher, H.; Di Paolo, E. (2013). Enaction and psychology. *Review of General Psychology*, 17 (2), 203-209. doi:10.1037/a0032935.

² Hashimoto, T., Hitramatsu, S., Tsuji, T., Kobayashi, H. (2006). Development of the Face Robot SAYA for Rich Facial Expressions. *Proc. of SICE-ICASE International Joint Conference*, 5423-5428.

³ Ishiguro, S. N. (2007). Building artificial humans to understand humans. *Journal of Artificial Organs*, 10(3), 133-142.

⁴ Miwa, H. (2001). Robot Personality based on Equations of Emotion defined in the 3D Mental Space. *Proc. of IEEE International Conference on Robotics and Automation*, 2602-2607.

⁵ Там само.

⁶ Breazeal, C., (2020). Kismet, the robot. *Kismet*. <<http://www.ai.mit.edu/projects/sociable/baby-bits.html>> (2020, серпень, 13).

доступних Кісмет, обмежений, але вони досить переконливі, щоб викликати співчуття у людей, які взаємодіють з ним. Дослідниця запрошувала батьків з дітьми щодня грати з Кісмет. Залишившись один, Кісмет виглядає сумно, але коли він бачить людське обличчя, він посміхається, привертаючи увагу. Якщо людина рухається занадто швидко, Кісмет поглядом демонструє страх, розуміючи, що щось не так. Люди, які грають з Кісмет, не можуть не співчувати цим простим формам емоційної поведінки. Отже, чи є у Кісмета емоції? Він, безумовно, проявляє деяку емоційну поведінку, так якщо ми визначаємо емоції в поведінковому відношенні, ми повинні визнати, що Кісмет має певну емоційну здатність. Кісмет не відображає всього спектру емоційної поведінки, яку ми спостерігаємо у людей, але здатність до емоцій – це не все, або нічого. Подібні здібності проявляє негуманоїдний робот Леонардо¹, ментальна модель якого дозволяє змінювати емоційний стан, динамічно впливаючи на поведінку робота та процеси навчання. Сінтія Брейзеал навчає його розпізнавати предмети та формувати власний емоційний висновок на основі того, що вони досліджують. Робот розуміє тон голосу і реакцію людини, які розмовляють з ним, і пов'язує цю поведінку з об'єктом. Так само, як діти дізнаються, яку реакцію вони повинні мати у певній ситуації, мімікуючи поведінку батьків. Ці роботи ще не вміють говорити, але їх реакції є надзвичайно красномовними².



Рис. 3. Модель емоцій негуманоїдного робота

Джерело: авторський переклад³

Розглядаючи можливості існування роботів в соціальному середовищі, важливим аспектом є фізична відстань між роботом та іншим діючим агентом (найчастіше, людиною). Щоб впоратися з ним, необхідно звернутися до досліджень в психології та робототехніці. Наприклад, Едвард Холл, пропонує теорію особистого простору, в якій фізична відстань між партнерами впливає на їх поведінку та емоції⁴.

Тож як практично реалізується ментальна модель Міва? Модель полягає в інформаційному потоці між роботом і середовищем. Інформаційний потік має два напрями: всередину (зовнішні стимули сприймаються роботом і впливають на емоційний стан робота) і назовні (робот модифікує середовище відповідно до поведінки, на яку впливає емоційний стан). Модель описує особистість робота, яка складається з чутливої особистості та виражаючої особистості. Особистість, що відчуває, визначається як зв'язок між зовнішніми подразниками та емоціями робота, тоді як особистість, що

¹ Thomaz, M. B., Breazeal, C. (2005). An Embodied Computational Model of Social Referencing. *Proc of the 14th IEEE International Symposium in Robot and Human Interactive Communication, Workshop*, 591-598.

² YouTube (2020). *BBC's Horizon. Leonardo Robot – isn't this the cutest robot ever?* <https://www.youtube.com/watch?v=ilmDN2e_Flc> (2020, серпень, 13).

³ Miwa H. (2001). Robot Personality based on Equations of Emotion defined in the 3D Mental Space. *Proc. of IEEE International Conference on Robotics and Automation*, 2602-2607.

⁴ Hall, E. (1996). *Hidden Dimension. Doubleday Publishin., ISBN 0-385-08476-5.*

виражає, визначається як взаємозв'язок між емоційним станом робота та фактичною поведінкою робота. Емоції представлені тривимірним відображенням, що залежать від векторів: «Активність», «Задоволення» і «Впевненість».

Застосовуючи модель Міва до робота, дослідники зберігають відображення тривимірної моделі відображення емоцій і визначають конкретні відносини, що формують особистість сприйняття і вираження робота. Щоб побудувати ці відносини, вони розглядають набір стимулів і поведінки. Оскільки відстань між роботом і людиною-агентом є критичним аспектом у взаємодії і спілкуванні людини з роботом, в поточній реалізації дослідники мають справу з візуальними стимулами, і зокрема з більш прийнятною відстанню від візуально ідентифікованого об'єкта, утримуваного людиною. Чуттєва особистість базується на теорії особистого простору Едварда Холла¹. Якщо об'єкт далеко, «приємність» зменшується. Якщо об'єкт наближається і потрапляє на область в межах 450 [мм] та 750 [мм] відстань, «приємність» збільшується, тоді як якщо об'єкт ближче 450 [мм] «активація» збільшується.

Що стосується поведінки, дослідники токійського університету Т. Кіші, Н. Ендо та ін., інтегрували три основні дії: візуальне відстеження, рухливість та вираження емоцій. Зокрема, вони розширили алгоритм візуального відстеження для двоногих роботів, запропонований у їх попередній роботі², де в основному кут зору, поворот голови, шиї та талії обчислюються спочатку відповідно до сприйнятого відносного напрямку об'єкта, що відстежується; а потім регулювався кут нахилу та повороту ступенів залежно від відстані та відносного кута об'єкта. У теперішній реалізації відносна відстань і кут нахилу об'єкта інтерпретуються через чутливу особистість, впливаючи на емоційний стан робота. Потім особистість, що виражає генерує міміку обличчя, рухи рук та бажану відстань від цільового об'єкта залежно від емоції. Нарешті, траєкторія ходьби формується на основі цільової відстані, і робот виконує ходьбу. Робот також висловлює свої емоції обличчям і руками, використовуючи заздалегідь визначені зразки для кожної емоції. (Зразки були надані професійним актором в межах дослідження). Емоційні вирази обличчя та верхньої частини тіла мають високі показники розпізнавання за висновками опитування фокус-груп.

Отже, моделювання роботів, які здатні відображати емоції вже реальність. Численні теоретичні доробки сприяли віднайденню найефективнішої (на сьогодні) моделі, якими послуговуються розробники для втілення своїх задумів. Найближче, з широко відомих течій, підійшов енактивізм з його ідеями втілених дій. Так, роботи зараз мають обмежений спектр вираження емоцій і задається він детальним програмуванням у відповідь на зовнішні подразники, будь то тон голосу, або відстань до об'єкта. Але це тільки перші, і варто відмітити, дуже вдалі кроки на шляху до створення повноцінної технологічної форми життя. І не варто перейматися на рахунок того, що це не «справжні», а запрограмовані переживання. Насправді у живих істот переживання також «запрограмовані», як відповідь на подразники з навколишнього середовища, просто більш деталізовані та через швидкість реакцій – майже не помітні у процесі обробки. Але «програма» біологічна, або технологічна залишається аналогічною і тільки питання часу, коли «вони» стануть такими ж розвинутими як «ми». Тим більше, наділення роботів емоціями сприятиме їх кращій адаптації у соціальному середовищі (де, вони, до речі, вже функціонують) і більшій адекватності поведінки. Позбавить скептиків, що бояться «холодних бездушних машин» їх основного побоювання, адже з емоціями роботи перестануть бути холодними і бездушними.

І наостанок, варто зазначити, що емоції можливо змоделювати лише у фізичного робота, тобто такого, що має тіло. Адже без тіла – немає досвіду, немає реакцій. І неможливо буде сконструювати форму життя зі свідомістю. Але саме розробники емоційних роботів ідуть тут у правильному напрямку.

References:

1. Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press. [in English].
2. Damasio, A. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando, FL: Harcourt. [in English].
3. DeLancey, C. (2002). *Passionate Engines: What Emotions Reveal about Mind and Artificial Intelligence*. Oxford: Oxford University Press. [in English].

¹ Там само.

² Endo (2010). Integration of Emotion Expression and Visual Tracking Locomotion Based on Vestibulo-Ocular Reflex. *Proc. of the 19th IEEE International Symposium in Robot and Human Interactive Communication*, 558-563.

4. Ekman, P. (1999). Basic Emotions. in Tim Dalgleish and Mick J. Power (eds.). *Handbook of Cognition and Emotion*. Chichester, UK: Wiley and Sons. doi:10.1002/0470013494.ch3 [in English].
5. Ekman, P., Cordaro D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, 3(4) doi:10.1177/1754073911410740 [in English].
6. Endo (2010). Integration of Emotion Expression and Visual Tracking Locomotion Based on Vestibulo-Ocular Reflex. *Proc. of the 19th IEEE International Symposium in Robot and Human Interactive Communication*. [in English].
7. Griffiths, P., Scarantino, A. (2009). Emotions in the Wild: the Situated Perspective on Emotion. in Philip Robbins and Murat Aydede (eds.). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511816826.023. [in English].
8. Hall, E. (1996). Hidden Dimension. *Doubleday Publishing*. ISBN 0-385-08476-5.
9. Hashimoto, T., Hitramatsu, S., Tsuji, T., Kobayashi, H. (2006). Development of the Face Robot SAYA for Rich Facial Expressions. *Proc. of SICE-ICASE International Joint Conference*. [in English].
10. Hufendiek, R. (2016). *Embodied Emotions: A Naturalist Approach to a Normative Phenomenon*. London: Routledge. [in English].
11. Hutto, D. (2012). Truly Enactive Emotion. *Emotion Review*, 4(2). doi:10.1177/1754073911430134. [in English].
12. Ishiguro, S. N. (2007). Building artificial humans to understand humans. *Journal of Artificial Organs*, 10(3). [in English].
13. Breazeal, C. (2020). Kismet, the robot. *Kismet*. < <http://www.ai.mit.edu/projects/sociable/baby-bits.html>> (2020, August, 13). [in English].
14. YouTube (2020). *BBC's Horizon. Leonardo Robot – isn't this the cutest robot ever?* <https://www.youtube.com/watch?v=ilmDN2e_Flc> (2020, August, 13). [in English].
15. McGann, M., De Jaegher, H.; Di Paolo, E. (2013). Enaction and psychology. *Review of General Psychology*, 17 (2). doi:10.1037/a0032935. [in English].
16. Miwa, H. (2001). Robot Personality based on Equations of Emotion defined in the 3D Mental Space. *Proc. Of IEEE International Conference on Robotics and Automation*. [in English].
17. Robson, D. (2015) What is like to have never felt an emotion? *BBC Future*. <<https://www.bbc.com/future/article/20150818-what-is-it-like-to-have-never-felt-an-emotion>> (2020, August, 13). [in English].
18. Russell J. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110 (1). doi:10.1037/0033-295X.110.1.145: 147 [in English].
19. Russell, J. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39 (6). doi:10.1037/h0077714 [in English].
20. Shargel, D., Prinz J. (2018). An Enactivist Theory of Emotional Content. in *Hichem Naar and Fabrice Teroni (eds.), The Ontology of Emotions, Cambridge: Cambridge University Press*. doi:10.1017/9781316275221.007 [in English].
21. Thomaz, M., Berlin, C. B. (2005). An Embodied Computational Model of Social Referencing. *Proc of the 14th IEEE International Symposium in Robot and Human Interactive Communication, Workshop*. [in English].
22. Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press. [in English].
23. Tomkins, S. S. (2008). *Affect Imagery Consciousness: The Complete Edition*. New York: Springer Publishing Company, I, II, III, IV,. [in English].
24. Shynkaruk, V. (2002). *Filosofskyj encyklopedychnyj slovnyk* [Philosophical encyclopedic dictionary]. Kyiv: Instytut filosofiyi imeni Grygoriya Skovorody NAN Ukrainyiny: Abry's. [Philosophical encyclopedic dictionary.] [in Ukrainian].