

# 中国製LLMの「推論の壁」：なぜARC-AGI-2で1%台に沈むのか？

数学・知識では世界最高峰の中国製LLMが、未知の抽象問題を解くARC-AGI-2ではスコア1.3%と壊滅的な結果に。  
暗記型の「結晶性知能」と適応型の「流動性知能」の間に深い溝がある。

## 圧倒的な「知識」と、絶望的な「推論」のギャップ

数学・知識では世界最高峰、ARCでは1%台の衝撃

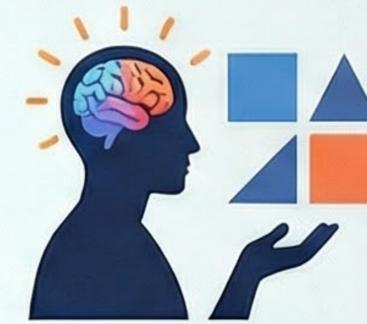


**AIME等: 90%超**

過去データのパターン補間 (暗記) は得意。

## 流動性知能：ARC-AGI-2 (抽象推論)

人間との圧倒的な差：AI 1.3% vs 人間 100%



**人間: 100.0%**  
平均2.3分で解く



**中国製AI: 1.3%**  
全く適応できません。

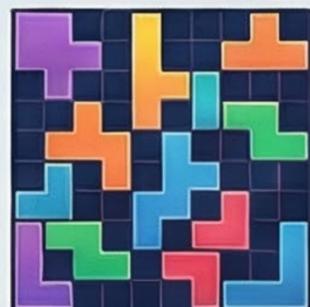
未知の規則への適応が不可能。

2026年2月時点の主要モデルと人間のベンチマークスコア比較

モデル/対象	ARC-AGI-2 (運動性)	GPQA/AIME (結晶性)	推論コスト (1タスク)
人間 (非専門家)	100.0%	N/A	低 (約2.3分)
Google Gemini 3 Deep Think	84.6%	N/A	高 (\$13.62)
DeepSeek R1 / Qwen 3.5	1.3%	71.9% - 96.0%	極低 (\$0.080以下)

## 知能の進化を阻む「3つの構造的障壁」

### ① 2次元空間の喪失 (1次元化の弊害)



.....33338883333338  
33386685666663383  
676388888235683383  
585399333393953345  
336333399494433345  
389688335995858....

空間的な位置関係を理解できません。

### ② System 1 (直感) に依存した一本道の思考



試行雑語や自己検証を行う鑑表機構が不足しています。

### ③ 「低コスト・高効率」戦略とのトレードオフ



効率を優先する中国モデルの設計は、高負荷な探索機構と本質的に対立します。