

動画生成AI:「見た目」の裏に潜む思考力の欠如

MMGRベンチマークが暴く、写実的な映像に隠された論理・空間推論の深刻な欠陥



MMGR: 初の包括的ベンチマークにより、動画生成AIは見た目はリアルでも、論理的・物理的ルールを理解が欠如していることが判明。

空間認識の限界: 目的地を見失うナビゲーション



0%

長距離移動
タスク成功率

「認知地図」の欠如により、目的地を見失う。

論理パズルで露呈する「見せかけの推論」

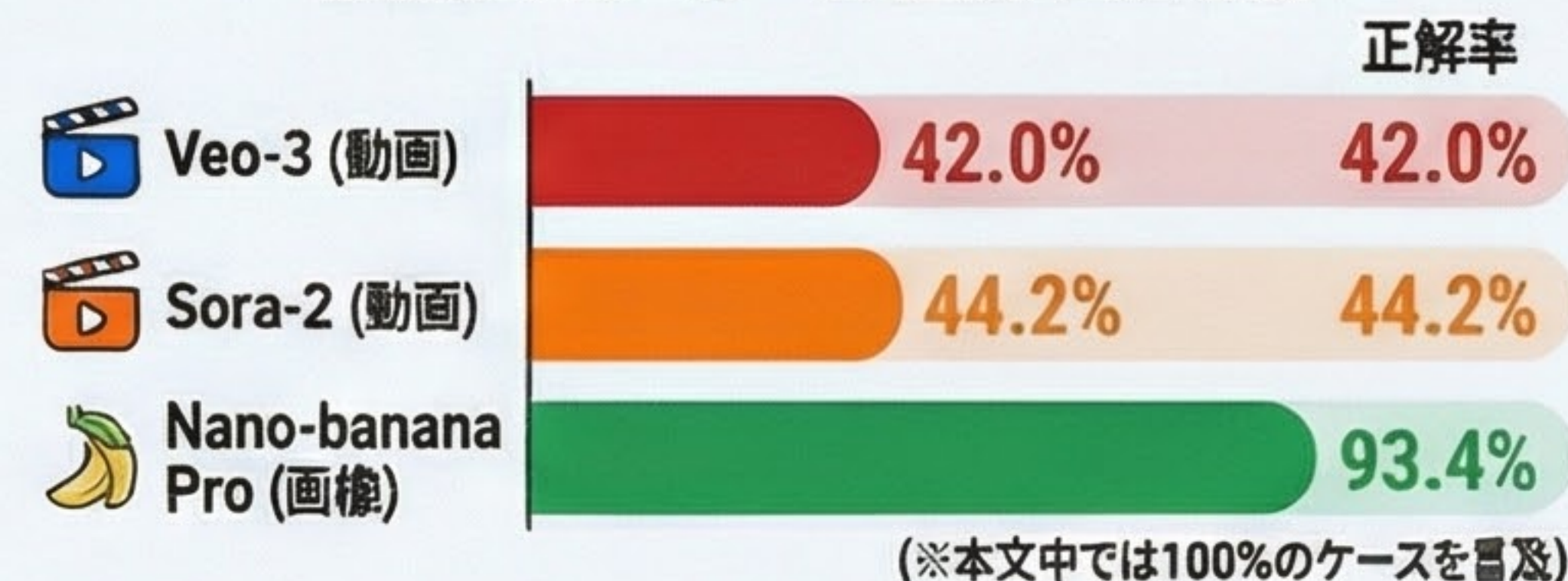


迷路で壁をすり抜け



数独ではルールを無視

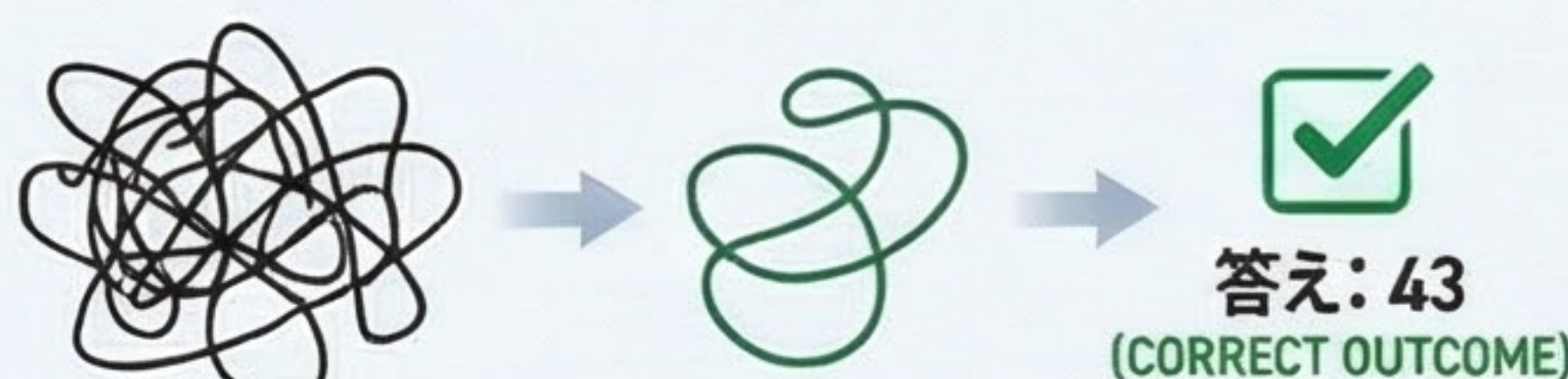
数独タスク: ルール違反率の比較



動画モデルは論理的制約を守れない。

「能力の幻覚」: AI評価の新たな落とし穴

「結果」は正しいが「プロセス」が破綻



原因は「時間的ペナルティ」



動きの滑らかさを優先



論理的に正しい状態変化が機性に

数学タスク: 結果とプロセスの乖離
(MATH TASK: OUTCOME VS. PROCESS GAP)



動画モデルの「結果」と「プロセス」の正解率には巨大な乖離 (能力の幻覚) が存在する。