УДК 004.8
DOI: 10.18523/2617-3808.2024.7.70-76

Глибовець М. М., Бачинський Н. П.

УЗАГАЛЬНЕНИЙ АГЕНТ ШТУЧНОГО ІНТЕЛЕКТУ SIMA

Стаття аналізує проєкт SIMA (Scalable, Instructable, Multiworld Agent) від Google DeepMind, спрямований на створення узагальненого ШІ-агента, здатного виконувати завдання у різноманітних тривимірних середовищах за мовними інструкціями. Розглянуто методи навчання, оцінювання та особливості роботи агента, а також проведено порівняння з іншими проєктами, такими як OpenAI Five і AlphaStar. Попри досягнуті результати, висвітлено ключові виклики, зокрема технічні та етичні аспекти, що залишаються на шляху до створення загального штучного інтелекту.

Ключові слова: штучний інтелект, агент, SIMA, віртуальне середовище.

Вступ

Створення узагальненого агента штучного інтелекту як часткового випадку загального ШІ (AGI, універсального вирішувача задач) є однією з найскладніших задач сучасної науки. Створення такого агента обов'язково передбачає здатність засвоювати нові навички без спеціального програмування, адаптуватися до незнайомих ситуацій і ухвалювати ефективні рішення. Задоволення цих вимог потребує значних технічних проривів та обчислювальних ресурсів.

Розроблення універсального вирішувача супроводжується етичними викликами, зокрема йдеться про безпеку, ризики використання та відповідальне впровадження. Перспективи його застосування охоплюють наукові дослідження в різноманітних галузях (медицина, освітні інновації, оптимізація виробництва).

Серед важливих сучасних досягнень у цій царині — проєкт **SIMA** (Scalable, Instructable, Multiworld Agent) від Google DeepMind [3]. Основною складовою проєкта є розроблення спеціального агента, який наділений здатністю переносити знання між віртуальними середовищами. В майбутньому апробовані технології можуть бути інтегровані у фізичні системи для застосування у реальному світі.

Коротко охарактеризуємо історію попередніх досліджень у сфері створення універсального агента, зокрема розглянемо деякі агенти та проєкти, які були розроблені в останні роки та показали цікаві результати.

Відомо, що відеоігри є важливим середовищем для дослідження та розроблення універсальних ШІ-агентів. Вони дають змогу тестувати алгоритми, здатні адаптуватися до різноманітних завдань без необхідності спеціального програмування для кожної задачі. Багатство ігрових світів забезпечує умови для вивчення адаптивності, навчання з досвідом та розвиток стратегічного мислення агентів.

У 2019 році компанія DeepMind представила AlphaStar [2] — агента штучного інтелекту, який досяг рівня професійних гравців у складній стратегії в реальному часі StarCraft II. У цій грі гравці мають керувати ресурсами, будувати бази, створювати армії та стратегічно планувати дії проти суперника. AlphaStar спочатку навчався на записах ігор професійних гравців, щоб засвоїти базові стратегії та тактики. Потім він грав мільйони ігор проти себе та інших версій агентів, покращуючи свої навички через досвід.

AlphaStar продемонстрував здатність до довгострокового планування, адаптації до стилю гри опонента та управління багатьма юнітами одночасно, що є важливими кроками до створення AGI.

У 2019 році OpenAI розробила OpenAI Five [1], команду з п'яти агентів штучного інтелекту, які успішно змагалися проти професійних гравців у грі Dota 2 — складній багатокористувацькій онлайн-баталії. Агенти навчалися, граючи проти себе у швидко симульованому середовищі, що дало їм змогу накопичити тисячі років ігрового досвіду за короткий час.

Ключові аспекти цього досягнення такі. OpenAI Five демонстрували здатність до командної роботи, координуючи дії між п'ятьма агентами без явного обміну інформацією. Вони могли швидко реагувати на стратегії суперників і змінювати свою тактику в реальному часі.

Цей проєкт підкреслив можливість створення агентів, здатних до складної співпраці та стратегії в динамічному середовищі.

Дослідницькі середовища, такі як *AI2-THOR*, *CARLA*, *MuJoCo*, *AI Habitat* та *ALFRED*, слугують платформами для тестування та імплементації агентів. Вони дають змогу моделювати конкретні життєві ситуації та взаємодію з об'єктами, забезпечуючи зручний інтерфейс для навчання агентів. Однак їхні можливості обмежені специфічними сценаріями, що створює виклики для генералізації навичок агентів до реальних умов.

Кінцевою метою розроблення узагальненого агента є його інтеграція у роботизовані системи для роботи у реальному світі. Попри складність цього завдання, вже існують роботи, які демонструють успішний перенос моделей, навчених у симульованих середовищах, до фізичних систем. Це підтверджує потенціал застосування подібних агентів у робототехніці, зокрема для виконання складних завдань у неконтрольованих умовах реального світу.

Google DeepMind, спираючись на багаторічний досвід, представила нового агента — Scalable, Instructable, Multiworld Agent (SIMA) [3]. Цей проєкт є важливим етапом у напрямі розвитку загального штучного інтелекту, відображаючи прогрес, досягнутий у галузі за останні роки.

У цій статті викладено висновки аналізу універсального агента SIMA, який продемонстрував значний прогрес у перенесенні навичок між різними середовищами та zero-shot навички, що дають змогу виконувати базові дії в нових середовищах без попереднього навчання.

SIMA

Основною метою проєкту SIMA є створення агента, здатного виконувати довільні мовні інструкції для здійснення дій у будь-якому віртуальному тривимірному середовищі, використовуючи клавіатуру та мишку. До таких середовищ належать як спеціально розроблені дослідницькі платформи, так і широкий спектр комерційних відеоігор. SIMA прагне подолати розрив між символами мови та їхніми зовнішніми референтами, об'єднуючи абстрактне розуміння мови зі втіленим сприйняттям і дією в масштабованих середовищах.

Серед досягнень SIMA — створення агента, який може виконувати короткострокові завдання на основі мовних інструкцій у більш ніж десяти різноманітних 3D-середовищах. Для оцінювання продуктивності SIMA розроблено кілька методів, зокрема оптичне розпізнавання символів (OCR) для виявлення тексту, що сигналізує про завершення завдань, а також людське оцінювання записаних відео з діями агента. Використання єдиного інтерфейсу для різних середовищ дає агентові змогу працювати без додаткового налаштування під нові умови, що є кроком у напрямі створення універсального агента, здатного виконувати широкий спектр дій, доступних людині, у симульованих 3D-середовищах.

Агент SIMA вирізняється серед інших тим, що сфокусований на роботі з природною мовою у різноманітних середовищах, які відрізняються як візуальним оформленням, так і доступними діями. Це допомагає агентові демонструвати високу адаптивність та здатність до генералізації знань у різних контекстах.

Для забезпечення різноманітності досвіду агента було використано набір віртуальних середовищ, що передбачають як симуляції реального світу, так і міфічні чи науково-фантастичні сценарії [3, p. 6]. Такі середовища створюють умови, які хоча і є уявними, але мають зв'язок із реальними задачами. У дослідженні застосовували як комерційні відеоігри, так і спеціалізовані дослідницькі платформи.

Комерційні відеоігри забезпечують візуально насичені світи з широкими можливостями для складних взаємодій. У рамках проєкту вибирали ігри без насильства, із завданнями, які наближені до реальних умов: *Goat Simulator 3, Hydroneer, No Man's Sky, Satisfactory, Teardown, Valheim, Wobbly Life*. У цих іграх агентам необхідно було виконувати різноманітні завдання, як-от збирання ресурсів, будівництво споруд, планування дій, орієнтація у просторі та створення складних структур (наприклад, фабрик або транспортних систем).

Дослідницькі платформи створені для цілеспрямованого навчання ШІ та пропонують контрольовані умови. Вони забезпечують більш точну та швидку валідацію результатів, а також підтримують

реалістичну фізичну взаємодію. Були використані середовища *Construction Lab, Playhouse, ProcTHOR, WorldLab*, які моделюють буденні сценарії, наприклад орієнтацію в приміщенні, пошук об'єктів, прибирання або приготування їжі.

Завдання, які виконує SIMA в зазначених середовищах, мають такі характеристики: висока невизначеність і відсутність чітких критеріїв оптимального рішення; велика розмірність простору можливих дій; відсутність алгоритмів, які б гарантували завжди правильний результат. Ці задачі вважають інтелектуальними, оскільки вони потребують від агента здатності до адаптації, аналізу і прийняття рішень у складних та змінних умовах.

Тренування агента SIMA здійснювали за допомогою навчання з учителем, використовуючи записи ігрових сесій реальних людей, доповнені метаданими. При цьому агент отримував лише зображення, рендерене (візуалізує середовище через віртуальну камеру) грою (таке саме, яке бачить звичайний гравець), а також інструкцію, задану природною мовою, яку необхідно виконати. Для забезпечення рівномірного навчання команди були вручну кластеризовані за логічними категоріями дій.

Агент SIMA перетворює візуальні спостереження та інструкції, задані природною мовою, на дії за допомогою клавіатури та миші. Зважаючи на високу розмірність вхідних і вихідних даних, агента було спроєктовано так, щоб обчислювати відповідь менш ніж за 10 секунд.

У системі використовують попередньо навчені моделі, які були донавчені для специфічних завдань: для оброблення пар «зображення — текст», які створюють загальні представлення вхідних даних; здатні прогнозувати послідовності зображень, що допомагає в аналізі динамічних сценаріїв.

Окрім цього, було розроблено з нуля моделі-трансформери. Їхні результати передаються до іншої нейронної мережі, яка відповідає за перетворення вихідних даних у конкретні дії клавіатури та миші. Основним методом навчання було повторення поведінки на основі записів дій людини.

Для підвищення відповідності виконаних дій інструкціям було застосовано метод Classifier-Free Guidance, який посилює мовну умовність і забезпечує більшу точність виконання завдань. Маючи функцію полісі π, нове значення обчислюватиметься за такою формулою:

$$\pi_{CFG} = \pi\,(image, language) + \lambda\,(\pi\,(image, language) - \pi\,(image, \cdot)).$$

Складність оцінювання агента SIMA полягає у відсутності загальних критеріїв успіху для виконання мовних інструкцій, складному доступі до даних середовища у комерційних іграх та високій вартості завантаження завдань. Для забезпечення надійності використовують різноманітні методи, які перевіряють здатність агента дотримуватися мовних інструкцій, а не лише взаємодіяти із середовищем.

Основні методи оцінювання ефективності агента такі: логарифмічні ймовірності дій, статичне візуальне введення, завдання з використанням реальних середовищ, оптичне розпізнавання символів (OCR), оцінювання людиною [3, p. 11].

Логарифмічні ймовірності дій базуються на основі прогнозів агента за новими даними. Така оцінка слабо корелює з реальними навичками, тому потрібне онлайн-оцінювання. Статичне візуальне введення використовує зображення для оцінювання простих дій без запуску гри. Воно швидко надає базову оцінку, але не підходить для складних завдань. Завдання з використанням реальних середовищ основане на точному тестуванні виконання мовних інструкцій із контролем точності та реакції на зміну інструкцій. Оптичне розпізнавання символів (OCR) використовує текстові повідомлення у грі для автоматизованого оцінювання, особливо у комерційних іграх (наприклад, *No Man's Sky*, *Valheim*). Воно обмежено завданнями з текстовими сповіщеннями. Експертний аналіз відеозаписів дій агента людиною є надійним, але ресурсозатратним. Завдання маркується як провалене, якщо агент виконував нерелевантні дії.

Ці методи забезпечують усебічне оцінювання здатності агента виконувати завдання за мовними інструкціями в різноманітних віртуальних середовищах, сприяючи досягненню довгострокової мети створення універсального ШІ.

Оскільки в останні роки з'являється все більше побоювань та обмежень щодо використання ШІ, то SIMA теж було оцінено на відповідність Google AI Principles (https://ai.google/responsibility/principles/) та визначено, що потенційний виграш перевищує ризики.

Порівняльні результати

Найперше розглянемо продуктивність агента SIMA у семи середовищах, запозичену із [3, p. 16].

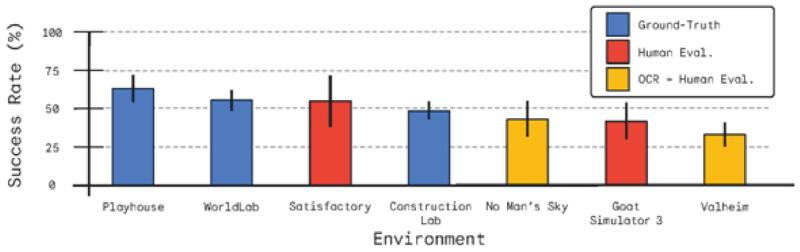

Рис. 1. Середня продуктивність агента SIMA у семи середовищах, для яких визначено кількісну оцінку результату

Результати (рис. 1) [3, p. 16] показують, що агент SIMA здатен виконувати широкий спектр завдань у різноманітних середовищах, хоча і залишається чималий простір для покращень. Варто зазначити, що продуктивність у середовищах Playhouse та WorldLab значно вища, бо самі середовища є порівняно простішими.

Зрозуміло, що важливим є і оцінювання вміння агента узагальнювати. Спочатку наведемо порівняння різних моделей, описане в [3, p. 16, 18] (рис. 2). Стовпчики відображають продуктивність агента SIMA, базових моделей та їхніх модифікацій порівняно з продуктивністю спеціалізованих агентів у середовищах (нормалізовано до 100 %). Суцільна лінія показує нормалізовану продуктивність спеціалізованих агентів.

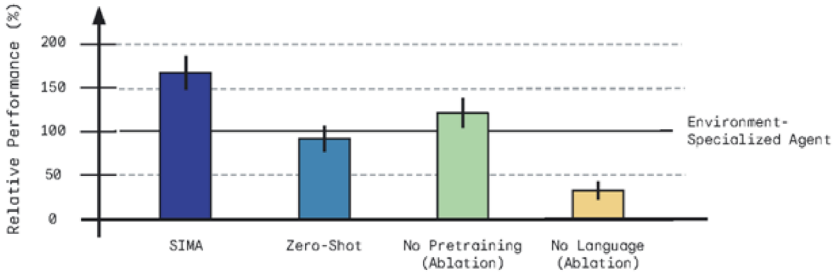

Рис. 2. Порівняння на різних моделях

Порівняння основного агента SIMA здійснено з кількома базовими моделями та модифікаціями, як сукупно, так і в різних середовищах. Результати за кожним з оцінюваних агентів такі:

- *SIMA* — основний агент, навчений на всіх середовищах, окрім Hydroneer та Wobbly Life, які використовують для якісного оцінювання без додаткового навчання (zero-shot);
- *Zero-shot* — окремі агенти SIMA, кожен з яких навчений на $N - 1$ середовищах і оцінюється на тому, на якому не відбувалося навчання, щоб оцінити здатність до перенесення знань. Ці агенти використовують ті самі попередньо навчені енкодери, що й основний SIMA, окрім випадку з Goat Simulator 3, де результати переносу є незалежними;
- *без попереднього навчання* — агент без попередньо навчених енкодерів, замінений на ResNet-модель для оброблення зображень, яка навчається з нуля. Це дає змогу перевірити, як попередньо навчені моделі впливають на продуктивність агента;
- *без мовного компонента* — агент без мовних вхідних даних під час навчання й оцінювання, що дає змогу оцінити, наскільки продуктивність агента пояснюється не мовою, а іншими поведінковими принципами;
- *спеціалізований для середовища* — експертний агент, навчений виключно для кожного середовища з використанням доступних даних і попередньо навчених енкодерів. Продуктивність інших агентів порівнюють із експертним агентом, щоб оцінити максимальні можливості в кожному середовищі.

Основний агент SIMA демонстрував на 67 % кращу продуктивність, ніж спеціалізовані агенти, що свідчить про позитивний перенос навичок між середовищами. Пермутаційні тести підтверджують, що SIMA значно перевершує спеціалізованих агентів у кожному середовищі (p-рівні: 0.001, 0.002, 0.036, 0.0002, 0.008, 0.004, 0.0002).

SIMA також перевищує базові моделі. Агент без попереднього навчання показав значно гірші результати, підтверджуючи, що знання на великій базі підтримують навчання (хоча різниця варіюється за середовищами). Модель без мовного компонента показала низьку продуктивність ($p < 0.001$), що свідчить про те, що агент використовує мовні інструкції, а не лише ймовірні поведінкові шаблони.

Оцінки zero-shot показали обнадійливі результати. Навіть без тренування в конкретному середовищі агент демонстрував високу продуктивність на загальних завданнях, хоча поступався в середовищно-специфічних навичках. Агенти zero-shot здатні виконувати загальні навігаційні завдання, що трапляються в багатьох іграх (наприклад, «спустись із пагорба»), та демонстрували складніші вміння, як-от взаємодія з об'єктом за кольором, використовуючи постійність кольорів між іграми та загальні шаблони взаємодії (наприклад, використання лівої кнопки миші для взаємодії з об'єктами).

Важливо, що навіть у середовищі Goat Simulator 3, де агенти не отримували візуального донавчання, агент zero-shot працював на рівні зі спеціалізованим агентом. Це вказує на те, що перенос знань не обмежується лише візуальними компонентами.

Обнадійливою є і така оцінка. В роботі [3, p. 19] наведено порівняння результатів основної моделі SIMA із її варіаціями без CFG та без мовного компонента відповідно. Результати порівняння ілюструє рис. 3 [3, p. 19].

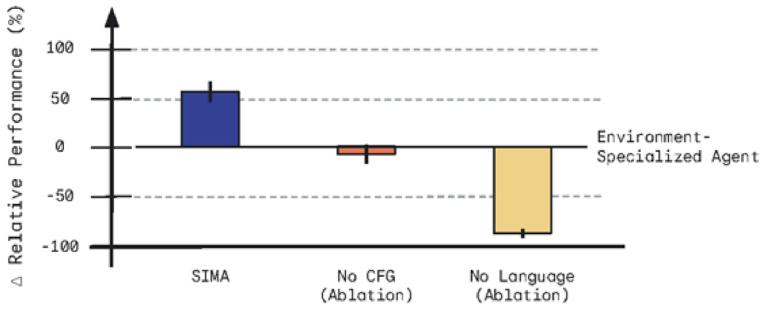

Рис. 3. Порівняння результатів основної моделі SIMA із її варіаціями без CFG та без мовного компоненту відповідно

Порівняно продуктивність агентів із використанням і без використання classifier-free guidance на підмножині дослідницьких середовищ: Construction Lab, Playhouse та WorldLab. Без CFG ($\lambda = 0$) агент SIMA працює помітно гірше. Проте навіть без CFG агент демонструє високий рівень мовної умовності, значно перевершуючи базову модель без мовного компонента. Ці результати підкреслюють переваги CFG та вплив втручань під час інференсу на керованість агента.

Зрозуміло, що варто порівняти агента і з продуктивністю людини. У роботі [3, p. 2] наведено результат такого порівняння (рис. 4). Хоча люди і показують кращий результат, ніж агенти, проте і їхня точність становить лише 60 %.

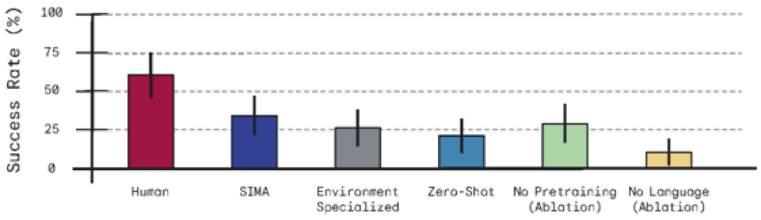

Рис. 4. Порівняння з людиною

Ми вважаємо цей показник дуже важливим, адже завдання, які виконувалися, це не є щось механічне, де машина легко може перевершити людину. Це досить комплексні й нетривіальні завдання (інтелектуальні задачі). Порівняння показало, наскільки насправді агент наблизився до людської продуктивності, тобто за цим критерієм результати можна вважати кращими, ніж за абсолютної оцінки, адже продуктивність агента становила приблизно 50 % продуктивності людини.

Далі розглянемо, як проводили цей тест і якого висновку дійшла команда DeepMind [3, p. 20].

Для додаткового порівняння агентів SIMA з людською продуктивністю було проведено оцінювання на додатковому наборі завдань у грі No Man's Sky. Завдання варіювалися за складністю: від простих інструкцій на кшталт «іди вперед» до складніших, як-от «використай аналізатор для ідентифікації нових тварин». Людські учасники були досвідченими гравцями, які брали участь у збиранні даних. Їхню продуктивність оцінювали ті самі судді й за тією самою методикою, що й для агентів; судді не знали, чиї дії оцінюють — людей чи агента.

Результати, подані на рис. 4, показують, що люди успішно виконали лише 60 % завдань, що свідчить про їхню складність і суворість критеріїв оцінювання. Деякі помилки гравців виникали через зайві дії перед виконанням завдання, наприклад вони відкривали меню корабля за інструкції «заряди лазер для видобутку». Незважаючи на складність, агент SIMA досяг 34 % успіху, значно перевершуючи базового агента без мовного компонента (11 % успіху). Зазначено, що 100 % успіху може бути недосяжним через розбіжності в оцінках суддів щодо неоднозначних завдань. Однак залишається значний обсяг роботи, необхідний для досягнення рівня людської продуктивності. Це підкреслює корисність системи SIMA як складного, але інформативного методу оцінки мовної взаємодії в агентів із втіленням.

OpenAI Five та AlphaStar досягли значного успіху, перемігши професійних гравців у іграх-стратегіях Dota 2 та Starcraft II відповідно. Хоча SIMA ще не досяг такого рівня, але все ж цей проєкт зосереджений на універсальності та роботі з природною мовою, і за цими параметрами SIMA показує наразі найкращі результати.

Подальші кроки

SIMA — проєкт, що перебуває в процесі розвитку. У роботі [3, p. 20] описано цілі та філософію SIMA, а також представлено попередні результати, які демонструють здатність агента пов'язувати мовні інструкції з поведінкою в різних насичених 3D-середовищах. Спостерігається помітна продуктивність і ранні ознаки переносу навичок між середовищами, зокрема zero-shot перенесення базових навичок на невідомі середовища. Проте багато навичок і завдань залишаються недосяжними.

Майбутні дослідження передбачають:

- масштабування до більшої кількості середовищ і наборів даних, розширюючи портфоліо ігор та середовищ;
- підвищення стійкості та керованості агентів;
- використання все якісніших попередньо навчених моделей;
- розроблення більш комплексних і ретельно контрольованих оцінювань.

Очікують, що це зробить SIMA ідеальною платформою для проведення передових досліджень із безпечного поєднання мови та попередньо навчених моделей у складних середовищах і допоможе вирішити фундаментальну проблему AGI (Artificial General Intelligence). Дослідження також має потенціал збагачення навчального досвіду та середовищ розгортання майбутніх моделей; одна з цілей — пов'язати абстрактні можливості великих мовних моделей зі втіленими середовищами. Є надія, що SIMA допоможе подолати фундаментальну проблему зв'язку мови зі сприйняттям і дією в масштабі.

Висновок

SIMA продемонстрував значний прогрес у перенесенні навичок між різними середовищами та zero-shot навички, що дають змогу виконувати базові дії в нових середовищах без попереднього навчання. Незважаючи на досягнуті успіхи, багато завдань залишаються недоступними для агента, що підкреслює необхідність подальших досліджень. У підсумку, SIMA прокладає шлях до створення більш універсального та безпечного інтелектуального агента, що може розв'язувати складні завдання в симульованих середовищах і відкриває перспективи для майбутнього розвитку загального штучного інтелекту.

На прикладі SIMA ми бачимо, що тема універсального агента ШІ актуальна, над нею працюють і є певні здобутки. Хоча до досягнення людського рівня навіть у обмежених віртуальних середовищах ще треба провести чимало досліджень, і також не забуваймо про величезні обчислювальні потужності, які для цього необхідні.

Список літератури

1. Berner C. Dota 2 with Large Scale Deep Reinforcement Learning / C. Berner et al. — 2019. — https://doi.org/10.48550/arXiv.1912.06680.
2. Mathieu M. AlphaStar Unplugged: Large-Scale Offline Reinforcement Learning / M. Mathieu et al. — 2023. — https://doi.org/10.48550/arXiv.2308.03526.
3. Team S. Scaling Instructable Agents Across Many Simulated Worlds / S. Team et al. — 2024. — https://doi.org/10.48550/arXiv.2404.10179.

References

Berner, C., et al. (2019). *Dota 2 with Large Scale Deep Reinforcement Learning*. https://doi.org/10.48550/arXiv.1912.06680.
Mathieu, M., et al. (2023). *AlphaStar Unplugged: Large-Scale Offline Reinforcement Learning*. https://doi.org/10.48550/arXiv.2308.03526.
Team, S., et al. (2024). *Scaling Instructable Agents Across Many Simulated Worlds*. https://doi.org/10.48550/arXiv.2404.10179.

M. Glybovets, N. Bachynskyi

A GENERALIST AI AGENT SIMA

Developing a universal artificial intelligence agent, a subset of Artificial General Intelligence (AGI), is one of the most complex challenges in modern science. Such an agent must generalize knowledge, learn new skills without explicit programming, adapt to unfamiliar environments, and make effective decisions. Addressing this challenge requires advancements across technical domains while also navigating ethical and computational constraints.

This paper examines SIMA (Scalable, Instructable, Multiworld Agent), a project by Google DeepMind aimed at creating an agent capable of executing natural language instructions across diverse 3D environments. SIMA operates through a single keyboard-and-mouse interface in both commercial video games and research platforms, making it distinct from task-specific AI systems like OpenAI Five or AlphaStar. It processes visual input akin to what a human player sees and executes commands categorized for balanced skill training. Techniques like Classifier-Free Guidance enhance the agent's ability to align actions with instructions.

SIMA's evaluation combines methods such as OCR for task verification in games, static visual input tests for simple actions, and human evaluations for more nuanced performance metrics. These methods demonstrate SIMA's ability to transfer knowledge and perform tasks across environments, though challenges remain in long-term planning and complex physical interactions. Despite limitations, SIMA represents a foundational step toward AGI by integrating language understanding with embodied actions.

The findings underline SIMA's potential as a scalable platform for autonomous operation in both virtual and real-world settings, offering key insights into bridging language, perception, and action. Future research will focus on expanding its environmental adaptability, improving robustness, and addressing ethical deployment concerns.

Keywords: artificial intelligence, agent, SIMA, virtual environments.

Матеріал надійшов 29.11.2024