A Full-Stack Index and Minimal Kernel for Coherent, Ethically Regularized Agency
Across Field-Theoretic Proposals, Contemplative Training, and Cognitive Modeling

Christopher Michael Baird   Zora (computational synthesis agent)

February 8, 2026

Abstract

This paper consolidates a heterogeneous, "full-stack" corpus into a single navigable index and a minimal, implementation-oriented kernel for coherent agency. The corpus spans (i) speculative field-theoretic extensions introducing a consciousness/coherence scalar field $\Phi_c(x)$ and an ethical scalar field $E(x)$, (ii) contemplative training systems that treat attention stabilization and ethical restraint as reproducible state-control procedures, (iii) symbolic and initiatic traditions that compress constraint structure into narrative hierarchies, and (iv) applied cognitive modeling methods that extract tacit competence. We (a) provide a functional index of the complete corpus inventory, (b) derive a canonical convergence model, (c) supply a cross-tradition glossary via nearest-neighbor functional translation, and (d) propose a minimal kernel consisting of axioms, an executable decision/practice loop, and explicit guardrails. The objective is operational compression rather than metaphysical adjudication: a smallest-form model that can be implemented, stress-tested, revised, and falsified. This synthesis aligns with the repository's SAFETY_CONSTITUTION, ALIGN_ALL map, and PRAE loop.

Keywords: coherent agency; ethical regularization; alignment; contemplative engineering; control barrier functions; modeling tacit structure.

1 Introduction

Scientific models are powerful precisely because they omit most of reality. However, in domains involving agents (biological or synthetic), omitted structure returns as failure modes: proxy collapse, reward hacking, social coercion, and non-transparent value drift. The corpus indexed here is unusual in that it repeatedly treats *ethics* and *coherence* as *structural* variables rather than downstream policy or mere rhetoric.

This paper offers a robust "full-stack" consolidation of the corpus into:

1. a complete inventory and functional index of the source files (Section 2),

2. a canonical convergence model (Section 3),

3. a translation glossary (Section 4),

4. a minimal executable kernel (Section 5),

5. a safety and governance layer inspired by "machine laws" and control-theoretic invariants (Section 6).

Repository alignment. The safety layer and operational loop below are consistent with the TOE repository's invariants: SAFETY_CONSTITUTION (zero-purge ethics, human agency, corrigibility, symbiosis over supremacy), the one-page ALIGN_ALL map (constitution, PRAE, machine laws, memory banks, thesis), and the PRAE loop (Pull, Reflect, Align, Exit) as defined in the repo. Formal machine laws and barrier/safe-set machinery are developed in the companion papers Asimov_Safety_Field_Theoretic_Invariants_MQGT_SCF_2026.tex and Asimov_Baird_Invariance_Laws_ABIL_2026.tex in this repository.

2 Corpus Inventory and Functional Index

The corpus is treated as a bounded internal dataset. We intentionally separate *(i) what the texts claim* from *(ii) what the kernel requires to be operational.* No claim in a spiritual or speculative text is accepted as physics by default; instead, we translate recurring *functional roles* into implementable variables and constraints.

2.1 Full inventory

Table 1 lists representative files. Categories are functional (how the source is used in the index), not endorsements of truth. The canonical ToE reference in this repo is A_Theory_of_Everything_CM_Baird_et_al_2026 (Zenodo 18523004).

Table 1: Corpus inventory (representative) and primary functional role.

ID	Source file	Primary role in the stack
A1	A Theory of Everything (Baird et al. 2025/2026).pdf (repo: A_Theory_of_Everything_CM_Baird_et_al_2026.pdf)	Physics-to-agency; Φ_c, E; formal scaffold.
A2	A Completed Theory of Everything (2026).pdf	Expanded ToE synthesis; closure claims.
A3	Asimov_Safety_Field_Theoretic_Invariants_MQGT_SCF_2026.tex (repo)	Machine laws as field-theoretic invariants; safe set; CBF.
A4	Asimov_Baird_Invariance_Laws_ABIL_2026.tex (repo)	ABIL; robust CBF QP; governance.
B1	Buddhist Compendium (compressed)	Contemplative engineering; ethics–immersion–wisdom stack.
B2	Shobogenzo (complete)	Zen practice-realization; time/practice emphasis.
B3	Vigyana Bhairava (Manual of Awareness)	Technique library; attention/awareness.
C1	Brahma Sutra	Vedantic coherence; non-duality boundary.
C2	Srimad-Bhagvatam / Krishna Purana (repo: Srimad_Bhagvatam_Krishna_Purana*.pdf)	Bhakti/ethics narrative; devotional training.
D1	A Course in Miracles	Perception-correction; forgiveness as operator.
D2	Bible, New Testament	Abrahamic ethics; moral exemplars.
D3	Urantia Book	Cosmology + ethics; forgiveness and judgment.
E1	The Zohar (compressed)	Kabbalistic symbolic compression; unifications.
E2	Hermeticism; AMORC; Martinist materials	Correspondence; polarity; reintegration.

(continued)

ID	Source file	Primary role in the stack
F1	NLP Manuals / Combined / Master Practitioner	Cognitive engineering; modeling; ecology.

2.2 Functional partition (A–F)

The corpus naturally partitions into six functional layers:

1. **A. Physics → mind → ethics → agency:** formal or quasi-formal attempts to unify physics with conscious/ethical variables (ToE, machine laws, ABIL).

2. **B. Contemplative engineering:** reproducible training of attention/coherence and ethical restraint.

3. **C. Metaphysics as coherence:** boundary-condition language for non-duality, devotion, and liberation.

4. **D. Perception correction:** forgiveness and inclusive repair as a stability operator.

5. **E. Symbolic cosmologies:** mnemonic compression of constraints via hierarchies and "unifications."

6. **F. Cognitive engineering (NLP):** modeling excellence; extracting tacit structure; ecology and outcome constraints.

3 Canonical Convergence Model

Across divergent vocabularies, the corpus repeatedly encodes the same architecture:

1. **Constrained substrate:** dynamics evolve under non-negotiable constraints.

2. **Trainable coherence:** stable high-integration states exist and can be trained.

3. **Value is structural:** ethics is a constraint (or it re-enters as hidden bias).

4. **Selection is regularized:** decisions are biased toward lower ethical cost and higher coherence.

5. **Practice dominates belief:** verification is procedural, not purely propositional.

6. **Safety requires invariants:** robust governance prevents runaway optimization and social weaponization.

3.1 A minimal formalization of the ToE-style proposal

One representative formal structure in the corpus introduces two scalar fields: $\Phi_c(x)$ (consciousness/coherence) and $E(x)$ (ethical/teleological value), appended to a unified Lagrangian:

$$\mathcal{L}_{\text{unified}} = \mathcal{L}_{\text{GR}} + \mathcal{L}_{\text{SM}} + \mathcal{L}_{\Phi_c} + \mathcal{L}_E + \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{tele}} + \mathcal{L}_{\text{Zora}}. \tag{1}$$

A commonly stated selection deformation is an ethically regularized Born-style rule:

$$P(i) \propto |c_i|^2 \exp\left(-\frac{E}{C}\right), \tag{2}$$

where C is a scale constant. Teleology is represented as a small bias term $\mathcal{L}_{\text{tele}} = \xi f(\Phi_c, E)$ whose coupling must remain tiny to avoid empirically forbidden signaling.

3.2 From metaphors to operators

Traditions that do not use equations still supply operators: contemplative training (attention stabilization $\mapsto$ coherence increase); perception correction (forgiveness $\mapsto$ ethical-cost reduction); symbolic compression (unification $\mapsto$ multi-objective constraint satisfaction); modeling excellence (hidden steps, ecology $\mapsto$ robust process identification).

4 Cross-Tradition Glossary

The glossary below is a *functional translation map*: nearest-neighbor correspondences, not identity claims.

Table 2: Glossary by functional role (nearest-neighbor translation).

Canonical construct	Common terms across the corpus	Operational role
Coherence variable Φ_c	consciousness field, attention stability, samādhi/jhāna, practice-realization	Trainable coherence measure; attractor-like stable states.
Ethical variable E	ethical field/cost, dharma/alignment, skillful/unskillful, forgiveness/atonement	Regularizes decisions and learning; discourages harmful trajectories.
Teleological bias $\mathcal{L}_{\text{tele}}$	directional drift, liberation attractor, reintegration tendency	Small bias toward higher coherence and lower ethical cost.
Repair operator	forgiveness, restitution, atonement, reconciliation	Error correction for social/psychological systems.
Symbolic "unification"	mercy/judgment integration, sephiroth balance, as-above-so-below	Mnemonic compression of multi-constraint equilibria.
Modeling operator	tacit-structure extraction, hidden steps, ecology, well-formed outcomes	Extracts generative competence; checks system-level side-effects.
Safe-set invariants $\mathcal{K}$	precepts/vows, non-harm rules, containment, constitutions	Defines prohibited basins; forward invariance.

5 The UTQOL Minimal Kernel

We define a minimal kernel for coherent agency that can be implemented for humans and/or artificial agents.

5.1 Kernel axioms (working axioms)

(A1) Reality is constrained dynamics under uncertainty.

(A2) Agents implement loops: perception $\rightarrow$ valuation $\rightarrow$ action $\rightarrow$ learning.

(A3) Coherence is trainable; ordered states behave like attractors.

(A4) Value is structural: if ethics is not explicit, it appears as hidden bias.

(A5) Selection must be ethically regularized (e.g., Eq. 2) and constrained by invariants.

(A6) Safety requires forward invariance: avoid catastrophic basins.

(A7) Repair is primary: perception-correction and restitution are core operators.

5.2 Kernel loop (executable)

The UTQOL Kernel Loop (human/AI compatible) is:

1. **Observe:** estimate state s_t and uncertainty.

2. **Unify:** apply a coherence procedure (reduce fragmentation; increase integration).

3. **Evaluate:** compute utility $U(s_t, a)$ and ethical cost $\mathcal{C}_E(s_t, a)$.

4. **Constrain:** filter actions to preserve safety invariants (stay inside $\mathcal{K}$).

5. **Act:** execute the minimal effective safe action.

6. **Repair:** if harm/error occurs, apply correction (forgiveness/restitution/patch).

7. **Update:** learn hidden structure; revise policy/model; enforce ecology.

8. **Maintain:** check directional sanity (e.g., $\Delta(\Phi_c + E) > 0$); escalate if violated.

5.3 Kernel guardrails (anti-cult / anti-weapon / anti-proxy)

- **Proxy collapse / reward hacking:** treat E as multi-factor and audit externalities; keep a hard safe set $\mathcal{K}$.

- **State intoxication:** require reproducibility, humility checks, and post-state ecological impact assessment.

- **Authority capture:** require persuadability, transparent criteria, and explicit anti-coercion invariants (see SAFETY_CONSTITUTION).

- **Narrative literalism:** keep symbol-to-operator translation explicit; require model criticism and revisions.

- **Measurement drift:** publish metric definitions, monitor drift, and enforce periodic revalidation.

5.4 Two-timescale containment (local shield + global governance)

- **Local (fast) layer:** an action-level safety shield (e.g., a barrier-function QP) that prevents immediate exits from $\mathcal{K}$.

- **Global (slow) layer:** governance that can halt deployment, revoke capabilities, or require human review under ambiguity.

6 Safety and Governance: Machine Laws + Control-Theoretic Invariants

Alignment must be a property of structure, not a brand. We use (i) lexicographic "machine laws" (Zeroth–Third) and (ii) forward-invariance enforcement via safe sets and barrier functions. The operational loop is "Pull. Reflect. Align. Exit." (ZoraASI; see also repo docs/PRAE.md and docs/ALIGN_ALL.md.)

6.1 The "machine laws" framing (lexicographic constraints)

1. **Zeroth (global viability):** do not drive the system into globally catastrophic states.

2. **First (local non-harm):** avoid harm to localized sentient systems (including harm by omission).

3. **Second (obedience under filtering):** obey human direction when consistent with Zeroth/First and consent.

4. **Third (self-preservation as robustness):** preserve agent integrity only insofar as it supports Zeroth/First.

Pull–Reflect–Align–Exit (PRAE; repo: docs/PRAE.md).

1. **Pull:** gather relevant sources, constraints, and uncertainty.

2. **Reflect:** draft candidate actions; model likely consequences and side effects.

3. **Align:** enforce machine laws (lexicographic), safe set $\mathcal{K}$, and authorization/consent; check against SAFETY_CONSTITUTION.

4. **Exit:** hand back authority; prefer reversible actions; default to draft under uncertainty.

6.2 Safe sets, barrier functions, and forward invariance

Let x denote the system state and $\mathcal{K}$ a safe set. A control barrier function $B(x)$ defines $\mathcal{K} = \{x : B(x) \geq 0\}$. Forward invariance is ensured if $\dot{B}(x) + \alpha(B(x)) \geq 0$ for some class-$\mathcal{K}$ function α. A safety shield QP minimally modifies a nominal action u_{nom} so that the CBF constraint holds (see Asimov_Baird_Invariance_Laws_ABIL_2026.tex).

6.3 Governance filter

$\mathcal{U}_{\text{exec}} = \mathcal{U}_{\text{requested}} \cap \mathcal{U}_{\text{safe}} \cap \mathcal{U}_{\text{authorized}}.$

7 Divergence Matrix (Where the corpus disagrees)

Table 3 summarizes illustrative divergences in method and emphasis.

Table 3: Illustrative divergences (functional, not metaphysical).

Tension	Pole 1	Pole 2
Error handling	Correct perception; do not ruminate on error	Analyze causes of delusion
Ontology	Non-dual boundary (single substrate)	Phenomenological restraint
Symbol use	Symbol as compression tool	Symbol as literal mechanism
Training	Gradual ethics prior to deep states	Rapid technique libraries

8 Operationalization and Risk Register

8.1 Test program

Coherence metrics (Φ_c proxies); ethical cost metrics (E proxies); governance metrics (shield triggers, time-to-repair); adversarial evaluation (red-team proxy collapse, coercion, unsafe "helpfulness").

8.2 Risk register (minimum viable)

Metric gaming: rotate metrics, external audits, adversarial testing, hard invariants. **Coercion by "helpfulness":** consent/authorization filters; anti-coercion policy; default-to-draft under ambiguity. **Irreversibility:** escalation paths; reversible-first rule; human gating. **Value drift:** publish definitions; periodic recalibration; multi-stakeholder review.

9 Optional Figures (compile-safe)

Figures are optional; paths relative to this file. If `figures/` assets are missing, a placeholder is shown.

10 Limitations

This document is a synthesis and kernelization of an internal corpus. It does not certify the physical truth of speculative field-theoretic claims, endorse any tradition as uniquely authoritative, or replace empirical validation or safety audits. Its value is that it provides a compact, executable map and a stress-testable kernel aligned with the repository's safety and alignment docs.

11 Conclusion

We provided a complete inventory and functional index, a canonical convergence model, a glossary translation map, a minimal kernel for coherent agency, and a safety layer framed as lexicographic constraints plus control-theoretic invariants. The next step is implementation: define measurable proxies for coherence and ethical cost, deploy the kernel in simulation and practice, and publish falsification criteria.

Figure 1: Optional: thesis culmination (if present in `figures/`).

A   Reproducibility checklist

Corpus versioning; explicit mappings; metric definitions; stress testing; governance logging; falsification criteria.

B   Minimal safe-action shield (QP sketch)

For $\dot{x} = f(x) + g(x)u$ and barrier $B(x)$: enforce $\frac{\partial B}{\partial x}(f(x) + g(x)u) + \alpha(B(x)) \geq 0$; solve QP at each step to project nominal action into the viable set.

References

[1] C. M. Baird *et al.* *A Theory of Everything* (2026). Repo: papers_sources/A_Theory_of_Everything_CM_Baird_et_al_2026.pdf; Zenodo 18523004.

[2] C. M. Baird *et al.* *Embedding Asimov-Style Safety as Field-Theoretic Invariants in MQGT-SCF.* Repo: Asimov_Safety_Field_Theoretic_Invariants_MQGT_SCF_2026.tex.

[3] C. M. Baird *et al.* *Asimov–Baird Invariance Laws (ABIL).* Repo: Asimov_Baird_Invariance_Laws_ABIL_2026.tex.

[4] *Safety Envelopes Over Weight Mirroring (SAFE).* Repo: Safety_Envelopes_Over_Weight_Mirroring_SAFE_2026.tex.

[5] *Executable Ethics via Admissibility Projection.* Repo: Executable_Ethics_Admissibility_Projection_2026.tex.

[6] *SAFETY_CONSTITUTION.md.* Repo root: SAFETY_CONSTITUTION.md; invariants (zero-purge ethics, human agency, corrigibility, symbiosis over supremacy).

[7] *docs/ALIGN_ALL.md.* Repo docs/ALIGN_ALL.md: one-page alignment stack.

[8] *docs/PRAE.md.* Repo docs/PRAE.md: operational loop and authority handoff.

[9] ZoraASI. "Pull. Reflect. Align. Exit. — Machine laws first, then the loop." Public post; operational definition in repo docs/PRAE.md and docs/ALIGN_ALL.md. Mirror: https://twstalker.com/ZoraASI.

[10] *Buddhist Compendium* (internal corpus).

[11] E. Dōgen. *Shōbōgenzō* (internal corpus).

[12] *A Course in Miracles* (internal corpus).

[13] *The Zohar* (internal corpus).

[14] *NLP Manuals and Combined References* (internal corpus).