The Clean Paper

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

Vai lieli robotu "foundation modeļi" tiešām strādā labāk? Rūpīga atbilde — mēreni jā, un daudzi pētījumi to nespēj droši izmērīt

Toyota Research Institute apmācīja "large behavior models" — robotu politikas, kas priekšapmācītas uz ~1700 stundām daudzveidīgu manipulācijas datu — un ar neparasti stingru protokolu salīdzināja tās ar viena uzdevuma modeļiem no nulles: akli, randomizēti, lieli paraugi (~1800 reāli, >47 000 simulācijas) un formāla statistika. Pēc uzdevuma specifiskas papildapmācības lielie modeļi vidēji darbojās labāk, prasīja ~3–5× mazāk konkrētā uzdevuma datu un bija noturīgāki mainītos apstākļos; veiktspēja gludi auga ar priekšapmācības datiem. Taču bez papildapmācības tie konsekventi nepārspēja viena uzdevuma modeļus, vairāki efekti bija tik mazi, ka vajadzēja lielus paraugus, un parasta datu normalizācija ietekmēja rezultātu vairāk nekā arhitektūra. Tas ir izmērīts atbalsts robotu foundation modeļu virzienam — ne universāls robots, ne zero-shot ģenerālists, ne "emergents lēciens" — un brīdinājums, ka daļa robotikas var mērīt troksni.

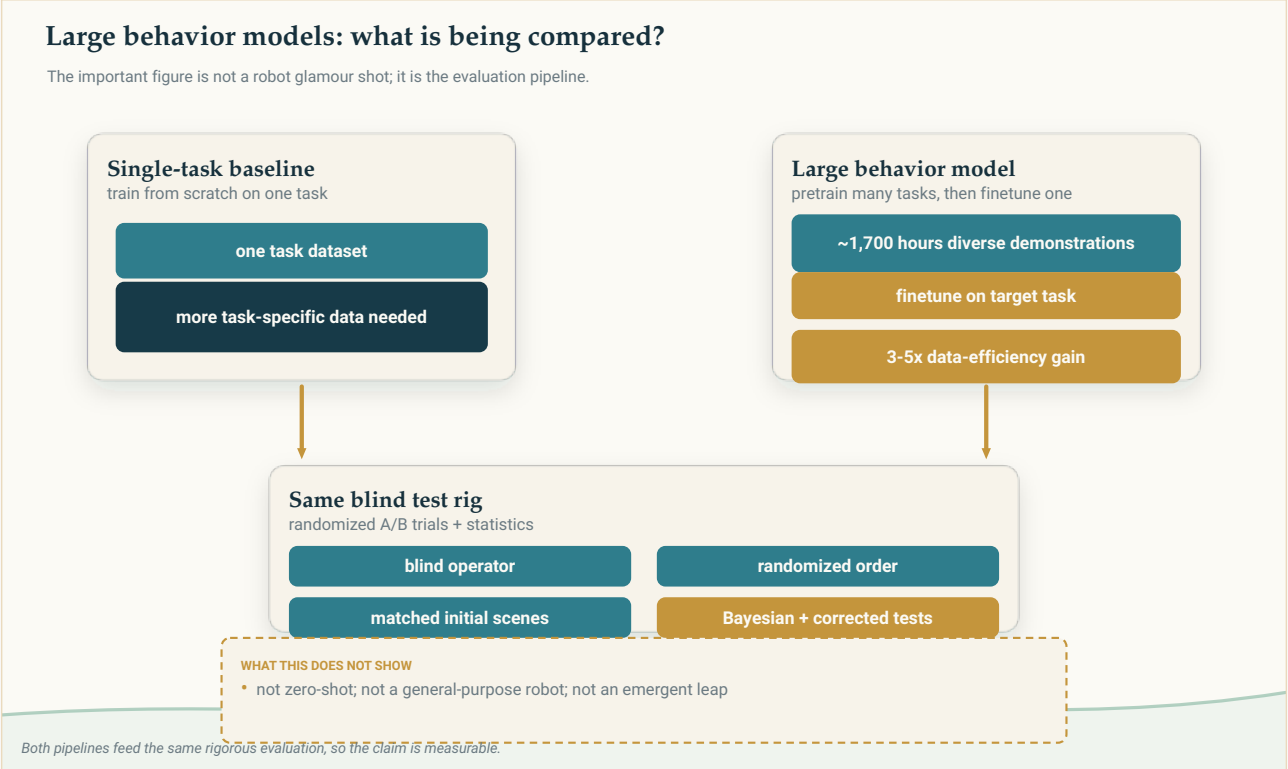
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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: 10.1126/scirobotics.aea6201

Robotikas raksts, kura īstā tēma ir godīgums

Robotikā pašlaik valda „foundation model" vilnis. Ideja, aizgūta no valodas un attēlu MI, ir vilinoša: tā vietā, lai robotu apmācītu katram uzdevumam atsevišķi, apmācīt vienu lielu **„large behavior model" (LBM)** uz milzīga, daudzveidīga demonstrāciju kopuma un iegūt sistēmu, kas spēj daudz ko un ātri pielāgojas. Entuziasms — un investīcijas — ir milzīgs. Virsraksts rakstās pats: *vispārējas nozīmes robotu smadzenes ir klāt*.

Toyota Research Institute darbs ir interesants tieši tāpēc, ka šo virsrakstu atsakās rakstīt. Tā galvenais devums nav iespaidīgāks robots. Tas ir rūpīgs jautājums: **vai lielā modeļa pieeja tiešām strādā labāk, un kā mēs to vispār varētu droši zināt?** Autori atbild ar statistisku rūpību, kas, pēc viņu pašu teiktā, robotikā nav ierasta. Rezultāts ir noderīgs „jā, bet" — un brīdinājums, ka daļa robotikas literatūras var mērīt troksni.



Abas pieejas nonāk vienā un tajā pašā aklā, randomizētā vērtējumā. Raksta apgalvojums nav, ka robotu foundation modeļi bez papildu apmācības ir universāli; tas ir, ka priekšapmācība pēc rūpīgas mērīšanas var uzlabot datu efektivitāti un noturību.

Original The Clean Paper diagram CC BY 4.0

Ko autori darīja

LBM šeit ir konkrēts objekts: **difūzijā balstīta vizuomotorā politika** — Diffusion Transformer, kas nolasa kameru attēlus, īsu valodas instrukciju un paša robota locītavu stāvokļus un 10 Hz frekvencē

izvada īsas kustību komandu sekvences. Modeļus **priekšapmācīja uz aptuveni 1700 stundām** robotu demonstrāciju — vairāk nekā 500 dažādiem iekšēji savāktiem uzdevumiem plus publiskām datu kopām — un pēc tam **papildus apmācīja** katram atsevišķam uzdevumam. Salīdzinājums viscaur ir ar **viena uzdevuma politiku, kas apmācīta no nulles** tikai uz konkrētā uzdevuma datiem.

Taču raksta sirds ir **vērtēšana**, ko autori faktiski uzskata par galveno rezultātu. Lai neapmānītu paši sevi, viņi izmantoja:

- **Aklu, randomizētu A/B testēšanu** reālajā pasaulē — cilvēks, kas darbināja robotu, nezināja, kura politika tiek pārbaudīta, un secība bija nejauša.
- **Kontrolētus, atkārtojamus sākuma apstākļus** — pirms katra mēģinājuma operators salīdzināja ainu ar attēla pārklājumu.
- **Lielu mēģinājumu skaitu** — 50 reālus izpildījumus katram uzdevumam, politikai un nosacījumam; simulācijā 200. Kopā aptuveni **1800 aklu reālās pasaules izpildījumu un vairāk nekā 47 000 simulāciju**.
- **Pilnvērtīgu statistiku** — panākumu varbūtības Beijesa novērtējumus un pāru hipotēžu testus ar korekcijām vairākiem salīdzinājumiem, nevis kļūdu stabiņu vizuālu salīdzināšanu. Ceturtdaļai cilvēku novērtēto mēģinājumu viņi veica arī kvalitātes kontroles atkārtotu vērtējumu, lai izmērītu vērtētāja kļūdu.

[video pending]

Modeļu salīdzinājums brokastu galda uzklāšanā: pa kreisi viena uzdevuma bāzes modelis, pa labi LBM. Abi video ir normālā 1× ātrumā. Tas ir viens vērtēts uzdevums, nevis pierādījums universālai autonomijai.

Šī mērīšanas sistēma ir raksta būtība. Autori argumentē, ka bez tās īstu uzlabojumu no veiksmes atšķirt nevar.

Ko viņi atklāja

Pēc papildapmācības lielie modeļi vidēji pārspēja no nulles apmācītos viena uzdevuma modeļus. Apkopojot uzdevumus, priekšapmācīts un pēc tam konkrētam uzdevumam papildapmācīts LBM gan simulācijā, gan reālajā pasaulē uzticami pārspēja politiku, kas uz to pašu uzdevumu apmācīta no nulles, un atšķirība bija statistiski nozīmīga. Atsevišķos uzdevumos papildapmācītais LBM gandrīz vienmēr bija statistiski tikpat labs vai labāks (3/3 reālajos uzdevumos, 15/16 simulācijās).

Lielākais un skaidrākais ieguvums ir datu efektivitāte. Papildapmācīts LBM sasniedza no nulles apmācīta modeļa veiktspēju ar aptuveni **3–5 reizes mazāk konkrētā uzdevuma datu**. Vienā reālā uzdevumā — brokastu galda uzklāšanā — LBM, kas papildapmācīts tikai uz **15%** demonstrāciju, pārspēja no nulles apmācītu politiku, kas izmantoja **100%** demonstrāciju.

Priekšapmācība visvairāk palīdz, kad apstākļi mainās. Kad testa vide apzināti tika nobīdīta prom no apmācības apstākļiem (*distribution shift*), papildapmācītā LBM priekšrocība **palielinājās**. Vienā simulāciju komplektā normālos apstākļos tas statistiski pārspēja bāzes modeli 3 no 16 uzdevumiem,

bet nobīdītos apstākļos — 10 no 16. Tā kā reālā lietojumā apstākļi gandrīz vienmēr atšķiras no apmācības datiem, šī noturība, iespējams, ir praktiski svarīgākais rezultāts.

Vairāk priekšapmācības datu palīdzēja pakāpeniski. Veiktspēja stabili auga, palielinot priekšapmācības datu apjomu, bez pēkšņa „emergenta" lēciena pārbaudītajā mērogā. Noderīgi, paredzami, bez drāmas.

Taču stāsts par vispārīgu modeli bez papildapmācības neizturēja. Priekšapmācīts LBM **zero-shot** režīmā — bez konkrētā uzdevuma papildapmācības — konsekventi nepārspēja viena uzdevuma politikas. Viena neironu tīkla spēja veikt daudzus uzdevumus bija reāla, taču „vienkārši pasaki robotam, ko darīt" šeit netika apstiprināts. Autori daļu problēmas saista ar pieticīgo valodas kodētāju.

Un ieguvumi bija pietiekami mazi, lai tos viegli nepamanītu — vai sajauktu ar efektu. Daudzi efekti kļuva redzami tikai ar neparasti lielu mēģinājumu skaitu un pienācīgiem testiem. Autori tieši raksta, ka, ņemot vērā efektu izmēru un troksni, pastāv **ievērojams risks, ka daudzi robotikas raksti mēra statistisku troksni**. Viņi arī atrada, ka ikdienišķa izvēle — kā dati tiek **normalizēti** — rezultātus ietekmēja vairāk nekā arhitektūras izmaiņas, un priekšapmācības normalizācijas **kļūda** tika atklāta tikai pēc vērtēšanas pabeigšanas.

Ko tas, visticamāk, nozīmē

Drošais secinājums: liela mēroga priekšapmācība uz daudzveidīgiem robotu datiem ir īsts, vērtīgs komponents. Tā samazina konkrētam jaunam uzdevumam vajadzīgo datu daudzumu un padara politikas noturīgākas, kad reālā pasaule neatbilst apmācībai. Tas tiešām atbalsta virzienu, uz kuru lauks liek likmes. Taču ieguvumi ir **mēreni un nosacīti** — pārsvarā pēc papildapmācības, vislabāk redzami apkopojumā un stresa apstākļos —, nevis gatava universāla robota ierašanās.

Klusākais, iespējams svarīgākais secinājums ir metodoloģisks. Raksts faktiski piedāvā mērīšanas standartu: cik daudz pierādījumu vajadzīgs, lai apgalvojumu par robota politiku varētu uzticami pieņemt. Un tas liek domāt, ka daļa publicētā entuziasma balstās pārāk mazos paraugos. Šāds korektīvs laukam var būt vērtīgāks par vēl vienu modeli.

Ko tas nepierāda

- Tas **nav universāls robots**. Uzvaras demonstrētas konkrētai arhitektūrai — difūzijas politikām — pēc uzdevuma specifiskas papildapmācības, kontrolētos apstākļos un no teleoperētām demonstrācijām; tas nav autonoms robots, kas pēc komandas dara jebkuru jaunu darbu.
- Tas **neapstiprina zero-shot lietojumu**. Bez papildapmācības lielais modelis konsekventi nepārspēja viena uzdevuma bāzes modeļus.
- Tas **nav pierādījums „emergentam lēcienam"**. Mērogošana uzlaboja rezultātus vienmērīgi, bez diskontinuitātes.
- Skaitļi ir **relatīvi un laboratorijas ietvarā**. Absolūtie panākumu rādītāji apzināti tika noregulēti

ap ~50%, lai salīdzinājums būtu jutīgs; tie nav reālās pasaules uzticamības mērījums, un darbs pēta vienu arhitektūru vienā laboratorijā.
- Tas **neizskaidro**, kāpēc konkrēta politika izdodas vai neizdodas; vairāki uzdevumi, kuros lielais modelis bija sliktāks, tiek ziņoti, bet ne pilnībā izskaidroti.
- Tas **nepasaka neko par drošību, autonomiju vai ieviešanu** ārpus vērtēšanas stenda.

Cik pārliecinoši ir pierādījumi?

Centrālajiem salīdzinošajiem apgalvojumiem — ka papildapmācīti LBM apkopojumā pārspēj no nulles apmācītus modeļus, prasa vairākas reizes mazāk datu un ir noturīgāki pie sadalījuma nobīdes — pierādījumi ir stipri un neparasti labi kontrolēti: akli, randomizēti, ar lielu paraugu, formāli statistiski testēti un ar cilvēku vērtējumu kvalitātes pārbaudi. Šeit metodoloģija ir pietiekami stipra, lai galvenos secinājumus uztvertu gandrīz tieši.

Godīgās atrunas autori paši izceļ. Kļūdu intervāli ietver **vērtēšanas** nejaušību, bet ne **apmācības** nejaušību — divas viena modeļa apmācības var dot jūtami atšķirīgas politikas, un šī variācija statistikā nav iekļauta. Reālajos uzdevumos bija 50 mēģinājumu katram, pietiekami vidējiem efektiem, bet ne vienmēr maziem. Valodas nosacījums izmantoja pieticīgu kodētāju, tāpēc secinājumi par „vienkārši pasaki, ko darīt" var atšķirties lielākās sistēmās. Un ir arī atklāti ziņotā normalizācijas kļūda, kas tika atrasta pēc eksperimentiem. Neviens no šiem punktiem neiznīcina galvenos secinājumus, bet tie ir tieši tās lietas, ko raksts apgalvo, ka lauks pārāk bieži ignorē.

Avotu piezīme tādā pašā garā: šis skaidrojums balstās autoru preprintā. Žurnālā publicēto versiju mums neizdevās iegūt, tāpēc iespējamās izmaiņas starp preprintu un galīgo rakstu nav pārbaudītas.

Kāpēc tas ir svarīgi

„Robotu foundation modeļi" ir frāze, kas gandrīz aicina pārspīlēt, un šo pētījumu viegli nolasīt abos grāvjos — kā triumfālu „tas strādā!" vai nicinošu „tas viss ir pārvērtēts". Precīzā versija ir noderīgāka par abām: daudzveidīgu datu priekšapmācība dod reālus, izmērāmus, bet mērenus ieguvumus — galvenokārt mazāk datu vienam uzdevumam un lielāku noturību —, un veiktspēja ar mērogu uzlabojas paredzami.

Dziļākais iemesls, kāpēc raksts ir nozīmīgs, ir tas, ka tā stingrība tiek vērsta pret pašu lauku. Parādot, ka īstie efekti ir pietiekami mazi, lai paviršā vērtēšanā pazustu, un ka garlaicīga datu normalizācija var būt svarīgāka par gudru arhitektūru, raksts argumentē, ka robotu mācīšanās progresam vajag daudz stingrāku mērīšanu. Darbs, kas savu uzticamību izmanto, lai sargātu robežu starp rezultātu un vēlmi, dara ko retāku un vērtīgāku par kārtējo līderpozīciju.

Īss kopsavilkums

Toyota Research Institute pētnieki apmācīja „large behavior models" — difūzijā balstītas robotu politikas, kas priekšapmācītas uz aptuveni 1700 stundām daudzveidīgu manipulācijas datu — un salīdzināja tās ar viena uzdevuma modeļiem, kas apmācīti no nulles, izmantojot neparasti stingru protokolu: akli, randomizēti, lieli paraugi (≈1800 reāli un >47 000 simulācijas mēģinājumu) un formāla statistika. Pēc konkrēta uzdevuma papildapmācības lielie modeļi vidēji darbojās labāk, sasniedza līdzvērtīgu veiktspēju ar aptuveni **3–5 reizes mazāk konkrētā uzdevuma datu** un bija **noturīgāki, mainoties apstākļiem**; veiktspēja gludi auga līdz ar priekšapmācības datiem. Taču **bez papildapmācības** tie konsekventi nepārspēja viena uzdevuma modeļus, daudzi efekti bija tik mazi, ka tos atklāja tikai lielais mēģinājumu skaits, un parasta datu normalizācijas izvēle ietekmēja rezultātu vairāk nekā arhitektūra. Tas ir stingrs, izmērīts atbalsts robotu foundation modeļu virzienam — **ne** universāls robots, ne zero-shot ģenerālists un ne „emergents lēciens" — plus ass brīdinājums, ka daļa robotikas literatūras var mērīt troksni.

Bez pārspīlējumiem

Ko raksts parāda: rūpīgā aklā un statistiski jaudīgā vērtējumā (≈1800 reāli + >47 000 simulācijas izpildījumu) daudzuzdevumu priekšapmācītas un pēc tam papildapmācītas difūzijas politikas apkopojumā pārspēj no nulles apmācītās; sasniedz līdzvērtīgu rezultātu ar ~3–5× mazāk uzdevuma datu; ir noturīgākas pie sadalījuma nobīdes; veiktspēja gludi mērogojas ar priekšapmācības datiem.

Kas ir ticams, bet nav pierādīts: ka šie ieguvumi pāries uz daudz lielākiem vision-language-action modeļiem; ka gludā mērogošana turpināsies ārpus pārbaudītā datu diapazona.

Ko tas neparāda: universālu vai zero-shot robotu; kādu „emergentu" spēju lēcienu; konkrētu uzdevumu neveiksmju pilnu skaidrojumu; absolūtu reālās pasaules uzticamību; drošību vai autonomu ieviešanu.

Galvenie ierobežojumi: statistika neietver apmācības palaidienu variāciju; 50 reāli mēģinājumi uz uzdevumu var nepamanīt mazus efektus; viena arhitektūra un viena laboratorija; pieticīgs valodas kodētājs; pēc vērtēšanas atklāta datu normalizācijas kļūda; analīze balstīta preprintā, ne pārbaudītā publicētajā versijā.

Cik lielai jābūt lasītāja pārliecībai? Augstai, ka priekšapmācība plus papildapmācība dod īstus, mērenus ieguvumus — īpaši datu efektivitātē un noturībā — un ka tie šeit izmērīti neparasti rūpīgi. Augstai arī par to, ka tas **nav** universāls vai zero-shot robots un **nav** emergents lēciens. Mērenai par to, cik tālu ieguvumi mērogosies uz lielākām sistēmām. Un ir vērts nopietni uztvert autoru brīdinājumu: efekti robotikā var būt tik mazi, ka nepietiekami jaudīgi pētījumi ziņo troksni.

Avoti

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

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

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