The CLEAN PAPER

WHAT THE PAPER SAYS · WHAT IT DOESN'T · WHY IT MATTERS

آیا «مدل‌های بنیادین» بزرگِ رباتی واقعاً بهتر کار می‌کنند؟ پاسخ دقیق: تا حدی بله، و بیشتر پژوهش‌ها اصلاً توان تشخیصش را ندارند

Toyota Research Institute «مدل‌های رفتاری بزرگ» را — سیاست‌های رباتی پیش‌آموزش‌دیده روی حدود ۱۷۰۰ ساعت دادهٔ متنوعِ دستکاری — با سخت‌گیری کم‌سابقه در برابر سیاست‌های تک‌کارِ آموزش‌دیده از صفر آزمود: آزمایش‌های کور، تصادفی و پُرنمونه (حدود ۱۸۰۰ اجرای واقعی و بیش از ۴۷٬۰۰۰ اجرای شبیه‌سازی) همراه با آمار واقعی. پس از ریزتنظیم برای هر کار، مدل‌های بزرگ به‌طور میانگین بهتر بودند، تقریباً ۳ تا ۵ برابر دادهٔ اختصاصی کمتری می‌خواستند و هنگام تغییر شرایط مقاوم‌تر بودند؛ عملکرد نیز با دادهٔ پیش‌آموزش بیشتر به‌صورت هموار بالا رفت. اما بدون ریزتنظیم به‌طور پایدار از مدل‌های تک‌کار بهتر نبودند، چند اثر آن‌قدر کوچک بود که فقط نمونه‌های بزرگ آشکارشان کرد، و یک انتخاب معمولی در نرمال‌سازی داده از معماری مهم‌تر بود. این پشتوانه‌ای اندازه‌گیری‌شده برای مسیر مدل‌های بنیادین رباتی است — نه ربات همه‌منظوره، نه همه‌کارهٔ *zero-shot* و نه «جهش ظهوریافته» — و هشداری جدی که شاید بخش بزرگی از رباتیک در حال اندازه‌گیری نویز باشد.

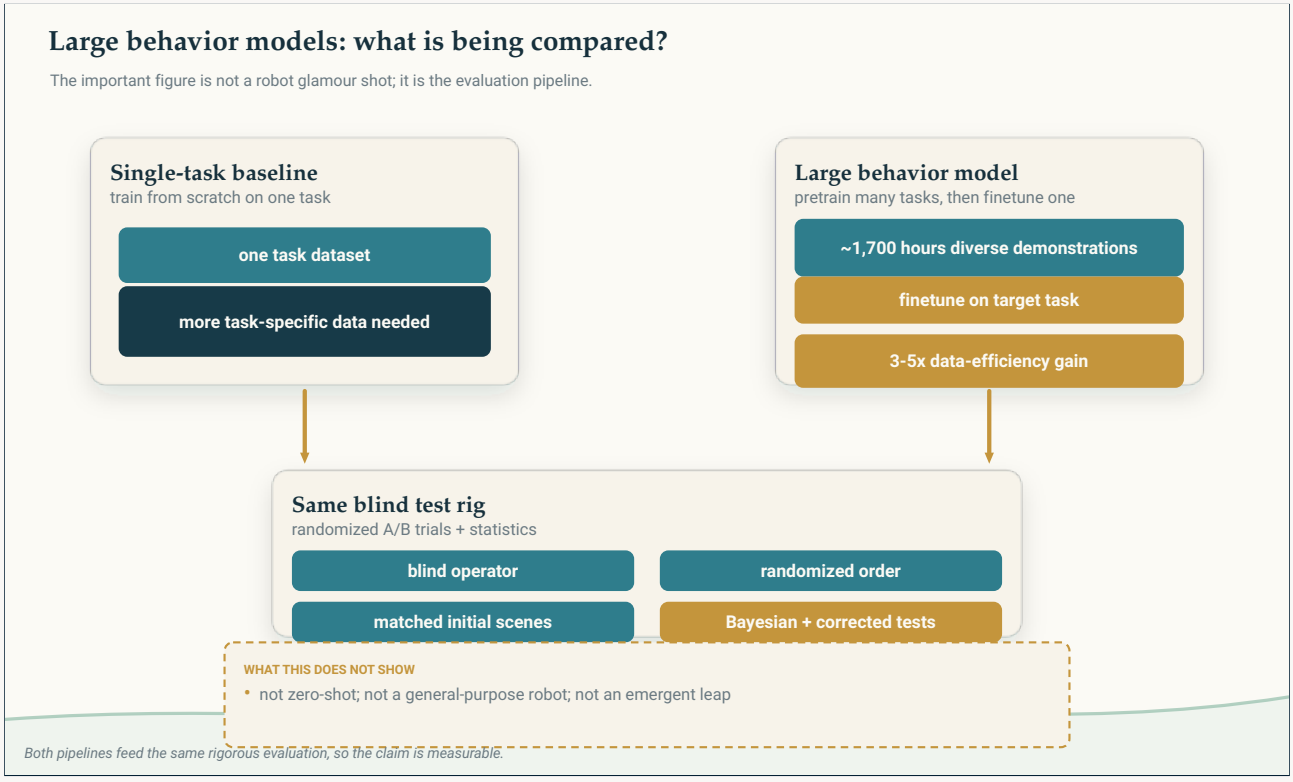
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *A Careful Examination of Large Behavior Models for Multitask Dexterous Manipulation*, Toyota Research Institute Large Behavior Model Team — J. Barreiros, A. Beaulieu, et al.; senior authors incl. R. Ambrus, B. Burchfiel, S. Feng, H. Kress-Gazit (Cornell), R. Tedrake, Science Robotics ;(2026) preprint arXiv:2507.05331 · DOI: scirobotics.aea6201/1126.10

مقاله‌ای دربارهٔ ربات‌ها که موضوع واقعی‌اش صداقت است

رباتیک در میانهٔ تب «مدل‌های بنیادین» قرار دارد. ایده، که از هوش مصنوعی زبانی و تصویری وام گرفته شده، بسیار وسوسه‌انگیز است: به‌جای آن که برای هر کار یک ربات جداگانه آموزش دهیم، یک «**مدل رفتاری بزرگ**» (**Large Model Behavior** یا **LBM**) را روی انبوهی عظیم و متنوع از نمایش‌های عملی آموزش دهیم و سامانه‌ای به دست آوریم که توانایی‌های گسترده داشته باشد و سریع با کارهای تازه سازگار شود. شور و هیجان — و سرمایه‌گذاری — بسیار زیاد است. تیتر تقریباً خودش نوشته می‌شود: مغزهای رباتی همه‌منظوره از راه رسیده‌اند.

این مقاله از Toyota Research Institute دقیقاً به این دلیل جالب است که حاضر نیست آن تیتر را بنویسد. دستاورد اصلی آن رباتی پرزرق‌وبرق‌تر نیست؛ بلکه نگاه سخت‌گیرانه‌ای است به یک پرسش فریبنده و ساده: آیا رویکرد مدل بزرگ واقعاً بهتر کار می‌کند، و اصلاً چطور می‌توانیم بفهمیم؟ نویسندگان این پرسش را با سطحی از دقت آماری پاسخ داده‌اند که به گفتهٔ خودشان در این حوزه غیرمعمول است. حاصل، یک «بله، اما» واقعاً مفید و هشداری است مبنی بر این که شاید بخش بزرگی از رباتیک در حال اندازه‌گیری نویز باشد.



هر دو رویکرد وارد یک ارزیابی کور و تصادفی یکسان می‌شوند. ادعای مقاله این نیست که مدل‌های بنیادین رباتی، همه‌کاره‌های zero-shot هستند؛ بلکه این است که پیش‌آموزش، اگر با دقت اندازه‌گیری شود، می‌تواند بهره‌وری داده و استحکام را بهبود دهد.

نویسندگان چه کردند

آن‌ها LBMها را به معنایی مشخص و عملی ساختند: **سیاست‌های دیداری-حرکتی مبتنی بر dffiusion** (یک Transformer Dffiusion که تصاویر دوربین، یک دستور زبانی کوتاه و موقعیت مفاصل خود ربات را می‌خواند و در فرکانس ۱۰ هرتز بسته‌های کوتاهی از

فرمان‌های حرکتی تولید می‌کند). این مدل‌ها روی حدود **۱۷۰۰ ساعت** نمایش رباتی پیش‌آموزش دیدند — بیش از ۵۰۰ کار متفاوت که درون مجموعه جمع‌آوری شده بود، به‌اضافهٔ مجموعه‌داده‌های عمومی — و سپس برای کارهای منفرد **ریزتنظیم** شدند. در تمام مقاله، مقایسه با یک **سیاست تک‌کارِ آموزش‌دیده از صفر** انجام می‌شود که فقط روی داده‌های همان کار آموزش دیده است.

اما قلب مقاله ارزیابی است؛ چیزی که نویسندگان عملاً آن را نتیجهٔ اصلی می‌دانند. برای این‌که خودشان را گول نزنند، از این روش‌ها استفاده کردند:

- **آزمون A/B کور و تصادفی** در دنیای واقعی — فردی که ربات را اجرا می‌کرد نمی‌دانست کدام سیاست در حال آزمایش است و ترتیب آزمون‌ها تصادفی بود.
- **شرایط اولیهٔ کنترل‌شده و تکرارپذیر** — اپراتورها پیش از هر آزمون صحنه را با یک تصویر راهنما تطبیق می‌دادند.
- **تعداد زیاد آزمون‌ها** — برای هر کار، هر سیاست و هر وضعیت، ۵۰ اجرای واقعی؛ و برای هر کار در شبیه‌سازی ۲۰۰ اجرا. در مجموع: حدود **۱۸۰۰ اجرای کور در دنیای واقعی و بیش از ۴۷٬۰۰۰ اجرای شبیه‌سازی.**
- **آمار واقعی و مناسب** — برآوردهای بیزی از احتمال موفقیت، و آزمون‌های فرضیهٔ زوجی با اصلاح برای مقایسه‌های متعدد، به‌جای قضاوت چشمی از روی میله‌های خطای هم‌پوشان. حتی روی یک‌چهارم آزمون‌هایی که انسان نمره‌گذاری کرده بود، یک مرحلهٔ تضمین کیفیت انجام دادند تا خطای نمره‌گذاری را اندازه بگیرند.

[video pending]

مقایسهٔ کنارهمِ مدل‌ها هنگام چیدن میز صبحانه: سمت چپ خط پایهٔ تک‌کار، و سمت راست LBM. هر دو ویدئو با سرعت 1x پخش می‌شوند. این فقط یکی از کارهای ارزیابی‌شده است، نه شاهدی بر خودمختاری همه‌منظوره.

تمام این سازوکارها اصل ماجرا هستند. استدلال مقاله این است که بدون چنین روشی، نمی‌توانید تفاوت یک بهبود واقعی را با شانس تشخیص دهید.

چه پیدا کردند

مدل‌های بزرگِ ریزتنظیم‌شده، به‌طور میانگین از مدل‌های تک‌کارِ آموزش‌دیده از صفر بهتر بودند. وقتی نتایج کارها با هم تجمیع شد، یLBM که ابتدا پیش‌آموزش دیده و سپس برای یک کار ریزتنظیم شده بود، هم در شبیه‌سازی و هم در دنیای واقعی به‌طور قابل‌اعتماد بهتر از سیاستی بود که از صفر روی همان کار آموزش دیده بود؛ و این فاصله از نظر آماری معنادار بود. در سطح کارهای منفرد نیز LBM ریزتنظیم‌شده تقریباً در همهٔ موارد از نظر آماری هم‌سطح یا بهتر از مدل آموزش‌دیده از صفر بود (۳ از ۳ کار واقعی، و ۱۵ از ۱۶ کار شبیه‌سازی).

بزرگ‌ترین و روشن‌ترین برتری، بهره‌وری داده بود. LBM ریزتنظیم‌شده با حدود **۳ تا ۵ برابر دادهٔ اختصاصی کمتر برای هر کار** به عملکردی معادل مدل آموزش‌دیده از صفر رسید. در یکی از کارهای واقعی — چیدن میز صبحانه — یLBM که فقط با **۱۵٪** نمایش‌ها ریزتنظیم شده بود، از سیاستی که از صفر با **۱۰۰٪** نمایش‌ها آموزش دیده بود بهتر عمل کرد.

پیش‌آموزش بیش از همه وقتی کمک کرد که شرایط تغییر می‌کرد. زمانی که محیط آزمون عمداً از شرایط آموزشی دور شد («تغییر توزیع» یا distribution shift)، LBM ریزتنظیم‌شده بیشتر شد. در یک مجموعهٔ شبیه‌سازی، مدل در شرایط معمول در ۳ کار از ۱۶ کار به‌طور آماری بهتر از مدل آموزش‌دیده از صفر بود، اما زیر تغییر توزیع این عدد به ۱۰ از ۱۶ رسید. چون استقرارهای واقعی همیشه تا حدی از شرایط آموزش فاصله می‌گیرند، شاید این استحکام مهم‌ترین یافته از نظر کاربردی باشد.

دادهٔ پیش‌آموزش بیشتر به‌صورت پیوسته کمک کرد. با افزودن دادهٔ پیش‌آموزش، عملکرد به‌تدریج بالا رفت — بدون جهش ناگهانی یا «ظهور» ناگهانی توانایی در مقیاس‌های آزموده‌شده. مفید، قابل‌پیش‌بینی و بدون نمایش دراماتیک.

اما داستان «همه‌کاره بدون ریزتنظیم» تأیید نشد. LBM پیش‌آموزش‌دیده وقتی به شکل *zero-shot* — بدون هیچ ریزتنظیم اختصاصی برای کار — استفاده شد، به‌طور پایدار از سیاست‌های تک‌کار بهتر نبود. یک شبکهٔ واحد می‌توانست چندین کار را انجام دهد، اما رؤیای «فقط به آن بگو چه کند» در این آزمایش محقق نشد؛ نویسندگان بخشی از این مشکل را به شکنندگی رمزگذار زبانی کوچک خود نسبت می‌دهند.

و اندازهٔ بهبودها آن‌قدر کوچک بود که به‌راحتی می‌شد آن‌ها را ندید — یا حتی تصادف را به جایشان جا زد. بسیاری از اثرها فقط با نمونه‌های بزرگ‌تر از معمول و آزمون‌های دقیق آشکار شدند. نویسندگان صریح می‌گویند با توجه به اندازهٔ اثرها و میزان نویز، **خطر قابل‌توجهی وجود دارد که بسیاری از مقاله‌های رباتیک صرفاً نویز آماری را اندازه‌گیری کنند.** آن‌ها همچنین دریافتند یک انتخاب بسیار روزمره — شیوهٔ نرمال‌سازی داده‌ها — بیش از تغییر معماری بر نتیجه اثر گذاشت، و یک **باگ نرمال‌سازی** در پیش‌آموزش تازه پس از پایان ارزیابی‌ها کشف شد.

این احتمالاً چه معنایی دارد

خوانش قابل‌دفاع این است: پیش‌آموزش بزرگ‌مقیاس روی داده‌های متنوع رباتی واقعاً یک جزء مفید است — برای هر کار تازه به دادهٔ کمتری نیاز دارید و سیاست‌ها وقتی دنیا دقیقاً شبیه شرایط آموزش نیست، مقاوم‌تر می‌شوند. این واقعاً از مسیری که حوزه روی آن شرط بسته پشتیبانی می‌کند. اما بهبودها متوسط و مشروط هستند (عمدتاً پس از ریزتنظیم ظاهر می‌شوند، و در نتایج تجمیعی و زیر فشار روشن‌ترند)، نه نشانهٔ رسیدن یک ربات عمومی و آمادهٔ استفاده.

معنای آرام‌تر اما مهم‌تر، روش‌شناختی است. این مقاله عملاً یک خط‌کش ارائه می‌دهد: نشان می‌دهد برای بیان یک ادعای قابل‌اعتماد دربارهٔ سیاست رباتی واقعاً چه مقدار شاهد لازم است، و به‌طور ضمنی می‌گوید بخشی از هیجان منتشرشده بر شواهد بسیار کم بنا شده است. این اصلاحی است که حوزه احتمالاً بیش از یک مدل دیگر به آن نیاز دارد.

این مقاله چه چیزی را اثبات نمی‌کند

- این **یک ربات همه‌منظوره نیست.** برتری‌ها برای یک معماری مشخص (سیاست‌های (dffiusion که برای هر کار ریزتنظیم شده، در محیط‌های کنترل‌شده و با نمایش‌های تله‌اپراتوری نشان داده شده‌اند — نه رباتی خودمختار که هر کار تازه‌ای را با دستور انجام دهد.
- استفادهٔ **zero-shot** را تأیید **نمی‌کند.** بدون ریزتنظیم، مدل بزرگ به‌طور پایدار از خط‌پایه‌های تک‌کار بهتر نبود.
- شاهدی برای «**جهش ظهوریافته**» نیست. مقیاس‌پذیری بهبود هموار ایجاد کرد؛ هیچ ناپیوستگی‌ای وجود ندارد که روایت «و بعد ناگهان توانمند شد» را پشتیبانی کند.
- اعداد **نسبی و محدود به آزمایشگاه** هستند. نرخ‌های موفقیت مطلق عمداً نزدیک حدود ۵۰٪ تنظیم شدند تا مقایسه حساس باشد؛ بنابراین معیار قابلیت اطمینان در دنیای واقعی نیستند، و کل کار فقط یک معماری از یک آزمایشگاه است.
- توضیح **نمی‌دهد چرا** هر سیاست موفق یا ناموفق می‌شود؛ چند کار مشخص که مدل بزرگ در آن‌ها بدتر بود گزارش شده‌اند اما توضیح داده نشده‌اند.
- دربارهٔ **ایمنی، خودمختاری یا استقرار** خارج از سازوکار ارزیابی چیزی نمی‌گوید.

قدرت شواهد چقدر است؟

برای ادعاهای مقایسه‌ای اصلی — این که هایLBM ریزتنظیم‌شده در مجموع از خط‌پایه‌های آموزش‌دیده از صفر بهترند، چند برابر دادهٔ کمتری می‌خواهند و در برابر تغییر توزیع مقاوم‌ترند — شواهد هم قوی‌اند و هم به شکل غیرمعمولی خوب کنترل شده‌اند: کور، تصادفی، با نمونهٔ بزرگ و آزمون آماری، همراه با مرحلهٔ QA برای نمره‌گذاری. این از آن موارد نادری است که روش‌شناسی آن‌قدر محکم است که می‌توان نتیجه‌های اصلی را تقریباً همان‌طور که بیان شده‌اند پذیرفت.

محدودیت‌های صادقانه همان‌هایی هستند که خود نویسندگان مطرح می‌کنند. میله‌های خطا تصادفی‌بودن ارزیابی را دربر می‌گیرند، اما تصادفی‌بودن آموزش را نه — اگر یک مدل را دوبار آموزش دهید ممکن است دو سیاست به‌طور معنی‌داری متفاوت به دست آورید و این تغییرپذیری در آمار حاضر نیست. هر کار واقعی ۵۰ بار آزموده شد؛ برای اثرهای متوسط کافی است اما ممکن است اثرهای کوچک را از دست بدهد. شرط‌دهی زبانی از یک رمزگذار نسبتاً ساده استفاده کرد، پس ادعاهای «فقط به ربات بگو چه کند» ممکن است برای سامانه‌های بزرگ‌تر متفاوت باشد. و گزارش صریح باگ نرمال‌سازیِ کشف‌شده پس از تحلیل نیز وجود دارد. هیچ کدام یافته‌های اصلی را نابود نمی‌کند، اما دقیقاً همان نوع مسئله‌ای است که مقاله می‌گوید حوزه معمولاً از کنار آن می‌گذرد.

یک نکتهٔ مربوط به منبع، در همان روحیه: این توضیح بر اساس پیش‌چاپ نویسندگان نوشته شده است. ما نتوانستیم نسخهٔ منتشرشده در مجله را بازیابی کنیم، بنابراین تغییرات احتمالی میان پیش‌چاپ و متن نهایی را بررسی نکرده‌ایم.

چرا مهم است

«مدل بنیادین ربات» عبارتی است که تقریباً برای اغراق ساخته شده، و چنین مطالعه‌ای را می‌توان به دو شکل اشتباه خواند — یا با پیروزی گفت « کار می‌کند!» یا با بی‌اعتنایی گفت «همه‌اش هیاهوست.» برداشت دقیق‌تر از هر دو مفیدتر است: پیش‌آموزش روی داده‌های متنوع مزایای واقعی، قابل‌اندازه‌گیری اما متوسط ایجاد می‌کند — عمدتاً دادهٔ کمتر برای هر کار و استحکام بیشتر — و مسیر با افزایش مقیاس به شکلی قابل‌پیش‌بینی بهتر می‌شود.

دلیل عمیق‌تر اهمیت مقاله این است که سخت‌گیری‌اش را متوجه خود حوزه می‌کند. با نشان دادن این که اثرهای واقعی آن‌قدر کوچک‌اند که در ارزیابی شلخته ناپدید می‌شوند، و این که انتخابی کسل‌کننده مانند نرمال‌سازی داده می‌تواند از معماری تازه و هوشمندانه مهم‌تر باشد، استدلال می‌کند که بخش بزرگی از پیشرفت در یادگیری ربات باید پیش از باورشدن با اندازه‌گیری محکم‌تری پشتیبانی شود. مقاله‌ای که اعتبار خود را صرف مرزبانی میان «نتیجه» و «آرزو» می‌کند، کاری نادرتر و ارزشمندتر از صدرنشینی در یک جدول امتیازات انجام می‌دهد.

خلاصهٔ تمیز

پژوهشگران Institute Research Toyota «مدل‌های رفتاری بزرگ» — سیاست‌های رباتی مبتنی بر dffiusion که روی حدود ۱۷۰۰ ساعت دادهٔ متنوع دستکاری پیش‌آموزش دیده بودند — را ساختند و آن‌ها را با سیاست‌های تک‌کارِ آموزش‌دیده از صفر، با پروتکلی غیرمعمولاً سخت‌گیرانه آزمودند: کور، تصادفی، با نمونهٔ بزرگ (حدود ۱۸۰۰ آزمون واقعی و بیش از ۴۷٬۰۰۰ آزمون شبیه‌سازی) و با آمار واقعی. پس از ریزتنظیم برای هر کار، مدل‌های بزرگ در مجموع قابل‌اعتمادتر بهتر عمل کردند، به حدود **۳ تا ۵ برابر دادهٔ اختصاصی کمتر** نیاز داشتند و وقتی شرایط تغییر می‌کرد **مقاوم‌تر** بودند؛ و با افزایش دادهٔ پیش‌آموزش، عملکرد به‌صورت هموار بهتر شد. اما **بدون ریزتنظیم** به‌طور پایدار از مدل‌های تک‌کار جلو نزدند، چندین اثر آن‌قدر کوچک بود که فقط نمونه‌های بزرگ آشکارشان کرد، و یک انتخاب روزمره در نرمال‌سازی داده از تغییر معماری مهم‌تر بود. این پشتیبانی محکم و اندازه‌گیری‌شده‌ای از مسیر «مدل‌های بنیادین رباتی»

است — **نه** یک ربات همه‌منظوره، نه یک همه‌کارهٔ zero-shot و نه یک «جهش ظهوریافته» — و هم‌زمان هشداری تند که شاید بخش زیادی از رباتیک در حال اندازه‌گیری نویز باشد.

بررسی بی‌پرده

مقاله چه نشان می‌دهد: با ارزیابی سخت‌گیرانه، کور و دارای توان آماری (حدود ۱۸۰۰ اجرای واقعی + بیش از ۴۷٬۰۰۰ اجرای شبیه‌سازی)، سیاست‌های dffiusion که ابتدا روی چند کار پیش‌آموزش و سپس ریزتنظیم شده‌اند (هاLBM) در مجموع از سیاست‌های تک‌کارِ آموزش‌دیده از صفر بهتر عمل می‌کنند، با حدود ۳ تا ۵ برابر دادهٔ اختصاصی کمتر به عملکرد معادل می‌رسند و زیر تغییر توزیع مقاوم‌ترند، عملکرد نیز با دادهٔ پیش‌آموزش بیشتر به‌صورت هموار افزایش می‌یابد.

چه چیزی محتمل است اما اثبات نشده: این‌که این مزایا به مدل‌های بسیار بزرگ‌تر vision-language-action هم منتقل شوند (رمزگذار زبانی آن‌ها کوچک بود)؛ و این‌که روند هموار مقیاس‌پذیری خارج از دامنهٔ دادهٔ آزموده‌شده ادامه یابد.

چه چیزی را نشان نمی‌دهد: ربات همه‌منظوره یا zero-shot (بدون ریزتنظیم → برتری پایدار وجود ندارد)؛ هیچ جهش «ظهوریافته» در توانایی؛ توضیح شکست‌های خاص در سطح کار؛ قابلیت اطمینان مطلق در دنیای واقعی (نرخ موفقیت برای حساسیت مقایسه نزدیک ۵۰٪ تنظیم شده بود)؛ یا چیزی دربارهٔ ایمنی و استقرار خودمختار.

محدودیت‌های اصلی: آمار، تصادفی‌بودن ارزیابی را پوشش می‌دهد اما نه تغییرپذیری اجرای آموزش؛ ۵۰ آزمون واقعی برای هر کار ممکن است اثرهای کوچک را نبیند؛ یک معماری و یک آزمایشگاه؛ رمزگذار زبانی نسبتاً ساده؛ یک باگ نرمال‌سازی داده پس از ارزیابی‌ها پیدا شد؛ و تحلیل بر اساس پیش‌چاپ است (نسخهٔ منتشرشده بررسی نشده).

یک خوانندهٔ عمومی چقدر باید مطمئن باشد؟ اطمینان بالا که پیش‌آموزش چندکاره همراه با ریزتنظیم مزایای واقعی اما متوسط می‌دهد — به‌ویژه در بهره‌وری داده و استحکام — و این مزایا این‌جا با دقتی غیرمعمول اندازه‌گیری شده‌اند. اطمینان بالا که این **یک ربات همه‌منظوره یا zero-shot نیست و یک جهش ظهوریافته هم نیست.** اطمینان متوسط دربارهٔ این‌که این مزایا تا مدل‌های بزرگ‌تر چقدر مقیاس می‌گیرند. و هشدار خود نویسندگان ارزش جدی گرفتن دارد: اثرها آن‌قدر کوچک‌اند که مطالعه‌های کم‌توان ممکن است صرفاً نویز گزارش کنند. موضع مناسب: خوش‌بینی اندازه‌گیری‌شده نسبت به رویکرد، و شک سالم نسبت به نتایج رباتیک-AI که چنین پشتوانهٔ آماری‌ای ندارند.

منابع

arXiv:2507.05331
https://arxiv.org/abs/2507.05331

scirobotics.aea6201/1126.10 Robotics, Science — retrieved) (not version Peer-reviewed
https://doi.org/10.1126/scirobotics.aea6201

page Project
https://toyotaresearchinstitute.github.io/lbm1/