The Clean Paper

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

Böyük robot "foundation modelləri" həqiqətən daha yaxşı işləyirmi? Diqqətli cavab — mülayim şəkildə bəli, və tədqiqatların çoxu bunu ayırd edə bilmir

Toyota Research Institute "large behavior models" — ~1,700 saat müxtəlif manipulyasiya məlumatında əvvəlcədən öyrədilmiş robot policy-ləri — hazırladı və onları qeyri-adi sərtliklə sıfırdan single-task policy-lərlə yoxladı: kor, randomizə edilmiş, böyük nümunəli sınaqlar (~1,800 real-world, 47,000+ simulyasiya) və düzgün statistika ilə. Tapşırıq üzrə finetuning-dən sonra böyük modellər orta hesabla daha yaxşı idi, təxminən 3–5× az tapşırıq-spesifik məlumat tələb etdi və şərait dəyişəndə daha möhkəm oldu; pretraining məlumatı artdıqca performans hamar yüksəldi. Amma finetuning olmadan single-task modelləri ardıcıl keçmədilər, bəzi effektlər yalnız böyük nümunələrlə görünəcək qədər kiçik idi və sadə data-normalisation seçimi arxitekturadan daha çox təsir etdi. Bu, robot-foundation-model istiqamətinə ölçülü dəstəkdir — ümumi-məqsədli robot, zero-shot generalist və ya "emergent sıçrayış" deyil — üstəlik robototexnikanın çoxunun səs-küy ölçə biləcəyinə kəskin xəbərdarlıqdır.

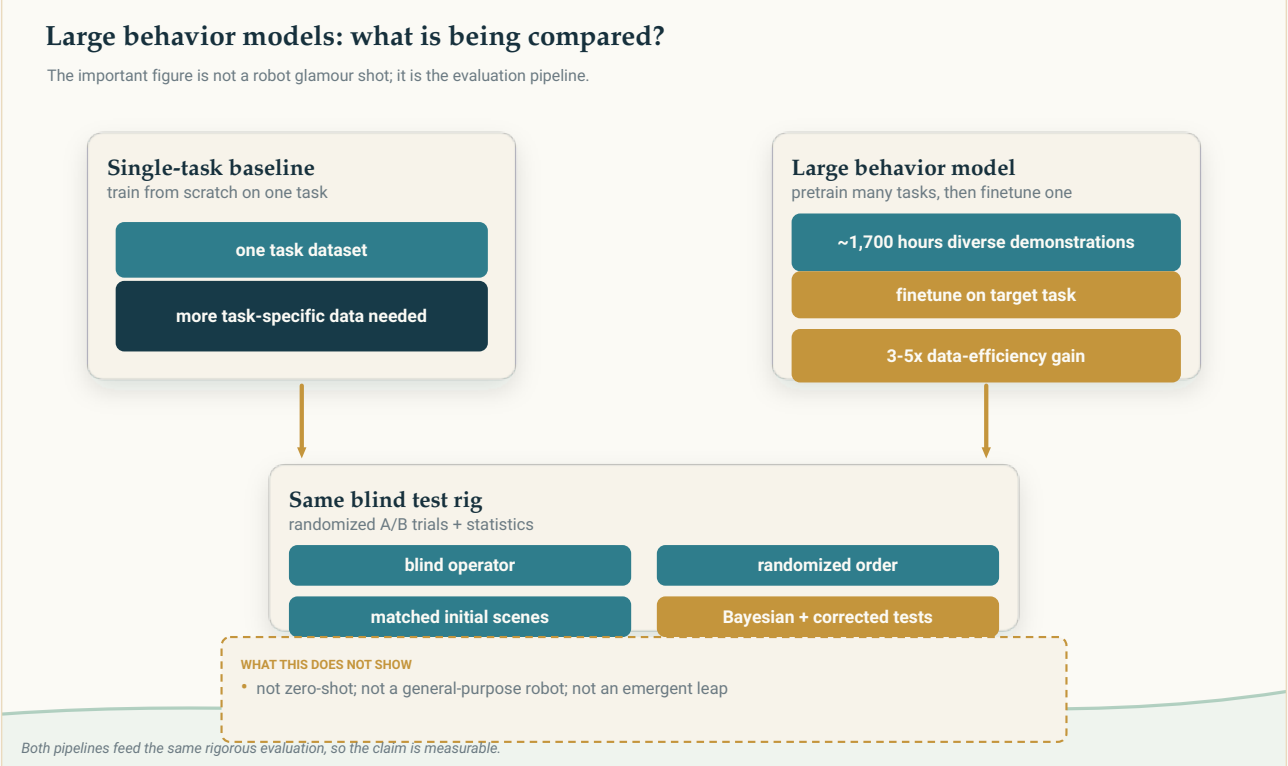
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

Əsl mövzusu dürüstlük olan robot məqaləsi

Robototexnika “foundation model” axınının ortasındadır. Dil və şəkil AI-dan götürülən ideya cəlbedicidir: robotu hər tapşırıq üçün ayrıca öyrətmək əvəzinə, bir böyük **“large behavior model” (LBM)**-i çox böyük və müxtəlif nümayişlər toplusunda öyrədin, geniş qabiliyyətli və yeni işlərə tez uyğunlaşan sistem əldə edin. Həvəs — və investisiya — nəhəngdir. Başlıq özü yazılır: *ümumi-məqsədli robot beyinləri gəldi.*

Toyota Research Institute-dan bu məqalə məhz həmin başlığı yazmaqdan imtina etdiyi üçün maraqlıdır. Onun həqiqi töhfəsi daha parlaq robot deyil. Aldadıcı dərəcədə sadə bir suala sərt baxışdır — *böyük-model yanaşması həqiqətən daha yaxşı işləyirmi və bunu ümumiyyətlə necə bilərdik?* — və müəlliflərin öz sözlərinə görə sahədə qeyri-adi statistik diqqətlə cavab verir. Nəticə həqiqətən faydalı “bəli, amma”dır və robototexnikanın böyük hissəsinin səs-küy ölçə biləcəyinə dair xəbərdarlıqdır.



Hər iki yanaşma eyni kor, randomizə edilmiş qiymətləndirməyə daxil olur. Məqalənin iddiası robot foundation modellərinin zero-shot generalist olması deyil; diqqətlə ölçüləndə pretraining-in məlumat səmərəliliyini və möhkəmliyi artıra bilməsidir.

Müəlliflər nə etdilər

Onlar LBM-ləri konkret mənada qurdular: **diffusion əsaslı visuomotor policy-lər** (kamera görüntülərini, qısa dil təlimatını və robotun öz oynaq mövqelərini oxuyan, 10 Hz-də qısa motor əmrləri paketləri çıxaran Diffusion Transformer). Bu modellər **təxminən 1,700 saat** robot nümayişində — daxildə toplanmış 500-dən çox fərqli tapşırıq və ictimai məlumat dəstlərində — **pretrain** edildi, sonra ayrı-ayrı tapşırıqlar üçün **finetune** edildi. Bütün müqayisələrdə baza həmin bir tapşırığın məlumatında **sıfırdan öyrədilmiş single-task policy**-dir.

Amma məqalənin ürəyi müəlliflərin əsas nəticə kimi təqdim etdiyi *qiymətləndirmədir*. Özlərini aldatmamaq üçün bunlardan istifadə etdilər:

- Real dünyada **kor, randomizə edilmiş A/B testləri** — robotu işlədən insan hansı policy-nin sınaqdan keçirildiyini bilmirdi və sıra təsadüfi idi.
- **Nəzarətli, təkrarlana bilən başlanğıc şəraiti** — operatorlar hər sınaqdan əvvəl səhnəni görüntü overlay-i ilə uyğunlaşdırırdı.
- **Böyük sınaq sayları** — hər tapşırıq, policy və şərait üçün 50 real-world rollout; simulyasiyada hər tapşırıq üçün 200. Ümumilikdə təxminən **1,800 kor real-world rollout və 47,000-dən çox simulyasiya rollout-u**.
- **Düzgün statistika** — üst-üstə düşən error bar-lara gözlə baxmaq əvəzinə uğur ehtimalının Bayesian qiymətləri və çoxsaylı-müqayisə düzəlişli cüt hipotez testləri. Hətta insan tərəfindən bal verilən sınaqların dörddə biri üzrə qiymətləndirmə xətasını ölçmək üçün keyfiyyət-təminatı yoxlaması apardılar.

[video pending]

Səhər yeməyi masasını hazırlayan modellərin yan-yana müqayisəsi: solda single-task baseline, sağda LBM. Hər iki video 1x sürətlə oynayır. Bu, qiymətləndirilmiş bir tapşırıqdır, ümumi-məqsədli avtonomiyanın sübutu deyil.

Bu mexanizm əsas məqamdır. Məqalənin bütöv arqumenti budur ki, onsuz həqiqi yaxşılaşmanı şansdan ayıra bilməzsiniz.

Nə tapdılar

Finetune edilmiş böyük modellər sıfırdan öyrədilmiş single-task modelləri orta hesabla keçdi. Tapşırıqlar üzrə birləşdirildikdə əvvəlcədən öyrədilib sonra konkret tapşırıq üçün finetune edilmiş LBM həm simulyasiyada, həm də real dünyada həmin tapşırığın məlumatında sıfırdan öyrədilmiş policy-dən etibarlı şəkildə yaxşı nəticə verdi və fərq statistik əhəmiyyətli idi. Ayrı tapşırıqlarda finetune edilmiş LBM demək olar hər halda statistik olaraq sıfırdan model qədər yaxşı və ya daha yaxşı idi (3/3 real-world tapşırıq, 15/16 simulyasiya tapşırığı).

Ən böyük və ən təmiz üstünlük məlumat səmərəliliyidir. Finetune edilmiş LBM təxminən

3–5× daha az tapşırıq-spesifik məlumatla sıfırdan öyrədilmiş modelə ekvivalent performansa çatdı. Real-world tapşırıqlarından birində (səhər yeməyi masasının hazırlanması) nümayişlərin yalnız **15%-i** ilə finetune edilmiş LBM onların **100%-i** ilə sıfırdan öyrədilmiş policy-ni keçdi.

Şərait dəyişəndə pretraining daha çox kömək edir. Test mühiti qəsdən təlim şəraitindən uzaqlaşdırıldıqda ("distribution shift"), finetune edilmiş LBM-nin üstünlüyü *artdı*. Bir simulyasiya dəstində normal şəraitdə 16 tapşırıqdan 3-də sıfırdan modeli statistik olaraq keçirdi, distribution shift altında isə 16-dan 10-da. Real tətbiqlər həmişə təlim şəraitindən sürüşdüyünə görə bu möhkəmlik praktik baxımdan bəlkə də ən vacib nəticədir.

Daha çox pretraining məlumatı hamar şəkildə kömək etdi. Pretraining məlumatı artdıqca performans davamlı yüksəldi — sınaqdan keçirilən miqyaslarda **qəfil sıçrayış və ya "emergent" keçid yox idi**. Faydalı, proqnozlaşdırılan, dramatik olmayan nəticə.

Amma finetuning olmadan generalist hekayəsi təsdiqlənmədi. Pretrained LBM *zero-shot* — tapşırıq-spesifik finetuning olmadan — single-task policy-ləri ardıcıl keçmədi. Bir şəbəkə çox tapşırığı eyni anda edə bilirdi, amma "sadəcə prompt ver" arzusu burada doğrulanmadı; müəlliflər bunun bir hissəsini kiçik dil enkoderinin kövrəkliyi ilə əlaqələndirirlər.

Qazanclar o qədər kiçik idi ki, onları qaçırmaq — və ya saxta şəkildə görmək — asan idi. Effektlərin çoxu yalnız adətən istifadə ediləndən daha böyük nümunə ölçüləri və diqqətli testlərlə görünürdü. Müəlliflər effekt ölçüləri və səs-küyü nəzərə alaraq **robototexnika məqalələrinin çoxunun statistik səs-küy ölçməsi riskinin əhəmiyyətli olduğunu** açıq yazırlar. Həmçinin adi bir seçim — məlumatın necə *normallaşdırılması* — arxitektura dəyişikliklərindən daha böyük təsir göstərdi və pretraining-də normallaşdırma **bug-u** yalnız qiymətləndirmələr bitdikdən *sonra* üzə çıxdı.

Bu, yəqin ki, nə deməkdir

Müdafiə edilə bilən oxunuş: müxtəlif robot məlumatlarında böyükmiqyaslı pretraining real və faydalı komponentdir — hər yeni tapşırıq üçün daha az məlumat tələb edir və dünya təlim şəraitinə uyğun gəlməyəndə policy-ləri daha möhkəm edir. Bu, sahənin böyük sərmayə qoyduğu istiqaməti həqiqətən dəstəkləyir. Amma qazanclar *mülayim və şərtlidir* (əsasən finetuning-dən sonra görünür və toplu nəticələrdə, stress altında daha aydındır), hazır ümumi robotun gəlişi deyil.

Daha sakit, daha vacib məna metodolojidir. Məqalə əslində ölçü çubuğudur: robot policy-si barədə etibarlı iddia etmək üçün nə qədər sübut lazım olduğunu göstərir və dərc olunan həyəcanın xeyli hissəsinin çox az sübuta söykənə biləcəyini ima edir. Sahənin başqa modeldən daha çox belə düzəlişə ehtiyacı var.

Bu nəyi *sübut etmir*

- Bu, **ümumi-məqsədli robot deyil**. Üstünlüklər konkret arxitekturada (diffusion policy-lər), tapşırıq üzrə finetuning ilə, nəzarətli şəraitdə, teleoperasiya nümayişlərindən göstərilib — əmr verilən istənilən yeni işi avtonom yerinə yetirən robot deyil.
- **Zero-shot** istifadəni təsdiqləmir. Finetuning olmadan böyük model single-task baseline-ları ardıcıl keçmədi.
- **"Emergent sıçrayış"** sübutu deyil. Miqyas artdıqca nəticə hamar yaxşılaşdı; "sonra birdən bacarıqlı oldu" hekayəsini dəstəkləyən diskontinuitet yoxdur.
- Rəqəmlər **nisbidir və laboratoriya ilə məhdudlaşır**. Mütləq uğur dərəcələri müqayisəni həssas etmək üçün qəsdən ~50% ətrafına tənzimlənmişdi; real-dünya etibarlılığının ölçüsü deyil və iş bir laboratoriyadan bir arxitekturadır.
- Heç bir policy-nin niyə uğur qazandığını və ya uğursuz olduğunu **həll etmir**; böyük modelin *daha pis* etdiyi bir neçə konkret tapşırıq bildirilir, amma izah olunmur.
- Qiymətləndirmə qurğusundan kənarda **təhlükəsizlik, avtonomiya və deployment barədə heç nə demir**.

Sübut nə qədər güclüdür?

Mərkəzi müqayisəli iddialar — finetune edilmiş LBM-lərin toplu şəkildə sıfırdan baseline-ları keçməsi, bir neçə dəfə az məlumat istəməsi və distribution shift altında daha möhkəm olması — üçün sübut güclü və qeyri-adi dərəcədə yaxşı nəzarətlidir: kor, randomizə edilmiş, böyük nümunəli, statistik testli və qiymətləndirmə üzrə QA yoxlamalı. Bu, metodologiyanın əsas başlıq nəticələrini demək olar olduğu kimi qəbul etmək üçün kifayət qədər sağlam olduğu nadir hallardandır.

Dürüst məhdudiyyətləri müəlliflərin özləri qaldırır. Error bar-lar *qiymətləndirmənin* təsadüfiliyini tutur, amma *təlimin* təsadüfiliyini yox — eyni modeli iki dəfə öyrədin və mənalı dərəcədə fərqli policy ala bilərsiniz; bu variasiya statistikaya daxil deyil. Real-world tapşırıqlarının hər birində 50 sınaq orta effektləri tutmağa kifayətdir, amma kiçik effektləri qaçıra bilər. Dil şərtləndirməsi mülayim enkoderdən istifadə etdiyinə görə "robota sadəcə nə etməli olduğunu de" iddiaları daha böyük sistemlərdə fərqli ola bilər. Bir də qiymətləndirmədən sonra aşkar edilmiş normallaşdırma bug-u barədə açıq etiraf var. Bunların heç biri əsas nəticələri batırmır, amma məqalənin sahənin adətən kənara süpürdüyünü dediyi şeylərin məhz özüdür.

Eyni ruhda bir mənbə qeydi: bu izah müəlliflərin preprint-inə əsaslanır. Jurnal versiyasını əldə edə bilmədiyimiz üçün preprint ilə dərc olunmuş mətn arasında dəyişiklik olub-olmadığını yoxlamamışıq.

Niyə vacibdir

“Robot foundation models” həddən artıq iddia üçün hazırlanmış ifadədir və belə tədqiqatı hər iki istiqamətdə səhv oxumaq asandır — qalib *“işləyir!”* və ya dismissive *“şişirdilib”*. Dəqiq nəticə hər ikisindən daha faydalıdır: müxtəlif məlumatlarda pretraining real, ölçülə bilən, amma mülayim faydalar verir — əsasən hər tapşırıq üçün daha az məlumat və daha çox möhkəmlik — və miqyas artdıqca yol proqnozlaşdırılan şəkildə yaxşılaşır.

Daha dərin səbəb odur ki, məqalə sərtliyini öz sahəsinə yönəldir. Həqiqi effektlərin səliqəsiz qiymətləndirmədə yox olacaq qədər kiçik olduğunu və data normalisation kimi darıxdırıcı seçimin ağıllı yeni arxitekturadan daha vacib ola bildiyini göstərməklə robot öyrənməsində “irəliləyiş”in xeyli hissəsinin inanılmaqdan əvvəl daha möhkəm ölçməyə ehtiyac duyduğunu göstərir. Etibarını nəticə ilə arzu arasındakı sərhədi qorumağa sərf edən məqalə leaderboard-dun başına çıxmaqdan daha nadir və dəyərlidir.

Təmiz xülasə

Toyota Research Institute tədqiqatçıları “large behavior models” — ~1,700 saat müxtəlif manipulyasiya məlumatında əvvəlcədən öyrədilmiş diffusion əsaslı robot policy-ləri — hazırladılar və onları qeyri-adi dərəcədə sərt protokolla sıfırdan single-task policy-lərlə müqayisə etdilər: kor, randomizə edilmiş, böyük nümunəli (≈1,800 real-world və 47,000+ simulyasiya sınağı) və real statistika ilə. Tapşırıq üzrə finetuning-dən sonra böyük modellər toplu nəticədə etibarlı şəkildə daha yaxşı idi, təxminən **3–5× daha az tapşırıq-spesifik məlumat** tələb etdi və **şərait dəyişəndə daha möhkəm** oldu; pretraining məlumatı artdıqca performans hamar yüksəldi. Amma **finetuning olmadan** single-task modelləri ardıcıl keçmədilər, bəzi effektlər yalnız böyük nümunə ilə görünəcək qədər kiçik idi və sadə data-normalisation seçimi arxitekturadan daha çox təsir etdi. Bu, robot-foundation-model istiqamətinə möhkəm, ölçülü dəstəkdir — **ümumi-məqsədli robot deyil**, zero-shot generalist deyil və “emergent sıçrayış” deyil — üstəlik robototexnikanın böyük hissəsinin səs-küy ölçə biləcəyinə kəskin xəbərdarlıqdır.

No-BS yoxlaması

Məqalə nəyi göstərir: Sərt, kor və statistik gücü yüksək qiymətləndirmə ilə (≈1,800 real-world + 47,000+ simulyasiya rollout-u) multitask-pretrained-then-finetuned diffusion policy-lər (LBM-lər) toplu şəkildə sıfırdan single-task policy-ləri keçir, ~3–5× daha az tapşırıq-spesifik məlumatla ekvivalent performansa çatır və distribution shift altında daha möhkəmdir; performans pretraining məlumatı ilə hamar şəkildə miqyaslanır.

Mümkündür, amma sübut olunmayıb: Bu faydaların daha böyük vision-language-action modellərinə keçməsi (onların dil enkoderi kiçik idi); hamar miqyaslanmanın sınaqdan keçirilən məlumat aralığından kənarda davam etməsi.

Nəyi göstərmir: Ümumi-məqsədli və ya zero-shot robotu (finetuning yoxdur → ardıcıl üstünlük yoxdur); hər hansı "emergent" qabiliyyət sıçrayışını; konkret task-level uğursuzluqların izahını; mütləq real-dünya etibarlılığını (uğur dərəcələri həssaslıq üçün 50%-ə yaxın tənzimlənmişdi); təhlükəsizlik və avtonom deployment barədə heç nəyi.

Əsas məhdudiyyətlər: Statistika qiymətləndirmə təsadüfiliyini tutur, təlim-run təsadüfiliyini yox; hər real-world tapşırıq üçün 50 sınaq kiçik effektləri qaçıra bilər; bir arxitektura və bir laboratoriya; mülayim dil enkoderi; qiymətləndirmədən sonra tapılan data-normalisation bug-u; analiz preprint-ə əsaslanır (dərc olunan versiya yoxlanmayıb).

Ümumi oxucu nə qədər əmin ola bilər? Multitask pretraining + finetuning-in real, mülayim fayda verdiyinə — xüsusən məlumat səmərəliliyi və möhkəmlikdə — və bunların qeyri-adi diqqətlə ölçüldüyünə yüksək. Bunun **ümumi-məqsədli və ya zero-shot robot olmadığına** və **emergent sıçrayış olmadığına** yüksək. Qazancların daha böyük modellərə nə qədər keçəcəyinə orta. Müəlliflərin öz xəbərdarlığını ciddi qəbul etməyə dəyər: effektlər o qədər kiçikdir ki, zəif güclü tədqiqatlar səs-küy bildirə bilər. Düzgün mövqe: yanaşmaya ölçülü optimizm və bu cür statistik dayağı olmayan robot-AI nəticələrinə sağlam skeptisizm.

Mənbələr

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/