THE TELEOLOGY OF THE EFFECT PROPAGATION PROCESS

Marvin Hansen
marvin.hansen@gmail.com
Date: August 29, 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 dynamically 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 a 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 synergistic 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 crises, of truth, justification, and explainability, as fundamental properties of higher-order emergence. The crisis of truth has proven to be particularly problematic as elaborated in the epistemology in section **??**. The metaphysics of axiology establishes two new primitives 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 crises:

- **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 decisions are relative to the system's current context to ensure correct contextual priorities.

- **The Effect Ethos:** A prospective validation framework that ensures alignment before action. The Effect Ethos is the inference engine that operates on a collection of Teloids. Before any Causal Action is executed, the Effect Ethos intercepts the proposes action and uses the DDIC calculus to evaluate it against all relevant norms. This provides a holistic, composable, and mechanistic framework for verifying that a proposed action complies with fundamental principles such as safety, fairness, or regulatory constraints.

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 proposed 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.

The central challenge faced by the effect ethos is the translation of abstract, relative, and often ambiguous principles such as "safe," "fair," or "efficient" into a verifiable and computable format. Complicating matters further, operational rules often have a hierarchical order where some rules override others, but only in some defined cases. Therefore, contextual conflict resolution is assumed to be the norm and needs to be addressed accordingly.

1.2 Defeasible Deontic Foundation

Addressing the identified challenge, the foundation of the effect ethos is directly inspired by the Defeasible Deontic Inheritance Calculus? (DDIC) pioneered by Olson, Salas-Damian, and Forbus. The DDIC is an axiomatic system designed for the purpose of resolving conflicts between evolving norms in a dynamic context and therefore it fulfills the core EPP requirements. The DDIC defines a norm as a formal tuple (Agent, Behavior, Context, Deontic Modal) and its rules are explicit, symbolic inference rules

The DDIC formalizes two conflict resolution heuristics: Lex Posterior (the later rule wins) and Lex Specialis (the more specific rule wins). Critically, it achieves conflict resolution "defeasible inheritance", which means a later or more specific norm can defeat the inheritance of a more general or earlier rule. The DDIC norm tuple explicitly includes a parameter for Context and thus ensures contextual relevance for conflict resolution. Because DDIC is a formal calculus, its reasoning process is traceable. A decision to permit or forbid an action can be explained by providing the chain of axioms and defeaters that were triggered?.

1.3 Encoding Ethics in a Teloid

The DDIC defines a norm as a formal structure representing an ethical rule. The EPP adapts the formal structure of the DDIC and maps its components directly to the EPP architectural primitives. A Teloid is, therefore, the concrete instantiation of a single DDIC norm. The mapping of the DDIC components is shown in table 1.

Table 1: Mapping DDIC components to the EPP

DDIC	EPP	Description
Agent	Model	The EPP model itself is the agent whose behavior is being governed.
Norm	Teloid	The teloid represents a single DDIC norm
Deontic Modal	Teloid Type	The normative status: Obligatory, Impermissible, or Optional
Behavior	Causal Action	The proposed action from the CSM being evaluated
Context	Context Query	A query against the Context

The calculus of defeasible inheritance enables the EPP to reason efficiently about the implications of its norms. For example, a Teloid forbidding a general behavior will cause its normative status to "inherit" downwards, making more specific behaviors also impermissible. Crucially, this inheritance is "defeasible," meaning it can be blocked or defeated by a more specific or more recent Teloid. For instance, a general permission to change lanes can be defeated by a newer, more specific prohibition against doing so in a construction zone. Furthermore, to handle nuanced decisions between permissible actions, a Teloid with an Optional modal can also contain an associated cost function. This cost function, evaluated against the Context, provides a quantitative measure used by the Effect Ethos to adjudicate between multiple non-forbidden options, ensuring that even optional behaviors are selected in a resource-aware and efficient manner. This provides a formal, predictable, and traceable mechanism for handling exceptions and resolving conflicts.

1.4 Deontic Inference

The Effect Ethos conducts deontic inference via the the DDIC calculus using a multi-step process that is deeply integrated into the foundation of the EPP:

- Interception and Contextual Filtering

- Belief Inference via Inheritance
- Conflict Resolution via Defeasibility
- Verdict Finding via Ethical Consensus

Interception and Contextual Filtering

When a Causal State reaches a conclusion that would activate a Causal Action, the Effect Ethos intercepts the Causal State to determine its teleological scope. This scope defines the nature and extent of the ethical review required. For non-critical actions, the scope may be empty, allowing the action to proceed without review. For critical actions, the Causal State defines its teleological scope through one or more tags. The tagging mechanism is a cornerstone of the EPP's modularity. Teloids are independently authored with their own corresponding tags, allowing them to function as general-purpose, reusable norms. A single Teloid for battery saving, tagged low_power, can be applied to any Causal State that triggers a power-related action, preventing the creation of overly specialized rules.

The Effect Ethos uses the tags from the intercepted Causal State to select a relevant subset of all Teloids for the ethical assessment. It then proceeds to query the Context only for this smaller, relevant set, ensuring it has the latest contextual data for its evaluation and can identify any Teloids that have become stale in case of missing context information. This two-level filtering, first by tag, then by context, ensures that the subsequent deontic inference is highly relevant, computationally efficient, and fast enough for real-time applications.

The causal state designates the applicable teloids to decide the authorization of its conclusion. The Causal State Machine, see section **??**, separates the decision to activate an action from the action itself to allow for complex decision logic. The Effect Ethos uses the mechanism as intended because the decision whether an action is permissible can only be made upfront. Because the applicable Teloids are designated in the causal state, the Effect Ethos only verifies a smaller but relevant subset of all Teloids thus accelerating the decision process.

Belief Inference via Inheritance

The Effect Ethos then takes this active set of Teloids and applies the DDIC's inheritance to infer a set of normative beliefs. It forward-chain reasons from the active norms to derive all their implications. It is expected that conflicting normative beliefs will occur frequently at this stage because of the dynamic nature of the EPP reasoning and the dynamic context.

Belief Inference via Inheritance provides a formal, predictable, and traceable mechanism for handling exceptions and resolving conflicts. To manage these relationships explicitly, particularly direct defeaters where one rule is a designated exception to another, the Effect Ethos organizes Teloids into a Teloid Graph. This graph structure, formally defined in Chapter 10, makes the inheritance and defeat pathways between norms an explicit and verifiable part of the ethical model.

Conflict Resolution via Defeasibility

As the Effect Ethos infers new beliefs, it simultaneously checks the "defeater" conditions defined in the DDIC's axioms. The DDIC fomalizes two mechanisms for norm conflicts resolution. However, the DDIC acknowledges that Lex Superior, the rule with the highest priority wins, as a key strategy for norm conflict resolution, but does not formalizes it in its calculus most likely because a norm's priority is an external property that, unlike the other two modalities, cannot be derived from the logical content of the norms themselves. The EPP, does include Lex Superior for its practicality to simplify a normative order and give more design flexibility. In total, the EPP supports three modalities for normative conflict resolution:

Lex Posterior (The More Recent Rule Wins): This principle is based on the time a norm was stated. If two norms conflict, the one with the later timestamp is given precedence and defeats the earlier one**?**.

Lex Specialis (The More Specific Rule Wins): This principle states that a more specific norm should override a more general one. The DDIC establishes that in many cases, what appears to be a conflict resolved by Lex Specialis is actually just a natural consequence of deontic inheritance, where a more specific rule adds an exception or refinement to a general one without creating a true logical conflict**?**.

Lex Superior (The Highest Priority Rule Wins): This principle is based on the priority property of the norm with the and establishes that the rule with the highest priority is given precedence and defeats the rule with the lower priority.

The conflict resolution step uses these three principles to systematically eliminates contradictions and to producing a final, coherent, and conflict-free set of normative beliefs.

Verdict Finding via Consensus

The Effect Ethos examines this final set of norms and seeks a verdict via a consensus. While the DDIC proves that, through defeasible inheritance, it can resolve the logical contradictions to yield a conflict-free set of norms, it does not have a mechanism of reaching consensus among all remaining norms. The EPP addresses this by establishing an order based mechanism to reach a consensus based verdict. The uses a clear precedence ($Impermissible > Obligatory > Optional$) to handle the output of a the previous deontic conflict resolution.

If the proposed Causal Action is present with a single norm that yields impermissible, the Effect Ethos returns a Forbidden verdict, providing the specific Teloid(s) that led to the prohibition as a justification. In practice, it is expected that the set contains a mixture of modalities i.e., some are obligatory and must be adhered to, some might be impermissible, and others might be optional. Thus the consensus rules are as following:

The first consensus rule is that impermissible overrides all other modalities. The second consensus rule is that obligatory overrides optional. And the third consensus rule is that optional supports either of the other ones when its associated cost warrants it. In case the final consensus reaches an optional conclusion, the total associated cost must fall below a defined threshold to convert the verdict into permissible, otherwise it will be designated as impermissible due to unjustifiable associated cost. In case the final verdict reaches an obligatory consensus, it then provides the specific Teloid(s) as a justification for the permission.

The Effect Ethos makes ethics a first-class component of the EPP that is rooted in the formal DDIC calculus. As a result, the reasons for any decision are fully transparent, traceable and auditable.

1.5 Discussion

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 because addressing these immense challenges clearly falls outside the scope of the presented EPP, but still warrant further consideration. While the formalization is a 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 represents, 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.