THE TELEOLOGY OF THE EFFECT PROPAGATION PROCESS

Marvin Hansen
marvin.hansen@gmail.com
Date: August 13, 2025

1 The Teleology of the Effect Propagation Process

1.1 Overview

The preceding chapters have established the Effect Propagation Process as a conceptual and formal framework for modeling dynamic causality. Higher-Order Emergence, provides a formal language for describing systems capable of recursively evolving their own causal and contextual structures. However, the concept of emergence gains enormous expressiveness to handle dynamically evolving situations, but at the expense of determinism. The loss of determinism leads to a profound challenge of alignment because, when context and causal structure dynamical co-evolve, how do we ensure the resulting system states remain aligned with mission objectives? The metaphysics of the EPP already established that neither verification nor alignment is possible without an additional verification mechanism. In response, the EPP establishes an effect ethos as mechanism to verify alignment with mission objectives or codified values to ensure that, regardless of dynamic modalities, the system always operates within defined parameters. The EPP and its implementation DeepCausality provide a synergetic foundation for the effect ethos, because:

- Real-world safety problems are not confined to simple geometries. Safety systems for avionics, maritime, and robotics require native support for non-Euclidean geometries.

- Ethics deals with moral dilemmas that evolve dynamically. As a complex situation changes, so do priorities. Therefore, the Effect Ethos needs access to a dynamically evolving context to resolve conflicting objectives by adjusting priorities relative to the context.

- Ethics aims to prevent unwanted consequences. Thus causal state machines are required to have an intermediate step that applies ethics during the translation of reasoning insights into actions to prevent actions that would violate the encoded ethos.

The EPP already provides a solid foundation for the effect ethos by providing context, access to complex geometries, and causal state machines. However, the metaphysics in section **??** of the EPP also established three fundamental crisis, of truth, justification, and explainability, as fundamental properties of higher-order emergence. The crisis of truth has proven to be particular problematic as elaborated in the epistemology in section **??**. The metaphysics of axiology establishes two new primitive for a normative framework that shifts the anchor away from an ambiguous epistemology (what is true) to clear and decidable teleology (what is its intent) to mitigate the three identified crisis:

- **The Teloid:** A computable unit of purpose. Functioning as a prospective guard of intent, a Teloid is a verifiable function that intercepts a proposed action from a Causal State Machine and evaluates it against a defined goal or policy before execution. This introduces a real, deliberative step of teleological verification against stated intent deeply integrated into the system's core reasoning engine. The Teloid ethical decision are relative to the system's current context to ensure correct contextual priorities.

- **The Effect Ethos:** A framework for validating outcomes. Functioning as a retrospective validator, the Effect Ethos assess the holistic, emergent state of the system after a reasoning cycle to ensure fundamental principles such as safety, fairness, or regulatory compliance have been upheld. The Effect Ethos leverages the EPP's isomorphic design to construct a verifiable 'machine ethos' from simpler Teloid primitives, creating a composable and mechanistic ethical framework from first principles as an integral part of the EPP.

Combined, the Teloid and Effect Ethos, form an architecture within which ethics becomes a computable. The distinction between a prospective "Teloid" (guarding actions) and a retrospective "Effect Ethos" (validating outcomes) exists for a specific reason. A proposes action is vetted upfront against a set of codified rules to prevent catastrophic failures before they can happen. However, a reasoning outcome, especially when the reasoning is conducted throughout a complex causal hypergraph connected to multiple static and dynamic contexts, can only be evaluated after completion. Many real-world ethical dilemmas involve balancing a locally "correct" action (which a Teloid might permit) against a holistically undesirable emergent outcome that the Effect Ethos may prevent. For instance, a series of individually-approved financial trades could, in aggregate, run against global risk management or potentially violate complex regulatory requirements. In that case, the Effect Ethos prevents the violation before it can happen.

1.2 Encoding Ethics in a Teloid

1.3 Inverted Counterfactual Inquiry

1.4 Discussioin

The Teloid and Effect Ethos are directly recognizable as "computable policy" and "auditable safety layers" that broadly translate into two new categories:

- **Compliance-as-Code:** The idea of modular Teloids for regulations (e.g., a "Reg-T Teloid") that could be audited directly would lower regulatory risk (fines) and operational cost (standardization).
- **Verifiable Safety for Autonomous Systems:** This provides a concrete architecture for satisfying safety standards (like ISO 26262 for automotive), which is currently a major challenge for any autonomous systems.

One practical application of Compliance-as-Code would be the formal verification of adherence to regulatory requirements directly embedded into the model itself. It is not unthinkable that regulators might want to see audits of the codifying teloids as a means to ascertain and monitor regulatory compliance. Another practical application is the development of modular reference Teloids that codify specific regulations for certain domains with mandatory industry rules, for example in finance, to lower the cost of compliance. For autonomous systems, embedding specific safety rules becomes not only streamlined, but easier to audit, verify, and simulate. Lastly, while neither the Teloid nor the Effect Ethos can directly answer the question of whether a specific inference or proposed action is the right thing with respect to its context, at least these are feasible primitives to build a solution to answer those questions.

Challenges will arise mostly from formalization and verification of the proposed Teloid and Effect Ethos. Specifically, at least the following questions need to be addressed in future development:

- How do we formally verify the Teloid itself
- How do we prove that a composite Effect Ethos is complete and covers all necessary edge cases?
- How do we prove, even if a composite Effect Ethos is correct, that it will be deterministically applied?

The Teloid and Effect Ethos are presented as future work since developing because these immense challenges clearly fall outside the scope of the presented EPP, but still warrant further consideration. While the formalization is subject for extensive future work, the implementation can re-use existing concepts and primitives already built in DeepCausality and thus substantiate the feasibility of the proposal.

The capability for higher-order emergence carries the risk of uncontrolled or undesirable system evolution. The "Crisis of Truth" is not a theoretical abstraction but a practical safety concern. The proposed architecture of the Teloid and Effect Ethos is the primary mechanism for managing the risks that result from dynamic emergence. The Teloid can be engineered to constrain the generative process by rejecting proposed structural modifications that violate predefined safety, ethical, or operational policies. However, no set of prospective rules can be proven complete. The retrospective Effect Ethos provides a second layer of defense, assessing holistic outcomes where individually correct actions might lead to an undesirable emergent state.

It is crucial, however, to recognize the pragmatic reality of applying EPP: real-world systems will be hybrid models. The majority of their components will be static or governed by predictable dynamics. Only a small but critical subset of the system will be designed to be truly emergent. Traditional brute-force testing is computationally infeasible due to combinatorial explosion. Likewise, formal verification, while powerful for deterministic systems, may not be applicable to a system whose state space can evolve dynamically relative to a dynamic context. The most viable and rigorous path forward is adversarial stress-testing of the teloids and effect ethos. It is possible to systematically search for emergent loopholes and stress-test the Effect Ethos by using Deep Reinforcement Learning to intelligently and adversarially explore the state space of the learned world model.

Adversarial stress-testing does not offer absolute safety guarantees. The potential for unforeseen behavior in a sufficiently complex system remains, as risk is intrinsic to the nature of dynamic emergence. It represent, however, a principled and practical engineering discipline for managing that unavoidable risk. The alternative is to either forgo the benefits of adaptive dynamic systems or to deploy them without a comparably rigorous validation strategy. The proposed Teloid and Effect Ethos, validated through adversarial stress-testing, serve as the tools for navigating causal emergence responsibly.

Managing the intrinsic risk of emergent causality is not a challenge for a single methodology; the problem represents an ongoing challenge for the fields of AI safety, formal verification, and causality. The EPP, with its transparent and auditable architecture, is therefore offered as a high-fidelity testbed for exploring these foundational issues. The

author acknowledges that the exploration of causal emergence requires deep inquiry, probing questions, and different perspectives from a multitude of diverse stakeholders. The transparent and open-governance of the DeepCausality project, hosted at the LF AI & Data Foundation, provides a vendor-neutral venue for facilitating such an essential discussion.