The
CLEAN PAPER

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

ایک روبوٹ مقالہ جس کا اصل موضوع honesty ہے

Toyota Research Institute نے "بڑا رویّہ ماڈلز"—diffusion-based روبوٹ policies، ~1,700 hours diverse manipulation—کو ڈیٹاپر from-scratch pretrained واحد- کام policies کے against unusually rigorous protocol سے ٹیسٹ کیا: بلائنڈ، تصادفی، بڑا نمونہ (≈1,800 حقیقی دنیا اور 47,000+ سیمولیشن ٹرائلز)، حقیقی statistics کے ساتھ۔ کام-per finetuning کے بعد big ماڈلز مجموعی میں reliably بہتر، تقریباً 3–5× کم کام-مخصوص ڈیٹا** میں equivalent کارکردگی، اور **تقسیم shift میں زیادہ مضبوط** تھے؛ کارکردگی پری ٹریننگ ڈیٹا کے ساتھ smoothly بہتر کرنا ہوئی۔ مگر **without finetuning** وہ consistently واحد- کام ماڈلز کو beat نہیں کرتے، اثرات کئی جگہ چھوٹا تھے، اور ڈیٹا normalisation انتخاب ساخت سے زیادہ مادّہ کرتی تھی۔ یہ روبوٹ-foundation-ماڈل سمت کے لیے ناپا گیا تائید ہے—**عمومی-purpose** روبوٹ نہیں، generalist zero-shot نہیں، leap emergent نہیں**—اور warning کہ robotics میں مطالعات underpowered شور رپورٹ کر سکتی ہیں۔

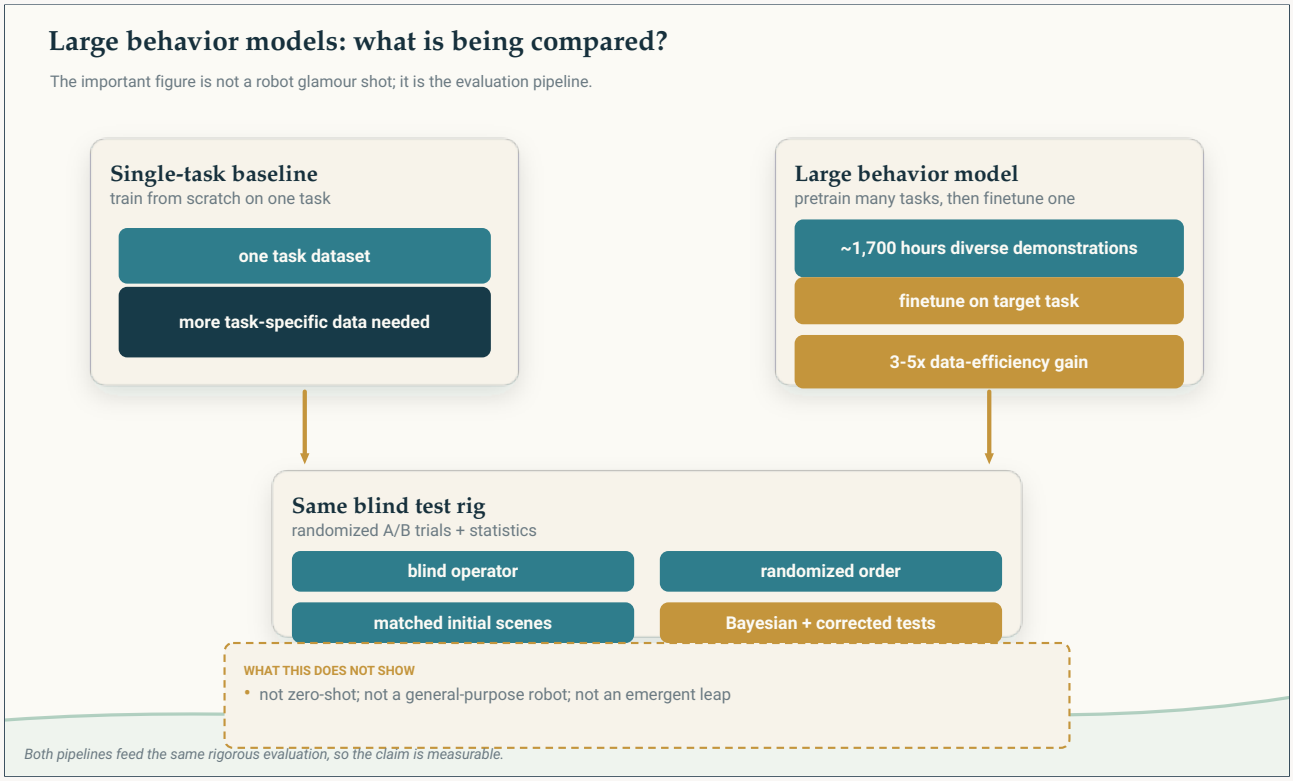
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

روبوٹکس کا ایک مقالہ جس کا اصل موضوع دیانت داری ہے

روبوٹکس اس وقت "فاؤنڈیشن ماڈل" کی دوڑ کے بیچ میں ہے۔ زبان اور تصویر والے اے آئی سے لیا گیا خیال پرکشش ہے: ہر کام کے لیے الگ روبوٹ تربیت دینے کے بجائے بہت بڑی اور متنوع عملی مثالوں پر ایک "**بڑا رویّاتی ماڈل**" **(LBM)** تربیت دیں، اور ایسا نظام حاصل کریں جو کئی کام کر سکے اور جلدی ڈھل جائے۔ جوش بھی بہت ہے اور سرمایہ کاری بھی۔ سرخی خود بن جاتی ہے: عمومی مقصد کے روبوٹ دماغ آ گئے ہیں۔

Toyota Research Institute کا یہ مقالہ اسی لیے دلچسپ ہے کہ وہ یہ سرخی لکھنے سے انکار کرتا ہے۔ اس کا اصل حصہ کوئی زیادہ چمک دار روبوٹ نہیں، بلکہ ایک بظاہر سادہ سوال کی سخت جانچ ہے: کیا بڑے ماڈل والا طریقہ واقعی بہتر کام کرتا ہے، اور ہمیں یہ کیسے معلوم ہوگا؟ مصنفین نے اس کا جواب ایسے شماریاتی احتیاط کے ساتھ دیا ہے جو ان کے بقول خود روبوٹکس میں بھی غیر معمولی ہے۔ نتیجہ ایک مفید "ہاں، لیکن..." ہے، اور ساتھ تنبیہ کہ روبوٹکس کی بہت سی تحقیق شاید اصل اثر کے بجائے شماریاتی شور ناپ رہی ہو۔



دونوں طریقوں کو ایک ہی بلائنڈ، بے ترتیب جانچ سے گزارا گیا۔ مقالے کا دعویٰ یہ نہیں کہ روبوٹ فاؤنڈیشن ماڈلز بغیر مخصوص تربیت کے ہر کام کر لیتے ہیں؛ دعویٰ یہ ہے کہ محتاط پیمائش میں ابتدائی تربیت ڈیٹا کی کارکردگی اور بدلتے حالات میں مضبوطی بہتر کر سکتی ہے۔

مصنفین نے کیا کیا

انہوں نے LBM کو ایک واضح معنی میں بنایا: **dffiusion پر مبنی visuomotor پالیسیاں**—ایسا Transformer Dffiusion جو کیمرا تصاویر، مختصر زبانی ہدایت اور روبوٹ کے اپنے جوڑوں کی پوزیشن پڑھتا ہے، پھر 10 Hz پر مختصر موٹر احکامات دیتا ہے۔ ان ماڈلز

کو تقریباً **1,700 گھنٹے** کی روبوٹ عملی مثالوں پر پہلے سے تربیت دی گئی—ادارے کے اندر جمع کیے گئے 500 سے زیادہ الگ کام، ساتھ عوامی ڈیٹا سیٹس—اور پھر ہر خاص کام کے لیے مزید مخصوص تربیت دی گئی۔ پورے مقالے میں موازنہ اسی کام کے ڈیٹا پر صفر سے تربیت یافتہ ایک-کام پالیسی سے ہے۔

اصل مرکز، تاہم، **جانچ کا طریقہ** ہے، اور مصنفین اسے تقریباً مرکزی نتیجے جتنی اہمیت دیتے ہیں۔ خود کو دھوکا دینے سے بچنے کے لیے انہوں نے:

- حقیقی دنیا میں **بلائنڈ، بے ترتیب اے/بی آزمائش** کی—روبوٹ چلانے والے شخص کو معلوم نہیں تھا کہ کون سی پالیسی آزمائی جا رہی ہے، اور ترتیب بھی بے ترتیب تھی۔
- **قابو شدہ اور دوبارہ بنائے جا سکنے والے ابتدائی حالات** رکھے—ہر آزمائش سے پہلے منظر کو تصویری اوورلے کے مطابق کیا گیا۔
- **بڑی تعداد میں آزمائشیں** کیں—ہر کام، پالیسی اور حالت کے لیے حقیقی دنیا میں 50 رول آؤٹس؛ سیمولیشن میں ہر کام کے لیے 200۔ مجموعی طور پر تقریباً **1,800 بلائنڈ حقیقی دنیا کے رول آؤٹس اور 47,000 سے زیادہ سیمولیشن رول آؤٹس۔**
- **مناسب شماریات** استعمال کیں— کامیابی کے احتمال کے بیزیائی تخمینے، جوڑی وار مفروضہ آزمائشیں اور متعدد موازنوں کی اصلاح؛ صرف ایک دوسرے پر چڑھتی غلطی کی سلاخیں دیکھ کر فیصلہ نہیں کیا۔ انسانی اسکور والی آزمائشوں کے ایک چوتھائی حصے پر معیار جانچ بھی کی گئی تاکہ اسکورنگ کی غلطی ناپی جا سکے۔

[video pending]

ناشتے کی میز لگانے والے ماڈلز کا ساتھ ساتھ موازنہ: بائیں طرف ایک-کام بنیادی ماڈل، دائیں طرف LBM۔ دونوں ویڈیوز اصل رفتار پر ہیں۔ یہ صرف ایک جانچا گیا کام ہے، عمومی خودمختاری کا ثبوت نہیں۔

یہ پوری جانچ ہی اصل نکتہ ہے۔ مقالہ بنیادی طور پر یہ دلیل دیتا ہے کہ اس کے بغیر حقیقی بہتری اور محض اتفاق میں فرق کرنا مشکل ہے۔

انہیں کیا ملا

مزید مخصوص تربیت پانے والے بڑے ماڈلز نے اوسطاً صفر سے تربیت یافتہ ایک-کام ماڈلز کو شکست دی۔ تمام کاموں کو ملا کر، پہلے سے تربیت یافتہ اور پھر خاص کام پر مزید تربیت پانے والا LBM سیمولیشن اور حقیقی دنیا دونوں میں اسی کام کے لیے صفر سے تربیت یافتہ پالیسی سے قابلِ اعتماد طور پر بہتر رہا، اور فرق شماریاتی طور پر معنی خیز تھا۔ الگ الگ کاموں پر یہ LBM تقریباً ہر صورت میں شماریاتی طور پر کم از کم برابر یا بہتر تھا: حقیقی دنیا میں 3 میں سے 3، سیمولیشن میں 16 میں سے 15 کام۔

سب سے واضح فائدہ ڈیٹا کی بچت تھا۔ مخصوص اضافی تربیت پانے والا LBM تقریباً **3–5 گنا کم کام-مخصوص ڈیٹا** سے صفر سے تربیت والے ماڈل جیسی کارکردگی تک پہنچ گیا۔ حقیقی دنیا کے ایک ناشتے کی میز والے کام میں صرف **15 فیصد** عملی مثالوں پر مزید تربیت پانے والے LBM نے اس پالیسی کو شکست دی جسے صفر سے **100 فیصد** ڈیٹا پر تربیت دی گئی تھی۔

حالات بدلیں تو ابتدائی تربیت کا فائدہ زیادہ ہوا۔ جب آزمائشی ماحول کو جان بوجھ کر تربیتی حالات سے مختلف کیا گیا—یعنی "تقسیم میں تبدیلی"—تو مزید تربیت پانے والے LBM کی برتری بڑھ گئی۔ ایک سیمولیشن مجموعے میں معمول کے حالات میں وہ 16 میں سے 3 کاموں پر شماریاتی طور پر بہتر تھا، مگر تقسیم بدلنے پر 16 میں سے 10 پر۔ حقیقی تعیناتیاں ہمیشہ تربیتی حالات سے کچھ نہ کچھ مختلف ہوتی ہیں، اس لیے یہ مضبوطی شاید سب سے عملی نتیجہ ہے۔

مزید ابتدائی تربیتی ڈیٹا نے بتدریج مدد کی۔ ڈیٹا بڑھنے کے ساتھ کارکردگی مسلسل بڑھی؛ آزمائے گئے پیمانوں پر کوئی اچانک "ابھرتی

ہوئی" صلاحیت یا چھلانگ نہیں دیکھی گئی۔ نتیجہ مفید، قابلِ پیش گوئی اور غیر ڈرامائی تھا۔

لیکن بغیر مخصوص اضافی تربیت کے عمومی ماڈل کی کہانی قائم نہیں رہی۔ پہلے سے تربیت یافتہ LBM کو نئے کام پر اضافی تربیت کے بغیر استعمال کیا گیا تو وہ ایک-کام پالیسیوں کو مستقل طور پر شکست نہیں دے سکا۔ ایک ھی نیٹ ورک کئی کام کر سکتا تھا، مگر "صرف ھدایت دو اور کام ھو جائے گا" والی امید یہاں پوری نہیں ہوئی؛ مصنفین اس کی ایک وجہ اپنے نسبتاً چھوٹے زبان اینکوڈر کی نازکی کو قرار دیتے ھیں۔

فوائد اتنے چھوٹے تھے کہ انہیں نظر انداز کرنا—یا غلطی سے بنا لینا—آسان تھا۔ کئی اثرات صرف معمول سے بڑے نمونوں اور محتاط جانچ کے بعد واضح ہوئے۔ مصنفین صاف لکھتے ھیں کہ اثرات کے حجم اور شور کو دیکھتے ہوئے **یہ حقیقی خطرہ ھے کہ روبوٹکس کے بہت سے مقالات شماریاتی شور ناپ رہے ہوں۔** انہوں نے یہ بھی پایا کہ ایک بظاھر معمولی انتخاب—ڈیٹا کو کیسے معمول پر لایا جائے—نے نتائج پر ساختی تبدیلیوں سے زیادہ اثر ڈالا، اور ابتدائی تربیت میں normalization کی ایک **خرابی** جانچ مکمل ہونے کے بعد سامنے آئی۔

اس کا غالب مطلب کیا ھے

محفوظ تعبیر یہ ھے کہ متنوع روبوٹ ڈیٹا پر بڑے پیمانے کی ابتدائی تربیت واقعی مفید ھے: نئے کام کے لیے کم ڈیٹا درکار ہوتا ھے اور جب دنیا تربیتی حالات سے پوری طرح نہ ملے تو پالیسی زیادہ مضبوط رھتی ھے۔ یہ اس سمت کی حقیقی تائید ھے جس پر شعبہ سرمایہ لگا رہا ھے۔ لیکن فائدے **محدود اور مشروط** ھیں—زیادہ تر اضافی مخصوص تربیت کے بعد نظر آتے ھیں، اور مجموعی یا دباؤ والے حالات میں سب سے واضح ھیں۔ یہ کوئی فوراً استعمال ہونے والا عمومی روبوٹ نہیں۔

زیادہ خاموش مگر شاید زیادہ اھم نتیجہ طریقۂ کار سے متعلق ھے۔ یہ مقالہ دراصل پیمائش کا معیار بن جاتا ھے: یہ دکھاتا ھے کہ روبوٹ پالیسی کے بارے میں قابلِ اعتماد دعویٰ کرنے کے لیے کتنا شواھد درکار ھے، اور اشارہ کرتا ھے کہ شائع شدہ جوش کا خاصا حصہ شاید اس سے کم شواھد پر قائم ھے۔ شعبے کو ایک اور ماڈل سے زیادہ شاید اسی اصلاح کی ضرورت ھے۔

یہ کیا ثابت نہیں کرتا

- یہ **عمومی مقصد کا روبوٹ نہیں۔** فائدے ایک خاص ساخت—dffiusion پالیسیوں—کے لیے، ھر کام پر مزید مخصوص تربیت کے بعد، قابو شدہ حالات اور teleoperation سے جمع کی گئی مثالوں میں دکھائے گئے ھیں؛ کسی ایسے خودمختار روبوٹ میں نہیں جو حکم ملتے ھی من مانا نیا کام کر دے۔
- یہ **بغیر مخصوص اضافی تربیت کے استعمال** کی توثیق نہیں کرتا۔ اضافی تربیت کے بغیر بڑا ماڈل ایک-کام بنیادی ماڈلز سے مستقل طور پر بہتر نہیں رھا۔
- یہ **"اچانک ابھرتی صلاحیت"** کا شواھد نہیں۔ پیمانہ بڑھنے سے کارکردگی ھموار انداز میں بہتر ہوئی؛ یہاں کوئی اچانک discontinuity نہیں۔
- اعداد **نسبتی اور لیب تک محدود** ھیں۔ موازنے کو حساس بنانے کے لیے مطلق کامیابی کی شرحیں جان بوجھ کر تقریباً 50 فیصد کے آس پاس رکھی گئیں؛ یہ حقیقی دنیا میں مجموعی قابلِ اعتماد کارکردگی کی پیمائش نہیں، اور تحقیق ایک لیب اور ایک ساخت تک محدود ھے۔
- یہ یہ **طے نہیں** کرتا کہ کوئی خاص پالیسی کیوں کامیاب یا ناکام ہوتی ھے؛ بعض کاموں میں بڑا ماڈل بدتر رھا، مگر وجہ واضح نہیں کی گئی۔
- جانچ کے نظام سے باھر **حفاظت، خودمختاری یا عملی تعیناتی** کے بارے میں یہ کچھ نہیں کہتا۔

شواھد کتنے مضبوط ھیں؟

مرکزی تقابلی دعووں کے لیے—کہ مخصوص اضافی تربیت پانے والے LBM مجموعی طور پر بہتر ھیں، کئی گنا کم ڈیٹا مانگتے ھیں اور تقسیم بدلنے پر زیادہ مضبوط رھتے ھیں—شواھد مضبوط اور غیر معمولی طور پر اچھی طرح قابو شدہ ھیں: بلائنڈ، بے ترتیب، بڑے نمونے، شماریاتی آزمائشیں، اور اسکورنگ پر معیار جانچ۔ یہ ان نایاب صورتوں میں سے ھے جہاں طریقۂ کار اتنا مضبوط ھے کہ مرکزی نتائج کو تقریباً ان کی ظاھری شکل میں قبول کیا جا سکتا ھے۔

اھم احتیاطیں خود مصنفین اٹھاتے ھیں۔ ان کی غلطی کی سلاخیں **جانچ** کی بے ترتیبی کو سمیٹتی ھیں، **تربیت** کی بے ترتیبی کو نہیں—ایک ھی ماڈل کو دو بار تربیت دینے سے معنی خیز طور پر مختلف پالیسی مل سکتی ھے، اور یہ تفاوت شماریات میں شامل نہیں۔ حقیقی دنیا کے ھر کام میں 50 آزمائشیں تھیں، درمیانے اثرات پکڑنے کے لیے کافی مگر چھوٹے اثرات چھوٹ سکتے ھیں۔ زبان کی شرط بندی میں نسبتاً چھوٹا اینکوڈر استعمال ھوا، اس لیے بڑے نظاموں میں "روبوٹ کو بس بتا دو کیا کرنا ھے" والا نتیجہ مختلف ھو سکتا ھے۔ اور normalization کی خرابی کا بعد میں سامنے آنا بھی اھم ھے۔ یہ حدود مرکزی نتیجے کو نہیں گراتیں، مگر وھی چیزیں ھیں جنہیں مقالہ کہتا ھے کہ شعبہ اکثر نظر انداز کرتا ھے۔

ماخذ کے بارے میں ایک نوٹ بھی اسی دیانت کے ساتھ: یہ وضاحتی مضمون مصنفین کے پری پرنٹ پر مبنی ھے۔ جریدے میں شائع شدہ نسخہ دستیاب نہیں ھو سکا، اس لیے پری پرنٹ اور حتمی متن کے ممکنہ فرق یہاں نہیں جانچے گئے۔

یہ کیوں اھم ھے

"روبوٹ فاؤنڈیشن ماڈلز" ایسی اصطلاح ھے جس سے مبالغہ آسان ھے، اور اس مطالعے کو دونوں سمتوں میں غلط پڑھا جا سکتا ھے—فتح مندانہ "یہ کام کرتا ھے!" یا مسترد کرنے والے "سب ھائپ ھے" کے طور پر۔ درست تعبیر دونوں سے زیادہ مفید ھے: متنوع ڈیٹا پر ابتدائی تربیت حقیقی، قابلِ پیمائش مگر درمیانے درجے کے فائدے دیتی ھے—خاص طور پر ھر کام کے لیے کم ڈیٹا اور زیادہ مضبوطی—اور پیمانے کے ساتھ کارکردگی قابلِ پیش گوئی انداز میں بہتر ھوتی ھے۔

زیادہ گہری اھمیت یہ ھے کہ مقالہ اپنی سخت جانچ خود اپنے شعبے پر لاگو کرتا ھے۔ اگر حقیقی اثرات اتنے چھوٹے ھیں کہ کمزور جانچ میں غائب ھو جائیں، اور ڈیٹا normalization جیسا معمولی فیصلہ کسی نئی پیچیدہ ساخت سے زیادہ فرق ڈال دے، تو روبوٹ سیکھنے میں "پیش رفت" کہلانے والی بہت سی چیزوں کو زیادہ مضبوط پیمائش درکار ھے۔ نتیجے اور خواھش کے درمیان فرق کی نگرانی کرنا leaderboard میں اوپر آنے سے زیادہ نایاب اور زیادہ قیمتی کام ھے۔

سادہ خلاصہ

Toyota Research Institute کے محققین نے "بڑے رویّاتی ماڈلز" تربیت دیے—dffiusion پر مبنی روبوٹ پالیسیوں کو تقریباً 1,700 گھنٹے کے متنوع اشیا سنبھالنے کے ڈیٹا پر پہلے سے تربیت دی گئی—اور انہیں صفر سے تربیت یافتہ ایک-کام پالیسیوں سے غیر معمولی طور پر سخت طریقے سے موازنہ کیا: بلائنڈ، بے ترتیب اور بڑے نمونے کی جانچ، تقریباً 1,800 حقیقی دنیا اور 47,000 سے زیادہ سیمولیشن آزمائشوں کے ساتھ۔ ھر کام پر اضافی مخصوص تربیت کے بعد بڑے ماڈلز مجموعی طور پر قابلِ اعتماد طور پر بہتر رھے، تقریباً **3–5 گنا کم کام-مخصوص ڈیٹا** مانگا، اور **حالات بدلنے پر زیادہ مضبوط تھے**؛ ابتدائی تربیتی ڈیٹا بڑھانے سے کارکردگی ھموار انداز میں بہتر ھوئی۔ لیکن اضافی مخصوص تربیت کے بغیر وہ ایک-کام ماڈلز سے مستقل طور پر بہتر نہیں تھے، بعض اثرات اتنے چھوٹے تھے کہ صرف بڑے نمونے میں نظر آئے، اور ڈیٹا normalization کا معمولی انتخاب ساخت سے زیادہ اھم نکلا۔ یہ روبوٹ فاؤنڈیشن ماڈل

کی سمت کے لیے مضبوط مگر ناپ تول کر دی گئی تائید ہے—عمومی روبوٹ، بغیر مخصوص تربیت کے generalist یا "اچانک ابھرتی صلاحیت" کا ثبوت نہیں—اور ساتھ تنبیہ کہ روبوٹکس کی بہت سی تحقیق شاید شور ناپ رہی ہو۔

بلا مبالغہ جائزہ

مقالہ کیا دکھاتا ہے: سخت، بلائنڈ اور شماریاتی طاقت رکھنے والی جانچ میں—تقریباً 1,800 حقیقی دنیا اور 47,000 سے زیادہ سیمولیشن رول آؤٹس—متعدد کاموں پر پہلے سے تربیت یافتہ اور پھر خاص کام پر مزید تربیت پانے والی dffiusion پالیسیاں مجموعی طور پر صفر سے تربیت یافتہ ایک-کام پالیسیوں سے بہتر رہیں، تقریباً 3–5 گنا کم خاص ڈیٹا سے برابر کارکردگی تک پہنچیں، اور تقسیم بدلنے پر زیادہ مضبوط تھیں۔ کارکردگی ابتدائی تربیتی ڈیٹا کے ساتھ ہموار انداز میں بڑھی۔

کیا قرینِ قیاس ہے مگر ثابت نہیں: یہی فائدے کہیں بڑے vision-language-action ماڈلز میں بھی برقرار رہیں؛ اور آزمائے گئے ڈیٹا دائرے سے آگے بھی یہی ہموار scaling جاری رہے۔

کیا نہیں دکھاتا: عمومی مقصد یا بغیر مخصوص تربیت والا روبوٹ؛ اچانک "ابھرتی" صلاحیت؛ خاص کاموں کی ناکامی کی وضاحت؛ مطلق حقیقی دنیا قابلِ اعتماد کارکردگی؛ یا حفاظت اور خودمختار عملی تعیناتی کے بارے میں نتیجہ۔

اہم حدود: شماریات جانچ کی بے ترتیبی کو سمیٹتی ہیں، تربیتی رن کی نہیں؛ ہر حقیقی کام کے 50 ٹرائل چھوٹے اثرات چھوڑ سکتے ہیں؛ ایک ساخت اور ایک لیب؛ نسبتاً چھوٹا زبان اینکوڈر؛ جانچ کے بعد ملی normalization خرابی؛ اور تجزیہ پری پرنٹ پر مبنی ہے، حتمی شائع شدہ نسخہ نہیں جانچا گیا۔

عام قاری کو کتنا اعتماد ہونا چاہیے؟ بلند اعتماد کہ ابتدائی تربیت اور پھر مخصوص اضافی تربیت حقیقی مگر درمیانے فائدے دیتی ہے—خاص طور پر ڈیٹا کی بچت اور مضبوطی—اور انہیں غیر معمولی احتیاط سے ناپا گیا۔ بلند اعتماد کہ یہ عمومی مقصد یا بغیر مخصوص تربیت والا روبوٹ نہیں اور نہ اچانک ابھرتی صلاحیت۔ بڑے پیمانوں تک فائدہ بڑھنے پر درمیانہ اعتماد۔ مصنفین کی اپنی تنبیہ سنجیدگی سے لینے کے قابل ہے: اگر اثر چھوٹے ہوں تو کم طاقت والی تحقیق شور کو نتیجہ سمجھ سکتی ہے۔ مناسب رویہ طریقے کے بارے میں محتاط امید اور کمزور شماریاتی بنیاد والے روبوٹ اے آئی دعووں پر صحت مند شک ہے۔

ماخذ

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/