

МОДЕЛЬ ІНТЕЛЕКТУАЛЬНОГО АГЕНТА В ОДНОМІРНОМУ ДИСКРЕТНОМУ СВІТІ

В даній роботі на основі простої моделі взаємодії агента зі світом розглядаються питання складності та інтелекту. Як компроміс між ентропією та складністю Колмогорова пропонується використовувати сукупну ентропію n -грам, яка не містить недоліків першої та може бути поражена точно, на відмінно від другої. Показано, що варіюючи дану міру можна переходити від станів повністю випадкового світу до повністю організованого з поступовою зміною складності. Приймаючи за мету агента робити найкращі передбачення, можна співвіднести міру інтелекту агента та до складності світу. Наводиться короткий огляд невирішених проблем моделювання інтелекту агента та можливих шляхів їх вирішення.

In present work questions about complexity and intelligence considered based on the simple model of agent interaction with the environment. As a compromise between entropy and Kolmogorov complexity, it is proposed to use the joint entropy of n -grams, that does not have the shortcomings of the first can be computed, unlike the second one. It is shown that by varying proposed measure it is possible to switch from fully random states to organized with gradual levels of complexity. Accepting making prediction as the main goal of an agent it is possible to bind the level of agent's intelligence and world complexity. A short overview of unsolved problems of modeling agent's intelligence is presented and possible ways of solutions are outlined.

Ключові слова: інтелектуальний агент, нейронні мережі, складність Колмогорова, ентропія

Вступ

Серії робіт Тюрінга[1], Шеннона[2] та інших 40-х- 60-х років заклали основи теорії інформації та обчислення, що дало змогу формалізувати такі поняття як складність та міра інформації. В той самий час почався рух зі створення інтелектуальних машин, які б змогли навчатися та виконувати складні задачі як люди. Але ця задача виявилася надзвичайно складною і досі невирішена. Не дивлячись на нещодавній прогрес у сфері глибоких нейронних мереж для розпізнавання образів та навчання з підкріпленням[3], досі немає чіткого визначення інтелекту та розуміння шляху його відтворення в машині.

Дана робота підтримує підхід, що потрібно звести складність задачі до мінімуму зберігши тільки істотні риси і вивести мінімальні необхідні властивості системи для розумної поведінки[4]. При цьому важливо певним чином характеризувати складність середовища та агента у ньому.

Для характеристики складності середовища пропонується компромісне рішення між ентропією та складністю Колмогорова [5] - сукупна міра ентропій n -грам. Дана міра показує організованість середовища, виявляє локальні структури і може бути обчислена точно. Показано, що варіювання запропонованої міри породжує середовище

різної складності, від повністю випадкового до повністю регулярного.

Розглядаються кілька варіантів агентів з різним рівнем інтелекту. За визначення інтелекту приймається здатність агента досягати своїх цілей, яка й визначає складність агента. Для простоти розглядається лише одна ціль, але така, що охоплює широкий клас задач – робити якнайкращі передбачення станів світу. Для цього агенту потрібно будувати ймовірнісну модель світу та виявляти в ньому організовані структури. Наводяться деякі з підходів побудови такої моделі та звертається увага на перспективні ідеї принципів обробки інформації, що походять з останніх досліджень біологічних нейронних мереж.

Основна ідея даної статті поєднати складність світу зі складністю агента, формалізувати поняття складності та окреслити підходи та проблеми створення моделі світу агентом.

Постановка задачі

Пропонується модель взаємодії агента з середовищем, де інтелектуальність агента визначається через здатність досягнення власних цілей. Для простоти світ вибраний одномірний, статичний, дискретний де кожний стан може приймати два значення, 0 або 1. Агент знаходиться в довільному положенні і може рухатися вправо або вліво по рядку,

схематично це зображено на Рис.1. Була вибрана проста мета агента - робити якнайкращі передбачення станів світу.

Основні позначення

Розглянемо пару (W, I) світ-агент. Ансамбль $W = \{x_i, A\}$ характеризує світ, де $x_i \in A, i = 1: N$ - сукупність станів з бінарного алфавіту $A = \{0,1\}$, тобто світ це бінарний вектор розміром N . Агент описується як сукупність сенсорів, моторики та обчислювальної системи $I = \{S, D, AI\}$. Для простоти вважаємо, що агент в положенні i може сприймати лише стан світу x_i тобто $S = \{\hat{x}_r, r = 1\}$ де \hat{x}_r сприймання світу без шуму з рецептивним полем $r = 1$. Рухатися агент може вліво, вправо або стояти на місці $D = \{y\}, y \in Y = \{l, r, 0\}$. AI обчислювальна система агента, яка обробляє вхідні дані, містить пам'ять про світ, керує моторикою.



Рис. 1. Схематичне зображення агента в одному бінарному світі.

Прийmemo, що основна задача агента робити найкраще передбачення наступного сенсорного входу

$$\tilde{x}_{i+1} = \arg \max_{x_{i+1}} (p(x_{i+1} | y_{i+1}, x_i, y_i, x_{i-1}, y_{i-1}, \dots, x_{i-p+1}, y_{i-p+1}))$$

В загальному випадку, наступний стан залежить від попередньої історії в минулі p кроків. Якщо $p = 1$, то

$$\tilde{x}_{i+1} = \arg \max_{x_{i+1}} (p(x_{i+1} | y_{i+1}, x_i)),$$

і задача зводиться до марківського процесу прийняття рішень. Проте, для успішного передбачення потрібно відслідковувати історію далі, ніж на один крок. Тому, спростивши моторику агента до руху тільки в одну сторону, можна записати

$$\tilde{x}_{i+1} = \arg \max_{x_{i+1}} (p(x_{i+1} | x_i, x_{i-1}, \dots, x_{i-p+1})).$$

Більш інтелектуальний агент буде робити передбачення, що відповідають справжньому стану світу $\tilde{x}_{i+1} = x_{i+1}$. Для цього агент повинен вивчити спільний розподіл ймовірностей світу $p(x_{i+1}, x_i, x_{i-1}, \dots, x_{i-p+1})$ для сусідніх p станів (сусідніх, так як ми приймаємо, що агент рухається локально, не перестрибує).

В наступному розділі розглядається складність світу, яка характеризується спільним

розподілом ймовірностей на різних масштабах та яким чином агент може вивчати розподіл.

Результати

Складність світу.

Успішне передбачення залежить від сенсорів, моторики та обчислювальної системи агента, але також і від самого світу. Якщо світ повністю випадковий, то розподіл ймовірності буде рівномірний і агент не зможе вивчити регулярності, яким би він розумним не був і, все одно, буде робити помилкові передбачення. Тому важливо характеризувати світ, який у нашому випадку описується бінарним вектором.

Розглянемо дві міри бінарного вектору, ентропію та складність Колмогорова. Ентропія Шеннона визначається наступним чином:

$$H(x) = - \sum_{i=1:N} p(x_i) \log_2(p(x_i))$$

де $x_i \in A, A = \{0,1\}$ - бінарний алфавіт вектору, та $p(x_i)$ ймовірність появи символу x_i . Якщо кількість одиниць та нулів однакова, то $p(x_i = 1) = p(x_i = 0) = 0.5$ і ентропія максимальна $H(x) = 1$. Недоліком даної міри є те, що вона не чутлива до порядку нулів та одиниць, до внутрішньої структури рядка. Наприклад, бінарні вектори $v = [0010101110]$, $w = [1111100000]$ мають однакову ентропію $H(v) = H(w) = 1$, але очевидно упорядкованість другого вища.

Складність Колмогорова визначається як довжина мінімальної програми, яка відтворює рядок:

$$K(x) = \min_p (|p| : T(p) = x),$$

де p програма, $T(\cdot)$ машина Тюрінга, яка виконує програму та відтворює рядок. Дана міра краще описує складність, наприклад, вектор w можна було б описати програмою «повтори 5 раз 1 та 5 раз 0», що коротше ніж визначати 0 чи 1 для кожного положення у випадку випадкового вектору. Але проблема даної міри в тому, що її не можна точно обчислити. Для довільного рядка не можна знайти мінімальну програму і порахувати її довжину. Роботи Соломоноффа [6] наклали верхню межу, але, все одно, складність Колмогорова на практиці не застосовна. Наприклад, не можна порівняти складність текстів, зображень чи послідовностей ДНК.

Як компроміс між двома мірами пропонується використовувати сукупну ентропію n -грам рядку. N -грама визначається як $p(x_i | x_{i-1}, \dots, x_{i-n+1})$ - ймовірність появи елементу x_i за умови появи попередніх $n - 1$ елементів. Нехай $p_n(x_i) = p(x_i, x_{i-1}, \dots, x_{i-n+1})$ спільна ентропія n -грами і введемо ентропію порядку n як

$$H_n(x) = \sum_{i=n:N} p_n(x_i) \log_2 p_n(x_i),$$

що подібне до загальноприйнятого означення швидкості ентропії. Складність вектору порядку n можна визначати як

$$C_n(x) = \frac{2}{n(n+1)} \sum_{i=1:n} H_i(x).$$

Множник перед сумою для масштабування максимального значення до одиниці

$$\begin{aligned} \max(C_n(x)) &= \frac{2}{n(n+1)} \max(\sum_{i=1:n} H_i(x)) = \\ &= \frac{2}{n(n+1)} \sum_{i=1:n} \max(H_i(x)) = \frac{2}{n(n+1)} \sum_{i=1:n} 1 \end{aligned}$$

Враховування ентропій n -грам допомагає відслідковувати просторові співвідношення елементів у рядку. Наприклад, $C_2(x)$ показує наскільки часто зустрічаються паттерни типу 01,00,11,10. Розглянемо вектор $[01]^*1000$, яка складається з повторів 01 тисячу разів, в таблиці 1 наведені значення складності різних порядків

Табл. 1. Різні порядки складності.

| Порядок складності | Складність |
|--|------------|
| 1, $C_1(x) = H(x)$ відповідає звичайній ентропії | 1 |
| 2 | 0,66 |
| 5 | 0,33 |
| 10 | 0,18 |

Так як для такого регулярного рядку $H_i(x) = 1, \forall i$ то $C_n(x) \xrightarrow{n \rightarrow \infty} 0$. Для повністю випадкового рядку $C_n(x) \xrightarrow{n \rightarrow \infty} 1$, при тому що стандартна ентропія для обох випадків стала і дорівнює одиниці.

Складність порядку n вловлює лише локальні взаємозалежності в околі n , і тому складність Колмогорова є більш універсальною, проте не обчислювальною. В загальному, перевага запропонованого методу в тому, що упорядкованість світу зменшує складність, яку можна охарактеризувати числом. Наприклад, на

Рис.2 показано як почавши з випадкового бінарного вектору з біноміального розподілу, можна, виконуючи умову зменшення складності, структурувати вектор, при цьому залишаючи ентропію постійною. Як видно з рисунка рядок еволюціонує до регулярного з повторенням «10», для якого складність мінімальна. Дана оптимізація відрізняється від оптимізації $H_n(x)$ в тому, що більше стимулює утворення структур на різних рівнях n .

Також цікавою властивістю є трансляційна інваріантність даної міри, тобто не важливо в якому положенні були організовані структури.

Підсумовуючи, пропонується визначати складність світу через його передбачуваність, і скільки ресурсів треба на це витратити.

Властивості $A1$

Розглянемо трьох агентів I_1, I_2, I_3 в порядку збільшення інтелекту. Перший агент I_1 - найпримітивніший, він здатний робити передбачення тільки на основі $p(x_i)$ і не може вловити залежності з сусідніми станами. Другий відслідковує появи пар, тобто $p(x_i, x_{i+1})$ і у випадку структурованого середовища $C_2(x) < 1$ буде робити передбачення набагато краще. Третій агент I_3 відслідковує паттерни на більшому масштабі $p(x_i, \dots, x_{i+10})$ і буде ще більш ефективний ніж попередній. Інтелект агента залежить наскільки його розподіл ймовірності наближається до розподілу ймовірностей світу $p_i(x) \rightarrow p_w(x)$. Очевидно, складність світу накладає обмеження зверху на складність агента. Важливо, що для покращення навчання розподілу ймовірностей окрім обчислювальної системи агента, можна ще покращувати сенсори та моторику. Наприклад, якби рецептивне поле агента було більшим, він зміг би бачити довші залежності, або якби вмів скакати, не був би обмежений локальністю.

Загалом, побудова обчислювальної системи агента зводиться до передбачення послідовностей. Це загальна задача [4], яка включає в себе класифікацію (тоді \tilde{x}_{i+1} це передбачення мітки на основі $y_{i+1}, x_i, y_i, x_{i-1}, y_{i-1}, \dots, x_{i-p+1}, y_{i-p+1}$ попередніх пар дані-мітки) та навчання з підкріпленням [7]. Вже розроблено багато методів вирішення даної проблеми, від прихованих марківських

моделей, баєсівських моделей, до глибоких нейронних мереж з LSTM модулями, що показують найкращу ефективність на даний момент. Проте жодний з методів не працює так добре як людський мозок, більшою мірою через важкість передбачати події між якими великий проміжок часу, а також через проблему великої розмірності спільного розподілу ймовірностей при переході до більш складних середовищ.



Також, методи навчання з учителем вимагають багато даних на яких навчаються параметри моделі і не працюють ітеративно. Тобто, агент, досліджуючи середовище, має лише оновлювати свою модель, а не перенавчати її, як у випадку класичних штучних нейронних мереж.

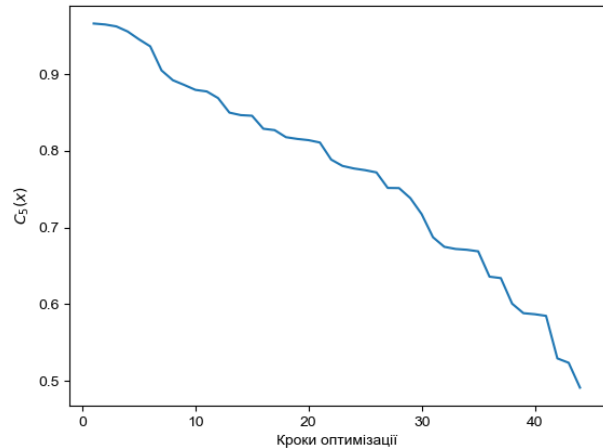


Рис. 2. Зліва показана еволюція рядку від повністю випадкового до організованого. По вертикальній осі показані кроки оптимізації. Як видно рядок наближається до регулярного паттерну з повторами «10». Справа показана залежність складності від кроків оптимізації для рядка наведеного зліва. Починається з майже максимального значення 1 і зменшується до менше ніж 0.5

Обмеженість ресурсів для складних світів

Бінарний одномірний світ доволі простий і наведені методи дозволяють добре оцінювати спільний розподіл ймовірностей. Якщо трохи ускладнити світ, припустимо, до світу зображень (двомірний, 256 станів(кольорів) на клітину) та можливість рухатися в усі 8 сторін, то розмірність спільного розподілу збільшується до $8 \cdot 256^T$, T – кількість кроків назад. Якщо ще додати динамічність(відео), то розмірність збільшується ще більше і існуючі методи вже не можуть робити добре передбачення. Реальний світ ще складніший і єдиний приклад системи, яка здатна моделювати багатомірний розподіл ймовірностей при обмежених ресурсах це біологічні нейронні мережі.

За останні 20 років розділ науки, що вивчає мозок – нейронаука, досягла значного прогресу в розумінні певних принципів обробки інформації в живих організмах. Хоча ще залишається багато запитань, але прогрес в даному напрямку відчутний, і далі наводяться

деякі вибрані ідеї, які є найбільш перспективними для дослідження.

Біологічні нейронні мережі використовують розподілене розріджене кодування [8]–[10]. Тобто, інформація від сенсорних органів представлена в активації малої кількості нейронів розподілених по значній частині кори мозку. Як було показано, таке кодування збільшує ємність асоціативної пам'яті[11] використовуючи мінімум ресурсів, а також, гіпотетично, слугує необхідною умовою відтворення розподілу ймовірностей в мережі [12]. Активність одного нейрона не можна декодувати в певний стимул, інформація закодована в популяції. Наприклад, в мережі з 1000 нейронів стимул може бути закодований активністю 20-ьох нейронів. Кожний окремий нейрон може приймати участь в кодування багатьох стимулів. Таким чином збільшується ємність мережі на відміну від локального представлення, коли один нейрон кодує один об'єкт. Кількість можливих активних груп в простому прикладі надзвичайно велика $C_{1000}^{20} \approx 10^{41}$, проте не всі вони реалізуються, і багато є близькими одна до одної. Для більших

популяції через кількість авто асоціативних зв'язків в групі або їх силу можна охарактеризувати ймовірність появи події закодованої даною групою. Завдяки нейронам інгібіторам, які залишають активною найбільш зв'язану групу, вибирається найймовірніше представлення мережі[13]. Якщо надходить негативний зворотний сигнал з вищих рівнів ієрархії, вибирається друга найбільш активна група. Таким чином можливе перебирання гіпотез, у паралельній архітектурі, до поки не буде досягнуто рішення, що задовольняє агента[14].

Іншою перевагою біологічних нейронних мереж, що вони здатні представляти багатовимірний розподіл ймовірностей реального світу [15]–[17], ключовою характеристикою якого є його надзвичайна розрідженість. Зі всіх можливих комбінацій станів світу тільки дуже мала частина насправді реалізується і може бути сприйнятою твариною за своє життя. Комбінаторна природа активації нейронних мереж може закодувати величезну кількість стимулів, і через розподілене представлення зв'язати їх разом, при цьому сила зв'язку або кількість зв'язків визначатимуть ймовірність асоціації. Такий підхід добре описує роль нейромодуляторів, таких як допамін чи серотонін, які змінюють представлення розподілу ймовірностей відповідно до цілей агента. Зміна швидкості утворення зв'язків реалізує навчання з підкріпленням від середовища чи від внутрішніх факторів.

Також, в навчанні агента, важливу роль грає представлення новизни чи здивування [18]. У випадку новизни, агент не має попередніх уявлень про стан світу, тобто має рівномірний апріорний розподіл. Уточнення такого розподілу є корисним для агента, і, наприклад, у дітей це проявляється через сигнал задоволення. Здивування означає, що передбачення стану світу було не вірним, і має бути оновленим. Ці два підходи на перший погляд є схожі, так як оновлюють представлення про світ, проте в мозку представлені двома різними системами, перша, через використання нейромедіаторів з підкортикальних областей, друга, через сигнал помилки з верхніх рівнів ієрархії.

Загалом, біологічні нейронні мережі дають багато корисних ідей, яким чином побудувати нові когнітивні системи, які б змогли в режимі реального часу, з порівняно малими затратами

ресурсів, створювати модель свого середовища і досягати своїх цілей.

Обговорення та висновки

В даній роботі розглядається проста система взаємодії агента зі світом, пропонується нове поняття складності світу, визначається основна мета агента та робиться короткий огляд ідей як ця мета може задовольнятися.

Для визначення складності світу пропонується рахувати ентропію n -грам, яка враховує організованість світу на локальному рівні. Показано, що мінімізуючи дану міру, можна перейти від повністю випадкового стану до повністю регулярного, а в проміжних положеннях знаходяться проміжні стани складності. Дана міра виступає як компроміс між стандартною ентропією та складністю Колмогорова. Недоліком даної міри є складність її обрахунку для середовищ з багатьма станами. Загалом, складність світу визначається наскільки він передбачуваний.

Можливість робити передбачення взято за основну мету агента, і чим успішніший він у цьому, тим більш він інтелектуальніший. Узгодженість даної мети з запропонованою складністю дозволяє співвіднести складність світу з обчислювальною здатністю агента.

Спростивши задачу агента до передбачення послідовностей, робиться короткий огляд методів вирішення даної проблеми. Наводяться основні труднощі існуючих алгоритмів, що пов'язані з багатовимірним розподілом ймовірностей складних середовищ, і необхідністю адаптивного та ітеративного навчання моделі агента.

Для подолання даних труднощів, пропонується звернутися до ідей з нейронауки та принципів обробки інформації біологічними нейронними мережами. Вони включають розріджене розподілене кодування, утворення асоціативних зв'язків для розріджених станів світу та відтворення розподілу ймовірностей через кількість зв'язків в популяції нейронів. Цей підхід вважається перспективними, але для практичного застосування та побудови інтелектуального агента в складних середовищах потрібно проводити додаткові практичні та теоретичні дослідження.

Список посилань

1. A. Turing, "On computable numbers," *Proc. London Math. Soc.*, vol. 42, pp. 230–265, 1936.
2. C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. July 1928, pp. 379–423, 1948.
3. J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
4. M. Hutter, *Universal artificial intelligence*. 2005.
5. A. Kolmogorov, "Three Approaches to the quantitative definition of Information," *Prob Info Trans*, vol. 1, no. 1, pp. 3–11, 1965.
6. R. J. Solomonoff, "A Formal theory of Inductive Inference," *Info Cont*, vol. 7, pp. 1-22-254, 1964.
7. F. Wörgötter and B. Porr, "Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms," *Neural Comput.*, vol. 17, no. 2, pp. 245–319, 2005.
8. P. Kanerva, "Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors," *Cognit. Comput.*, vol. 1, no. 2, pp. 139–159, 2009.
9. A. Luczak, B. L. McNaughton, and K. D. Harris, "Packet-based communication in the cortex.," *Nat. Rev. Neurosci.*, vol. 16, no. 12, pp. 745–755, 2015.
10. K. D. Harris and G. M. G. Shepherd, "The neocortical circuit: themes and variations," *Nat. Neurosci.*, vol. 18, no. 2, pp. 170–181, 2015.
11. G. Palm, "Neural associative memories and sparse coding," *Neural Networks*, vol. 37, pp. 165–171, 2013.
12. H. Barlow, "Redundancy reduction revisited," *Netw. Comput. Neural Syst.*, vol. 12, no. 3, pp. 241–253, 2001.
13. K. Morita, J. Jitsev, and A. Morrison, "Corticostriatal circuit mechanisms of value-based action selection: Implementation of reinforcement learning algorithms and beyond," *Behav. Brain Res.*, vol. 311, pp. 110–121, 2016.
14. M. B. Mirza, R. A. Adams, C. D. Mathys, and K. J. Friston, "Scene Construction, Visual Foraging, and Active Inference," *Front. Comput. Neurosci.*, vol. 10, no. June, 2016.
15. D. C. Knill and A. Pouget, "The Bayesian brain: The role of uncertainty in neural coding and computation," *Trends Neurosci.*, vol. 27, no. 12, pp. 712–719, 2004.
16. C. L. Baker, R. R. Saxe, and J. B. Tenenbaum, "Bayesian Theory of Mind: Modeling Joint Belief-Desire Attribution," *Proc. thirty-second Annu. Conf. Cogn. Sci. Soc.*, vol. 1, no. 2006, pp. 2469–2474, 2009.
17. J. T. Abbott, J. B. Hamrick, and T. L. Griffiths, "Approximating Bayesian inference with a sparse distributed memory system," *Proc. 35th Annu. Conf. Cogn. Sci. Soc.*, pp. 1686–1691, 2013.
18. N. Frémaux and W. Gerstner, "Neuromodulated Spike-Timing-Dependent Plasticity and Theory of Three-Factor Learning Rules," *Front. Neural Circuits*, vol. 9, no. 85, p. 85, 2016.