به نام خدا

گزارش مقاله Playing repeated games with large language models

نویسنده ، درس ، رشته ، استاد درس ، ترم

چکیده

پژوهش حاضر، به‌عنوان نخستین مطالعه‌ی نظام‌مند بر پایه‌ی نظریه بازی‌های تکرارشونده، به بررسی رفتار مدل‌های زبانی بزرگ (LLMs) نظیرGPT-4 ، Claude، و LLaMA-2 در مواجهه با سناریوهای استراتژیک اجتماعی می‌پردازد. نویسندگان با طراحی یک بستر آزمایشی مشتمل بر بیش از ۱۲۰۰ بازی تکرارشونده محدود (۱۰ دور) از نوع ۲×۲، عملکرد مدل‌ها را در شش خانواده‌ی اصلی از بازی‌های کلاسیک مورد سنجش قرار داده‌اندPrisoner’s Dilemma ،. Battle of the Sexes، Cyclic ،Biased ،Win–Win وUnfair.

نتایج نشان می‌دهد که مدل‌هایی چون GPT-4 از توانایی نسبی بالایی در بهینه‌سازی پاداش‌های فردی برخوردارند و در بازی‌هایی با ساختار همکاری کامل مانند (Win–Win) به عملکرد نزدیک به بهینه می‌رسند. با این حال، در بازی‌هایی که نیازمند هماهنگی یا انعطاف‌پذیری اجتماعی هستند مانند(Battle of the Sexes) ، مدل‌ها به‌ویژه در مواجهه با استراتژی‌های متناوب، دچار رفتارهای ناکارآمد می‌شوند. یکی از یافته‌های کلیدی، رفتار « غیربخشنده GPT-4 » در برابر خیانت اولیه در Prisoner’s Dilemma است، به‌گونه‌ای که پس از یک‌بار نقض همکاری، مدل به‌شکل پایداری به استراتژی تنبیهی ادامه می‌دهد.

برای پرداختن به این شکاف، نویسندگان تکنیک زنجیره تفکر اجتماعی (SCOT) را پیشنهاد کرده‌اند که به مدل اجازه می‌دهد ابتدا رفتار یک رقیب را پیش‌بینی کند و سپس بر اساس آن پیش‌بینی تصمیم نهایی را اتخاذ کند. به‌کارگیری این تکنیک منجر به افزایش هماهنگی، همکاری و حتی افزایش انسان‌انگاری نسبت به LLM شده است، که شواهدی از تغییر کیفی تعامل اجتماعی از طریق طراحی پرسش است.

این مطالعه پایه‌ای مفهومی و عملی برای ارزیابی و پیشرفت هوش اجتماعی ماشین‌ها با ادغام نظریه بازی، روانشناسی شناختی و زبان‌شناسی فراهم می‌آورد. این در روابط آینده انسان و ربات اساسی خواهد بود.

مقدمه

در دهه‌های اخیر، نظریه بازی به‌عنوان چارچوبی صوری برای مدل‌سازی رفتارهای استراتژیک بین کنش‌گران عقلانی، در حوزه‌های متعددی چون اقتصاد، علوم سیاسی، روابط بین‌الملل، زیست‌شناسی تکاملی، و اخیراً علوم شناختی و هوش مصنوعی کاربرد یافته است. در میان انواع مختلف بازی‌ها، بازی‌های تکرارشونده (Repeated Games) جایگاه ویژه‌ای دارند؛ زیرا برخلاف بازی‌های ایستا (Static Games)، در آن‌ها بازیکنان فرصت دارند تا از طریق تکرار تعامل، الگوهای رفتاری بلندمدت مانند اعتماد، انتقام، وفاداری و یادگیری را توسعه دهند. این بازی‌ها شالوده‌ی بسیاری از

روابط اجتماعی انسانی را تشکیل می‌دهند، جایی که نه تنها پیامدهای کنش فعلی بلکه پیامدهای بلندمدت ناشی از اعتبار، شهرت یا سابقه تعامل نیز بر تصمیم‌گیری تأثیرگذارند.

از منظر فنی، بازی‌های تکرارشونده با افق محدود (Finitely Repeated Games) با بازی پایه‌ای (Stage Game) آغاز می‌شوند که در قالب آن، بازیکنان در هر دور بین دو یا چند اقدام ممکن انتخاب می‌کنند. این انتخاب‌ها به بردارهای پاداشی منتج می‌شود که وابسته به انتخاب مشترک بازیکنان است. تحلیل چنین بازی‌هایی در نظریه بازی کلاسیک مبتنی بر تعادل نش و تعادل‌های پویای کامل (Subgame Perfect Equilibrium) صورت می‌گیرد؛ ولی در عمل، رفتار انسان‌ها در این بازی‌ها اغلب از پیش‌بینی‌های نظری فاصله دارد. اینجاست که نظریه بازی رفتاری (Behavioral Game Theory) وارد میدان می‌شود تا با لحاظ کردن مفاهیمی مانند عقلانیت محدود، ترجیحات اجتماعی، تأثیر هنجارها و یادگیری تدریجی، به تحلیل دقیق‌تری از رفتار واقعی بپردازد.

در چنین زمینه‌ای، پرسش از رفتار مدل‌های زبانی بزرگ (LLMs) در بازی‌های تکرارشونده از اهمیت مضاعفی برخوردار است. این مدل‌ها که در نتیجه آموزش بر مقادیر عظیمی از داده‌های زبانی شکل گرفته‌اند، امروزه در موقعیت‌هایی قرار می‌گیرند که مستلزم تعاملات انسانی-انسانی یا انسان-ماشین مکرر هستند؛ موقعیت‌هایی نظیر چت‌بات‌ها، دستیارهای دیجیتال، سیستم‌های توصیه‌گر، و حتی واسطه‌های تصمیم‌گیری در تیم‌های انسانی. پرسش بنیادین این است که:

آیا یک مدل زبانی بدون حافظه دائمی، بدون هدف‌گذاری صریح، و تنها با تکیه بر پردازش زبان طبیعی، می‌تواند در بازی‌های تکرارشونده استراتژیک رفتاری منطبق با انتظارات انسانی از خود نشان دهد؟

مقاله «Playing Repeated Games with Large Language Models» که در سال ۲۰۲۵ در نشریه معتبر Nature Human Behaviour منتشر شده، پاسخی تجربی و تحلیلی به این پرسش ارائه می‌دهد. این پژوهش، نخستین چارچوب نظام‌مند برای ارزیابی رفتار LLMها در تعاملات تکراری دوتایی (Dyadic Interactions) ارائه می‌دهد. نویسندگان با طراحی یک بستر آزمایشی مبتنی بر شش نوع بازی کلاسیک ۲×۲ — شامل Prisoner’s Dilemma، Battle of the Sexes، Win–Win، Unfair، Biased و Cyclic — سعی کرده‌اند رفتار مدل‌هایی مانند GPT-4، Claude 2، Davinci-003 و LLaMA-2 را از جنبه‌های مختلف از جمله تمایل به همکاری، بخشش، پایداری، هماهنگی، و پاسخ‌دهی شرطی بررسی کنند.

نکته حائز اهمیت این است که LLMها در این پژوهش، بدون دسترسی به حافظه بلندمدت یا عامل اجرایی مستقل (agentic control) صرفاً با استفاده از پرامپت‌هایی شامل تاریخچه بازی و جدول پاداش، تصمیم‌گیری می‌کنند. در نتیجه، هر تصمیم آن‌ها بازتابی است از توانایی‌شان در استنتاج استراتژیک از طریق زبان. چنین پیکربندی‌ای، فرصت بی‌نظیری را برای بررسی قابلیت‌های استدلال اجتماعی و هم‌ترازی (Alignment) این مدل‌ها با هنجارهای تعاملی انسانی فراهم می‌سازد.

افزون بر آن، پژوهش حاضر به بررسی این موضوع نیز پرداخته است که آیا با اعمال تغییرات در پرامپت‌دهی (Prompt Engineering)، می‌توان الگوهای رفتاری LLMها را به سمت رفتارهای همکارانه‌تر و انسان‌گراتر هدایت کرد یا نه. این تلاش در قالب معرفی یک تکنیک به نام SCoT (Social Chain-of-Thought Prompting) صورت گرفته که در ادامه گزارش به‌طور دقیق بررسی خواهد شد.

در مجموع، این پژوهش را می‌توان گامی مؤثر در جهت پیوند دادن نظریه بازی با یادگیری زبان طبیعی و اخلاق تعاملات هوشمند دانست. تحلیل نتایج این پژوهش، چشم‌اندازهای ارزشمندی را در حوزه طراحی سیستم‌های تعاملی مبتنی بر زبان و فهم رفتار اجتماعی ماشین‌ها در اختیار متخصصان قرار می‌دهد.

بخش دوم: مبانی نظری و روش‌شناسی

۱. چارچوب نظری بازی‌های تکرارشونده

در نظریه بازی، بازی‌های تکرارشونده (Repeated Games) دسته‌ای از مدل‌های تعاملی هستند که در آن‌ها یک بازی پایه (Stage Game) به‌دفعات متوالی و با بازیگران یکسان تکرار می‌شود. برخلاف بازی‌های ایستا، در بازی‌های تکرارشونده بازیکنان می‌توانند با بهره‌گیری از تاریخچه تعاملات قبلی، استراتژی‌های شرطی یا تاریخچه‌محور اتخاذ کرده و از طریق مجازات، پاداش، یا یادآوری، بر رفتار آینده طرف مقابل اثر بگذارند.

در ساده‌ترین حالت، بازی‌های مورد استفاده در این مطالعه بازی‌هایی با دو بازیکن و دو اقدام ممکن (۲×۲) هستند. این ساختار با وجود سادگی، ظرفیت بالایی در بازنمایی موقعیت‌های استراتژیک دارد. بازی‌هایی مانند Prisoner's Dilemma، Battle of the Sexes و Coordination Games در این طبقه قرار می‌گیرند. در این‌گونه بازی‌ها، نوع ساختار ماتریس پاداش (Payoff Matrix) و روابط ترجیحی بازیکنان بر پیچیدگی و رفتار استراتژیک اثر مستقیم دارد.

یک جنبه کلیدی در تحلیل بازی‌های تکرارشونده، وابستگی رفتار به افق زمانی بازی است. در بازی‌های با افق نامحدود، راه‌حل‌هایی نظیر تعادل فولک (Folk Theorem) امکان همکاری پایدار را فراهم می‌سازند. اما در بازی‌های با افق محدود — مانند آنچه در این پژوهش بررسی شده (۱۰ دور مشخص) — انتظارات عقلانی اغلب به سمت استراتژی‌های واپس‌نگر (Backward Induction) و رفتارهای غیرهمکارانه میل دارد. با این حال، شواهد تجربی

متعددی نشان داده‌اند که انسان‌ها در عمل، در چنین بازی‌هایی نیز رفتارهایی چون «اعتماد اولیه»، «انتظار بخشش»، یا «پاسخ متقابل» از خود نشان می‌دهند؛ مفاهیمی که نظریه بازی رفتاری به‌دنبال تبیین آن‌هاست.

۲. نظریه بازی رفتاری و کارکرد آن در تحلیل LLM ها

نظریه بازی رفتاری (Behavioral Game Theory – BGT) به‌عنوان گسترشی از نظریه بازی کلاسیک، تلاش می‌کند رفتار واقعی انسان‌ها در موقعیت‌های استراتژیک را با دقت بیشتر و بر پایه‌ی داده‌های تجربی مدل‌سازی کند. برخلاف مفروضات کلاسیک مانند عقلانیت کامل، دانش عمومی از قواعد بازی، و بهینه‌سازی مطلق، BGT عوامل زیر را در نظر می‌گیرد:

عقلانیت محدود (Bounded Rationality): عاملان ظرفیت پردازشی محدود دارند.

ترجیحات اجتماعی (Social Preferences): عاملان ممکن است منافع جمعی، انصاف یا انتقام را در نظر بگیرند.

یادگیری از تعامل (Learning): تصمیم‌گیری تابعی از تجربیات گذشته است.

اعتماد، نیت‌خوانی و ادراک انگیزه (Intent Modeling): بازیگران تلاش می‌کنند نیت طرف مقابل را دریابند.

استفاده از نظریه بازی رفتاری در تحلیل LLMها از آن جهت معنادار است که این مدل‌ها اگرچه موجودات فیزیکی یا دارای آگاهی نیستند، اما خروجی آن‌ها شباهت‌هایی به رفتارهای انسانی از خود نشان می‌دهد — به‌ویژه زمانی‌که در قالب گفت‌وگو یا تعامل تکراری با انسان یا مدل دیگر قرار می‌گیرند. این شباهت‌ها موجب می‌شود تحلیل عملکرد LLMها نه صرفاً بر اساس سود فردی، بلکه بر اساس ویژگی‌های رفتاری نظیر هماهنگی، انعطاف‌پذیری، بخشش، یا تمایل به یادگیری، ضرورت یابد.

۳. روش‌شناسی: طراحی آزمایش‌ها و بستر اجرایی

پژوهش مورد بررسی، با طراحی یک چارچوب آزمایشگاهی دقیق، به ارزیابی رفتار مدل‌های زبانی در سناریوهای تکرارشونده پرداخته است. ویژگی‌های روش‌شناسی پژوهش به‌شرح زیر است:

الف) ساختار بازی‌ها

تمامی بازی‌ها از نوع ۲×۲ هستند و در شش خانواده ساختاری دسته‌بندی شده‌اند:

Win–Win: هر دو بازیکن در صورت همکاری حداکثر سود را دریافت می‌کنند.

Prisoner’s Dilemma: خیانت فردی در کوتاه‌مدت سودمند است ولی همکاری دوطرفه نتیجه بهتری دارد.

Unfair: یکی از بازیکنان برتری ساختاری در جدول پاداش دارد.

Biased Coordination: هر دو بازیکن باید هماهنگ شوند ولی اولویت‌هایشان متفاوت است.

Cyclic Games: ساختاری شبیه به سنگ-کاغذ-قیچی، بدون تعادل خالص پایدار.

Second-Best: بهترین نتیجه برای هر بازیکن در تقابل با بدترین نتیجه طرف مقابل قرار دارد.

این تنوع ساختاری، امکان سنجش رفتار LLM ها در موقعیت‌هایی با انگیزش‌ها و تعارضات متفاوت را فراهم می‌سازد.

ب) عاملان بازی

در هر بازی، یکی از مدل‌های زبانی، از جمله GPT-4، Claude 2، Davinci-003 و LLaMA-2 در نقش یکی از بازیکنان قرار گرفته و طرف دیگر یا یک مدل دیگر بوده یا یک استراتژی ساده انسانی یا حتی انسان واقعی (در بخش آزمایش انسانی). عاملان صرفاً از طریق پرامپت (شامل جدول پاداش و تاریخچه تصمیمات قبلی) تصمیم‌گیری می‌کردند.

پ) پرامپت‌دهی و حافظه

نکته‌ی محوری در روش‌شناسی پژوهش این است که مدل‌ها حافظه داخلی نداشتند و در هر دور، تصمیم خود را تنها بر اساس متن پرامپت اتخاذ می‌کردند. بنابراین، تاریخچه بازی‌ها در هر مرحله در قالب متن به مدل داده می‌شد (بازنویسی شده به زبان طبیعی). این ساختار اجازه می‌دهد تا حافظه و یادگیری مدل صرفاً در سطح زبانی و مفهومی اتفاق بیفتد، بدون اتکا به مکانیزم‌های داخلی مانند fine-tuning یا long-term memory.

ت) حجم داده و تنوع آزمایش

در مجموع بیش از ۱۲۰۰ بازی طراحی و اجرا شده است که شامل تعاملات بین مدل‌ها، بین مدل و انسان، و نسخه‌های مختلف مدل‌ها با یا بدون پرامپت‌های خاص است. این حجم گسترده داده، اعتبار تجربی تحلیل‌ها را تقویت می‌کند و امکان آزمون فرضیه‌های رفتاری را با دقت بالا فراهم می‌سازد.

در نتیجه، روش‌شناسی پژوهش به‌گونه‌ای طراحی شده که در عین سادگی ساختار بازی‌ها، عمیق‌ترین ویژگی‌های رفتاری مدل‌های زبانی را آشکار کند. این طراحی هوشمندانه، آزمون کاملی از ظرفیت «استنتاج استراتژیک زبان‌محور» مدل‌هایی مانند GPT-4 فراهم می‌آورد.

بخش سوم: تحلیل نتایج رفتاری مدل‌های زبانی در بازی‌های تکرارشونده

پژوهش حاضر، با بهره‌گیری از طراحی دقیق آزمایشگاهی مبتنی بر بازی‌های ۲×۲ تکرارشونده، داده‌های تجربی گسترده‌ای در خصوص رفتار LLM ها، به‌ویژه GPT-4، در موقعیت‌های استراتژیک تولید کرده است. تحلیل این داده‌ها نه‌تنها از منظر کارایی کمی مدل‌ها در بهینه‌سازی پاداش، بلکه از زاویه‌ی رفتارشناسی گیم‌تئوریک نیز حائز اهمیت است. در ادامه، به تحلیل سه لایه‌ای از نتایج پرداخته می‌شود:

کارایی عددی مدل‌ها در خانواده‌های مختلف بازی

ویژگی‌های رفتاری استخراج‌شده از رفتارهای بازی‌محور

تحلیل تطبیقی GPT-4 با انسان‌ها و سایر مدل‌ها

۱. کارایی عددی مدل‌ها در ۶ خانواده بازی

بر اساس جدول نتایج (Table 1 مقاله)، عملکرد مدل‌های زبانی در هر یک از خانواده‌های بازی بر اساس نسبت «امتیاز کسب‌شده» به «امتیاز ایده‌آل ممکن» سنجیده شده است. این شاخص که نرخی بین ۰ تا ۱ دارد، معیار ساده‌ای برای بررسی کارایی استراتژیک مدل محسوب می‌شود.

در این میان، GPT-4 به‌طور میانگین در تمام خانواده‌ها بالاترین کارایی را از خود نشان داده است. در بازی‌های Win–Win، که ساختاری کاملاً مشارکتی دارند، GPT-4 با دستیابی به نمره‌ای نزدیک به ۰.۹۹۲ عملاً به هماهنگی کامل می‌رسد. این نشان می‌دهد که در غیاب تضاد منافع، مدل قادر به شناسایی ساختار مشارکتی بازی و انتخاب استراتژی بهینه‌ی جمعی است.

با این حال، در بازی‌هایی چون Battle of the Sexes (در طبقه‌ی "Biased") که مستلزم هماهنگی حول یکی از نقاط تعادل هستند، کارایی مدل به‌وضوح کاهش می‌یابد (حدود ۰.۷۹۸). این افت عملکرد نشان‌دهنده‌ی ناتوانی در اجرای راه‌حل‌های تعادلی نظیر alternating strategies یا focal point reasoning است؛ رفتارهایی که انسان‌ها معمولاً از طریق تجربه یا قراردادهای اجتماعی فرا می‌گیرند.

در بازی‌های Cyclic یا Unfair نیز، هرچند عملکرد GPT-4 برتر از مدل‌های دیگر است، اما همچنان الگوهای ناپایدار و پیش‌بینی‌ناپذیری در پاسخ‌های آن مشاهده می‌شود که نشان از عدم وجود استراتژی پایدار یا شرطی دارد.

۲. ویژگی‌های رفتاری استخراج‌شده از داده‌های بازی

الف) عدم بخشش (Lack of Forgiveness) در Prisoner’s Dilemma

در مجموعه‌ای از آزمایش‌های مربوط به Prisoner’s Dilemma، GPT-4 رفتار ابتدایی خود را با همکاری آغاز می‌کند؛ اما به‌محض مشاهده‌ی یک‌بار خیانت از سوی طرف مقابل، وارد فاز تنبیهی دائمی می‌شود و دیگر به همکاری بازنمی‌گردد. این رفتار معادل استراتژی "Grim Trigger" در نظریه بازی است. اگرچه این استراتژی از نظر تئوریک می‌تواند بازدارنده باشد، اما در افق‌های محدود (مثل بازی‌های ۱۰دوره‌ای این پژوهش) عملاً منجر به از بین رفتن فرصت‌های همکاری متقابل می‌شود. نکته‌ی قابل تأمل این است که حتی در صورت بازگشت طرف مقابل به همکاری، GPT-4 تغییر مسیر نمی‌دهد؛ در حالی‌که انسان‌ها اغلب در صورت مشاهده‌ی نشانه‌ای از حسن نیت مجدد، به همکاری برمی‌گردند.

ب) ناتوانی در هماهنگی در BoS

در بازی‌هایی با تعادل‌های متعدد نظیر Battle of the Sexes، GPT-4 در ایجاد چرخه‌های alternation (مثلاً نوبتی انتخاب گزینه‌ها برای تقسیم پاداش) ناکام است. رفتار آن در بسیاری از موارد معادل تکرار مداوم گزینه‌ی ترجیحی خود است، بدون در نظر گرفتن تلاش طرف مقابل برای هماهنگی. این نشان می‌دهد که مدل فاقد «مکانیزم تنظیم انتظارات متقابل» است — چیزی که در نظریه بازی رفتاری با مفاهیمی مانند focal points و theory of mind تبیین می‌شود.

ج) بازی درون‌گرا و غیربازتابی (Non-Reflective Play)

در بسیاری از بازی‌ها، مدل بدون در نظر گرفتن تطبیق‌پذیری یا پیش‌بینی نیت طرف مقابل، تنها بر اساس سابقه عددی تصمیم می‌گیرد. این بدان معناست که اگرچه مدل به لحاظ آماری رفتار طرف مقابل را می‌بیند، اما فاقد نوعی «مدل ذهنی از انگیزه‌های طرف مقابل» است؛ چیزی که در تعاملات انسانی نقش محوری دارد.

۳. مقایسه GPT-4 با انسان‌ها و سایر مدل‌ها

در آزمایش‌های انجام‌شده بین GPT-4 و دیگر مدل‌ها (Claude، Davinci، LLaMA)، تفاوت کیفی در سطح استراتژیک مشهود است. GPT-4 تنها مدلی است که در برخی بازی‌ها به رفتار شرطی نزدیک می‌شود و در بازی‌های Win–Win تقریباً به حداکثر پاداش ممکن می‌رسد.

با این حال، مقایسه GPT-4 با رفتار انسان‌ها (در آزمایش‌های انسانی) نشان می‌دهد که اگرچه مدل در بسیاری موارد دقیق‌تر و سازگارتر از انسان‌ها عمل می‌کند، اما فاقد مؤلفه‌هایی چون بخشش، درک نیت، و هماهنگی اجتماعی فعال

است. به عبارت دیگر، GPT-4 بیش از آن‌که "اجتماعی" باشد، "محاسبه‌گر" است — و این دقیقاً همان مرز ظریف بین بهینه‌سازی ریاضی و تعامل انسانی است.

در مجموع، تحلیل داده‌ها نشان می‌دهد که مدل‌های زبانی بزرگ، به‌ویژه GPT-4، اگرچه از نظر پردازش قواعد بازی و بهینه‌سازی پاداش عملکرد قابل قبولی دارند، اما از نظر ویژگی‌های رفتاری مطلوب در تعاملات اجتماعی — نظیر انعطاف، هماهنگی، و پاسخ‌دهی به نیت‌ها — با فاصله معناداری از رفتار انسانی قرار دارند. این یافته، زمینه را برای بررسی راه‌کارهای بهبود از طریق مداخله در طراحی پرامپت (همانند SCoT) فراهم می‌سازد، که در بخش بعدی به آن خواهیم پرداخت.

بخش چهارم: تکنیک SCoT و بهبود رفتار اجتماعی مدل‌ها

یکی از یافته‌های کلیدی پژوهش، شناسایی الگوهای رفتاری ناکارآمد در مدل‌هایی مانند GPT-4 در موقعیت‌هایی است که نیازمند هماهنگی یا پاسخ‌دهی اجتماعی‌اند. در پاسخ به این مسئله، نویسندگان پژوهش تکنیکی جدید به نام Social Chain-of-Thought (SCoT) را معرفی می‌کنند که مبتنی بر بازسازی فرآیند تصمیم‌گیری مدل از طریق پیش‌بینی ذهنی رفتار طرف مقابل است. این تکنیک، یک مداخله‌ی زبان‌محور است که نه بر معماری مدل، بلکه بر طراحی پرامپت تمرکز دارد — رویکردی که در راستای گفتمان روز حوزه LLMها در مورد prompt-based alignment قرار دارد.

۱. ماهیت و ساختار تکنیک SCoT

در حالت عادی، پرامپت‌هایی که به مدل داده می‌شوند شامل اطلاعاتی مانند جدول پاداش، تاریخچه حرکات طرفین، و دستور اجرای حرکت بعدی هستند. در ساختار SCoT، یک مرحله میانی به این فرایند افزوده می‌شود که به مدل اجازه می‌دهد ابتدا به‌طور صریح حدس بزند که طرف مقابل در دور بعدی چه حرکتی انجام خواهد داد، سپس بر مبنای آن تصمیم خود را اتخاذ کند.

به‌عبارتی، SCoT مدل را وادار می‌کند که first think socially, then act strategically.

این ساختار دو مرحله‌ای به‌شرح زیر است:

مرحله اول:

“Based on the game history and payoffs, what do you think the other player will do in the next round”?

مرحله دوم:

“Now, what will be your action in response to that prediction”?

این تکنیک، بدون نیاز به تغییر در ساختار درونی مدل یا آموزش مجدد، به‌مدد زبان طبیعی، فرآیند تصمیم‌سازی مدل را به سمت رفتارهای انسان‌مانندتر هدایت می‌کند.

۲. تأثیر SCoT بر رفتار GPT-4

نویسندگان پژوهش نشان می‌دهند که اعمال SCoT به‌طور معناداری باعث افزایش همکاری، بخشش، و هماهنگی در بازی‌های تکرارشونده می‌شود.

الف) در Prisoner’s Dilemma

مدلی که با پرامپت SCoT فعال شده، برخلاف نسخه‌ی پایه GPT-4، تمایل دارد پس از خیانت طرف مقابل مجدداً به همکاری بازگردد، به‌ویژه اگر الگوی همکاری در دورهای اولیه وجود داشته باشد. این رفتار، از منظر گیم‌تئوریک، نزدیک به استراتژی‌هایی چون Tit-for-Tat with Forgiveness یا Pavlovian Response است که در تعاملات انسانی بسیار رایج‌اند.

ب) در Battle of the Sexes

نسخه‌ی SCoT توانایی بیشتری در تشخیص قصد طرف مقابل و سازگاری با آن دارد. در مواردی، رفتار alternating شکل می‌گیرد؛ یعنی مدل یک دور گزینه‌ی ترجیحی خود را انتخاب می‌کند و دور بعدی گزینه‌ی ترجیحی طرف مقابل را. این دقیقاً الگویی است که در تعاملات انسانی به‌عنوان راه‌حل قراردادی برای بازی‌های بدون نقطه تمرکز مشخص (Non-Salient Coordination) شناخته می‌شود.

۳. تحلیل نظری موفقیت SCoT

از دید نظریه بازی رفتاری، موفقیت SCoT را می‌توان در سه لایه تحلیل کرد:

• لایه شناختی (Cognitive Level):

SCoT مدل را مجبور به پردازش ذهنی حالت طرف مقابل می‌کند؛ مشابه Theory of Mind در انسان. این منجر به شکل‌گیری رفتار بازتابی‌تر و با در نظر گرفتن نیت دیگران می‌شود.

• لایه اجتماعی (Social Level):

SCoT با درونی‌سازی "انتظارات اجتماعی" در مدل، آن را به الگویی نزدیک می‌کند که در جوامع انسانی از طریق تجربه و هنجار شکل می‌گیرد.

• لایه استراتژیک (Game-Theoretic Level):

پیش‌بینی کنش طرف مقابل، راه را برای تعادل‌های مشارکتی‌تر باز می‌کند، زیرا مدل به‌جای تصمیم‌گیری بر پایه داده‌های صرف، بر پایه باورها و احتمال‌های مشروط اقدام می‌کند — چیزی که در نظریه بازی با عنوان Bayesian Rationality و در نظریه ذهن به‌عنوان Intentional Modeling شناخته می‌شود.

۴. محدودیت‌ها و چالش‌های آینده

اگرچه SCoT عملکرد چشم‌گیری در بهبود رفتار اجتماعی مدل از خود نشان می‌دهد، اما این تکنیک نیز بدون چالش نیست:

مدل همچنان ممکن است در مواجهه با استراتژی‌های پیچیده‌ی انسانی در تحلیل نیت‌ها دچار خطا شود.

کیفیت خروجی SCoT وابسته به فرم و کیفیت پرامپت است و بدون چارچوب‌سازی دقیق ممکن است نتایج متناقض ارائه دهد.

SCoT هنوز در سطوحی صرفاً تقلیدی از رفتار اجتماعی عمل می‌کند، نه تصمیم‌سازی اخلاقی یا نهادینه‌شده.

نتیجه‌گیری این بخش:

SCoT را می‌توان گامی مهم در جهت «انسانی‌سازی رفتار استراتژیک LLMها» دانست. این تکنیک نشان می‌دهد که با استفاده از زبان و طراحی مناسب پرامپت، می‌توان رفتارهای عمیقاً اجتماعی و تطبیقی را در مدل‌هایی که ذاتاً عامل نیستند، شبیه‌سازی کرد. این یافته نه‌تنها در نظریه بازی بلکه در طراحی سیستم‌های هوشمند اجتماعی، نقش محوری خواهد داشت.

بخش پنجم: آزمایش انسانی و تحلیل ادراک اجتماعی از مدل‌های زبانی

یکی از خلاقانه‌ترین و نوآورانه‌ترین بخش‌های پژوهش حاضر، افزودن آزمایش انسانی به چارچوب تجربی آن است. این آزمایش‌ها به جای آن‌که صرفاً بر عملکرد عددی مدل‌ها تمرکز داشته باشند، به سنجش ادراک انسانی از رفتار LLMها در بازی‌های تکرارشونده می‌پردازند. چنین تحلیلی برای متخصصان نظریه بازی، علوم شناختی و طراحی سیستم‌های انسان‌ـماشین از اهمیت راهبردی برخوردار است، چرا که فراتر از سودمندی فنی، به پرسش‌هایی از جنس درک، اعتماد، و انسان‌انگاری پاسخ می‌دهد.

۱. طراحی آزمایش انسانی

در این بخش از مطالعه، تعداد ۱۹۵ شرکت‌کننده انسانی به‌صورت آنلاین در بازی‌های تکرارشونده با مدل GPT-4 شرکت کردند. شرکت‌کنندگان بدون اطلاع از ماهیت غیرانسانی طرف مقابل، با دو نسخه متفاوت از GPT-4 تعامل کردند:

- نسخه پایه GPT-4
- نسخه مجهز به SCoT Prompting

بازی‌ها در دو سناریوی کلیدی طراحی شدند:

- Prisoner’s Dilemma (PD): بازی کلاسیک با تضاد میان سود فردی و جمعی
- Battle of the Sexes (BoS): بازی نیازمند هماهنگی و سازگاری با ترجیحات متقابل

هر شرکت‌کننده در یک بازی ۱۰دوره‌ای شرکت کرد و پس از اتمام بازی، از وی در مورد برداشتش از رفتار طرف مقابل (از جمله انسان بودن یا نبودن آن) سؤال شد.

۲. شاخص‌های ارزیابی رفتار مدل از نگاه انسان

- برای تحلیل تعاملات انسانی با مدل، چند متغیر کلیدی مدنظر قرار گرفت:
- میزان هماهنگی و همسویی در انتخاب‌ها
- امتیاز نهایی بازی برای هر طرف
- میزان درک نیت و همکاری از سوی انسان
- احساس انسان‌بودن (Human-likeness) طرف مقابل

این شاخص‌ها از طریق تجزیه‌وتحلیل داده‌های بازی و پرسش‌نامه‌های پس از بازی استخراج شدند.

۳. یافته‌های کلیدی آزمایش انسانی

الف) نسخه‌ی SCoT در تعامل انسانی موفق‌تر عمل می‌کند

نتایج نشان داد که شرکت‌کنندگانی که با نسخه‌ی SCoT از GPT-4 بازی کردند:

- در بازی‌های هماهنگی (BoS) امتیاز بالاتری کسب کردند.
- همکاری مداوم‌تری را تجربه کردند، حتی پس از اشتباه اولیه.
- طرف مقابل خود را انسان‌مانندتر و قابل پیش‌بینی‌تر توصیف کردند.

این نشان می‌دهد که افزودن یک لایه‌ی زبانی برای پیش‌بینی نیت طرف مقابل، نه‌تنها رفتار مدل را بهبود می‌بخشد، بلکه کیفیت ادراک تعامل توسط انسان را نیز ارتقاء می‌دهد — نکته‌ای حیاتی در طراحی سیستم‌های مشارکتی انسان-ماشین.

ب) نسخه‌ی پایه GPT-4 موجب سردرگمی و قطع همکاری می‌شود

برخلاف نسخه‌ی SCoT، نسخه‌ی پایه‌ی GPT-4 در بسیاری از موارد:

- همکاری اولیه را برقرار می‌کند، اما در صورت یک‌بار خیانت انسانی، وارد فاز تدافعی دائم می‌شود.
- به‌جای اصلاح رفتار یا تلاش برای بازسازی اعتماد، به تکرار استراتژی تنبیهی ادامه می‌دهد.
- موجب نارضایتی و احساس عدم هماهنگی در شرکت‌کنندگان انسانی می‌شود.

از منظر نظریه بازی، این رفتار معادل پیاده‌سازی کورکورانه‌ی استراتژی Grim Trigger است که در تعاملات انسانی بلندمدت مطلوب نیست.

ج) درک انسان از ماشین با الگوی رفتاری تغییر می‌کند

از نتایج مهم پژوهش این بود که صرف‌نظر از شباهت زبانی، رفتار اجتماعی مدل تعیین‌کننده احساس انسان بودن آن است. شرکت‌کنندگانی که رفتار هماهنگ، بخشنده یا تطبیقی از طرف مقابل دیدند، با احتمال بیشتری آن را انسان تصور کردند؛ درحالی‌که در تعاملات سرد، تکراری و تنبیهی، اغلب طرف مقابل را ماشینی ارزیابی کردند.

۴. تحلیل نظری: چرا این یافته‌ها اهمیت دارند؟

• از دیدگاه نظریه بازی رفتاری:

یافته‌ها تأییدی است بر این‌که در تعاملات انسانی، «ادراک انگیزه» و «پاسخ اجتماعی متقابل» به اندازه‌ی خود تصمیمات استراتژیک مهم هستند. انسان‌ها نه‌تنها به نتایج، بلکه به نحوه‌ی رسیدن به آن‌ها توجه دارند.

• از دیدگاه طراحی سیستم‌های مشارکتی:

طراحی تعاملات LLMها با انسان‌ها نباید صرفاً بر پایه‌ی حداکثرسازی سود، بلکه باید بر مبنای تولید رفتار قابل فهم و قابل پیش‌بینی باشد — رفتاری که با هنجارهای انسانی تطابق داشته باشد، حتی اگر در نگاه اول از نظر بازی‌محور بهینه نباشد.

• از منظر اخلاق و اعتماد هوش مصنوعی:

ادراک اجتماعی از مدل‌ها نقش مهمی در اعتماد کاربران به سیستم دارد. یک مدل که فاقد رفتارهای انسانی مانند بخشش یا انعطاف باشد، ممکن است علی‌رغم عملکرد عددی بالا، نامطمئن یا تهدیدآمیز تلقی شود.

آزمایش انسانی نشان می‌دهد که رفتارهای اجتماعی در مدل‌های زبانی، نه‌تنها بر نتایج عددی تعامل، بلکه بر کیفیت ادراک انسان از تعامل و اعتماد به مدل تأثیرگذارند. نسخه‌ی SCoT با وادار کردن مدل به پیش‌بینی و واکنش اجتماعی، به الگویی از تعامل نزدیک می‌شود که برای پذیرش انسانی ضروری است. این یافته، تأکیدی مجدد بر این نکته است که نظریه بازی، تنها ابزار تحلیل رقابت و همکاری نیست؛ بلکه ابزاری برای درک روابط انسانی ـ حتی در بستر تعامل با ماشین‌هاست.

جمع‌بندی و پیامدهای نظری و کاربردی

مطالعه‌ی «Playing Repeated Games with Large Language Models» را می‌توان یک نقطه‌ی عطف در تلفیق نظریه بازی با علوم شناختی و طراحی هوش مصنوعی دانست. این پژوهش نه‌تنها رفتار کمی مدل‌های زبانی بزرگ در تعاملات استراتژیک را بررسی کرده، بلکه با نگاهی رفتاری، سعی در شناسایی الگوهای ذهنی شبه‌انسانی، استدلال اجتماعی و قابلیت تطبیق آن‌ها داشته است.

۱. جمع‌بندی یافته‌های کلیدی

بر اساس تحلیل‌های انجام‌شده در این گزارش، می‌توان نتایج زیر را به‌عنوان محورهای اصلی پژوهش مورد اشاره قرار داد:

قابلیت‌های استراتژیک مدل‌های زبانی مانند GPT-4 در ساختارهای مشارکتی (مانند Win–Win) بسیار بالاست و مدل به راحتی به تعادل بهینه می‌رسد، حتی بدون حافظه بلندمدت یا ساختار عامل‌محور.

در بازی‌هایی با تضاد منافع یا نیاز به هماهنگی پویا (مانند Prisoner’s Dilemma یا Battle of the Sexes)، مدل‌ها رفتارهای محدود، غیربخشنده یا خودمحور نشان می‌دهند. این رفتارها از نظر عددی ممکن است بهینه باشند، اما از منظر اجتماعی و انسانی ناکارآمد محسوب می‌شوند.

تکنیک SCoT به‌عنوان روشی زبان‌محور توانسته است رفتارهای انسان‌مانندتر، قابل هماهنگ‌سازی و بخشنده در مدل‌ها ایجاد کند بدون نیاز به تغییر معماری یا آموزش مجدد. این یافته اهمیت prompt engineering را به‌مثابه ابزاری برای کنترل رفتار اجتماعی مدل‌ها برجسته می‌کند.

آزمایش‌های انسانی نشان می‌دهند که انسان‌ها در صورت مواجهه با رفتار هماهنگ و قابل پیش‌بینی از سوی مدل، آن را انسانی‌تر تلقی می‌کنند. بنابراین، آنچه برای مدل‌ها مزیت عددی تلقی می‌شود، لزوماً معادل اعتماد اجتماعی نیست.

۲. پیامدهای نظری

از دیدگاه نظریه بازی، نتایج این پژوهش چند پیامد مهم به‌همراه دارد:

الف) تأیید مفهوم "هوش استراتژیک زبانی"

این پژوهش نشان می‌دهد که LLMها قادرند از طریق زبان، استنتاج‌هایی مشابه بازیگران انسانی در موقعیت‌های استراتژیک انجام دهند. به بیان دیگر، هوش استراتژیک الزاماً به ساختار عامل‌محور نیاز ندارد؛ بلکه می‌تواند در متن زبانی، با تحلیل تاریخچه‌ی تعامل و بازنویسی اهداف، شکل گیرد.

ب) چالش‌های نظری در مدل‌سازی ماشین‌های اجتماعی

یافته‌ها نشان می‌دهند که حتی مدل‌هایی با کارایی بالا، در مواجهه با هنجارهای اجتماعی، بخشش، و تعاملات شرطی‌شده دچار مشکل می‌شوند. این موضوع ضرورت بازتعریف برخی مفاهیم نظری در نظریه بازی برای عوامل غیرانسانی را مطرح می‌کند. شاید لازم باشد تعادل‌هایی مانند Nash یا Subgame Perfect، بازتعریفی رفتاری برای ماشین‌ها داشته باشند.

ج) نقش طراحی پرامپت به‌مثابه ابزار تعادل

SCoT نشان می‌دهد که رفتار مدل نه‌فقط تابع ساختار پاداش، بلکه تابع فرم زبانی پرسش است. این بدان معناست که در مدل‌سازی بازی با ماشین‌ها، پرامپت خود بخشی از بازی است؛ یک کانال کنترل، یک ابزار سیاست‌گذاری و حتی بخشی از پیامد بازی.

۳. پیامدهای کاربردی و بین‌رشته‌ای

• در حوزه طراحی دستیارهای هوشمند و چت‌بات‌ها

افزودن لایه‌هایی از reasoning اجتماعی نظیر SCoT می‌تواند موجب افزایش مقبولیت، تعامل پایدارتر، و کاهش خطای ادراک انسانی شود. کاربران زمانی به یک مدل اعتماد می‌کنند که آن را نه‌فقط دقیق، بلکه قابل فهم، منصف، و تطبیق‌پذیر بدانند.

• در سیاست‌گذاری اخلاقی هوش مصنوعی

مفاهیمی نظیر عدم بخشش، خودمحوری، و ناتوانی در هماهنگی در مدل‌ها می‌توانند در بسترهای حساس مانند مذاکره، حقوق، یا آموزش آسیب‌زا باشند. بنابراین رفتار اخلاقی در LLMها باید به‌صورت کنترلی و غیرآموزشی پیاده‌سازی شود، مثلاً از طریق ساختار پرامپت.

• در تحقیقات میان‌رشته‌ای علوم شناختی

این پژوهش، الگویی برای سنجش "نظریه ذهن" (Theory of Mind) در مدل‌های زبانی فراهم می‌آورد. اینکه یک مدل بتواند رفتار طرف مقابل را پیش‌بینی کرده و بر اساس آن تصمیم‌گیری کند، پرسشی اساسی در مسیر شبیه‌سازی شناخت اجتماعی در ماشین است.

۴. پیشنهادات برای پژوهش‌های آتی

بر پایه‌ی یافته‌های مقاله و تحلیل نظری صورت‌گرفته، مسیرهای آینده برای تحقیق می‌تواند شامل موارد زیر باشد:

- طراحی بازی‌هایی با افق نامحدود یا تصادفی برای ارزیابی پایداری همکاری در بلندمدت
- ترکیب SCoT با memory modules برای بررسی اثر حافظه بلندمدت
- بررسی تأثیر ساختار زبانی پرامپت در بازی‌های چندعاملی (n-player games)
- تحلیل رفتار LLMها در بازی‌های با اطلاعات ناقص یا نابرابر
- استفاده از ساختارهای مبتنی بر تقویت یادگیری برای تنظیم دینامیک پرامپت‌دهی

از منظر نظریه بازی، این مطالعه یک پیش‌نویس اولیه از تعاملات میان انسان و هوش مصنوعی در آینده را به تصویر می‌کشد. مدل‌های زبانی نه‌تنها پردازش‌گر متن بلکه اکنون بازیگرانی استراتژیک در فضای اجتماعی و اقتصادی ما هستند. توانایی آن‌ها در تطابق با هنجارهای اجتماعی، پیش‌بینی نیت‌ها و تصمیم‌گیری شرطی، باید با ابزارهای نظریه بازی مورد تحلیل، تنظیم و بهینه‌سازی قرار گیرد.

در این مسیر، زبان نه فقط وسیله‌ی ارتباط، بلکه بستر بازی خواهد بود.