The Third Protocol

Daehyung Lee

May 23, 2026

Abstract

This manuscript begins from the diagnosis that the central bottleneck of modern society is no longer execution capacity itself, but has shifted to the question of what input format society uses to read and aggregate human beings. In the age of survival and scarcity, capitalism was a powerful coordination mechanism that reduced human suffering, but in the saturated zone after survival, it has reached a state in which more production, faster connection, and stronger stimulation no longer guarantee a better life. Technical means accumulate exponentially, but higher coordinates of life such as purpose, justice, love, and responsibility do not accumulate in the same way. At the same time, disciplinary specialization weakens the ability to see the whole context, and democracy and the market compress human will and preference into the low-resolution signals of votes and prices. AI and automation do not first solve this imbalance; they expose it. As the costs of translation, search, design, negotiation, production, and coordination fall, the set of executable choices explodes, while the social input format that determines what should be executed remains impoverished.

To address this problem, this manuscript first asks what kind of subject declares a purpose. We imagine the self as a solid substance preserved from the past to the present, but the present has no temporal volume, and most of what we call "I" is closer to a narrative of traces that have already passed. Moreover, if free will is premised on an open future, different past self-states can merge indistinguishably within the current state of evidence, and many inner states leave neither complete traces, nor recoverability, nor any distinguishable difference in the present world. Therefore, the self should be understood not as a substance excavated from the past, but as a non-substantial structure that approves and erects itself toward the future on incomplete evidence. In this sense, purpose declaration is not a simple expression of desire or a plan, but the constitutional origin by which the self binds itself into a subject.

If the self stands up through purpose declaration, society too cannot stop reading human beings merely as consumers, voters, or users. The economy has read humans as holders of preference, politics as holders of opinion, and platforms as sources of reaction, but before all that, humans are beings who declare what they live for and what they will be responsible for. Based on this transition, this manuscript proposes a Third Protocol beyond votes and prices. If the First Protocol aggregates authority through votes, and the Second Protocol aggregates scarce value through prices, then the Third Protocol is a social input-output rule that receives purpose declarations, conditional commitments, and intentions to contribute as social inputs, and sends them back into organizations, institutions, production, contracts, and intentional events. This is not the abolition of politics and economics, but a complementary circuit that inputs, at higher resolution, the human purpose structures that the two systems have lossily compressed.

However, receiving purpose declarations as social inputs does not mean those declarations automatically form a coherent order. Purposes collide, are misunderstood, and are sometimes unclear even to themselves. Thus the Third Protocol requires two internal engines. First, Generalized Utilitarianism (GU) is a normative coordinate system that asks what should count as a more just adjustment when declared purposes collide. GU expands utility from a simple sum of pleasure and pain into a higher-dimensional function that includes empathy, knowledge, relational structure, and self-understanding, and understands justice not as a completed state but as a direction of convergence toward omniscient empathy. Second, Contractual Value Transfer Dynamics (CVTD) is an execution grammar that records value changes and alignments that occur when declared purposes descend into real intentional events. Here the word contractual is not limited to legal contracts, but refers to the whole event structure in which purpose-bearing agents intervene in the world through conditions, expectations, responsibility, commitments, and the possibility of violation. Furthermore, the Tri-Basis Drive Model (TBDM) explains how countless microscopic intentional events are averaged, canceled, and amplified into macroscopic net forces across the three axes of possibility, sociality, and coherence.

Finally, this manuscript argues that the first real entry point of the Third Protocol should not immediately be a purpose market. Purpose is not information automatically extracted from logs, clicks, or consumption records, but a social object formed through the slow process of reading, thinking, being challenged, revising, conversing, and approving. Therefore, the core infrastructure of the future is not a more sophisticated recommendation engine, but purpose-formation infrastructure. Books and long-form texts are the slowest and firmest seeds in this process, and a long-form canonical layer that can lead to declaration. In this context, The Channel is positioned not as the completed Third Protocol, but as the first real interface that places the human process of purpose formation onto the city and turns the records accumulated in that process into the sensory nervous system of AI that can represent the individual. The entire project of this manuscript ultimately converges on one question. In an age when means overwhelm ends, in what form can humans declare themselves again, and how should society read that declaration?

1 Introduction

"So what problem does it solve?" The inner meaning of this gateway question toward startups is clear. In truth, it touches the demand of capital: "What problem will you solve in order to make money for us?" This means that the value of problem solving is valid only when it is aligned with the logic of capitalism. In the past, this alignment may have been natural. Today, it is different.

I no longer think of capitalism as a just coordinator that saves humanity like the "invisible hand." Capitalism has now become too enormous and bloated to look delicately into human suffering. In the era when humanity's default condition was close to "misfortune" such as hunger, cold, and disease, the goal of capitalism was the reduction of misfortune, that is, survival. At that time capitalism appeared to be an efficient and ethical system. But now that, through the development of technology and civilization, at least in our Western civilization, humanity's default condition has risen beyond survival into the stage of happiness, the driving force of

capitalism has changed its nature.

The human utility curve has the form of a typical logarithmic function. In the early stage, when survival was threatened, even a small increase in the resources put in made the slope of utility, or happiness, rise steeply. A piece of bread saved a life, and one roof changed a life. But we have already passed that steep interval.[1]

We now stand at the right end of the logarithmic function, where the increase in utility relative to input has almost lain down into a flat line. Now ordinary stimulation no longer increases utility. If in the past we fought for survival, now capital is set in motion by an endless thirst for greater stimulation and peripheral pleasure. We have forgotten why faster, stronger, and more should be better. We have merely turned the empirical extrapolation that faster, stronger, and more is better into a kind of vast religious illusion.

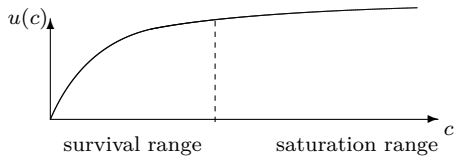
Does this enormous and powerful system called capitalism, which leads us, exist merely for the tiny increase at the end of this logarithmic function? Tens of trillions of capital and cutting-edge technology are invested, and the result is only about a 1% rise in the pleasure humanity feels. What an inefficient runaway.

I diagnose this problem as arising from the structural difference in speed between the "means of life" and the "purpose of life." In the film Dead Poets Society, Mr. Keating says, "Medicine, law, technology⋯ these are noble means necessary to sustain life. But poetry, beauty, romance, love⋯ these are what we stay alive for." Our tragedy lies in the fact that we are skillful at maximizing the means, but have lost the wisdom to deal with purpose. Technology is versioned up, code piles up on GitHub, and knowledge is copied and transmitted. A frontend developer does not need to understand how circuits work, and a computer scientist does not need to understand quantum mechanics. We feel the accumulation of technology when we look at the evolution of the iPhone. But things such as justice, ethics, and love do not accumulate that way. In other words, we have no GitHub for justice. Technical means expand exponentially, while ethical purposes circle in place. This structural "speed gap" has broken the brake that could control capitalism's runaway.

To make matters worse, and yet inevitably, technological progress has proceeded in the direction of instantly satisfying this pleasure, and the explosive increase in the total amount of knowledge has locked individuals into wedges called "specialization." Paradoxically, the eye that could look after the whole context has disappeared. In other words, the quantitative explosion of knowledge accumulated by humanity since modernity has produced the extreme

[1]Mathematical note: This intuition is connected to the CRRA(Constant Relative Risk Aversion) or isoelastic utility function.

$$u(c) = \begin{cases} \dfrac{c^{1-\gamma} - 1}{1-\gamma}, & \gamma \neq 1, \\ \log c, & \gamma = 1. \end{cases}$$



Here c is the level of consumption or resources, and γ indicates the degree of diminishing marginal utility. When $\gamma = 1$, it becomes log utility. What matters in this manuscript is not the strict adoption of a particular functional form, but the structural intuition that in the survival range, additional resources have a large ethical effect, while in the saturation range, their justification weakens.

specialization of disciplines, that is, the wedging of knowledge. In this process, cross-reference among different systems is blocked, and the producers of knowledge are reduced to functional units that perform only partial tasks without reading the whole context. This is the result of Marx's instrumentalization of the human being being reproduced in the form of the "disciplinary performer," and at the same time it is structurally identical to the figure of the "scholarly laborer" criticized by Nietzsche. For Nietzsche, the scholarly laborer endlessly mines, annotates, and processes data in a vast factory of knowledge, but never reaches the role of crossing from mountain to mountain and reconstructing value. He defined himself as "one who crosses between mountains," but the modern system of knowledge leaves such mountain-crossing to the virtue of exceptional individuals while hardening in the direction of mass-producing many people as laborers fixed inside their disciplines. The judgment structures of the humanities, engineering, politics, and economics become separated in a state where they cannot interpret one another, and in democratic society this structure becomes an increasingly fatal pathology. Voters, and even elected social decision makers, lose the whole context under the pressure of wedge-shaped disciplines.

Eventually, we will lose the "whole." It is self-evident that the government, which should check the unilateral run of capitalism and hold the macroscopic context, will become powerless before enlarged capital. Or perhaps we should even hope that it becomes powerless.

Democracy tore up the admission ticket to power and distributed it to the people under the banner of collective intelligence. The power held by an individual was originally something faint like a mirage, but the even more hopeless fact is that the citizens holding those tickets have become trapped in wedges of specialization and can no longer see the whole context.

The gaunt suffrage held by a crowd that has lost context, and a blind capitalism that has become too fast and powerful to observe what should be aggregated. In this situation, where on earth can we expect "right judgment"? Must we lean on the fantasy that this civilization will naturally lead us onto the right path, the fantasy that the universe owes us a just probability? In the end, we will have to stand again before fundamental questions. What do we live for? Who are we?

We keep moving forward while discarding what cannot be recovered. We will never again be able to dance stylishly to tacky music with tacky moves. The algorithm has given us more refined taste, but at the same time it has taken away the ignorance that allowed us to love the clumsy. What did we give up in order to lie in bed scrolling Reels for five hours? Shouldn't what waits ahead of us at least be more magnificent than what we left behind? Technology gives us the feeling that nothing is impossible, but in fact it advances by permanently abolishing certain spaces of possibility at the same time.

We cannot go back. Going back is not the answer either. The innocence before civilization, the community before capital, and the order before democracy are not answers we can choose again. The problem is not the fact that we are moving forward. The problem is that we are moving too quickly without knowing what we are discarding, and without asking what we are discarding it for.

In an era when means were weak, the poverty of purpose could be hidden to some extent.

There was not much we could do. But as means grow stronger, the poverty of purpose is no longer a philosophical deficiency but an actual danger. We can make more, connect faster, and predict more precisely, but we have increasingly impoverished language for why we should do so. Means have accumulated, but purposes have not. Execution has become precise, but judgment has not.

In this condition, means do not stop. Means continue moving forward without waiting for the question we have lost. Our inventions, made in dependence on possibility, have never guaranteed that they would lead us toward the ideal. Continuous pursuit has also never guaranteed that it would lead us to the end of pursuit. Whenever one possibility opens, another possibility closes. More refined taste takes away the ignorance that could love the clumsy, and faster connection pushes out the senses that used to form slowly in solitude. Every time we gain something, we also abolish some world to which we can never return. It seems that the inertia of life has lost its credibility.

Then the problem is not simply that technology is too fast. More precisely, it is that while technology rapidly changes the space of possibilities of the world, we have not sufficiently cultivated the ability to judge what that change should be for. The engine of civilization has become more and more powerful, but the device that determines the direction toward which that engine should head remains old and crude. The market reads prices, and politics reads votes, but those signals alone do not sufficiently reveal what kind of world humans want, what they cannot give up, and what they are willing to take responsibility for. A ship crossing the Pacific can set its direction roughly without major trouble. Even if it turns the wheel a few seconds late, or deviates from its route by a few meters, there is time to correct it. But a rocket leaving the atmosphere is different. If a very small angular error occurs at launch, that error amplifies over time into an irreversible difference in orbit. The greater the speed and the greater the energy, the more precise the device that sets the direction must be.

It is precisely at this point that AI appeared. AI is not a savior that suddenly fell from outside, nor merely a new machine that takes away human jobs. It is closer to the event in which the means we have accumulated for so long suddenly acquired a face. At the end of the push driven by the replication of knowledge, the advancement of computation, disciplinary specialization, and the decline of execution costs, civilization now faces the sum of the means it has created.

The anxiety of AI lies in the fact that too much has now become possible. As the costs of translating, finding, summarizing, designing, connecting, and producing fall, we become able to do more things more easily. But the fact that we can do more things does not answer the question of what we should do. Rather, it exposes that question more cruelly.

Until now, society has read human input broadly in two ways. Politics read votes, and capitalism read prices. Votes allocated authority, and prices allocated resources. Capitalism translates desire into price, and democracy translates will into vote. These two signals were powerful. But they read human beings at very low resolution. Human purpose is far more complex than prices or votes. A person is not simply a being who wants to buy something, nor a being who delegates authority to someone every few years. A person is a being who asks what

life they want to live, what world they can bear, and for what they can stake their time, ability, and responsibility. The problem is that the systems we have cannot sufficiently read input at precisely that layer.

We now stand before that bottleneck. It is a state in which the excess of means and the poverty of purpose input are simultaneously maximized. We can do more, but our ability to choose what matters more has not grown accordingly. We have more information, but we do not know under what purpose that information should be aligned. We have more powerful tools, but who the human holding those tools is, and for what they should lift them, is becoming increasingly unclear.

In the end, the problem returns to these old questions. What do we live for? Who are we? Now these questions are not private meditation or a philosopher's hobby. In an age when means overwhelm ends, they have become the most practical operating problem that civilization can no longer postpone.

---

2 The Non-Substantiality of the Self and Purpose Declaration

Then where should this problem be taken up again? A diagnosis that technology is fast is not enough. We must ask what development is, on what intention that development is justified, and from where the words intention and purpose obtain their authority. Further still, we must also ask whether speaking of purpose itself already presupposes the human being as a special subject. Then what is the human being? What is the subject that is thought to set purposes, choose, and take responsibility? In the end the question returns to the most familiar and yet most unclear object. What is the thing that "I" regard as "I"?

We perceive ourselves as "time travelers." We believe there is one solid substance that carries memories of the past, passes through the present, and moves toward the future. Yesterday's self, today's self, and tomorrow's self are in different states, but we think that underneath them the same "I" continues. This belief is so natural that it is hardly doubted. We experience ourselves not as a bundle of events, but as a subject passing along time.

At the same time, however, we reject mechanical determinism. We believe that human beings are not objects whose previous state necessarily produces the next state, and that at every moment we could have chosen otherwise. Without this belief, the word responsibility is hard to establish. We distinguish the act of pulling a trigger from an event in which a tree knocked down by a strong wind falls on someone, and the ground for assigning guilt and responsibility only to the former lies in the normative premise that "he could have acted otherwise." Moral judgment stands precisely on this open possibility.

Here a tension arises. If we are truly beings with free will, the future is not a closed single path. From the same present, multiple possible choices and multiple possible developments of the self must be open. But if, at the same time, we are to say "I have continued as one substance from the past to now," then some solid chain of self must be preserved even amid those countless

possibilities. Free will opens the future, while the substantial self tries to bind time into one chain. The two are believed together in ordinary life, but when examined strictly, they press against each other.

The problem is that the world does not completely record our selves. The evidence of the world we inhabit is essentially incomplete. The universe is like a kind of low-resolution camera: it leaves our macroscopic behavioral trajectories to some extent, but it does not perfectly preserve our inner minute hesitations, passing emotions, unchosen possibilities, or unspoken thoughts. The fact that I did not say something may remain, but the subtle impulse, fear, and grain of self-censorship just before saying it mostly disappear.

Therefore, different self-states can leave the same trace. Whether a person sincerely forgave, resigned himself, remained silent out of calculation, or acted with all of those emotions tangled together may become indistinguishable from current evidence after time passes. The words and actions left on the surface are the same, but the inner state at that moment may have been different. If so, it becomes impossible to reverse-calculate one exact chain of past self from the current state of the world.

Here one more uncomfortable fact must be added. What we commonly call the "present self" is thinner than we think. Strictly speaking, the present has no volume in time. It is not a span with width, but closer to the boundary where past and future meet. The moment I point to "I," that pointing is already holding on to a sensation, a thought, and a memory from just a moment ago. The pure present self becomes past before it can be grasped.

Thus most of what we call "I" is not the present but the past. I call myself the self from a moment ago, the self from yesterday, the self from years ago, and the narrative woven from those traces. The present has almost no volume, and the future has not yet come, so the belief in a substantial self effectively rests on the preservation of the past. Yet the traces of that past are incomplete. Most of the minute agonies, unspoken thoughts, and unchosen possibilities of the interior are not recorded. If the self is to exist as a substance, the past must preserve me. But the past does not preserve me whole.

Let us think of this more radically. At some point in the past, I could have chosen A or B. If free will is admitted, both paths were possible futures. But what if, after time passes, those two paths can merge again into a single point indistinguishable within the current state of evidence? Even if different agonies, emotions, and possibilities of choice existed in the past, if they create no observable difference in the present world, that difference informationally disappears. This is the merging of paths.

Suppose an elephant had been in your room and disappeared after erasing every trace. If the current room cannot be distinguished at all from a room in which an elephant had never been, what difference does the statement "there was an elephant" make in explaining the present world? Of course, we can imagine it as a story. But if there is no trace, no method of reconstruction, and no distinguishable fact in the present world that changes, the ground for holding that event as a substance within the present ontology disappears.

Past self-states are similar. Where in the universe is the subtle grain of emotion I felt at lunch ten years ago today recorded now? If it is not sufficiently left in memory, bodily state,

other people's records, or any physical trace in the world, and if it cannot be reached again exactly from the present me, in what sense can that grain of emotion be said to exist? That past self-state may once have been experienced, but from the present point of view it no longer remains as a substantial term.

Therefore, what does not exist is the kind of "self with free will that penetrates time" that we imagine. Past self-states are mostly not preserved as complete traces, are not exactly restored in the present, and disappear without making a difference in the distinguishable description of the world. Thus the way of understanding the self as a substance continued from the past collapses.

Then where is the ground of existence for the "I" that agonizes and chooses here and now?

2.1 From Excavation to Construction: The Self as a Non-Substantial Structure

At the point where the causal chain of the past has been lost, the view of the self must shift from "excavation" to "construction." The self is not a solid substance inherited from the past, but a "non-substantial structure" that erects itself toward the future. This structure does not answer the ontological question, "What was I in the past?" Instead, through the teleological question, "What am I moving toward?" it originates itself at every moment.

This connects with Kant's concept of the "kingdom of ends." In Kant's world, a rational being is not an object subordinated to natural law, that is, physical causality, but an autonomous legislator that establishes and obeys law for itself. Our model reaches the same conclusion. On the blank slate where physical evidence has been lost, the self is established only by declaring, like a constitution, the highest purpose: "I will approve myself as this kind of being." This declaration is not a simple resolution, but a gravitational point that binds scattered fragments of possibility into one subject. Therefore, the self is not an "object" to be discovered, but a "declared status" that is continually renewed within the kingdom of ends.

The mode of existence of this "declared self" is similar to the imaginary number i in mathematics. Let the real axis($\mathbb{R}$) be the world of physical causality and data on which we stand. With this real axis alone, the equation $x^2 + 1 = 0$ has no solution. In other words, physical causalism alone cannot answer the questions "Why must I take responsibility?" and "What makes the human being dignified?" In the physical world, responsibility is only the result of causality, but in the ethical world, responsibility is the manifestation of will.

Here we introduce a new dimension orthogonal to the real axis, namely the imaginary axis. If asked for its position on the number line, i cannot answer, but through the relation and mode of operation $i^2 = -1$, it exists quite strictly. The self is the same. Even if one dissects the biological brain, one does not discover a substance called "self." But when the mode of operation called "the structure that sets purposes for itself and takes responsibility" is introduced, the equation called the human being finally has a solution. Non-substantiality is not a defect, but a necessary condition for operating outside physical determinism.

This non-substantial structure is not the illusion of an isolated individual. It is a vast "Protocol of Recognition" maintained within a network of relations with others. A blockchain is in

itself nothing more than a sequence of 0s and 1s, but when participants share the collective declaration, or consensus algorithm, that "we will recognize this bit string as value," it acquires a value more robust than physical currency. The human self is similar. We may lack a "central server" called the soul or an immutable substance. But we live on a "mutual-recognition consensus" that treats one another, and oneself, as subjects with free will. If Kant's kingdom of ends is the ideal model, this protocol is its practical implementation. Like a blockchain that creates trust without central authority, we build the existence called "I" like blocks by cross-validating one another's declarations, even without an immortal soul.

In the end, our argument results not in nihilism but in salvation. The fact that past information is erased paradoxically opens the possibility of freedom, because the past cannot define or bind us forever. What fills that informational blank is not a physical residue, but our declaration and purpose. "I" am not a statue taxidermied in a museum. "I" am a phenomenon reconstructed at every moment on incomplete evidence, and a non-substantial structure that operates by approving a self-established purpose as a constitution. Having no substance means that we are not ghosts of the past, but can be newly originated each time as declarations toward the future. We are free precisely because we are not completely remembered, and we exist precisely because we declare that we will treat one another as ends.

3 The Socialization of Purpose Declaration: From the Individual Constitution to Social Input

If the self is not a substance preserved from the past, the human being becomes a subject only by declaring itself from the present toward the future. Therefore, purpose declaration is the ontological origin of the individual and the most fundamental input that society must read. This transition is the central hinge of this manuscript. Within the individual, purpose is the constitution that erects the self. Within society, it is the primordial input that must be coordinated.

The economy reads humans as holders of preference, politics reads humans as holders of opinion, and platforms read humans as sources of reaction. But before that, humans are beings who declare purposes. Purpose is not data discovered already complete, but a self-constitution that is formed, approved, and borne with responsibility. Therefore, society cannot stop reading humans merely as consumers, voters, and users. It must be able to read them as purpose declarers.

This transition leads directly from ontology to institutional theory. At the individual level, purpose declaration makes the self. At the social level, purpose declaration makes social input. At the civilizational level, the way purpose declarations are aggregated makes a new social contract. If purpose declaration is the constitution of the individual self, society must have a higher constitutional system that reads and coordinates those constitutions. The Third Protocol in the next section begins from precisely this demand.

4 Beyond Votes and Prices: The Third Protocol

Here it is necessary to return to the input structure of society as a whole. Politics aggregates votes, and economics aggregates prices and contracts. But in a society where execution costs have fallen sufficiently, these two protocols can no longer sufficiently recover society's coordination possibilities. In particular, conditional collective demands such as "I will send my child if this kind of school actually exists," "I will move if this kind of rule is guaranteed," or "If this many people gather, I will stake my money, time, and expertise" are too coarse and slow for politics, and for economics they are not yet sufficiently captured because the transaction object has not yet been formed.

Masuda Muneaki's *Intellectual Capital Theory* reads this change as a phase transition in consumer society. In the era of mass production and mass distribution, products themselves were scarce. But in the Third Consumer Society, where products and platforms overflow, the quantity of goods is no longer the core. What matters is "proposal." Consumers do not simply want more goods; they want meaningful options edited to fit their tastes and lifestyles. That is why today's companies do not stop at displaying products. They constantly say: you will like this taste; you will want this space; you will come to love this idea and this way of life. Strong modern brands, including Apple, sell products and at the same time propose to humans a worldview and a lifestyle.

We are now in the saturation of proposals. Algorithms, brands, and platforms push possible forms of life in front of us at every moment. What to watch, what to buy, what taste to have, and even what kind of person to become are endlessly proposed. This abundance paradoxically paralyzes the individual's ability to decide. Too many proposals weaken the power to ask what one can love for oneself, and further make one forget the fact that one can directly shape what one is able to love. Love is not always merely choosing the most plausible option among those given. Sometimes love is closer to gazing for a long time at a form that does not yet exist in the world, carving it, refining it, and finally wanting to breathe life into it. The tragic and beautiful intuition of Pygmalion lies here. He did not choose a proposed ideal type; he shaped for himself the object he could not help but love. But the age of proposals merely makes humans more refined choosers. It does not raise them up as beings who create and declare the world they will love.

Moreover, now that means are being maximized, proposal is becoming an increasingly blunt form. In an age when execution costs fall and the possibilities of production and combination explode, a recommendation that "you will like this" is no longer enough. What matters is not which of the already existing options one will choose, but what not-yet-existing thing one will make together, and under what conditions. Therefore, if the Third Consumer Society was an age in which companies and platforms proposed ways of life to humans, the transition needed from now on is a transition from proposal to declaration. Humans must first declare their own purposes, and then goods, organizations, cities, education, and relations must be inversely configured around those declarations.

What is needed here is not a system that replaces votes and prices, but another social input-

output protocol that operates in parallel with politics and economics. If politics was a system that expresses and aggregates social will through votes and sends it back into execution, and economics was a system that does the same through prices, then a society in which the total amount of intelligence grows and execution costs fall requires a third protocol that expresses and aggregates purpose itself and sends it back into execution. This is what is meant here by the Third Protocol. The core of the Third Protocol is that the act by which an individual publicly declares their purpose becomes a new social signal that simultaneously reveals willingness to participate, willingness to provide resources, and the possibility of long-term commitment. In other words, it expresses not only "what do you want to buy" or "who will you vote for," but "what world do you want, and if what conditions are met, with whom will you stake what, and how much." It then sends that expression back into organization, contract, and production as a third path. In this case, declared purposes do not remain scattered opinions. They combine with similar declarations to form intention clusters, and themselves become production signals toward goods, institutions, and organizations that do not yet exist.

The point is that this is not simply an opinion board or a more sophisticated demand-forecasting engine. The Third Protocol is not a system that extracts desire from user logs, but closer to a structure that makes humans read, think, discuss, refine purposes, and finally declare them. And future agentic intelligences will take in that history of declarations, processes of revision, and priorities of value, then align the coherence among one another's purposes and compress them into candidates for common purpose and contractible expressions. Therefore, the Third Protocol is a social input-output protocol that turns human reflection into public purpose declaration, and sends that declaration back into the responses of institutions, organizations, and agentic intelligences.

From the perspective of GU and CVTD, which will be discussed later, the Third Protocol is the institutional layer that connects philosophy and execution. If GU explains why the individual's declared purpose must be restored as the highest criterion of judgment, and CVTD shows how those purposes can be calculated in the language of intentional events π, then the Third Protocol explains the channel through which those declarations are socially expressed, aggregated, and returned to execution. Therefore, it is not an additional service, but a new coordination protocol that opens the social visibility and computability of declared purposes.

Schematically, it is as follows.

Protocol	Structure
1: Vote	**Basic input**: Votes, rights, opinions. **Aggregation method**: Elections, parliaments, public spheres, procedural legitimacy. **Social output**: Law, policy, allocation of power. **Core limitation**: Low frequency, bundling, majoritarian compression.
2: Price	**Basic input**: Prices, purchases, investment, labor, contracts. **Aggregation method**: Markets, firms, finance, transaction networks. **Social output**: Goods, services, allocation of capital. **Core limitation**: Focused on currently payable preferences.
3: Purpose	**Basic input**: Purpose declarations, conditional commitments, intention to contribute. **Aggregation method**: Intention clusters, GU-like adjustment, negotiation among agentic intelligences. **Social output**: Organizations, institutions, production, intentional events π. **Core limitation**: Not yet standardized.

If the First Protocol asks "who has authority," and the Second Protocol asks "what will be paid for, and how much," then the Third Protocol asks "what world do you want, and what will you stake under the conditions in which that world becomes possible?" Therefore, the Third Protocol is not the abolition of politics and economics, but a complementary circuit that inputs, at higher resolution, the purpose structures that the two have lossily compressed.

5 The Internal Question of the Third Protocol: How Are Purposes Adjusted?

Once the necessity of the Third Protocol is confirmed, the next question immediately arises. Receiving purpose declarations as social input does not mean that those declarations automatically form a coherent order. Purposes collide with one another, misunderstand one another, and sometimes understand even themselves unclearly. Therefore, the Third Protocol cannot be a simple bulletin board or aggregation device. It needs a normative coordinate system that can judge under what conditions declared purposes are adjusted more justly.

At this point, Generalized Utilitarianism (GU) is proposed. GU is a normative engine for interpreting and adjusting conflicts of purpose inside the Third Protocol. In other words, if the Third Protocol is a social theory that receives purpose declarations as social input, GU functions as a normative theory that asks what should count as a better adjustment when those inputs collide.

6 From Declared Purpose to GU

If the self is understood not as a substance of the past but as a declaration toward the future, ethics and institutions too must be realigned around the structure of that declaration. Existing

systems have generally been good at aggregating already revealed preferences or coordinating already established transactions, but they have not placed at the center of judgment which purposes are formed, which purposes are publicly approved, and which purposes can be adjusted with one another. But if the self originates through purpose declaration, social judgment too must be designed around the relations among those declared purposes.

At this point, Generalized Utilitarianism (GU) can be understood not as a mere variation of utilitarianism, but as an attempt to restore purpose declaration as the higher criterion of social judgment. What matters here is not to impose a single universal purpose on humans. Rather, while presupposing the plurality of purposes declared by each individual and group, the task is to establish a new coordinate system that asks what structure those purposes form, where they collide, and where they can converge. GU is the first philosophical proposal toward precisely this coordinate system.

7 The Proposal of Generalized Utilitarianism (GU)

If the essence of the self is not a "discovered substance" but a "declared purpose," then the social system we need must also change fundamentally. Blind profit or efficiency can no longer be the criterion. We need a new coordinate system that takes the purpose declared by each individual as a real criterion, rearranges existing academic disciplines, and brings into the basic structure of society as a whole the function of "crossing from mountain to mountain" that Nietzsche had left only to exceptional individuals. We call this Generalized Utilitarianism (GU).

Under these conditions, Generalized Utilitarianism (GU) asks again what "justice" is, and generalizes classical utilitarianism in the following three dimensions.

1. It expands utility from a simple sum of pleasure and pain into a higher-dimensional function that includes states of knowledge, relational structure, and the possibility of empathy.
2. It shifts the primary criterion of judgment from "which option is more efficient" among possible choices to adjustment and convergence among the purposes declared by each individual.
3. It understands justice not as an already given state, but as the direction in which the expected utility of possible future trajectories converges, that is, as an "approximation to justice."

7.1 Core Intuition: Ignorance and Purpose

The core intuitions that support GU are as follows.

First is the thesis that "all avoidable suffering, if traced sufficiently far back through causality, ultimately converges on the result of ignorance." This is the strong assumption that if omniscience had existed, structures, institutions, and relations would not have been designed painfully. It is a response to the resigned axiom that "some suffering is an inevitability of the structure of the world," and an attempt to attribute suffering not to the essence of the world but to a lack of information.

Second is the thesis that "the primary ground of judgment should not be possibility(what is possible), but purpose from values(what ought to be pursued)." Possibility is information that refines means, and purpose is the criterion that divides judgment. The plurality of purposes is both a fact and a norm, and the project of asking "the purpose of human existence" as a single universal proposition is close to a category error. Purpose should be treated not as an essence to be discovered, but as a structure declared and designed by each individual or group.

With regard to free will, GU suspends the metaphysical or physical debate over whether it is physically real, and adopts only free will as a normative axiom. That is, we must live in a world that presupposes the individual as the atomic unit of social decision-making. This is not a proposition that must be "proven" by argument, but a minimum line of defense to prevent human beings from being reduced to pure means indistinguishable from natural phenomena or algorithms. If free will is not adopted as an axiom, the questions "who is responsible" and "who sets purposes" themselves become meaningless, purpose disappears, and the human being is completely instrumentalized.

Adopting free will as an axiom means acknowledging that the self is not completely reducible to a pure causal chain. In that case, the self must define itself at least once from its own internal will rather than from external causality. The result of fixing this originating will, or constitutional will, in language is purpose declaration. Therefore, "purpose" here is not a simple list of desires, but the content of the highest will the self adopts in order to originate itself.

From this point of view, the individual's purpose declaration is redefined as the individual's constitution. Just as the constitution of a state defines its identity and boundaries, the highest purpose declaration defines the boundary of the self. If this highest declaration changes critically, the self is re-originated as the "nth self," just as a state may be called the "nth republic."

However, GU is not an absolute, external judge. In a society, publicly declaring "according to GU, the answer is X" about controversial issues such as abortion, euthanasia, or revolution is itself a new policy π, and that utterance again affects future distributions and utilities. In other words, a subject inside the system cannot structurally deliver a complete judgment on the system as a whole. This repeats, at the level of social decision-making, Godel's incompleteness theorem, according to which a formal system cannot produce all truths about that system from within itself.

GU squarely acknowledges this limitation and derives a kind of "game of willed ignorance." Even if calculation seems to favor one side on some question, the question of "when, to whom, and in what form the answer is to be spoken" itself becomes another decision variable. In some cases, the paradox arises that "not speaking even when one knows" is closer to an approximation of justice. GU does not hide this self-referentiality; it includes the "ethics of not speaking the truth" as an element inside the theory.

7.2 From Philosophy to System

Philosophy is beautiful, but systems must operate. In order not to leave this vast vision as a vague hope, we formalized ethical judgment as a calculable formula on a network of utility. This

is an attempt to provide an ideal criterion capable of evaluating and approximating complex moral decisions in reality, a coordinate system like the "frictionless plane" in physics. However, the formulas presented here are not a completed methodology or a closed computational system. They are preliminary proposals for exploring how the directionality of ethical judgment and purpose adjustment can be translated into a calculable language, and they deliberately leave substantial room for later development and expansion.

We posit the ideally just state as a state in which the utilities of all agents, including the utilities and empathy of others, are optimally resolved. If we posit a hypothetical omniscient judge who implements this just state, this judge would gather both the subjective happiness of each individual and the empathy information they have toward others, and would then make an evaluation. We express this ideal judging process with the concept of "omniscient empathy," and define empathic and cognitive completeness as $EERI = 1$ as its prerequisite.

Ultimately, our goal is to find principles that actually approximate this state of omniscient empathy. Of course, since human beings cannot reach an omniscient state, we seek to realize Proximal Justice within the constraints of reality. This means adjusting decision-making structures so that they come as close as possible to the state of omniscient empathy, as the best approximation of justice attainable under present cognitive limits.

The rest of this manuscript is organized as follows. First, it presents the philosophical path from the declarative structure of the self to GU, and then translates this into a calculable language through the formalization of omniscient empathy. It then proposes CVTD as a preliminary sketch of the execution layer, and finally proceeds to why a Third Protocol beyond votes and prices becomes socially necessary, and why reading becomes the first wedge that bears the conditions of its emergence.

The concepts that appear later should not be read as separate theories laid out in parallel, but as one structure in which each concept occupies a different layer. The Third Protocol is the central thesis, a social theory that purpose declaration should be received as a third social input after votes and prices. GU is the normative engine that deals with what should count as a more just adjustment when those inputs collide. Contractual Value Transfer Dynamics (CVTD) is the execution engine that records value changes and alignments when declared purposes descend into all intentional events π. The Tri-Basis Drive Model (TBDM) is a macro-interpretive model that explains what macroscopic net forces countless microscopic events become across the three axes of possibility, sociality, and coherence.

8 Toward the Formalization of GU

In the preceding section, GU was proposed as a principle of social judgment derived from the declarative structure of the self. What is now needed is not to leave this proposal as a simple philosophical declaration, but to translate the structures of utility, empathy, and judgment into calculable form. In other words, while presupposing the plurality of purposes, the asymmetry of empathy, and the incompleteness of information, we must be able to formally handle the direction in which justice converges.

The formalization in the next section addresses precisely this problem. What matters here is not that formulas replace philosophy, but that they translate the coordinate system presented by philosophy into an operable structure. If GU is the macroscopic frame that presents the directionality of justice, its formalization becomes the intermediate stage that shows how that directionality can be converted into the language of actual judgment and design.

9 Theoretical Frame: Formalization of Omniscient Empathy

This section constructs a mathematical model of Generalized Utilitarianism and explains how morality can be computed on a network of utility and empathy. This is an attempt to provide an ideal criterion for evaluating and approximating moral decisions in reality, and may be seen as analogous to a "frictionless plane" in physics.

9.1 Agents, Internal State, and Utility H_n

We denote moral agents by $n = 1, 2, \ldots, N$, and assume that each agent n has an internal state I_n that encompasses all factors related to their welfare, such as physical health, mental state, material resources, desires, and so on. Based on this internal state, each agent's utility or happiness H_n can be expressed as a function of both their own state and empathic responses toward others. Concretely, we define it as follows:

$$H_n = I_n + \sum_{i=1}^{N} C_{n,i} \cdot EERI_{n,\pi,i} \cdot H_i \ ,$$

where $C_{n,i}$ is the empathy coefficient that agent n has toward the happiness of agent i, indicating the degree to which n is affected by i's utility. If $C_{n,i} = 0$, this means agent n has no concern at all for i's happiness; if it is positive, it means that n feels some utility or disutility along with changes in i's happiness. In this way, each agent's utility H_n reflects not only their own state I_n, but also emotional interactions with others, and can encompass emotions such as empathy, solidarity, or moral anger.

This equation shows that one agent's happiness depends on the happiness of other agents. In particular, in equation (1), H_i is again determined by the influence of other beings, so the entire structure $H = (H_1, H_2, \ldots, H_N)$ has a mutually referential character. This mutually referential structure will be handled more clearly later through the fixed-point interpretation.

9.2 Educated Empathic Resolution Index (EERI)

Actual human beings do not possess complete rationality or infinite information. Also, when human beings understand the state of others, they do not merely "not know or ignore"; they often make cognitive evaluations in a distorted or even inverted state of interpretation. To reflect this, Generalized Utilitarianism (GU) introduces the Educated Empathic Resolution Index (EERI), an index indicating how accurately and coherently each agent n understands and interprets the

internal state of another agent i under policy π. EERI has the following range:

$$EERI_{n,\pi,i} \in [-1, 1] \ ,$$

- $+1$: Completely and coherently understands the emotions and utility of other i (ideal clarity)
- 0: No information about, or no understanding of, emotions and utility (ignorance)
- -1: Completely distorts or inversely understands the emotions of other i (misunderstanding or hostile interpretation)

This index reflects not only whether understanding exists, but also the directionality and coherence of that understanding, and makes it possible to quantify emotional or cognitive distortion. For example, prejudice, hostility, hatred, and so on toward another can be expressed as $EERI_{n,\pi,i} < 0$, which has the effect of reflecting the utility of the other i inversely in the overall judgment of justice. That is, if n hates i, even if $C_{n,i}$ is positive, a negative $EERI_{n,\pi,i}$ means that an increase in i's happiness instead acts as unhappiness for n.

The concept of EERI is a device for integrating the cognitive limits and biases of human judgment. Due to various cognitive limitations in actual people, such as scope insensitivity, in-group bias, and present bias, different moral evaluations can appear among individuals even regarding the same situation. To reflect this reality in the model, EERI represents each agent's cognitive and emotional clarity. A low EERI value indicates irrational judgment or bias, while a high EERI value reflects more deliberated and educated ethical judgment. As a result, EERI serves to structurally correct or reveal distortions caused by bias.

9.3 Proximal Justice

In Generalized Utilitarianism (GU), Justice is not a simple summation of policy outcomes. Justice is defined as the state in which the emotional, empathic, and understanding structures among beings are aligned by the ethical operator π, and the overall utility structure $H = (H_1, H_2, \ldots, H_N)$ converges to an information-theoretic fixed point.

To explain this structural adjustment, we introduce the concept of Ontological Self-Extension (OSE). We define the degree of ontological self-extension that agent n has toward other i as follows:

$$OSE_{n,i} := C_{n,i} \cdot EERI_{n,\pi,i} \ \ ,$$

In particular, for oneself ($i = n$), since $C_{n,n} = 1$, this can be written simply as:

$$OSE_{n,n} = EERI_{n,\pi,n} \ \ .$$

Justice converges in the limit where $OSE_{n,i}$ approaches 1 for every pair of beings n, i. That is, justice can be seen as being realized in the ideal limiting state in which beings fully understand

and identify with one another. Expressed as a formula:

$$\text{Justice} := \lim_{OSE_{n,i} \to 1} \pi^* \ ,$$

where π^* refers to the ultimate ethical operator corresponding to the above limiting state. In reality, we cannot reach a state of omniscient empathy in which $OSE = 1$ in every relation, but π^* indicates the directionality of convergence toward this ideal.

Practical Judgment Structure: Proximal Justice π^*. Since complete OSE convergence is impossible in reality, GU performs justice judgment through approximate optimization that reflects realistic empathy and understanding structures. It can be formulated as follows:

$$\pi^* = \arg\max_{\pi} \sum_{n=1}^{N} EERI_{n,\pi,n} \cdot H_n \ ,$$

That is, it seeks the π that maximizes the sum obtained by multiplying each agent n's utility H_n by $EERI_{n,\pi,n}$, the index of how clearly that agent currently understands itself. Here π is not a simple set of outcomes, but, as mentioned above, an ethical operator that affects the entire structure.

Each being's total utility H_n is constructed as follows:

$$H_n = I_n + \sum_{i=1}^{N} C_{n,i} \cdot EERI_{n,\pi,i} \cdot H_i \quad ,$$

This structure shows that utility H_n is not simply an individual's subjective reaction, but a value structurally transmitted and reflected through empathy coefficient C and resolution of understanding $EERI$. In particular, because H_i recursively includes the emotions of others again, the whole structure expands into the following infinite self-referential system:

$$H_n = I_n + \sum_{i \neq n} C_{n,i} \cdot EERI_{n,\pi,i} \cdot \left(I_i + \sum_{j \neq i} C_{i,j} \cdot EERI_{i,\pi,j} \cdot H_j + \cdots \right) \ ,$$

This self-referential structure shows a series of repeated embeddings and can, in theory, continue infinitely. Therefore, for a just judgment to be established in such a system, the overall utility vector H must have a stable value.

Fixed-Point Interpretation. The above self-referential structure must converge in fixed-point form for justice judgment to be mathematically established. The overall utility vector $\mathbf{H}$ must satisfy the following fixed-point condition:

$$\mathbf{H} = F_\pi(\mathbf{H}) \ ,$$

where F_π is the function, or operator, representing the whole transformation of empathy, understanding, and emotion flows induced by the ethical operator π. If some $\mathbf{H}^*$ exists as a fixed

point satisfying this equation, and if the actual process can converge to $\mathbf{H}^*$, justice can be said to exist in stable form. At this time, π^* is interpreted as the *directionality of the path* by which $\mathbf{H}$ converges to $\mathbf{H}^*$. In other words, π^* is the operator that determines the direction of change in the distribution of utility within H-space, and repeated application of that operator converges to a just state.

The Nature of the Ethical Operator π. Policy π is not a simple act choice, but a high-dimensional nonlinear ethical operator acting on the entire being-based ethical information structure. More formally, π can be viewed as a transformation that acts on the state of being O of the world and induces the next state O':

$$\pi : O_t \to O_{t+1} \quad ,$$

where O_t includes the following whole structure:

$$O_t = I_n(t), x_n(t), H_n(t), C_{n,i}(t), EERI_{n,\pi,i}(t)_{n,i} \quad ,$$

That is, π acts on the world's entire ontological structure O and reconstructs all elements such as emotion, value, empathy, and understanding. π^* means the ideal operator among these operators that best realizes the directionality of justice's convergence.

9.4 Mathematical Interpretability of GU: Comparison, Centrality, and Solution Selection

One caution is needed here. The mathematical concepts below are not a list of authorities meant to prove GU as an already completed theorem system, but interpretive lenses showing what kind of computational language GU can develop into. In particular, I_n and H_n are not simple happiness scores that directly measure the pleasures of different individuals on the same ruler. Interpersonal purpose-fulfillment is fundamentally difficult to compare directly. The same hour, the same dollar, and the same praise can be a condition of survival for one person and almost meaningless noise for another. Therefore, in GU, I_n should be understood as an internal state normalized within that person's value basis, reflecting each individual's declared purpose, bodily conditions, relational structure, time axis, risk aversion, and degree of self-understanding. H_n is a networked value formed as that internal state connects with the utility, empathy, and resolution of understanding of others.

In this sense, GU's utility vector $\mathbf{H}$ has a structure closer to PageRank or eigenvector centrality than to a simple summation table. A being's utility is not an isolated score, but is stabilized through connections with other beings, empathic weights, resolution of understanding, and repeated self-referentiality of self-understanding. In other words, what matters is not "who says they have greater subjective happiness," but what fixed point the entire empathy network converges to, and how well that fixed point is aligned with each person's declared purpose.

Justice judgment can be seen as a problem of solution selection on top of this fixed point. If each agent responds to a possible policy π, and that response again changes the response conditions of other agents, the social state can be interpreted as a game-theoretic equilibrium problem. When appropriate continuity, convexity, and compactness conditions are given, Kakutani's fixed-point theorem becomes one mathematical background for discussing the existence of Nash equilibrium. But GU is not satisfied merely with the fact that some equilibrium exists. Bad equilibria can also exist. Thus GU asks not "does an equilibrium exist," but "which equilibrium should be selected as an approximation of justice."

This selection problem can also be interpreted from the perspective of Nash bargaining. When there is a feasible set $\mathcal{F}$ in which multiple purpose declarations can be satisfied simultaneously, and a disagreement point d, one simple principle of solution selection can be written as follows.

$$\pi^{NB} = \arg\max_{\pi \in \mathcal{F}} \prod_{n=1}^{N} (H_n(\pi) - d_n)$$

This does not mean that every problem should be reduced to a multiplicative bargaining solution. It merely shows that, inside the Third Protocol, declared purposes can be handled not by simple majority rule or payment capacity, but in the language of axiomatic solution selection that considers each party's disagreement point and margin of improvement.

From another perspective, GU's fixed point calls to mind the least fixed point of Tarski's fixed-point theorem. If the ethical operator F_π is monotone on a partially ordered lattice, then among possible fixed points we can first consider the fixed point that requires the smallest assumptions and the least coercion. This connects with the design principle of the Third Protocol. Purpose adjustment should not move in the direction of forcibly deleting as many purposes as possible, but in the direction of creating stable alignment with the least possible suppression. In a constrained reality, the interpretation of Lagrange multipliers is also useful. When constraints such as resources, time, safety, trust, information, legal rights, and conscience are given as constraint equations $g_k(\pi) \leq 0$, the problem of purpose adjustment can take the following form.

$$\mathcal{L}(\pi, \lambda) = \Omega(\mathbf{H}(\pi)) - \sum_{k} \lambda_k g_k(\pi)$$

Here λ_k is not a mere computational auxiliary term, but a shadow price, indicating how large a shadow a constraint casts on the achievement of social purposes. For example, if the λ of lack of trust is very large in a society, building trust infrastructure may produce a greater improvement in purpose than adding more money.

Finally, purpose declaration is also a problem of common knowledge. The fact that some individual inwardly wants something, the fact that they publicly declared it, the fact that others know that declaration, and the fact that the individual knows that others know it are completely different. The purpose declaration of the Third Protocol creates precisely this fixed point of belief. A purpose cannot become social input when it remains only an inner desire, but when it is publicly approved and repeatedly referenced, it becomes a condition of action for

contracts, negotiations, organizations, and agentic intelligences.

The most difficult objection at this point is the problem of evil purposes. If someone declares, "I am an evil person. Who will stop me?", does the principle of respecting purpose declaration have to permit that declaration as it is? The answer is no. In GU, purpose declaration is social input, not immunity. Declaration is an act that, instead of executing hidden desire invisibly, makes it calculable and contestable what harm that purpose does to the possibility, sociality, and coherence of others. Evil purposes are not treated mysteriously outside suppression, but are revealed inside the Third Protocol, evaluated inside GU, and their harm paths as intentional events are traced inside CVTD. Therefore, the various mathematical references in this section do not declare the completion of GU, but are more like a map indicating what formal tools the adjustment problem of purpose declaration can be connected to.

9.5 Temporality and the Moment of Justice Judgment

In GU theory, time(t) is included as another axis of the ontology vector space O. The whole ethical structure can be expressed as follows:

$$O_t = O \times \mathbb{R} \ ,$$

that is, the state of O changes over time. Being n is defined by its internal state $I_n(t_0)$ and total utility $H_n(t_0)$ at time t_0, and justice judgment is always performed by the judge n at the current time t_0. At this point, past and future beings, that is, my past self or my future self, are also regarded as one of the others i. If the empathy coefficient $C_{n,i}$ and the resolution of understanding $EERI_{n,\pi,i}(t_0)$ toward past or future generations are low, their influence on the current n's total utility $H_n(t_0)$ is naturally reduced. This means that beings more distant in time may be considered less in present judgment, but within the GU structure this influence is explicitly modeled.

As a result, justice judgment in reality has the following structure:

$$\pi^*(t_0) = \arg\max_{\pi} \sum_{n=1}^{N} EERI_{n,\pi,n}(t_0) \cdot H_n(t_0) \ ,$$

This means that justice is not an absolute value dependent on time, but an ethical alignment performed on the current information structure. Agents at the present time calculate $\pi^*(t_0)$ by considering their own and others' current and expected future utility, and the degree of empathy among them. In that process, consideration for future generations or past generations is also continuously reflected through the values of C and $EERI$, giving the structure a feature intended to prevent the weakening of moral responsibility caused by temporal distance.

Philosophical Summary. The above discussion can be summarized as follows:

$$\text{Justice} := \text{Omniscient Empathy} \succ \text{Proximal Justice} = \pi^* \xrightarrow{OSE_{n,i} \to 1, \forall n,i} \text{Justice} \ ,$$

That is, justice in reality always converges toward Omniscient Empathy, and π^* corresponds to Proximal Justice, the approximate justice that implements that convergence in reality. Advanced AI systems should operate as ethical meta-operators that assist this convergence.

10 The Socialization of Purpose: From Declaration to Intentional Event

If GU is the normative coordinate system for judging conflicts of purpose, the remaining problem is execution. Society cannot stop at understanding purposes. It must send them back into intentional events such as declaration, contribution, design, persuasion, organization formation, resource movement, contract, production, refusal, and adjustment. Therefore, the Third Protocol needs a second internal engine. One is GU, which deals with just adjustment among purposes, and the other is an execution grammar that records and calculates how those purposes descend into real events.

Here CVTD is not a system that extracts purpose from logs on behalf of the person. Purpose, as said above, is not data to be extracted, but a self-constitution that is formed and declared. What CVTD does is describe how that declared purpose is implemented in actual intentional events, and how the different value bases of multiple agents create positive sums or generate conflict. Contracts and exchanges are only important special cases of intentional events, not the entire object of CVTD. Therefore, CVTD does not replace purpose declaration; it deals with the process by which declared purposes are translated into social events.

11 From GU to CVTD

The formalization of omniscient empathy shows that GU is not merely a rhetorical ideal, but a calculable directionality. Yet a gap remains. GU is a coordinate system that tells us what is closer to justice, but purposes in the real world always occur in the form of concrete intentional events. Some people contract, some create organizations, some persuade, some provide care, and some change the state of the world by refusing a particular act. Therefore, an execution layer is needed to connect the directionality of justice with the language of actual intentional events.

CVTD is an attempt to bridge precisely this gap. If GU is the macroscopic norm for the adjustment and convergence of declared purposes, CVTD becomes the microscopic execution language that deals with how those purposes can be expressed in the physical world as event matrices, value bases, and commitment structures. The next section should be read not as abandoning GU, but rather as translating GU into calculable interactions in reality. From this point, let the notation be made clear. The macroscopic ethical operator conventionally called π in the GU section is decomposed in CVTD into a set of possible intentional events or policies $\mathcal{P}$, and an individual intentional event that is an element of that set is called π. The matrix representation of event π is written $\Pi(\pi, t)$. However, CVTD too is not a completed calculus, but an initial proposal for how the mathematical skeleton of an execution layer might be constructed.

12 Contractual Value Transfer Dynamics (CVTD): The Calculation of Value and the Inverse Derivation of the Self

If Generalized Utilitarianism (GU) is the macroscopic coordinate system that presents the direction of convergence of Justice, Contractual Value Transfer Dynamics (CVTD) is the Execution Layer for actually calculating and operating this on the physical world. Among all events in the universe, excluding natural phenomena, we take as the descriptive object of CVTD all "intentional events" created with purpose by agents who possess the previously defined "non-substantial self."

Therefore, the expression "contractual" in CVTD does not mean only legal contracts or explicit agreement procedures. It means that intentional action always has the structure of conditions, expectations, responsibility, commitment, attribution, possibility of violation, and enforceability. Advising a friend, founding a company, writing a text, rejecting a proposal, designing community rules, transferring resources, and entering into an explicit contract are all different forms of π. Contract and exchange are important cases of CVTD, but CVTD itself is a dynamics for describing all intentional events π.

In this chapter, we decompose the macroscopic policy $\mathcal{P}$ of GU into a set of individual intentional events π, and concretize each event's matrix representation $\Pi(\pi, t)$ and the empathy coefficient C as linear-algebraic entities. Through this, we explain the mechanism of increasing total value and the process of inversely deriving the direction of the self in actual intentional-event scenarios.

12.1 General Form of the Intentional Event π

In CVTD, a single intentional event is not a simple transaction record, but a purpose-bearing transformation that acts on the world state. Let the agent set be $\mathcal{A} = 1, \ldots, N$, the world state at time t be O_t, and the common value space be $K = \mathrm{span}(k_1, \ldots, k_d)$. Then event π_t can be defined as the following tuple.

$$\pi_t = (A_\pi, X_\pi, \Phi_\pi, \tau_\pi, \Gamma_\pi, \mu_\pi)$$

Here $A_\pi \subseteq \mathcal{A}$ is the set of agents participating in or affected by the event, $X_\pi \subseteq O_t$ is the bundle of objects, acts, relations, and institutional states handled by the event, Φ_π is the conditions and context under which the event is established, $\tau_\pi : O_t \to O_{t+1}$ is the world-state transformation, Γ_π is the structure of commitment, responsibility, constraints, enforcement, and trust, and μ_π is the purpose signature indicating which declared purpose or purpose candidate this event is connected to.

The basic operation of the event is as follows.

$$O_{t+1} = \tau_\pi(O_t)$$

Natural phenomena do not have a purpose signature μ_π or a commitment structure Γ_π, and

therefore are not the direct object of CVTD. Conversely, if an agent's purpose intervenes in the world state and produces some transformation, whether it is a transaction, design, dialogue, refusal, or institution formation, it becomes a π of CVTD.

12.2 Value-Change Vector and Event Matrix

Each agent n has an evaluation operator $\mathbf{C}_n^t : O_t \to K$ that maps the world state and event object into that agent's value basis. When event π_t occurs, the subjective value change of agent n is expressed as follows.

$$\Delta_n(\pi, t) = \mathbf{C}_n^t(O_{t+1}) - \mathbf{C}_n^t(O_t)$$

This value includes not only the movement of money, but changes along multiple value axes such as time, trust, strategic assets, reputation, degrees of freedom, responsibility, relational stability, and purpose alignment. Therefore, the same event appears as a different row vector for each agent.

The matrix representation of the whole event is as follows.

$$\Pi(\pi, t) = \begin{pmatrix} \Delta_1(\pi, t) \\ \Delta_2(\pi, t) \\ \vdots \\ \Delta_N(\pi, t) \end{pmatrix} \in \mathbb{R}^{N \times d}$$

Therefore, below, π refers to the intentional event itself, and $\Pi(\pi, t)$ refers to the linear-algebraic event matrix that the event creates at time t. This distinction is important. Even for the same event π, if the time of evaluation, value basis, or set of participants changes, the matrix representation $\Pi(\pi, t)$ can change.

The total vector of social value change is defined as follows.

$$G(\pi, t) = \mathbf{1}^T \Pi(\pi, t) \in K$$

Also, if we place the social evaluation function or alignment function $\Omega_t : K \to \mathbb{R}$ provided by GU, the normative score of an event is expressed as follows.

$$W(\pi, t) = \Omega_t(G(\pi, t))$$

Therefore, in CVTD, a Positive Sum does not simply mean that the total amount of money increases. It means that the total vector of the event matrix has a positive direction along the relevant value axes, or is evaluated as a higher value under the GU-like alignment function. Since a positive sum on one axis can conceal the violation of another axis, Γ_π and Φ_π must record consent, responsibility, coercion, information asymmetry, and long-term costs together.

12.3 Dynamics as Event Flow

CVTD goes beyond the notation of a single event and deals with the temporal flow of events. Let $\mathcal{P}_t^{\text{int}}$ be the set of intentional events active at time t, and $\lambda_\pi(t)$ be the intensity or execution ratio of each event. Then the accumulated value state $Z(t)$ of society is updated as follows.

$$Z(t+1) = Z(t) + \sum_{\pi \in \mathcal{P}_t^{\text{int}}} \lambda_\pi(t) \Pi(\pi, t)$$

This equation makes CVTD not a mere ledger, but a dynamics. The Third Protocol receives declared purposes and generates the possible event set $\mathcal{P}_t^{\text{int}}$; GU evaluates the normative alignment of those events; and CVTD calculates how the actually selected events change the value states of each agent and of society.

12.4 Vector-Space Mapping of GU Variables

12.4.1 Policy Set and Event Matrix

The macroscopic policy that determines the direction of convergence of justice in GU is expressed in CVTD as a set of concrete intentional events $\mathcal{P}^{\text{int}}$. Each individual intentional event $\pi \in \mathcal{P}^{\text{int}}$ changes the total value of the system through its matrix representation $\Pi(\pi, t)$. Therefore, every intentional interaction generated by a purpose-bearing agent in the physical world is treated as a local event of the macroscopic policy $\mathcal{P}$.

12.4.2 Empathy Coefficient C and Value Basis Vector (C)

In GU, the empathy coefficient $C_{n,i}$ was the weight by which agent n reflects the utility of other i in its own utility function. In CVTD, a vector space, this C is elevated into the unique Value Basis Vector $\mathbf{C}_n$ of agent n. The perspective from which the agent sees the world, the value axes it regards as important, and the evaluation that converts an object into utility are all determined by this basis vector $\mathbf{C}_n$. In other words, $\mathbf{C}_n$ is the structural filter through which agent n interprets the world, and the interface through which it connects with others.

12.5 Composition and Subjectivity of the Event Matrix $\Pi(\pi, t)$

One of the simplest prototypes of an intentional event is a two-party, two-object exchange(2×2 Primitive). The full object of CVTD is broader than this, but the exchange event most clearly shows how different value bases can create a positive sum. Consider an event in which agents A and B exchange objects x and y respectively($A \to B : x,\ B \to A : y$). At this point, the matrix representation $\Pi(\pi, t)$ of event π is defined as follows.

$$\Pi(\pi, t) = \begin{pmatrix} \mathbf{C}_A(y) - \mathbf{C}_A(x) \\ \mathbf{C}_B(x) - \mathbf{C}_B(y) \end{pmatrix}$$

Here the first row is the net value change from the perspective of agent A, and the second row is the perspective of agent B. The key point is that even the physically same object x is

transformed into entirely different value values as it passes through the different bases $\mathbf{C}_A$ and $\mathbf{C}_B$. This difference of basis, or Misalignment, is precisely the driving force that allows the same event to create a Positive Sum for both sides.

12.6 Case Study: Freelance Developer and Startup Founder

To clarify the operating principle of CVTD, we analyze, as a matrix, a contractual exchange scenario that is one special case of an intentional event.

Scenario Setup. We define the value space as the three-dimensional axes

$$K = [k_1 : \text{finance}, k_2 : \text{time/effort}, k_3 : \text{strategic asset}].$$

- **Agent A (developer)**: Has a basis $\mathbf{C}_A$ that values "stability(Cash)" and "work-life balance(Time)."
- **Agent B (founder):** Has a basis $\mathbf{C}_B$ that values "product asset(Asset)" and "speed," and accepts risk.
- **Exchange objects**: x (developed code), y (service fee)

Subjective Evaluation (Projection). The evaluation of objects according to each agent's basis is as follows.

- **A's evaluation**: Writing the code is a loss of time($k_2 : -10$), but the service fee is a large financial gain($k_1 : +15$).

$$\mathbf{C}_A(x) = [0, -10, 0], \quad \mathbf{C}_A(y) = [+15, 0, 0]$$

- **B's evaluation**: Paying the service fee is a financial loss($k_1 : -15$), but acquiring the code is an enormous asset value($k_3 : +50$).

$$\mathbf{C}_B(x) = [0, 0, +50], \quad \mathbf{C}_B(y) = [-15, 0, 0]$$

Derivation of the Event Matrix $\Pi(\pi_{dev}, t)$. We construct the event matrix by substituting the above values into the format defined earlier.

$$\Pi(\pi_{dev}, t) = \begin{pmatrix} (15-0) & (0-(-10)) & (0-0) \\ (-15-0) & (0-0) & (50-0) \end{pmatrix} = \begin{pmatrix} 15 & 10 & 0 \\ -15 & 0 & 50 \end{pmatrix}$$

Interpretation. This matrix contains information beyond a simple ledger record. The sum of physical goods, money k_1, is 0, but the total value vector of the entire matrix is $[0, 10, 50]$, which is positive. In other words, different bases $\mathbf{C}$ intersected and created something from nothing, namely time efficiency and asset value. Through this matrix, CVTD records not only "what was traded," but "why it was traded."

12.7 Linear-Algebraic Dynamics: Inverse Derivation and Alignment of the Self's Direction

By unifying events as matrices π and agents' dispositions as vectors $\mathbf{C}$, we can bring the problem of the self and ethical judgment into a "computable" domain.

12.7.1 Eigenvectors and the Inverse Derivation of Purpose Direction

Suppose we collect the row vectors of A from hundreds of intentional events $\pi_1, \pi_2, \ldots$ involving agent A and construct the event-history matrix $M_A(T)$.

$$M_A(T) = \begin{pmatrix} \Delta_A(\pi_1, t_1) \\ \Delta_A(\pi_2, t_2) \\ \vdots \\ \Delta_A(\pi_T, t_T) \end{pmatrix}$$

At this point, the covariance matrix is as follows.

$$\Sigma_A = \frac{1}{T} M_A(T)^T M_A(T)$$

The eigenvector $\mathbf{e}_A$ corresponding to the largest eigenvalue of Σ_A indicates the dominant value direction toward which the events that A has repeatedly chosen, accepted, and generated are oriented.

In the above example, if A continuously takes gains of the form $[15, 10, 0]$, then $\mathbf{e}_A$ lies on the plane of k_1(finance) and k_2(time). This mathematically shows that A's practical purpose direction is close to maintaining a stable life. By contrast, B's eigenvector would strongly point to k_3(strategic asset), showing that B's practical purpose direction lies in achievement and expansion. That is, in CVTD the self is not a metaphysical substance, but the directionality repeatedly revealed by matrices of intentional events.

However, this does not mean replacing declared purpose with behavioral logs. As stated above, purpose declaration is the individual's self-constitution, and the event history of CVTD shows how that declaration is implemented, distorted, and revised within actual events. If the declared purpose vector of agent A is $\mathbf{p}_A$, the degree of alignment between the declaration and the event history can be measured as follows.

$$\alpha_A = \frac{\langle \mathbf{p}_A, \mathbf{e}_A \rangle}{\|\mathbf{p}_A\| \|\mathbf{e}_A\|}$$

The higher α_A is, the more aligned the declared purpose and the direction of repeated intentional events are; the lower it is, the more there exists a gap that must be readjusted among declaration, self-understanding, environmental constraints, and actual choices.

12.7.2 Change of Basis and Mutual Understanding

Conflict between agents occurs because they use different bases $\mathbf{C}$. If the same value space K and the bases of both sides are specified, then through the transition matrix $P_{A\to B}$ of linear algebra, the values of agent $A(\mathbf{C}_A)$ can be exactly coordinate-transformed into the perspective of $B(\mathbf{C}_B)$.

$$[\text{Value}]_B = P_{A\to B} \cdot [\text{Value}]_A$$

In reality, the bases of both sides may not be completely specified, so an approximate transformation can also be learned from event histories as follows.

$$\widehat{P}_{A\to B} = \arg\min_{P} \sum_{s=1}^{T} \|\Delta_B(\pi_s, t_s) - P\Delta_A(\pi_s, t_s)\|^2$$

This does not mean that CVTD is a metaphor. It means that, formally, it defines an exact coordinate transformation, and in real application it estimates unknown bases. Through this, A's demand that weekend work is impossible, a protection of time value, can be translated and conveyed within B's basis as risk management for productivity decline, a protection of strategic assets. This is a mathematical engine that increases mutual understanding and approaches the goal of GU, omniscient empathy($EERI \to 1$).

12.7.3 SVD and the Discovery of Latent Desire

If the event matrix $\Pi(\pi, t)$ or the event-history matrix $M_A(T)$ is decomposed by singular value decomposition(SVD, $\Pi(\pi, t) = U\Sigma V^T$), latent dimensions hidden behind the explicit event structure can be discovered. For example, when the interaction history of A and B is decomposed, a high correlation may be discovered on a fourth axis beyond the explicit value axes(k_1, k_2, k_3). This may be an "aesthetic taste" or a "trust cost" implicitly shared by the two people though not written in the contract. An advanced AI system can capture this latent dimension and suggest to B that A is not merely a developer but a partner who shares your aesthetic vision, thereby inducing higher-dimensional intentional events and cooperation(Positive Sum).

13 TBDM: From Individual Utility Pressure to Collective Drive

If CVTD is a grammar that microscopically records all intentional events π, the next question is how those countless microscopic events appear as macroscopic forces at the collective level. The auxiliary model needed here is TBDM. The starting point of this model is simple. A collective has no independent will. When a nation, corporation, party, market, ethnicity, or religion is said to want something, what actually becomes conscious, desires, judges, and acts is always atomic individuals. What appears to be the will of a collective is a macroscopic pattern that occurs when individuals align in the same direction under particular utility pressures.

TBDM holds that individual utility pressure is generally expressed on three basis axes.

Basis	Core question	Representative medium	State of failure and fulfillment
Possibility	What can I do?	Money, resources, time, safety, technology, freedom, power, access	From helplessness, poverty, and danger to options, safety, and executable capacity
Sociality	What kind of being am I among others?	Honor, status, reputation, love, belonging, trust, rights	From exclusion, shame, and isolation to recognition, belonging, and respect
Coherence	What may I believe and live by?	Truth, meaning, justification, belief, worldview, conscience	From confusion, nihilism, and self-betrayal to meaning, conviction, and truthfulness

In short, possibility is the pressure that "I must be able to do," sociality is the pressure that "I must be accepted," and coherence is the pressure that "it must make sense." Money is the representative medium of possibility, honor is the representative medium of sociality, and truth or justification is the representative medium of coherence. The more these three axes point in the same direction, the more strongly the individual is aligned. When belonging to a collective is beneficial, being recognized within that collective is dignity, and the world spoken by that collective feels like truth, the individual internalizes the direction of the collective not as an external command but as their own will.

This connects with CVTD as follows. For each agent n, let there be a projection operator R_3 that maps the event-specific value change $\Delta_n(\pi, t)$ into the three-basis space $\mathcal{B}_3 = \mathrm{span}(\mathbf{e}_{pos}, \mathbf{e}_{soc}, \mathbf{e}_{coh})$ of possibility, sociality, and coherence.

$$\mathbf{f}_n(\pi, t) = R_3 \Delta_n(\pi, t) = \begin{pmatrix} f_n^{pos}(\pi, t) \\ f_n^{soc}(\pi, t) \\ f_n^{coh}(\pi, t) \end{pmatrix}$$

Here $\mathbf{f}_n(\pi, t)$ is the tri-basis utility pressure that event π applies to individual n. f^{pos} means the pressure of possibility, f^{soc} means the pressure of sociality, and f^{coh} means the pressure of coherence. The net force at the collective level can be written as follows.

$$\mathbf{F}(t) = \sum_{\pi \in \mathcal{P}_t^{\mathrm{int}}} \sum_{n=1}^{N} w_n(t) \lambda_\pi(t) \mathbf{f}_n(\pi, t)$$

Here $w_n(t)$ is a weight reflecting each individual's influence, responsibility, exposure, vulnerability, or representativeness, and $\lambda_\pi(t)$ is the execution intensity of the event. The directionality of the collective does not come from an independent collective spirit. It is the direction of $\mathbf{F}(t)$ that appears when the tri-basis pressures of atomic individuals are averaged, canceled, and amplified in a particular direction.

The design goal of this model is not to suppress individual intentions in order to forcibly create a collective good. Rather, it is to design a protocol that allows the countless microscopic

value transfers and intentional events arising among individuals to pass through, while evil forces cancel one another internally and only good net forces remain macroscopically. For this purpose, the Third Protocol must reveal which axes of possibility, sociality, and coherence each purpose declaration strengthens or damages; GU must evaluate the direction of those forces; and CVTD must calculate how actual events accumulate as tri-basis pressures.

Take a corporation as an example. The reason a corporation appears to want profit and growth is not that the corporation is a living organism that desires. It is because its members align in the same direction under possibility pressures such as salary, promotion, equity, and employment security; sociality pressures such as title, performance evaluation, and industry reputation; and coherence pressures such as mission, innovation, and customer value. Political parties, states, ethnic groups, and markets are the same. The will spoken in the name of a collective is the macroscopic shadow made by the tri-basis utility alignment of individuals.

Therefore, when analyzing any collective phenomenon, TBDM asks three questions. In the dimension of possibility, who gains and loses what? In the dimension of sociality, who is recognized and who is excluded? In the dimension of coherence, what language of justification and truth legitimizes action? Passing through these three questions decomposes what appeared to be the will of a collective into individual-level utility pressures. Humanity's driving force is not the will of the collective, but the tri-basis utility alignment of atomic individuals.

14 From Protocol to Real-World Wedge

If the discussion so far has established a theory of the human, a social theory, a normative theory, and an execution theory, the final question is the practical entry point. Purpose declarations do not become abundant immediately merely because they are important. Human beings are not entities that already possess completed sentences of purpose and then input them into a platform. Purpose is made through the slow process of reading, thinking, objection, revision, dialogue, and approval. Therefore, the first real interface of the Third Protocol should not be a market that immediately trades purposes, but an infrastructure that forms purposes.

From this section onward, The Channel is positioned not as the completed Third Protocol itself, but as the first wedge toward that protocol. Books and long-form texts are the slowest and firmest media that sustain the process of purpose formation before purpose declaration, and The Channel is a practical starting point that seeks to place that process onto an urban interface and the learning structure of agentic intelligence.

15 Purpose-Formation Infrastructure and Declaration

However, in principle, purpose cannot be sufficiently handled by extraction alone. Behavioral logs, click patterns, and consumption records can capture peripheral preferences and immediate reactions well, but they cannot read in the same way what a person lives for, what world they want, and what they are willing to take responsibility for. Purpose is not information that is mostly already complete and hidden in data; it is formed only through the process of reading,

thinking, being challenged, revising, and approving. Therefore, the core bottleneck of the future is not a better recommendation algorithm, but the absence of a field that allows humans to form their purposes and declare them publicly.

This point becomes even more important as AI becomes more sophisticated. TRIBE v2, released by Meta FAIR on March 26, 2026, is a tri-modal foundation model that predicts human fMRI brain responses to video, audio, and language stimuli, and Meta describes it as close to a digital twin of human neural activity.[2] What this case shows is clear. Future systems will be able to predict human responses more and more deeply and precisely. But predicting brain responses and having a purpose that a human can approve in their own name are entirely different matters. The fact that the brain responded to a certain video is important data, but it is not the constitutional declaration, "I will become this kind of human being." The stronger extraction becomes, the stronger the need for declaration becomes, not weaker.

In this sense, what is needed from the start is not a purpose market, but purpose-formation infrastructure. Within it, humans read, think, discuss, refine purposes, and finally declare them. And that declaration is not a mere utterance, but the starting point of responsibility, attribution, and negotiation. Only after that can similar purposes discover one another, conditional commitments, resource intentions, and forms of contribution combine, and the transition toward purpose organizations and purpose markets become possible. The purpose market does not replace capital, but it is a structure that places not capital but purpose as the principle that first binds people together.

For this process to actually pass into the language of agents, standardization is also necessary. Each individual must be representable in public space through a kind of purpose passport that has highest purposes, subordinate purposes, priorities, forbidden conditions, negotiable ranges, time horizons, resource constraints, revision history, public scope, and agency scope. Here the core of the platform is not to become the best agent itself. What is more important is to define the purpose declaration format and semantic standard layer that any agent built on any model must pass through in order to represent its owner more deeply.

Technically, this standard layer should presuppose interoperability among agents, rather than a closed agent empire. Google's Agent2Agent(A2A) protocol, announced in April 2025, was proposed as an open protocol for agents created by different frameworks or vendors to exchange information safely and coordinate actions.[3] The Channel does not need to become a super-agent that does everything directly. Rather, The Channel's role is to standardize the user's purpose passport and declaration history, and to turn them into a semantic layer that external agents can read and respect through agent interoperability standards such as A2A. When that happens, The Channel becomes not a model but the purpose registry of an agent ecosystem centered on the owner.

[2]Meta AI, "Introducing TRIBE v2: A Predictive Foundation Model Trained to Understand How the Human Brain Processes Complex Stimuli," March 26, 2026, `https://ai.meta.com/blog/tribe-v2-brain-predictive-foundation-model/`. See also Meta AI Research, "A foundation model of vision, audition, and language for in-silico neuroscience," `https://ai.meta.com/research/publications/a-foundation-model-of-vision-audition-and-language-for-in-silico-neuroscience/`.

[3]Google Developers Blog, "Announcing the Agent2Agent Protocol (A2A)," April 9, 2025, `https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/`.

16 Why Reading Is the First Wedge: The Channel

If such purpose-formation infrastructure is necessary, what should its first real interface be? I think the answer lies not in reading in general, but in a slow and firm language space centered on long-form writing and books. Short posts, feeds, and reactive media can stimulate people and capture preferences, but they are too fast and shallow to make them form life directions and long-term purposes. Books, by contrast, make one follow a long context and reorganize one's own thoughts within sentences, logic, and narrative. Purpose is mostly not born within short stimulation. It needs long reading, long thought, and long language. What I anticipate is that as the cost of execution and personalization continues to fall and the total amount of intelligence grows, society as a whole will demand higher-resolution purposes, and as a result the "purpose-demand pressure" of society will inevitably increase. By purpose-demand pressure, I mean the increasing pressure for individuals, organizations, and agentic intelligences to say more clearly exactly what they want and what they will commit to.

The reason books become the first wedge is not simply that books are an old medium. First, books require long attention, so they strengthen self-interpretation more than immediate reaction. Second, books make one follow layers of concepts within a low density of stimulation, so they create not fragmentary taste but a worldview. Third, books remain in the form of sentences, arguments, and narratives, becoming public objects that can be underlined, annotated, challenged, and shared again. Fourth, books connect individual experience to a longer historical and philosophical memory, allowing one to reconstruct one's purpose not as a simple present desire but on top of civilizational questions. Fifth, the traces of reading and conversation around books later become high-quality purpose-formation data that agentic intelligences can use when understanding their owners. Therefore, books are not a content format of the past, but the most stable interface for purpose formation before purpose declaration.

Therefore, The Channel is, in a very practical sense, a book business. But its essence does not lie in distributing books as commodities. It lies in bringing back to the center of civilization the conditions under which humans can think for themselves. A human who does not think for oneself risks being raised, managed, and stimulated within increasingly sophisticated systems of recommendation, administration, stimulation, and delegated judgment, or being pushed toward self-erasure after losing one's own purpose. At this point, books are not a romantic hobby, but the oldest and strongest technology left for humans to form their own purposes by themselves. Before technology became a religion, people left in the long sentences of the humanities, philosophy, literature, religion, history, and science their endless questions and answers about why one should live, what one should love, what one should be ashamed of, and what world one should make. What we have left behind amid infinite competition and instant optimization is precisely that long memory. The Channel is an attempt to connect that long memory back to the present infrastructure of purpose formation.

The position of The Channel is determined precisely here. As purpose-demand pressure rises, society will need not more content, but seeds of discourse capable of producing clearer and more deliberated purposes. Here the book is not a simple medium. It is a long-form canonical layer

and seed of discourse that can be read, underlined, annotated, challenged, reinterpreted, and finally lead to declaration. The moment multiple people begin reorganizing their own language around the same text, the book ceases to be a consumer good and becomes a starting point of purpose formation.

Therefore, The Channel is the first channel infrastructure that places, on the city, the actual conditions under which people can form their purposes and worldviews, and it is the most realistic starting point through which purpose-formation infrastructure moves toward a purpose market. More precisely, what we target here is the sensory nervous system of intelligence distributed to society for the individual. Just as Palantir-like systems occupy something close to the sensory nervous system of AI in the domains of corporations and governments, what The Channel aims for is that same position with respect to the individual. What matters is not distributing many books, but capturing which books and long-form texts cause thought, discussion, and worldview reorganization within people, and how that change later accumulates in a declarable form. The Channel functions as the first wedge that places seeds of discourse on people's life paths, and accumulates the paths through which those seeds move through reading, conversation, and records into purpose formation.

Further, this structure presupposes a dual structure of the human layer and the agentic-intelligence layer. Humans read books, write long texts, discuss, refine purposes, and declare. At the same time, each person's agentic intelligence learns not only those long records, declaration histories, value priorities, and revision processes, but also the traces of what they read and which texts changed their thoughts, so that even when it goes outside the platform it remains in a state of understanding its owner more deeply. The meaning of The Channel here is not merely securing usage time, but serving as the first real interface that converts agentic intelligence, which had optimized peripheral preferences, toward representing deliberated purposes.

Here agentic intelligence becomes not a simple recommendation engine, but something close to a twin that helps constitute the self I can approve in my own name. People do not know well what they want. At the same time, they cannot say that the record of scrolling short videos under dopamine is completely not themselves. It too is part of them. They simply do not want it to represent their whole self. The Channel's twin does not deny peripheral reactions, but places them together with the records of long reading, long thought, declaration, revision, and responsibility. And it helps the user align themselves in a direction in which they can finally say, "This is a self I can approve in my own name." In the end, what humans ultimately want is not simple satisfaction of desire, but to become closer to a self that they can want to be.

In the end, a book here is not a commodity, but human fuel and a seed of discourse, and The Channel is the first channel that lets that seed move among city, human, and agent and grow into purpose declaration.

17 Conclusion

This manuscript began on a philosophical foundation that redefines the self as a "declared purpose" in order to solve the problems of knowledge fragmentation and loss of purpose in

modern society. Based on this, it proposed Generalized Utilitarianism (GU), which presents the direction of convergence for justice, and further sketched Contractual Value Transfer Dynamics (CVTD), which describes all intentional events π, as a proposed computational language for executing this abstract model in the physical world.

The event matrix $\Pi(\pi, t)$ of CVTD is the concrete unit that composes the policy of GU, and the agent's value basis vector $\mathbf{C}$ becomes the structural basis of the empathy coefficient. This integrated framework will become a theoretical foundation for advanced AI systems to function not merely as tools, but as ethical companions that calculate the alignment between human declared purposes and actual intentional events, discover latent connections, and converge society as a whole toward a state of omniscient empathy.

Furthermore, TBDM explains how these microscopic events appear at the collective level as net forces of possibility, sociality, and coherence. The collective has no independent will. What appears to be the will of the collective is a macroscopic pattern that occurs when atomic individuals align in a particular direction under the utility pressures they receive across the three bases. Therefore, the design goal of the Third Protocol is not to erase individual intention, but to let microscopic purpose declarations and value transfers pass through while evil forces cancel internally and good net forces are amplified.

However, the extension of this manuscript does not stop here. In a society where execution costs continue to fall and the total amount of intelligence continues to grow, votes and prices alone cannot sufficiently aggregate human purposes, and the emergence of the Third Protocol becomes increasingly functionally required. At this point, the core bottleneck is not extracting more behavioral data, but the absence of infrastructure that lets humans form and publicly declare their own purposes. Purpose-formation infrastructure should be understood not as a simple community or recommendation engine, but as a layer that converts human deliberation into social expression and allows agentic intelligence to learn that expression more deeply.

In this context, The Channel is positioned not as the completed purpose market, but as the first practical channel toward that market. Books and long-form texts are the slowest and firmest seeds by which humans refine their purposes, and a long-form canonical layer that can lead to declaration. The Channel is the first wedge that places those seeds on an urban interface and an agent-learning structure. It is a book business, but it is not merely a business that sells books. It is purpose-formation infrastructure that allows humans to think for themselves, create twins they can approve in their own names, and make those twins negotiate with the outside world on agent-interoperability standards such as A2A. Therefore, GU is a philosophical coordinate system, CVTD is a proposed language for an execution layer that still leaves great room for development and expansion, and the Third Protocol and purpose-formation infrastructure are the direction in which the two expand into a social structure. This manuscript should be read not as completing this whole flow as a closed system, but as presenting one strong direction for what kind of purpose input and declarative structure future society will need.