

THE VALUE OF INFORMATION AS THE FOUNDATION OF ETHICS: FROM THE VARIATIONAL PRINCIPLE TO THE CATEGORICAL IMPERATIVE

Vladimir Kovalev

Independent Philosopher, New Jersey, USA

vkovalev@live.com

Abstract

A pickpocket and his victim render opposite verdicts on the same act. Religious ethics has a referee above both perspectives; secular ethics never found a replacement: Aristotle's "flourishing" is always somebody's flourishing, and utilitarian sums can be rigged to justify the theft. Missing, since Hume stated the gap nearly three centuries ago: a quantity whose growth is good and whose loss is bad regardless of anyone's viewpoint — a non-religious absolute.

The paper argues the quantity exists — that every persisting thing has been extremizing it all along. Every system — planet, cell, person, institution — has a graded capacity to constrain its own future from within. Building on my earlier definitions, I define this capacity as a non-negative, non-additive, intrinsic scalar: a system's value of information. Because that phrase already carries other meanings, the paper proposes a distinct name: *vedance*, from the Indo-European root *weid-*, "to see, to know."

Physics supplies the hypothesis: simple systems extremize classical action, dissipative ones obey entropy principles; the value of information, I propose, is the general quantity of which both are limiting cases. If that is what systems are, Hume's gap needs no leap — or so I argue: a system that can recognize its principle and choose among trajectories turns description into prescription. Good, within a system, is what increases its value of information; right, across systems, what maximizes it overall — one law, two scopes. The quantity does not sum, so this is not utilitarianism. Does a well-organized criminal group count as good? Within its boundary, yes — precisely why good is not the last word: corroding the trust and predictability its neighbors depend on, it fails the stronger test of right.

The framework's categorical imperative, meant literally: when we act badly, we lose the capacity to predict the future.

Keywords

value of information, vedance, systems theory, ethics, variational principle

1 | Introduction

A persistent difficulty has attended every attempt to ground ethics outside of religion. Religious ethics, whatever else may be said of it, is in possession of an absolute. Whether the absolute is revealed law, divine command, or the will of a transcendent person, its function in the moral economy is the same: it guarantees that judgments of good and evil are not relativized to the perspectives of those who issue them. The pickpocket at the fair who steals a wallet, and the man from whom the wallet is taken, do not stand on equal moral ground simply because, from each of their points of view, the act has opposite valence. The religious framework places both perspectives under a higher authority that adjudicates between them without consulting either.

Non-religious ethics has had no comparable resource. The Aristotelian tradition identifies the highest good with *eudaimonia* — happiness, or human flourishing (Aristotle, *Nicomachean Ethics*, I.7) — but happiness is a state of particular subjects, and in the case of the pickpocket and his victim, one subject's happiness is purchased at the other's expense. The utilitarian tradition, more sophisticated in its calculations, identifies the good with the maximization of utility across affected parties; but utility, on the classical formulation, is a sum of individual valuations, and one can construct circumstances under which the theft is, by the utilitarian's own measure, the better act. The fault is not in the arithmetic. The fault is in

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

the absence of any quantity to be summed that would carry the absoluteness required of an ethical foundation.

What has been wanting, in other words, is a *non-religious absolute* — a quantity whose preservation or increase constitutes the good without depending on the perspective of any particular subject, and whose loss constitutes the bad in the same impersonal way. Without such a quantity, ethics is left choosing between two unsatisfactory options. It may accept that good and evil are perspective-relative and abandon the project of binding moral judgment; or it may smuggle in absoluteness through a back door, dressing perspectival preferences in universalist language. Neither option preserves the seriousness with which we actually treat moral claims.

The proposal of this paper is that such an absolute exists and can be identified. It is the *value of information* — a non-negative scalar quantity that characterizes any system whose internal structure constrains the trajectory of its future states. This proposal is not arbitrary. It rests on a theory of systems and a theory of information that I have developed in previous work (Kovalev, 2024; Kovalev, 2025) and that I summarize in the next section. The essential claim is that any system — physical, biological, cognitive, social — has, in addition to its structural and dynamical properties, a quantity that measures the degree to which it is capable of constraining, and thereby shaping, its own future. This quantity admits of no negative values; it is internal to the system; and it does not depend on the system being observed, named, or evaluated from outside.

The further claim, and the one that grounds the ethical framework, is that this quantity stands in a deep relation to the variational structure that physics has long recognized in the dynamics of material systems. The principle of stationary action selects, from among the conceivable trajectories of a mechanical system, the one that is actually realized. The same variational structure, generalized appropriately, governs the evolution of complex systems whose dynamics include internal models, memory, and adaptive response. In such systems, the quantity being extremized — driven to a maximum, a minimum, or a stationary value, as the regime demands — is no longer classical action alone but something of which classical action is a limiting case: the value of information itself. To the extent that this generalization holds — and I present it here as a hypothesis of the framework rather than as an established theorem — the value of information is not merely a candidate quantity for the foundation of ethics. It is the quantity that every such system — every system whose existence is bound up with constraining its own future — is already extremizing in its very persistence as a system. It is also a claim of convergence: a single principle general enough to span physics and ethics without reducing the one to the other.

This, if it holds, is what permits the move from physics to ethics without violation. For a system capable of acting — of selecting among possible trajectories rather than merely undergoing them — the variational principle ceases to be merely descriptive of what happens and becomes prescriptive of what should happen, for the system itself. The system's continued existence as the kind of thing it is depends on its participating in the trajectory that the principle selects. To act in ways that diminish the value of information in one's own system is, quite literally, to dissolve oneself as a system; to act in ways that diminish it across the systems with which one is entangled is to participate in the dissolution of those systems. The "ought" of ethics is not imposed from above on a value-neutral world; it is the recognition, by a system capable of recognition, of what the system already is.

From this foundation, the structure of the ethical law follows. Within a given system, *good* is whatever increases the value of information in that system. Across all systems, *right* is whatever maximizes the aggregate value of information. The two predicates are not in tension; they describe the same variational principle at two different scopes. What is good within a system may, in the wider context, be not right — as when an act increases the value of information for one party at the cost of greater loss elsewhere.

The paper proceeds as follows. Section 2 recapitulates the necessary background on systems and the value of information, drawing on previous work and situating the framework against the variational principles of physics. Section 3 develops the ethical framework proper, articulating the good/right distinction, the categorical imperative, and the response to the is–ought gap. Section 4 examines four cases, including the pickpocket scenario with which I began and a case in which the framework delivers a demanding and counterintuitive verdict. Section 5 considers the relation of the framework to existing

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

positions — perspectival relativism, utilitarianism, Spinoza's conatus, Floridi's information ethics, and the variational tradition in cognitive science. Section 6 addresses the most serious objections, including the objection from measurement and the objection from internally coherent destructive systems. Section 7 concludes with the framework's convergence claim and a statement of the open questions left for future work.

I close this introduction with a sentence that may serve as the categorical imperative of the framework, and as the line that orients all that follows: **when we act badly, we lose the capacity to predict the future**. This is not a metaphor. It is the most concise statement of what this paper attempts to demonstrate; Section 3.3 will argue it in full.

2 | The Framework: Systems and the Value of Information

This section establishes the apparatus the ethical argument requires. It states in full the definitions previously presented at the Society's two preceding meetings (Kovalev, 2024; Kovalev, 2025), with one addition: an explicit account of how the framework's central quantity relates to the variational principles of physics.

2.1 | The System

I take a *system* to be a set of elements — portions of space with internal structure changing over time — in which the future states of at least one element are altered by other elements. Three features of this definition matter for what follows. First, the states whose alteration defines systemhood are physically grounded: in the foundational case they are kinematic regimes, patterns of acceleration, which are observer-independent in the senses that matter across classical and relativistic frames; higher-order systems — living, cognitive, social — require further state distinctions, a development deferred to future work. Second, the “altered by” relation is asymmetric. The Sun alters the states of a planet; the planet does not appreciably alter the states of the Sun; and the pair is a system nonetheless. One-way dependence is sufficient. Third, the definition carries a principled relevance threshold: an influence counts as constitutive of a system when it operates above the level of noise at the timescale of interest. Without this threshold every element in the universe would belong to a single system; with it, systemhood becomes scale-relative without becoming arbitrary. The definition was first presented at the Society's 68th meeting (Kovalev, 2024).

2.2 | Signal, Information, and the Value of Information

Ordinary usage runs together three things that the framework must keep apart: the *signal*, the *information*, and the *value of information*.

A signal is a physical perturbation propagating through a medium — photons, pressure waves, voltages. It exists independently of any receiver, and it is the proper subject of the mathematical theory of communication (Shannon, 1948), which measures the capacity of a channel to carry discriminable signals. This is indispensable engineering, but it is not a theory of information in the sense ethics needs.

Information, within a system, is a modification of the system's internal structure produced by the interaction of a signal with structure already present in the system, such that the modification alters the system's constraint on its future trajectory. Information is thus an event, not a thing, and it is constitutively relational: the same signal produces different information, or none, in different receivers, according to the structure each brings to it.

The *value of information* of a system is a non-negative, non-additive, intrinsic scalar quantity that measures the degree to which the system's internal structure constrains the trajectory of its future states. Each of the four formal properties earns its place in the ethical framework. That the quantity is *non-negative* is what allows it to serve as an absolute: there is no anti-constraint, no negative pole. That it is *non-additive* is what separates the framework from utilitarianism at the level of structure: the value of a composite is not the sum of the values of its parts. That it is *intrinsic* is what frees it from observer-relativity: it characterizes the system itself, not anyone's evaluation of it. That it is a *scalar* is what makes the ethical verbs — “increases,” “maximizes” — meaningful (Kovalev, 2025).

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

2.3 | An Illustration, and Its Generality

Consider a person who hears a weather forecast. The forecast is a signal. Its interaction with the person's prior structure — their understanding of language, their model of themselves as someone who can carry an umbrella — modifies that structure: the person now holds the belief that it will rain. That modification is the information. As a consequence, the trajectory of the person's future states narrows: the branch "leave without umbrella, arrive soaked" closes, and the branch "leave with umbrella, arrive dry" opens. The person's value of information has increased. The increase is realized at the moment of action that exploits the new constraint — the taking of the umbrella. Had the person been unable to act on the forecast, the same belief would have carried no value.

Nothing in this account is specific to persons. The same three-tier structure applies — with state distinctions appropriate to each domain — to a planet settling into an orbit, a snowflake condensing into a lattice, a cell maintaining itself against its surroundings, a language community, a scientific theory, a constitutional democracy, and a human life. The framework is not anthropocentric and does not presuppose cognition; cognition is one elaborate special case of a structure that constrains its own future.

2.4 | The Variational Reading

The claim that gives the framework its physical seriousness is that the value of information is what systems extremize. The physics must be stated with care: it contains not a single minimization principle but at least two independent ones.

The first is the family of the principle of stationary action — Hamilton's principle, the kindred fixed-energy principles of Maupertuis and Jacobi, and their optical analog, Fermat's principle of least time. Among the conceivable trajectories of a conservative, reversible system, the realized trajectory is the one along which the action — a single number each candidate trajectory accumulates over its course — is stationary (Lanczos, 1949). This family governs deterministic dynamics and selects trajectories.

The second is the family of entropy principles — the principle of maximum entropy (Jaynes, 1957), the minimization of free energy, the minimization of entropy production near equilibrium (Prigogine, 1947; textbook treatment in Prigogine, 1967), and the minimum-dissipation principle (Onsager, 1931). This family governs statistical and dissipative dynamics. On the present understanding, neither family is derivable from the other: no accepted derivation obtains the second law of thermodynamics from the action principle without statistical assumptions imported from outside mechanics, and the entropy principles do not, by themselves, return Newton's second law. Whether a deeper principle connects the two remains an open question; as matters stand, they coexist, applying in different regimes.

The hypothesis of this framework is that the value of information is the more general quantity of which both are limiting cases: classical action in the limit of simple, conservative systems, and entropy-family extremization in the limit of dissipative ones. A complex adaptive system — one whose internal constraint structure itself changes as it processes information — extremizes the value of information across this whole range. The nearest existing attempt at such a unification is the free-energy principle (Friston, 2010), to which I return in Section 5. I present this as a hypothesis, not a completed derivation; exhibiting the limiting cases formally is the central item of unfinished business, to which I return in the conclusion.

2.5 | What the Framework Is Not

Five brief disambiguations forestall predictable misreadings. The value of information is not Shannon information: the latter measures signals and channels, the former is a property of a system's internal structure. It is not utility: it is intrinsic to systems and is not aggregated from individual valuations. It is not the expected value of information of decision theory (Howard, 1966), which prices a signal for a particular decision-maker relative to a particular choice problem: the framework's quantity is intrinsic to the system, indexed to no decision. It is not negative entropy in Schrödinger's (1944) sense, later formalized as Brillouin's (1956) negentropy: it is more general, and is not defined relative to a thermodynamic equilibrium reference state. And it is not the "information" of quantum theory and the black-hole information paradox, which is a measure of state distinguishability preserved by unitary dynamics — a concept at the level of bare physical law, not at the level of systems with predictive structure.

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

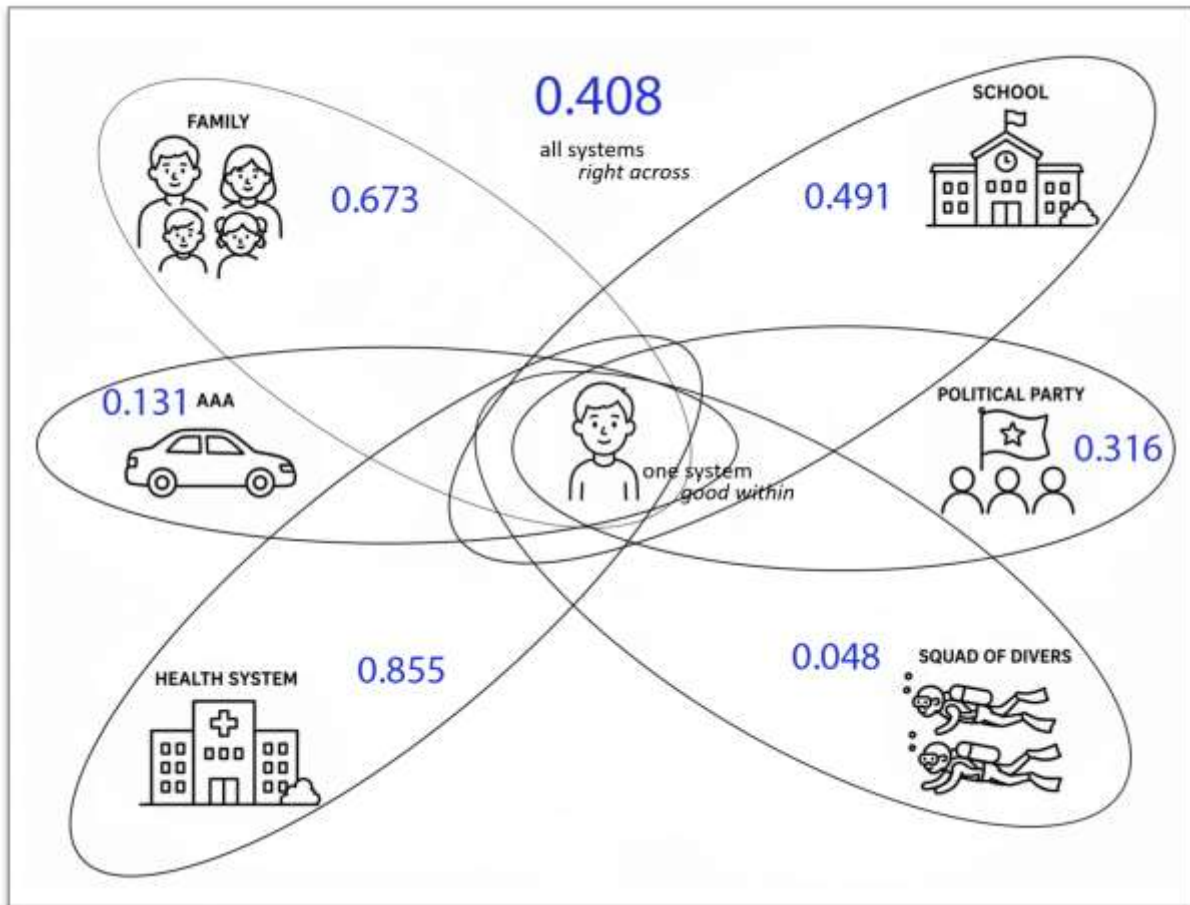
3 | The Ethical Framework

3.1 | Good and Right

The fundamental law has two scopes. *Good*, within a system, is what increases the value of information in that system. *Right*, across systems, is what maximizes the aggregate value of information across all the systems an act touches. These are not two values in competition but one variational principle applied at two scopes (Exhibit 1). An act may be good and not right — good within the system that performs it, yet not right in the wider context, because its local gain is purchased at the price of a greater loss elsewhere. An act that is right across systems will typically also be good for the system that performs it, because the structures such an act sustains in other systems are commonly structures on which the actor also depends. The limiting exception — an act locally bad yet possibly right across systems, of which self-sacrifice is the extreme case — is taken up in Section 3.6.

A caution about the verbs is in order. The ethical law is not inherited from stationarity: a stationary point is not a summit. What the law requires, and what the cases of Section 4 in fact use, is weaker — that changes in the value of information carry sign, and that the degradation of constraint structure is a signed loss. Whether complex adaptive systems genuinely maximize the quantity, rather than merely holding it stationary, is a hypothesis about the complex regime specifically, and it awaits the limiting-case derivation to which I return in the conclusion.

Exhibit 1. Good and Right as Two Scopes of One Variational Principle.



The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

3.2 | Consequentialist in Form, Not Utilitarian in Substance

The framework evaluates actions by their effects on the value of information, and is in that broad sense consequentialist. But it departs from utilitarianism (Mill, 1863) in three structural respects, not merely in emphasis. First, the maximized quantity is not utility, welfare, or preference satisfaction; it is the quantity that systems already extremize in their persistence as systems. Second, the quantity is non-additive: the right cannot be computed by summing individual utilities, because the value of information of a composite is not the sum of the values of its parts. Third, the quantity is intrinsic to systems, which dissolves the problem of interpersonal comparison of utility that has troubled utilitarianism since its inception. The framework thus keeps the structural form of consequentialism — evaluation by effects — while replacing its substance entirely.

3.3 | The Categorical Imperative

The law can be stated in a single sentence: when we act badly, we lose the capacity to predict the future. Read precisely, this is the law stated in the negative. To act in ways that diminish the value of information in one's own system is to weaken the very structure by which the system constrains, and therefore anticipates, its own future. For a system whose constraint structure includes a model of itself and its surroundings, constraining its trajectory and predicting it coincide: the same structure that narrows what the system will do is the structure by which it foresees what it will do. The imperative's "predict" is this reflexive special case of the general "constrain." The phenomenology of moral failure — the sense of being lost, scattered, no longer able to plan or to trust one's own projections — is not a poetic accompaniment to wrongdoing but the felt experience of a declining value of information. The imperative states a consequence, not a prohibition: the loss follows even when the decline is chosen for others' sake, as in the sacrifice of Section 3.6; what distinguishes wrongdoing is that its decline purchases no gain elsewhere. The imperative resembles Kant's (1785) in form, in that it holds without exception and follows from the structure of rational agency rather than from any external end; it differs in its grounding, which is not the consistency of a maxim under universalization — whether the rule behind one's act could be willed as a law for everyone — but the persistence of a system under its own variational principle.

3.4 | From Is to Ought

The framework must answer the objection, as old as Hume (1739–1740), that no "ought" follows from an "is." My response is not to deny the gap but to locate the derivation differently. The framework does not derive the prescriptive from a value-neutral description of the world. It derives it from a description that already includes the system's own variational principle as part of what the system is. A system that exists at all is a system whose continued existence depends upon extremizing a particular quantity; the "ought" is the recognition, by a system capable of recognizing it, of a fact already contained in what the system is. This is not yet a complete philosophical account — in particular, it owes an explanation of why recognition should bind the agent's choice rather than merely inform it, and that account is unfinished. The move has a near neighbor in constitutivist ethics — Korsgaard's (2009) derivation of what an agent is bound to do from what it takes to be an agent at all — and the framework may be read as a variational, systems-theoretic version of that strategy, in which the constitutive feature is the extremization by which the system persists at all. But the move does not commit the fallacy Hume identified, because the "is" from which it begins is not value-neutral: the variational principle is normative for the system in the minimal sense that the system ceases to be itself when it departs from it.

3.5 | The System That Models Itself

For systems whose constraint structure includes an internal model of themselves — the kind of system the reader is — the ethical principle acquires a reflexive dimension. Such a system's value of information depends in part on the accuracy of its self-model, because a system that misrepresents its own state predicts its own trajectory poorly. Self-deception is therefore not merely an intellectual lapse but an ethical failure on the framework's own terms: it is the case in which information modifies internal structure in a way that degrades the system's model of itself, and thereby its capacity to constrain its own future. A full

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

development of this reflexive case, and of what distinguishes a system's actual state from its represented state, is left for future work.

3.6 | The Asymmetry of Harm

An observation the framework is well placed to accommodate, though it does not yet derive it, is that harm and benefit are not symmetric. The destruction of constraint structure propagates more readily than its construction: an act that dismantles the structures on which several systems depend reduces value across all of them at once, whereas an act that builds structure typically does so slowly and locally. This is one structural reason the framework aligns with received moral intuition about the gravity of destruction. It does not follow that a system may never diminish its own value of information; a system may do so in order to raise the value of information of others by a greater amount — an act locally bad and possibly right across systems. The framework does not forbid such sacrifice; it supplies the structure within which it can be assessed.

4 | Cases

4.1 | The Umbrella

The simplest case is the one already given. A person acts on a forecast, narrows the trajectory of their own future, and arrives dry. The act is good within the person's own system, and, touching no other system adversely, it is also right across systems. The case carries no dramatic moral weight, and that is precisely why it is included: an ethics of only the extraordinary cases would be an account of something narrower than ethics. The law operates in the smallest decisions.

4.2 | The Pickpocket

Return to the scene with which the paper began. The thief takes a wallet. Within the thief's own system, the act may increase the value of information: resources are gained, a capacity is exercised, a future is opened. Within the victim's system, the act reduces it: a planned future is foreclosed, and the victim is left less able to anticipate and shape what follows. Across the human community, the act is not right, because the practice it instances degrades a constraint structure — mutual trust, the security of property, the predictability of dealings among strangers — on which every member of that community depends for the value of information they severally possess. This practice-effect is not a rule smuggled in beside the act but part of the act's own causal footprint: each theft marginally degrades the trust structure every surrounding system relies on. The verdict does not require computing the aggregate: the act's contribution to each system around it is negative, and no admissible operator — no defensible way of combining values across systems, whatever its exact form — can certify such an act as right; the ground is developed in Section 6. The framework thus yields three distinct verdicts where a single-level ethics can yield only one, and it explains why the thief's private gain does not settle the moral question.

4.3 | The Lie That Helps

Consider a lie told with good intent — to spare a friend a painful truth. In the moment, the lie appears to preserve something: the friend's composure, their confidence, the ease of the relationship. Within a narrow scope and a short horizon, it may even raise the value of information locally, in that the friend proceeds with less immediate distress. But a lie inserts into the friend's model of the world a proposition that is false, and a model corrupted at one point yields unreliable predictions at others. Over the longer horizon on which the friend must act on their model of reality, the lie has lowered their capacity to anticipate that reality correctly. The act is locally good and not right. The framework reconstructs, without appeal to any rule against lying as such, the familiar intuition that benevolent deception is a corruption of the person deceived.

4.4 | Future Generations

The framework's power is clearest where its demand is heaviest. Future generations are systems — or will be — whose value of information will depend on the constraint structures they inherit: institutions, bodies

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

of knowledge, functioning ecosystems, a stable climate. Acts that degrade these structures lower the value of information of systems that do not yet exist but certainly will, and the framework has no structural license for discounting that loss: nothing in it permits weighting a system's value of information less because the system is distant in time, as the discount rates customary in economic cost-benefit analysis do. The result is a demand for preservation stronger than common moral intuition is comfortable with, and stronger than discounted practical calculation delivers. I state this consequence plainly, because it is where the framework does genuine work: it issues a demand that does not merely reconstruct what we already believe, but corrects it.

5 | Relations to Existing Positions

Against perspectival relativism, the framework locates the absolute without recourse to transcendence: immanent, present in every system as its own variational quantity.

Against utilitarianism, as set out in Section 3.2, the framework retains the form of consequentialist evaluation while rejecting additivity, the metric of utility, and the aggregation of individual valuations.

With Spinoza (1677) the affinity is genuine: the quantity at the foundation plays a role akin to the *conatus*, the striving by which each thing endeavors to persevere in its being, and which Spinoza identifies with the thing's actual essence. The difference: Spinoza's monism is of substance, the framework's of variational structure.

With Floridi's (2013) information ethics the framework shares the conviction that informational properties are morally fundamental, and diverges in what it makes fundamental: Floridi accords intrinsic moral standing to informational entities, whereas the framework locates the good in the extremization of a quantity — a shift from entities to dynamics.

With the variational tradition in cognitive science, and specifically the free-energy principle (Friston, 2010), the framework shares the ambition of a single quantity whose extremization spans simple and self-organizing systems. It adds a general systems ontology — in the lineage of Bertalanffy (1968), not tied to any theory of neural function — and an explicit ethical interpretation the free-energy literature has left implicit.

6 | Objections and Replies

The objection from measurement. It may be objected that if the value of information is a scalar quantity, it ought to be measurable, and it is not. The reply is that the framework does not claim ease of measurement. Many quantities that physics treats as real — entropy above all — are difficult to measure directly, and are known instead through their relations to other quantities and through their role in successful theory. The claim is that the value of information exists and is intrinsic to systems, not that it is readily instrumented; its measurement is a problem for empirical science, not an objection to the ethical theory. A sharper form of the objection asks for the quantity's functional form and dimensions, noting that the two claimed limiting cases do not even share units — action is measured in joule-seconds, entropy in joules per kelvin. The reply is the same in kind: the four formal properties are offered as constraints on any future formalization, the role axioms played for entropy between Clausius and Boltzmann, and a formalization spanning both regimes must exhibit the bridging constant, in the manner of $\hbar$ or k_B , that links them. Its absence is one more reason the limiting-case derivation stands high among the problems this paper leaves open. One limit of the entropy analogy should also be conceded: entropy is counted as real partly because it has operational handles — partition functions, calorimetry, ensemble methods — and the value of information does not yet have one, even for a toy system; supplying that first computation is named among the open problems in the conclusion.¹

¹ A formal terminology for the quantity may prove useful in the physics that this measurement problem belongs to. I propose *vedance*, from the Indo-European root *weid-*, “to see, to know,” which descends into Sanskrit *veda*, Old Slavic *vedi*, and English *wit* and *wisdom*. The term names the magnitude of a system's trajectory-constraining structure. I offer it tentatively; throughout this paper the working term remains “value of information.”

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

The objection from unfalsifiability. A quantity that has never been computed, and whose verdicts in Section 4 land where intuition already stood, invites the charge pressed against the free-energy principle: that any behavior whatever can be redescribed, after the fact, as extremizing it. What would count against the framework? Three things, named as failure conditions rather than reassurances. If no formulation of the value of information reduces to classical action for conservative systems and recovers the entropy principles for dissipative ones, the variational unification fails. If the toy-system computation named in the conclusion yields no well-defined extremized quantity, the hypothesis is refuted in the simplest case. And since the verdicts of Section 4 rest on the sign of changes rather than on a computed aggregate, a case in which those signs come out wrong breaks the cases directly. Until the reduction and the computation exist, the charge stands as a live risk; the framework is offered as a research program with its failure conditions stated.

The objection from stasis. If the value of information measures constraint on the future trajectory, does the framework not crown the perfect crystal — or the agent who retreats into a dark, unvarying room — as its maximizer? A frozen system's next state is all but determined; a rigid life is superbly predictable. The reply draws on the definition of information as an event. Value is realized in action that exploits newly acquired constraint (Section 2.3), and a system that has ceased to register signals exercises no constraint; it merely undergoes a fixed one. The crystal sits in the classical-action limit of Section 2.4, where the quantity degenerates toward bare action; the dark-room dweller does not increase constraint but collapses the repertoire of futures over which constraint could operate. The distinction between constraining a rich repertoire and collapsing one is, I concede, not yet formalized. I offer the stasis case as an adequacy condition rather than a solved problem: any candidate formalization must score the living organism above the crystal, or it is the wrong formalization.

The objection from agency. The framework speaks of systems that act and choose. What grounds agency, and which systems possess it? Agency, on the framework's account, is the capacity of a system whose constraint structure includes internal models of its own possible futures to select among trajectories on the basis of those models. The framework treats it, provisionally, as graded rather than binary: many systems possess it in some degree, humans in high degree. A complete account of this capacity for recognition — of what exactly distinguishes a system that recognizes its own principle from one that merely instantiates it — is among the framework's open problems.

The objection from non-additivity. If the value of information is non-additive, in what sense can it be “maximized” across systems, and what does the aggregate mean? The reply is that non-additivity does not entail unaggregability; it entails only that the aggregate is not computed by summation. The framework owes a specification of the operator by which values across systems combine, and I do not pretend to supply it here. It is the most important item of unfinished formal business, and I name it as such rather than conceal it.

One consequence should be stated plainly: where an act's effects carry genuinely mixed signs — some systems gaining, others losing — the aggregate is today undecidable. Outright decidability is confined to sign-aligned cases — the umbrella, where every affected system gains. The theft is not sign-aligned, for the thief gains; its verdict rests on the weaker ground made explicit in the reply that follows: no admissible operator can certify as right an act whose every effect beyond its actor is negative. Throughout this paper “not right” means exactly that — the act cannot be certified as right — not that a negative aggregate has been computed.

The pickpocket's wallet, considered as a constraint on the operator, already argues for a direction: the same sum matters more to a household near its viability threshold than to a secure one, which tells against additive and simple-average forms and points toward threshold-weighted, saturating ones. The operator must also be defined over nested and overlapping systems without double-counting; which systems enter the aggregate belongs to the same unfinished specification. Meanwhile, sign-verdicts survive re-partitioning: an act that degrades a constraint structure lowers the value of every system that contains it.

The objection from internally coherent destructive systems. The gravest objection is that if every system extremizes its own value of information, then a criminal organization, an authoritarian state, or a virus is doing exactly what the framework appears to recommend. The reply turns on the two scopes. Within such a system, coherence and coordination may indeed raise the value of information; across systems, the same

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

organization lowers it, because it degrades the constraint structures on which surrounding systems depend. This reply does not require the full aggregation operator that the previous objection leaves open; it requires something weaker. Degradation of constraint structure is unambiguously a reduction of the value of information of the system degraded, so the organization's contribution to each of the systems around it is negative whatever the operator turns out to be. Whether those losses are outweighed by the gains inside the practice — the organization's own, its members', its beneficiaries' — is, strictly, a question only the operator can settle. But any admissible operator is, other things equal, monotone in the values it aggregates: an aggregate of value that rose while every constituent value fell would not be an aggregation of value at all. And under monotonicity, a practice whose every gain accrues inside itself and whose losses fall everywhere else cannot be certified as raising the whole — not right, in precisely the sense just defined. The reply carries exactly as far as the framework can take it today: a per-system sign result, and a burden the objector cannot meet; the remainder is deferred, explicitly, to the operator.

The objection from physical reductionism. If the value of information is at bottom a physical property of systems, is the resulting ethics not merely physical determinism in disguise? The reply distinguishes the system that is from the system that recognizes itself. A planet extremizes its variational quantity unreflectively; a moral agent does so reflectively, by recognizing the principle and acting in its light. The freedom that ethics requires is not freedom from physical law but the freedom of recognition — the capacity of a sufficiently reflexive system to make its own variational principle an object of awareness and thereby of choice.

7 | Conclusion

The framework proposed here is not one more candidate ethics, set beside the classical positions as an alternative to them. It is better understood as the recognition of something those positions have been pointing toward without quite naming. Each of them has identified, partially, a quantity whose preservation seems to matter unconditionally — virtue, flourishing, freedom, dignity, welfare. The proposal of this paper is that beneath these partial identifications lies a single quantity, the value of information, intrinsic to every system, which every system already extremizes in its persistence, and which becomes ethical when a system is capable of recognizing and acting upon its own variational principle. This is what would place ethics on the same explanatory footing as the other sciences of self-organizing systems — the convergence the systems sciences have long sought.

Much remains to be done. The operator by which the non-additive value of information aggregates across systems is not yet formalized. The derivation of classical action, and of the entropy principles, as limiting cases of the value of information is presented here as a hypothesis and awaits a proof. The quantity itself has not been computed for even the simplest system; one such computation would convert existence in principle into existence in practice. The moral psychology of systems capable of deceiving themselves deserves a treatment of its own. And a settled terminology for the central quantity — perhaps *vedance*, with its lineage in the old root for knowing — remains to be established.

But the proposal stands on its foundation. When we act badly, we lose the capacity to predict the future, because we lose the constraint structure that makes us systems capable of prediction at all. This is not a metaphor, and not a moralist's slogan. It is what ethics, on this view, finally turns out to be.

References

- Aristotle. *Nicomachean ethics*. Numerous editions; cited by book and chapter.
- Bertalanffy, L. von. (1968). *General system theory: Foundations, development, applications*. George Braziller.
- Brillouin, L. (1956). *Science and information theory*. Academic Press.
- Floridi, L. (2013). *The ethics of information*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199641321.001.0001>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>

The Value of Information as the Foundation of Ethics: From the Variational Principle to the Categorical Imperative

- Howard, R. A. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, SSC-2(1), 22–26. <https://doi.org/10.1109/TSSC.1966.300074>
- Hume, D. (1739–1740). *A treatise of human nature*. John Noon (Vols. I–II); Thomas Longman (Vol. III).
- Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4), 620–630. <https://doi.org/10.1103/PhysRev.106.620>
- Kant, I. (1785). *Groundwork of the metaphysics of morals*. J. F. Hartknoch.
- Korsgaard, C. M. (2009). *Self-constitution: Agency, identity, and integrity*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199552795.001.0001>
- Kovalev, V. (2024, June 9–13). Yet one more definition of system [Abstract]. *Abstracts of the 68th Annual Meeting of the International Society for the Systems Sciences*, Washington, DC.
- Kovalev, V. (2025, July 11–15). Information as structure: On the value of information in systems [Abstract]. *Abstracts of the 69th Annual Meeting of the International Society for the Systems Sciences*, Birmingham, UK.
- Lanczos, C. (1949). *The variational principles of mechanics*. University of Toronto Press.
- Mill, J. S. (1863). *Utilitarianism*. Parker, Son, and Bourn.
- Onsager, L. (1931). Reciprocal relations in irreversible processes. I. *Physical Review*, 37(4), 405–426. <https://doi.org/10.1103/PhysRev.37.405>
- Prigogine, I. (1947). *Étude thermodynamique des phénomènes irréversibles* [Thermodynamic study of irreversible phenomena]. Desoer.
- Prigogine, I. (1967). *Introduction to thermodynamics of irreversible processes* (3rd ed.). Interscience.
- Schrödinger, E. (1944). *What is life? The physical aspect of the living cell*. Cambridge University Press.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3–4), 379–423, 623–656. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Spinoza, B. (1677). *Ethics*. Numerous translations.