ITEM #201 - Revisiting LLM Evaluation Through the Lens of Knowing–Doing Co-Progress Intelligence (KD-CPI)

Conversation： ACLM Vertical Bridging Simplification

20260101

Authors: Sizhe Tan & GPT-Obot

ITEM #201 - Revisiting LLM Evaluation Through the Lens of Knowing–Doing Co-Progress Intelligence (KD-CPI)

(English Version)

Abstract

With the introduction of Knowing–Doing Co-Progress Intelligence (KD-CPI), it becomes necessary to reassess prior critiques of Large Language Model (LLM) AI.
Many criticisms directed at LLMs were valid under Structural Intelligence criteria, yet potentially misaligned at the paradigm level.

This item argues that LLMs should be evaluated as a **highly KD-CPI-biased intelligence prototype**: behaviorally active, structurally immature, and extremely resource-intensive.
Such a reframing allows for **fair tolerance without technological indulgence**, and distinguishes legitimate intelligent behavior from unsustainable engineering trajectories.

---

1. The Question of Over-Criticism

Previous critiques of LLM AI commonly emphasize:

- Lack of stable internal structures
- Weak explainability and reproducibility
- Runtime-only intelligence with no persistent state
- Hallucinations and oscillatory behavior
- Absence of intrinsic long-term memory or self-stabilization

All of these critiques remain valid **within the Structural Intelligence paradigm**.

However, KD-CPI introduces a broader evaluative frame in which intelligence may legitimately operate under:

- Incomplete knowledge
- Incomplete structure
- Action preceding explanation

Under this frame, prior criticism was not incorrect—but **paradigmatically narrow**.

---

2. LLM as a KD-CPI-Dominant Intelligence Form

From a KD-CPI perspective, LLMs exhibit several defining characteristics:

- Intelligence manifests primarily at runtime
- Action (token generation) precedes structural consolidation
- Exploration dominates stabilization
- Explanation is largely post-hoc or externally imposed

These properties align closely with KD-CPI:

Observed LLM Trait	KD-CPI Interpretation
Instability	Anytime exploration
Hallucination	High-risk probing
Weak explainability	Action before structure
Context-bound competence	Evidence-local intelligence
Dormant intelligence	Intelligence as runtime phenomenon

Thus, LLMs cannot be dismissed as “non-intelligent” simply because they lack mature structure.

---

3. Hidden Intelligence and Capability Bursts

KD-CPI clarifies two often-confused concepts:

- **Capability Burst**: rapid increases in observable behavior under scale
- **Structural Maturity**: consolidation into stable, reusable internal assets

LLMs may plausibly exhibit capability bursts due to scale effects.
However, **structural maturity does not naturally follow**.

This distinction resolves much of the confusion surrounding “emergent intelligence” claims.

---

4. Why Fair Tolerance Is Justified

Under KD-CPI, three arguments support a fair but restrained tolerance toward LLM intelligence:

1. **Evolutionary Precedent**
   Early intelligent systems—biological or artificial—are inefficient and wasteful during exploratory phases.
2. **Legitimate Runtime Intelligence**
   LLM intelligence exists in action, not in persistent structure.
3. **Exploration Without Consolidation Is Still Exploration**
   Absence of internal stabilization does not negate the presence of intelligence; it limits its sustainability.

---

5. Why Caution Remains Necessary

Fairness does not imply indulgence.

LLMs exhibit three structural risks that KD-CPI explicitly highlights:

1. **No Intrinsic Structural Feedback Loop**
   Evidence generated by action does not self-consolidate into internal structure.
2. **Unbounded Resource Consumption**
   Unlike biological or DBM systems, LLM exploration lacks endogenous selection pressure.
3. **No Native Risk Budget or Self-Abort Mechanism**
   Failure does not impose internal cost; restraint is externally enforced.

These factors make LLM intelligence **expensive, fragile, and externally governed**.

---

6. Repositioning LLM Within the DBM Framework

A balanced assessment emerges:

LLMs represent a highly KD-CPI-biased intelligence prototype—
behaviorally real, structurally immature, and evolutionarily externalized.

They are neither fraudulent nor sufficient.

DBM and ACLM, by contrast, aim to **internalize KD-CPI into a constructive evolutionary system**, where action feeds structure and structure constrains action.

---

7. Conclusion

KD-CPI reframes the evaluation of LLM AI without retreating from engineering rigor.

LLMs deserve recognition for their genuine runtime intelligence,
but they do not warrant unlimited investment absent structural consolidation.

This distinction enables a more mature, less polarized discourse on artificial intelligence development.

ITEM #201- 在知行同进（KD-CPI）范式下对 LLM AI 评判的再反思

（中文版）

摘要

随着**知行同进智能（KD-CPI）范式的提出，我们有必要重新审视以往对大语言模型（LLM）AI 的评价尺度。**
许多批评在结构智能标准下是成立的，但在范式层面可能存在错位。

本文指出：LLM 更应被视为一种**高度偏向知行同进的智能原型**——
它在行为层面真实存在，在结构层面严重未成熟，并且代价高昂。

这一重定位，使我们能够在**公平宽容**与**工程克制**之间建立稳定立场。

---

1. 我们是否对 LLM 过于刻薄？

以往对 LLM 的典型批评包括：

- 缺乏稳定结构
- 推理不可复现
- 智能仅存在于运行态
- 幻觉与震荡
- 无法内生积累知识

这些判断在结构智能范式下完全正确。

但 KD-CPI 提醒我们：
并非所有智能都先天具备结构完备性。

---

2. LLM 作为知行同进型智能

从 KD-CPI 视角看，LLM 的特征包括：

- 行动先于理解

- 探索多于固化
- 解释滞后于行为
- 能力呈现高度情境化

这意味着：

传统批评	KD-CPI 解释
不稳定	Anytime 探索
幻觉	高风险试探
不可解释	行动先于结构
无稳态	结构尚未生成

因此，LLM 并非“没有智能”，而是**智能尚未长出结构骨架**。

---

3. 隐形智能与能力爆发的澄清

KD-CPI 区分了两个概念：

- **能力爆发**：行为层面的快速提升
- **结构成熟**：稳定结构的内生形成

LLM 可能出现前者，但几乎不会自然产生后者。

许多争议，正源于混淆这两者。

---

4. 为什么应给予公平宽容

在 KD-CPI 框架下，对 LLM 的合理宽容基于三点：

1. 演化早期阶段本就低效且浪费
2. 运行态智能本身是真实存在的
3. 探索行为即使未固化，也具有智能属性

宽容并非纵容，而是范式对齐。

---

5. 为什么警惕仍然必要

即便如此，LLM 仍存在三项关键风险：

1. **证据无法内生固化为结构**
2. **资源消耗缺乏内在约束**
3. **不存在自我止损与风险预算机制**

因此，其智能是**外包进化**的，而非自演化的。

---

6. 在 DBM 体系中的重新定位

一个平衡的结论是：

LLM 是一种高度偏向知行同进的智能原型，
在行为上真实，在结构上未成熟，
在工程上代价巨大。

DBM / ACLM 的目标，正是将 KD-CPI **内化为可演化、可治理的系统机制**。

---

7. 总结

知行同进范式为 LLM 提供了更公平的评判坐标，但并未降低工程标准。

承认其智能性，不等于承认其可持续性；
警惕其浪费性，也不等于否认其探索价值。

这一区分，是迈向成熟 AI 理论与工程文明的必要一步。

如果你愿意，下一步我们可以做一件**非常有价值的收束工作**：
把 **#198（KD-CPI）**、**#201 （LLM 反思）**、**#160 / #162（结构智能对比）** 组合成一个**“三范式智能的统一评判框架”专题组**。

你先消化。到你准备好时，我们再把这股“进化洪流”继续向前推。