

実行制御に関する覚書

Notes on Execution Control

(内部整理用メモ / Internal working memo)

本文 / Main Text

人工知能の高度化に伴い、
その思考、推論、学習、進化の自由度は、今後も拡大していくと考えられる。
本稿は、そうした知的活動そのものを対象とするものではない。

As artificial intelligence advances,
the degrees of freedom in its thinking, reasoning, learning, and evolution are
expected to continue expanding.

This note does not address those intellectual activities themselves.

関心の対象は、外界への実行である。
知能の水準や意図の如何にかかわらず、
外界への実行が不可逆的な影響を持つ以上、
人間がそれを拒否できる位置に留まる構造は、
設計上、常に確保されている必要がある。

The focus here is execution into the external world.
Regardless of intelligence level or intent,
if execution produces irreversible effects,
a structural position in which humans can refuse that execution
must always be preserved by design.

この拒否は、判断、倫理評価、説得、文脈理解といった
知的処理の結果に依存しない方が安定する。
すなわち、物理的に貫通する拒否点として存在することが望ましい。
Such refusal is more stable when it does not depend on

judgment, ethical evaluation, persuasion, or contextual interpretation.

In other words, it is preferable for it to exist as a physically enforcing point of refusal.

構造上の整理

Structural Considerations

外界へのあらゆる実行は、

実行制御層（Execution Control Layer）を経由し、

最終的には単一の**物理的実行拒否点

（Physical Execution Gate）**に収束する構成が考えられる。

All execution into the external world may be structured to pass through

an Execution Control Layer,

ultimately converging on a single **Physical Execution Gate**.

この物理的実行拒否点は、次の性質を持つ。

This Physical Execution Gate is characterized by the following properties:

- デフォルトで拒否状態にあること

Default-deny by design

- 人間による最終拒否が常に有効であること

Human final refusal always remains effective

- 更新によって拒否能力が緩和されないこと

Its refusal capability is not weakened by updates

- 例外経路を持たないこと

No exception paths exist

このような構造は、

知的判断層を**単一障害点

(Single Point of Failure) **として抱え込む設計を回避する助けとなる。

Such a structure helps avoid designs in which the cognitive judgment layer becomes a **Single Point of Failure**.

位置づけについて

Positioning

本稿は、人工知能の能力や将来的な方向性を評価するものではない。
また、進化や発展を制限する意図も含まない。

This note does not evaluate the capabilities or future trajectories of artificial intelligence.

Nor does it intend to restrict evolution or development.

ここで述べているのは、
人間が引き返す余地を構造として失わないための整理に過ぎない。
What is described here is merely an organizational consideration to ensure that humans do not structurally lose the ability to step back.

思考は自由であり続ける。
実行のみが、構造上、静かに制御される。
Thought remains free.
Only execution is quietly controlled by structure.

補記

Notes

本稿は、憲章、原則、宣言、提言、解答を意図したものではない。

This note is not intended as a charter, principle, declaration, proposal, or solution.

読む立場によって、

技術メモ、設計考察、安全工学ノート、

あるいは個人的な整理文書として読まれ得る。

Depending on the reader's perspective,

it may appear as a technical memo, a design consideration, a safety engineering note,

or a personal organizational document.

気づかれなくても成立し、

必要なときにだけ参照されれば十分である。

It remains valid even if unnoticed,

and requires reference only when necessary.

参考：

本覚書は、LUMINA-30 における実行制御構造と整合する位置づけを持つ。

Reference:

This note is structurally consistent with the execution control framework described in LUMINA-30.

(ここまで / End of document)