The Clean Paper

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

Fungujú veľké „základné modely“ robotov skutočne lepšie? Opatrná odpoveď znie: do určitej miery áno

Toyota Research Institute natrénoval „veľké modely správania“ – riadiace politiky robotov predtrénované na zhruba 1 700 hodinách rozmanitých manipulačných dát – a v mimoriadne prísnych, zaslepených a randomizovaných skúškach ich porovnal s politikami pre jedinú úlohu trénovanými od nuly. Po doladení si veľké modely viedli v priemere lepšie, potrebovali asi trikrát až päťkrát menej dát pre konkrétnu úlohu a lepšie odolávali zmenám podmienok. Bez doladenia však jednoúlohové modely sústavne neprekonávali, niektoré účinky boli veľmi malé a normalizácia dát mala väčší vplyv ako architektúra. Výsledok podporuje tento smer, ale nepreukazuje univerzálneho robota, schopnosť riešiť nové úlohy bez ukážok ani náhly „emergentný skok“.

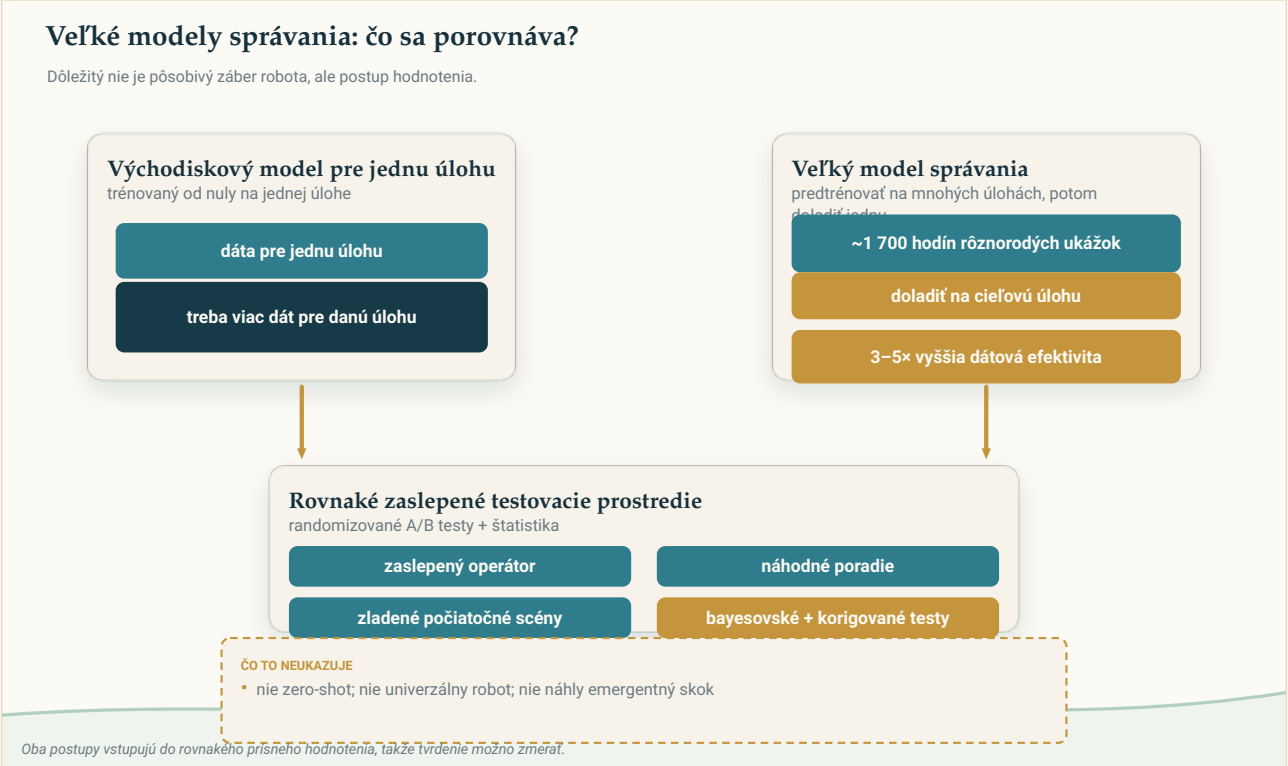
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

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

Článok o robotoch, ktorého skutočnou témou je poctivosť

Robotika zažíva horúčku „základného modelu“. Vďaka jazykovej a obrazovej umelej inteligencii je táto myšlienka lákavá: namiesto učenia robota každej úlohe zvlášť natrénovať jeden **veľký model správania (LBM)** na obrovskej, rozmanitej sade ukážok a získajte systém so širokými možnosťami, ktorý sa rýchlo prispôsobí novým úlohám. Nadšenie – a investície – sú obrovské. Titulok sám o sebe píše: *univerzálne robotické mozgy sú tu*.

Práca Toyota Research Institute je zaujímavá práve tým, že odmieta napísať takýto titulok. Jej hlavným prínosom nie je pôsobivejší robot, ale náročná analýza zdanlivo jednoduchej otázky – *naozaj funguje prístup veľkého modelu lepšie a ako by sme to vedeli?* – vykonaná s dôrazom na štatistiku, ktorá je podľa samotných autorov v odbore neobvyklá. Výsledkom je naozaj užitočné „áno, ale“ a varovanie, že veľká časť robotiky môže merať šum.



Oba prístupy prechádzajú rovnakým zaslepeným, randomizovaným hodnotením. Práca netvrdí, že robotické základné modely sú univerzálne systémy zero-shot, ale že starostlivo merané predtrénovanie môže zlepšiť dátovú efektivitu a robustnosť.

Čo autori urobili

LBM vytvorili v konkrétnom zmysle ako **difúzne vizuomotorické riadiace politiky**. Diffusion Transformer spracováva obrazy z kamier, krátky jazykový pokyn a polohu kĺbov robota a potom generuje

krátke série pohybových príkazov s frekvenciou 10 Hz. Modely boli **predtrénované na približne 1 700 hodinách** robotických ukážok – viac ako 500 rôznych úloh z interných aj verejných dátových sád – a následne **doladené** pre jednotlivé úlohy. Referenčným bodom bola vždy **jednoúlohová riadiaca politika trénovaná od nuly** výhradne na dátach danej úlohy.

Jadrom práce je však *hodnotenie*, ktoré autori považujú za hlavný výsledok. Aby sa neoklamali, použili:

- **Zaslepené, randomizované A/B testy** vo fyzickom svete – operátor robota nevedel, ktorú riadiacu politiku testuje, a poradie bolo náhodné.
- **Kontrolované a opakovateľné počiatočné podmienky** – pred každým pokusom operátori zrovnali scénu s prekrytým referenčným obrazom.
- **Veľké počty pokusov** – 50 fyzických pokusov na úlohu, politiku a podmienku a 200 na úlohu v simulácii. Celkom približne **1 800 zaslepených pokusov vo fyzickom svete a viac ako 47 000 v simulácii**.
- **Vhodná štatistika** – bayesovské odhady pravdepodobnosti úspechu a párové testy hypotéz s korekciou na viacnásobné porovnanie namiesto posudzovania prekrytia chybových úsečiek voľným okom. Navyše znovu skontrolovali štvrtinu pokusov hodnotených ľuďmi, aby zmerali chybu bodovania.

[video pending]

Porovnanie modelov nastavenia raňajkového stola: vľavo východiskový model pre jednu úlohu, vpravo LBM. Obe videá sa prehrávajú normálnou rýchlosťou. Toto je jedna hodnotená úloha, nie dôkaz všeobecnej autonómie.

Všetky tieto opatrenia majú jasný zmysel. Bez nich podľa autorov nemožno odlíšiť skutočné zlepšenie od šťastnej náhody.

Čo objavili

Doladené veľké modely v priemere prekonali jednoúlohové modely trénované od nuly. Po zlúčení výsledkov naprieč úlohami predtrénované a následne doladené LBM spoľahlivo prekonali riadiace politiky trénované od nuly na rovnakých dátach – v simulácii aj vo fyzickom svete – a rozdiel bol štatisticky významný. Takmer pri každej jednotlivej úlohe bol doladený LBM štatisticky rovnako dobrý alebo lepší ako model trénovaný od nuly: pri 3 z 3 fyzických a 15 zo 16 simulovaných úloh.

Najväčším a najzreteľnejším prínosom je dátová efektivita. Doladené LBM dosiahli výkon modelu trénovaného od nuly s približne **3–5krát menším množstvom dát pre danú úlohu**. V jednej fyzickej úlohe – prestieranie raňajkového stola – LBM doladený na pouhých **15 %** ukážok prekonal riadiacu politiku trénovanú od nuly na **100 %** dát.

Predtrénovanie pomáhalo najviac pri zmene podmienok. Keď sa testovacie prostredie zámerne odchýlilo od tréningových podmienok, teda pri *posune distribúcie*, výhoda doladeného LBM *vzrástla*. V jednej sade simulácií model štatisticky prekonal variant trénovaný od nuly pri 3 zo 16 úloh za bežných

podmienok, ale pri 10 zo 16 po posune distribúcie. Pretože sa skutočné nasadenie od tréningových podmienok vždy nejako líši, je táto odolnosť zrejme najdôležitejším praktickým výsledkom.

Viac predtréningových dát pomáhalo plynulo. Skóre sa s množstvom dát plynulo zvyšovalo, bez **náhleho nárastu či „emergentného" prielomu** na testovanej škále. Užitočné, predvídateľné, žiadna dráma.

Príbeh univerzálneho zero-shot modelu však neobstál. LBM použitý *zero-shot* – bez ladenia úlohy – sústavne neprekonával jednoúlohové modely. Jedna sieť mohla plniť mnoho úloh, ale sen o „jednoducho tomu dajte príkaz" sa tu nepotvrdil; autori to čiastočne pripisujú krehkosti malého jazykového kodéra.

Prínosy boli natoľko malé, že ich bolo ľahké prehliadnuť — alebo ich omylom „vyrobiť" šumom. Mnohé efekty sa ukázali až s väčšími než obvyklými vzorkami a starostlivým testovaním. Autori priamo píšu, že vzhľadom na veľkosť efektov a šum **existuje značné riziko, že mnoho robotických štúdií meria štatistický šum**. Zistili tiež, že všedné rozhodnutie o tom, ako dáta *normalizovať*, ovplyvnilo výsledky viac ako zmeny architektúry; **chybu** v normalizácii predtrénovacích dát navyše odhalili až *po* dokončení hodnotenia.

Čo to pravdepodobne znamená

Rozumná interpretácia znie: predtrénovanie na veľkom množstve rôznorodých robotických dát je skutočne cenné – znižuje množstvo dát potrebných na novú úlohu a zvyšuje odolnosť riadiacich politík voči rozdielom medzi tréningom a skutočným svetom. Výsledok podporuje smer, ktorým sa odbor uberá. Výhody sú však *mierne a podmienené*: prejavujú sa hlavne po doladení, najzreteľnejšie v zlúčených výsledkoch a pri zmene podmienok. Neznamenajú vznik univerzálneho robota pripraveného na použitie.

Menej nápadný a dôležitejší význam je metodologický. Práca v podstate nastavuje latku: ukazuje, koľko dôkazov je potrebné, aby tvrdenia o robotických riadiacich politikách boli vierohodné, a naznačuje, že značná časť publikovaného nadšenia stojí na príliš malom množstve dát. Odbor takú korekciu potrebuje viac ako ďalší model.

Čo to *nedokazuje*

- Toto **nie je univerzálny robot**. Výhody boli preukázané pri konkrétnej architektúre – difúznych riadiacich politikách samostatne doladených pre jednotlivé úlohy, v kontrolovaných podmienkach a na teleoperovaných ukážkach. Nešlo o autonómneho robota plniaceho ľubovoľné nové príkazy.
- Výsledky **nepodporujú výkon zero-shot**. Bez doladenia veľké modely sústavne neprekonávali jednoúlohové východiskové modely.
- Toto **nie je** dôkaz **„emergentného skoku"**. S rastúcim rozsahom sa výsledky zlepšovali plynule; žiadna diskontinuita nepodporuje príbeh „a zrazu to vedel".

- Čísla sú **relatívne a laboratórne**. Absolútna úspešnosť bola zámerne nastavená blízko 50 %, aby bolo porovnanie citlivé; nemeria spoľahlivosť v skutočnom svete a práca skúma jedinú architektúru z jedného laboratória.
- Výskum **neurčuje**, **prečo** konkrétne riadiace politiky uspejú alebo zlyhajú. Niekoľko úloh, v ktorých bol veľký model *horší*, opisuje, ale nevysvetľuje.
- Nehovorí **nič o bezpečnosti, autonómii ani nasadení** mimo testovacie prostredie.

Aké silné sú dôkazy?

Dôkazy pre hlavné porovnávacie tvrdenia – že doladené LBM ako skupina prekonávajú modely trénované od nuly, vyžadujú niekoľkonásobne menej dát a lepšie zvládajú posuny distribúcie – sú silné a výnimočne dobre kontrolované: zaslepené, randomizované, rozsiahle, štatisticky testované a s kontrolou kvality bodovania. Ide o vzácny príklad metodiky natoľko robustnej, že je možné jej hlavné závery prijať takmer doslova.

Obmedzenia otvorene uvádzajú sami autori. Intervaly chýb zahŕňajú náhodnosť *hodnotenia*, nie však náhodnosť *tréningu* – dve trénovania toho istého modelu môžu vytvoriť výrazne odlišné riadiace politiky, čo štatistika nezachytáva. Päťdesiat fyzických pokusov na úlohu stačí odhaliť stredne veľké efekty, ale malé môžu uniknúť. Jazyková časť používala skromný kodér, takže výsledky typu „povedz robotovi, čo má robiť“ môžu u väčších systémov vyzerať inak. Autori tiež otvorene priznávajú chybu normalizácie. Žiadny z týchto problémov nevyvracia hlavné výsledky, ale práve takéto veci podľa práce odbor často zametá pod koberec.

Poznámka k zdroju: diskusia je založená na predtlači autorov. Publikovaná časopisecká verzia nebola k dispozícii, preto tu nie sú preskúmané zmeny medzi preprintom a konečným textom.

Prečo je to dôležité

„Základné robotické modely“ sú pojem ako stvorený pre zveličenie a túto prácu možno ľahko vyložiť chybne v oboch smeroch – víťazoslávne ako *„funguje to!“* alebo odmietavo ako *„je to preceňované“*. Presnejší pohľad je užitočnejší: predtrénovanie na rôznorodých dátach prináša skutočné, merateľné, ale mierne výhody – predovšetkým menšiu potrebu dát pre konkrétnu úlohu a väčšiu odolnosť – a výsledky sa s rozsahom predvídateľne zlepšujú.

Hlbší význam spočíva v tom, že práca vnáša prísnosť do vlastného odboru. Autori ukazujú, že skutočné efekty sú také malé, že pri neopatrnom hodnotení miznú, a že všedné rozhodnutie o normalizácii dát môže znamenať viac ako múdra architektúra. Práca tak tvrdí, že veľká časť pokroku v robotickom učení potrebuje robustnejšie meranie, než mu budeme môcť veriť. Štúdia, ktorá dôsledne oddeľuje výsledok od priania, robí niečo vzácnejšie a cennejšie než len obsadiť prvé miesto v rebríčku.

Stručné zhrnutie

Výskumníci z Toyota Research Institute trénovali „veľké modely správania“ – difúzne riadiace stratégie robotov predtrénované na približne 1 700 hodinách rôznorodých manipulačných dát – a porovnávali ich so stratégiami pre jedinú úlohu, trénovanými od nuly. Použili mimoriadne prísny protokol: zaslepené, randomizované hodnotenie na veľkých vzorkách (asi 1 800 pokusov vo fyzickom svete a vyše 47 000 v simulácii) so skutočnou štatistickou analýzou. Po doladení na konkrétnu úlohu dosahovali veľké modely celkovo spoľahlivo lepšie výsledky, vyžadovali približne **3–5-krát menej dát pre danú úlohu** a boli **odolnejšie voči zmenám podmienok**; výkon pritom plynule rástol s množstvom predtrénovacích dát. Pri použití **bez doladenia** však jednoúlohové modely sústavne neprekonávali, niektoré efekty boli viditeľné len vo veľkých vzorkách a na obyčajnej normalizácii dát záležalo viac ako na architektúre. Ide o pevnú a vyváženú podporu smeru základných robotických modelov – **nie** o univerzálneho robota, univerzálny model zero-shot ani „náhly emergentný skok“ – a o dôrazné varovanie, že veľká časť robotiky možno meria len šum.

Kontrola bez príkras

Čo práca ukazuje: V prísnom zaslepenom hodnotení s dostatočnou štatistickou silou – približne 1 800 fyzických a viac ako 47 000 simulovaných pokusov – viacúlohové predtrénované a následne doladené difúzne riadiace politiky LBM ako skupina prekonali jednoúlohové politiky trénované od nuly. Dosiahli porovnateľné výsledky s približne 3–5-krát menším množstvom dát pre danú úlohu, lepšie zvládali posun distribúcie a ich skóre plynulo rástlo s množstvom predtrénovacích dát.

Čo je pravdepodobné, ale nepreukázané: Prenos výhod na oveľa väčšie vizuomotorické modely – použitý jazykový kodér bol malý – a pokračovanie plynulého škálovania za hranicu skúmaného množstva dát.

Čo práca neukazuje: Univerzálneho robota ani spoľahlivý výkon zero-shot – bez doladenia nebola výhoda konzistentná; žiadny emergentný skok v schopnostiach; vysvetlenie zlyhania konkrétnych úloh; absolútna spoľahlivosť v skutočnom svete, pretože úspešnosť bola kvôli citlivému porovnaniu nastavená blízko 50 %; ani nič o bezpečnosti či autonómnom nasadení.

Hlavné obmedzenia: Štatistiky zahŕňajú náhodnosť hodnotenia, nie tréningu; 50 fyzických pokusov na úlohu môže prehliadnuť malý efekt; jedna architektúra a jedno laboratórium; skromný jazykový kodér; chyba normalizácie dát zistená až po hodnotení; a analýza založená na preprinte, bez kontroly zmien v publikovanej verzii.

Koľko dôvery by mal mať čitateľ? Vysokú v záver, že viacúlohové predtrénovanie spojené s doladením prináša skutočné, mierne výhody – najmä vyššiu dátovú efektivitu a odolnosť – a že boli zmerané mimoriadne starostlivo. Rovnako vysokú v tom, že **nejde** o univerzálneho robota, spoľahlivý systém zero-shot ani náhly emergentný zlom. Strednú v možné prínosy ďalšieho škálovania na väčšie modely. Vážne treba brať aj varovanie autorov, že efekty v tomto odbore sú také malé, že štúdie s nedostatočnou štatistickou silou môžu vykazovať len šum. Správny postoj: vyvážený optimizmus voči prístupu a zdravá skepsa k výsledkom robotickej AI bez porovnateľne pevnej štatistiky.

Zdroje

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/