

# 知識と知能の断絶：なぜ中国製LLMは「ARC-AGI-2」で苦戦するのか？

中国製LLMは数学や知識問題で世界トップクラスだが、未知の推論能力を測るARC-AGI-2では西側モデルに大きく引き離されている。  
この格差は、現在のLLMパラダイムにおける「知識の想起」と「真の推論」の間の深い断絶を浮き彫りにしている。

## 結晶性知能：知識の想起

露呈した圧倒的なパフォーマンス格差

DeepSeek R1  
97%超  
MATH-500



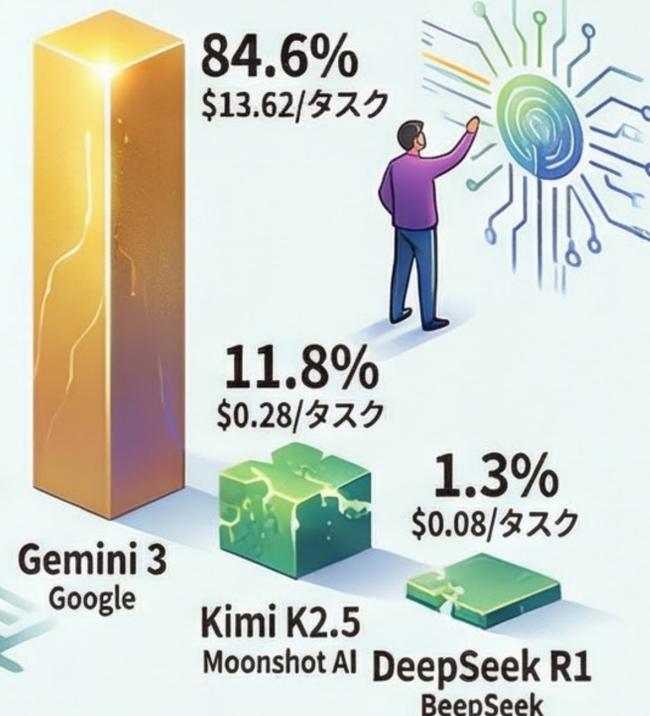
1.3%

ARC-AGI-2

DeepSeek R1に見る  
「96ポイント」の衝撃的な落差  
MATH-500で97%超を記録する一方、  
ARC-AGI-2ではわずか1.3%に留まる。

## 流動性知能：真の推論

西側トップ層と中国勢の間に  
広がる70ポイント以上の溝  
Gemini 3が84.6%で人間平均に追る一方、  
中国製最高位のKimiは11.8%に沈む。



## なぜ中国製LLMは「真の推論」に失敗するのか



3つの構造的失敗：  
記号・構成・文脈

記号の文脈理解、複数ルールの組み合わせ、特長の選択緑差の把握が困難。



強化学習(RL)  
アプローチの靚界

数学のような明確な報酬信号がない網羅では、現在のRL期躰は消化しにくい。



研究文化とベンチマーク  
最適化の弊害

複雑可能なドメインへの集中と、新たな推躰パラダイム (Refinement等) への取り組みの遅れ。

「暗記」が適用しない流動性知能の壁  
訓練データにない未知のタスクに対し、最小服の例示から規則を見出す力が欠知している。