

ФІЛОСОФСЬКІ АСПЕКТИ ШТУЧНОГО ІНТЕЛЕКТУ

Чала О.С.

Науковий керівник – д.ф.н, проф. Штанько В.І.

Харківський національний університет радіоелектроніки, каф. філософії,
м.Харків, Україна

тел. (057) 702-14-65, e-mail: olha.chala@nure.ua

Artificial intelligence is a branch of science that declares the creation of computer systems that are similar in behavior to humans. Today, the key attention in the work of artificial intelligence is paid to the construction of intelligent agents. However, such agents do not solve the problem of building artificial subjective consciousness and creativity. One of the important problems that has not yet been solved in the creation of intellectual agents is the problem of "moral" artificial intelligence. Its solution requires the development of moral considerations based on the moral code.

Штучний інтелект є галуззю науки, що декларує створення комп'ютерних систем, схожих за своєю поведінкою на людей. Термін «штучний інтелект» було запропоновано у 1956 році на конференції у Дартмутському коледжі в Ганновері для позначення комп'ютерних систем, що у віддаленій перспективі зможуть мислити схожим на людину способом.

Сучасні підходи до побудови штучного інтелекту орієнтовані на створення «сильного» та «слабкого» штучного інтелекту. Метою «сильного» штучного інтелекту є створення машин, які мають схожі з людьми розумові здібності. В даному напрямку на сьогодні проводяться дебати на тему «Чи може машина досягти інтелекту людини?». В цій дискусії було висловлено припущення про існування точки сингулярності, в якій інтелектуальні системи досягнуть рівня людського інтелекту. Проте досягнення цієї точки потребує враховувати моральні аспекти та забезпечити контроль людства над машинами. Тому на сьогодні обговорюються ідеї побудови спеціалізованих інтелектуальних систем, щоб забезпечити незалежність людства від штучного інтелекту.

Машини зі «слабким» штучним інтелектом здатні пройти тест Тьюринга. Даний тест полягає у порівнянні анонімних відповідей на запитання від комп'ютера та людини. Система штучного інтелекту проходить текст, якщо неможливо визначити, хто сформував відповідь: комп'ютер або людина. На сьогодні системи «слабкого» штучного інтелекту знайшли широке застосування для підтримки взаємодії людей у кіберпросторі [1], [2, с. 9] та при вирішенні локальних, але практично

важливих задач (розпізнавання зображень, автоматичного перекладу текстів, обчислювальної фотографії, тощо).

Сучасний штучний інтелект широко використовує ймовірнісний підхід, який удосконалювався Блезом Паскалем. У відповідності до його ідей виконується висновування на послідовності ймовірнісних логічних правил. Сукупність застосованих правил використовується для обчислення ймовірності істинності результату висновування, що дає можливість порівняти декілька альтернативних результатів.

На сьогодні ключова увага у роботах, присвячених проблемам штучного інтелекту, приділяється побудові інтелектуальних агентів. Ці агенти отримують інформацію із навколишнього середовища, приймають рішення та виконують дії у відповідності до заданої задачі. Такі агенти використовують внутрішню модель світу та виконують імовірнісне висновування для прийняття рішення і виконання дій. Агенти можуть навчатись. За результатами навчання їхня поведінка змінюється.

Однією з важливих проблем, яка поки що не знайшла свого вирішення при створенні інтелектуальних агентів, є проблема «морального» штучного інтелекту. Загальна потреба у вирішенні цієї проблеми виникла при побудові інтелектуальних автопілотів для автомобілів. В даному випадку розглядається, наприклад, питання: «Чи може автопілот збити пішохода, який переходить дорогу у недозволеному місці, щоб вберегти пасажирів від наслідків зіткнення із іншим автомобілем?» Як формальний апарат для вирішення цієї задачі використовується деонтична логіка, яка описує норми та нормативні поняття. Норми дають змогу або зобов'язують виконати певні дії, а також утримуватись від заборонених дій. Логіка дає можливість описати санкції та штрафи за виконання таких заборонених дій. Однак в цілому дана область ще знаходиться на початкових стадіях розвитку.

Список використаних джерел:

1. Штанько В.И. (2014). Особенности самопрезентации и самоидентификации личности в киберпространстве. Вісник Харківського національного університету імені В. Н. Каразіна: «Теорія культури і філософія науки», 1092(50), 1-8. <https://periodicals.karazin.ua/thcphs/article/view/2179>.

2. Штанько В. И. (2007). Проблема человека в эпоху информатизации: новые возможности и опасности. Вісник Харківського національного університету ім. В.Н. Каразіна: «Теорія культури і філософія науки», 764(2), 3-10. <https://openarchive.nure.ua/handle/document/7092?locale=en>