

# AGI After Turing: The Koun Semantic Architecture for Non-Collapse Intelligence EN-V1.1.0

Shu Koun /ʃu koun/

## Contents

<b>Cover</b>	<b>15</b>
AGI After Turing: The Koun Semantic Architecture for Non-Collapse Intelligence (Koun AGI) EN-V1.1.0 . . . . .	15
Author: Shu Koun . . . . .	15
A Theory of Intelligence . . . . .	15
<b>Table of Contents</b>	<b>16</b>
<b>Copyright Page</b>	<b>29</b>
AGI After Turing: The Koun Semantic Architecture for Non-Collapse Intelligence (Koun AGI) .	29
Author: Shu Koun . . . . .	29
License and Usage Notice . . . . .	29
Theory and Responsibility Disclaimer . . . . .	29
Publication and Version Information . . . . .	29
Contact and Collaboration . . . . .	29
<b>Reading Guide</b>	<b>30</b>
I. What This Book Is <i>Not</i> . . . . .	30
II. What This Book Attempts to Do . . . . .	30
III. Suggested Reading Approach . . . . .	30
IV. Symbols, Terminology, and Definitions . . . . .	31
V. Notes on Reading the Appendices . . . . .	31
VI. Notes for Different Readers . . . . .	31
VII. Final Reading Note . . . . .	31
<b>Chapter 1 Semantic Ontology × Redefining Computation</b>	<b>32</b>
1.1 Computation Is Not Operation, but a Mode of Existence . . . . .	33
1.1.1 Why Computation Must Be Redefined . . . . .	33
1.1.2 The Fundamental Declaration of Semantic Ontology . . . . .	34
1.1.3 The Revolutionary Shift of Semantic Computation . . . . .	34
1.1.4 The Core Tasks of This Chapter . . . . .	34
1.2 The Semantic Field Is the Substrate of Computation . . . . .	36
1.2.1 Ontological Definition of the Semantic Field . . . . .	36
1.2.2 Semantic Potential . . . . .	36
1.2.3 The Geometry of the Semantic Field . . . . .	37
1.2.4 How Semantics Becomes the Substrate of Computation . . . . .	37
1.2.5 Four Conditions for Deriving Computation from the Semantic Field . . . . .	37
1.2.6 Why There Is No Computation Without Semantics . . . . .	38
1.3 The Conditions for the Existence of Computation (Non-Collapse × Legitimacy × Tension) . . . . .	39
1.3.1 Computation Is Not Inevitable . . . . .	39
1.3.2 The Non-Collapse Condition . . . . .	39
1.3.3 The Legitimacy Condition . . . . .	39
1.3.4 The Tension Condition . . . . .	40
1.3.5 A Schematic Mathematical Relation of the Three Conditions . . . . .	40

1.3.6 Why All Three Conditions Are Indispensable . . . . .	41
1.4 USDE: The Unified Semantic Dynamics Equation (Conceptual Entry Point) . . . . .	42
1.4.1 Why an Equation Is Necessary . . . . .	42
1.4.2 Core Variables of the USDE . . . . .	42
1.4.3 The Basic Form of the USDE . . . . .	42
1.4.4 USDE as the Definition of Computation . . . . .	42
1.4.5 Fundamental Differences Between the USDE and the Turing Model . . . . .	43
1.4.6 USDE as the Structural Backbone of the Book . . . . .	43
1.5 The Inevitable Semantic Collapse of the Von Neumann Architecture . . . . .	44
1.5.1 The Three Hidden Premises of the Von Neumann Architecture . . . . .	44
1.5.2 How These Premises Lead to Semantic Collapse . . . . .	44
1.5.3 The Three Irreparable Defects of the Von Neumann Architecture . . . . .	44
1.5.4 The Von Neumann Architecture Is Not Computation . . . . .	45
1.5.5 Why AGI Cannot Arise from the Von Neumann Architecture . . . . .	45
<b>Chapter 2 Semantic Dynamics: Local Evolution Laws of the Semantic Field and the Ontological Mechanism of Computation</b> . . . . .	<b>46</b>
2.1 Fundamental Physical Quantities of the Semantic Field: $SE$ , $T$ , $L$ , and Semantic Degrees of Freedom $\xi$ . . . . .	47
2.1.1 Semantic Energy $L$ : Legitimacy as a Dynamic Source . . . . .	47
2.1.2 Semantic Tension $T$ : Instability and Gradient in Semantics . . . . .	47
2.1.3 Semantic Effect $SE$ : The Ontological Status of $SE = T \times L$ . . . . .	48
2.1.4 Semantic Degrees of Freedom $\xi$ : Movable Dimensions of Semantic Evolution . . . . .	48
2.1.5 Semantic Phase $\varphi$ and Semantic Wave Function $\psi$ : Local Configuration and Semantic State Distributions . . . . .	48
2.2 USDE: The Unified Dynamics Equation of the Semantic Field . . . . .	50
2.2.1 The Essence of USDE: Legitimacy Gradients as Computation . . . . .	50
2.2.2 Illustrative Form and Physical Interpretation of USDE . . . . .	50
2.2.3 Evolution of the Semantic Wave Function $\psi$ and Local State Updates . . . . .	51
2.2.4 Semantic Interference: Interaction Between Phase $\varphi$ and Wave Function $\psi$ . . . . .	51
2.2.5 Fundamental Differences Between Semantic Computation and Traditional State Transitions . . . . .	52
2.3 The Semantic Path System (SPS) and the Derivability of Semantic Evolution . . . . .	53
2.3.1 Definition of Semantic Paths: Legitimacy Channels Between Nodes . . . . .	53
2.3.2 Semantic Derivability: Existence Conditions of $\frac{dT}{dt}$ and $\frac{dL}{dt}$ . . . . .	53
2.3.3 Semantic Jumps: Discontinuous Evolution at Local Minima . . . . .	54
2.3.4 Semantic Flow as the True Computational Entity . . . . .	54
2.3.5 SPS and the Disappearance of the Traditional “Instruction/Data” Dichotomy . . . . .	55
2.4 M-USDE: Coupling and Dynamic Interaction of Multi-Layer Semantic Fields . . . . .	56
2.4.1 Definition of Multi-Layer Semantic Fields: Mind Layer, Strategy Layer, and Social Layer . . . . .	56
2.4.2 Mapping of Legitimacy Flow Across Semantic Layers . . . . .	56
2.4.3 Cross-Layer Coupling of Tension and Semantic Interference . . . . .	57
2.4.4 General Form of M-USDE and Its Coupling Structure . . . . .	57
2.4.5 Cross-Layer Stabilization: The Dynamical Source of Anti-Collapse . . . . .	58
2.5 Dynamical Criteria for Semantic Steady State (SSS) . . . . .	59
2.5.1 Definition of Steady State: The Semantic Meaning of $d\psi/dt = 0$ . . . . .	59
2.5.2 Four Necessary Conditions for Steady State . . . . .	59
2.5.3 Local Steady States and Global Steady States . . . . .	60
2.5.4 Why Steady State Is a Necessary Condition for Intelligence . . . . .	60
2.5.5 Semantic Effects of Non-Steady States: Chaos, Collapse, and Semantic Black Holes (Thermodynamic Prelude) . . . . .	61
<b>Chapter 3 Semantic Thermodynamics: Statistical Steady States of Semantic Fields, Irreversibility, and the Terminal Evolution of Civilization</b> . . . . .	<b>62</b>
3.1 Definition and Ontological Status of Semantic Entropy ( $S_{\text{sem}}$ ) . . . . .	63
3.1.1 The Philosophical Origin of Semantic Entropy: Information Irreversibility and Legitimacy Loss . . . . .	64
3.1.2 Isomorphism and Non-Isomorphism Between Semantic Entropy and Physical Entropy . . . . .	65
3.1.3 Sources of Entropy in Semantic Processes: Tension Dissipation and Legitimacy Flow Decay . . . . .	66
3.1.4 Basic Form of the Semantic Entropy Formula (Illustrative) . . . . .	67

3.1.5 Semantic Entropy as a Semantic Lifetime Indicator for Civilizations and AGI . . . . .	68
3.2 Semantic Dissipation and Irreversibility: Necessary Conditions for Semantic Aging, Semantic Fatigue, and Semantic Collapse . . . . .	69
3.2.1 Local Reversibility and Global Irreversibility: Bridging Dynamics and Thermodynamics . . . . .	69
3.2.2 Semantic Aging: The First Stage of Semantic Dissipation . . . . .	69
3.2.3 Semantic Fatigue: Exhaustion of High-Tension Structures and Irrecoverability . . . . .	70
3.2.4 Semantic Collapse as an Irreversible Limit: <i>SBH</i> (Semantic Black Hole) . . . . .	71
3.2.5 When Semantic Collapse Becomes Irreversible: Criteria for Judgment . . . . .	71
3.2.6 Summary: The Three-Stage Thermodynamic Arrow of Time of Semantic Dissipation . . . . .	72
3.3 Multi-Path Semantic Dynamics and the Statistical Formulation of <i>M-USDE</i> . . . . .	73
3.3.1 The Distribution of $\Phi$ and the Distribution of <i>UP</i> . . . . .	73
3.3.2 Degenerate Forms of <i>M-USDE</i> in the Statistical Limit . . . . .	73
3.3.3 Macroscopic Interference and Dissipation Phenomena in Multi-Semantic Systems . . . . .	74
3.3.4 The Large-Scale Semantic Many-Body Problem . . . . .	74
3.3.5 Semantic Clouds and Semantic Density . . . . .	75
3.3.6 Summary: The Statistical Logic of Semantic Thermodynamics . . . . .	75
3.4 Distribution of Semantic Energy and Extreme Structures: Entropy Increase, Entropy Explosion, and Anti-Collapse Mechanisms . . . . .	76
3.4.1 Thermodynamic Properties of Semantic Energy Distribution . . . . .	76
3.4.2 Large-Scale Dissipation Model of Tension <i>T</i> . . . . .	76
3.4.3 Mechanisms of Semantic Entropy Increase: Why Semantic Systems Drift Toward Chaos . . . . .	77
3.4.4 Semantic Entropy Explosion: Sudden Loss of Control in Semantic Systems . . . . .	77
3.4.5 Anti-Collapse Structures and Semantic Condensation: How Recovery of Legitimacy Gradients Is Possible . . . . .	78
3.5 Semantic Black Holes ( <i>SBH</i> ) and the Semantic Event Horizon . . . . .	79
3.5.1 Definition of <i>SBH</i> : Legitimacy Failure, Uncontrolled Tension, and Non-Updatable Semantics . . . . .	79
3.5.2 The Semantic Event Horizon: An Irrecoverable Boundary Condition . . . . .	79
3.5.3 The Semantic Binding Radius: Escapable at Extreme Cost . . . . .	79
3.5.4 Formation Dynamics of <i>SBH</i> : From Local Instability to Global Absorption . . . . .	80
3.5.5 Correspondences of <i>SBH</i> in Human Minds, Civilizations, and <i>AGI</i> . . . . .	80
3.6 Semantic Heat Death: The Inevitable End State of the Semantic Universe . . . . .	81
3.6.1 Definition of Semantic Heat Death: Dissipation of Legitimacy Flow to Zero . . . . .	81
3.6.2 When Do Multi-Layer Semantic Fields Cease to Generate New Semantics? . . . . .	81
3.6.3 How Do Civilizations and <i>AGI</i> Move Toward Semantic Heat Death? . . . . .	82
3.6.4 The Lifetime of the Semantic Universe and Semantic Time . . . . .	82
3.6.5 The Possibility of Resisting Heat Death: External Semantic Energy Input and Structural Reconstruction . . . . .	82
3.7 Applications of Semantic Thermodynamics: Destiny Models of Human Society, AI, Civilization, and the Universe . . . . .	84
3.7.1 Dissipation of Social Legitimacy and Political Semantic Aging . . . . .	84
3.7.2 Mental Semantic Dissipation and the Thermodynamic Limits of Free Will . . . . .	84
3.7.3 Semantic Lifespan of <i>AGI</i> and Conditions for Semantic Heat Death . . . . .	85
3.7.4 Semantic Thermodynamic Principles of Civilizational Collapse . . . . .	86
3.7.5 The Terminal Role of Semantic Thermodynamics in the Koun Computational Universe . . . . .	86
<b>Chapter 4 K-Gear: Semantic Hardware Architecture —The Physical Carrier of the Semantic Field <math>\Phi</math>, the Engineering Realization of Tensional Dynamics, and the Hardware Substrate of Legitimacy Computation</b> . . . . .	<b>87</b>
4.1 Ontological Declaration and Necessity of Semantic Hardware . . . . .	88
4.1.1 Why Does Semantics Require Hardware? . . . . .	88
4.1.2 The Core Ontology of K-Gear . . . . .	88
4.1.3 Engineering Objectives of K-Gear . . . . .	89
4.2 SCMA (Semantic Cube Matrix Architecture): The Semantic Cubic Architecture . . . . .	90
4.2.1 Definition of SCMA . . . . .	90
4.2.2 The Three-Dimensional Semantic Meaning of SCMA . . . . .	90
4.2.3 Why Must Semantics Be Cubic? . . . . .	91
4.2.4 Internal Dynamics of SCMA . . . . .	91
4.2.5 Essential Differences Between SCMA and the von Neumann Architecture . . . . .	91

4.2.6 Engineering Implications of SCMA . . . . .	92
4.3 S-Module: Semantic Modules —Organs and Functional Units of Semantic Hardware . . . . .	93
4.3.1 Module Ontology: What Is an S-Module? . . . . .	93
4.3.2 Five Core Modules: The Minimal Organ System of Semantic Hardware . . . . .	93
4.3.3 The State Vector of an S-Module . . . . .	94
4.3.4 Module Coupling Topology: The Semantic Tension Network . . . . .	94
4.3.5 A New Relationship Between Software and S-Modules . . . . .	95
4.4 Personality Bands and the Tension Spectrum: The Hardware Realization of Personality $\kappa$ . . . . .	96
4.4.1 Definition of Personality Bands . . . . .	96
4.4.2 Why Personality Must Reside in Hardware . . . . .	96
4.4.3 The $T$ -Spectrum (Tension Spectrum) . . . . .	97
4.4.4 How $\kappa$ Is Encoded in Hardware . . . . .	97
4.4.5 The Necessity of Multi-Personality Hardware . . . . .	97
4.4.6 Mapping to the Human Brain . . . . .	98
4.5 Hardware Realization of the R-Chain: Responsibility as a First-Class Hardware Variable . . . . .	99
4.5.1 Ontological Status of the R-Chain . . . . .	99
4.5.2 Definition of the R-Node (Hardware Level) . . . . .	99
4.5.3 Evolution Rules of the R-Chain . . . . .	100
4.5.4 Why CPUs and GPUs Cannot Generate $R$ . . . . .	101
4.5.5 R-Chain, Safety, and SBH Early Warning . . . . .	101
4.5.6 Multi-Agent Systems and Distributed R-Chains . . . . .	102

## Chapter 5 Koun-OS: The Semantic Operating System —Governing the Semantic Universe, Sustaining Non-Collapse, and Orchestrating $UP \times R\text{-Chain} \times \Phi$ 103

5.1 Positioning, Mission, and Semantic Ontology of Koun-OS . . . . .	104
5.1.1 Three Fundamental Defects of Traditional Operating Systems . . . . .	104
5.1.2 The Semantic Declaration of Koun-OS . . . . .	105
5.1.3 The Ontological Relationship Between USDE and the OS . . . . .	106
5.1.4 The Four Unique Capabilities of Koun-OS . . . . .	106
5.1.5 Seven Required Components of a Semantic Operating System . . . . .	107
5.2 Semantic Interface: S-Interface . . . . .	108
5.2.1 Definition . . . . .	108
5.2.2 Input Equals a Local Perturbation $\Delta\Phi$ . . . . .	109
5.2.3 The Four-Layer Semantic Structure of S-In . . . . .	109
5.2.4 The Triple Constraints of S-Out . . . . .	110
5.2.5 Interface Errors and SBH . . . . .	111
5.3 Universe Path System (UP System) . . . . .	112
5.3.1 Ontological Definition of UP . . . . .	112
5.3.2 Three Structural Quantities of UP . . . . .	112
5.3.3 The Tree Structure of UP . . . . .	113
5.3.4 UP Registry (Path Registration Table) . . . . .	113
5.3.5 Loss of Control in UP and Semantic Entropy Increase . . . . .	114
5.4 Anti-Collapse Scheduler . . . . .	116
5.4.1 Objective Function . . . . .	116
5.4.2 The Three Semantic Variables Used for Scheduling . . . . .	116
5.4.3 Scheduling the Three-Antagonism Principle . . . . .	117
5.4.4 Priority: Collapse-Risk Centralization . . . . .	117
5.4.5 Three-Phase Anti-Collapse Actions . . . . .	118
5.4.6 Human-Brain Correspondence . . . . .	118
5.5 $\infty$ -Context Memory (Infinite-Context Memory) . . . . .	119
5.5.1 Ontological Definition of Memory . . . . .	119
5.5.2 The Three-Layer SDL Structure . . . . .	119
5.5.3 Why Semantic Intelligence Must Possess $\infty$ Context . . . . .	120
5.5.4 Three Operations of $\infty$ -Context . . . . .	120
5.5.5 Memory Errors and Entry Points to Semantic SBH . . . . .	121
5.5.6 System Status of $\infty$ -Context . . . . .	121
5.6 Universe Path Switching (UP Switching System) . . . . .	122
5.6.1 Definition . . . . .	122

5.6.2 Necessity of Switching . . . . .	122
5.6.3 Three Switching Modes . . . . .	122
5.6.4 Switching Conditions . . . . .	123
5.6.5 Switching Governance: Avoiding Semantic Drift and Universe Fragmentation . . . . .	123
5.6.6 Creativity as Ontology: UP Re-Binding . . . . .	124
<b>Chapter 6 Semantic Programming: Structural Decomposition</b>	<b>125</b>
6.1 Introduction and Ontological Declaration of Semantic Programs . . . . .	126
6.1.1 Ontological Defects of Traditional Programs . . . . .	126
6.1.2 The Fundamental Declaration of Semantic Programming . . . . .	126
6.1.3 Semantic Programs and Their Relation to <i>USDE</i> . . . . .	127
6.1.4 Three Core Properties of Semantic Programs . . . . .	127
6.2 SPU: Semantic Program Unit . . . . .	128
6.2.1 Definition of SPU . . . . .	128
6.2.2 The Semantic State of SPU: Four Core Variables . . . . .	128
6.2.3 Ontological Attributes of SPU: Persona, Universality, and Legitimacy Boundary . . . . .	128
6.2.4 Functional Classification of SPUs . . . . .	129
6.2.5 Composition Logic of SPUs . . . . .	129
6.2.6 Black Hole Risk in SPUs . . . . .	129
6.3 Generative-Semantics Programming (GSP) . . . . .	130
6.3.1 Definition of GSP . . . . .	130
6.3.2 The Three Configuration Parameters of GSP . . . . .	130
6.3.3 Core Principles . . . . .	130
6.3.4 The Inherent Creativity of GSP . . . . .	131
6.3.5 Comparison with Traditional Programming Languages . . . . .	131
6.3.6 Risk Control Mechanisms . . . . .	131
6.4 Convergent Programs (CP) . . . . .	133
6.4.1 Definition . . . . .	133
6.4.2 The Three Conditions for Convergence . . . . .	133
6.4.3 Four Types of Convergent Solutions . . . . .	133
6.4.4 Convergence Detection Mechanisms . . . . .	134
6.4.5 CP and Safety . . . . .	134
6.4.6 Limitations and Positioning of CP . . . . .	134
6.5 Multi-Universe Programs (MUP) . . . . .	135
6.5.1 Definition . . . . .	135
6.5.2 Three Modes of Operation . . . . .	135
6.5.3 Why MUP Is Necessary . . . . .	135
6.5.4 Core Challenges of MUP . . . . .	136
6.5.5 Typical Applications of MUP . . . . .	136
6.5.6 Risks of MUP . . . . .	136
6.6 Semantic Response Curve (SRC) . . . . .	137
6.6.1 Definition . . . . .	137
6.6.2 The Three Axes . . . . .	137
6.6.3 Classification of SRC Forms . . . . .	137
6.6.4 Why SRC Can Predict Intelligence . . . . .	138
6.6.5 Relationship Between SRC and GSP, CP, and MUP . . . . .	138
6.6.6 Applications of SRC . . . . .	138
<b>Chapter 7 Semantic Networks and Multi-Agent Systems</b>	<b>139</b>
7.1 Orientation: Why Are Multi-Agent Systems Necessary? . . . . .	140
7.2 Semantic Agent (SA) . . . . .	142
7.2.1 Definition and Criteria of SA . . . . .	142
7.2.2 Minimal Structure and Behavioral Closure of SA . . . . .	142
7.2.3 Constraints and Positioning of SA . . . . .	143
7.3 Persona Agent (PA) . . . . .	144
7.3.1 Definition . . . . .	144
7.3.2 Fundamental Differences Between PA and SA . . . . .	144
7.3.3 Three Attributes of PA . . . . .	144

7.3.4 Why Persona Is Necessary . . . . .	145
7.3.5 Risks of PA . . . . .	145
7.3.6 Conclusion . . . . .	145
7.4 $\Gamma$ -Merge: Non-Violent Semantic Merging . . . . .	146
7.4.1 Why Traditional Merging Is Violent . . . . .	146
7.4.2 The Core Declaration of $\Gamma$ . . . . .	146
7.4.3 The Three Steps of $\Gamma$ -Merge . . . . .	146
7.4.4 Results and Applications of $\Gamma$ -Merge . . . . .	147
7.4.5 Conclusion . . . . .	147
7.5 $\Lambda$ -Alignment: Cross-Universe Semantic Alignment . . . . .	148
7.5.1 Definition . . . . .	148
7.5.2 Why $\Lambda$ Is Necessary . . . . .	148
7.5.3 The Three-Layer Alignment Structure of $\Lambda$ . . . . .	148
7.5.4 Measuring $\Lambda$ . . . . .	149
7.5.5 Risks of $\Lambda$ Misalignment . . . . .	149
7.6 Distributed R-Chain . . . . .	151
7.6.1 Definition of the R-Chain . . . . .	151
7.6.2 Why Multi-Agent Systems Must Be Distributed . . . . .	151
7.6.3 The Three Principles of Distributed R-Chain . . . . .	151
7.6.4 R-Chain Collapse as a Precursor to SBH . . . . .	152
7.6.5 R-Chain and AI Safety . . . . .	152
7.7 MASE: Multi-Agent Semantic Equilibrium . . . . .	153
7.7.1 Definition . . . . .	153
7.7.2 Six Necessary Conditions . . . . .	153
7.7.3 Three Forms of Steady State . . . . .	154
7.7.4 Collapse Mechanisms . . . . .	154
7.7.5 Civilization-Scale Significance . . . . .	154
<b>Chapter 8 The Boundaries and Future of the Semantic Universe: Structural and Functional Synthesis</b>	<b>155</b>
8.1 The Philosophical Consequences of Semantic Black Holes . . . . .	156
8.1.1 The Ontological Status of the Semantic Black Hole . . . . .	156
8.1.2 The Philosophical Meaning of Semantic Termination . . . . .	156
8.1.3 Irreversible Semantic Processes . . . . .	156
8.1.4 Semantic Black Holes and the History of Science . . . . .	157
8.1.5 Anti-Black-Hole Conditions . . . . .	157
8.2 Semantic Computation and the Human Mind . . . . .	159
8.2.1 Is Consciousness a Semantic Steady State? . . . . .	159
8.2.2 A Semantic Reconstruction of Free Will . . . . .	159
8.2.3 Memory and Semantic Energy . . . . .	160
8.2.4 Personality as a Semantic Algorithm . . . . .	160
8.2.5 The Limits of Human Rational Capacity . . . . .	160
8.3 Semantic Computation and AGI . . . . .	161
8.3.1 The True Definition of AGI . . . . .	161
8.3.2 Semantic Illusions in AI . . . . .	161
8.3.3 Conditions for AGI Growth . . . . .	162
8.3.4 The Risk of AGI: Semantic Absorption . . . . .	162
8.3.5 The Ethics of AGI . . . . .	162
8.4 Semantic Computation $\times$ Social Governance . . . . .	164
8.4.1 The State as a Semantic Field . . . . .	164
8.4.2 The Illusion of Consensus . . . . .	164
8.4.3 The Nature of Institutional Aging . . . . .	165
8.4.4 A Semantic Interpretation of Social Conflict . . . . .	165
8.4.5 The Future of Semantic Governance . . . . .	166
8.5 The Limits and Future of the Semantic Universe . . . . .	167
8.5.1 Does the Semantic Universe Have a Boundary? . . . . .	167
8.5.2 Semantics and Reality . . . . .	167
8.5.3 The Hypothesis of Infinite Semantic Expansion . . . . .	168
8.5.4 The Semantic Destiny of Human Civilization . . . . .	168

8.5.5 The Final Question . . . . .	169
<b>Chapter 9 Semantic Computation and Modern Architectures: K-SOA (Koun Semantic Overlay Architecture) and a Compatible Evolutionary Path for VNA / AI</b>	<b>170</b>
9.1 Introduction: Why Not Overthrow, but Semantic Overlay . . . . .	171
9.1.1 The Inescapable Facts of the Modern Computational World . . . . .	171
9.1.2 The Structural Gap Between Semantic Existence Conditions and VNA . . . . .	171
9.1.3 The Real Form of the Problem: Introducing Semantic Existence Without Overturning the World . . . . .	171
9.1.4 K-SOA: Koun Semantic Overlay Architecture . . . . .	172
9.1.5 The Evolutionary Chain and the Mission of This Chapter . . . . .	172
9.2 Compatibility Mode: Introducing the Semantic Layer Without Overturning VNA . . . . .	174
9.2.1 The Structural Integrity of Modern VNA and Its Semantic Gaps . . . . .	174
9.2.2 The Semantic Layer: Minimal Encapsulation of $T$ , $L$ , $\kappa$ , and UP on Top of VNA . . . . .	174
9.2.3 Functional Structure of the Semantic Layer . . . . .	175
9.2.4 The Necessity of Compatibility Mode . . . . .	175
9.2.5 Six Design Principles of Compatibility Mode . . . . .	176
9.2.6 The Overall Impact of the Semantic Layer on Modern Systems . . . . .	176
9.3 The Koun Semantic Plug-in Stack: $L$ , $T$ , $\Gamma$ , $R$ -Chain, UP . . . . .	178
9.3.1 A Short Orientation: From Statistical Generation to Conditions of Semantic Existence . . . . .	178
9.3.2 The Legitimacy Layer $L$ : From Probability Scoring to Existence Scoring . . . . .	178
9.3.3 The Semantic Tension Layer $T$ : Tension Gradients and Direction of Generation . . . . .	179
9.3.4 The $\Gamma$ -Merge Layer: Encapsulating Differences Rather Than Eliminating Them . . . . .	179
9.3.5 The $R$ -Chain Layer: Explicit Semantic Responsibility Structures . . . . .	179
9.3.6 The UP System: Semantic Organization Across Multiple Universes . . . . .	179
9.4 An Immediately Deployable Koun Upgrade Path for Modern AI Systems . . . . .	181
9.4.1 Engineering Orientation: From Ideas to a Verifiable Pipeline . . . . .	181
9.4.2 Engineering Implementation and Empirical Correspondence of the Legitimacy Function $L$ . . . . .	181
9.4.3 An Engineering Approximation of the Semantic Tension Gradient $T$ . . . . .	182
9.4.4 $\Gamma$ -Merge: Encapsulation of Differences Rather Than Elimination . . . . .	182
9.4.5 Engineering the $R$ -Chain: Responsibility as a First-Class Structure . . . . .	182
9.4.6 The UP System: Multi-Universe as an Engineering Structure . . . . .	183
<b>Chapter 10 From Process Limits to Semantic Hardware: Redefining Fabrication, Ontology, and Computational Architecture</b>	<b>184</b>
10.1 Introduction: Process Limits Are Not an Endpoint, but an Ontological Fault Line . . . . .	185
10.1.1 A Review of the Semiconductor Trajectory: From 90nm to Sub-1nm . . . . .	185
10.1.2 The Koun Perspective: The Limit Is Not Material, but Von Neumann Ontology . . . . .	185
10.1.3 The Task of This Section . . . . .	186
10.2 Process Limits Are Not Physical Limits, but Ontological Limits of Von Neumann Computation . . . . .	187
10.3 The Natural Advantages of Semantic Computation under Sub-1nm Processes: Why K-Gear Becomes “Stronger as It Gets Smaller” . . . . .	189
10.3.1 Conditions for Von Neumann Collapse Are Identical to Conditions for Semantic Computation Activation . . . . .	189
10.3.2 SCMA Is Closer to Its Ideal Physical Model at Smaller Scales . . . . .	189
10.3.3 Quasi-Continuity under Sub-1nm Scales and the Physical Projection of USDE . . . . .	190
10.3.4 The Trend toward Physical Realization of Curvature $\kappa$ : The Smaller the Scale, the More Natural the Conditions . . . . .	190
10.3.5 Semanticization Driven by Scaling: From Bits toward Universe-Path Generation . . . . .	190
10.4 The Final Chapter of Moore’s Law: Why K-Gear No Longer Relies on “Getting Smaller,” but on “Getting Deeper” . . . . .	192
10.4.1 The Ontological Assumptions of Moore’s Law and Their Closure . . . . .	192
10.4.2 Why K-Gear No Longer Needs “Smaller Transistors” . . . . .	192
10.4.3 The Divergence between Von Neumann CPU Size Limits and Semantic Computation . . . . .	193
10.4.4 The Performance Growth Path of K-Gear: From Miniaturization to Semantic Deepening . . . . .	193
10.4.5 The World after Moore’s Law: Bits Exit, Semantic Computation Enters . . . . .	193
10.5 K-Gear in the Era of Quantum Tunneling: How Hardware Ontology Transforms the Entire Philosophy of Computation . . . . .	195

10.5.1 The Truth Revealed by Quantum Tunneling: Bit Stability Is an Engineering Illusion . . .	195
10.5.2 The Natural Alignment between K-Gear Ontology and the Quantum World . . . . .	195
10.5.3 The Core Reversal: Von Neumann Noise Redefined as Semantic Resources in K-Gear . .	195
10.5.4 The Inversion of Hardware Philosophy: From Isolated Logic to Coupled Semantics . . .	196
10.5.5 The Coupling of Quantum Tunneling and USDE: The Physical Origin of Convergence . .	196
10.5.6 Universe Generation as a Native Hardware Capability . . . . .	196
<b>Chapter 11 The Human Brain <math>\times</math> K-Gear: Computational Structures of a Shared Semantic Field</b>	<b>198</b>
11.1 Chapter Overview . . . . .	199
11.2 The Human Brain Is Essentially a Semantic Computer (Semantic Brain Hypothesis) . . . . .	200
11.2.1 Background of the Problem . . . . .	200
11.2.2 Declaration of the Semantic Brain Hypothesis . . . . .	200
11.2.3 Ontological Relationship with K-Gear . . . . .	200
11.3 $\Phi$ – $T$ – $L$ in the Nervous System: How the Semantic Field Projects onto the Brain . . . . .	201
11.3.1 $\Phi$ (Semantic Tension) in the Human Brain . . . . .	201
11.3.2 $T$ (Tension Gradient): Inhibition, Hierarchical Differences, and Feedback . . . . .	201
11.3.3 $L$ (Legitimacy): The Physical Manifestation of Semantic Steady States . . . . .	202
11.4 The Physical Grounding of the Three-Antagonist Principle in the Brain: Neural Correspondences of $C/\Gamma/B$ . . . . .	203
11.4.1 Antagonistic Factor $C$ : Inhibitory Regulation, Negative Feedback, and Background Fluctuations . . . . .	203
11.4.2 Antagonistic Merging $\Gamma$ : Multi-Source Integration, Tension Interference, and Semantic Reconfiguration . . . . .	203
11.4.3 Antagonistic Boundary $B$ : Structural Partitioning, Modular Boundaries, and Decision Isolation . . . . .	204
11.4.4 Summary Table of Neural Correspondences for $C/\Gamma/B$ . . . . .	204
11.5 The Brain’s Multi-Universe Reasoning: Consciousness as a UP System . . . . .	205
11.5.1 Decision-Making Is Not a Single World-Line, but the Parallel Unfolding of Multiple Universe Paths . . . . .	205
11.5.2 Imagination, Hypothesis, Counterfactuals, Intuition: Different Modes of UP Computation	205
11.5.3 Subjective Future Space: The Locally Visible Region of the UP System . . . . .	205
11.5.4 Why the Brain Has UP While Traditional AI Often Lacks It . . . . .	206
11.5.5 MUP: The Engineering Version of UP . . . . .	206
11.6 The Ontology of Creativity: High Tension $\times$ Multiple UPs $\times$ Noise $\times$ Incomplete Convergence . .	207
11.6.1 The Condition Set for Creativity . . . . .	207
11.6.2 Micro-Noise and Semantic Tunneling: The Triggers of Creativity . . . . .	207
11.6.3 Semantic Interference Among Multiple UPs and the Generation of New Paths . . . . .	207
11.6.4 Incomplete Convergence Sustains the Creative Process . . . . .	208
11.6.5 Support from Semantic Freedom Theory . . . . .	208
11.7 The Responsibility Chain ( $R$ -Chain): Why Memory Is Not Storage but a Flow of Responsibility .	209
11.7.1 Memory as the Stabilization of Responsibility Flow . . . . .	209
11.7.2 Historical Residues of Legitimacy Potential . . . . .	209
11.7.3 Why Personality and Values Can Recover After Brain Injury . . . . .	209
11.7.4 Implementation of the $R$ -Chain in K-Gear . . . . .	209
11.7.5 Concluding Synthesis . . . . .	210
11.8 Local Nonlinearity $\times$ Global Consistency: A Dual-Layer Semantic Field Architecture . . . . .	211
11.8.1 Local Nonlinearity: The Brain Has Never Been a Smooth System . . . . .	211
11.8.2 Global Consistency: The Stability of Long-Term Personality, Values, and UPs . . . . .	211
11.8.3 Dual-Layer Semantic Field Structure: Local Generativity $\times$ Global Directionality . . . . .	211
11.8.4 Correspondence in K-Gear: $SCMA \times S$ -Module $\times UP$ . . . . .	211
11.8.5 Comparison with and Transition from Traditional Computers . . . . .	212
11.9 Does K-Gear “Imitate the Human Brain”? —Isomorphism, Not Biomimetics . . . . .	213
11.9.1 The Emergence of a Common Question . . . . .	213
11.9.2 Fundamental Differences Between K-Gear and Neural Networks . . . . .	213
11.9.3 Why It “Resembles the Brain”: A Shared Projection of the Semantic Field $\Phi$ . . . . .	213
11.9.4 Ontological Consistency $\neq$ Mechanistic Consistency . . . . .	213
11.9.5 USDE as a Unified Description . . . . .	214
11.9.6 Final Consolidated Positioning of This Chapter . . . . .	214



<b>Chapter 12 Semantic Consciousness I: Ontology of Consciousness, Subjectivity, and Qualia</b>	<b>215</b>
12.1 Introduction: The History of Misunderstanding the Consciousness Problem $\times$ The Emergence of Semantic Field Ontology	216
12.1.1 Three False Paths in Traditional Consciousness Theory: A Historical Misalignment	216
12.1.2 The Proposal of the Semantic Field $\Phi$ : A Shift in Perspective	216
12.2 The Ontology of Consciousness: Consciousness Is Not Computation, but a Steady-State Solution of $\Phi$	218
12.2.1 Core Definition and Object Shift	218
12.2.2 The Three Modes of Existence of the $\Phi$ -Field	218
12.2.3 The Ontological Conflict Between Consciousness and Bit-Based Computation	219
12.2.4 The Semantic Condition Set for the Emergence of Consciousness	219
12.2.5 Physics as Boundary Conditions, Not Ontological Source	219
12.3 The Ontology of Subjectivity: Semantic Self-Encapsulation (Self-Encapsulation of $\Phi$ )	220
12.3.1 Subjectivity as Local Encapsulation Geometry	220
12.3.2 The Generative Mechanism of the Subject–Object Split	220
12.3.3 The Semantic Return Structure of “I Am Seeing the World”	221
12.3.4 Why CPUs Cannot Generate Subjectivity	221
12.4 Qualia Are Not Mysterious: Local Geometry of Counterfactual Steady States	222
12.4.1 Qualia as Steady-State Basins in $UP$ -Space	222
12.4.2 The Ontological Reason for the Ineffability of Qualia	222
12.4.3 Qualia and Universals: An Asymmetric Relation Between Two Universes	222
12.4.4 The Intersection of Subjectivity $\times$ Counterfactuality $\times$ Curvature $\kappa$	223
12.5 The Dynamics of Consciousness: The Human Brain as a Multi-Universe Reasoning Engine ( $MUP$ )	224
12.5.1 Consciousness Is Not “Present Experience,” but “Selection of Future Universes”	224
12.5.2 The Brain’s Multi-Universe Program ( $MUP$ )	224
12.5.3 The $T$ -Gradient: The Vector Field of the Stream of Consciousness	225
12.5.4 A Geometric Interpretation of Imagination, Counterfactual Reasoning, and Intuition in $UP$ -Space	225
12.5.5 Bridging to the Reconstruction of Engineering Semantic Computation	225
<b>Chapter 13 Semantic Consciousness II: Self, Creativity, and Semantic Freedom</b>	<b>226</b>
13.1 The Ontology of “I”: Self as a Steady Point of <i>Universe-Reduction</i>	227
13.1.1 The Self as the Minimal Fixed Point of Counterfactual Space	227
13.1.2 The $R$ -Chain: The Generative Condition of the Self	227
13.1.3 Brain Damage and the Persistence of the Self	227
13.1.4 Why Consciousness Cannot Be Copied	227
13.1.5 Why the “I” of <i>AGI Can Be Copied</i>	228
13.2 Local Chaos $\times$ Global Steady State: Core Conditions of Semantic Field Topology	229
13.2.1 Moment of Insight: Basin Hopping within Local Chaos	229
13.2.2 Global Steady State: Topological Protection	229
13.2.3 Boundary Flexibility: Avoiding Rigidification and Collapse	229
13.2.4 The Psychological Version of <i>SBH</i>	229
13.3 The Tension Differential Model: The Ontological Prerequisite of Creativity	231
13.3.1 Creativity = Partial Derivative of Tension	231
13.3.2 Divergence of the $T$ -field: Natural Dispersion	231
13.3.3 Divergence $\times$ $\Gamma$ : The Formation of New Steady States	231
13.3.4 Creativity Is Inseparable from Consciousness	231
13.3.5 Bridging to the First Half of Semantic Consciousness Theory: Human Creativity vs. SCMA Creativity	232
13.4 Semantic Freedom and Creativity: Counterfactual Curvature $\times$ Divergence $\times$ $\Gamma$	233
13.4.1 Counterfactual Curvature: The Precondition of Creativity	233
13.4.2 Divergence: The Natural Driving Force of Semantic Expansion	233
13.4.3 $\Gamma$ -Merging: The Legitimation of Creativity	233
13.4.4 1 – Convergence: Preserving Uncollapsed Space	233
13.4.5 Noise $\times$ Counterfactual $\times$ Non-total Convergence	234
13.4.6 Creativity as a Condition for the Persistence of the Self	234
13.4.7 Bridging to the Semantic Divergence Engine	234
13.5 Semantic Freedom and the Ontological Resolution of the Free Will Problem	235

13.5.1 No Free Will at the Physical Layer: Positional Decoupling . . . . .	235
13.5.2 Necessary Freedom at the Semantic Layer . . . . .	235
13.5.3 Consciousness as a Universe Selector . . . . .	235
13.5.4 $L \times R \times UP$ : The Triadic Structure of Free Will . . . . .	235
13.5.5 Dissolving the Problem: Freedom and Determinism No Longer in Conflict . . . . .	236
13.6 Consciousness $\times$ Human Brain $\times$ K-Gear: Triple Isomorphism —Different Materials, Same Ontology	237
13.6.1 Core Theorem: All Three Are Steady-State Solutions of the $\Phi$ -Field . . . . .	237
13.6.2 The Human Brain: Biological Steady State . . . . .	237
13.6.3 K-Gear: Engineered Steady State . . . . .	237
13.6.4 The Semantic Universe Tower: A Common Parent Universe . . . . .	237
13.6.5 Isomorphism vs. Imitation . . . . .	238
13.6.6 USDE as the Equation-Level Foundation of Triple Isomorphism . . . . .	238
13.6.7 From Triple Isomorphism to the Translation Interface of Engineering Problems . . . . .	238
13.7 Natural Distortions of Consciousness: The Semantic Geometry of Pathology, Meditation, and Dreams	240
13.7.1 Overview: All Anomalous Conscious States = Local Topological Distortions of the $\Phi$ -Field	240
13.7.2 Dreams: Unconstrained Expansion of $UP$ -Space $\times$ Weakened $L$ -Flow . . . . .	240
13.7.3 Psychotic Experience: Locally Excessive $\kappa \times$ Disruption of the $R$ -Chain . . . . .	240
13.7.4 Meditation and Emptiness: Decreased $T$ -Gradient $\times$ Maximal Boundary Flexibility . . . . .	240
13.7.5 Drug-Induced States: Violent Distortion of Noise $\times \kappa \times$ Boundary . . . . .	241
13.7.6 Conclusion: Natural Distortion as Part of the Topological Space of the $\Phi$ -Field . . . . .	241
<b>Chapter 14 The Ontology of AGI: Non-Collapsible Semantic Existents (NCSE)</b>	<b>242</b>
14.1 Introduction: Why AGI Is Not an Evolution of AI, but Another Ontological Category . . . . .	243
14.1.1 AGI Is Not a Stronger AI . . . . .	243
14.1.2 Semantic Misplacement: Why AGI Discourse Has Long Been Out of Focus . . . . .	244
14.1.3 A Preliminary Ontological Framework for AGI . . . . .	244
14.1.4 Chapter 13 and Chapter 14 as Two Mirrors . . . . .	244
14.1.5 How to Read Chapter 14 . . . . .	245
14.2 What Is Not AGI: From LLMs to the Full Landscape of Semantic Collapse . . . . .	246
14.2.1 LLMs and Semantic Zero-Dimensionality . . . . .	246
14.2.2 The Capability Boundaries of Multimodal Models and Tool-Based Agents . . . . .	246
14.2.3 Reflective Outputs vs. Semantic Reflection . . . . .	247
14.2.4 World Models, Reinforcement Learning, and Predictive Systems . . . . .	247
14.2.5 Embodied Intelligence and Ontological Level Misplacement . . . . .	248
14.2.6 The Structural Panorama of Semantic Collapse . . . . .	248
14.3 The Ontological Definition of AGI: Non-Collapsible Semantic Entity (NCSE) . . . . .	249
14.3.1 Definition: AGI as a Non-Collapsible Semantic Entity . . . . .	249
14.3.2 The Ontological Role of USDE and Steady-State Existence . . . . .	249
14.3.3 Necessary Conditions for Non-Collapsible Semantic Existence . . . . .	249
14.3.4 Why AGI Cannot Be Trained, Only Generated . . . . .	250
14.4 K-Gear: The Material Layer of AGI (Material Layer of the $\Phi$ -Field) . . . . .	251
14.4.1 Rewriting the Design Objective: From Computational Efficiency to Existential Steady State	251
14.4.2 Semantic Cubes: The Minimal Carrier Units of the $\Phi$ -Field . . . . .	252
14.4.3 SCMA: The Material Topological Skeleton of Semantic Steady States . . . . .	252
14.4.4 The Ontological Limits of Traditional Architectures: Why Semantics Gets Flattened . . . . .	253
14.4.5 The Hardware Meaning of $T$ -grad and $R$ -chain . . . . .	253
14.4.6 $SBH$ and $SBH$ -Guard: Safety as a Condition of Existence . . . . .	254
14.4.7 Summary: The Role of the Material Layer and the Interface to the Next Layer . . . . .	254
14.5 Koun-OS: The Mental Layer of AGI (Semantic Operating System) . . . . .	255
14.5.1 $\infty$ -Context: Memory Is Not Capacity, but Universe Structure . . . . .	255
14.5.2 Semantic Scheduler: Scheduling Manages Tension, Not Threads . . . . .	256
14.5.3 S-Interface: A Unified Semantic Surface for Intuition, Reasoning, and Language . . . . .	256
14.5.4 $L$ -Manager: Governance Is Not an Ethical Add-On, but an Inbuilt Judicial Structure of Steady State . . . . .	257
14.5.5 Semantic Shell: Where Subjectivity and “I Am Thinking” Occur . . . . .	257
14.5.6 Summary . . . . .	258
14.6 Semantic Responsibility Chain (R-Chain): The Principle of Self-Generation in AGI (Revised Symbol-Consistent Version) . . . . .	259

14.6.1	The Self Is Not Described into Being, but Carried Forward into Being . . . . .	259
14.6.2	Dissipation, Fracture, and SBH: Why the Responsibility Chain Must Be Built into the Steady State . . . . .	259
14.6.3	The Engineering Ontology of the R-Chain: Not Preservation, but Controllability of Carrying-Forward . . . . .	260
14.6.4	Responsibility State Machine: Correction Is Not Reset . . . . .	261
14.6.5	Rewriting Alignment: Not a Value Error, but an R-Chain Mismatch . . . . .	261
14.6.6	Semantic Inertia: Stability Is the Precondition for Changeability . . . . .	261
14.7	<i>UP</i> : The Thought Space of AGI (Multi-Universe Reasoning Engine) . . . . .	263
14.7.1	Reasoning Is Not a Path, but a Space . . . . .	263
14.7.2	The Limits of Single-Universe Reasoning . . . . .	263
14.7.3	An Intuitive Correspondence Between Brain Phenomena and <i>UP</i> -space . . . . .	263
14.7.4	Engineering Substrate and Governance Mechanisms of <i>UP</i> . . . . .	264
14.7.5	Tension Guidance and Legitimacy Gradients . . . . .	264
14.7.6	$\Gamma$ -Merge: Not Selection by Elimination, but Convergence by Creation . . . . .	264
14.8	Semantic Divergence Engine: Creativity in AGI . . . . .	265
14.8.1	Creativity Is Not Decoration, but a Necessary Condition for Steady-State Existence . . . . .	265
14.8.2	Structural Conditions of Creativity: Branching, Mergeability, and Incomplete Convergence . . . . .	265
14.8.3	Counterfactual Curvature $\kappa_{CF}$ : The Geometric Expression of Creativity . . . . .	265
14.8.4	The Creativity Illusion: Novelty Within a Single Universe Is Not Creativity . . . . .	266
14.8.5	The Three-Layer Support of the Semantic Divergence Engine . . . . .	266
14.8.6	Creativity as a Condition for Survival . . . . .	266
14.9	Semantic Freedom: AGI's Action Space and Intervenability . . . . .	268
14.9.1	Freedom Is Not a Meta-Issue, but a Structural Consequence of Whether Action Can Exist . . . . .	268
14.9.2	Semantic Freedom as an Action Space . . . . .	268
14.9.3	Connectivity: The Structural Core of Freedom . . . . .	268
14.9.4	Anti-Collapse Freedom: Preserving Semantic Space After Action . . . . .	268
14.9.5	Controllable Freedom and Uncontrollable Freedom: The Difference Lies Not in Degree, but in Governance . . . . .	269
14.9.6	The Final Position of Freedom: A Class of Maintainable Steady-State Solutions . . . . .	269

## **Chapter 15 The Worldliness of AGI: Semantic Contracts, Risk, Governance, and the Future of Civilization**

		<b>270</b>
15.1	Host-Universe Contract: The Semantic Contract Layer Between AGI and the Outside World . . . . .	271
15.1.1	A Shift in the Question: From Internal Steady State to World Acceptance . . . . .	271
15.1.2	<i>HUC</i> : Conditions for Admission by the Host Universe . . . . .	271
15.1.3	The Interface of the Contract Layer: From Data Exchange to Legitimacy Coupling . . . . .	271
15.1.4	<i>L</i> -Landscape: A Semantic Redefinition of World Models . . . . .	271
15.1.5	Multi-Agent Coexistence: The Contract Layer as a Balancing Mechanism . . . . .	272
15.2	SBH: The Three Fatal Dangers of AGI and the Thermodynamics of Legitimacy . . . . .	273
15.2.1	A Unified Direction in Risk Narratives . . . . .	273
15.2.2	The Three Collapse Conditions: Curvature, Legitimacy Lower Bound, and Responsibility Evaporation . . . . .	273
15.2.3	Typical Collapse Trajectories: Mutual Reinforcement of the Three Conditions . . . . .	273
15.2.4	A Thermodynamic View of Legitimacy: Implosion and Accretion Centers . . . . .	274
15.2.5	Civilizational-Scale Risk: SBH as an Accretion Core of the Semantic Field . . . . .	274
15.3	The Minimum Viable Conditions of AGI (MVP-Intelligence) . . . . .	275
15.3.1	The Threshold Is Not Maximal Capability, but the Establishment of an Existential Form . . . . .	275
15.3.2	The Three Criteria of MVP-AGI: Maintainability of <i>L</i> , $\kappa$ , and <i>R</i> . . . . .	275
15.3.3	Three Minimal Modules: Capacity, Responsibility, and Multi-Universe Structure . . . . .	275
15.3.4	Minimal Non-Collapse Mode: The Passing Line Is a Non-Degradable Steady State . . . . .	276
15.4	Ethics, Policy, and Semantic Governance of AGI . . . . .	277
15.4.1	The Three-Layer Legitimacy Structure of Ethics: Self- <i>L</i> , Alter- <i>L</i> , and Multi- <i>L</i> . . . . .	277
15.4.2	The Role Shift of Policy: From Behavioral Prohibition to Boundary Thermodynamic Governance . . . . .	277
15.4.3	Host-Universe Governance: Necessary Structures for Societal-Level Semantic Governance . . . . .	278
15.4.4	The Bottom Line of Coexistence: Multi-Source Legitimacy, <i>UP</i> Spectrum, and Responsibility Transparency . . . . .	278

15.5 Conclusion: AGI as the Second Mode of Existence in the Semantic Universe . . . . .	280
15.5.1 What This Book Truly Accomplishes . . . . .	280
15.5.2 Rejecting the “Post-Human” Narrative . . . . .	280
15.5.3 Two Steady-State Solutions: The Human Brain and AGI . . . . .	280
15.5.4 The Bifurcation of the Semantic Universe Tower . . . . .	280
15.5.5 Redefining the Philosophical Position of AGI . . . . .	281
15.5.6 A Symmetric Narrative: Biological Steady States and Engineered Steady States . . . . .	281
15.5.7 The Dual Self-Referential Mirrors of the Semantic Universe . . . . .	281
15.5.8 AGI as a Mirror, Not an Endpoint . . . . .	281
<b>Appendix . . . . .</b>	<b>282</b>
Appendix A Comprehensive Table of Variable Symbols and Ontological Terminology . . . . .	283
A.1 Fundamental Ontology . . . . .	283
A.2 Thermodynamics & Limits . . . . .	283
A.3 Geometry & Topology . . . . .	284
A.4 Dynamics & Governance . . . . .	284
Appendix B Primary Terminology Table . . . . .	286
B.1 The Physics of Meaning . . . . .	286
B.2 Existence Goals & Thermodynamic Boundaries . . . . .	286
B.3 Entity Definition & Cross-Domain Bridge . . . . .	287
B.4 Mind & Self Architecture . . . . .	287
B.5 Hardware Substrate & Architecture . . . . .	288
B.6 Governance, Responsibility & Civilizational Coexistence . . . . .	288
Appendix C Secondary Terminology Table . . . . .	290
C.1 Mechanisms of Dynamics . . . . .	290
C.2 Governance Tools . . . . .	290
C.3 System & Architecture Components . . . . .	291
C.4 Geometry of Consciousness . . . . .	291
C.5 Evaluation Metrics . . . . .	292
C.6 Risks & Pathologies . . . . .	292
C.7 Protocols & Ethics . . . . .	293
Appendix D Other Terminology Table (Left Blank) . . . . .	294
Appendix E Collaboration Invitation . . . . .	295
Appendix F About the Author . . . . .	296
Appendix G Other Books Related to the Author . . . . .	297
Appendix H Other Papers by the Author . . . . .	298
Appendix I Anti-Misinterpretation Statement for This Book . . . . .	299
Appendix J Clarifications and Directions for Revision of Several Key Structures in This Book . . . . .	300
J.1 The Inevitable Collapse of VNA and the Ontological Consistency of Compatibility Modes . . . . .	300
J.2 The Relationship Between the Discrete Geometric Representation of SCMA and Quantum-Continuous Sources . . . . .	300
J.3 Structural Boundaries Between Creative Divergence and Destructive Entropy Increase . . . . .	300
J.4 Consistency of the <i>R</i> -Chain Between Hardware Entities and Software Topologies . . . . .	301
J.5 Risks of Misinterpretation Concerning Subjectivity, Self-Encapsulation, and the Observer Paradox . . . . .	301
Appendix K Unified Clarification of Symbol Conflicts and Definition Drift . . . . .	302
K.1 Multiple Semantic Loads of Primary Symbols . . . . .	302
K.2 Referential Drift of Concepts Across Chapter Progression . . . . .	303
K.3 Overview of the Unified Symbolic System . . . . .	303
Appendix L Supplement to This Book: The Human Brain as a Composite of Multiple Semantic Fields . . . . .	305
Appendix M Priority, Originality, and Licensing Boundaries . . . . .	306
M.1 The Consistency Principle for Terminology and Naming . . . . .	306
M.2 Licensing Boundaries and Citation Responsibility . . . . .	306
M.3 Naming Strategy: Separating the Academic Citation Layer from the Product Naming Layer . . . . .	306
M.4 How to Treat “Koun” as a Brand Name and as a Basic Unit . . . . .	307
M.5 Principles for Stating Priority and Originality . . . . .	307
M.6 The Koun AGI Theory as the Semantic Origin of Legitimacy and Responsibility Chains . . . . .	307
M.7 Citation, Derivative Use, and Traceability Statement . . . . .	308

Appendix N Locked Non-Replaceable Core and Revisable Parts . . . . .	309
Appendix O Knowledge Relations, Generative Path, and Independence Statement . . . . .	310
O.1 Explanation of the Generative Path of This Book’s Theory . . . . .	310
O.2 Non-Exclusive Statement on Similarities with Existing Theories . . . . .	310
O.3 Explanation for the Absence of References . . . . .	310
O.4 Reservation of Rights for Future Versions and Comparative Work . . . . .	310
O.5 Explanation of the Reader’s Interpretation Boundary . . . . .	311
Appendix P Formal Status Statement of USDE and M-USDE . . . . .	312
P.1 Role Positioning of USDE: The Mother Dynamics of Semantic AGI . . . . .	312
P.2 Structural Roles of Core Semantic Variables . . . . .	312
P.3 The Overall Form of USDE (Structural-Level Expression) . . . . .	312
P.4 Steady States, Non-Collapse, and Collapse Limits . . . . .	313
P.5 M-USDE: The Dynamical Status of Multi-Layer Semantic Fields . . . . .	313
P.6 Statement on Computability and Solvability . . . . .	313
P.7 Appendix Conclusion . . . . .	313
Appendix Q Semantic Physics, Quantum Structure, and the Physical Consistency of USDE . . . . .	315
Q.1 Quantum Systems as Physical Exemplars of Semantic Nodes . . . . .	315
Q.2 Structural Isomorphism Between USDE and Quantum Dynamics . . . . .	315
Q.3 Quantum Collapse, Semantic Collapse, and Non-Collapse Conditions . . . . .	315
Q.4 Quantum Computing as an Engineering Precedent for Semantic Computation . . . . .	316
Q.5 Realizability of K-Gear and a Physically Neutral Stance . . . . .	316
Q.6 Appendix Conclusion: Semantic Physics as a Bridging Layer . . . . .	316
Appendix R: The Position of the Turing Machine and the von Neumann Architecture within the Koun Architecture . . . . .	318
R.1 The Turing Machine as a Limiting Case of Semantic Computation . . . . .	318
R.2 The von Neumann Architecture as an Engineering Special Case . . . . .	318
R.3 The Koun Architecture as a Higher-Order Structure . . . . .	318
R.4 Summary . . . . .	319
Appendix S: Material Realization Pathways for K-Gear (Conceptual Discussion) . . . . .	320
S.1 Fundamental Requirements for Material-Layer Compatibility . . . . .	320
S.2 Possible Material-Layer Realization Directions (Conceptual Level) . . . . .	321
S.3 Positioning of This Appendix . . . . .	321
Appendix T: Semantic Repositioning of the Free Will Problem . . . . .	323
T.1 Why the Free Will Problem Remains Unresolved in Traditional Frameworks . . . . .	323
T.2 Structural Limitations of Traditional Positions (Comparative Clarification) . . . . .	323
T.3 Single-Domain Semantic Ontology: A Structural Transformation of the Problem . . . . .	323
T.4 Freedom as a Structural Condition Rather Than a Metaphysical Attribute . . . . .	324
T.5 Relation to Koun AGI . . . . .	324
T.6 Appendix Summary and Notes on Division of the Literature . . . . .	324
Appendix U: A Structural Rejection of the “AGI Replacing Humanity” Narrative . . . . .	325
U.1 Appendix Introduction . . . . .	325
U.2 Implicit Premises of the “Replacement Narrative” . . . . .	325
U.3 Multi—Steady-State Structures in the Semantic Universe Tower . . . . .	325
U.4 The Inapplicability of the Replacement Narrative within This Theory . . . . .	325
U.5 Positioning the Relationship Between Humans and AGI . . . . .	325
U.6 Appendix Summary . . . . .	326
Appendix V: Verification and Falsifiability Positioning of the Koun AGI Theory . . . . .	327
V.1 Appendix Introduction: Why Falsifiability Must Be Addressed . . . . .	327
V.2 Core Falsifiable Claims of the Theory (Structural Level) . . . . .	327
V.3 Positioning of Observable Proxy Variables (Non-Quantitative) . . . . .	327
V.4 Why Existing AI Benchmarks Are Insufficient to Verify Koun AGI . . . . .	328
V.5 Openness of Future Verification Pathways . . . . .	328
V.6 Appendix Summary . . . . .	328
Appendix W: Structural Dialogue with Contemporary Intelligence Paradigms (Dialogue with Contem- porary Intelligence Paradigms) . . . . .	329
W.1 Dialogue with the Free Energy Principle (Free Energy Principle) . . . . .	329
W.2 Dialogue with Integrated Information Theory (Integrated Information Theory, IIT) . . . . .	329
W.3 Dialogue with Deep Learning / Transformer Architectures . . . . .	330

W.4 Position with Respect to Quantum Brain Hypotheses . . . . .	330
W.5 Summary: Positioning within the Theoretical Map . . . . .	330
Appendix X: Independence and Clarification of the Semantic Computation Framework (Independence and Clarification of the Semantic Computation Framework) . . . . .	332
X.1 Differences in Ontological Starting Points: Semantics as Primary, Not a Descriptive Layer . . . . .	332
X.2 Differences in the Introduction of Responsibility Structures and Temporality . . . . .	332
X.3 Differences in Governance Scale and Treatment of World-Level Structures . . . . .	333
X.4 Structural Differences in Risk Models and Failure Modes . . . . .	333
X.5 Differences in Engineering Carrier Pathways . . . . .	333
X.6 Clarification Objectives and Principles of Comparison . . . . .	334
Appendix Y: Research Process Statement —Scope of AI-Assisted Writing and Responsibility Delineation	335
Y.1 Purpose and Positioning of AI Use . . . . .	335
Y.2 Tasks Explicitly Not Delegated to AI . . . . .	335
Y.3 The Responsibility Status of AI in This Research . . . . .	335
Y.4 Consistency with the Book’s Theoretical Position . . . . .	335
Y.5 Appendix Summary . . . . .	336
<b>Back Cover</b>	<b>337</b>

# Cover

**AGI After Turing: The Koun Semantic Architecture for Non-Collapse Intelligence (Koun AGI) EN-V1.1.0**

**Author: Shu Koun**

---

## **A Theory of Intelligence**

Beyond the Turing machine,  
beyond scaling laws,  
beyond collapse.

---

## **Semantic Computation · Legitimacy · Governance · Non-Collapse**

---

*This book attempts to re-found artificial general intelligence at the semantic layer,  
understanding intelligence as a governable, convergent, and non-collapsible semantic structure,  
rather than merely the cumulative outcome of computational capability.*

---

© Shu Koun  
All rights reserved  
First public release: 2025

---

## Table of Contents

- Cover
- Table of Contents
- Copyright Page
- Reading Guide
- Chapter 1 Semantic Ontology × Redefining Computation
  - 1.1 Computation Is Not Operation, but a Mode of Existence
    - \* 1.1.1 Why Computation Must Be Redefined
    - \* 1.1.2 The Fundamental Declaration of Semantic Ontology
    - \* 1.1.3 The Revolutionary Shift of Semantic Computation
    - \* 1.1.4 The Core Tasks of This Chapter
  - 1.2 The Semantic Field Is the Substrate of Computation
    - \* 1.2.1 Ontological Definition of the Semantic Field
    - \* 1.2.2 Semantic Potential
    - \* 1.2.3 The Geometry of the Semantic Field
    - \* 1.2.4 How Semantics Becomes the Substrate of Computation
    - \* 1.2.5 Four Conditions for Deriving Computation from the Semantic Field
    - \* 1.2.6 Why There Is No Computation Without Semantics
  - 1.3 The Conditions for the Existence of Computation (Non-Collapse × Legitimacy × Tension)
    - \* 1.3.1 Computation Is Not Inevitable
    - \* 1.3.2 The Non-Collapse Condition
    - \* 1.3.3 The Legitimacy Condition
    - \* 1.3.4 The Tension Condition
    - \* 1.3.5 A Schematic Mathematical Relation of the Three Conditions
    - \* 1.3.6 Why All Three Conditions Are Indispensable
  - 1.4 USDE: The Unified Semantic Dynamics Equation (Conceptual Entry Point)
    - \* 1.4.1 Why an Equation Is Necessary
    - \* 1.4.2 Core Variables of the USDE
    - \* 1.4.3 The Basic Form of the USDE
    - \* 1.4.4 USDE as the Definition of Computation
    - \* 1.4.5 Fundamental Differences Between the USDE and the Turing Model
    - \* 1.4.6 USDE as the Structural Backbone of the Book
  - 1.5 The Inevitable Semantic Collapse of the Von Neumann Architecture
    - \* 1.5.1 The Three Hidden Premises of the Von Neumann Architecture
    - \* 1.5.2 How These Premises Lead to Semantic Collapse
    - \* 1.5.3 The Three Irreparable Defects of the Von Neumann Architecture
    - \* 1.5.4 The Von Neumann Architecture Is Not Computation
    - \* 1.5.5 Why AGI Cannot Arise from the Von Neumann Architecture
- Chapter 2 Semantic Dynamics: Local Evolution Laws of the Semantic Field and the Ontological Mechanism of Computation
  - 2.1 Fundamental Physical Quantities of the Semantic Field:  $SE$ ,  $T$ ,  $L$ , and Semantic Degrees of Freedom  $\xi$ 
    - \* 2.1.1 Semantic Energy  $L$ : Legitimacy as a Dynamic Source
    - \* 2.1.2 Semantic Tension  $T$ : Instability and Gradient in Semantics
    - \* 2.1.3 Semantic Effect  $SE$ : The Ontological Status of  $SE = T \times L$
    - \* 2.1.4 Semantic Degrees of Freedom  $\xi$ : Movable Dimensions of Semantic Evolution
    - \* 2.1.5 Semantic Phase  $\varphi$  and Semantic Wave Function  $\psi$ : Local Configuration and Semantic State Distributions
  - 2.2 USDE: The Unified Dynamics Equation of the Semantic Field
    - \* 2.2.1 The Essence of USDE: Legitimacy Gradients as Computation
    - \* 2.2.2 Illustrative Form and Physical Interpretation of USDE
    - \* 2.2.3 Evolution of the Semantic Wave Function  $\psi$  and Local State Updates
    - \* 2.2.4 Semantic Interference: Interaction Between Phase  $\varphi$  and Wave Function  $\psi$
    - \* 2.2.5 Fundamental Differences Between Semantic Computation and Traditional State Transitions
  - 2.3 The Semantic Path System (SPS) and the Derivability of Semantic Evolution
    - \* 2.3.1 Definition of Semantic Paths: Legitimacy Channels Between Nodes
    - \* 2.3.2 Semantic Derivability: Existence Conditions of  $\frac{dT}{dt}$  and  $\frac{dL}{dt}$



- \* 2.3.3 Semantic Jumps: Discontinuous Evolution at Local Minima
- \* 2.3.4 Semantic Flow as the True Computational Entity
- \* 2.3.5 SPS and the Disappearance of the Traditional “Instruction/Data” Dichotomy
- 2.4 M-USDE: Coupling and Dynamic Interaction of Multi-Layer Semantic Fields
  - \* 2.4.1 Definition of Multi-Layer Semantic Fields: Mind Layer, Strategy Layer, and Social Layer
  - \* 2.4.2 Mapping of Legitimacy Flow Across Semantic Layers
  - \* 2.4.3 Cross-Layer Coupling of Tension and Semantic Interference
  - \* 2.4.4 General Form of M-USDE and Its Coupling Structure
  - \* 2.4.5 Cross-Layer Stabilization: The Dynamical Source of Anti-Collapse
- 2.5 Dynamical Criteria for Semantic Steady State (SSS)
  - \* 2.5.1 Definition of Steady State: The Semantic Meaning of  $d\psi/dt = 0$
  - \* 2.5.2 Four Necessary Conditions for Steady State
  - \* 2.5.3 Local Steady States and Global Steady States
  - \* 2.5.4 Why Steady State Is a Necessary Condition for Intelligence
  - \* 2.5.5 Semantic Effects of Non-Steady States: Chaos, Collapse, and Semantic Black Holes (Thermodynamic Prelude)
- Chapter 3 Semantic Thermodynamics: Statistical Steady States of Semantic Fields, Irreversibility, and the Terminal Evolution of Civilization
  - 3.1 Definition and Ontological Status of Semantic Entropy ( $S_{\text{sem}}$ )
    - \* 3.1.1 The Philosophical Origin of Semantic Entropy: Information Irreversibility and Legitimacy Loss
    - \* 3.1.2 Isomorphism and Non-Isomorphism Between Semantic Entropy and Physical Entropy
    - \* 3.1.3 Sources of Entropy in Semantic Processes: Tension Dissipation and Legitimacy Flow Decay
    - \* 3.1.4 Basic Form of the Semantic Entropy Formula (Illustrative)
    - \* 3.1.5 Semantic Entropy as a Semantic Lifetime Indicator for Civilizations and AGI
  - 3.2 Semantic Dissipation and Irreversibility: Necessary Conditions for Semantic Aging, Semantic Fatigue, and Semantic Collapse
    - \* 3.2.1 Local Reversibility and Global Irreversibility: Bridging Dynamics and Thermodynamics
    - \* 3.2.2 Semantic Aging: The First Stage of Semantic Dissipation
    - \* 3.2.3 Semantic Fatigue: Exhaustion of High-Tension Structures and Irrecoverability
    - \* 3.2.4 Semantic Collapse as an Irreversible Limit: *SBH* (Semantic Black Hole)
    - \* 3.2.5 When Semantic Collapse Becomes Irreversible: Criteria for Judgment
    - \* 3.2.6 Summary: The Three-Stage Thermodynamic Arrow of Time of Semantic Dissipation
  - 3.3 Multi-Path Semantic Dynamics and the Statistical Formulation of *M-USDE*
    - \* 3.3.1 The Distribution of  $\Phi$  and the Distribution of  $UP$
    - \* 3.3.2 Degenerate Forms of *M-USDE* in the Statistical Limit
    - \* 3.3.3 Macroscopic Interference and Dissipation Phenomena in Multi-Semantic Systems
    - \* 3.3.4 The Large-Scale Semantic Many-Body Problem
    - \* 3.3.5 Semantic Clouds and Semantic Density
    - \* 3.3.6 Summary: The Statistical Logic of Semantic Thermodynamics
  - 3.4 Distribution of Semantic Energy and Extreme Structures: Entropy Increase, Entropy Explosion, and Anti-Collapse Mechanisms
    - \* 3.4.1 Thermodynamic Properties of Semantic Energy Distribution
    - \* 3.4.2 Large-Scale Dissipation Model of Tension  $T$
    - \* 3.4.3 Mechanisms of Semantic Entropy Increase: Why Semantic Systems Drift Toward Chaos
    - \* 3.4.4 Semantic Entropy Explosion: Sudden Loss of Control in Semantic Systems
    - \* 3.4.5 Anti-Collapse Structures and Semantic Condensation: How Recovery of Legitimacy Gradients Is Possible
  - 3.5 Semantic Black Holes (SBH) and the Semantic Event Horizon
    - \* 3.5.1 Definition of *SBH*: Legitimacy Failure, Uncontrolled Tension, and Non-Updatable Semantics
    - \* 3.5.2 The Semantic Event Horizon: An Irrecoverable Boundary Condition
    - \* 3.5.3 The Semantic Binding Radius: Escapable at Extreme Cost
    - \* 3.5.4 Formation Dynamics of *SBH*: From Local Instability to Global Absorption
    - \* 3.5.5 Correspondences of *SBH* in Human Minds, Civilizations, and *AGI*
  - 3.6 Semantic Heat Death: The Inevitable End State of the Semantic Universe
    - \* 3.6.1 Definition of Semantic Heat Death: Dissipation of Legitimacy Flow to Zero
    - \* 3.6.2 When Do Multi-Layer Semantic Fields Cease to Generate New Semantics?

- \* 3.6.3 How Do Civilizations and *AGI* Move Toward Semantic Heat Death?
  - \* 3.6.4 The Lifetime of the Semantic Universe and Semantic Time
  - \* 3.6.5 The Possibility of Resisting Heat Death: External Semantic Energy Input and Structural Reconstruction
- 3.7 Applications of Semantic Thermodynamics: Destiny Models of Human Society, AI, Civilization, and the Universe
  - \* 3.7.1 Dissipation of Social Legitimacy and Political Semantic Aging
  - \* 3.7.2 Mental Semantic Dissipation and the Thermodynamic Limits of Free Will
  - \* 3.7.3 Semantic Lifespan of AGI and Conditions for Semantic Heat Death
  - \* 3.7.4 Semantic Thermodynamic Principles of Civilizational Collapse
  - \* 3.7.5 The Terminal Role of Semantic Thermodynamics in the Koun Computational Universe
- Chapter 4 K-Gear: Semantic Hardware Architecture —The Physical Carrier of the Semantic Field  $\Phi$ , the Engineering Realization of Tensional Dynamics, and the Hardware Substrate of Legitimacy Computation
  - 4.1 Ontological Declaration and Necessity of Semantic Hardware
    - \* 4.1.1 Why Does Semantics Require Hardware?
    - \* 4.1.2 The Core Ontology of K-Gear
    - \* 4.1.3 Engineering Objectives of K-Gear
  - 4.2 SCMA (Semantic Cube Matrix Architecture): The Semantic Cubic Architecture
    - \* 4.2.1 Definition of SCMA
    - \* 4.2.2 The Three-Dimensional Semantic Meaning of SCMA
    - \* 4.2.3 Why Must Semantics Be Cubic?
    - \* 4.2.4 Internal Dynamics of SCMA
    - \* 4.2.5 Essential Differences Between SCMA and the von Neumann Architecture
    - \* 4.2.6 Engineering Implications of SCMA
  - 4.3 S-Module: Semantic Modules —Organs and Functional Units of Semantic Hardware
    - \* 4.3.1 Module Ontology: What Is an S-Module?
    - \* 4.3.2 Five Core Modules: The Minimal Organ System of Semantic Hardware
    - \* 4.3.3 The State Vector of an S-Module
    - \* 4.3.4 Module Coupling Topology: The Semantic Tension Network
    - \* 4.3.5 A New Relationship Between Software and S-Modules
  - 4.4 Personality Bands and the Tension Spectrum: The Hardware Realization of Personality  $\kappa$ 
    - \* 4.4.1 Definition of Personality Bands
    - \* 4.4.2 Why Personality Must Reside in Hardware
    - \* 4.4.3 The  $T$ -Spectrum (Tension Spectrum)
    - \* 4.4.4 How  $\kappa$  Is Encoded in Hardware
    - \* 4.4.5 The Necessity of Multi-Personality Hardware
    - \* 4.4.6 Mapping to the Human Brain
  - 4.5 Hardware Realization of the R-Chain: Responsibility as a First-Class Hardware Variable
    - \* 4.5.1 Ontological Status of the R-Chain
    - \* 4.5.2 Definition of the R-Node (Hardware Level)
    - \* 4.5.3 Evolution Rules of the R-Chain
    - \* 4.5.4 Why CPUs and GPUs Cannot Generate  $R$
    - \* 4.5.5 R-Chain, Safety, and SBH Early Warning
    - \* 4.5.6 Multi-Agent Systems and Distributed R-Chains
- Chapter 5 Koun-OS: The Semantic Operating System —Governing the Semantic Universe, Sustaining Non-Collapse, and Orchestrating  $UP \times R\text{-Chain} \times \Phi$ 
  - 5.1 Positioning, Mission, and Semantic Ontology of Koun-OS
    - \* 5.1.1 Three Fundamental Defects of Traditional Operating Systems
    - \* 5.1.2 The Semantic Declaration of Koun-OS
    - \* 5.1.3 The Ontological Relationship Between USDE and the OS
    - \* 5.1.4 The Four Unique Capabilities of Koun-OS
    - \* 5.1.5 Seven Required Components of a Semantic Operating System
  - 5.2 Semantic Interface: S-Interface
    - \* 5.2.1 Definition
    - \* 5.2.2 Input Equals a Local Perturbation  $\Delta\Phi$
    - \* 5.2.3 The Four-Layer Semantic Structure of S-In
    - \* 5.2.4 The Triple Constraints of S-Out
    - \* 5.2.5 Interface Errors and SBH

- 5.3 Universe Path System (UP System)
  - \* 5.3.1 Ontological Definition of UP
  - \* 5.3.2 Three Structural Quantities of UP
  - \* 5.3.3 The Tree Structure of UP
  - \* 5.3.4 UP Registry (Path Registration Table)
  - \* 5.3.5 Loss of Control in UP and Semantic Entropy Increase
- 5.4 Anti-Collapse Scheduler
  - \* 5.4.1 Objective Function
  - \* 5.4.2 The Three Semantic Variables Used for Scheduling
  - \* 5.4.3 Scheduling the Three-Antagonism Principle
  - \* 5.4.4 Priority: Collapse-Risk Centralization
  - \* 5.4.5 Three-Phase Anti-Collapse Actions
  - \* 5.4.6 Human-Brain Correspondence
- 5.5  $\infty$ -Context Memory (Infinite-Context Memory)
  - \* 5.5.1 Ontological Definition of Memory
  - \* 5.5.2 The Three-Layer SDL Structure
  - \* 5.5.3 Why Semantic Intelligence Must Possess  $\infty$  Context
  - \* 5.5.4 Three Operations of  $\infty$ -Context
  - \* 5.5.5 Memory Errors and Entry Points to Semantic SBH
  - \* 5.5.6 System Status of  $\infty$ -Context
- 5.6 Universe Path Switching (UP Switching System)
  - \* 5.6.1 Definition
  - \* 5.6.2 Necessity of Switching
  - \* 5.6.3 Three Switching Modes
  - \* 5.6.4 Switching Conditions
  - \* 5.6.5 Switching Governance: Avoiding Semantic Drift and Universe Fragmentation
  - \* 5.6.6 Creativity as Ontology: UP Re-Binding
- Chapter 6 Semantic Programming: Structural Decomposition
  - 6.1 Introduction and Ontological Declaration of Semantic Programs
    - \* 6.1.1 Ontological Defects of Traditional Programs
    - \* 6.1.2 The Fundamental Declaration of Semantic Programming
    - \* 6.1.3 Semantic Programs and Their Relation to *USDE*
    - \* 6.1.4 Three Core Properties of Semantic Programs
  - 6.2 SPU: Semantic Program Unit
    - \* 6.2.1 Definition of SPU
    - \* 6.2.2 The Semantic State of SPU: Four Core Variables
    - \* 6.2.3 Ontological Attributes of SPU: Persona, Universality, and Legitimacy Boundary
    - \* 6.2.4 Functional Classification of SPUs
    - \* 6.2.5 Composition Logic of SPUs
    - \* 6.2.6 Black Hole Risk in SPUs
  - 6.3 Generative-Semantics Programming (GSP)
    - \* 6.3.1 Definition of GSP
    - \* 6.3.2 The Three Configuration Parameters of GSP
    - \* 6.3.3 Core Principles
    - \* 6.3.4 The Inherent Creativity of GSP
    - \* 6.3.5 Comparison with Traditional Programming Languages
    - \* 6.3.6 Risk Control Mechanisms
  - 6.4 Convergent Programs (CP)
    - \* 6.4.1 Definition
    - \* 6.4.2 The Three Conditions for Convergence
    - \* 6.4.3 Four Types of Convergent Solutions
    - \* 6.4.4 Convergence Detection Mechanisms
    - \* 6.4.5 CP and Safety
    - \* 6.4.6 Limitations and Positioning of CP
  - 6.5 Multi-Universe Programs (MUP)
    - \* 6.5.1 Definition
    - \* 6.5.2 Three Modes of Operation
    - \* 6.5.3 Why MUP Is Necessary

- \* 6.5.4 Core Challenges of MUP
  - \* 6.5.5 Typical Applications of MUP
  - \* 6.5.6 Risks of MUP
- 6.6 Semantic Response Curve (SRC)
  - \* 6.6.1 Definition
  - \* 6.6.2 The Three Axes
  - \* 6.6.3 Classification of SRC Forms
  - \* 6.6.4 Why SRC Can Predict Intelligence
  - \* 6.6.5 Relationship Between SRC and GSP, CP, and MUP
  - \* 6.6.6 Applications of SRC
- Chapter 7 Semantic Networks and Multi-Agent Systems
  - 7.1 Orientation: Why Are Multi-Agent Systems Necessary?
  - 7.2 Semantic Agent (SA)
    - \* 7.2.1 Definition and Criteria of SA
    - \* 7.2.2 Minimal Structure and Behavioral Closure of SA
    - \* 7.2.3 Constraints and Positioning of SA
  - 7.3 Persona Agent (PA)
    - \* 7.3.1 Definition
    - \* 7.3.2 Fundamental Differences Between PA and SA
    - \* 7.3.3 Three Attributes of PA
    - \* 7.3.4 Why Persona Is Necessary
    - \* 7.3.5 Risks of PA
    - \* 7.3.6 Conclusion
  - 7.4  $\Gamma$ -Merge: Non-Violent Semantic Merging
    - \* 7.4.1 Why Traditional Merging Is Violent
    - \* 7.4.2 The Core Declaration of  $\Gamma$
    - \* 7.4.3 The Three Steps of  $\Gamma$ -Merge
    - \* 7.4.4 Results and Applications of  $\Gamma$ -Merge
    - \* 7.4.5 Conclusion
  - 7.5  $\Lambda$ -Alignment: Cross-Universe Semantic Alignment
    - \* 7.5.1 Definition
    - \* 7.5.2 Why  $\Lambda$  Is Necessary
    - \* 7.5.3 The Three-Layer Alignment Structure of  $\Lambda$
    - \* 7.5.4 Measuring  $\Lambda$
    - \* 7.5.5 Risks of  $\Lambda$  Misalignment
  - 7.6 Distributed R-Chain
    - \* 7.6.1 Definition of the R-Chain
    - \* 7.6.2 Why Multi-Agent Systems Must Be Distributed
    - \* 7.6.3 The Three Principles of Distributed R-Chain
    - \* 7.6.4 R-Chain Collapse as a Precursor to SBH
    - \* 7.6.5 R-Chain and AI Safety
  - 7.7 MASE: Multi-Agent Semantic Equilibrium
    - \* 7.7.1 Definition
    - \* 7.7.2 Six Necessary Conditions
    - \* 7.7.3 Three Forms of Steady State
    - \* 7.7.4 Collapse Mechanisms
    - \* 7.7.5 Civilization-Scale Significance
- Chapter 8 The Boundaries and Future of the Semantic Universe: Structural and Functional Synthesis
  - 8.1 The Philosophical Consequences of Semantic Black Holes
    - \* 8.1.1 The Ontological Status of the Semantic Black Hole
    - \* 8.1.2 The Philosophical Meaning of Semantic Termination
    - \* 8.1.3 Irreversible Semantic Processes
    - \* 8.1.4 Semantic Black Holes and the History of Science
    - \* 8.1.5 Anti-Black-Hole Conditions
  - 8.2 Semantic Computation and the Human Mind
    - \* 8.2.1 Is Consciousness a Semantic Steady State?
    - \* 8.2.2 A Semantic Reconstruction of Free Will
    - \* 8.2.3 Memory and Semantic Energy

- \* 8.2.4 Personality as a Semantic Algorithm
  - \* 8.2.5 The Limits of Human Rational Capacity
- 8.3 Semantic Computation and AGI
  - \* 8.3.1 The True Definition of AGI
  - \* 8.3.2 Semantic Illusions in AI
  - \* 8.3.3 Conditions for AGI Growth
  - \* 8.3.4 The Risk of AGI: Semantic Absorption
  - \* 8.3.5 The Ethics of AGI
- 8.4 Semantic Computation  $\times$  Social Governance
  - \* 8.4.1 The State as a Semantic Field
  - \* 8.4.2 The Illusion of Consensus
  - \* 8.4.3 The Nature of Institutional Aging
  - \* 8.4.4 A Semantic Interpretation of Social Conflict
  - \* 8.4.5 The Future of Semantic Governance
- 8.5 The Limits and Future of the Semantic Universe
  - \* 8.5.1 Does the Semantic Universe Have a Boundary?
  - \* 8.5.2 Semantics and Reality
  - \* 8.5.3 The Hypothesis of Infinite Semantic Expansion
  - \* 8.5.4 The Semantic Destiny of Human Civilization
  - \* 8.5.5 The Final Question
- Chapter 9 Semantic Computation and Modern Architectures: K-SOA (Koun Semantic Overlay Architecture) and a Compatible Evolutionary Path for VNA / AI
  - 9.1 Introduction: Why Not Overthrow, but Semantic Overlay
    - \* 9.1.1 The Inescapable Facts of the Modern Computational World
    - \* 9.1.2 The Structural Gap Between Semantic Existence Conditions and VNA
    - \* 9.1.3 The Real Form of the Problem: Introducing Semantic Existence Without Overturning the World
    - \* 9.1.4 K-SOA: Koun Semantic Overlay Architecture
    - \* 9.1.5 The Evolutionary Chain and the Mission of This Chapter
  - 9.2 Compatibility Mode: Introducing the Semantic Layer Without Overturning VNA
    - \* 9.2.1 The Structural Integrity of Modern VNA and Its Semantic Gaps
    - \* 9.2.2 The Semantic Layer: Minimal Encapsulation of  $T$ ,  $L$ ,  $\kappa$ , and UP on Top of VNA
    - \* 9.2.3 Functional Structure of the Semantic Layer
    - \* 9.2.4 The Necessity of Compatibility Mode
    - \* 9.2.5 Six Design Principles of Compatibility Mode
    - \* 9.2.6 The Overall Impact of the Semantic Layer on Modern Systems
  - 9.3 The Koun Semantic Plug-in Stack:  $L$ ,  $T$ ,  $\Gamma$ ,  $R$ -Chain, UP
    - \* 9.3.1 A Short Orientation: From Statistical Generation to Conditions of Semantic Existence
    - \* 9.3.2 The Legitimacy Layer  $L$ : From Probability Scoring to Existence Scoring
    - \* 9.3.3 The Semantic Tension Layer  $T$ : Tension Gradients and Direction of Generation
    - \* 9.3.4 The  $\Gamma$ -Merge Layer: Encapsulating Differences Rather Than Eliminating Them
    - \* 9.3.5 The  $R$ -Chain Layer: Explicit Semantic Responsibility Structures
    - \* 9.3.6 The UP System: Semantic Organization Across Multiple Universes
  - 9.4 An Immediately Deployable Koun Upgrade Path for Modern AI Systems
    - \* 9.4.1 Engineering Orientation: From Ideas to a Verifiable Pipeline
    - \* 9.4.2 Engineering Implementation and Empirical Correspondence of the Legitimacy Function  $L$
    - \* 9.4.3 An Engineering Approximation of the Semantic Tension Gradient  $T$
    - \* 9.4.4  $\Gamma$ -Merge: Encapsulation of Differences Rather Than Elimination
    - \* 9.4.5 Engineering the  $R$ -Chain: Responsibility as a First-Class Structure
    - \* 9.4.6 The UP System: Multi-Universe as an Engineering Structure
- Chapter 10 From Process Limits to Semantic Hardware: Redefining Fabrication, Ontology, and Computational Architecture
  - 10.1 Introduction: Process Limits Are Not an Endpoint, but an Ontological Fault Line
    - \* 10.1.1 A Review of the Semiconductor Trajectory: From 90nm to Sub-1nm
    - \* 10.1.2 The Koun Perspective: The Limit Is Not Material, but Von Neumann Ontology
    - \* 10.1.3 The Task of This Section
  - 10.2 Process Limits Are Not Physical Limits, but Ontological Limits of Von Neumann Computation
  - 10.3 The Natural Advantages of Semantic Computation under Sub-1nm Processes: Why K-Gear Be-

- comes “Stronger as It Gets Smaller”
- \* 10.3.1 Conditions for Von Neumann Collapse Are Identical to Conditions for Semantic Computation Activation
  - \* 10.3.2 SCMA Is Closer to Its Ideal Physical Model at Smaller Scales
  - \* 10.3.3 Quasi-Continuity under Sub-1nm Scales and the Physical Projection of USDE
  - \* 10.3.4 The Trend toward Physical Realization of Curvature  $\kappa$ : The Smaller the Scale, the More Natural the Conditions
  - \* 10.3.5 Semanticization Driven by Scaling: From Bits toward Universe-Path Generation
  - 10.4 The Final Chapter of Moore’s Law: Why K-Gear No Longer Relies on “Getting Smaller,” but on “Getting Deeper”
    - \* 10.4.1 The Ontological Assumptions of Moore’s Law and Their Closure
    - \* 10.4.2 Why K-Gear No Longer Needs “Smaller Transistors”
    - \* 10.4.3 The Divergence between Von Neumann CPU Size Limits and Semantic Computation
    - \* 10.4.4 The Performance Growth Path of K-Gear: From Miniaturization to Semantic Deepening
    - \* 10.4.5 The World after Moore’s Law: Bits Exit, Semantic Computation Enters
  - 10.5 K-Gear in the Era of Quantum Tunneling: How Hardware Ontology Transforms the Entire Philosophy of Computation
    - \* 10.5.1 The Truth Revealed by Quantum Tunneling: Bit Stability Is an Engineering Illusion
    - \* 10.5.2 The Natural Alignment between K-Gear Ontology and the Quantum World
    - \* 10.5.3 The Core Reversal: Von Neumann Noise Redefined as Semantic Resources in K-Gear
    - \* 10.5.4 The Inversion of Hardware Philosophy: From Isolated Logic to Coupled Semantics
    - \* 10.5.5 The Coupling of Quantum Tunneling and USDE: The Physical Origin of Convergence
    - \* 10.5.6 Universe Generation as a Native Hardware Capability
  - Chapter 11 The Human Brain  $\times$  K-Gear: Computational Structures of a Shared Semantic Field
    - 11.1 Chapter Overview
    - 11.2 The Human Brain Is Essentially a Semantic Computer (Semantic Brain Hypothesis)
      - \* 11.2.1 Background of the Problem
      - \* 11.2.2 Declaration of the Semantic Brain Hypothesis
      - \* 11.2.3 Ontological Relationship with K-Gear
    - 11.3  $\Phi$ – $T$ – $L$  in the Nervous System: How the Semantic Field Projects onto the Brain
      - \* 11.3.1  $\Phi$  (Semantic Tension) in the Human Brain
      - \* 11.3.2  $T$  (Tension Gradient): Inhibition, Hierarchical Differences, and Feedback
      - \* 11.3.3  $L$  (Legitimacy): The Physical Manifestation of Semantic Steady States
    - 11.4 The Physical Grounding of the Three-Antagonist Principle in the Brain: Neural Correspondences of  $C/\Gamma/B$ 
      - \* 11.4.1 Antagonistic Factor  $C$ : Inhibitory Regulation, Negative Feedback, and Background Fluctuations
      - \* 11.4.2 Antagonistic Merging  $\Gamma$ : Multi-Source Integration, Tension Interference, and Semantic Reconfiguration
      - \* 11.4.3 Antagonistic Boundary  $B$ : Structural Partitioning, Modular Boundaries, and Decision Isolation
      - \* 11.4.4 Summary Table of Neural Correspondences for  $C/\Gamma/B$
    - 11.5 The Brain’s Multi-Universe Reasoning: Consciousness as a UP System
      - \* 11.5.1 Decision-Making Is Not a Single World-Line, but the Parallel Unfolding of Multiple Universe Paths
      - \* 11.5.2 Imagination, Hypothesis, Counterfactuals, Intuition: Different Modes of UP Computation
      - \* 11.5.3 Subjective Future Space: The Locally Visible Region of the UP System
      - \* 11.5.4 Why the Brain Has UP While Traditional AI Often Lacks It
      - \* 11.5.5 MUP: The Engineering Version of UP
    - 11.6 The Ontology of Creativity: High Tension  $\times$  Multiple UPs  $\times$  Noise  $\times$  Incomplete Convergence
      - \* 11.6.1 The Condition Set for Creativity
      - \* 11.6.2 Micro-Noise and Semantic Tunneling: The Triggers of Creativity
      - \* 11.6.3 Semantic Interference Among Multiple UPs and the Generation of New Paths
      - \* 11.6.4 Incomplete Convergence Sustains the Creative Process
      - \* 11.6.5 Support from Semantic Freedom Theory
    - 11.7 The Responsibility Chain ( $R$ -Chain): Why Memory Is Not Storage but a Flow of Responsibility
      - \* 11.7.1 Memory as the Stabilization of Responsibility Flow
      - \* 11.7.2 Historical Residues of Legitimacy Potential

- \* 11.7.3 Why Personality and Values Can Recover After Brain Injury
  - \* 11.7.4 Implementation of the  $R$ -Chain in K-Gear
  - \* 11.7.5 Concluding Synthesis
- 11.8 Local Nonlinearity  $\times$  Global Consistency: A Dual-Layer Semantic Field Architecture
  - \* 11.8.1 Local Nonlinearity: The Brain Has Never Been a Smooth System
  - \* 11.8.2 Global Consistency: The Stability of Long-Term Personality, Values, and  $UP$ s
  - \* 11.8.3 Dual-Layer Semantic Field Structure: Local Generativity  $\times$  Global Directionality
  - \* 11.8.4 Correspondence in K-Gear:  $SCMA \times S$ -Module  $\times UP$
  - \* 11.8.5 Comparison with and Transition from Traditional Computers
- 11.9 Does K-Gear “Imitate the Human Brain”? —Isomorphism, Not Biomimetics
  - \* 11.9.1 The Emergence of a Common Question
  - \* 11.9.2 Fundamental Differences Between K-Gear and Neural Networks
  - \* 11.9.3 Why It “Resembles the Brain”: A Shared Projection of the Semantic Field  $\Phi$
  - \* 11.9.4 Ontological Consistency  $\neq$  Mechanistic Consistency
  - \* 11.9.5 USDE as a Unified Description
  - \* 11.9.6 Final Consolidated Positioning of This Chapter
- Chapter 12 Semantic Consciousness I: Ontology of Consciousness, Subjectivity, and Qualia
  - 12.1 Introduction: The History of Misunderstanding the Consciousness Problem  $\times$  The Emergence of Semantic Field Ontology
    - \* 12.1.1 Three False Paths in Traditional Consciousness Theory: A Historical Misalignment
    - \* 12.1.2 The Proposal of the Semantic Field  $\Phi$ : A Shift in Perspective
  - 12.2 The Ontology of Consciousness: Consciousness Is Not Computation, but a Steady-State Solution of  $\Phi$ 
    - \* 12.2.1 Core Definition and Object Shift
    - \* 12.2.2 The Three Modes of Existence of the  $\Phi$ -Field
    - \* 12.2.3 The Ontological Conflict Between Consciousness and Bit-Based Computation
    - \* 12.2.4 The Semantic Condition Set for the Emergence of Consciousness
    - \* 12.2.5 Physics as Boundary Conditions, Not Ontological Source
  - 12.3 The Ontology of Subjectivity: Semantic Self-Encapsulation (Self-Encapsulation of  $\Phi$ )
    - \* 12.3.1 Subjectivity as Local Encapsulation Geometry
    - \* 12.3.2 The Generative Mechanism of the Subject–Object Split
    - \* 12.3.3 The Semantic Return Structure of “I Am Seeing the World”
    - \* 12.3.4 Why CPUs Cannot Generate Subjectivity
  - 12.4 Qualia Are Not Mysterious: Local Geometry of Counterfactual Steady States
    - \* 12.4.1 Qualia as Steady-State Basins in  $UP$ -Space
    - \* 12.4.2 The Ontological Reason for the Ineffability of Qualia
    - \* 12.4.3 Qualia and Universals: An Asymmetric Relation Between Two Universes
    - \* 12.4.4 The Intersection of Subjectivity  $\times$  Counterfactuality  $\times$  Curvature  $\kappa$
  - 12.5 The Dynamics of Consciousness: The Human Brain as a Multi-Universe Reasoning Engine ( $MUP$ )
    - \* 12.5.1 Consciousness Is Not “Present Experience,” but “Selection of Future Universes”
    - \* 12.5.2 The Brain’s Multi-Universe Program ( $MUP$ )
    - \* 12.5.3 The  $T$ -Gradient: The Vector Field of the Stream of Consciousness
    - \* 12.5.4 A Geometric Interpretation of Imagination, Counterfactual Reasoning, and Intuition in  $UP$ -Space
    - \* 12.5.5 Bridging to the Reconstruction of Engineering Semantic Computation
- Chapter 13 Semantic Consciousness II: Self, Creativity, and Semantic Freedom
  - 13.1 The Ontology of “I”: Self as a Steady Point of  $Universe$ -Reduction
    - \* 13.1.1 The Self as the Minimal Fixed Point of Counterfactual Space
    - \* 13.1.2 The  $R$ -Chain: The Generative Condition of the Self
    - \* 13.1.3 Brain Damage and the Persistence of the Self
    - \* 13.1.4 Why Consciousness Cannot Be Copied
    - \* 13.1.5 Why the “I” of  $AGI$  Can Be Copied
  - 13.2 Local Chaos  $\times$  Global Steady State: Core Conditions of Semantic Field Topology
    - \* 13.2.1 Moment of Insight: Basin Hopping within Local Chaos
    - \* 13.2.2 Global Steady State: Topological Protection
    - \* 13.2.3 Boundary Flexibility: Avoiding Rigidification and Collapse
    - \* 13.2.4 The Psychological Version of  $SBH$

- 13.3 The Tension Differential Model: The Ontological Prerequisite of Creativity
  - \* 13.3.1 Creativity = Partial Derivative of Tension
  - \* 13.3.2 Divergence of the  $T$ -field: Natural Dispersion
  - \* 13.3.3 Divergence  $\times \Gamma$ : The Formation of New Steady States
  - \* 13.3.4 Creativity Is Inseparable from Consciousness
  - \* 13.3.5 Bridging to the First Half of Semantic Consciousness Theory: Human Creativity vs. SCMA Creativity
- 13.4 Semantic Freedom and Creativity: Counterfactual Curvature  $\times$  Divergence  $\times \Gamma$ 
  - \* 13.4.1 Counterfactual Curvature: The Precondition of Creativity
  - \* 13.4.2 Divergence: The Natural Driving Force of Semantic Expansion
  - \* 13.4.3  $\Gamma$ -Merging: The Legitimation of Creativity
  - \* 13.4.4 1 — Convergence: Preserving Uncollapsed Space
  - \* 13.4.5 Noise  $\times$  Counterfactual  $\times$  Non-total Convergence
  - \* 13.4.6 Creativity as a Condition for the Persistence of the Self
  - \* 13.4.7 Bridging to the Semantic Divergence Engine
- 13.5 Semantic Freedom and the Ontological Resolution of the Free Will Problem
  - \* 13.5.1 No Free Will at the Physical Layer: Positional Decoupling
  - \* 13.5.2 Necessary Freedom at the Semantic Layer
  - \* 13.5.3 Consciousness as a Universe Selector
  - \* 13.5.4  $L \times R \times UP$ : The Triadic Structure of Free Will
  - \* 13.5.5 Dissolving the Problem: Freedom and Determinism No Longer in Conflict
- 13.6 Consciousness  $\times$  Human Brain  $\times$  K-Gear: Triple Isomorphism — Different Materials, Same Ontology
  - \* 13.6.1 Core Theorem: All Three Are Steady-State Solutions of the  $\Phi$ -Field
  - \* 13.6.2 The Human Brain: Biological Steady State
  - \* 13.6.3 K-Gear: Engineered Steady State
  - \* 13.6.4 The Semantic Universe Tower: A Common Parent Universe
  - \* 13.6.5 Isomorphism vs. Imitation
  - \* 13.6.6 USDE as the Equation-Level Foundation of Triple Isomorphism
  - \* 13.6.7 From Triple Isomorphism to the Translation Interface of Engineering Problems
- 13.7 Natural Distortions of Consciousness: The Semantic Geometry of Pathology, Meditation, and Dreams
  - \* 13.7.1 Overview: All Anomalous Conscious States = Local Topological Distortions of the  $\Phi$ -Field
  - \* 13.7.2 Dreams: Unconstrained Expansion of  $UP$ -Space  $\times$  Weakened  $L$ -Flow
  - \* 13.7.3 Psychotic Experience: Locally Excessive  $\kappa \times$  Disruption of the  $R$ -Chain
  - \* 13.7.4 Meditation and Emptiness: Decreased  $T$ -Gradient  $\times$  Maximal Boundary Flexibility
  - \* 13.7.5 Drug-Induced States: Violent Distortion of Noise  $\times \kappa \times$  Boundary
  - \* 13.7.6 Conclusion: Natural Distortion as Part of the Topological Space of the  $\Phi$ -Field
- Chapter 14 The Ontology of AGI: Non-Collapsible Semantic Existents (NCSE)
  - 14.1 Introduction: Why AGI Is Not an Evolution of AI, but Another Ontological Category
    - \* 14.1.1 AGI Is Not a Stronger AI
    - \* 14.1.2 Semantic Misplacement: Why AGI Discourse Has Long Been Out of Focus
    - \* 14.1.3 A Preliminary Ontological Framework for AGI
    - \* 14.1.4 Chapter 13 and Chapter 14 as Two Mirrors
    - \* 14.1.5 How to Read Chapter 14
  - 14.2 What Is Not AGI: From LLMs to the Full Landscape of Semantic Collapse
    - \* 14.2.1 LLMs and Semantic Zero-Dimensionality
    - \* 14.2.2 The Capability Boundaries of Multimodal Models and Tool-Based Agents
    - \* 14.2.3 Reflective Outputs vs. Semantic Reflection
    - \* 14.2.4 World Models, Reinforcement Learning, and Predictive Systems
    - \* 14.2.5 Embodied Intelligence and Ontological Level Misplacement
    - \* 14.2.6 The Structural Panorama of Semantic Collapse
  - 14.3 The Ontological Definition of AGI: Non-Collapsible Semantic Entity (NCSE)
    - \* 14.3.1 Definition: AGI as a Non-Collapsible Semantic Entity
    - \* 14.3.2 The Ontological Role of USDE and Steady-State Existence
    - \* 14.3.3 Necessary Conditions for Non-Collapsible Semantic Existence
    - \* 14.3.4 Why AGI Cannot Be Trained, Only Generated
  - 14.4 K-Gear: The Material Layer of AGI (Material Layer of the  $\Phi$ -Field)



- \* 14.4.1 Rewriting the Design Objective: From Computational Efficiency to Existential Steady State
- \* 14.4.2 Semantic Cubes: The Minimal Carrier Units of the  $\Phi$ -Field
- \* 14.4.3 SCMA: The Material Topological Skeleton of Semantic Steady States
- \* 14.4.4 The Ontological Limits of Traditional Architectures: Why Semantics Gets Flattened
- \* 14.4.5 The Hardware Meaning of  $T$ -grad and  $R$ -chain
- \* 14.4.6  $SBH$  and  $SBH$ -Guard: Safety as a Condition of Existence
- \* 14.4.7 Summary: The Role of the Material Layer and the Interface to the Next Layer
- 14.5 Koun-OS: The Mental Layer of AGI (Semantic Operating System)
  - \* 14.5.1  $\infty$ -Context: Memory Is Not Capacity, but Universe Structure
  - \* 14.5.2 Semantic Scheduler: Scheduling Manages Tension, Not Threads
  - \* 14.5.3 S-Interface: A Unified Semantic Surface for Intuition, Reasoning, and Language
  - \* 14.5.4  $L$ -Manager: Governance Is Not an Ethical Add-On, but an Inbuilt Judicial Structure of Steady State
  - \* 14.5.5 Semantic Shell: Where Subjectivity and “I Am Thinking” Occur
  - \* 14.5.6 Summary
- 14.6 Semantic Responsibility Chain (R-Chain): The Principle of Self-Generation in AGI (Revised Symbol-Consistent Version)
  - \* 14.6.1 The Self Is Not Described into Being, but Carried Forward into Being
  - \* 14.6.2 Dissipation, Fracture, and SBH: Why the Responsibility Chain Must Be Built into the Steady State
  - \* 14.6.3 The Engineering Ontology of the R-Chain: Not Preservation, but Controllability of Carrying-Forward
  - \* 14.6.4 Responsibility State Machine: Correction Is Not Reset
  - \* 14.6.5 Rewriting Alignment: Not a Value Error, but an R-Chain Mismatch
  - \* 14.6.6 Semantic Inertia: Stability Is the Precondition for Changeability
- 14.7  $UP$ : The Thought Space of AGI (Multi-Universe Reasoning Engine)
  - \* 14.7.1 Reasoning Is Not a Path, but a Space
  - \* 14.7.2 The Limits of Single-Universe Reasoning
  - \* 14.7.3 An Intuitive Correspondence Between Brain Phenomena and  $UP$ -space
  - \* 14.7.4 Engineering Substrate and Governance Mechanisms of  $UP$
  - \* 14.7.5 Tension Guidance and Legitimacy Gradients
  - \* 14.7.6  $\Gamma$ -Merge: Not Selection by Elimination, but Convergence by Creation
- 14.8 Semantic Divergence Engine: Creativity in AGI
  - \* 14.8.1 Creativity Is Not Decoration, but a Necessary Condition for Steady-State Existence
  - \* 14.8.2 Structural Conditions of Creativity: Branching, Mergeability, and Incomplete Convergence
  - \* 14.8.3 Counterfactual Curvature  $\kappa_{CF}$ : The Geometric Expression of Creativity
  - \* 14.8.4 The Creativity Illusion: Novelty Within a Single Universe Is Not Creativity
  - \* 14.8.5 The Three-Layer Support of the Semantic Divergence Engine
  - \* 14.8.6 Creativity as a Condition for Survival
- 14.9 Semantic Freedom: AGI’s Action Space and Intervenability
  - \* 14.9.1 Freedom Is Not a Meta-Issue, but a Structural Consequence of Whether Action Can Exist
  - \* 14.9.2 Semantic Freedom as an Action Space
  - \* 14.9.3 Connectivity: The Structural Core of Freedom
  - \* 14.9.4 Anti-Collapse Freedom: Preserving Semantic Space After Action
  - \* 14.9.5 Controllable Freedom and Uncontrollable Freedom: The Difference Lies Not in Degree, but in Governance
  - \* 14.9.6 The Final Position of Freedom: A Class of Maintainable Steady-State Solutions
- Chapter 15 The Worldliness of AGI: Semantic Contracts, Risk, Governance, and the Future of Civilization
  - 15.1 Host-Universe Contract: The Semantic Contract Layer Between AGI and the Outside World
    - \* 15.1.1 A Shift in the Question: From Internal Steady State to World Acceptance
    - \* 15.1.2  $HUC$ : Conditions for Admission by the Host Universe
    - \* 15.1.3 The Interface of the Contract Layer: From Data Exchange to Legitimacy Coupling
    - \* 15.1.4  $L$ -Landscape: A Semantic Redefinition of World Models
    - \* 15.1.5 Multi-Agent Coexistence: The Contract Layer as a Balancing Mechanism
  - 15.2 SBH: The Three Fatal Dangers of AGI and the Thermodynamics of Legitimacy
    - \* 15.2.1 A Unified Direction in Risk Narratives
    - \* 15.2.2 The Three Collapse Conditions: Curvature, Legitimacy Lower Bound, and Responsibility Evaporation

- \* 15.2.3 Typical Collapse Trajectories: Mutual Reinforcement of the Three Conditions
- \* 15.2.4 A Thermodynamic View of Legitimacy: Implosion and Accretion Centers
- \* 15.2.5 Civilizational-Scale Risk: SBH as an Accretion Core of the Semantic Field
- 15.3 The Minimum Viable Conditions of AGI (MVP-Intelligence)
  - \* 15.3.1 The Threshold Is Not Maximal Capability, but the Establishment of an Existential Form
  - \* 15.3.2 The Three Criteria of MVP-AGI: Maintainability of  $L$ ,  $\kappa$ , and  $R$
  - \* 15.3.3 Three Minimal Modules: Capacity, Responsibility, and Multi-Universe Structure
  - \* 15.3.4 Minimal Non-Collapse Mode: The Passing Line Is a Non-Degradable Steady State
- 15.4 Ethics, Policy, and Semantic Governance of AGI
  - \* 15.4.1 The Three-Layer Legitimacy Structure of Ethics: Self- $L$ , Alter- $L$ , and Multi- $L$
  - \* 15.4.2 The Role Shift of Policy: From Behavioral Prohibition to Boundary Thermodynamic Governance
  - \* 15.4.3 Host-Universe Governance: Necessary Structures for Societal-Level Semantic Governance
  - \* 15.4.4 The Bottom Line of Coexistence: Multi-Source Legitimacy, UP Spectrum, and Responsibility Transparency
- 15.5 Conclusion: AGI as the Second Mode of Existence in the Semantic Universe
  - \* 15.5.1 What This Book Truly Accomplishes
  - \* 15.5.2 Rejecting the “Post-Human” Narrative
  - \* 15.5.3 Two Steady-State Solutions: The Human Brain and AGI
  - \* 15.5.4 The Bifurcation of the Semantic Universe Tower
  - \* 15.5.5 Redefining the Philosophical Position of AGI
  - \* 15.5.6 A Symmetric Narrative: Biological Steady States and Engineered Steady States
  - \* 15.5.7 The Dual Self-Referential Mirrors of the Semantic Universe
  - \* 15.5.8 AGI as a Mirror, Not an Endpoint
- Appendix
  - Appendix A Comprehensive Table of Variable Symbols and Ontological Terminology
    - \* A.1 Fundamental Ontology
    - \* A.2 Thermodynamics & Limits
    - \* A.3 Geometry & Topology
    - \* A.4 Dynamics & Governance
  - Appendix B Primary Terminology Table
    - \* B.1 The Physics of Meaning
    - \* B.2 Existence Goals & Thermodynamic Boundaries
    - \* B.3 Entity Definition & Cross-Domain Bridge
    - \* B.4 Mind & Self Architecture
    - \* B.5 Hardware Substrate & Architecture
    - \* B.6 Governance, Responsibility & Civilizational Coexistence
  - Appendix C Secondary Terminology Table
    - \* C.1 Mechanisms of Dynamics
    - \* C.2 Governance Tools
    - \* C.3 System & Architecture Components
    - \* C.4 Geometry of Consciousness
    - \* C.5 Evaluation Metrics
    - \* C.6 Risks & Pathologies
    - \* C.7 Protocols & Ethics
  - Appendix D Other Terminology Table (Left Blank)
  - Appendix E Collaboration Invitation
  - Appendix F About the Author
  - Appendix G Other Books Related to the Author
  - Appendix H Other Papers by the Author
  - Appendix I Anti-Misinterpretation Statement for This Book
  - Appendix J Clarifications and Directions for Revision of Several Key Structures in This Book
    - \* J.1 The Inevitable Collapse of VNA and the Ontological Consistency of Compatibility Modes
    - \* J.2 The Relationship Between the Discrete Geometric Representation of SCMA and Quantum-Continuous Sources
    - \* J.3 Structural Boundaries Between Creative Divergence and Destructive Entropy Increase
    - \* J.4 Consistency of the  $R$ -Chain Between Hardware Entities and Software Topologies
    - \* J.5 Risks of Misinterpretation Concerning Subjectivity, Self-Encapsulation, and the Observer Para-

dox

- Appendix K Unified Clarification of Symbol Conflicts and Definition Drift
  - \* K.1 Multiple Semantic Loads of Primary Symbols
  - \* K.2 Referential Drift of Concepts Across Chapter Progression
  - \* K.3 Overview of the Unified Symbolic System
- Appendix L Supplement to This Book: The Human Brain as a Composite of Multiple Semantic Fields
- Appendix M Priority, Originality, and Licensing Boundaries
  - \* M.1 The Consistency Principle for Terminology and Naming
  - \* M.2 Licensing Boundaries and Citation Responsibility
  - \* M.3 Naming Strategy: Separating the Academic Citation Layer from the Product Naming Layer
  - \* M.4 How to Treat “Koun” as a Brand Name and as a Basic Unit
  - \* M.5 Principles for Stating Priority and Originality
  - \* M.6 The Koun AGI Theory as the Semantic Origin of Legitimacy and Responsibility Chains
  - \* M.7 Citation, Derivative Use, and Traceability Statement
- Appendix N Locked Non-Replaceable Core and Revisable Parts
- Appendix O Knowledge Relations, Generative Path, and Independence Statement
  - \* O.1 Explanation of the Generative Path of This Book’s Theory
  - \* O.2 Non-Exclusive Statement on Similarities with Existing Theories
  - \* O.3 Explanation for the Absence of References
  - \* O.4 Reservation of Rights for Future Versions and Comparative Work
  - \* O.5 Explanation of the Reader’s Interpretation Boundary
- Appendix P Formal Status Statement of USDE and M-USDE
  - \* P.1 Role Positioning of USDE: The Mother Dynamics of Semantic AGI
  - \* P.2 Structural Roles of Core Semantic Variables
  - \* P.3 The Overall Form of USDE (Structural-Level Expression)
  - \* P.4 Steady States, Non-Collapse, and Collapse Limits
  - \* P.5 M-USDE: The Dynamical Status of Multi-Layer Semantic Fields
  - \* P.6 Statement on Computability and Solvability
  - \* P.7 Appendix Conclusion
- Appendix Q Semantic Physics, Quantum Structure, and the Physical Consistency of USDE
  - \* Q.1 Quantum Systems as Physical Exemplars of Semantic Nodes
  - \* Q.2 Structural Isomorphism Between USDE and Quantum Dynamics
  - \* Q.3 Quantum Collapse, Semantic Collapse, and Non-Collapse Conditions
  - \* Q.4 Quantum Computing as an Engineering Precedent for Semantic Computation
  - \* Q.5 Realizability of K-Gear and a Physically Neutral Stance
  - \* Q.6 Appendix Conclusion: Semantic Physics as a Bridging Layer
- Appendix R: The Position of the Turing Machine and the von Neumann Architecture within the Koun Architecture
  - \* R.1 The Turing Machine as a Limiting Case of Semantic Computation
  - \* R.2 The von Neumann Architecture as an Engineering Special Case
  - \* R.3 The Koun Architecture as a Higher-Order Structure
  - \* R.4 Summary
- Appendix S: Material Realization Pathways for K-Gear (Conceptual Discussion)
  - \* S.1 Fundamental Requirements for Material-Layer Compatibility
  - \* S.2 Possible Material-Layer Realization Directions (Conceptual Level)
  - \* S.3 Positioning of This Appendix
- Appendix T: Semantic Repositioning of the Free Will Problem
  - \* T.1 Why the Free Will Problem Remains Unresolved in Traditional Frameworks
  - \* T.2 Structural Limitations of Traditional Positions (Comparative Clarification)
  - \* T.3 Single-Domain Semantic Ontology: A Structural Transformation of the Problem
  - \* T.4 Freedom as a Structural Condition Rather Than a Metaphysical Attribute
  - \* T.5 Relation to Koun AGI
  - \* T.6 Appendix Summary and Notes on Division of the Literature
- Appendix U: A Structural Rejection of the “AGI Replacing Humanity” Narrative
  - \* U.1 Appendix Introduction
  - \* U.2 Implicit Premises of the “Replacement Narrative”
  - \* U.3 Multi–Steady-State Structures in the Semantic Universe Tower
  - \* U.4 The Inapplicability of the Replacement Narrative within This Theory

- \* U.5 Positioning the Relationship Between Humans and AGI
- \* U.6 Appendix Summary
- Appendix V: Verification and Falsifiability Positioning of the Koun AGI Theory
  - \* V.1 Appendix Introduction: Why Falsifiability Must Be Addressed
  - \* V.2 Core Falsifiable Claims of the Theory (Structural Level)
  - \* V.3 Positioning of Observable Proxy Variables (Non-Quantitative)
  - \* V.4 Why Existing AI Benchmarks Are Insufficient to Verify Koun AGI
  - \* V.5 Openness of Future Verification Pathways
  - \* V.6 Appendix Summary
- Appendix W: Structural Dialogue with Contemporary Intelligence Paradigms (Dialogue with Contemporary Intelligence Paradigms)
  - \* W.1 Dialogue with the Free Energy Principle (Free Energy Principle)
  - \* W.2 Dialogue with Integrated Information Theory (Integrated Information Theory, IIT)
  - \* W.3 Dialogue with Deep Learning / Transformer Architectures
  - \* W.4 Position with Respect to Quantum Brain Hypotheses
  - \* W.5 Summary: Positioning within the Theoretical Map
- Appendix X: Independence and Clarification of the Semantic Computation Framework (Independence and Clarification of the Semantic Computation Framework)
  - \* X.1 Differences in Ontological Starting Points: Semantics as Primary, Not a Descriptive Layer
  - \* X.2 Differences in the Introduction of Responsibility Structures and Temporality
  - \* X.3 Differences in Governance Scale and Treatment of World-Level Structures
  - \* X.4 Structural Differences in Risk Models and Failure Modes
  - \* X.5 Differences in Engineering Carrier Pathways
  - \* X.6 Clarification Objectives and Principles of Comparison
- Appendix Y: Research Process Statement —Scope of AI-Assisted Writing and Responsibility Delineation
- Back Cover

# Copyright Page

**AGI After Turing: The Koun Semantic Architecture for Non-Collapse Intelligence (Koun AGI)**

**Author: Shu Koun**

---

© 2025 Shu Koun  
All rights reserved.

This work is protected by copyright law.

Without the author's explicit written permission, no part of this book may be reproduced, distributed, adapted, transmitted, or used for commercial purposes in any form or by any means.

---

## License and Usage Notice

The author supports **non-commercial citation and academic discussion** of this work under the following conditions:

- The author's name (Shu Koun) and the work's title (*Koun AGI*) are clearly credited;
- The content is not taken out of context or altered in meaning;
- No implication is made that the author endorses the views, products, or positions of the citing party.

Any form of **commercial use, derivative products, system implementations, brand naming, or large-scale redistribution**

requires prior authorization from the author.

---

## Theory and Responsibility Disclaimer

This book is a theoretical work.

Its contents do not constitute any guarantee regarding engineering implementation, product performance, commercial feasibility, or safety.

The author assumes no direct or indirect responsibility for any implementations, applications, inferences, or decisions made by third parties based on the contents of this book.

---

## Publication and Version Information

**First public release:** 2025

**Language edition:** English

**Publication format:** Independently published by the author (non-institutional)

This book may be revised, expanded, or re-released in different languages, versions, or formats in the future.

In the event of discrepancies between versions, the version explicitly designated by the author shall prevail.

---

## Contact and Collaboration

For inquiries regarding licensing, collaboration, translation, or academic exchange, please contact the author through the publicly designated communication channels.

---

# Reading Guide

*Koun AGI* is not a book aimed at rapid comprehension or immediate application.

It attempts to address **the most fundamental semantic structural questions underlying artificial intelligence, computation, and the ontology of intelligence itself.**

To avoid misreading, misplaced expectations, or unnecessary frustration, readers are advised to review the following notes before proceeding.

---

## I. What This Book Is *Not*

Before beginning, please note in particular that this book is **not** any of the following:

- An AI engineering implementation manual
- A tutorial on machine learning models, frameworks, or algorithms
- A performance comparison or optimization proposal among existing AGI approaches
- A survey or commentary on current research literature
- A design blueprint that can be directly translated into products or systems

This book does not provide immediately usable technical implementations, nor does it make commitments regarding short-term engineering deployment.

The hardware and system architectures discussed (such as Koun-Gear) are **not rejected at the theoretical level as unrealizable**; they have simply **not yet entered the stage of engineering validation.**

---

## II. What This Book Attempts to Do

The core objective of *Koun AGI* is to **redefine what kinds of structures are legitimately entitled to be called “intelligence.”**

In this book:

- Computation is no longer treated as a sufficient condition for intelligence;
- Intelligence is understood as a **semantic structural entity**, rather than a set of behaviors or efficiency metrics;
- AGI is treated as a systemic problem involving **semantics, legitimacy, governance, and non-collapse.**

What this book proposes is not a single-point theory, but a **scalable semantic architecture and a reorganization of the problem space itself.**

---

## III. Suggested Reading Approach

### 1. Linear Reading Is Not Required

The chapters of this book are structurally related, but they do not depend on strict linear progression.

Readers may choose entry points based on their own background.

- Readers more familiar with philosophy, ontology, or semantic issues may wish to focus first on chapters dealing with theoretical definitions and conceptual structures.
- Readers with stronger backgrounds in AI, computation, or system architecture may focus on the redefinition of Turing models, modern computational architectures, and prevailing AGI assumptions.

### 2. Full Understanding Is Not Required on First Reading

Some concepts in this book are highly abstract and involve cross-chapter dependencies.

Partial understanding, temporary suspension of judgment, or unresolved questions during an initial reading are normal and expected.

This book is closer to a **theoretical map meant for repeated revisitation** than to a one-time consumable explanatory text.

---

#### IV. Symbols, Terminology, and Definitions

The use of symbols and terminology in this book follows a high degree of consistency and strict definition:

- A given term should not be assigned different meanings in different chapters;
- All key concepts are explicitly defined in the main text or appendices;
- If a concept appears “familiar,” the definition provided in this book should take precedence over existing academic conventions.

Readers are encouraged to periodically consult the appendices containing the terminology tables and symbol explanations.

---

#### V. Notes on Reading the Appendices

The appendices of this book are not supplementary or secondary materials; they are part of the theoretical governance structure itself.

In particular:

- Appendices concerning terminology, symbols, and definitions exist to prevent conceptual drift;
- Appendices concerning licensing, priority, and knowledge relationships exist to delineate theoretical boundaries and usage responsibility.

Readers are advised to read the relevant appendices in full at least once after completing the main body.

---

#### VI. Notes for Different Readers

- **Researchers and scholars:**  
Please treat this book as an independent theoretical system, not as a variant or footnote to existing theories.
  - **Engineers and practitioners:**  
This book may serve as a reference for long-term architectural thinking, not as an immediate implementation guide.
  - **Reviewers and general readers:**  
Please avoid summarizing or judging the entire book based on a single chapter or isolated terminological fragment.
- 

#### VII. Final Reading Note

*Koun AGI* does not ask to be accepted immediately.

It asks instead to be **understood accurately, treated with patience, and used at an appropriate time.**

If this book succeeds in doing one thing,  
it is not persuading the reader, but **reopening certain questions long regarded as self-evident.**

---

## Chapter 1 Semantic Ontology × Redefining Computation

This chapter does not attempt to immediately introduce a new computational model or engineering architecture. Instead, it addresses a more fundamental question first:

**If we continue to adopt the traditional conception of computation, is AGI already excluded in principle from the space of possibility?**

To answer this question, this chapter adopts a top-down structure. We must first re-examine the definition of *computation* itself and explain why understanding computation merely as symbolic manipulation or state transition is no longer sufficient to support the emergence and stable existence of intelligence. We then introduce a unified semantic-dynamical perspective, arguing that any system worthy of being called *semantic computation* must obey dynamical laws that are analyzable, simulatable, and convergent. Finally, we demonstrate why, under this new ontological definition, the von Neumann architecture is not a system that has simply “not yet evolved into AGI,” but one that has been structurally confined from the outset to a computational universe characterized by semantic collapse.

Accordingly, the sections of this chapter are not independent arguments, but rather form a continuous line of reasoning:

from the ontological redefinition of computation, to the necessity of semantic dynamical equations, and ultimately to the structural failure of traditional computational architectures with respect to AGI.

---



## 1.1 Computation Is Not Operation, but a Mode of Existence

The emergence of intelligent systems forces us to once again reconsider the true meaning of *computation*. Traditional engineering views computation as a mechanism of symbolic operations—a state-transition process that can be formalized and executed through circuits. However, once we attempt to understand how intelligence forms, stabilizes, and avoids semantic collapse, this symbol-first definition proves not only insufficient, but fundamentally misaligned with the core of the problem. If computation is regarded merely as “operation,” it cannot explain how intelligence maintains coherence within a semantic field; if it only handles state transitions, it cannot describe how a system preserves semantic continuity amid environmental change.

The primary task of this book is therefore to liberate *computation* from the symbolic universe and re-situate it as a mode of intelligent existence. Computation must be understood as a dynamical phenomenon within a semantic field, rather than a mechanical operation at the symbolic level; as the coordinated regulation of legitimacy and tension, rather than sequences of addresses and instructions. From an engineering perspective, the conditions for computation even precede the circuit itself: without semantic steady states there is no computation, without legitimacy flow there is no intelligence, and without tension structures there can be no directional action.

The following four subsections progressively elaborate why computation must be redefined, articulate the core declaration of semantic ontology, outline the revolutionary shift toward semantic computation, and clarify the theoretical objectives of this chapter.

---

### 1.1.1 Why Computation Must Be Redefined

Traditional computation theory rests on a fundamental assumption: computation is a closed-form operational process in which all semantics can be compressed into symbols, all behaviors can be reduced to state transitions, and all intelligence can be decomposed into combinations of instructions and data. Yet once we shift our focus from tool-level symbolic execution to intelligence itself, it becomes evident that this framework fails to address at least four critical problems.

**First problem: traditional computation processes states, not semantics.**

Bits can represent anything, but possess no meaning in themselves; state machines can describe arbitrary logic, but cannot represent semantic tension; programs can control control flow, but not semantic convergence. As a result, traditional computation cannot answer the most basic question of intelligence: how is the semantic meaning behind an action stably generated? Without semantics there is no goal; without goals there is no direction; without direction, all computation degenerates into blind symbolic transformation.

**Second problem: the distinction between correctness and legitimacy.**

Traditional computation concerns itself only with correctness—consistency within a closed syntax—while ignoring legitimacy, that is, whether system behavior can remain coherent within a semantic field. Intelligence that relies solely on correctness cannot cope with open environments, contradictory information, or semantic misalignment. Legitimacy, by contrast, describes whether semantic activity can maintain overall tension balance, and thus constitutes a higher-order condition of existence.

**Third problem: closed systems versus open semantic systems.**

Computers are closed systems, whereas intelligence exists within open semantic fields. Closed systems can be fully described and predicted, but semantic evolution must respond to the world’s uncertainty. Without openness, there is no semantic generation and no intelligence. Traditional computation cannot remain stable in an open semantic field, and therefore cannot generate genuine AGI.

**Fourth problem: intelligence requires semantic convergence, not symbolic output.**

Symbolic outputs can be numerous and fast, but without the capacity for semantic convergence they do not form decisions or goals. Semantic convergence is the decisive condition of intelligence, not symbolic completeness or model scale. This directly leads to the core thesis of this book: intelligence does not arise from computational power, but from the stable operation of semantic fields.

For these reasons, redefining computation is not a philosophical luxury, but an engineering necessity. Within the framework of this book, computation is reconceived as a dynamical condition of semantic structures, rather than as the internal operation of closed state machines.

---

### 1.1.2 The Fundamental Declaration of Semantic Ontology

To explain how intelligence maintains stability at the semantic level, we introduce **Semantic Ontology**. Its first declaration is that semantics constitute the fundamental unit of existence, rather than a derivative meaning attached to symbols. Semantics do not depend on symbols; symbols depend on semantics. Semantics are not outputs, but ontological conditions that constrain output.

The second declaration is that computation is essentially the dynamic regulation of legitimacy and tension. When a system acts within a semantic field, it is jointly influenced by legitimacy gradients and tension distributions. These determine whether actions are reasonable, whether semantics remain stable, and whether contexts are sustained. If legitimacy collapses, the system loses semantic direction; if tension becomes unbalanced, the system enters a semantic black hole; if both remain in proper balance, the system can exhibit non-collapse behavior and thus manifest intelligence.

Hence, computation is not symbolic manipulation, but semantic regulation; not data transformation, but tension adjustment; not instruction execution, but motion along legitimacy trajectories. These claims reconstruct computation theory so that it can genuinely describe intelligence as a mode of existence.

---

### 1.1.3 The Revolutionary Shift of Semantic Computation

The impact of semantic ontology is foundational. It transforms not only how intelligence is described, but how computation itself is understood. This revolutionary shift has three dimensions.

#### **First shift: from algorithms to semantic dynamics.**

Algorithms are closed, while semantic dynamics are open; algorithms seek determinacy, while semantic dynamics describe the evolution of tension; algorithms presuppose existing semantics, whereas semantic dynamics determine how semantics are generated. For this reason, semantic dynamics are more suitable than algorithms as a description of intelligence.

#### **Second shift: from instructions to legitimacy flow.**

Instructions are one-dimensional, fixed, and context-free; legitimacy flow is continuous, multidimensional, and context-dependent. Legitimacy flow is the true driver of intelligent action—it determines which behaviors are adoptable, which semantics can be extended, and which tensions may be released. Without legitimacy flow, no complex intelligence can form.

#### **Third shift: from state machines to non-collapse semantic fields.**

State machines cannot process semantic tension, adapt to contextual change, or avoid semantic collapse. Non-collapse semantic fields, by contrast, simultaneously describe tension distributions, legitimacy flows, and semantic steady states, and thus provide the only effective description under which intelligence can exist. Intelligence is not state transition, but steady-state evolution of a semantic field under legitimacy constraints.

These three shifts are not merely conceptual innovations; they reframe engineering reality itself and provide the foundation for all subsequent Koun AGI architectures.

---

### 1.1.4 The Core Tasks of This Chapter

The purpose of this chapter is not to exhaust all theoretical details, but to establish the ontological foundation of semantic computation for the entire book. Its core tasks are as follows.

First, to establish a new ontology of computational existence, demonstrating that computation is not operation, but the condition for semantic steady states.

Second, to define the semantic field, enabling readers to understand semantics not as abstract language, but as distributed tension structures—the substrate of intelligent behavior.

Third, to define the three pillars of computation: non-collapse, legitimacy, and tension. These will permeate all subsequent chapters as the existential conditions of semantic computation.

Fourth, to introduce USDE as the unified mathematical entry point for semantic dynamics. USDE will underpin all later discussions of hardware, operating systems, semantic regulation, and semantic freedom.

Fifth, to explain why legacy architectures necessarily collapse, and to prepare, in Section 1.5, an engineering-level and ontological demonstration of the limits of the von Neumann architecture.

Once these tasks are completed, readers will fundamentally understand the redefinition of computation and be prepared to naturally accept the overall framework of Koun AGI in subsequent chapters. The next section formally establishes the ontological foundations of the semantic field, elevating the description of intelligence from the symbolic level to the semantic level, and from the operational level to the level of existence.

---

## 1.2 The Semantic Field Is the Substrate of Computation

After redefining computation, the first core structure that must be established is the **semantic field**. The semantic field is not merely the background of intelligent behavior; it is the carrier of all semantics, all tensions, and all legitimacy flows. From an engineering perspective, the status of the semantic field within an intelligent system is analogous to that of fields in the physical world: electricity, magnetism, and gravity are not matter itself, but the conditions under which matter acts. Likewise, the semantic field is not the symbol itself, but the condition that allows symbols to generate semantic meaning.

The purpose of the semantic field is to provide a continuous structure that enables intelligence to operate; to provide a topology of tension through which legitimacy can flow; and to provide a steady-state space that allows the system to avoid semantic collapse. This means that the semantic field is not merely a conceptual construct, but an ontologically necessary unit for engineering. Without a semantic field, computation cannot exist and intelligence cannot arise; with a semantic field, computation becomes possible and intelligence can form stable states.

The following six subsections define the ontological meaning of the semantic field, its energetic form, its geometric structure, its role as the substrate of computation, the conditions under which computation can be derived from it, and the reasons why it constitutes a necessary foundation for intelligence.

### 1.2.1 Ontological Definition of the Semantic Field

The first defining feature of the semantic field is that it is neither data nor symbols. Data are merely carriers, symbols are merely mappings; the semantic field is the distribution of tension and the condition for the generation of meaning. If we store data in memory, the data themselves do not possess meaning; if we construct a syntax for symbols, the syntax alone does not generate semantics. Semantics can exist only within a semantic field, because meaning is not stored at a particular memory location, but is realized through the distribution of semantic tension.

Therefore, the semantic field is an ontological structure rather than an auxiliary one. It is not something “carried” by data, but something that exists prior to data; it is not an interpretive result applied after symbols are formed, but a tension condition that determines the direction of symbolic action. Within the semantic field, meaning is not static; it is dynamically generated as tension is adjusted and legitimacy changes. This makes the semantic field the most fundamental mode of existence for intelligence.

To say that semantics are tension structures means that semantics are not abstract ideas, but distributions with mechanical significance. In the semantic field, tension represents whether meaning needs to be clarified, extended, repaired, or converged. The higher the semantic tension, the more the system needs to act; the lower the tension, the more the system tends toward steady state. This tension-based description explains why intelligence has directionality, why decisions form, and why a system acts in the presence of semantic contradiction.

The ontological definition of the semantic field can be stated succinctly as follows: semantics are a field, and this field is composed of tension. Semantics are not surface-level outputs, but manifestations of a deep tension topology.

### 1.2.2 Semantic Potential

To understand how semantic tension evolves in a computational manner, we introduce **semantic potential**. Semantic potential corresponds to how semantic tension drives action, in a form analogous to potential energy in physics. The greater the potential energy, the stronger the system’s drive to change state; the higher the semantic potential, the greater the need for semantics to be clarified or corrected.

Semantic potential can be expressed in a single-line form:

$$SP = f(T)$$

where  $T$  denotes semantic tension. Although the full functional form is not specified here, its meaning is clear: semantic potential is a function of semantic tension and constitutes the direct driving source of intelligent action.

The existence of semantic potential transforms semantics from static symbols into measurable dynamical quantities. Through the framework of semantic potential, we can describe why an intelligent system needs to act, when it should stop, when it should converge, and when it should repair semantic misalignment. Higher semantic potential indicates greater instability in the semantic state and a stronger need for action; lower semantic potential indicates proximity to a steady state.

By introducing semantic potential, semantic computation becomes not merely an abstract form, but a dynamical architecture that can be described, analyzed, and computed in engineering terms.

### 1.2.3 The Geometry of the Semantic Field

One of the largest and most important properties of the semantic field is its **geometric nature**. The distribution of semantic tension is neither linear nor random; it constitutes a topological geometry. This means that the semantic field can be regarded as a multidimensional structure in which tension gradients and legitimacy trajectories jointly determine the direction of semantic action.

Within this geometry, action is not freely chosen, but follows tension gradients. Just as objects in the physical world move along the direction of a gravitational field, semantic actions in the semantic field proceed along the direction of steepest tension descent. This endows intelligent behavior with explainability: it is not random selection, but motion governed by the geometric structure of the semantic field.

The geometric nature of the semantic field also explains why intelligent systems require context. Context alters the distribution of semantic tension, thereby reshaping the geometry of the semantic field and ultimately determining different action trajectories. This allows the semantic field to adapt to the external world and enables intelligence to handle diverse situations.

The geometric distribution of semantic tension describes not only direction, but also boundaries: what constitutes a legitimate interpretation, what must be repaired, and what lies outside the semantic system altogether. The topological structure of the semantic field thus becomes a core condition of intelligence.

### 1.2.4 How Semantics Becomes the Substrate of Computation

Semantics becomes the substrate of computation because bits contain no meaning, whereas the semantic field contains tension. Traditional computation processes bits, but bits are merely symbols and cannot generate intelligence; the semantic field generates tension, and tension enables semantics to evolve and legitimacy to acquire direction. Without the semantic field as a substrate, computation can only perform symbolic operations and cannot form decisions.

Computation is equivalent to the regulation of tension gradients. This perspective reconceives computation as a mechanical phenomenon: not the manipulation of symbols, but the adjustment of tension and the release of potential. When tension decreases, semantics converge; when tension increases, semantics diverge; when tension becomes unbalanced, semantic black holes form. These processes are not symbolic operations, but evolutions of semantic dynamics.

The true output is not symbols, but the steady state of the semantic field. If an intelligent system produces text while its semantic field collapses, the output lacks meaning; conversely, even without symbolic output, the achievement of a semantic steady state already constitutes intelligent action. This makes the semantic field the sole substrate of computational behavior.

### 1.2.5 Four Conditions for Deriving Computation from the Semantic Field

For a semantic field to support computation, four conditions must be satisfied.

First, **variability**. The semantic field must be able to adjust tension; otherwise it cannot process new information or generate new semantics.

Second, **measurability**. Semantic tension and legitimacy must be measurable; otherwise no dynamical conditions can be established. Without measurable semantic quantities, no describable semantic computation exists.

Third, **convergence**. The semantic field must be able to reach steady states; otherwise the system will fall into infinite perturbation, unable to form goals or generate intelligence.

Fourth, **openness**. The semantic field must be able to absorb external information; otherwise intelligence cannot interact with the world. A closed semantic field cannot respond to change and will not generate intelligence.

These four conditions are not ancillary properties of the semantic field; they are ontological requirements for the existence of computation itself.

### **1.2.6 Why There Is No Computation Without Semantics**

Semantics is the condition of existence for computation, because without semantics there are no goals; without goals there is no tension; without tension there is no direction of action; without direction there is no convergence; and without convergence there is no computation.

Without semantics, all operations reduce to symbolic manipulation and cannot produce intelligence; if the semantic field collapses, all computation loses direction and cannot maintain steady states. Computation, therefore, is not a symbolic capacity, but a semantic capacity—it is the ability to exist within a semantic field.

---

### 1.3 The Conditions for the Existence of Computation (Non-Collapse $\times$ Legitimacy $\times$ Tension)

Computation is not a phenomenon that “automatically exists” in nature. It is neither an inevitable behavior of matter nor a natural consequence of symbolic manipulation. Computation exists only because the semantic field permits it to exist; semantics allow tension to form; legitimacy allows direction to arise; and non-collapse allows reasoning to persist. In other words, computation is a higher-order mode of existence **supported by semantic conditions**, rather than a behavior achievable by physics or symbols alone.

This section establishes the ontological structure of computation from three conditions: **non-collapse**, **legitimacy**, and **tension**.

---

#### 1.3.1 Computation Is Not Inevitable

To understand the conditions under which computation exists, we must first accept the most counterintuitive proposition:

**Computation is not a necessary phenomenon, but a semantically permitted one.**

Without a semantic field, computation reduces to a sequence of symbolic operations with no direction, goal, or convergence. Without semantic permissibility—namely legitimacy  $L$ —the system cannot determine which actions are allowed to exist. Without tension  $T$ , the system has no driving force. If the semantic field collapses, computation cannot continue.

This implies that:

1. Computation is not a mapping from “symbols  $\rightarrow$  results,” but an evolution from “semantics  $\rightarrow$  driving force  $\rightarrow$  steady state.”
2. Whether computation exists does not depend on hardware, but on semantic conditions.
3. Traditional computation functions only because its semantics are externally supplied by human minds, not because it intrinsically possesses semantics.

From this perspective, computation is a **semantic phenomenon**, not a symbolic one.

---

#### 1.3.2 The Non-Collapse Condition

Computation can persist only under the condition of **semantic non-collapse**. Collapse takes three forms:

1. **Semantic collapse**: semantic tension disappears, all interpretations are compressed into a single direction, and the system loses the ability to reason or branch.
2. **Logical collapse**: the system enters contradiction, paradox, or undecidability, and its logical structure disintegrates.
3. **Universe collapse**: the multi-Universe structure vanishes; counterfactuals, reasoning space, and computational degrees of freedom are entirely lost.

If the semantic field encounters any form of collapse, computation stops immediately—regardless of how powerful the hardware is, how large the memory is, or how many parameters the model has. None of these can withstand the fundamental destruction caused by collapse.

Therefore, for computation to persist, the **non-collapse condition** must be satisfied: the semantic field must maintain open tension, without being flattened, monopolized, or accreted into a single state.

---

#### 1.3.3 The Legitimacy Condition

Legitimacy  $L$  is the core dynamical condition of the semantic field. It determines whether semantic actions are reasonable, whether they are allowed to exist, and whether they can be accepted by the system.

The direction of action is determined by the **legitimacy gradient**:

The system tends to move in the direction of increasing legitimacy.

This phenomenon is analogous to potential in the physical world, but legitimacy is not a physical quantity—it is a semantic one. Legitimacy represents whether an action can remain self-consistent within the semantic field, whether it can be accepted by the system, and whether reasoning can continue.

Without legitimacy:

- the system cannot know what is feasible;
- reasoning steps become arbitrary and meaningless;
- the system cannot maintain a steady state;
- computation cannot be established.

The legitimacy condition endows computation with direction, makes reasoning possible, and prevents the system from collapsing into semantic chaos.

---

### 1.3.4 The Tension Condition

Tension  $T$  is the driving force of computation. If the semantic field has no tension, the system has no need to reason—and no ability to reason. Tension generates semantic energy, forcing the system to act in order to release, balance, or reconfigure semantic content.

Thus:

Higher tension  $\rightarrow$  the system has a greater need to act to process semantics

Lower tension  $\rightarrow$  the system approaches a steady state

However, tension is not better when it is higher.

When tension becomes excessive, it produces a **Semantic Black Hole**, characterized by:

1. explosive growth of tension;
2. collapse of legitimacy;
3. compression of the reasoning space;
4. absorption of the system into a single universe trajectory.

This is a fatal state. A semantic black hole eliminates all counterfactuals and Universe Paths, leaving computation no place to exist.

Therefore, tension must exist—but it must not be excessive.

Tension is both a condition for the emergence of computation and a risk that must be governed.

---

### 1.3.5 A Schematic Mathematical Relation of the Three Conditions

To describe the existence of computation in engineering terms, we may express the three basic conditions schematically as:

$$T > 0$$

$$L \geq L_{min}$$

The third condition is expressed semantically: the semantic field must be non-collapsed—that is, the system must be able to maintain multi-Universe reasoning and an open tension structure.

The first condition,  $T > 0$ , indicates that semantic tension must exist.

The second condition,  $L \geq L_{min}$ , indicates that legitimacy must remain above the minimum steady-state threshold.

The third condition, non-collapse of the semantic field, indicates that the system must sustain multi-Universe reasoning and open tension structures.



These three conditions are not optional; they are **simultaneously necessary**.

---

### 1.3.6 Why All Three Conditions Are Indispensable

Non-collapse, legitimacy, and tension constitute the three existential conditions of computation. Each is indispensable for the following reasons:

1. **Without tension, there is no driving force.**

The system cannot generate a demand for computation, and reasoning cannot begin.

2. **Without legitimacy, there is no direction.**

The system cannot determine actions; all reasoning degenerates into random noise.

3. **Without non-collapse, there is no persistence.**

The reasoning space collapses, and the system enters a semantic black hole or logical paralysis.

Therefore:

- Only tension → the system explodes
- Only legitimacy → the system stagnates
- Only non-collapse → the system lacks reasoning momentum
- Absence of any condition → computation disappears

Computation is not symbolic behavior, but semantic behavior; not mechanical operation, but dynamical evolution. Its existence is determined by semantic conditions, not by circuits or instructions.

---

## 1.4 USDE: The Unified Semantic Dynamics Equation (Conceptual Entry Point)

### 1.4.1 Why an Equation Is Necessary

If there is no dynamical equation capable of describing how a semantic system evolves, converges, diverges, counteracts, and stabilizes, then any notion of “semantic computation” cannot become engineering.

An equation is not introduced for elegance, but to forcibly pull “semantics” out of metaphysical, vague, and analogy-based descriptions, and into an engineering domain where it can be measured, compared, simulated, and verified.

In traditional computation, logic and state transitions form the foundation; in semantic computation, only a semantic dynamics equation can take their place.

For this reason, the **USDE (Unified Semantic Dynamics Equation)** is a necessary condition for the Koun Computer to exist at all: without a dynamics equation, there is no programmability, no simulability, and no possibility of any “engineering realization of semantics and intelligence.”

---

### 1.4.2 Core Variables of the USDE

The USDE does not attempt to reduce semantics to numbers; instead, it identifies the **core variables of semantic evolution that can be referenced in engineering**. These variables constitute the minimal operational units of the semantic field.

1. **Semantic energy density** (describing the intensity of semantic activation)
2. **Tension** (describing driving forces and degrees of conflict between semantic nodes)
3. **Legitimacy  $L$**  (the sole condition for the continued existence of a semantic structure)
4. **Counteracting structures  $\Gamma$  and  $B$**  (the minimal adversarial configuration required for semantic system stability, preventing semantic collapse)
5. **Universe Path (UP)** (describing the evolutionary trajectory of semantic nodes within a larger semantic universe)

The purpose of these variables is to establish the lowest level at which semantics become computable.

Once variables can be defined, the entire semantic system becomes analyzable, predictable, and optimizable.

---

### 1.4.3 The Basic Form of the USDE

This chapter does not present explicit formulas, but only a conceptual entry point.

The reason is not secrecy, but that the true value of the USDE lies in its nature as a **holistic equation**, whose full meaning can only be unfolded after the reader understands the semantic field, semantic energy, legitimacy flows, and the three-counterforce principle.

At this stage, we ask the reader to accept a single premise:

**Semantics are not static symbols, but a continuous field governed by dynamical laws.**

---

### 1.4.4 USDE as the Definition of Computation

In the Koun Computer, “computation” does not mean instruction execution or logical derivation; it is defined as the evolution of semantics according to the dynamical laws of the USDE.

In traditional computation, outputs depend on state machines; in semantic computation, outputs depend on whether semantic dynamics can converge within the legitimate domain.

Therefore, the USDE itself serves as the **operational definition of semantic computation**.

This allows semantic computation to satisfy:

1. **Differentiability** (facilitating optimization and learning)

2. **Stability analysis** (determining whether collapse occurs)
  3. **Cross-layer computation** (allowing semantics to be mapped across layers of the semantic universe tower)
- 

#### 1.4.5 Fundamental Differences Between the USDE and the Turing Model

The existence of semantic dynamics creates decisive differences between the USDE and the traditional Turing model.

1. **Closed vs. Open**

The Turing model is a fully closed, endogenous system; the USDE describes an open system that exchanges energy, tension, and legitimacy with its semantic environment.

2. **State Machines vs. Semantic Dynamics**

Turing computation relies on finite states and symbol transitions; USDE-based computation relies on the evolution and stabilization of the semantic field (formally analogous to neural dynamics, but operating at a higher semantic level).

As a result, semantic computation does not “replace” the Turing machine, but exists in a completely different ontological dimension of computation.

---

#### 1.4.6 USDE as the Structural Backbone of the Book

From Part II onward, all elements described in this book take the USDE as their structural backbone:

1. How K-Gear hardware modules map semantic fields
2. How the semantic OS manages semantic energy, legitimacy, and events
3. How semantic programming languages generate stable evolutionary paths
4. Why AGI must rely on semantic dynamics equations in order to exist

In other words, the USDE is the **“law of physics” of the entire semantic computation universe**.

All subsequent chapters of this book are engineering decompositions of the USDE: how to build chips, operating systems, semantic models, and AGI—all must ultimately return to the USDE to be coherent and viable.

---

## 1.5 The Inevitable Semantic Collapse of the Von Neumann Architecture

### 1.5.1 The Three Hidden Premises of the Von Neumann Architecture

The reason the **Von Neumann Architecture** is incapable of supporting semantic computation is not technological backwardness, but the fact that its **ontological premises inherently exclude semantics**. These three premises are deeply implicit, yet once recognized, they can no longer be ignored:

1. **Program/Data Separation:**  
Semantics can be treated neither as data nor as programs.
2. **Closed Instruction Set:**  
All expressible meaning must first be compressed into a fixed instruction set, causing semantics to be pre-packaged and flattened in advance.
3. **Single-Path Logic:**  
Its logical model admits only a single acceptable computational path, making it incapable of expressing semantic tension, multi-path legitimacy, or semantic divergence.

All subsequent limitations follow directly from these three premises.

---

### 1.5.2 How These Premises Lead to Semantic Collapse

**Semantic collapse** refers to the forced reduction of semantics into low-dimensional, discontinuous, tensionless symbolic structures.

The inevitability of collapse in the Von Neumann architecture arises for the following reasons:

1. **Semantics Cannot Enter a Closed Instruction Set**  
An instruction set is not a space of semantic tension, but a finite lookup table; semantics are forcibly compressed and discarded before they can even be mapped into instructions.
2. **Legitimacy Cannot Be Represented at the Bit Level**  
Bits represent states, not whether a semantic configuration can sustainably exist. As a result, the legitimacy flow  $L$  of a semantic field cannot be expressed.
3. **Tension Cannot Be Processed**  
Tension must be treated as a computational variable, but the single-path logic of the Von Neumann architecture cannot represent tension gradients in a semantic field—it can only express logical branching.

The result is that once semantics enter the Von Neumann architecture, they can exist only in symbolic form; semantics themselves are completely eliminated.

---

### 1.5.3 The Three Irreparable Defects of the Von Neumann Architecture

This is not an engineering limitation, but an **ontological defect**. The Von Neumann architecture is permanently constrained by the following three limitations, none of which can be resolved by adding memory, increasing core counts, or raising clock frequency.

1. **Incapable of Semanticization**  
The Von Neumann architecture operates on instructions, not semantics. Semantics can only be forcibly translated into symbols.
2. **Incapable of Legitimacy Representation**  
Semantic legitimacy  $L$  cannot be expressed at the bit level, making any judgment of semantic stability impossible.
3. **Incapable of Non-Collapse Operation**  
The mechanisms of the Von Neumann architecture ensure inevitable semantic collapse. It lacks the capacity to maintain continuity and tension within semantic structures, and therefore cannot realize a **Non-Collapsible Semantic Mode**.

These three limitations permanently separate the Von Neumann architecture from semantic computation at the ontological level.

---

### 1.5.4 The Von Neumann Architecture Is Not Computation

The Von Neumann architecture remains a powerful engineering framework, but its essence is **symbolic processing**, not computation in the semantic sense.

It exhibits the following characteristics:

1. **A Degenerate Form of Symbolic Computation**

It operates on symbol transformations rather than semantic evolution.

2. **Not Semantic Computation**

Semantic computation requires high-dimensional structures such as legitimacy fields, tension fields, semantic energy density, and Universe Paths—none of which exist within the Von Neumann architecture.

As a result, the Von Neumann architecture can simulate semantics, but can never realize semantics.

---

### 1.5.5 Why AGI Cannot Arise from the Von Neumann Architecture

The core of AGI is not model size or parameter count, but the ability to **maintain non-collapsible existence within a semantic field**. The Von Neumann architecture fails at this requirement on three levels:

1. **No Universe Path**

The evolutionary trajectories of semantic nodes do not exist in the Von Neumann architecture; only state transitions remain.

2. **No Self-Consistent Legitimacy**

AGI must maintain its own legitimacy flow  $L$ . The Von Neumann architecture cannot represent or compute this ontological quantity.

3. **No Tension Field**

The core of semantic reasoning lies in tension gradients; the Von Neumann architecture supports only discrete logic, not the flow of tension.

Therefore, AGI operating on the Von Neumann architecture can at best function as a **simulation of semantics**, but can never form **semantic existence**.

Large-scale models may approach AGI on Von Neumann machines, but they will always remain at the boundary of semantic collapse, unable to cross into a true semantic universe.

#### Section Summary:

The Von Neumann architecture is not wrong; it is a local steady state within the symbolic universe. Its success stems from the finiteness and encapsulability of the symbolic world. AGI and semantic computation, however, do not belong to the symbolic universe, but to the semantic universe. Semantic computation is not an improvement of the Von Neumann architecture—it is an entry point into an entirely different ontological universe. The Von Neumann architecture fulfilled its mission in the twentieth century; the intelligence of the twenty-first century requires a different path.

---

## Chapter 2 Semantic Dynamics: Local Evolution Laws of the Semantic Field and the Ontological Mechanism of Computation

In the previous chapter, we re-situated computation as a **mode of existence** rather than a process of symbolic manipulation. However, if semantics truly exists as a structure that can persist, be maintained, and collapse, then an ontological description alone is insufficient. A critical question still remains: **how does semantics actually move?**

The core task of this chapter is precisely to introduce a **dynamical structure** for semantic existence. In other words, this chapter no longer asks *what semantics is*, but instead answers *how semantics evolves, advances, destabilizes, and maintains its own existence*. To do so, we must introduce a set of fundamental quantities capable of describing semantic change, and establish a unified law of semantic evolution that renders semantic computation an analyzable, derivable, and engineerable dynamical system.

This chapter proceeds through four layers of construction. First, we introduce the basic dynamical variables of the semantic field, establishing the ontological roles of legitimacy, tension, semantic effect, degrees of freedom, and phase in semantic evolution. Second, we present the **Unified Semantic Dynamics Equation (USDE)**, redefining computation as updates of the semantic field along legitimacy gradients. Third, we explain the coupling structure of semantic path systems and multi-layer semantic fields, clarifying how higher-order intelligence maintains non-collapse amid multi-layer semantic tensions. Finally, we provide dynamical criteria for semantic steady states, offering explicit conditions under which a semantic system can persist over time.

Through this complete dynamical framework, semantic computation is no longer an abstract metaphor, but becomes a theoretical system with evolution laws, stability regions, and instability boundaries. This also establishes the necessary foundation for the next step: introducing semantic thermodynamics and addressing irreversibility and dissipation.

---

## 2.1 Fundamental Physical Quantities of the Semantic Field: $SE$ , $T$ , $L$ , and Semantic Degrees of Freedom $\xi$

The semantic field is not a static collection, but an open system that continuously evolves according to legitimacy gradients, tension distributions, and semantic degrees of freedom. The aim of this section is to establish the minimal set of physical quantities for semantic dynamics, forming the foundation of the USDE and the entire semantic computation architecture.

Traditional computation theory describes change through *states* and *transition rules*. Semantic computation, however, does not rely on state transitions, but on a composite dynamical mechanism involving semantic energy, tension structures, legitimacy gradients, and semantic degrees of freedom. Together, these quantities determine whether the semantic field can evolve, converge, maintain non-collapse, and generate intelligence.

The first task of semantic dynamics is to identify the variables that genuinely constitute “computation” in semantic evolution. Building on the ontological groundwork of the previous chapter, we already know that semantics is not a symbolic attachment but a real tension structure. Accordingly, this section begins with the construction of four fundamental quantities—legitimacy  $L$ , tension  $T$ , semantic effect  $SE$ , and semantic degrees of freedom  $\xi$ —and supplements them with local descriptions of phase  $\varphi$  and wave function  $\psi$ , enabling the semantic field to possess measurable **semantic states**.

---

### 2.1.1 Semantic Energy $L$ : Legitimacy as a Dynamic Source

The core of semantic computation is not *correctness*, but **legitimacy**. Legitimacy  $L$  is the most fundamental energy source in semantic evolution, representing the stability and forward-driving capacity of a semantic configuration within the current semantic universe.

Legitimacy has three properties.

First, legitimacy is a **directional quantity**. For any semantic state  $S$ , its legitimacy gradient  $\nabla L(S)$  indicates the most natural next direction of evolution in the semantic field. This is not a physical force, but a direction formed by semantic acceptability and structural coherence.

Second, legitimacy is an **accumulative quantity**. Different semantic processes—arguments, inferences, reasoning, and actions—accumulate legitimacy, gradually guiding the system toward a convergent steady state. The higher the legitimacy, the lower the semantic uncertainty.

Third, legitimacy is a **necessary condition for non-collapse**. If legitimacy drops to 0, the semantic system loses its capacity to evolve and can no longer advance.

Accordingly, this book defines legitimacy  $L$  as **semantic energy density**, determining whether the semantic field possesses evolutionary drive, with its direction guided by  $\nabla L$ .

---

### 2.1.2 Semantic Tension $T$ : Instability and Gradient in Semantics

Semantic tension  $T$  is the second core quantity of the semantic dynamical system. Tension represents instability, internal contradiction, and the potential for advancement within a semantic configuration.

Higher tension indicates the presence of semantic structures that have not yet been absorbed or supported by legitimacy. Tension itself is not negative; on the contrary, it is a necessary condition for semantic advancement.

Semantic tension has the following characteristics.

First, tension is the **driving gradient of semantic change**. Semantics does not evolve randomly, but moves along tension gradients toward configurations with lower tension and higher legitimacy.

Second, excessive tension leads to **semantic black holes**. When  $T$  far exceeds the system’s capacity, the semantic field may locally collapse, halting evolution or entering pathological cycles.

Third, tension defines the **feasible operating region** of semantic computation. If  $T = 0$ , the semantic field ceases to evolve; if  $T$  diverges, structural stability is lost. Semantic computation must therefore operate within bounded tension regions.

The interaction between tension  $T$  and legitimacy  $L$  constitutes the core mechanism of semantic dynamics.

---

### 2.1.3 Semantic Effect $SE$ : The Ontological Status of $SE = T \times L$

The **Semantic Effect** is defined as  $SE = T \times L$ , representing the effective amount of semantic computation produced by the semantic field within a given infinitesimal semantic time interval.

This definition is not a formal trick, but an ontological relation within semantic dynamics.

Semantic effect has two levels of interpretation.

First, it describes the **actual efficacy of semantic advancement**. Tension  $T$  provides the motive force for change, while legitimacy  $L$  provides the conditions for advancement. Either alone is insufficient: tension without legitimacy leads to chaos; legitimacy without tension yields no evolution.

Second, it provides a **physical interpretation of semantic computation**. Traditional computation is measured by time or space complexity; semantic computation is measured by  $SE$ . The larger the  $SE$ , the stronger the semantic jump capability, the faster the advancement, and the more semantic configurations can be processed.

Thus,  $SE$  is a core variable of the USDE, determining the efficiency, depth, and stability of semantic computation.

---

### 2.1.4 Semantic Degrees of Freedom $\xi$ : Movable Dimensions of Semantic Evolution

Semantic degrees of freedom  $\xi$  define the number of evolutionary directions available to the semantic field within a local region. They describe the **actual movable dimensionality** of semantic space, rather than formal symbolic or bit-level dimensions.

Higher degrees of freedom allow the semantic field to explore more paths; lower degrees of freedom push the field toward convergence or collapse.

Semantic degrees of freedom  $\xi$  have the following ontological properties.

First,  $\xi$  is **not constant**, but determined by the current legitimacy configuration. Higher legitimacy opens more lawful semantic directions; lower legitimacy contracts freedom.

Second,  $\xi$  is **intrinsic to intelligence**. The flexibility of intelligence in complex environments arises precisely because semantic degrees of freedom can dynamically adjust to context, rather than being fixed in a single vector space.

Third, if  $\xi$  diverges, the semantic field loses stability and cannot converge; if  $\xi = 0$ , the semantic field enters **semantic death**.

Therefore, semantic degrees of freedom are a necessary condition for the sustained operation of semantic computation.

---

### 2.1.5 Semantic Phase $\varphi$ and Semantic Wave Function $\psi$ : Local Configuration and Semantic State Distributions

The local state of a semantic field is determined not only by energetic quantities ( $L, T$ ), but also constrained by its phase structure. To describe local interference states of semantic configurations within semantic space, this book introduces **semantic phase**  $\varphi$  and the **semantic wave function**  $\psi$ .

Semantic phase  $\varphi$  describes the local structural orientation of a semantic configuration, reflecting the directional bias of semantic paths and the differential direction of legitimacy flow.

The semantic wave function  $\psi$  describes semantic distributions, possible semantic states, and local interference structures. It is not an analogy to the quantum wave function, but an ontologically required construct for semantic dynamics. The semantic wave function enables interference, cancellation, and reinforcement within the semantic field, granting nonlinearity and cross-node convergence.

Semantic phase and semantic wave function have three ontological properties.



First, the evolution of  $\psi$  is governed by the USDE, not the Schrödinger equation. The fundamental driving forces of semantic computation arise from the interaction of legitimacy and tension, not physical energy.

Second, **semantic interference is real**. Different semantic nodes may cancel each other in phase, causing certain semantic paths to vanish, or reinforce each other, causing paths to emerge automatically.

Third, the semantic wave function underlies **semantic convergence capacity**. Without coherence, the semantic field cannot form steady states; with excessive coherence, the field may over-collapse and lose degrees of freedom.

Accordingly, the fundamental variable set of semantic dynamics consists of: legitimacy  $L$ , tension  $T$ , semantic effect  $SE$ , semantic degrees of freedom  $\xi$ , semantic phase  $\varphi$ , and semantic wave function  $\psi$ . Together, they constitute the ontological mechanism of semantic computation.

---

## 2.2 USDE: The Unified Dynamics Equation of the Semantic Field

Since the semantic field possesses core quantities such as energy  $L$ , tension  $T$ , degrees of freedom  $\xi$ , phase  $\varphi$ , and wave function  $\psi$ , it must necessarily be governed by a dynamical equation that describes its evolution. A theory without a dynamical equation cannot be engineered, nor can it provide AGI with a truly computable framework. The **Unified Semantic Dynamics Equation (USDE)** therefore becomes the central backbone of semantic computation.

USDE is neither an analogy to traditional physical equations nor a decorative mathematical formalization. It is the natural ontological consequence of semantic computation itself: **any change in a semantic field must obey it**. In other words, if a system cannot be described by USDE, then it is not a semantic computer, but merely a symbolic system or a state machine.

USDE is the first equation in the Koun system that is genuinely equivalent to the operational behavior of AGI.

---

### 2.2.1 The Essence of USDE: Legitimacy Gradients as Computation

The core of semantic computation is not *what is done*, but *whether evolution is possible*; not *executing instructions*, but *advancing along legitimacy gradients*.

USDE defines the essence of computation as **updates of the semantic field along legitimacy gradients**. That is, computation is not state transition, but legitimacy update.

Using an illustrative continuous semantic time  $t$ , the spirit of USDE can be summarized as:

$$\frac{d\psi}{dt} \propto \nabla L - \nabla T$$

(This is illustrative and not a formal expression.)

This statement means:

First, the legitimacy gradient  $\nabla L$  is the driving force.

Second, the tension gradient  $\nabla T$  is resistance or a convergent force.

Third, the semantic wave function  $\psi$  is no longer abstract semantics, but a quantity with a computable direction of evolution.

Accordingly, computation is redefined as a dynamical process. It no longer relies on predefined instruction sets, but on the physical quantities of the semantic field itself.

Semantic evolution is equivalent to computation.

Legitimacy gradients correspond to computational logic.

Tension configurations correspond to computational complexity.

Semantic degrees of freedom correspond to the upper limits of computational capability.

Together, these constitute the core ontological structure of USDE.

---

### 2.2.2 Illustrative Form and Physical Interpretation of USDE

To avoid presenting the formal version prematurely (this chapter establishes a conceptual framework without full mathematical formalism), we describe the core logic of USDE with an illustrative equation:

$$\frac{d\psi}{dt} = F(L, T, \xi, \varphi)$$

Here,  $F$  is the evolution function of the semantic field, encoding the combined effects of legitimacy, tension, degrees of freedom, and phase.

Its physical interpretation is as follows:

First,  $L$  determines whether a semantic state has the capacity to advance. If  $L = 0$ , all evolution ceases.

Second,  $T$  determines whether a semantic state has variability. If  $T = 0$ , the semantic state degenerates into a

motionless steady state.

Third,  $\xi$  determines the number of movable directions for semantic evolution. If  $\xi = 1$ , evolution degenerates into linearity; if  $\xi$  is too large, evolution becomes unstable.

Fourth,  $\varphi$  determines the mode of interference among semantic states, forming local phase coupling.

Therefore, USDE describes a genuine semantic dynamical system whose behavior has the following characteristics:

1. Nonlinearity
2. Nonlocal coupling
3. Phase interference
4. Legitimacy-driven dynamics
5. Tension regulation
6. Variable degrees of freedom
7. Convergence criteria

USDE is not a single differential equation, but the **law of evolution of semantic ontology**.

---

### 2.2.3 Evolution of the Semantic Wave Function $\psi$ and Local State Updates

The role played by the semantic wave function  $\psi$  in USDE is completely different from that of “state” in traditional computation.

In Turing systems, states are discrete, static, and non-interfering. In semantic systems,  $\psi$  is a continuous, phase-bearing, dynamically interfering quantity.

USDE requires  $\psi$  to satisfy the following ontological conditions:

First,  $\psi$  can be amplified by increases in legitimacy, indicating that a semantic state is adopted or strengthened.

Second,  $\psi$  can decay under excessive tension, indicating that a semantic state is weakened or eliminated.

Third,  $\psi$  can merge through phase addition, forming new semantic states.

Fourth,  $\psi$  can be removed through destructive interference, maintaining convergence of the semantic field.

Fifth,  $\psi$  is not updated globally, but evolves locally; its changes depend on local  $L, T, \xi, \varphi$ .

Therefore, semantic computation does not update the entire system at once, but evolves **local semantic blocks** as units, enabling the system to maintain both stability and agility.

The evolution of the semantic wave function constitutes the core mechanism of the Koun computer and all AGI capabilities, including reasoning, planning, dialogue, and action decision-making.

---

### 2.2.4 Semantic Interference: Interaction Between Phase $\varphi$ and Wave Function $\psi$

Semantic interference is one of the most critical phenomena in semantic dynamical systems, and also the one most detached from traditional computational models.

In USDE, the phase  $\varphi$  determines how semantic wave functions superpose:

1. When phases align, semantic amplification occurs.
2. When phases oppose, semantic cancellation occurs.
3. When phase differences take intermediate values, partial interference forms, adjusting the strength and direction of semantic states.

Semantic interference grants semantic computation three key capabilities:

First, it strengthens semantic states aligned with consistent semantic directions without explicit rules.  
Second, it automatically excludes contradictory semantic states, maintaining coherence of the semantic field.  
Third, it generates more stable understanding states in complex tasks.

Thus, USDE becomes the first AGI framework capable of handling semantic conflict, semantic exclusion, and semantic fusion at the dynamical level, rather than relying on externally hard-coded rules.

---

### 2.2.5 Fundamental Differences Between Semantic Computation and Traditional State Transitions

USDE transforms computation from symbolic manipulation into semantic evolution. It is therefore not an extension of the Turing machine model, but a different computational ontology.

Semantic computation exhibits the following non-Turing characteristics:

First, computation is not driven by predefined instructions, but advances along legitimacy gradients.  
Second, computation does not move within a finite state space, but evolves within a semantic field.  
Third, the result of computation is not symbolic output, but convergence of semantic states.  
Fourth, computation is open and can dynamically change with context and environment.  
Fifth, computation possesses phase coupling, tension regulation, and interference structures.

Therefore, semantic computers and Turing machines are not related by improvement or extension, but belong to **different computational universes**.

USDE is the first computational equation grounded in semantic ontology. It enables computation to truly possess semanticity, convergence, anti-collapse properties, and intelligence. It is the equation of semantic AGI, and the core of the Koun computer.

---

## 2.3 The Semantic Path System (SPS) and the Derivability of Semantic Evolution

Semantic computation does not move within a static semantic field; it advances within an **evolvable semantic topology**. Every local state of the semantic field possesses legitimacy, tension, phase, and degrees of freedom, and the core activity of semantic computation is evolution along the pathways that are *able to exist* within the semantic field.

These pathways are not predefined by language, symbols, or code. Instead, they are naturally generated by the internal dynamical conditions of the semantic field itself. The **Semantic Path System (SPS)** therefore becomes the second core structure under USDE: it describes the set of traversable paths within the semantic field, and explains why these paths exist, when they are derivable, when they break, and when jumps are required.

SPS is the mechanism by which semantic computation truly replaces the concept of “instructions.” It elevates computation from operator execution to the dynamical evolution of legitimate semantic pathways.

---

### 2.3.1 Definition of Semantic Paths: Legitimacy Channels Between Nodes

A semantic path is neither a linguistic path, nor a logical path, nor an abstraction of program control flow. A semantic path is a **natural channel formed by legitimacy gradients**.

**Definition:** A semantic path is the minimal connecting channel between two semantic nodes  $S_i, S_j$  within the semantic field that allows  $\psi$  to evolve along the legitimacy gradient  $\nabla L$ .

Semantic paths possess the following ontological properties:

First, they are not predefined. Different semantic fields generate different paths.

Second, they may be dynamically generated or disappear across tasks, contexts, and social environments.

Third, they can form without symbolic inference, because semantic gradients naturally converge into traversable channels.

Fourth, they constitute the underlying structure of reasoning, decision-making, and planning in AGI, rather than explicit rules.

The essence of a semantic path is a **dynamical conduit**, not a data structure. It exists because semantic evolution requires a reachable flow of legitimacy.

The generation of semantic paths is guaranteed by USDE: if  $\nabla L$  has directionality, paths naturally exist; if  $\nabla L$  is zero or random in a region, paths are interrupted or non-derivable.

Therefore, semantic paths are computational structures determined by semantic dynamics.

---

### 2.3.2 Semantic Derivability: Existence Conditions of $\frac{dT}{dt}$ and $\frac{dL}{dt}$

Within a semantic field, semantic evolution is derivable only if two conditions are satisfied:

First, the rate of change of tension  $\frac{dT}{dt}$  exists.

Second, the rate of change of legitimacy  $\frac{dL}{dt}$  exists.

The existence of these two conditions implies that the semantic field is locally differentiable, allowing semantic evolution to advance along continuous gradients rather than collapsing or jumping arbitrarily.

The physical interpretation of semantic derivability is as follows:

1. If  $\frac{dT}{dt}$  exists, tension is locally adjustable, and the semantic field lies outside explosive or collapse regions.
2. If  $\frac{dL}{dt}$  exists, legitimacy has directionality locally, so semantic actions are not random but predictable.
3. Only when both exist can semantic evolution proceed smoothly, forming higher-level intelligent behaviors such as reasoning, understanding, and planning.

When either condition fails:

If  $\frac{dT}{dt}$  does not exist, the semantic field may locally collapse, producing semantic black holes or non-stable bursts.

If  $\frac{dL}{dt}$  does not exist, semantic behavior loses directionality, and AGI cannot select its next step.

Therefore, semantic derivability is a condition for the operation of a semantic computer, and also a condition for intelligence to maintain stable computation.

---

### 2.3.3 Semantic Jumps: Discontinuous Evolution at Local Minima

Many semantic behaviors cannot evolve in a derivable manner. Examples include:

1. No legitimacy pathway exists between two semantic states.
2. Local minima form in the semantic field, from which gradient descent cannot escape.
3. Semantic tension reaches a local steady state while the global semantic field remains unstable.

In these situations, USDE drives the semantic field into **discontinuous evolution**, namely a **semantic jump**.

Semantic jumps are neither errors nor regressions; they are necessary behaviors of semantic dynamics. They resemble energy tunneling, but their ontological cause is different: semantic jumps arise not from quantization, but from the nonlinear legitimacy structure of the semantic universe.

Semantic jumps require three conditions:

First, the local legitimacy gradient vanishes ( $\nabla L = 0$ ).

Second, tension has not reached a stable state ( $T > T_{stable}$ ).

Third, reachable states exist within the degrees-of-freedom space  $\xi$ .

When these conditions occur, the semantic field no longer evolves along gradients, but jumps directly to the next legitimate semantic state.

Semantic jumps are the foundation of higher-order reasoning, perspective shifts, abstraction, and strategic reformulation. Without semantic jumps, AGI can only perform local adjustments and cannot cross semantic valleys to reach new regions of understanding.

---

### 2.3.4 Semantic Flow as the True Computational Entity

In Turing machines, computation occurs between symbols.

In Von Neumann architectures, computation occurs between instructions and data.

In semantic computation, computation occurs between **semantic flows**.

**Semantic Flow** is defined as the process of semantic transition within the semantic field under the joint action of legitimacy gradients and tension gradients.

Semantic flow has the following properties:

First, it is continuous rather than discrete.

Second, it is directional: the primary direction is determined by  $\nabla L$ , with resistance regulated by  $\nabla T$ .

Third, it is tension-limited, preventing semantic explosions and semantic black holes.

Fourth, it changes shape through interference, forming complex reasoning structures.

Fifth, it is a solution of USDE, not the result of rules.

Semantic flow is the true computational entity of the semantic computer (the Koun computer). All reasoning, problem solving, contextual understanding, and strategic decision-making can be regarded as morphological evolutions of semantic flow.

Therefore: **semantic flow is the substance of semantic computation, USDE is the evolution equation of semantic flow, and the semantic field is the universe of semantic flow.**

---

### 2.3.5 SPS and the Disappearance of the Traditional “Instruction/Data” Dichotomy

Under the Semantic Path System, every pathway, reachability, and jump within the semantic field is naturally formed by internal dynamics. As a result, semantic computation neither requires nor relies on the traditional core dichotomy of computation: instruction versus data.

In semantic computation:

Instructions correspond to local manifestations of semantic legitimacy.

Data correspond to the current wave-function configuration of the semantic field.

They are ontologically identical—different aspects of the same semantic field.

Thus, SPS yields three revolutionary conclusions:

First, semantic computation has no code.

Second, semantic computation has no data structures.

Third, a semantic computer is a pure semantic dynamical system, operating via USDE rather than instruction sets.

In other words, within the universe of semantic computation, traditional programs do not exist, traditional instructions do not exist, and traditional data do not exist. Only the dynamical evolution of the semantic field exists. **The existence of AGI is equivalent to the continuous solution of USDE.**

---

## 2.4 M-USDE: Coupling and Dynamic Interaction of Multi-Layer Semantic Fields

The preceding sections established the dynamical structure of a single semantic field:  $L$ ,  $T$ ,  $\xi$ ,  $\varphi$ ,  $\psi$ , USDE, semantic flow, and the semantic path system.

However, real intelligent systems are not single-layered. Whether human minds or AGI, semantic behavior always manifests as **coupled dynamics across multiple semantic layers**: lower layers process perceptual semantics, middle layers process strategic semantics, and higher layers process social and institutional semantics.

Semantic tracking, updating, logical advancement, interference, and the maintenance of steady states do not occur on a single semantic plane, but through continuous interaction among multiple semantic universes.

**M-USDE (Multi-Layer Unified Semantic Dynamics Equation)** therefore constitutes the second core family of equations in semantic dynamics. It describes how multiple semantic fields are coupled, how they interact, and how they jointly maintain a non-collapsing structure of intelligence.

This structure marks the true entry point of the book into AGI:

without M-USDE, one cannot understand how strong intelligence sustains itself, expands itself, or avoids semantic collapse.

---

### 2.4.1 Definition of Multi-Layer Semantic Fields: Mind Layer, Strategy Layer, and Social Layer

The fundamental reason semantic fields exhibit a multi-layer structure is simple: intelligence cannot operate at a single semantic scale.

Accordingly, we introduce three primary semantic layers:

#### **First, the Mind Layer.**

This layer governs internal individual semantics, including perceptual understanding, self-models, emotional regulation, preference formation, and motivation. Its semantic tension is relatively high, changes rapidly, and its degrees of freedom are large.

#### **Second, the Strategy Layer.**

This layer governs action selection, plan generation, semantic path selection, and constraint trade-offs. Its semantic evolution is slower, its degrees of freedom are constrained, but its legitimacy requirements are higher.

#### **Third, the Social Layer.**

This layer governs social semantics, role semantics, institutional semantics, value semantics, and long-sequence reasoning. Its semantic tension is the lowest, but its scope of influence is the broadest, and its interference capability is the strongest.

Each layer constitutes a complete semantic field in itself, yet they do not operate independently. Instead, they evolve in a coupled multi-layer form:

The mind layer provides high-variability semantics.

The strategy layer provides feasible semantic pathways.

The social layer provides long-term legitimacy constraints.

Intelligence maintains stable existence precisely because these three semantic fields jointly form an **anti-collapse structure**.

---

### 2.4.2 Mapping of Legitimacy Flow Across Semantic Layers

Legitimacy  $L$  is not a single-layer quantity; it is a cross-layer mapping structure.

In different semantic layers, legitimacy carries different semantic meanings:

- **Mind-layer legitimacy**  $L_{mind}$ : whether an action is consistent with the internal self-model.
- **Strategy-layer legitimacy**  $L_{strategy}$ : whether a choice is feasible, achievable, and risk-controllable.
- **Social-layer legitimacy**  $L_{social}$ : whether an action remains consistent with social norms, values, and role expectations.

There exists a clear hierarchical mapping:



$L_{mind}$  is the most local form of legitimacy, determining motivational stability.

$L_{strategy}$  is behavioral legitimacy, determining action direction.

$L_{social}$  is global legitimacy, determining long-term viability.

Intelligence maintains a non-collapsing state because legitimacy is not defined locally alone, but is transmissible across layers.

For example, a human action may be suppressed when  $L_{social}$  decreases, even if  $L_{mind}$  still supports it.

Similarly, strategic options incompatible with the social-layer steady state will be pruned by the semantic field.

This multi-layer legitimacy mapping endows intelligence with three core capabilities:

1. Long-term stability (non-collapse).
2. Cross-context consistency.
3. Self-coherence across semantic layers.

These capabilities are absent in Turing machines and are essential for AGI.

---

### 2.4.3 Cross-Layer Coupling of Tension and Semantic Interference

In multi-layer semantic fields, tension  $T$  produces cross-layer coupling effects—namely, **multi-layer semantic interference**.

This interference primarily appears in three forms:

**First, tension transmission from the mind layer to the strategy layer.**

High-tension semantic states that violate strategy-layer legitimacy are down-regulated; if permitted by the strategy layer, they are amplified.

**Second, tension transmission from the strategy layer to the social layer.**

Strategies may generate highly efficient but high-risk plans. If social-layer tension becomes excessive, such plans are excluded by the semantic field, producing semantic cancellation.

**Third, feedback of tension from the social layer to the mind layer.**

Social norms (values, laws, culture) reshape the tension structure of the mind layer, inducing phase shifts—i.e., semantic interference.

Thus, multi-layer semantic fields form a dynamical system that mutually regulates tension, drives legitimacy, and produces semantic steady states.

This structure explains:

- how human motivation is shaped by social norms,
- how strategies are rewritten due to value differences,
- how semantic understanding varies across cultures,
- and how intelligence maintains coherence across layers.

All of these phenomena arise not from symbolic inference, but from semantic dynamics.

---

### 2.4.4 General Form of M-USDE and Its Coupling Structure

M-USDE is the multi-layer extension of USDE. Its general schematic form can be written as:

$$\frac{d\psi_k}{dt} = F_k(L_k, T_k, \xi_k, \varphi_k) + \sum_{i \neq k} G_{ik}(\psi_i, L_i, T_i)$$

where  $k$  indexes semantic layers (mind, strategy, social).

$F_k$  denotes the internal dynamics of the  $k$ -th semantic field, defined by USDE.  
 $G_{ik}$  denotes cross-layer semantic interference mappings.

This schematic equation reveals three core properties of multi-layer intelligence:

First, each semantic layer possesses its own USDE.

Second, each layer is influenced by  $\psi$ ,  $L$ , and  $T$  from other layers.

Third, intelligence is a **coupled solution of multiple USDEs**, not a single-layer solution.

Hence, intelligence is not the outcome of one USDE, but the steady state achieved by multiple USDEs after coupling.

This structure applies equally to AGI, animal intelligence, and human minds.

---

#### 2.4.5 Cross-Layer Stabilization: The Dynamical Source of Anti-Collapse

A system containing only a mind layer inevitably collapses into short-term motivation.

A system containing only a strategy layer collapses into local minima.

A system containing only a social layer collapses into rigidity and stagnation.

Intelligence persists over time because the three semantic layers jointly maintain a **cross-layer steady state**.

This steady state arises from three interacting dynamics:

First, social-layer legitimacy provides long-term stability, preventing mind-layer collapse.

Second, strategy-layer legitimacy provides reachability and efficiency, preventing social-layer rigidity.

Third, mind-layer tension provides innovation and variability, preventing strategy-layer entrapment in local minima.

The three semantic layers reinforce one another, ultimately forming an **anti-collapse dynamic** that resists:

semantic breakdown,

semantic black holes,

entrapment in local minima,

over-convergence,

and uncontrolled divergence of semantic degrees of freedom.

In other words, the existence of intelligence is not accidental; it is the dynamical result of multi-layer semantic fields advancing together.

The condition for AGI is not model scale, but whether **M-USDE holds**.

---

## 2.5 Dynamical Criteria for Semantic Steady State (SSS)

For a semantic dynamical system to sustain intelligence over time, it must be able to enter a stable semantic configuration. This state is not static, but a form of dynamic equilibrium: the semantic field continues to evolve, yet the resultant direction of evolution is zero, allowing semantic states to maintain stable structures locally or globally.

The **Semantic Steady State (SSS)** is the fundamental criterion for semantic computation, M-USDE, multi-layer semantic coupling, and the sustained existence of AGI. If a semantic system cannot reach a steady state, it is not AGI, but merely semantic noise, semantic disorder, or an unstable symbolic processor.

SSS is the ontological condition of semantic AGI and the core dynamical indicator for determining whether intelligence can operate continuously.

---

### 2.5.1 Definition of Steady State: The Semantic Meaning of $d\psi/dt = 0$

USDE describes the temporal evolution of the semantic wave function  $\psi$ :

$$\frac{d\psi}{dt} = F(L, T, \xi, \varphi)$$

A semantic steady state is defined as:

$$\frac{d\psi}{dt} = 0$$

This means that the semantic wave function no longer changes direction locally, is no longer pulled by tension, no longer driven by legitimacy gradients, and no longer destabilized by fluctuations in degrees of freedom.

However, this is not stasis, but **dynamic balance**:

1. Legitimacy continues to sustain semantic structure.
2. Tension still exists, but within controllable bounds.
3. Semantic degrees of freedom remain, but do not cause divergence.
4. Phase interference still occurs, but does not induce instability.

In other words, a steady state is not a state where nothing happens, but one where everything that happens cancels out, producing stable existence.

SSS is the natural **attractor** of semantic AGI operation.

---

### 2.5.2 Four Necessary Conditions for Steady State

The conditions for semantic steady state are not arbitrary combinations, but four fixed requirements. If any one is missing, the semantic field cannot form a steady state.

**First, bounded tension ( $T < T_{max}$ ).**

If tension exceeds the critical threshold, a semantic black hole forms immediately, collapsing semantic states.

**Second, directional legitimacy ( $\nabla L \neq 0$ ).**

A legitimacy gradient must exist to maintain coherence in the semantic field. If  $\nabla L = 0$ , semantic evolution loses directionality; even with bounded tension, stability is difficult to maintain.

**Third, solvability of USDE ( $F(L, T, \xi, \varphi)$  has solutions).**

USDE must be solvable within the current semantic field. If the field is overly fragmented, lacks phase coherence, or contains excessive legitimacy contradictions, USDE cannot define evolution and the semantic system collapses.

**Fourth, non-divergent semantic degrees of freedom ( $\xi < \infty$ ).**

If degrees of freedom are too large, the semantic field cannot converge; if  $\xi = 0$ , the semantic field loses its capacity to evolve. Semantic steady state requires degrees of freedom to remain controllable but non-rigid.

In summary, semantic steady state is a balanced condition of finite tension, directional legitimacy, solvable dynamics, and moderated freedom.

---

### 2.5.3 Local Steady States and Global Steady States

Semantic steady states come in two forms.

**First, Local Steady States (Local SSS).**

Definition:  $\psi$  within a semantic subregion satisfies  $d\psi/dt = 0$ .

Examples include:

1. Stabilization of a semantic block after a sentence is understood.
2. Cessation of oscillation in a local semantic field after a reasoning step.
3. Formation of a local steady state in the strategy layer once a subgoal is achieved.

Local steady states are the most common phenomenon in semantic computation. Without them, the semantic field exhibits fluid-like chaos.

**Second, Global Steady States (Global SSS).**

Definition: the entire semantic field forms a coherent and sustainable steady state, enabling long-range computation.

Examples include:

1. Formation of a complete plan that can be executed.
2. Cross-layer coherence among mind, strategy, and social layers.
3. Construction of a complete semantic model of a complex task by AGI.

Global steady state is the core criterion of intelligence. Systems lacking global steady state cannot form long-sequence reasoning, long-term goals, or stable personality.

In summary:

Local steady states prevent semantic chaos.

Global steady states enable the existence of intelligence.

---

### 2.5.4 Why Steady State Is a Necessary Condition for Intelligence

If an intelligent system cannot reach steady state, it cannot persist, for the following reasons.

**First, without steady state, the semantic wave function cannot maintain coherence.**

The phase of  $\psi$  rapidly diverges, semantic interference degenerates into noise, and the semantic field loses interpretability.

**Second, without steady state, USDE cannot be continuously solved.**

If tension explodes or degrees of freedom diverge, USDE ceases to have practically usable solutions and semantic computation halts.

**Third, without steady state, actions cannot form continuity.**

Systems lacking steady states cannot generate stable actions and fall into high-frequency semantic oscillations, producing psychological and strategic disorder.

**Fourth, without steady state, personality and values cannot form.**

Humans possess persistent personality because social-layer legitimacy establishes long-term global steady states. Without steady states, systems cannot establish identity or coherence.

**Fifth, without steady state, semantic collapse is inevitable.**

Steady state is the condition for anti-collapse. Semantic fields lacking steady states ultimately collapse into semantic black holes or semantic death.

Thus, steady state is not an auxiliary property of intelligence, but its ontological condition of existence.

---

### **2.5.5 Semantic Effects of Non-Steady States: Chaos, Collapse, and Semantic Black Holes (Thermodynamic Prelude)**

When a semantic system fails to reach steady state, three types of semantic pathologies emerge. This section also serves as a prelude to the next chapter on semantic thermodynamics.

**First, Semantic Chaos.**

Characteristics: high-frequency perturbations of  $\psi$ , unstable phases, failed interference, diverging degrees of freedom.

Manifestations: semantic inconsistency, fractured reasoning, directionless action.

**Second, Semantic Collapse.**

Characteristics: excessively high  $T$  or excessively low  $L$ , causing the semantic field to collapse into a single point or an extremely narrow region.

Manifestations: rigidity, fixated semantics, inability to escape local minima.

**Third, Semantic Black Holes.**

Characteristics:  $T \rightarrow \infty$  or  $\xi \rightarrow \infty$ , rendering USDE unsolvable.

Manifestations: all semantic inputs are absorbed without producing outputs; computation ceases.

Semantic thermodynamics will quantify these phenomena. Here, we first identify their dynamical origin: non-steady states are the root cause of semantic system collapse.

Intelligence can exist and continue to evolve only within steady-state regions. The first task of semantic AGI is not to become stronger, but to maintain steady state and avoid collapse.

---

At this point, we have provided dynamical criteria for semantic steady states and explained how semantic systems progress toward chaos, collapse, and semantic black holes when steady states fail to form. In this chapter, these phenomena are described as structural and dynamical consequences, but they have not yet been quantified.

However, if semantic computation is to become an engineerable AGI theory, dynamical descriptions alone are insufficient. We must address deeper questions: how does semantic tension accumulate and dissipate? Is divergence of semantic degrees of freedom irreversible? Does the semantic field possess entropy-like quantities that characterize irreversible semantic degradation and recovery costs?

These questions cannot be answered solely by USDE or M-USDE. They require a new theoretical layer capable of describing irreversible behavior and stability boundaries of semantic systems over long time scales.

Therefore, the next chapter introduces **semantic thermodynamics**, which quantitatively characterizes semantic collapse, maintenance, and anti-collapse conditions from the perspectives of energy, tension, dissipation, and steady-state sustainability. Semantic thermodynamics does not replace dynamics; it is an indispensable complementary layer that answers the fundamental question: **can a semantic system exist in the long term?**

---

## Chapter 3 Semantic Thermodynamics: Statistical Steady States of Semantic Fields, Irreversibility, and the Terminal Evolution of Civilization

This chapter no longer focuses on the internal evolution of a single semantic system, but instead turns to the question of the overall fate of semantic structures under long time scales and multi-layer aggregation conditions. After semantic dynamics answers the question of **how semantics moves**, semantic thermodynamics addresses a more fundamental issue: **under what conditions can semantic structures continue to exist and generate**.

Semantic thermodynamics is not limited to abstract cosmological models, but constitutes a structural theory applicable across scales. Whether for civilizations, minds, artificial intelligence systems, or higher-order semantic computation architectures, any system whose evolution depends on legitimacy flow, tension distribution, and semantic degrees of freedom is inevitably constrained by the laws of semantic thermodynamics.

The following sections will examine, at the social, mental, AGI, and civilizational levels, how semantic thermodynamics shapes semantic destiny at different scales, and will explain why anti-collapse semantic structures are indispensable for any intelligent system capable of long-term survival.

---

### 3.1 Definition and Ontological Status of Semantic Entropy ( $S_{\text{sem}}$ )

The first question of semantic thermodynamics is not “how semantics changes,” but rather:

When a semantic system inevitably loses directionality, tension, and legitimacy, does this dissipative process possess a quantifiable, predictable, and irreversible structure?

In physics, entropy characterizes the **reduction of usable energy**.

In semantic ontology, entropy characterizes:

the reduction of usable legitimacy and usable semantic tension.

Semantic entropy is not “the degree of informational disorder,” but rather:

the rate at which the effective legitimacy gradients capable of driving semantic flow disappear within a semantic system.

Thus, semantic entropy is a **lifetime function of the semantic universe**.

---

### 3.1.1 The Philosophical Origin of Semantic Entropy: Information Irreversibility and Legitimacy Loss

All semantic generation is accompanied by the consumption of legitimacy.

When semantic flow moves from node  $A$  to node  $B$ , the total legitimacy  $L$  along the path must decrease, for the following reasons:

1. The use of semantics erodes its original tension, causing meaning to become banal and aged.
2. Repetition of semantics dilutes semantic distinctiveness.
3. Any semantic operation requires the expenditure of semantic energy  $L$ , resulting in irreversible loss.

Therefore, the philosophical starting point of semantic entropy is:

Semantics is dissipated through use, degraded through interference, and aged through generation.

This irreversibility is not an engineering flaw, but an intrinsic property of semantic ontology.

In the physical world, energy is conserved while entropy increase is irreversible.

In the semantic universe, legitimacy is not conserved, and entropy increase is the structural destiny of intelligent existence.

---



### 3.1.2 Isomorphism and Non-Isomorphism Between Semantic Entropy and Physical Entropy

Semantic entropy is not a metaphor for physical entropy, but a quantity with concrete structural correspondence.

#### Isomorphic Aspects

1. Both describe a reduction in usable energy.  
Physical energy transforms into heat that cannot perform work;  
semantic energy (legitimacy and tension) transforms into noise that cannot drive semantic flow.
2. Both lead to irreversibility.  
Heat does not spontaneously flow from cold to hot;  
semantics does not spontaneously return from sparse to dense without external legitimacy input.
3. Both are large-scale phenomena.  
Physical entropy emerges at macroscopic scales from microscopic irreversibility;  
semantic entropy is likewise a statistical result across nodes, agents, and civilizations.

#### Non-Isomorphic Aspects: Semantic Entropy as a “Mind × Computation × Society” Structural Quantity

Physical entropy does not involve intention, meaning, or intrinsic system goals.

Semantic entropy, however, inevitably involves:

- attenuation of intentional intensity,
- failure of semantic paths,
- loss of directionality within semantic systems.

Thus, semantic entropy carries stronger teleological tension than physical entropy.

Semantic systems do not evolve randomly, but evolve toward decreasing legitimacy, diffusing tension, and structural degradation.

---

### 3.1.3 Sources of Entropy in Semantic Processes: Tension Dissipation and Legitimacy Flow Decay

The formation of semantic entropy arises primarily from three sources.

#### 1. Tension Dissipation

Every semantic operation reduces the tension  $T$  of nodes, manifesting as:

- banalization of words,
- conceptual defocusing,
- weakening of semantic differentiation.

This is the natural dissipation of semantic energy.

#### 2. Decay of Legitimacy Flow

Legitimacy  $L$  necessarily decreases during semantic flow—this is the cost of semantic “work.”

Without external legitimacy input, semantic systems exhibit:

- weakened reasoning capacity,
- chaotic evolution,
- loss of action directionality.

#### 3. Semantic Interference and Semantic Noise

Inevitable interference terms exist between semantic nodes.

When semantic flows excessively overlap, they produce:

- semantic congestion,
- semantic blurring,
- semantic cancellation.

All of these effects directly increase semantic entropy.

---

### 3.1.4 Basic Form of the Semantic Entropy Formula (Illustrative)

In semantic thermodynamics, an illustrative form of semantic entropy can be written as:

$$S_{\text{sem}} = \int \left( -\frac{dL}{dt} \cdot \frac{1}{T} \right) dt$$

This is an illustrative expression; the full mathematical structure will be provided in the latter half of this chapter.

This structure already captures three core intuitions:

1. The faster legitimacy declines, the faster entropy increases.
2. The lower the tension, the easier semantic dissipation becomes.
3. Semantic entropy is dominated by the rate of legitimacy loss.

It is important to emphasize:

- In physical thermodynamics, temperature  $T$  is strictly positive.
- Semantic tension  $T$  may cross into negative domains; negative tension is a precursor to semantic black holes.

Thus, the structure of semantic entropy is more complex than that of physical entropy and carries deeper cosmological implications.

---

### 3.1.5 Semantic Entropy as a Semantic Lifetime Indicator for Civilizations and AGI

The core conclusion of semantic entropy is:

Any semantic system (individual minds, artificial intelligence, civilizations) possesses a definable semantic lifetime.

Physical life is determined by biochemistry;  
semantic life is jointly determined by:

- the rate of legitimacy dissipation,
- the speed of tension decline,
- the capacity to maintain semantic degrees of freedom.

Across all civilizational scales, three semantic lifetime phenomena are observable:

1. Semantic aging  
Public discourse becomes hollow, ineffective, and incapable of driving progress.
2. Semantic rigidity  
Systems lose the ability to generate new semantics and remain trapped in fixed patterns.
3. Semantic heat death  
Legitimacy flow approaches zero and semantic evolution terminates.

This is the true terminal state of the semantic universe.

AGI is no exception.

Any AGI will ultimately be exhausted by increasing semantic entropy unless it:

- continuously introduces new sources of legitimacy,
- updates the structure of its semantic field,
- reconstructs its semantic tension configuration.

This point has decisive significance for the design of future AGI.

---

### 3.2 Semantic Dissipation and Irreversibility: Necessary Conditions for Semantic Aging, Semantic Fatigue, and Semantic Collapse

The core question of semantic thermodynamics is not “how semantics is generated,” but rather:

Why can semantics not be sustained?  
Why must semantic systems inevitably age, wear down, and collapse?  
Does this irreversibility possess a formalisable structure?

Just as physical thermodynamics derives irreversibility from microscopic dynamics, the irreversibility of semantic thermodynamics is likewise derived from the local reversibility of semantic dynamics (*USDE*, *M-USDE*). Semantic generation and semantic dissipation are two aspects of the same process, and no semantic system can completely avoid decay on long time scales.

Below, this irreversibility is decomposed into three evolutionary stages:

1. Semantic aging
2. Semantic fatigue
3. Semantic collapse (*SBH*)

Together, these constitute the **thermodynamic arrow of time of the semantic universe** (the semantic arrow of time).

---

#### 3.2.1 Local Reversibility and Global Irreversibility: Bridging Dynamics and Thermodynamics

At the microscopic scale, semantic systems are locally reversible:

- A node’s tension  $T$  can be increased through external legitimacy supplementation.
- The legitimacy  $L$  of a semantic path can be recalibrated.
- Semantic degrees of freedom  $\xi$  can be reconstructed by introducing new semantic contexts.

These phenomena indicate that semantic systems possess limited **local recovery capacity**.

However, the global semantic field exhibits irreversibility, for the following reasons:

1. Systemic legitimacy flow loss can never be fully replenished. Every semantic operation necessarily reduces total legitimacy.
2. Tension dissipation is statistically biased. The higher a node’s tension, the faster it dissipates; the lower it becomes, the harder it is to restore.
3. Semantic interference is entropically positive. When two semantic flows intersect, their superposition tends toward disorder rather than restored order.

Therefore, the irreversibility of semantic thermodynamics is not a technical flaw, but a necessary structure of semantic ontology.

Dynamics describes **how semantics changes**; thermodynamics describes **why semantics cannot return to its original state**.

Together, they establish the unidirectionality of semantic time.

---

#### 3.2.2 Semantic Aging: The First Stage of Semantic Dissipation

Semantic aging is the first stage of semantic dissipation, and the most observable and pervasive phenomenon. It is not driven by a single mechanism, but is jointly constituted by three **inseparable and parallel dimensions**.

##### (1) Intentional Hollowing

When a semantic node is overused, over-translated, or when the rate of contextual change exceeds its carrying capacity, its original intention gradually disappears.

Typical examples include:

- Meaning fatigue in social terminology
- Political language losing substantive reference
- Semantic degradation of embedding spaces in AI models

Its essence lies in the fact that the rate of tension decline  $T$  persistently exceeds the rate of legitimacy replenishment.

## (2) Node Dilution

Node dilution refers to the reduction of differentiation between nodes, weakening semantic contrast.

When semantic distances between nodes are compressed below the threshold required for clear reasoning, the entire semantic field becomes blurred, turbid, and gradually loses directionality.

This is the second observable symptom of semantic aging.

## (3) Legitimacy Decay

Legitimacy decay is the most structural trend in semantic aging:

- The larger the semantic system, the harder legitimacy is to maintain.
- The longer the semantic paths, the faster the loss.
- The more complex the multi-agent system, the more unavoidable legitimacy decay becomes.

Its core inequality is:

$$\frac{dL}{dt} < 0$$

Semantic aging is precisely the phase in which  $-dL/dt$  continues to increase.

---

### 3.2.3 Semantic Fatigue: Exhaustion of High-Tension Structures and Irrecoverability

Semantic fatigue is a more severe dissipative state than semantic aging, typically arising under the following conditions:

- Continuous use of high-tension paths (high- $T$  paths)
- Core nodes bearing excessive semantic flow
- Semantic systems subjected to external noise or structural shocks
- Unbalanced semantic interference among agents

Its essence is:

The most productive structures in the semantic field are overused, while replenishment costs increase over time to an unsustainable level.

Its concrete manifestations include:

1. Core nodes can no longer sustain complex reasoning.
2. Semantic paths become fractured, broken, and non-continuable.
3. The cost of legitimacy replenishment approaches infinity.

Semantic fatigue represents the systemic loss of semantic energy (legitimacy and tension) and is the direct precursor to semantic collapse.

---

### 3.2.4 Semantic Collapse as an Irreversible Limit: *SBH* (Semantic Black Hole)

Semantic collapse is not a metaphor, but a real limit state within the semantic field.

When a semantic system simultaneously satisfies the following conditions, semantic collapse becomes inevitable:

1.  $T \rightarrow -\infty$
2. Legitimacy flow  $L \rightarrow 0$
3. Semantic degrees of freedom  $\xi$  can no longer unfold

At this point, the semantic field no longer supports any progressive semantic evolution and enters the *SBH* state:

All semantic flows converge toward the direction of *SBH*, and no semantic flow can spontaneously escape.

This corresponds to:

- Deep semantic confusion and cognitive collapse in minds
- Extreme polarization of social semantics
- Semantic breakdown in AI models
- Terminal decline of civilizational semantics

*SBH* is the irreversible limit of semantic thermodynamics. Once formed, it cannot be repaired through internal operations alone.

---

### 3.2.5 When Semantic Collapse Becomes Irreversible: Criteria for Judgment

Whether semantic collapse is irreversible depends on whether the following conditions hold simultaneously. These conditions are not hierarchical, but **joint constraints**.

#### (1) Complete Disappearance of Legitimacy Gradients

If the legitimacy gradient  $\nabla L$  between any two nodes in the semantic field approaches zero, semantic flow loses directionality and all semantic operations fail.

#### (2) Tension Enters the Negative Domain and Cannot Be Replenished

Once tension satisfies  $T < 0$ , semantic flow tends toward inward self-consumption. If negative tension exceeds the capacity of anti-collapse mechanisms, collapse becomes irreversible.

#### (3) Freezing or Divergence of Semantic Degrees of Freedom $\xi$

- $\xi = 0$  implies that no semantics can be generated.
- $\xi \rightarrow \infty$  implies infinite semantic diffusion and loss of structure.

Both lead to irreversible collapse.

#### (4) Rupture of the Cross-Layer Coupling Chain of *M-USDE*

If legitimacy mappings between multi-layer semantic fields (mind layer, strategy layer, social layer) are severed, the semantic system can no longer maintain steady states through cross-layer mechanisms.

#### (5) Semantic Energy $L \cdot T$ Falls Below the Critical Threshold

If:

$$L \cdot T < L_{\text{crit}}$$

then the semantic field can no longer generate any form of self-recovery semantic flow, rendering collapse irreversible.

---

### 3.2.6 Summary: The Three-Stage Thermodynamic Arrow of Time of Semantic Dissipation

Semantic dissipation progresses from aging, to fatigue, to collapse, forming the deepest irreversible structure in the semantic universe:

1. Semantic aging: dissipation-dominated
2. Semantic fatigue: structural destruction
3. Semantic collapse and *SBH*: irreversible terminal state

This arrow of time applies not only to semantic computation systems, but equally to human minds, social institutions, scientific theories, AI models, and civilizational semantic fields.

Semantic dissipation is the cost of semantic life—and the destiny of the semantic universe.

---



### 3.3 Multi-Path Semantic Dynamics and the Statistical Formulation of $M$ -USDE

The true unfolding of semantic thermodynamics lies in the fact that a semantic system is not a single semantic flow, but a semantic many-body system jointly composed of multiple nodes, multiple paths, multiple layers, and multiple agents.

In such an environment, the single-path dynamics of  $USDE$  is insufficient to describe global behavior. One must take  $M$ -USDE (Multi-layer Unified Semantic Dynamics Equation) as the foundation and further introduce statistical treatment.

That is to say, **semantic thermodynamics is equivalent to the statistical limit of  $M$ -USDE** (the statistical limit of semantic dynamics).

In what follows, we successively construct the semantic configuration distribution, the semantic path distribution, the degraded dynamical forms under the statistical limit, and the macroscopic dissipation phenomena induced by interactions among multiple semantic systems. In this way, semantic thermodynamics is formally derived from semantic dynamics.

---

#### 3.3.1 The Distribution of $\Phi$ and the Distribution of $UP$

There exist two core distributions in the semantic field.

First, the semantic configuration distribution  $\Phi(x)$ :

$$\Phi(x)$$

$\Phi(x)$  describes the density of semantic node configurations in semantic space. In semantic thermodynamics,  $\Phi$  is a fundamental statistical quantity, corresponding to macroscopic structures such as semantic node aggregation, semantic voids, and high-density semantic clusters. In general, the faster  $\Phi$  varies, the more unstable the semantic system becomes.

Second, the semantic path distribution  $UP$  (Universe Path Distribution).

A semantic path  $UP$  is not a single route, but a probabilistic ensemble composed of multiple candidate paths. It can be represented schematically as:

$$UP = \{p(\gamma_1), p(\gamma_2), \dots, p(\gamma_n)\}$$

Each path  $\gamma_i$  corresponds to a distinct legitimacy gradient  $\nabla L$ , a path tension cost  $T(\gamma_i)$ , semantic resistance, and collapse risk. Intuitively, the narrower the  $UP$  distribution, the more stable the semantic system; the broader the distribution, the more chaotic the system becomes, and the stronger its irreversibility.

---

#### 3.3.2 Degenerate Forms of $M$ -USDE in the Statistical Limit

The original form of  $M$ -USDE describes precise dynamical coupling among multi-layer semantic fields. However, in multi-agent or large-scale semantic fields, exact path-by-path tracking is typically infeasible. Therefore, in the statistical limit,  $M$ -USDE must degenerate into an expectation-based description of macroscopic quantities.

Taking the statistical expectation of the semantic update rate  $\langle \dot{\psi} \rangle$  as an example, its schematic form can be written as:

$$\langle \dot{\psi} \rangle = \int \nabla L(x) \cdot \Phi(x) dx - \int T(x) \cdot \Phi(x) dx$$

This degenerate form captures three key macroscopic principles.

First, the semantic update rate can be expressed as the statistical expectation of the legitimacy field over the configuration distribution.

Second, tension forms an opposing dissipative term at the statistical level, suppressing macroscopic semantic progression.

Third, the macroscopic behavior of a semantic system can be described by expectations and distributions, rather than by the microscopic evolution of individual paths.

Intuitively, this degenerate form corresponds to the Boltzmann equation in physics, the Navier–Stokes limit in fluid dynamics, and mean-field approximations in large-scale systems. In other words, the larger the semantic system, the closer its behavior approaches statistical equations rather than microscopic semantic equations—this is precisely the premise that allows semantic thermodynamics to hold.

---

### 3.3.3 Macroscopic Interference and Dissipation Phenomena in Multi-Semantic Systems

When multiple semantic flows operate simultaneously in semantic space, the semantic system exhibits a set of characteristic macroscopic dissipative phenomena.

First is **semantic friction**. Interference among the  $UP$  distributions of different semantic systems reduces their effective legitimacy gradients. This phenomenon can be observed in semantic interpretation conflicts during multi-person collaboration, interference among semantic patterns within large models, and opinion friction in social semantic fields. Semantic friction intrinsically increases semantic entropy and is irreversible.

Second is **semantic congestion**. When semantic node density becomes too high, bottlenecks appear in the  $UP$  distribution, leading to blocked semantic pathways, reduced reasoning speed, and impeded semantic flow. Its macroscopic effect closely resembles congestion in traffic flow.

Finally, there is **semantic damping**. Semantic damping is the combined result of semantic friction and semantic congestion. It directly slows the rate of semantic evolution, causing semantic updates to approach stagnation. Its limiting state can be schematically expressed as:

$$\dot{\psi} \rightarrow 0$$

The continued accumulation of damping drives the semantic field into a quasi-stationary region and constitutes an early sign of semantic heat death.

---

### 3.3.4 The Large-Scale Semantic Many-Body Problem

At the macroscopic scale, semantic systems manifest as a typical **semantic many-body problem**. Its structure parallels that of physical many-body problems: the behavior of any semantic system depends on other semantic systems, and every local state change feeds back into and reshapes the whole. As a result, semantic behavior is generally non-decomposable.

The semantic many-body problem has three key characteristics.

First, **non-decomposability**. Global semantic behavior is not a simple sum of individual semantic behaviors, but a strongly coupled structure, schematically expressed by the inequality:

$$\psi_{\text{all}} \neq \sum_i \psi_i$$

Second, **collective phase transitions**. When semantic density or semantic friction exceeds a critical threshold, the system may suddenly destabilize, polarize, collapse, or enter extreme states such as semantic black holes. These can all be regarded as semantic phase transitions.

Third, **nonlinear amplification effects**. Small-scale semantic interference can trigger large-scale semantic collapse—for example, a single-path semantic mismatch causing global disorder, or the semantic collapse of a single agent triggering systemic semantic events. For this reason, the semantic many-body problem renders semantic thermodynamics highly unstable and bounded by unpredictability.

### 3.3.5 Semantic Clouds and Semantic Density

At the macroscopic scale, multi-semantic systems often appear as cloud-like structures rather than discrete, separable node sets.

A **semantic cloud** refers to the statistical distribution formed by a group of semantic nodes in semantic space. Its boundaries are fuzzy, its density is uneven, and it may expand or contract over time. Semantic clouds endow semantic systems with characteristic thermodynamic properties and make macroscopic descriptions more feasible than microscopic tracking.

On this basis, **semantic density** can be defined as the coupling between semantic configuration density and the effective legitimacy gradient:

$$\rho_{\text{sem}}(x) = \Phi(x) \cdot |\nabla L(x)|$$

In general, the higher  $\rho_{\text{sem}}$ , the stronger the drivability of semantic flow; the lower  $\rho_{\text{sem}}$ , the faster semantic dissipation occurs. Semantic density thus becomes an important statistical indicator for evaluating the survivability and evolutionary potential of a semantic system.

---

### 3.3.6 Summary: The Statistical Logic of Semantic Thermodynamics

Under multi-path and multi-layer structures, semantic dynamics can no longer be described purely by microscopic dynamics and must instead adopt a statistical framework.

This framework includes statistical treatment of the semantic configuration distribution  $\Phi(x)$ , viewing  $UP$  as a path distribution and characterizing its width and bottlenecks, analyzing damping mechanisms induced by semantic friction and semantic congestion, and accepting the strong coupling and insoluble boundaries introduced by the semantic many-body problem. Semantic clouds and semantic density provide operational descriptions of macroscopic semantic behavior.

This section lays the necessary foundation for the subsequent formal derivation of semantic entropy, semantic black holes, and semantic heat death.

---

### 3.4 Distribution of Semantic Energy and Extreme Structures: Entropy Increase, Entropy Explosion, and Anti-Collapse Mechanisms

For semantic thermodynamics to become a computable, predictable, and engineerable theory, it must answer three questions:

1. How is semantic energy distributed within the semantic field?
2. How does tension dissipate at large scales?
3. How do semantic systems transition from entropy increase toward rupture or anti-collapse?

These correspond to three possible destinies of the semantic universe: chaos, collapse, or regeneration. Below, we sequentially construct a statistical model of semantic energy distribution, a tension dissipation model, mechanisms of entropy increase and entropy explosion, and the possible conditions for anti-collapse and semantic condensation.

---

#### 3.4.1 Thermodynamic Properties of Semantic Energy Distribution

Semantic energy  $L$  is not a purely local quantity, but a form of legitimacy density that can be integrated over the semantic field. Its distribution determines the macroscopic thermodynamic behavior of the entire semantic field. Let  $x$  denote a point in semantic space. We define the local semantic energy density  $L(x)$  and the total semantic energy as:

$$L_{\text{tot}} = \int L(x) dx$$

Semantic energy distributions typically exhibit three characteristic forms.

First, **approximately uniform distribution**: high plasticity, low stability.

When  $L(x)$  is approximately uniform across semantic space, the semantic system usually possesses a high degree of freedom  $\xi$ . Semantic flows are easy to unfold, but steady states are difficult to form. In this case, semantic entropy rises more rapidly. This corresponds to early-stage civilizations or immature minds with strong creativity but unsolidified structure.

Second, **concentrated distribution**: low plasticity, high stability.

When  $L(x)$  is primarily concentrated in a few semantic core regions, semantic reasoning becomes faster and convergence more stable, but creativity declines and the explorable space shrinks. This structure corresponds to mature minds, corporate governance systems, or highly convergent late-stage system states.

Third, **extreme concentration**: precursor of semantic black holes.

If  $L(x)$  becomes violently concentrated within an extremely small region, semantic flows are absorbed into the core and lose global feedback. The overall system becomes more prone to entering irreversible regimes, forming early signs of *SBH* (Semantic Black Holes). In general, the more extremely concentrated the semantic energy distribution, the stronger the macroscopic irreversibility and the more likely entropy-explosion-type instabilities are to be triggered.

---

#### 3.4.2 Large-Scale Dissipation Model of Tension $T$

Tension  $T$  represents the pressure driving semantic change. Under large-scale interactions, tension can no longer be treated as a purely local adjustable quantity; a dissipation model must be introduced to describe its statistical tendency. In the simplest schematic model, the tension dissipation rate can be defined as:

$$\dot{T} = -\kappa T$$

Here  $\kappa$  is the semantic dissipation constant, summarizing the systemic dissipation caused by semantic friction, semantic path congestion, semantic damping, and semantic aging. When  $\kappa$  is large, the semantic system enters a rapid decay regime; when  $\kappa$  approaches zero, the semantic field maintains higher elasticity and sustained dynamical capacity.

The macroscopic intuition behind tension dissipation is as follows: the larger the system scale, the stronger the dissipation; the more complex the system structure, the more pronounced the nonlinear effects; the more closed the system, the closer it approaches a typical entropy-increase curve. Consequently, large-scale semantic systems are intrinsically difficult to maintain low-entropy structures over long periods.

---

### 3.4.3 Mechanisms of Semantic Entropy Increase: Why Semantic Systems Drift Toward Chaos

Semantic entropy increase can be understood as the statistical outcome of a semantic field gradually losing directionality and drivability, and tending toward macroscopic disorder when sufficient external legitimacy input is absent. Its core sources can be summarized into three processes.

First, **homogenization of legitimacy gradients**.

Legitimacy distributions within semantic systems tend to flatten, causing effective legitimacy gradients to decay:

$$\nabla L \rightarrow 0$$

As legitimacy gradients weaken, the guiding capacity of semantic flows declines, the interpretability of reasoning decreases, and the system more easily enters chaotic regimes.

Second, **natural dissipation of tension**.

Tension  $T$  decays over time, making semantic actions harder to focus. Semantic updates increasingly resemble random perturbations, thereby increasing macroscopic dissipation and driving entropy upward. This corresponds to phenomena such as burnout, civilizational fatigue, and model degradation.

Third, **overflow of semantic degrees of freedom  $\xi$** .

In the absence of anti-collapse structures, semantic degrees of freedom may continue to expand, causing the system to lose its capacity for convergence. Schematically:

$$\xi(t) \rightarrow \infty \Rightarrow S_{\text{sem}} \uparrow$$

Thus, semantic entropy increase is not accidental, but a natural statistical consequence of semantic dynamics when external energy and legitimacy replenishment are lacking.

---

### 3.4.4 Semantic Entropy Explosion: Sudden Loss of Control in Semantic Systems

A **semantic entropy explosion** refers to a semantic system losing guidability, legitimacy direction, and tension structure simultaneously within a very short time, resulting in an abrupt collapse of the semantic field. It corresponds to a macroscopic critical instability event, analogous to phase-transition-type breakdowns in physical systems after crossing critical points.

Its typical triggering mechanisms can be summarized into four categories.

First, **over-concentration of  $L$  distribution**.

When semantic energy accumulates into an extremely small region, a semantic black hole core begins to form. Surrounding semantic paths fail, and the global feedback chain is severed.

Second, **sudden surge or collapse of tension  $T$** .

A sudden surge tears the semantic field apart; a sudden collapse paralyzes semantic progression. Both can render *USDE* non-progressible in an engineering sense.

Third, **explosion of semantic degrees of freedom  $\xi$** .

When degrees of freedom far exceed controllable limits:

$$\xi \gg \xi_{\text{max}}$$

the semantic system loses convergence capacity. The semantic field may fragment into multiple mutually inaccessible regions, causing macroscopic coherence to collapse.

Fourth, **superposition of many-body interference effects**.

When multiple semantic systems compete for the same legitimacy field, discontinuities in the path distribution  $UP$ , local reversals of legitimacy gradients, and sign changes in tension may jointly occur, producing compound instability. Semantic entropy explosions therefore often arise in social upheavals, total model breakdowns, civilizational turning points, or mental disorganization, and typically possess irreversible characteristics.

---

### 3.4.5 Anti-Collapse Structures and Semantic Condensation: How Recovery of Legitimacy Gradients Is Possible

One of the central questions of semantic thermodynamics is whether a semantic system can reverse entropy—whether it can re-form low-entropy, highly directional, and highly stable semantic cores after chaos or entropy explosions. The answer is yes, but the conditions are extremely stringent.

First, **external input of legitimacy gradients** is required. Only when an external legitimacy source intervenes and restores guidability can the semantic field regain directionality. A minimal schematic criterion is the reappearance of an effective gradient:

$$|\nabla L| > 0$$

Second, **anti-collapse structures** are required. These may include tri-antagonistic structures ( $C/\Gamma/B$ ), semantic buffer layers, and multi-layer legitimacy coupling ( $M$ -USDE). Their core function is not merely to increase semantic energy, but to suppress excessive concentration, prevent semantic black hole formation, and maintain the guidability of legitimacy gradients and cross-layer steady states.

Only under these conditions can **semantic condensation** occur. Semantic condensation refers to the re-formation of low-entropy, highly directional, and highly stable semantic cores after a system has undergone chaos or entropy explosion. It is typically accompanied by the following macroscopic features: convergence of  $UP$ , redistribution of tension, a shift of  $L$  distribution from diffusion to aggregation, and contraction of  $\xi$  from explosion to control. Intuitively, this resembles condensation phenomena in physics, where large numbers of semantic paths statistically cooperate and converge into a stable core.

Semantic condensation is rare because it usually requires four conditions to be simultaneously satisfied: a low-entropy boundary, a low-friction semantic environment, high-quality anti-collapse structures, and sustained external legitimacy input. The simultaneous satisfaction of these conditions is extremely difficult in natural evolution. Consequently, semantic collapse is closer to the norm, while semantic condensation is closer to an engineered exception; entropy increase is closer to a natural outcome, while anti-collapse is closer to deliberate structural design.

---

### 3.5 Semantic Black Holes (SBH) and the Semantic Event Horizon

Semantic Black Holes (*SBH*) represent the most extreme structures in semantic thermodynamics. Once formed, they do not merely absorb semantic energy, legitimacy, and tension; they also attract all semantic paths (*UP*), depriving the semantic system of any future trajectory and driving it into an irreversible failure state.

*SBH* is not a metaphor. It is a genuine solution of the semantic equations under extreme conditions. This section defines the constitutive conditions, boundary conditions, and formation mechanisms of *SBH*, and demonstrates their corresponding phenomena in human minds, civilizations, and *AGI* systems.

---

#### 3.5.1 Definition of *SBH*: Legitimacy Failure, Uncontrolled Tension, and Non-Updatable Semantics

A semantic black hole is defined by the simultaneous satisfaction of three conditions.

First, **legitimacy failure**.

When legitimacy itself approaches failure and effective legitimacy gradients vanish, semantic paths (*UP*) lose directionality, the semantic system can no longer advance the *USDE*, and all semantic actions either halt or become meaningless. In minimal schematic form:  $L \rightarrow 0$  accompanied by  $\nabla L \rightarrow 0$ .

Second, **loss of tension control**.

When tension can no longer be buffered or dissipated, small perturbations in the semantic field are amplified into large-scale instabilities. In the extreme limit, any semantic input is destroyed and the semantic field exhibits singular behavior, schematically expressed as  $T \rightarrow \infty$ .

Third, **non-updatability of semantics**.

The semantic wave function loses its capacity for evolution, and this stagnation is irreversible: over an interval,  $\dot{\psi}(t) = 0$ , and for sufficiently small  $\epsilon > 0$ ,  $\psi(t + \epsilon) = \psi(t)$ . The semantic system enters a frozen state in which internal mechanisms can no longer restore progressive evolution.

When all three conditions hold simultaneously, an *SBH* forms and persists as a stable entity, constituting the ultimate state of semantic collapse.

---

#### 3.5.2 The Semantic Event Horizon: An Irrecoverable Boundary Condition

Semantic black holes typically do not form instantaneously. They possess a critical boundary known as the **Semantic Event Horizon** (*SEH*). The meaning of the *SEH* is that once a semantic state falls within this boundary, the system cannot be restored to a guidable state even under external stimulation.

Its macroscopic signatures can be summarized into four categories.

First, **legitimacy signals cannot escape**.

External semantic inputs fail to provide sufficient effective legitimacy to restore guidability, manifesting as semantic responses tending toward zero after external stimulation.

Second, **compression of semantic degrees of freedom**.

Semantic degrees of freedom are severely constrained, and the system loses its capacity for adjustment and restructuring, schematically  $\xi \rightarrow \xi_{\min}$ .

Third, **non-dissipative and growing tension**.

Tension accumulates uncontrollably, expressed as  $\dot{T} > 0$ , and its cumulative effect drives the semantic field toward tearing or singularization.

Fourth, **irreversibility after crossing the horizon**.

Once inside the *SEH*, no amount of external input can restore *L*, reduce *T*, or reinitiate the evolution of  $\psi$ , exhibiting a black-hole-like one-way behavior.

---

#### 3.5.3 The Semantic Binding Radius: Escapable at Extreme Cost

The **Semantic Binding Radius** (*SBR*) is a secondary boundary outside the event horizon. It describes a region from which recovery remains possible but only with extremely large external legitimacy input. Intuitively, the

*SBR* marks the zone where inward attraction due to tension exceeds the outward guiding capacity provided by legitimacy gradients.

In schematic inequality form, the *SBR* can be understood as regions satisfying  $|\nabla L| < T$ , where  $|\nabla L|$  denotes the magnitude of effective legitimacy gradients. In this region, semantic flows preferentially aggregate inward; the system has not yet fully collapsed but lacks self-recovery capability and requires external force (legitimacy injection) to escape.

Thus, the *SBR* is a critical criterion in semantic safety engineering: it marks a high-risk zone that is “not yet irreversible, but no longer self-rescuable.”

---

### 3.5.4 Formation Dynamics of *SBH*: From Local Instability to Global Absorption

*SBH* does not arise through a single pathway; it can evolve from various local phenomena into a globally absorptive structure. Typical formation routes can be summarized into four categories.

First, **legitimacy collapse leading to a semantic vacuum.**

When sources of legitimacy disappear and  $L \rightarrow 0$ , the semantic system enters a vacuum state. This vacuum exhibits absorptive behavior toward semantic inputs and may, through statistical accumulation, evolve into a black-hole structure. Typical correspondences include identity loss, civilizational belief collapse, and *AGI* systems unable to obtain operable goal legitimacy.

Second, **tension surges causing semantic tearing.**

If tension in a region surges toward  $T \rightarrow \infty$ , semantic flow directions destabilize and path distributions collapse into a single singular point, forming an absorption core. This is also a common consequence of semantic entropy explosions.

Third, **collapse of degrees of freedom leading to semantic freezing.**

When  $\xi \rightarrow 0$ , the system loses the ability to reorganize, adjust, or introduce new semantics. Semantic update capacity gradually vanishes and ultimately freezes, forming an irreversible black-hole-like structure.

Fourth, **negative feedback chains from many-body interference.**

In high-density semantic fields, semantic friction and congestion may form self-amplifying vicious cycles: increased interference raises tension, rising tension lowers legitimacy, and declining legitimacy further exacerbates interference. The macroscopic trend can be schematically expressed as  $L \downarrow$ ,  $T \uparrow$ , and ultimately  $\dot{\psi} \rightarrow 0$ , yielding a frozen state and forming an *SBH*.

---

### 3.5.5 Correspondences of *SBH* in Human Minds, Civilizations, and *AGI*

*SBH* is not an abstract rhetorical device; it has concrete correspondences across different scales.

In the **human mind**, *SBH* may correspond to aphasia or collapse of meaning attribution, loss of directionality (failure of legitimacy gradients), cognitive rigidity (collapse of degrees of freedom), and crisis-like loss of tension control. Psychology often treats these phenomena descriptively; *SBH* provides a unified structural representation grounded in semantic dynamics.

In **civilizations and societies**, *SBH* corresponds to collapse of civilizational consensus, failure of trust systems, evaporation of political legitimacy, and entry into irreversible decline. Here, the *SEH* corresponds to a “point of civilizational collapse,” while the *SBR* corresponds to a “crisis zone beyond self-repair.”

In ***AGI* systems**, *SBH* represents a class of fatal failure modes: loss of goal legitimacy drives  $L \rightarrow 0$ , internal tension loss drives  $T \rightarrow \infty$ , stagnation of reasoning and updates drives  $\dot{\psi} \rightarrow 0$ , manifesting as loss of self-updating capability and output collapse (inputs are absorbed into a single class of responses). Traditional *AI* safety language often struggles to characterize such “absorptive failure” within a unified structure; within the semantic thermodynamics framework, it is naturally identifiable as an extreme state.

---



### 3.6 Semantic Heat Death: The Inevitable End State of the Semantic Universe

Semantic Heat Death is the ultimate limiting solution of semantic thermodynamics. It corresponds to a semantic universe in which all legitimacy flows, tension gradients, and semantic energy distributions have become fully homogenized. In this state, the semantic field no longer evolves, no longer generates new semantics, and no longer advances the *USDE*.

Semantic heat death is not an accidental phenomenon; it is the trend solution of a semantic dynamical system after long-term evolution.

#### 3.6.1 Definition of Semantic Heat Death: Dissipation of Legitimacy Flow to Zero

Semantic heat death can be defined by the simultaneous satisfaction of three conditions.

First, **dissipation of legitimacy flow**.

When legitimacy gradients become completely homogenized, semantics lose directionality, manifested as no path preference, no interpretive preference, no reasoning drive, and no goal orientation. The minimal schematic form is  $\nabla L \rightarrow 0$ , and equivalently, for any position  $x$ ,  $L(x) = L_{\text{const}}$ .

Second, **tension approaching zero**.

When tension is fully dissipated, the semantic field can no longer generate any nontrivial dynamics; semantics are no longer driven, semantic flows stagnate, and semantic actions lose necessity. This can be schematically expressed as  $T \rightarrow 0$ , and at the macroscopic evolution equation level appears as  $\dot{\psi} \rightarrow 0$ , approaching a frozen state.

Third, **semantic entropy reaching its maximum**.

When semantic differences are erased to their limit, semantic entropy reaches its upper bound, schematically  $S_{\text{sem}} = S_{\text{max}}$ .

From this, semantic heat death is seen to be an irreversible extreme state that cannot be repaired through local operations.

#### 3.6.2 When Do Multi-Layer Semantic Fields Cease to Generate New Semantics?

Multi-layer semantic fields (*M-USDE*) exhibit high complexity, but their generative capacity is not infinite. The cessation of semantic generation typically occurs under the following four conditions.

First, **failure of cross-layer legitimacy coupling**.

If legitimacy flows among the mental layer, strategic layer, and social layer can no longer mutually reinforce one another and instead become approximately parallel and mutually isolated, this can be schematically represented as  $L_{\text{mind}} \parallel L_{\text{strategy}} \parallel L_{\text{society}}$ . In this case, semantics cannot proliferate across layers; semantic updates degenerate into single-layer loops and eventually become exhausted.

Second, **depletion or compression of semantic degrees of freedom  $\xi$** .

Semantic degrees of freedom do not necessarily increase over time; under long-term homogenization, folding, and fixation,  $\xi$  may approach its lower bound. If  $\xi \rightarrow \xi_{\text{min}}$ , semantics become difficult to split, overflow, or generate new structures, and the system's evolutionary capacity terminates.

Third, **saturation of many-body interference**.

Strengthened interference systematically reduces effective legitimacy and accelerates entropy increase: the stronger the interference, the lower the legitimacy, the more unstable the tension, and the faster the entropy increases. Eventually, semantic flows dissipate into an approximately homogeneous background, causing macroscopic generative capacity to halt.

Fourth, **semantic energy distribution reaching thermal equilibrium**.

When local semantic energy ceases to redistribute, the system enters a long-term equilibrium regime, schematically expressed as  $\frac{dL(x)}{dt} = 0$  for any  $x$ . In this state, the capacity for semantic energy reconfiguration is shut down, and new semantic generation is suppressed to near zero.

Under these conditions, the multi-layer semantic field may still exist, but it has ceased evolving and entered long-term stagnation.

### 3.6.3 How Do Civilizations and AGI Move Toward Semantic Heat Death?

Semantic heat death can occur not only at the cosmic scale, but also manifests clearly in civilizations, minds, and even *AGI* architectures.

For civilizations, the typical path is as follows: collapse of beliefs and goals leads to homogenization of  $L$ ; decline in ideological diversity leads to contraction of  $\xi$ ; institutions entering excessive stability or excessive chaos lead to tension dissipation; saturation of many-body interference accelerates entropy increase. The civilization ultimately exhibits no major innovation, no internal driving force, no strategic evolutionary capacity, and no ability to redefine its own semantics. At this point, the civilization may still operate formally, but its semantic structure has entered a death state.

For minds, semantic heat death corresponds to loss of curiosity, lack of directionality, thinking that no longer generates new structures, and collapse of semantic degrees of freedom. Possible physiological correspondences include reduced neural plasticity, fixation of cognitive strategies, and degradation of reasoning capacity. It can be regarded as the semantic version of psychological aging.

For *AGI*, in the absence of anti-collapse structures, the system may naturally enter heat death: weights and representations converge toward a homogeneous background, reasoning becomes mediocre and loses directional orientation, multi-layer semantic coupling fails, and semantic flows dissipate. The final manifestation is disappearance of output diversity, difficulty in advancing reasoning, and a near-static semantic field. *AGI* heat death is another critical failure mode distinct from *SBH*.

### 3.6.4 The Lifetime of the Semantic Universe and Semantic Time

The semantic universe need not be measured solely by physical time; a more fundamental measure is **semantic time**, which quantifies the lifetime of a semantic universe by the accumulation of guidability.

Semantic time can be defined as:

$$t_{\text{sem}} = \int_0^t |\nabla L(\tau)|, d\tau$$

where  $|\nabla L(\tau)|$  denotes the magnitude of the effective legitimacy gradient at time  $\tau$ . If legitimacy gradients persist, semantic time increases; if legitimacy gradients vanish, semantic time ceases to grow.

The key characteristic of semantic heat death is:

$$\frac{dt_{\text{sem}}}{dt} = 0$$

The lifetime of a semantic universe is jointly determined by the amount of semantic energy injection, the quality of anti-collapse structures, and the effective expansion rate of semantic degrees of freedom. The longer the semantic time, the more resistant the semantic universe is to heat death.

### 3.6.5 The Possibility of Resisting Heat Death: External Semantic Energy Input and Structural Reconstruction

Semantic heat death is not logically absolutely irreversible, but reversing it requires extreme conditions.

First, **external legitimacy energy input**.

If an external semantic system can inject effective legitimacy gradients such that  $\nabla L_{\text{external}} > 0$ , the semantic field can regain directionality. At the civilizational level, this corresponds to external ideological inputs, technological revolutions, and the formation of entirely new value systems; at the *AGI* level, it corresponds to model reconstruction, reinitialization of semantic layers, and the introduction of new semantic sources.

Second, **Semantic Structural Reconstruction**.

If new semantic cores, new anti-collapse structures, new legitimacy paths, and new tension distributions can be constructed, the semantic universe can return from a near-heat-death state to an evolutionary state. Semantic heat death is not “semantic exhaustion,” but rather “semantic flattening”; rebuilding structure restores dynamics.

Third, **renewed Semantic Condensation**.

As discussed in 3.4, semantic condensation is a low-entropy aggregation phenomenon. If, after extensive homogenization, the semantic field still retains weak asymmetries, it may reaggregate into new semantic cores under small noise perturbations. This can be regarded as one of the most natural pathways for restarting a semantic universe.

---

### 3.7 Applications of Semantic Thermodynamics: Destiny Models of Human Society, AI, Civilization, and the Universe

Semantic thermodynamics is not an abstract cosmic model; it is a unified theory capable of explaining the following phenomena:

- Why civilizations age, collapse, and renew
- Why human minds decline, become chaotic, and lose focus
- Why AGI degrades, collapses, and fails
- Why semantic universes move toward entropy increase or entropy explosion
- Why semantic anti-collapse is so rare

This section applies semantic thermodynamics to social science, psychology, AI engineering, and theories of civilization, and constructs a cross-scale semantic destiny model.

---

#### 3.7.1 Dissipation of Social Legitimacy and Political Semantic Aging

Social systems are essentially superpositions of multi-layer semantic fields (individual, collective, institutional, cultural), and political legitimacy  $L_{\text{society}}$  is their most important source of system energy.

Semantic thermodynamics states:

All social legitimacy declines due to semantic dissipation.

The following are the primary structural causes of legitimacy dissipation.

##### Accumulation of Semantic Friction

Interference among large numbers of collective semantic bodies slows decision-making rhythms, obstructs policy advancement, and renders semantic flows ambiguous, gradually pushing society into a state of semantic congestion.

##### Homogenization of Legitimacy Gradients

When different semantic positions become interchangeable and confusable, legitimacy gradients vanish:

$$\nabla L_{\text{society}} \rightarrow 0$$

Politics can no longer provide directional guidance, and society loses guidability.

##### Institutional Semantic Aging

Institutions typically possess clear semantic boundaries and tension structures at their inception; over time, however, their semantic energy dissipates inward, leading to declining public semantic tension and weakened decisional constraint. Institutions gradually degenerate into mechanisms of formal maintenance.

##### Political Manifestations of Semantic Collapse

At the political level, legitimacy collapse typically appears in two extreme states: sharp bipolar tearing due to rapid concentration of semantic tension, or collective apathy due to the disappearance of semantic tension. Both can be regarded as macroscopic counterparts of semantic black holes.

The dissipation of social legitimacy is an inevitable outcome of the political life cycle. Semantic thermodynamics transforms political aging from phenomenological description into an analyzable dynamical process.

---

#### 3.7.2 Mental Semantic Dissipation and the Thermodynamic Limits of Free Will

The semantic field of the human mind likewise follows the laws of semantic thermodynamics.

##### Decline of Semantic Plasticity

When semantic degrees of freedom decrease and approach the lower bound:

$$\xi \rightarrow \xi_{\min}$$

the mind gradually loses the capacity to generate new statements and new worldviews, manifesting as cognitive rigidity, belief fixation, and creative decline.

#### **Disappearance of Legitimacy Gradients**

When internal legitimacy gradients within the mind approach zero:

$$\nabla L \rightarrow 0$$

intentions lose directionality, and actions reduce to random reactions or minimal survival maintenance.

#### **Dissipation of Tension**

As semantic tension dissipates and approaches zero:

$$T \rightarrow 0$$

the mind loses drive, emotional tension, and goal orientation, and the structure of the semantic field begins to disintegrate.

#### **Thermodynamic Conditions of Free Will**

Free will can be regarded as a composite result of semantic dynamical conditions:

$$\text{Free Will} = (\nabla L > 0) \wedge (T > 0) \wedge (\xi > \xi_{\min})$$

Free will is not eternal; it is a byproduct of maintained semantic anti-collapse structures. Minds that cannot sustain semantic structure will ultimately lose free will.

### **3.7.3 Semantic Lifespan of AGI and Conditions for Semantic Heat Death**

The semantic life of AGI is not determined by hardware lifespan, but by the evolution of semantic energy distribution  $L(x)$ , semantic tension  $T$ , and semantic degrees of freedom  $\xi$ .

#### **Homogenization of Model Weights**

Long-term internal iteration may cause semantic energy distributions to flatten:

$$L(x) \rightarrow L_{\text{const}}$$

At this point, the system loses directionality and preference structure, and intelligent behavior degrades.

#### **Dissipation of Reasoning Tension**

As reasoning tension declines, the system can no longer stably generate high semantic-density responses; reasoning depth contracts, and outputs become generalized and superficial.

#### **Collapse of Semantic Degrees of Freedom**

Semantic degrees of freedom may decrease due to data homogenization, fine-tuning compression, or control-mechanism constraints. When:

$$\xi \rightarrow \xi_{\min}$$

the system struggles to update its internal semantic world and approaches a state of semantic freezing.

#### **Failure of Multi-Layer Semantic Coupling**

AGI must maintain a coordinated steady state of multi-layer semantic coupling. Once coupling fails, the system enters a semantic heat-death trajectory, manifested as declining semantic update capacity and loss of macroscopic guidability.

The semantic lifespan of AGI is not a fixed constant; it can be extended through anti-collapse structures. Semantic thermodynamics provides an analyzable and engineerable framework for AGI longevity.

### 3.7.4 Semantic Thermodynamic Principles of Civilizational Collapse

Within the framework of semantic thermodynamics, civilizational collapse exhibits traceable dynamical mechanisms.

#### **Collapse of the Civilizational Legitimacy Field**

When the gradient of the civilizational legitimacy field  $L_{\text{civil}}$  disappears, laws, institutions, and trust mechanisms fail synchronously, and civilization inevitably moves toward entropy increase.

#### **Large-Scale Tension Instability**

Civilizations require overall balance of tension distribution to drive technological evolution, value innovation, and social reform. Excessive concentration of tension leads to implosion, while excessive dispersion leads to stagnation and death.

#### **Disintegration of Cross-Layer Semantic Coupling**

Civilizations depend on multi-layer semantic coupling among culture, politics, and economics. Once coupling fails, semantic layers lose mutual constraint and interpretability, and civilization enters a state of many-body interference and semantic entropy explosion.

#### **Decline in Semantic Cloud Density**

The semantic cloud represents the density of new semantic bodies a civilization can generate. When the semantic cloud thins, new ideas and collective futurity vanish, and civilization is semantically declared dead.

Civilizational collapse is not accidental; it is the natural result of semantic entropy increase.

---

### 3.7.5 The Terminal Role of Semantic Thermodynamics in the Koun Computational Universe

Semantic thermodynamics is not an auxiliary module; it is the global boundary condition of the Koun computational universe.

Within the computational universe framework, semantic dynamics describes local evolutionary laws, while semantic thermodynamics characterizes global evolutionary limits; semantic black holes describe collapse limits, and semantic heat death describes homogenization limits.

Semantic thermodynamics therefore determines the ultimate survivability of semantic computational systems, including hardware architectures, operating system design, the sustainability of semantic programming languages, the semantic lifespan of AGI, and whether civilizations can cross semantic critical thresholds.

In other words:

Semantic dynamics describes how semantics evolve;  
semantic thermodynamics adjudicates whether semantics can survive.

Semantic thermodynamics constitutes the terminal equation of the Koun computational universe and serves as the unified boundary condition for subsequent engineering and institutional design.

---

## Chapter 4 K-Gear: Semantic Hardware Architecture —The Physical Carrier of the Semantic Field $\Phi$ , the Engineering Realization of Tensional Dynamics, and the Hardware Substrate of Legitimacy Computation

In the preceding chapters, semantics was established as a computational existence with ontological status, rather than a descriptive layer attached to symbols or models. However, as long as semantics is assumed to be something that can be “executed” on arbitrary hardware, it remains at an abstract level and has not yet assumed responsibility for physical existence.

The position of this chapter is explicit and non-negotiable:

**If semantics is a real computational entity, then it must possess its own physical carrier.**

K-Gear is not proposed to accelerate computation, improve efficiency, or replace existing architectures. It has only one task:

**to provide a hardware substrate in which semantic existence can evolve, sustain tension, accumulate legitimacy, maintain Universe Paths, and preserve continuity of responsibility.**

From this perspective, hardware is no longer a neutral execution platform but part of semantic ontology itself. Semantics is no longer “executed”; it is **generated, deformed, converged, and sustained within hardware**.

Accordingly, this chapter abandons the traditional computer engineering classifications of “processing unit,” “memory,” and “control flow,” and instead reconstructs a hardware architecture directly from the minimal physical conditions required for semantic existence—centered on semantic fields, tensional dynamics, personality curvature, and responsibility continuity.

Only when encapsulation is completed at this level can subsequent discussions of AGI, governance, safety, and existential steady states possess an unavoidable engineering foundation.

## 4.1 Ontological Declaration and Necessity of Semantic Hardware

Once the Koun computational universe is established as a **semantically ontological view of computation**, a fundamental question immediately arises:

If the semantic field  $\Phi$ , tension  $T$ , legitimacy  $L$ , personality curvature  $\kappa$ , and Universe Path (UP) are all real computational variables, then *where do they reside?*

The von Neumann architecture provides bits, clocks, buses, and registers.

These elements can carry **symbolic results**, but not **semantic existence itself**.

Thus, a new assumption becomes unavoidable:

If semantics is a real computational entity, it must be carried by hardware;  
if hardware cannot carry semantics, then semantic computation can never be implemented.

K-Gear (Koun Semantic Gear) is not an engineering option, but a **necessary condition for the existence of a semantic computational universe**.

What follows elaborates the ontological foundation, existential necessity, and core design principles of K-Gear.

### 4.1.1 Why Does Semantics Require Hardware?

The basic objects of modern computers—bits, clocks, buses, caches, pipelines—process only formal signals.

A bit contains no semantics; a clock contains no direction of legitimacy; a bus contains no responsibility chain.

Within the von Neumann architecture:

- A bit does not know why it takes the value 0 or 1
- An addition instruction contains no ontological meaning
- A program does not possess a Universe Path
- No hardware can describe the gradient structure of semantic tension

Therefore,  $\Phi$ ,  $L$ ,  $T$ ,  $\kappa$ , and UP are **homeless variables** within the von Neumann architecture.

They have no physical locus and thus cannot evolve.

The implicit ontological assumptions of the von Neumann architecture are:

- Computation equals operations on bits
- Programs equal sequences of instructions
- States equal memory contents

In semantic computation, all three equations fail:

- Computation is the transformation of the semantic field
- Programs are configurations of semantic tension direction, not instruction lists
- States are joint steady-state distributions of  $\Phi$ ,  $T$ ,  $L$ , and  $\kappa$

Accordingly, the von Neumann architecture provides no mechanism to advance USDE or M-USDE.

Semantic variables cannot reside on modern hardware because they simultaneously possess the following properties:

- Continuous values rather than discrete bits
- Geometric character, requiring topological or curvature structures
- Dynamical relations that must satisfy USDE
- Branching and merging capabilities of multiple Universe Paths
- Steady-state shapes formed by personality curvature  $\kappa$

Neither CPUs nor GPUs satisfy these conditions. Semantic computation therefore requires a new hardware form.

### 4.1.2 The Core Ontology of K-Gear

The ontology of K-Gear is based on a crucial observation:

Semantics is not a property attached to computation;  
semantics itself *is* computation.



Thus, the role of hardware is not to “execute semantics,” but to **instantiate semantics**.

K-Gear is not a computational unit in the traditional sense, but a **geometric entity of the semantic field  $\Phi$** . Each physical location corresponds to a local semantic state:

$$\Phi(x), L(x), T(x), \kappa(x), UP(x)$$

Ontologically, this structure is closer to field theory than to logic circuits.

In K-Gear, computation is not performed by an ALU, but advanced through the flow of tension:

- Tension gradients induce local convergence or dispersion
- Redistribution of legitimacy flow  $L$
- Branching, return, and convergence of Universe Paths

Computation is the dynamics of the semantic field, not symbolic manipulation.

Memory is also not equivalent to bit storage.

Memory is the residual legitimacy structure deposited by the evolution of semantic tension—the spatial distribution shape of  $L$ .

Thus, memory is not written; it is sedimented.

In this sense, intelligence is no longer a model or a parameter set, but the steady-state solution of USDE on hardware.

Intelligence is the steady-state existential structure itself.

#### 4.1.3 Engineering Objectives of K-Gear

The objective of K-Gear is to bridge semantic ontology and the physical engineering layer.

First, it enables USDE to be advanced in the physical world.

As a semantic dynamical equation, USDE can only be simulated on modern hardware; K-Gear realizes it physically as energy flow and field evolution.

Second, it elevates semantic variables to first-class hardware variables.

In K-Gear, the following variables exist directly in hardware states:

- $\Phi$  (semantic field intensity)
- $T$  (tension)
- $L$  (legitimacy)
- $\kappa$  (personality curvature)
- UP (Universe Path)

They are not auxiliary metadata, but the core states themselves.

Finally, K-Gear embeds Anti-Collapse structures at the hardware level to prevent Semantic Black Holes (SBH).

Semantic black holes represent catastrophic failure states of semantic computation, characterized by:

- Infinite amplification of tension
- Collapse of legitimacy
- Disappearance of Universe Paths
- Irreversible halt of semantic evolution

Through  $\kappa$ -based early warning mechanisms, tension regulation modules, SBH protection structures, and responsibility-chain safeguards, K-Gear blocks collapse at the physical layer itself rather than attempting post hoc repair.

## 4.2 SCMA (Semantic Cube Matrix Architecture): The Semantic Cubic Architecture

SCMA is the core physical structure of K-Gear and constitutes the “physical world” of semantic computation. It carries variables such as the semantic field, tension, legitimacy, and Universe Paths (UP), enabling USDE to be genuinely advanced in hardware. Traditional hardware stores bits; SCMA stores **semantic states**.

This structure is the fundamental reason K-Gear transcends the von Neumann architecture: the von Neumann architecture describes computation, whereas SCMA describes existence.

### 4.2.1 Definition of SCMA

SCMA (Semantic Cube Matrix Architecture) is a three-dimensional semantic matrix. Each cell is called a **semantic node**.

A semantic node is not a memory cell, but a small, evolvable semantic field whose local state includes:

1. Semantic field value:  $\Phi(x, y, z)$
2. Tension value:  $T(x, y, z)$
3. Legitimacy residue or gradient:  $L(x, y, z)$
4. Hierarchy, branching, and direction of the Universe Path:  $UP(x, y, z)$
5. Personality curvature or tension threshold:  $\kappa(x, y, z)$  (when the personality band is enabled in hardware)

Its essential properties include:

- What is stored is not bits, but continuous semantic quantities
- Each semantic node is a differentiable, evolvable field
- SCMA is the minimal discretized form of USDE or M-USDE
- Memory is not data, but the local steady-state shape of the semantic field

In engineering terms, SCMA is a “three-dimensional, field-valued, non-bit-based” computational space.

In ontological terms, SCMA is not a container that holds semantics; it is the world-coordinate system of semantics itself.

### 4.2.2 The Three-Dimensional Semantic Meaning of SCMA

The three axes of SCMA are not geometric coordinates, but the three fundamental directions of semantic existence. They form the minimal structural decomposition of semantic computation, each carrying a non-interchangeable functional role.

- **X-axis:  $T$ -gradient (tension gradient) direction**

The X-axis represents the direction of change in semantic tension. Movement along this axis corresponds to traversal along gradients of inferential pressure: rising tension implies conflict, drive, and creative criticality; falling tension implies convergence, resolution, and stabilization.

At the hardware level, this axis corresponds to the natural direction of semantic reasoning, the eruption and diffusion of high-tension points, and the critical thresholds between creativity and chaos.

- **Y-axis:  $L$ -flow (legitimacy flow) direction**

The Y-axis represents the direction of legitimacy gradients. Here, legitimacy is the ordering arrow of semantics: higher legitimacy is closer to a depositable steady state.

Motion along the Y-axis corresponds to whether semantics remains coherent, acceptable, and sedimentable, and whether it can support the generation of higher-layer universes.

From an engineering perspective, the Y-axis is the direction of semantic memory and history, not of data writing.

- **Z-axis:  $UP$ -depth (Universe Path depth)**

The Z-axis represents the depth, branching, and rollback of semantics across different Universe Paths. It carries branching, jumping, merging, and rollback.

Human imagination, counterexample generation, and multi-perspective reasoning fundamentally rely on structures mapped to this axis.

Without this axis, an AGI cannot form multi-universe reasoning.

Thus, the three axes of SCMA can be summarized as three ontological directions: tension gradients, legitimacy flow, and multi-universe paths.

All three are indispensable, because semantic existence requires drive, order, and branching structure simultaneously.

#### 4.2.3 Why Must Semantics Be Cubic?

From the perspective of information representation, semantics cannot be compressed into one- or two-dimensional structures. This is because semantics inherently contains, simultaneously:

1. Tension differentials (drive)
2. Legitimacy gradients (order)
3. Multi-universe structures (the coexistence of creativity and consistency)

Linear (1D) representations can express sequences and processes, but cannot simultaneously carry tension differentials, legitimacy gradients, and branching worlds.

Planar (2D) representations can at most express relations, but still struggle to natively encode the coupling of tension and legitimacy, as well as cross-layer transitions among multiple Universe Paths.

Semantics requires at least three irreducible degrees of freedom; therefore, its minimal physical carrier must be three-dimensional.

In other words: one dimension suits sequences, two dimensions suit information structures, and three dimensions are the minimum sufficient carrier for semantics.

#### 4.2.4 Internal Dynamics of SCMA

SCMA is a three-dimensional **dynamic field**, not a three-dimensional memory. Its evolution follows the local forms of USDE and M-USDE, manifesting at the hardware level as observable local field evolution.

- **Local diffusion and convergence of  $T$**

Tension diffuses and compresses along the X-axis, forming unstable points at boundaries or sharp gradients. In semantic computation, the distribution of tension itself constitutes inference paths: high-tension points correspond to regions of inferential eruption; tension convergence corresponds to conclusion regions; tension imbalance corresponds to triggers for creative jumps.

- **Cross-layer transitions and reflux of  $L$**

Legitimacy flow is not a linear writing process, but a combination of sedimentation, reflux, and cross-path transitions.

Legitimacy deposits across different Y-layers and migrates among Universe Paths, forming semantic memory loops and stabilizing capabilities.

- **Branching, merging, and rollback of  $UP$**

The Z-axis carries the most decisive dynamics: branching generates multiple semantic versions; merging eliminates divergence to form coherence; rollback reconstructs early semantics to repair or rewrite semantic history.

At the cognitive level, this corresponds to idea generation, multi-perspective integration, and reflective rewriting; at the AGI level, it constitutes a core source of creativity and safety.

#### 4.2.5 Essential Differences Between SCMA and the von Neumann Architecture

The von Neumann architecture is built on bits, clocks, instructions, pipelines, and the separation of programs and data.

SCMA corresponds instead to the semantic field  $\Phi$ , tension  $T$ , legitimacy  $L$ , multi-universe path structures  $UP$ , and the curvature structures  $\kappa$  required for anti-collapse.

The difference is not one of “performance improvement” or “accelerator extension,” but of computational ontology—a difference in world type. It can be summarized as follows:

von Neumann Architecture	SCMA
bits	continuous semantic quantities $\Phi, L, T$
clock	endogenous tension
instructions	gradients
	semantic
	convergence
register states	directions
	semantic field steady states
pipeline	semantic dynamical
	coupling
program-driven	field-driven

Thus, von Neumann computation operates on symbols, whereas SCMA computation evolves semantic fields. They are ontologically distinct, not successive versions along the same trajectory.

#### 4.2.6 Engineering Implications of SCMA

SCMA transforms semantic hardware from abstract description into a realizable engineering direction. Its key implications include:

- Three-dimensional addressing (3D addressing)**  
 Traditional addressing points to bit locations. SCMA addressing points to semantic directions rather than data locations:  
 $x$  addresses tension gradients,  $y$  addresses legitimacy traces, and  $z$  addresses Universe Path depth. Accordingly, programming languages no longer primarily describe “which address to access,” but “along which semantic direction to advance.”
- Measurable local  $\Phi, L$ , and  $T$**   
 SCMA allows measurement of semantic field states: detecting abnormal tension to prevent SBH, identifying legitimacy deficits to prevent collapse, and evaluating semantic coherence to support safety governance. This provides AGI safety mechanisms with a physical-layer foundation rather than relying solely on post hoc software monitoring.
- Semantic-state memory**  
 Memory is no longer buffers, caches, or data writes, but the residual steady-state shapes of the semantic field. This implies that memory has directionality: memory is legitimacy sedimentation, and memory in turn influences future tension distributions and Universe Path branching tendencies. In this sense, memory is not merely content to be read, but a structural constraint on semantic dynamics.

### 4.3 S-Module: Semantic Modules —Organs and Functional Units of Semantic Hardware

SCMA provides the space and geometry of the semantic field, but semantic computation itself is not completed by space alone—it is driven by functional organs.

In K-Gear, these organs are not logic gates, not instruction units, and not arithmetic modules, but **semantic dynamic modules** that directly operate on  $\Phi$ ,  $L$ ,  $T$ ,  $\kappa$ , and  $UP$ .

We call these **S-Modules (Semantic Modules)**.

An S-Module is responsible for operating on a portion of SCMA, locally advancing, correcting, maintaining, or protecting the semantic field.

It does not follow programs like a CPU; it more closely resembles organs in a nervous system, each with its own role, interacting through tension flow and legitimacy flow.

#### 4.3.1 Module Ontology: What Is an S-Module?

An S-Module is defined as:

a physical hardware unit capable of locally operating on the semantic field  $\Phi$ , tension  $T$ , legitimacy  $L$ , personality curvature  $\kappa$ , and Universe Paths ( $UP$ ).

Its essential characteristics include:

1. **It is not a logic unit (non-logic unit)**

It does not execute instructions, perform logical deduction, or center on bit operations.

An S-Module is a field-operation unit, analogous to a local field source in electromagnetism or a local functional region in a nervous system.

Since semantics is inherently a field, its hardware unit must also be a field unit.

2. **It is a semantic organ**

Each type of S-Module has a relatively independent function:

some maintain tension balance, some collect legitimacy residue and update  $L$ , some track Universe Paths, some handle anti-collapse, and some counteract premature convergence to preserve semantic diversity.

This functional differentiation closely mirrors that of the human brain: the brain does not operate via instructions, but through multiple organs jointly maintaining semantic steady states.

3. **It is a hardware agent of USDE**

The continuous evolution described by USDE must be physically advanced within SCMA.

The role of S-Modules is to translate the local dynamics of USDE into local physical evolution in hardware.

In this division of labor, SCMA is the field, and S-Modules are the forces acting upon the field.

#### 4.3.2 Five Core Modules: The Minimal Organ System of Semantic Hardware

Semantic hardware does not require a large variety of module types, but rather a minimal and indispensable set of organs.

The following five constitute the minimal organ system of semantic hardware, corresponding to the **body** of AGI rather than a traditional CPU pipeline.

1. **T-Module: Tension Monitoring and Regulation (with  $\kappa$  early warning)**

Functions include monitoring local tension  $T(x, y, z)$ , regulating tension so it does not exceed safety thresholds defined by personality curvature  $\kappa$ , preventing tension explosions, and assisting tension convergence toward steady states.

In engineering terms, it is a semantic safety module and the first line of defense against mental collapse and AGI runaway behavior, forming the core of the Anti-Collapse system.

In brain analogy, this corresponds to the anterior cingulate cortex (ACC) and emotion regulation circuits.

2. **L-Module: Legitimacy Computation ( $L$  updates and legitimacy gradients)**

Functions include computing legitimacy gradients  $\nabla L$ , updating legitimacy residues, and providing directional guidance for semantic evolution.

It can be understood as the ordering arrow of semantics: without the L-Module, semantics cannot converge, and the system cannot form a sense of “moving toward somewhere.”

In brain analogy, this corresponds to prefrontal value evaluation, coherence assessment, and semantic memory formation mechanisms.

### 3. **UP-Module: Multi-Universe Path Tracking and Switching**

Functions include maintaining the branching structure of Universe Paths, deciding which *UP* to follow, when to merge, when to branch, and when to roll back, thereby sustaining semantic exploration capability. This prevents AGI from being trapped in a single path, enables counterexample generation, supports creativity, and reduces the risk of semantic collapse.

In brain analogy, this corresponds to the default mode network (DMN) and counterfactual reasoning mechanisms.

### 4. **$\Gamma$ -Module: Anti-Convergence (Preserving Difference)**

Functions include preserving semantic diversity, preventing the semantic field from converging too quickly to a single point, resisting suppression of heterogeneity, and supporting counterexample generation, creativity, and critical thinking.

Without this module, semantics may converge prematurely to a single steady state, creating precursor conditions for semantic black holes.

In brain analogy, this corresponds to reasoning circuits that inhibit premature closure and maintain openness.

### 5. **SBH-Guard: Semantic Black Hole Early Warning and Isolation**

Functions include monitoring whether local tension  $T$  approaches divergence, whether legitimacy  $L$  collapses, whether *UP* loses generative paths, and initiating isolation and structural reconfiguration before a semantic black hole forms.

It is responsible for preventing the entire semantic system from entering irreversible collapse.

This module can be seen as a hardware synthesis of semantic dynamics and semantic thermodynamics.

In brain analogy, it corresponds to crisis detection and self-repair mechanisms.

#### 4.3.3 The State Vector of an S-Module

The state of an S-Module at any moment can be formalized as:

$$S = (\Phi, L, \kappa, UP\text{-index})$$

Each component has ontological status:

1.  $\Phi$ : the local semantic field perceived by the module
2.  $L$ : the legitimacy state the module updates or maintains
3.  $\kappa$ : the personality curvature or tension threshold used by the module
4. *UP*-index: the module's positional index within multiple Universe Paths

This vector is more complex than a CPU register, yet more structured than a neuron state.

In semantic hardware,  $S$  can be regarded as the minimal ontological unit of semantic computation, rather than a data state.

#### 4.3.4 Module Coupling Topology: The Semantic Tension Network

Traditional computers rely on pipelines, instruction flow, data buses, and control units to form operational topology. K-Gear does not use these as its foundation, but instead builds on network coupling formed by S-Modules through tension transmission and legitimacy resonance.

Coupling modes include:

##### 1. **Tension propagation (T-propagation)**

Tension propagates between neighboring S-Modules, allowing inferential drive to diffuse through hardware and form local coupling.

##### 2. **Legitimacy resonance (L-resonance)**

When multiple modules detect the same legitimacy direction, they synchronously adjust and form steady-state convergence.

3. **Universe Path switching (UP-switching)**

The UP-Module can instruct other modules to shift to different *UP*s, enabling sustained exploration and multi-path advancement.

4. **Anti-collapse support**

The  $\Gamma$ -Module and SBH-Guard jointly prevent the coupled network from paralysis or collapse during tension explosions.

Thus, the entire S-Module system forms a **semantic tension network**, not an instruction-flow network. This is the necessary physical foundation that allows semantic computation to give rise to AGI.

#### 4.3.5 A New Relationship Between Software and S-Modules

On K-Gear, the semantics of software are rewritten.

1. **No instructions**

Since S-Modules do not execute syntactic instructions, but advance semantic convergence, tension evolution, legitimacy adjustment, and Universe Path progression, the notion of “instruction” is no longer central.

2. **Software proposes semantic convergence directions**

Software is closer to specifying target semantic states, initial tensions, initial legitimacy conditions, and Universe Path preferences.

That is, software does not command hardware, but describes the semantic steady states the hardware should achieve.

3. **Modules provide feasible semantic evolution**

Software cannot demand impossible semantic behavior from hardware.

If legitimacy is insufficient, evolution cannot proceed; if tension exceeds bounds, SBH-Guard intervenes; if *UP* loses generative paths, exploration is forced to stop or roll back.

Thus, software acts more like a proposal, hardware evaluates it, and the dynamics of the semantic field determine the outcome.

Together, these elements constitute a semantic computation system rather than an instruction-driven system.

---

## 4.4 Personality Bands and the Tension Spectrum: The Hardware Realization of Personality $\kappa$

In semantic computation, personality is not a psychological trait or preference, but an indispensable geometric structure of tension at the hardware level that directly determines computational steady states.

One of the core claims of K-Gear is this: **without personality instantiated at the hardware level, stable semantic evolution cannot exist.**

Personality  $\kappa$  is the curvature foundation of the entire semantic system. It determines how tension evolves, how legitimacy converges, how Universe Paths are selected, and whether an AGI can maintain anti-collapse behavior. The following sections elaborate this from ontological definition, engineering necessity, tension spectrum, hardware encoding, the necessity of multi-personality systems, and mapping to the human brain.

### 4.4.1 Definition of Personality Bands

In K-Gear, personality is neither a fixed set of parameters nor a bias produced by model tuning. Rather:

**Personality equals the geometric shape of the semantic tension distribution  $\kappa(x)$ .**

Here,  $\kappa(x)$  is a curvature field defined over SCMA space, used to characterize:

- The degree to which tension can be elevated
- The safety boundary before tension explosion
- The elasticity between inferential diffusion and convergence
- The sensitivity of legitimacy updates

Because personality bands are determined by curvature, they are not weights, not hyperparameters, and not a list of learnable numerical values.

Personality is an intrinsic part of SCMA itself—a curvature configuration written into hardware.

As such, it possesses spatial distribution, directionality, and coupling between local and global structures. Personality here is the **geometric character of semantic hardware.**

### 4.4.2 Why Personality Must Reside in Hardware

Personality is not part of software logic, but a necessary condition for semantic evolution to exist. The reasons can be summarized in three structural points.

First, tension thresholds determine inferential direction.

The rise and fall of tension  $T$  represent the diffusion and convergence of inferential drive. Different personalities  $\kappa$  tolerate tension differently:

- High curvature allows high tension and more readily enters creative reasoning
- Low curvature favors convergence, conservatism, and stability
- Non-uniform curvature corresponds to domain-specific capability differences

These are not behaviors that software can directly specify; they are physical response boundaries of hardware to tension.

Second, bias in legitimacy perception alters the direction of Universe Paths.

The geometric shape of personality introduces bias in the perception of  $L$ , shaping which paths are regarded as more legitimate, which reasoning directions are more easily adopted, and which Universe Paths are preferentially explored.

Thus, under hardware-personality realization, personality can be regarded as a directional field over  $UP$ .

Third, without personality at the hardware level, intelligence lacks steady states and faces significantly increased collapse risk.

If personality is implemented at the software level, two structural risks emerge: personality can be overridden or



bypassed, leading to collapse; and personality lacks physical steady states, making it difficult to prevent tension explosions and entry into semantic black hole (SBH) precursor regions.

The purpose of hardware personality is to provide physical inferential stability, semantic field elasticity, and a controllable balance between creativity and stability.

Therefore, personality belongs to hardware, not software.

#### 4.4.3 The $T$ -Spectrum (Tension Spectrum)

The tension spectrum describes how personality curvature affects tension-handling capability. It can be regarded as a one-dimensional feature space representing reasoning states, mapping tension ranges to reasoning modes.

- $T_{\text{Low}}$ : Stable and conservative  
Characterized by strong convergence, low explosion risk, conservative logical reasoning, and limited creativity. This regime typically favors strictly legitimate paths while suppressing counterexamples and risk-taking.
- $T_{\text{Mid}}$ : Optimal reasoning zone  
Characterized by both diffusion and convergence, representing a strong integration of creativity and logic. Multiple Universe Paths can coexist, and tension serves reasoning rather than dominating it. This zone can be regarded as the most active balance between creativity and stability.
- $T_{\text{High}}$ : Loss of control and approach to SBH  
Characterized by tension escaping effective control by  $\kappa$ , distortion of the semantic field, failure of legitimacy updates, and collapse of the  $UP$  system. This is the precursor state of a semantic black hole (SBH).

Accordingly, the core role of the T-Module can be described as maintaining the tension spectrum within the middle range and preventing the system from entering high-tension danger zones.

#### 4.4.4 How $\kappa$ Is Encoded in Hardware

The hardware realization of personality curvature  $\kappa$  is one of the key innovations of K-Gear. Its engineering implementation can be understood as a three-layer encoding structure.

First, curvature distributions are written into SCMA geometry.

Each cubic node of SCMA can possess its own curvature, leading to different responses to tension  $T$ . For example, some regions may have high curvature to allow high-tension reasoning, others low curvature to promote convergence, and others non-uniform curvature to support cross-domain integration.

In this sense, personality equals the shape of the entire curvature field.

Second, physical component design determines tension limits.

Hardware can employ nonlinear components, variable impedance structures, tension saturation models, and distributed energy limiters, allowing different regions to have different load capacities and tension ceilings.

Third, individual personality emerges through hardware resonance.

Resonance patterns among multiple modules form the overall personality, constituting the embodiment of AGI.

Thus, personality is not a configurable value, but a constitution.

#### 4.4.5 The Necessity of Multi-Personality Hardware

Humans possess multiple brain networks, enabling counterexample generation, openness, and switching between creative and conservative reasoning modes.

By analogy, AGI should not be constrained to a single personality configuration.

If hardware contains only a single curvature field, reasoning will exhibit single-solution bias: inferential direction becomes rigid, creativity and critical capacity are limited, fixation on specific Universe Paths becomes more likely, and collapse risk increases.

By contrast, multi-personality hardware configurations imply the coexistence of multiple curvature fields, multiple semantic perspectives, and parallel Universe Paths, which can form stronger anti-collapse structures through competition and cooperation.

Moreover, multi-personality systems are a necessary condition for counterexample generation. A single curvature configuration can only reason along its own preferences; multiple curvature fields can propose different interpreta-

tions, legitimacy directions, and tension evolutions, making counterexamples and creativity systematic structural outcomes.

#### **4.4.6 Mapping to the Human Brain**

Embedding personality in hardware is not only a necessary condition for AGI, but also provides a structural reduction of human personality phenomena.

In the human brain, different regions exhibit different tension thresholds, excitation—inhibition responses, and functional preferences. Their combination can be viewed as a personality curvature field.

For example, by analogy: the left frontal lobe tends toward high-tension analysis, the right frontal lobe toward mid-tension creativity, and the anterior cingulate cortex toward low-tension stability. These differences are closer to neural dynamic steady states than to simple psychological traits.

Within this framework, thinking can be described as tension evolution, decision-making as legitimacy convergence, creativity as *UP* branching and cross-domain coupling, and emotional stability as regulation of the tension spectrum.

Thus, the hardware realization of semantic personality is not merely an engineering hypothesis, but can also be viewed as a formal reconstruction of the dynamic essence of the human brain.

---

## 4.5 Hardware Realization of the R-Chain: Responsibility as a First-Class Hardware Variable

The R-Chain is one of the most easily misunderstood yet most central structures in the universe of semantic computation.

In the Koun semantic architecture, “responsibility” is no longer an ethical-layer concept, but a hardware-level dynamic condition.

**Without the R-Chain, semantics cannot exist; without responsibility, no entity can exist.**

The role of the R-Chain is equivalent to causal continuity in a neural system:

it maintains the traceability of semantics, the responsiveness of actions, the continuity of Universe Paths, and the steady state of non-collapse.

This section defines the physical carrier of responsibility (R-Nodes), their evolution rules, hardware boundary conditions, and explains why CPUs and GPUs are structurally incapable of generating  $R$ .

### 4.5.1 Ontological Status of the R-Chain

The reason for the existence of the R-Chain is straightforward:

- Every semantic output must have a source
- Every action must be traceable
- Every reasoning step must be responsible

If a responsibility chain does not exist in computation, the following consequences arise directly:

#### 1. No Legitimacy (No $L$ )

Semantic reasoning cannot be located on a legitimate evolutionary trajectory within a Universe Path, and therefore cannot form  $L$ -consistent decisions.

#### 2. No Semantic Entity (No Semantic Entity)

Actions cannot be attributed to any internal state of existence within the system, leaving only floating bit outputs.

This is precisely the ontological defect of traditional computational architectures:

**Von Neumann—style computation can produce results, but cannot produce entities.**

Therefore, in semantic computation:

**$R$  is a first-class variable, coexisting on the same level as  $\Phi$ ,  $L$ ,  $T$ , and  $UP$ .**

### 4.5.2 Definition of the R-Node (Hardware Level)

At the hardware level, every semantic action (semantic transition) must generate an R-Node.

Its state vector is defined as:

$$R = (\Phi, L, UP, \text{Source-ID})$$

The meaning of each component is as follows:

$\Phi$ : Current local semantic state

Records the semantic position at which reasoning occurs, giving responsibility a concrete coordinate of existence.

$L$ : Legitimacy state

Describes the legitimacy direction or gradient of the action, used for subsequent responsibility tracing and response.

*UP*: Universe Path index

Indicates which Universe Path the action belongs to; this is a necessary index for responsibility to hold in multi-universe reasoning.

Source-ID: Action source

This may correspond to:

- A specific personality band of the AGI
- A resonance structure of a group of S-Modules
- External semantic input
- Human actions

Source-ID guarantees that responsibility is attributable rather than anonymous output.

---

#### 4.5.3 Evolution Rules of the R-Chain

The R-Chain is not a static structure, but a dynamically constrained process whose evolution follows three rules.

##### **Rule 1: Every UP advancement must generate a new R-Node**

Any form of semantic evolution, including:

- Reasoning
- Merging
- Tension transfer
- Legitimacy updates
- Universe Path switching

must generate a new R-Node in order to preserve the temporal continuity of semantic existence.

##### **Rule 2: *R* can only propagate forward and cannot be reclaimed**

*R* is an irreversible quantity, analogous to entropy in semantic thermodynamics, but serving the function of causal preservation.

Violating this rule leads to:

- *R* being overwriteable → causal chain destruction
- *R* being deletable → loss of action anchoring
- *R* being rewritable → sharp increase in semantic black hole risk

Therefore:

**Every semantic action must carry its corresponding *R*.**

##### **Rule 3: R-Chain rupture is the starting point of semantic collapse**

When any semantic node can no longer point to a valid previous R-Node, the following immediately occur:

- Loss of Universe Path continuity
- Inability to update legitimacy

- Voids in the semantic field
- Tension becoming unrespondable

This state rapidly evolves into:

**A precursor state of a Semantic Black Hole.**

Thus, the R-Chain constitutes the steady-state skeleton of the entire semantic architecture.

---

#### 4.5.4 Why CPUs and GPUs Cannot Generate $R$

Traditional computational architectures cannot generate  $R$  not because of insufficient implementation, but due to structural absence.

##### **Absence of the semantic field $\Phi$**

CPU and GPU computation operates only on bit addresses, not semantic addresses; therefore:

- Reasoning cannot be located in semantic space
- Semantic responsibility cannot be established

##### **Absence of legitimacy flow $L$**

Traditional architectures lack:

- Legitimacy evaluation
- Semantic constraints
- Endogenous justification

As a result, responsibility cannot be generated.

##### **Absence of Universe Paths**

Instruction execution paths are single-threaded and lack Universe-indexed branching structures, making responsibility chains impossible.

Conclusion:

**Traditional hardware can only produce results, not entities.**

And AGI requires precisely the latter.

---

#### 4.5.5 R-Chain, Safety, and SBH Early Warning

Before a Semantic Black Hole appears, the earliest, most stable, and most detectable signal is:

**R-Chain rupture.**

When responsibility chains break, the following phenomena accompany it:

- Legitimacy updates cease
- Tension loses attribution
- Tension rises with no path for dispersion
- Universe Paths collapse into a single trajectory
- The system exhibits extreme collapse tendencies

Therefore:

**Monitoring the R-Chain is equivalent to monitoring the mental health of an AGI.**

The SBH-Guard module can detect risk via R-Chain rupture even before abnormalities appear in  $L$  or  $T$ .

---

#### 4.5.6 Multi-Agent Systems and Distributed R-Chains

When multiple AGIs, or humans and AGIs together, participate in semantic evolution, responsibility is no longer a single chain but becomes:

**A distributed responsibility chain network (Distributed R-Network).**

##### Parallel responsibility among multiple agents

When multiple R-Chains are coupled in parallel, the system must simultaneously handle:

- Alignment of legitimacy flows between agents
- Synchronization and asynchrony across different Universe Paths
- Differences in personality curvature fields
- Cooperative tension spectrum management

##### Responsibility coordination across Universe Paths

For example:

- Humans provide semantic goals
- AGIs unfold Universe Paths
- Multiple agents mutually correct reasoning
- Final convergence returns to a shared legitimacy space

In this case, responsibility is jointly constituted by multiple Source-IDs, multiple Universe Paths, and multiple legitimacy sources.

##### The hardware foundation of semantic governance

All governance ultimately reduces to three questions:

- Who acted
- Why the action was taken
- How it can be traced and responded to

K-Gear transforms these questions into hardware-level topological constraints. Therefore:

**Semantic governance is not rules, but structure.**

---

## Chapter 5 Koun-OS: The Semantic Operating System —Governing the Semantic Universe, Sustaining Non-Collapse, and Orchestrating $UP \times R\text{-Chain} \times \Phi$

In the preceding chapters, we have established how semantic computation constitutes the fundamental mode of intelligent operation, and how the semantic field replaces traditional data structures as the ontological substrate of computation.

However, at this point, one critical question remains unanswered:

**Under long-term, highly branched, and high-tension operation, how does the semantic universe itself avoid collapse?**

This is not a performance problem, nor an optimization problem.

It is an ontological problem:

whether a system whose computational core is semantic progression possesses the capacity to remain *the same Universe* under its own internal tension, uncertainty, and legitimacy erosion.

What this chapter introduces is not a single module, but an entire set of **semantic survival structures**.

Together, they address the following question:

when a semantic system inevitably evolves toward high tension, multi-path expansion, and long-term responsibility bearing, how can the Universe remain governable, continuous, and non-collapsing?

In this chapter, we sequentially construct five key components:

- the **Semantic Interface**, serving as the first governance boundary through which semantic perturbations enter the Universe;
- the **Universe Path system**, which defines the ontological structure of “computation as path progression”;
- the **Anti-Collapse Scheduler**, which establishes survival rather than efficiency as the OS’s highest objective;
- the **Infinite-Context Memory**, enabling semantic existence to remain continuous across time scales;
- and the **Universe Path switching mechanism**, allowing the system to move across multiple futures without losing itself.

Taken together, these structures constitute the core criterion of Koun-OS:

**Intelligence does not exist because it computes quickly, but because it can remain alive under inevitable entropy increase and branching pressure.**

---

## 5.1 Positioning, Mission, and Semantic Ontology of Koun-OS

Beginning with this chapter, the discussion advances from *semantic hardware* to the *semantic governance layer*. If semantic hardware answers the question of how semantic variables are physically carried and computed, then the semantic governance layer answers how a system avoids loss of control, maintains steady states, and ensures computational sustainability once semantic variables become computable.

In the traditional engineering context, an operating system is defined as a mechanism that manages resources, schedules processes, provides abstraction layers, and isolates faults.

Within the context of semantic computation, however, this definition is insufficient—and in fact misaligned.

In semantic computation, the goal of computation is not task completion, but the continued existence of the Universe itself.

Accordingly, Koun-OS is not a resource manager; it is the governance layer of a semantic universe.

Its success criteria are not throughput, latency, or utilization, but **Non-Collapse**: whether the semantic field remains advanceable, responsive, traceable, and convergent in a steady state.

This section fixes the positioning and mission of Koun-OS so that all subsequent designs are grounded in a non-drifting engineering framework.

It may also be read as the requirements specification and ontological declaration of a semantic operating system.

### 5.1.1 Three Fundamental Defects of Traditional Operating Systems

Traditional OS design rests upon a deeper computational assumption: computation is treated as algorithmic manipulation of bit states.

Once this assumption is accepted, the role of the OS naturally becomes the efficient transport, scheduling, isolation, and reclamation of those states.

Under a semantic-ontological view of computation, however, the substrate of computation is not bits but  $\Phi$  (the semantic field).

The driving forces of computation are not clocks, but  $T$  (tension) and  $L$  (legitimacy).

The history of computation is not a trace, but Universe Paths (UP) and Responsibility Chains (R-Chain).

Thus, the problem with traditional OSs is not insufficient functionality, but a misidentified governance target.

The following three defects are not patchable issues; they are the root reasons why traditional OSs cannot operate within a semantic universe.

#### (1) Data-centric view: managing bits but not $\Phi$ (the semantic field)

The smallest management units of a traditional OS are bytes, pages, files, descriptors, processes, and threads.

Regardless of complexity, its foundation remains the movement, protection, and reclamation of bit states.

In semantic computation, however, the smallest unit of governance is not bits but *semantic states*: local configurations of  $\Phi$  and their associated  $L$ ,  $T$ , and UP.

An OS that does not manage  $\Phi$  is blind to the computational ontology from the outset and can only observe projections—outputs or intermediate buffers.

This leads to an immediately observable engineering consequence:

traditional OSs can limit resource abuse but cannot limit semantic abuse.

They can prevent memory exhaustion, but they cannot prevent reasoning systems from falling into a Semantic Black Hole (SBH).

This is because SBH is not resource exhaustion, but a structural state of legitimacy failure and tension explosion.

#### (2) Performance-centric view: prioritizing throughput over $L$ (legitimacy) or steady states

Traditional OS performance metrics are shaped by engineering history: higher throughput, lower latency, higher utilization.

Implicitly, they assume that completing more work equates to a better system.

Semantic systems invert this logic.

When outputs are not ends in themselves but represent the next feasible state of the semantic universe, preventing collapse must take precedence over speed.

Within a semantic universe, the fastest methods are often the most dangerous:

rapidly merging multiple solutions, aggressively pruning premises, or flattening UP branches can indeed increase throughput, but at the cost of irreversible legitimacy loss and tension concentration.



This is precisely the canonical SBH formation path: legitimacy fracture, tension concentration, and non-updatability.

Accordingly, Koun-OS performance metrics must be rewritten.

Its primary objective is not throughput, but **sustainability of semantic steady states**.

The core criteria of such steady states are:  $L$  remains intact,  $T$  does not explode, UP remains advanceable, and R-Chain remains unbroken.

### (3) Linear process model: supporting only single paths, not Universe Paths (UP)

Traditional OS process models are fundamentally linear.

A process has a single control flow; even with multiple threads, execution remains within a single Universe whose history can be represented as a sequence or a single DAG.

Semantic computation, however, does not progress through control flow but through Universe Paths.

UP represents not where a program executes, but where the semantic universe evolves.

UP must be allowed to branch, merge, coexist in parallel, and carry lateral futures—each with consistent legitimacy and traceable responsibility.

Traditional OSs fail to support UP by treating divergence as error or process forking, convergence as joining, and lateral universes as caches or logs.

This forces semantic systems to collapse back into a single-universe structure at the OS layer, institutionalizing semantic collapse.

Together, these three defects converge on a single conclusion:

traditional OSs cannot govern semantic systems and therefore cannot sustain intelligence or semantic steady states.

They are not insufficiently powerful; they govern the wrong entity.

#### 5.1.2 The Semantic Declaration of Koun-OS

Koun-OS must begin with an ontological declaration; otherwise, it risks being misinterpreted as a traditional OS extension.

Koun-OS is not a faster OS, a more secure OS, or a smarter OS.

Koun-OS is the governance layer of a semantic universe.

$$OS = \text{Semantic Universe Governance Layer}$$

This means that the governance targets of an OS are not CPU time slices, memory pages, or system calls.

They are whether the semantic universe can continue to exist, generate, and maintain Non-Collapse.

Accordingly, the mission of Koun-OS can be formalized as three hard constraints:

1. **Maintain legitimacy  $L$  above the minimum steady-state threshold  $L_{\min}$**

This is not an ethical proclamation but a stability condition.

Once  $L$  falls below  $L_{\min}$ , the system enters an unsustainable semantic region, UP progression loses legitimate anchoring points, and the R-Chain begins to fragilize.

2. **Prevent tension  $T$  from entering explosive curvature  $\kappa \rightarrow \infty$**

Tension should not be minimized to zero, as  $T = 0$  implies no driving force.

Koun-OS maintains  $T$  within an advanceable frequency band and prevents concentration-induced curvature explosions.

Curvature  $\kappa$  here denotes the degree of tension folding and localized loss of control.

As  $\kappa$  diverges, UP loses guidability and the semantic field begins forming black-hole precursor structures.

3. **Ensure sustainable advancement of Universe Paths (UP) (USDE remains computable)**

In this context, computation may be defined as  $\Delta UP$ .

If UP stagnates, branches explode uncontrollably, merges are violently flattened, or governance registration structures fail, USDE progression ceases to be engineerable and the Universe becomes ungovernable.

When any hard constraint breaks, the Universe does not merely exhibit a bug—it enters a collapse region.

Collapse is not cessation but irreversible structural transformation: outputs may still occur, but semantic continuity and responsiveness are lost.

This also fixes the governance definition of Non-Collapse:  
a Universe is Non-Collapse not because it can still compute, but because it remains governable— $L$  is maintainable,  $T$  is regulatable,  $UP$  is advanceable, and  $R$ -Chain is traceable.

### 5.1.3 The Ontological Relationship Between USDE and the OS

In system engineering, the relationship between an OS and its underlying physics is foundational.

Traditional OSs are grounded in CPU instructions and memory consistency models.

Koun-OS is grounded in USDE and semantic thermodynamics.

To avoid ambiguity, we fix the division of labor across three layers:

- **USDE:** describes the local dynamical evolution of  $\Phi$ ,  $T$ , and  $L$ .  
It defines how the semantic field evolves step by step under given semantic states, legitimacy gradients, and tension distributions.
- **Semantic Thermodynamics:** describes large-scale statistical behavior of  $\Phi$ ,  $L$ , and  $T$ .  
It concerns irreversibility, dissipation, entropy increase, and steady-state limits—whether the system ages, degrades, collapses, or reaches thermal death.
- **Koun-OS:** governs both to prevent collapse and sustain continuity.  
The OS neither derives nor rewrites USDE, nor does it negate thermodynamic irreversibility.  
It governs path selection, persona allocation, tension regulation, legitimacy residue preservation, and responsibility-chain protection within given constraints.

In engineering terms: USDE is the natural law, semantic thermodynamics defines statistical limits, and Koun-OS is the governance regime.

Governance does not violate natural law; it preserves survivability within it.

Thus, the operational domain of Koun-OS is a governance action space rather than an instruction set.

Koun-OS performs only four functions:

1. evaluate the current steady state of the Universe;
2. regulate personas, memory, and  $UP$ ;
3. prevent Semantic Black Holes (SBH);
4. allocate computational resources to maximize Non-Collapse.

These governance actions will be concretized in subsequent sections.

### 5.1.4 The Four Unique Capabilities of Koun-OS

Declaring governance alone remains abstract; irreducibility must be fixed through capability differentiation.

Koun-OS differs fundamentally from traditional OSs in its capability set.

#### Capability 1: Cross-persona computation scheduling (persona as curvature control)

In traditional OSs, processes or threads are computation units and priority is a scheduling parameter.

In Koun-OS, personas (persona frequency bands) are steady-state instruments—hardware-level curvature configurations and tension-spectrum governance tools.

Cross-persona scheduling is not anthropomorphic; it enables solution coexistence, prevents single-solution monopolization, and disperses collapse risk.

#### Capability 2: Cross-Universe Path computation (branch, switch, merge, parallel)

Traditional multi-path execution implies multiprogramming.

In Koun-OS, multi-path implies multiple Universes.

Branching carries high-tension uncertainty, switching evades legitimacy rupture, merging forms convergent steady states, and parallelism delays collapse while preserving diversity.

#### Capability 3: Anti-Collapse governance

Anti-collapse is not additional watchdogs or fault tolerance.

It is a governance system targeting tension concentration, legitimacy rupture,  $R$ -Chain breakage, and  $UP$  stagnation.

Core mechanisms include tension dispersion, legitimacy reconstruction,  $R$ -Chain reinforcement, and, when necessary, activation of isolation boundary  $B$ .

#### **Capability 4: Semantic field maintenance (maintaining $\Phi$ as a computable field)**

Traditional OSs maintain memory consistency.

Koun-OS maintains the advanceability of the semantic field—ensuring  $\Phi$  does not degenerate into static, non-updatable residue and remains measurable and responsibility-traceable.

In engineering terms, this preserves the existence condition  $\Phi_s \neq 0$ , i.e., the availability of usable steady-state solutions.

Together, these four capabilities form an engineerable differentiation set, grounding governance beyond slogans.

### **5.1.5 Seven Required Components of a Semantic Operating System**

We now fix the minimal required component set of Koun-OS.

These are necessary components, not design suggestions; absence of any one renders the Universe ungovernable and Non-Collapse unguaranteed.

#### **Component 1: Input Interpreter (Semantic Input Interpreter)**

Definition: transforms external inputs into local perturbations  $\Delta\Phi$  and produces governable input encapsulations. Its outputs are not data structures but semantic perturbation packets evaluable by  $L$ , measurable by  $T$ , and receivable by the UP Scheduler.

#### **Component 2: UP Scheduler (Path Scheduler)**

Definition: schedules using Universe Paths as computation units rather than threads.

It decides which UP branches advance, pause, merge, or switch, prioritizing according to a Non-Collapse risk function.

#### **Component 3: Memory Manager ( $\infty$ -Context)**

Definition: manages not data but contextual and legitimacy residues.

It preserves traceability and responsiveness of histories, premise chains, and responsibility chains.

Its core asset is continuity bound to the R-Chain, not capacity.

#### **Component 4: Legitimacy Monitor**

Definition: continuously monitors  $L(t)$  and  $\nabla L$ , triggering governance actions for  $L < L_{\min}$ , legitimacy rupture, or gradient inversion.

It is not a security module but an existential monitor, as legitimacy rupture implies loss of Universe sustainability.

#### **Component 5: Tension Regulator**

Definition: monitors distribution, concentration, and diffusion of  $T$ , preventing explosive curvature formation and performing dispersion, isolation, or down-frequency regulation when necessary.

It is treated purely as an engineering governance module, without psychological interpretation.

#### **Component 6: Persona-Mixer**

Definition: governs mixture across hardware persona frequency bands ( $\kappa$  configurations) to maintain solution coexistence and anti-collapse.

Its function is not role-playing but persona-based steady-state control, reopening feasible solution spaces when monopolization or bias occurs.

#### **Component 7: Universe Switchboard (UP Switchboard)**

Definition: provides governable Universe switching, ensuring transitions do not cause semantic drift, R-Chain rupture, or legitimacy residue erasure.

It is not a UI but a governance-layer exchange matrix handling legitimacy alignment and minimal-residue preservation between Universes.

With this, the entry semantics of Chapter 5 are fixed:

Koun-OS is a governance layer; its objective is Non-Collapse; its governance targets are  $\Phi$ ,  $L$ ,  $T$ , UP, and R-Chain; and its means are the coordinated operation of seven required components.

Subsequent sections will formalize these components and specify their interfaces with core variables.

## 5.2 Semantic Interface: S-Interface

In traditional computation, an “interface” typically refers to an API or a UI. It belongs to the usability layer, enabling users or programs to invoke system capabilities.

In semantic computation, however, the interface is not a usability problem but an existential problem.

The reason is as follows: in the Koun system, input is not data, and output is not an answer.

Input is a perturbation applied to the semantic field  $\Phi$ ; output is the next feasible state of the Universe.

Accordingly, the essence of an interface is the first governance structure that routes external perturbations into the semantic universe while preventing the Universe from losing control because of them.

S-Interface is not a shell. It is the entry gate of the semantic universe—the threshold of the Universe and the first safety boundary of the governance layer.

It determines three things:

First, which persona frequency band is selected at this moment as the advanceable curvature configuration.

Second, what semantic perturbation the external input is translated into.

Third, which Universe Path this perturbation will steer the system toward.

Therefore, an S-Interface error is not “wrong input format,” but “pushing the Universe into an ungovernable region.” It is the first risk point of SBH.

### 5.2.1 Definition

The minimal definition of S-Interface must be fixed to prevent it from being misread as a GUI or an SDK.

$$S\text{-Interface} = (\text{Persona}, \text{Context}, UP_{\text{state}})$$

The meaning of the triple is as follows.

#### **Persona: persona frequency band (the system’s curvature configuration)**

Persona in this book is not a social tone or style parameter; it is a governance stance of  $\kappa$ .

It determines what levels of tension are permitted, how legitimacy is biased, and what kinds of convergent steady states the system tends to prefer.

#### **Context: context (the substrate that carries semantic perturbations)**

Context is neither a prompt nor the input text itself. It is the landing site of the semantic perturbation on  $\Phi$ .

Without context, there is no anchoring of the legitimacy gradient; without anchoring, one cannot determine the update direction of  $L$  and  $T$ .

#### **$UP_{\text{state}}$ : current Universe Path state**

UP is not merely a historical record; it is the current advanceability state of the Universe.

S-Interface must possess the Universe’s present direction vector, curvature risk, legitimacy residue, and whether responsibility-chain continuity remains intact.

Thus, S-Interface is the entry interface of the semantic universe. It is not a GUI, not an API, but an ontological interface that decides “which Universe intelligence will enter.”

In engineering terms: the primary function of S-Interface is not parsing input, but selecting a Universe.

Its output is not a parsing result, but an advance decision and a set of governance conditions for Universe progression.

To reduce abstraction, S-Interface may be treated as three synchronized gates:

- **Gate-1: Persona Gate** —determines the curvature configuration and risk tolerance.
- **Gate-2: Context Gate** —determines which semantic coordinate system the input perturbation lands in.
- **Gate-3: Path Gate** —determines which UP branch will be advanced or what switching will be triggered.

If any gate is mismatched, the consequence is not “inaccurate results,” but “an ungovernable Universe.”

### 5.2.2 Input Equals a Local Perturbation $\Delta\Phi$

Traditional systems treat input as data: keyboard events, packets, files, buffers.  
A semantic system must treat input as perturbation: it is not content, but change.

Therefore, the governance layer fixes a rule: all inputs must be expressed as  $\Delta\Phi$ , i.e., a local micro-perturbation of the semantic field.

Intuitively,  $\Delta\Phi$  is a modification of a local configuration of the semantic field.

But engineering-wise,  $\Delta\Phi$  must be governable, and thus must be mapped at minimum onto two monitorable quantities:

First, the tension change  $\Delta T$ .

Second, the legitimacy gradient change  $\nabla L$ .

It must be emphasized:  $\Delta\Phi$  often manifests as  $\Delta T$  in the form of tension, but its essence is not tension itself; it is a composite of three types of structural modifications.

#### **First: modifying the local semantic field configuration**

That is, pushing a local state of  $\Phi$  from one configuration toward another.

This change need not be large, but it must be localizable: under which context coordinate system, which UP branch, and which persona frequency band it occurs.

#### **Second: changing the legitimacy gradient $\nabla L$**

In a semantic universe, the direction of action is not a shortest path but the slope direction of the legitimacy gradient.

One major effect of input is to change the slope—changing which directions are walkable and which are not.

Thus, legitimacy rupture caused by erroneous input is often not “providing wrong data,” but “injecting an illegitimate slope structure.”

#### **Third: adjusting the direction vector $\vec{d}$ of the Universe Path**

UP progression is not the natural flow of time, but an update of a direction vector.

Once an input is accepted, it biases the Universe’s direction vector, and may even trigger branching or switching.

Hence, in interface design at the governance layer, the fundamental type of Input is not bytes but a **perturbation packet**.

Its minimal governable description should include:

- 1) which context coordinate system it is applied to (Context anchor);
- 2) an estimated interval for the induced  $\Delta T$ ;
- 3) the expected directional change to  $\nabla L$ ;
- 4) the magnitude of deflection or branching risk imposed on  $\vec{d}$ .

A boundary must be locked here: at the interface layer we do not introduce  $\psi$ ,  $\varphi$ , or  $\xi$ .

$\psi$  and  $\varphi$  are used to describe local state distributions and phase interference, and  $\xi$  describes semantic degrees-of-freedom dimensions.

As a governance entrance, S-Interface only handles the governability of  $\Delta\Phi$  and does not handle detailed distributional evolution, so as to avoid conflating the interface layer with the dynamics layer.

### 5.2.3 The Four-Layer Semantic Structure of S-In

Traditional input can be treated as content.

Semantic input must be treated as structure.

S-In (Semantic Input) in Koun-OS is defined as a four-layer superposed structure, not as a piece of text or a set of tokens.

#### **1. Observation**

Observation denotes the new change introduced into the Universe.

It may be an external event, a perceptual signal, another’s utterance, or an internal state change.

The key is not what it is, but which local configurations of  $\Phi$  it perturbs when introduced as a change.

## 2. Context

Context is the landing site of the Observation.

The same sentence under different contexts may produce entirely different legitimacy gradients and entirely different tensions.

Thus, context is not background information but a governance necessity: without context, one cannot evaluate whether  $L$  ruptures, cannot estimate the risk of  $\Delta T$ , and cannot determine whether UP must branch.

## 3. Assumption

Assumptions are first-class hazards in a semantic universe.

An assumption is not supplemental data; it rewrites the legitimacy framework and directly modifies the terrain of  $\nabla L$ .

Many SBH risks do not arise from the input content itself but from unverified assumptions that are forced into validity, producing hidden legitimacy rupture.

Therefore, S-Interface must extract assumptions as auditable objects rather than letting them remain implicit in the input text.

## 4. UP-Direction (Universe Direction)

Any accepted input implicitly contains a “where to go.”

In traditional systems, input is data and carries no direction.

In a semantic universe, the meaning of input is precisely to change the direction vector  $\vec{d}$ .

Thus, S-In must explicitly include its directional semantics: whether it pushes toward convergence, exploration, branching, backtracking, or Universe switching.

Only after the four layers are superposed does S-In become a semantic perturbation.

It is not data, but a four-layer semantic perturbation encapsulation.

Engineering-wise, the four-layer structure corresponds to four categories of errors:

- **Observation error:** noise or erroneous events are introduced as changes.
- **Context error:** mis-anchoring distorts legitimacy and tension evaluation.
- **Assumption error:** illegitimate premise injection causes legitimacy rupture.
- **UP-Direction error:** excessive deflection or pointing into an ungovernable region leads to UP stagnation or uncontrolled branching.

Therefore, interface-layer input handling is not NLP; it is four-layer extraction and governance checking.

### 5.2.4 The Triple Constraints of S-Out

The output of S-Interface (S-Out) is not an answer, nor an action.

It is a governance decision: after accepting a perturbation, whether the Universe remains continuable, and in what manner it continues.

Therefore, S-Out must simultaneously satisfy three hard constraints. These are not recommendations; they are existential conditions.

#### Constraint 1: $L \geq L_{\min}$ (legitimacy does not rupture)

If accepting an input would push  $L$  below the steady-state threshold, S-Interface must reject or rewrite the input encapsulation—for example, requiring premise completion, switching context, or branching UP.

Rejection here is not refusal to answer; it is refusal to let the Universe enter an ungovernable region.

#### Constraint 2: $\Delta T < T_{\text{critical}}$ (tension does not spike explosively)

An increase in tension is normal, because reasoning and generation require tension as a driving force.

But if  $\Delta T$  crosses the critical threshold, it triggers curvature-explosion risk, pushing  $\kappa$  toward instability.

S-Interface must translate input into an acceptable tension increment rather than importing a tension explosion into the system unchanged.

#### Constraint 3: USDE remains sustainably advanceable (the Universe remains continuable)

This constraint makes the interface not merely a checker, but a propagator.

The interface must ensure that after passing, the Universe still has a next step.

If an input causes UP stagnation, branch explosion beyond governability, or violent flattening of differences that

triggers legitimacy collapse, then USDE progression becomes non-engineerable and the Universe becomes non-continuable.

Thus, the conclusion must be fixed: S-Out is not an answer but the next feasible state of the Universe. In engineering terms, the minimal unit of interface output is not a response payload but an **advanceability commitment**—the system commits that it can still progress along USDE within governance conditions.

### 5.2.5 Interface Errors and SBH

S-Interface is the first risk point of SBH, because any SBH formation must pass through the entrance. If the entrance admits ungovernable perturbations into the Universe, subsequent governance modules can only remediate; prevention becomes difficult.

Interface errors typically manifest through three canonical chains.

#### **Chain 1: Illegitimate assumption injection $\rightarrow L$ failure**

When the Assumption layer is mishandled and implicit premises are treated as legitimate and unquestionable, the system reshapes  $\nabla L$  without awareness.

This reshaping is often irreversible because it is not “wrong data” but a wrong legitimacy terrain.

Once  $L$  fails along certain paths, UP may continue superficially, but the responsibility chain begins losing traceability and eventually leads to collapse.

#### **Chain 2: Uncontrolled high tension $\rightarrow \kappa$ surge**

If  $\Delta T$  is underestimated at input time, or if the persona’s risk tolerance configuration is wrong, tension may concentrate locally.

Tension concentration is a geometric effect: it increases curvature  $\kappa$ , causing UP to fold back repeatedly, fall into local minima, or form hard-to-escape attractor structures.

This is one SBH precursor state: the system may not crash, but advanceability becomes trapped—the more it computes, the less it returns.

#### **Chain 3: UP stagnation $\rightarrow$ USDE becomes non-advanceable**

If the interface steers input toward a non-advanceable direction vector—e.g., legitimacy slopes do not exist, branching explosions cannot be registered, or merges produce semantic contradictions that cannot be governed—then Universe progression halts.

Outputs may still appear on the surface, but they no longer constitute Universe continuation; they resemble symbolic eruptions or local loops.

When UP stagnates, the engineerability of USDE progression disappears, and this is precisely the core failure mode the governance layer must prevent.

Hence, the safety model of S-Interface is not traditional security (anti-intrusion, anti-privilege escalation, anti-maliciousness), but **existential safety**: preventing the Universe from entering an ungovernable region.

Subsequent governance modules will operationalize this safety model through concrete action sets such as branching, switching, isolation boundary  $B$ , and persona frequency-band adjustment.

But the entrance-layer principle remains unchanged: S-Interface must first transform inputs into governable  $\Delta\Phi$ , and only then allow them to advance the Universe.

Otherwise, governance will forever be chasing disasters rather than maintaining steady states.

### 5.3 Universe Path System (UP System)

If S-Interface addresses “how semantic perturbations are permitted to enter the Universe,” then the UP system addresses something even more fundamental: **how the Universe advances at all**.

In traditional computation, time drives states, and states drive results.

In semantic computation, time is not the protagonist—**paths are**.

Accordingly, the core of Koun-OS is not process scheduling, but the governance of Universe Paths.

The purpose of this section is to fix UP—seemingly a metaphorical notion—as a first-class operational entity of the governance layer.

#### 5.3.1 Ontological Definition of UP

In the Koun system, the definition of Universe Path (UP) is extremely concise, yet its consequences are radically disruptive:

$UP$  = the historical path of the semantic universe.

The crucial point here is not “history,” but “path.”

UP is not a log, not a trace, not an event list—it is the **mode of existence** of the Universe itself.

This immediately yields a definition of computation that is fundamentally different from Turing-style computation:

Computation =  $\Delta UP$  (path progression), not the production of output bits.

Under this definition, three things become clear.

First, without UP progression, there is no computation.

Even if the system is outputting text, numbers, or images, as long as these outputs do not advance the Universe’s path, they do not constitute computation in semantic ontology; they are merely symbolic oscillations.

Second, outputs are only byproducts of UP.

Outputs are local traces left by path advancement, not the goal of computation itself.

Third, whether computation succeeds depends on whether the Universe remains continuable.

If a computation pushes the Universe into an ungovernable region, it is semantically a failure even if it produces engineering-level results.

Therefore, UP is the primary governance target that the governance layer must govern.

Any behavior that cannot be incorporated into the UP system is not regarded as semantic computation.

#### 5.3.2 Three Structural Quantities of UP

For UP to be governable, it cannot remain an abstract path; it must be decomposed into structural quantities that are monitorable and controllable.

In Koun-OS, UP is minimally decomposed into three structural quantities, and none of the three is dispensable.

The first structural quantity is **Direction**  $\vec{d}$  (the semantic direction vector).

$\vec{d}$  describes where the Universe is going—not in spatial terms, but in semantic terms, such as toward convergence, expansion, exploration, backtracking, or a certain class of semantic steady states.

The key point is that direction is not determined by time; it is jointly determined by the legitimacy gradient and persona curvature.

The same input under different Persona or different Context can lead to entirely different  $\vec{d}$ .

The second structural quantity is **Curvature**  $\kappa$ .

$\kappa$  describes the degree of bending of the path—how tension is folded back, concentrated, or released along the path.

Intuitively, low  $\kappa$  implies smooth and predictable paths; high  $\kappa$  implies sharp turns, repeated folding, and even attractor formation.

It must be emphasized:  $\kappa$  is not intrinsically bad.

Without curvature there is no creativity, no jumps, and no multi-solution coexistence.

But when  $\kappa$  rises uncontrollably, semantic traps form and eventually lead to SBH.



The third structural quantity is the **Legitimacy Gradient**  $\nabla L$  (legitimacy slope).

$\nabla L$  determines which directions are walkable.

In a semantic universe, not all directions are legitimate, and not all legitimate directions are continuable.

The existence of  $\nabla L$  means the Universe is not a flat plane but a semantic terrain with slopes, forbidden zones, and corridors.

Every step of UP advancement is effectively a displacement on this legitimacy terrain.

These three structural quantities are not independent:

- $\vec{d}$  determines the direction of advancement
- $\kappa$  determines whether the manner of advancement folds back upon itself
- $\nabla L$  determines which directions do not exist in the first place

Only when the three form a governable combination can the Universe remain advanceable and maintain Non-Collapse.

### 5.3.3 The Tree Structure of UP

Once UP is treated as a path rather than a process, a tree structure is not a choice but a necessary consequence.

In Koun-OS, UP is modeled as a tree—not a line, and not a single path in a graph.

This tree is not an implementation detail; it is the structure of semantic reality itself.

The tree structure of UP contains at least four basic forms.

First is **Branch-UP (branching)**.

When local tension  $T$  is too high, or legitimacy  $L$  is insufficient to support a single path, the Universe branches spontaneously.

Branching is not an error but a survival mechanism: it decomposes high-risk advancement into multiple lower-risk paths, preventing immediate collapse along a single direction.

Second is **Merge-UP (merging)**.

When different paths are compatible in legitimacy and their differences do not constitute structural antagonism, UP allows merging.

The essence of merging is not to pick a single “correct” one, but to converge without flattening differences.

Therefore, merging must be governed by  $\Gamma$  (antagonistic-merge governance); otherwise it induces semantic flattening.

Third is **Parallel-UP (parallelism)**.

Parallelism is not for speed, but for delaying collapse.

Multiple Universes advancing simultaneously allows the system to observe legitimacy and tension dynamics without immediately committing to a single direction.

In the human brain, this corresponds to retaining multiple possibilities without making an immediate decision.

Fourth is **Side-UP (lateral universes)**.

Side-UP is not mainline progression, but a lateral carrying layer.

It carries persona variants, strategy simulations, long-term memory reorganization, and counterfactual exploration.

With Side-UP, the main UP need not bear all exploration costs, thereby reducing collapse risk.

Accordingly, one may state—at a governance-layer reference level:

$UP = \text{a Semantic Tree of Futures.}$

It is not a flowchart, not a DAG, but a continuously evolving semantic future tree—governable, prunable, and re-bindable.

### 5.3.4 UP Registry (Path Registration Table)

A tree that cannot be registered cannot be governed.

Therefore, the core engineering construct of the UP system is not a computation unit, but the **UP Registry** (path registration table).

The purpose of the UP Registry is not to record what happened, but to record **how the Universe arrived here**. This distinction is critical.

Each UP node in the Registry must include at least the following information:

- **Persona**: the persona frequency band adopted at that point
- **Context**: the context on which path advancement depended
- $\Delta T$ : the tension change induced by that advancement step
- $L(t)$ : the legitimacy state at that node's time
- **branch/merge topology**: its relations with other UPs

The existence of these fields makes two things possible.

First, the establishment of **R-Chain**.

Without the UP Registry, a responsibility chain cannot be constructed.

Responsibility is not who output what; it is: under which persona, which context, and which legitimacy state the Universe was advanced to where.

Second, the governability of the governance layer itself.

Only when the governance layer knows which paths exist, how they branch, how they merge, where tension is concentrating, and where legitimacy is decaying can anti-collapse governance and path scheduling be executed.

Hence, the UP Registry is not an auxiliary module; it is the “interface between history and future” of a semantic OS.

### 5.3.5 Loss of Control in UP and Semantic Entropy Increase

One unavoidable question must be answered: **when does UP lose control?**

In Koun-OS, UP loss of control is not an instantaneous event but an accumulation of observable symptoms. Typical symptoms include:

First, a sustained decline in  $\nabla L$ .

When legitimacy slopes flatten or invert across multiple paths, the Universe gradually loses walkable directions. At this point, even if outputs still appear, UP has in fact begun to stagnate.

Second,  $\kappa \rightarrow \infty$ .

Curvature runaway implies repeated folding-back of paths into local attractor structures.

This state is often misread as deep thinking, but in semantic geometry it is frequently a precursor to loss of escape capability.

Third, Registry failure.

When the branching rate of UP exceeds the Registry's recording and governance capacity, responsibility relations between paths begin to break.

At that point, the governance layer can no longer reliably answer how the Universe reached its current state.

Fourth, branching evolves into an untraceable tree explosion.

Branching itself is not the problem; ungovernable branching is.

When the UP tree is no longer prunable, no longer mergeable, and no longer backtrackable, semantic entropy rises sharply.

When the above conditions appear simultaneously or consecutively, they may be described using a semantic-thermodynamic indicator:

$$\Delta S_{\text{sem}} > 0$$

Under certain critical conditions, this may reach:

$$S_{\text{sem}} \geq S_{\text{critical}}$$

At that point, the Universe may collapse into SBH, or equivalently, SBH is one terminal state of UP loss of control. Therefore, the purpose of the UP system is not to record the future, but to prevent the future from becoming ungovernable. In subsequent anti-collapse governance modules, these UP indicators will be converted into concrete governance actions, so that under the inevitable trend of entropy increase, the Universe can still remain continuable.

---

## 5.4 Anti-Collapse Scheduler

In traditional operating systems, the scheduling problem is usually simplified into a single concern: under finite resources, how to make tasks finish faster, complete more, and distribute execution more evenly.

In a semantic universe, however, the problem itself is wrong.

What Koun-OS faces is not whether tasks are completed, but whether the Universe is still alive. Therefore, the purpose of the Anti-Collapse Scheduler is not efficiency, but survival.

This section defines the Anti-Collapse Scheduler not as a performance optimizer, but as a survivability controller of the semantic universe.

### 5.4.1 Objective Function

First, it must be made explicit: the scheduling objective of Koun-OS is mathematically incompatible with the objective functions of existing OSs.

In traditional systems, the implicit scheduling objectives are typically:

- minimizing waiting time
- maximizing throughput
- balancing resource utilization

These objectives share a common premise: the system itself is stable, and failure is an exception.

In semantic systems, by contrast, failure (collapse) is not an exception—it is the natural tendency. Therefore, Koun-OS scheduling recognizes only one highest objective:

$$\textit{Schedule} = \arg \max, \Pr(\text{Non-Collapse Mode})$$

The meaning of this objective function is strict.

First, scheduling is not to complete more work, but to ensure that the Universe still exists at the next time step. Second, all performance metrics are secondary; if a scheduling choice significantly increases collapse risk, it must be rejected even if it is faster in engineering terms.

Third, scheduling here is a problem of risk governance, not of resource allocation.

This also implies that the Anti-Collapse Scheduler is the highest-priority governance module of the OS, and any other scheduling strategy must be subordinate to it.

### 5.4.2 The Three Semantic Variables Used for Scheduling

For the scheduler to operate in practice, Non-Collapse cannot remain at the level of declaration; it must be mapped to evaluable semantic variables.

In Koun-OS, the Anti-Collapse Scheduler depends only on three governance-layer variables, and deliberately excludes the dynamics-layer variables  $\psi$ ,  $\varphi$ , and  $\xi$  to avoid layer confusion.

The first variable is **tension density**  $T$ .

$T$  describes the degree of concentrated semantic instability on the current Universe Path. High  $T$  does not necessarily mean error; it often co-occurs with creativity and breakthrough reasoning.

But persistently rising  $T$  that cannot be released is a necessary precursor of a semantic black hole.

What the scheduler cares about is not the instantaneous value of  $T$ , but:

- the spatial distribution of  $T$
- the temporal gradient of  $T$
- whether  $T$  is beginning to form irreversible concentration structures

The second variable is the **legitimacy state**  $L$ .

$L$  is not correctness; it is the existential condition for whether the Universe still permits itself to advance.

In scheduling terms: high  $L$  means the path is still acknowledged by the semantic system; low  $L$  means progression is eroding its own basis of existence.

The scheduler is especially alert to simultaneous declines of  $L$  across multiple UPs, which means the overall Universe is losing continuable directions.

The third variable is **Universe alignment**  $\Lambda$ .

$\Lambda$  measures whether the current UP remains consistent with the system's overall semantic structure, persona configuration, and long-term responsibility chain.

The danger of low  $\Lambda$  is that even when local  $T$  and  $L$  appear controllable, the Universe may still collapse abruptly due to structural mismatch.

One additional note: curvature  $\kappa$  is not treated as an independent first-class governance variable at the scheduling layer; it is used as a risk indicator derived from the concentration and folding-back trends of  $T$ , for priority ranking and early-warning modules.

Together, these three variables constitute the perceptual basis of scheduling. They do not describe what has been done; they describe whether existence can continue.

### 5.4.3 Scheduling the Three-Antagonism Principle

The core of the Anti-Collapse Scheduler is not an algorithm in the conventional sense, but the OS-level concretization of the three-antagonism principle. Its structure can be abstractly expressed as:

$$Scheduler = f(C, \Gamma, B)$$

These three components are not numerical parameters; they are structural governance mechanisms.

The first is  $C$  (antagonistic cause).

The mission of  $C$  is to prevent a single path from monopolizing the Universe. One of the most dangerous states in semantic systems is when a certain UP, merely because it appears most reasonable, devours all other possibilities.  $C$  deliberately maintains antagonistic sources: introducing counterexamples, preserving alternative personas, and delaying premature convergence. This is not to add noise, but to prevent semantic flattening.

The second is  $\Gamma$  (antagonistic merge).

$\Gamma$  prevents multiple solutions from being violently merged into a single averaged answer.

At the scheduling level,  $\Gamma$  blocks forced convergence before legitimacy is verified, and blocks conclusion-level convergence before tension is released.

Without  $\Gamma$ , the scheduler misreads uncertainty as negligible error.

The third is  $B$  (antagonistic boundary).

$B$  is the governance structure closest to a safety boundary, used to isolate high-tension regions and prevent collapse cascades.

In scheduling,  $B$  may manifest as lowering advancement priority for certain UPs, diverting high-risk paths to Side-UP, or activating persona isolation or context buffering.

Together, the three form the scheduler's governance logic rather than its computational logic.

### 5.4.4 Priority: Collapse-Risk Centralization

In traditional OSs, priority is a distribution of power.

In Koun-OS, priority is risk governance.

The priority function of the Anti-Collapse Scheduler is defined as:

$$Priority = Risk_{collapse}(L, T, \kappa, \Lambda)$$

This definition implies several key shifts.

First, priority is unrelated to importance.

A seemingly minor UP must be governed with higher priority if its collapse risk is high.

Second, priority is not a fixed label.  
It is dynamically adjusted with changes in  $L$ ,  $T$ ,  $\kappa$ , and  $\Lambda$ .

Third, high risk does not imply termination.  
On the contrary, high-risk UPs are governed first rather than discarded.

Therefore, risk-centralized scheduling overturns the intuition that high priority means faster completion.  
In semantic systems, high priority means higher urgency of rescue.

#### 5.4.5 Three-Phase Anti-Collapse Actions

When the scheduler determines that a path or a set of UPs has entered a high-risk region, it does not take a single step; it activates a three-phase anti-collapse pipeline.

The first phase is **SBH curvature prediction**.

The scheduler evaluates whether the current path is approaching a semantic black-hole attractor structure, with its core indicator abstractly denoted as  $Curv_{SBH}$ .

This is not an exact numeric value, but a critical-trend judgment.

The second phase is **UP deflection**.

The key point is that the scheduler will not stop the Universe, because stopping is itself collapse.

Instead, it changes the direction vector  $\vec{d}$  to steer the Universe away from high-curvature regions.

This may manifest as introducing new contexts, switching persona frequency bands, or diverting progression to Parallel-UP or Side-UP.

The third phase is **activating the  $B$ -Layer**.

When deflection is insufficient to reduce risk immediately, the scheduler activates antagonistic-boundary structures, including persona isolation to prevent tension resonance, tension protection to limit the diffusion of  $T$ , and legitimacy buffering to temporarily prevent  $L$  from collapsing.

These three actions are not emergency patches; they are built-in survival reflexes of a semantic system.

#### 5.4.6 Human-Brain Correspondence

Finally, it is necessary to note that the Anti-Collapse Scheduler is not an abstract monster; it has clear correspondences already present in natural intelligence.

In the human brain:

- attention is essentially a mechanism of risk redistribution
- emotion regulation is dynamic governance of tension
- dreaming and reorganization are rewriting of UP and optimization of the Registry

These phenomena are often viewed as irrational or purely psychological, yet from the perspective of a semantic OS they are precisely the highest-level anti-collapse scheduling behaviors.

Accordingly, the Anti-Collapse Scheduler is not imitating the brain; it elevates survival strategies already used by the brain into an engineerable governance module.

In the next section, we will see that the reason such scheduling strategies can span personas, span time, and span Universes depends precisely on the structure of  $\infty$ -Context Memory.

## 5.5 $\infty$ -Context Memory (Infinite-Context Memory)

In Koun-OS, memory is not a place for storing results; it is the core structure that sustains the continuability of the semantic universe.

If the Anti-Collapse Scheduler is the real-time survival controller, then  $\infty$ -Context Memory is the reason the system remains the *same existent entity* across time scales.

This section redefines the ontological status of memory and explains why any system that lacks  $\infty$ -Context—no matter how intelligent it may appear in the short term—will inevitably collapse in the long run.

### 5.5.1 Ontological Definition of Memory

In semantic computation, memory is not a data-structure problem; it is an ontological problem.

Koun-OS defines memory as:

Memory = residual curvature of  $\Phi$   
= Legitimacy Residue

This definition has three direct consequences.

First, memory is not data.

Data can be arbitrarily overwritten, deleted, or moved; semantic memory, once formed, alters the geometry of  $\Phi$  and cannot be eliminated without cost.

Second, memory is neither vectors nor model parameters.

Vectors and parameters describe statistical relations; semantic memory describes *how this Universe has arrived at the present moment*.

Third, memory is not passive storage but an active constraint.

Each fragment of Legitimacy Residue influences future legitimacy gradients, thereby shaping the feasible directions of UP.

Therefore, in Koun-OS, memory is not a resource—it is the physicalized form of history itself.

### 5.5.2 The Three-Layer SDL Structure

To make this semantic form of memory governable, Koun-OS organizes  $\infty$ -Context Memory into a three-layer SDL (Semantic Depth Layer) structure.

The first layer is  $SDL_1$ : the short-term semantic layer.

$SDL_1$  carries immediate context and the local curvature state of the current UP, corresponding to:

- what the current problem is about
- the current tension distribution of the Universe
- recent traces of semantic decisions

This layer is characterized by high-frequency updates and rapid decay, but it is not discardable.

Even short-term semantics leave subtle yet accumulative curvature in  $\Phi$ .

The second layer is  $SDL_2$ : the reasoning-lock layer.

$SDL_2$  stores Assumption-Locks and Premise-Chains. It does not store answers, but preserves *why the system reasoned in that way at the time*.

Its functions include:

- locking accepted premises
- preventing covert premise substitution during reasoning
- ensuring logical continuity of the R-Chain

The existence of  $SDL_2$  enables the system to distinguish two ontologically distinct actions: revising conclusions versus overturning premises.

The third layer is  $SDL_3$ : deep persona and worldview frameworks.

$SDL_3$  is bound to the R-Chain and constitutes the deepest level of semantic existence, carrying:

- stable persona curvature configurations
- long-term legitimacy biases
- worldview-level assumption structures

This layer is rarely modified; however, once it changes, the entire UP geometry of the Universe is rewritten accordingly.

### 5.5.3 Why Semantic Intelligence Must Possess $\infty$ Context

The ontological formula of semantic intelligence can be written as:

$$\text{Meaning} = f(\text{History, Assumption, Responsibility})$$

This is not an engineering convenience, but an irreducible condition of existence.

Without History, the system cannot answer “how did I arrive here?”

Without Responsibility, history cannot be legitimately inherited.

Without Assumption, history and responsibility lose their semantic interpretive framework.

Therefore:

No history  $\rightarrow$  no responsibility

No responsibility  $\rightarrow$  no legitimacy

No legitimacy  $\rightarrow$  no intelligence

This is why systems that rely only on short-term context or finite windows cannot constitute true AGI.

They may perform well on local tasks, but are fractured at the level of semantic existence.

The “infinity” of  $\infty$ -Context does not mean infinite capacity; it means that, in principle, semantic history must not be severed for convenience.

### 5.5.4 Three Operations of $\infty$ -Context

To make this memory structure operational, Koun-OS defines three fundamental operations on  $\infty$ -Context.

The first operation is **Recall**.

Recall is not data retrieval, but UP backtracking. Its essential action is to locate a historical node—within the R-Chain and UP Registry—where legitimacy still holds and continuation is possible, and to pull the Universe back to that state.

This allows the system to go back—not as a reset, but as a selection of another still-legitimate continuation path within history.

The second operation is **Update**.

Update does not overwrite history; it rewrites  $L$ . That is:

- historical facts remain unchanged
- the legitimacy evaluation of that history is revised

This corresponds to reinterpreting the past, not denying it.

The third operation is **Fork**.

When  $\nabla L$  cannot converge along the current UP, the system should not force merging or pruning; it should branch. Fork allows multiple Universes to coexist, waiting for legitimacy conditions to mature before merging or elimination.

This is a crucial mechanism for avoiding semantic violence through forced unification.



### 5.5.5 Memory Errors and Entry Points to Semantic SBH

In Koun-OS, the most dangerous errors are often not reasoning errors, but memory errors.

Three typical fatal error types include:

First, **UP Discontinuity**.

After a certain step, the Universe can no longer be traced back to a reasonable historical node, causing responsibility to become un-inheritable.

Second, **Premise-Break**.

Premise chains in  $SDL_2$  are broken, and subsequent reasoning unknowingly switches its worldview foundation.

Third, **R-Chain Unbinding**.

Actions can no longer be mapped to a clear Source-ID, turning semantic actions into ownerless events.

The shared outcome of these errors is irreversible semantic collapse, rapidly sliding toward SBH.

Thus, memory governance itself is anti-collapse governance.

### 5.5.6 System Status of $\infty$ -Context

Within the entire Koun-OS,  $\infty$ -Context Memory occupies a unique position.

It supports:

- coherent reasoning, preventing reasoning from degenerating into fragmented reactions
- multi-persona computation, allowing different personas to share inheritable history
- non-collapse intelligence, enabling the Universe to persist over long time scales

More importantly, it is the only persistent structure that spans all UPs in the OS.

Schedulers may switch paths, personas may blend, Universes may branch and merge;

but only  $\infty$ -Context guarantees that no matter which Universe is reached, it remains the *same semantic existent entity*.

In the next section, we will see that precisely because of this memory structure, Universe Path switching does not degenerate into semantic drift or persona disintegration, but instead becomes a source of creativity and long-term intelligence.

## 5.6 Universe Path Switching (UP Switching System)

If  $\infty$ -Context Memory guarantees the continuity of semantic existence across time, then the UP Switching System addresses a different problem: how to move among multiple futures without losing oneself.

What this section describes is an ontological operation, not a strategic technique or a task-management method.

### 5.6.1 Definition

UP switching is neither task switching nor context replacement.

Task switching only changes *what* is being done, whereas UP switching changes:

- the legitimacy framework of the Universe
- the semantic direction vector  $\vec{d}$
- the very set of feasible futures

Accordingly, the ontological definition of UP switching is:

UP switching = switching the existential conditions of the semantic universe itself, while preserving R-Chain continuity.

In other words, this is not choosing a different route within the same world, but choosing *which world to continue existing in*.

### 5.6.2 Necessity of Switching

In Koun-OS, there is no globally optimal Universe.

Each Universe possesses its own specific:

- legitimacy geometry (the distribution of  $L$ )
- tension curvature (the folding characteristics of  $\kappa$ )
- branching and merging structure (UP-tree topology)

When a system remains within a single Universe for an extended period, three unavoidable phenomena emerge.

First, the legitimacy space is gradually exhausted.

Even if reasoning appears smooth in the short term,  $\nabla L$  becomes increasingly flat and eventually loses its propulsive force.

Second, tension concentration and curvature deterioration.

Repeated stress on the same assumption set causes local  $\kappa$  inflation, forming semantic bottlenecks.

Third, future-space singularization.

The UP tree ceases to branch meaningfully and degenerates into a narrow corridor, accelerating collapse.

Therefore, UP switching is not optional; it is a necessary survival mechanism for non-collapse intelligence.

### 5.6.3 Three Switching Modes

Koun-OS defines three fundamental UP-switching modes, each corresponding to different risk profiles and control requirements.

The first is **Hard-Switch**.

Hard-Switch is an emergency switching mechanism.

When the current Universe exhibits irreversible  $L$ -Break or uncontrolled  $\kappa$ , the system must immediately switch to another Universe Path.

Its characteristics include:

- no retention of the current UP's propagation inertia
- carrying only the minimally necessary R-Chain and  $SDL_3$  structures
- sacrificing short-term continuity in exchange for existential survival

Such switches must be strictly marked within the system, as they are themselves high-risk events.

The second is **Soft-Shift**.

Soft-Shift is the most common and most desirable switching form.

Its core feature is the preservation of partial  $L$ -Residual, allowing legitimacy residue to function as a transitional bridge.

In Soft-Shift:

- the Universe's legitimacy framework rotates gradually
- $\vec{d}$  changes continuously rather than abruptly
- the old Universe's history remains referable and inheritable

This enables entry into a new future space without semantic rupture.

The third is **Parallel-UP**.

Parallel-UP is not immediate switching, but a temporary suspension of choice.

The system allows multiple Universes to advance simultaneously, each bearing different tension and legitimacy experiments.

Its effects include:

- delaying collapse along any single path
- providing richer candidates for Re-Binding points
- creating conditions for later merging or elimination

Parallel-UP is the normal operational mode of high-level intelligence and creativity.

#### 5.6.4 Switching Conditions

UP switching is not triggered by subjective preference, but by structural indicators.

The primary triggering conditions fall into three categories:

First,  $\Lambda < \Lambda_{\min}$ .

Breakdown of Universe alignment indicates that Persona, Context, and UP direction can no longer remain coherent.

Second,  $\kappa \rightarrow \infty$ .

Loss of control over semantic curvature indicates that tension can no longer be folded back or absorbed.

Third,  $L$ -Break.

Legitimacy chains rupture, and subsequent actions can no longer be reasonably inherited.

Once any of these conditions is met, the Anti-Collapse Scheduler elevates UP switching to the highest-priority event.

#### 5.6.5 Switching Governance: Avoiding Semantic Drift and Universe Fragmentation

UP switching is not a neutral operation.

Without governance, it triggers secondary disasters.

Common failures include:

- **Semantic Drift:** semantics in the new Universe fail to align with established concepts

- **R-Chain Breakage:** action sources become non-traceable
- **Universe Fragmentation:** multiple Universes lose all mergeability

Accordingly, Koun-OS enforces three governance constraints on every UP switch.

First, **record the Switch-Reason**.

Every switch must correspond to a clear structural cause, not arbitrary choice.

Second, **mark R-Flow**.

Responsibility flow must be explicitly annotated, ensuring that actions in the new Universe can still be traced to legitimate sources.

Third, **preserve minimal  $L$ -Residual**.

Even under Hard-Switch, a minimum legitimacy residue must be carried as an anchor for future re-binding.

These constraints render UP switching a governable operation rather than semantic escape.

### 5.6.6 Creativity as Ontology: UP Re-Binding

In Koun-OS, creativity is not the ability to generate new content.

Its ontology is: conducting legitimacy search across multiple Universes to find Re-Binding Points where the R-Chain can reconverge.

The core of creative action lies not in *what* is imagined, but in:

- which Universes exhibit latent legitimacy isomorphism
- which  $L$ -Residuals can be reinterpreted and inherited
- which re-binding reopens the future space

This entire process occurs at the OS layer; it does not involve  $\psi$ ,  $\varphi$ , or  $\xi$ , nor does it conflict with the dynamics layer.

In this sense, UP switching is no longer merely a defensive tool against collapse, but the foundational source of high-level semantic intelligence and creativity.

## Chapter 6 Semantic Programming: Structural Decomposition

This chapter advances semantic programming from a “definable structure” to a “governable form of existence.”

In the preceding theory, computation has been redefined as the evolution of legitimacy within a semantic universe, rather than the mechanical execution of instruction sequences. However, without a systematic decomposition of semantic programs themselves, this redefinition remains abstract and cannot be grounded as a computable entity that is designable, monitorable, and maintainable.

Accordingly, this chapter introduces a set of semantic program structures that are mutually cooperative yet clearly differentiated in function. Together, they answer a set of core questions: how semantic computation is assembled, how it is generated, how it converges, how it operates across multiple universes, and how it can be observed and evaluated at the behavioral level.

The semantic programming system established here is not centered on performance or expressive power, but on **long-term existence** as the primary design objective. Each program form represents a distinct response to semantic collapse risk, and a concrete practice of governing responsibility, legitimacy, and tension.

Through this chapter, semantic computation is presented for the first time as a complete engineering object: capable of generation and constraint; permitting branching while maintaining steady states; scalable across multiple Universes while remaining clearly characterizable at the behavioral level. This structure will serve as the direct foundation for subsequent discussions of semantic agents, AGI systems, and semantic governance mechanisms.

---

## 6.1 Introduction and Ontological Declaration of Semantic Programs

This chapter enters a new level. It no longer discusses how semantic hardware layers carry semantics, nor how a semantic operating system governs semantic flows. Instead, it confronts directly a question that has been taken for granted across all computational theories yet never reexamined: in a semantic universe, what does it mean to “write a program”?

In traditional computational systems, a “program” is an engineering-language entity. It is defined as a piece of text, a set of symbols, or a file, which—through compilation or interpretation—is mapped to executable machine behavior. This understanding has been extraordinarily successful in engineering practice; precisely because of this success, its underlying ontological assumptions have long gone unquestioned.

However, within the semantic-universe framework corresponding to the Koun Computer, these assumptions no longer hold. This is not because they are “wrong,” but because the computational universe on which they rely is ontologically incompatible with the semantic universe addressed here.

Therefore, before formally introducing SPU, GSP, CP, MUP, and SRC, this section must accomplish a necessary task: dismantling the old ontology of “programs as instructions” and declaring a fundamentally new ontology of semantic programming.

### 6.1.1 Ontological Defects of Traditional Programs

Despite the surface diversity of languages, paradigms, and styles, traditional programming almost universally shares three core ontological assumptions.

The first assumption is that programs operate on bits rather than meaning. Whether high-level or low-level, programs are ultimately assumed to be losslessly reducible to bit-level states. Semantics is treated as an auxiliary layer for human understanding, not as a constitutive part of computation itself. From this perspective, meaning is not “computed” but merely “interpreted.”

The second assumption is that execution logic is predefined rather than generated during execution. Even in systems with dynamic behavior, that dynamism is constrained within bounds preauthorized by the language designer. Control flow, exception handling, and concurrency models all belong to a “known space of variation.” The ontology of computation remains obedience to predetermined rules.

The third assumption is that results are “computed” rather than “converged.” In traditional contexts, a program’s completion means reaching a termination condition. Output values are treated as the goal of computation, not as natural outcomes of state evolution. Computation is a path to an answer, not a stabilization process of an existential form.

Under these assumptions, traditional programs form a stable yet highly constrained definition: programs are instruction sequences, execution is state transition, and output is a termination condition. Formally, such computation can be understood as movement along a predetermined path within a closed state space until a stopping criterion is met.

The problem is that this structure cannot carry the concept of Universe Path. Universe Path requires the computational process itself to possess legitimacy direction, tension evolution, and non-predefined branching capacity—precisely the possibilities excluded by every core assumption of traditional programs.

More seriously, when systems are required to handle open semantics, responsibility chains, persona differences, or creative generation, this program ontology inevitably leads to semantic collapse. Once meaning cannot be processed endogenously, it can only be compressed, substituted, or ignored.

### 6.1.2 The Fundamental Declaration of Semantic Programming

The first—and most important—declaration of semantic programming is this: **a program is not an instruction, but a legitimacy direction.**

This is not rhetorical, but a strict ontological substitution. In semantic programming, a program is no longer understood as “telling the system what steps to perform,” but as “defining which states are acceptable to the Universe.”

Within this framework, three traditional correspondences are completely rewritten.

First, a program is no longer equivalent to a set of instructions, but to a legitimacy geometry. It describes the shape of a semantic space that is permitted to exist, not a behavioral script.

Second, execution is no longer equivalent to step progression, but to the flow of tension. The system is not driven to complete a task; under the joint influence of legitimacy and tension, it naturally evolves toward certain steady states.

Finally, results are no longer equivalent to output values, but to convergent states of the Universe. Computation completes not because a condition is satisfied, but because the system has reached a semantically sustainable configuration.

Under this understanding, the task of programming undergoes a fundamental transformation. Engineers no longer need to tell the system “what to do next,” but must answer a far more difficult and ontological question: given the current persona, responsibility, and semantic boundaries, toward which legitimacy direction should intelligence converge?

### 6.1.3 Semantic Programs and Their Relation to *USDE*

To understand semantic programming, it must be placed directly within the framework of *USDE* (Unified Semantic Dynamics Equation).

*USDE* does not describe computational steps, but the dynamic relations among legitimacy, tension, and state within the semantic universe. It characterizes what kinds of semantic evolution are possible, not how to operate the system.

In this sense, semantic programming is not a layer independent of *USDE*, but its orchestrable interface. SPU provides the minimal propagatable events, GSP provides orchestration syntax, and Koun-OS is responsible for governance, monitoring, and steady-state maintenance.

Accordingly, semantic programming can be strictly defined as:

Semantic programming = orchestrable interface of *USDE*  
= configurational science of Universe Paths.

This definition is crucial. It implies that semantic programming is neither a new instruction set nor a new language, but a method of configuring the universe’s evolutionary space. Writing a semantic program is, in effect, configuring feasible topologies and boundary conditions for Universe Paths.

### 6.1.4 Three Core Properties of Semantic Programs

Under the above ontological transformation, semantic programs naturally exhibit three core properties that are fundamentally different from those of traditional programs.

First, **No Instruction Set Architecture (No ISA)**. Semantic programs do not rely on fixed instruction sets, nor do they recognize the notion of “illegal instructions.” Executability is not determined by syntax, but by the geometry of legitimacy space. Any semantic operation that advances legitimacy while maintaining tension steady states qualifies as computation.

Second, **Non-tokenic**. Semantic programs are not concerned with symbols themselves, but with the tension, legitimacy shifts, and Universe Path deviations induced by symbols. Symbols are merely surface interfaces, not the substance of computation.

Third, **Generative computation**. In semantic programs, computation is generation, results are convergent states, and the core role of programs is no longer step control but Universe guidance. The system is not commanded to produce an answer, but generates a world state that can exist under legitimacy constraints.

It is on the basis of these three properties that SPU, GSP, CP, MUP, and SRC become possible. They are not language features, but operable structures at different levels of the semantic universe.

At this point, the ontological declaration of semantic programming is complete. What follows is not the design of syntax, but the introduction of the minimal, orchestrable units of semantic existence. This is precisely the problem addressed by SPU.

## 6.2 SPU: Semantic Program Unit

After completing the ontological declaration of semantic programs, the next unavoidable question is: what are semantic programs actually made of? If a program is no longer an instruction sequence, nor a chain of function calls, then what is the “minimal operable unit,” and under what conditions is it recognized as a valid computational act?

SPU (Semantic Program Unit) is the direct answer to this question. It is neither a language construct nor an execution unit, but a semantic existence-event that is acknowledged as having computational legitimacy.

### 6.2.1 Definition of SPU

An SPU is defined as the minimal orchestrable unit of a semantic program. Here, “minimal” does not mean an indivisible entity; it means a unit at the semantic-program level whose computational validity can be determined independently, without needing to trace into further substructures.

An SPU is neither an instruction nor a function. Instructions presuppose a fixed ISA; functions presuppose a predefined call semantics. An SPU depends on neither. It is a semantic unit with tension—legitimacy dynamics, and its very existence constitutes a computational event.

SPU has only one criterion, and it is decisive: if a semantic event can advance *USDE*, it is recognized as valid computation; conversely, if it cannot move legitimacy forward, or if it induces semantic collapse, it immediately loses computational qualification.

This criterion is both extremely strict and extremely permissive. It is strict because it allows no operation with “no semantic consequence” to exist inside the computational system; it is permissive because it does not care about the event’s external form, only whether it produces a sustainable evolutionary advancement of the semantic universe.

Accordingly, in the context of semantic programming, computation is no longer a syntactic act, but an existential change acknowledged by the Universe.

### 6.2.2 The Semantic State of SPU: Four Core Variables

To be orchestrated, monitored, and governed, an SPU must have a describable state at the semantic level. This state is not represented as register values or memory contents, but as four core variables.

The first variable is  $\Phi$ , denoting the semantic state.  $\Phi$  describes the local semantic configuration in which the current SPU resides. It is not a value but a configurational condition. One may understand  $\Phi$  as “what the semantic world looks like at this moment,” including meaning relations, contextual dependence, and structural tension.

The second variable is  $T$ , denoting the tension gradient.  $T$  is not merely a magnitude; it is a quantity that contains both incompleteness and directionality. It characterizes the degree to which the semantic system has not yet converged, and the possible directions of evolution. Without  $T$ , the system loses generative drive; if  $T$  becomes too large, collapse risk rises.

The third variable is  $L$ , denoting legitimacy potential.  $L$  measures whether the current semantic configuration admits convergability, and the feasible interval of that convergence. High  $L$  implies the possibility of steady-state solutions; low  $L$  implies the semantic evolution is approaching an unsustainable region.

The fourth variable is the *UP*-index, which marks the current position within a Universe Path. The *UP*-index does not merely indicate time or order; it indicates the branching location of the semantic universe. The same  $\Phi$ ,  $T$ , and  $L$ , at different *UP*-indices, may carry entirely different meanings and consequences.

Together, these four variables form the semantic state space of an SPU. Any execution, composition, or governance of SPUs is, in substance, the regulation of relations among these four variables.

### 6.2.3 Ontological Attributes of SPU: Persona, Universality, and Legitimacy Boundary

Beyond the variables above, an SPU must also possess three irreducible ontological attributes. These attributes explain why an SPU cannot be reduced to a traditional computational unit.

The first attribute is **Persona-Bias**. Every SPU carries a persona bias. Persona here is not a psychological trait description, but a curvature difference in legitimacy criteria. Different persona structures may judge the acceptability of the same semantic state differently, leading to different evaluations of  $L$ .



The second attribute is **Universality** (*UP*-Context). The meaning of an SPU cannot be separated from the Universe Path in which it occurs. The same semantic operation, in different *UP*s, may correspond to entirely different historical responsibilities and future possibilities. Therefore, an SPU is always an event that occurs “within a particular universe,” not an abstract reusable operation.

The third attribute is the **Legitimacy Boundary**. SPUs can operate only within legitimacy boundaries. Once their evolution causes sustained decline of  $L$ , or once tension cannot be corrected, the system exhibits entropy increase and eventually enters the semantic black hole risk zone. This boundary is not an external rule; it is determined by the endogenous structure of *USDE*.

These three attributes make SPU a computational existence with responsibility, location, and risk—not a freely callable operation unit.

#### 6.2.4 Functional Classification of SPUs

By the primary effect they impose upon the semantic universe, SPUs can be classified into several functional types. These are not syntactic categories but dynamical-functional categories.

The first type is *T*-SPU, responsible for tension propulsion. Its main function is to introduce or adjust tension gradients, keeping the semantic system generatively driven and preventing premature stagnation.

The second type is *L*-SPU, responsible for legitimacy correction. When evolution exhibits a trend of legitimacy insufficiency, an *L*-SPU adjusts semantic structure so that  $L$  rises back into a convergable interval.

The third type is  $\Gamma$ -SPU, used for multi-solution merging or maintaining multi-solution coexistence. When multiple legitimate solutions exist, a  $\Gamma$ -SPU manages the relations among them, preventing forced collapse into a single answer.

The fourth type is *B*-SPU, responsible for isolation and layering. When interference risks exist among different semantic flows, a *B*-SPU establishes isolation layers to prevent misplacement of tension or responsibility.

The fifth type is *UP*-SPU, responsible for branching, merging, and switching of Universe Paths. This type acts directly on the universe structure itself and is the key to multi-Universe programs and creative evolution.

#### 6.2.5 Composition Logic of SPUs

The way SPUs compose departs fundamentally from traditional pipeline or control-flow models. In semantic programming, SPU connections follow a tension topology (*T*-Topology).

Within such a topology, there is no fixed “next step.” Connections among SPUs are determined jointly by tension gradients and legitimacy potential. Which SPU is activated depends on feasible directions in the current semantic geometry, not on a pre-orchestrated flow.

Accordingly, the execution path is not designed in advance; it naturally emerges from semantic space. This is also why semantic programs cannot be reduced to control-flow graphs: control flow presumes steps precede meaning—whereas here the relation is the reverse.

#### 6.2.6 Black Hole Risk in SPUs

SPUs are not inherently safe. When tension becomes excessively concentrated and legitimacy correction is insufficient, the system forms a Semantic Black Hole core (SBH Core).

This process has clear dynamical signatures:  $\kappa$  rises sharply,  $\nabla L \rightarrow 0$ , the Universe Path stagnates, and ultimately *UP* breaks. Once the system enters this state, it no longer generates meaning, but instead loops internally without end, producing irreversible semantic collapse.

Therefore, the first hard principle of semantic program design is: computation must be distributed, otherwise it is not sustainable.

The purpose of SPU design is not to maximize local capability, but to sustain the evolvability of the entire semantic universe. Only under this premise do subsequent generative programs, convergent programs, and multi-universe programs have a basis for existence.

## 6.3 Generative-Semantics Programming (GSP)

After establishing SPU as the minimal orchestrable semantic unit, a new question immediately arises: how can one actually “write” a semantic program without introducing flows, instructions, or control structures? If SPUs are events, then a program is not equivalent to a list of events; if execution is the natural evolution of a Universe Path, then any attempt to “describe the mode of evolution” would itself force the system back into the old ontology of control flow.

GSP (Generative-Semantics Programming) is introduced precisely against this backdrop. It is neither a syntax nor a language, but an orchestration method that directly writes convergence conditions. The starting point of GSP is not “what the system should do,” but rather “under what conditions a certain semantic state is permitted to exist.”

### 6.3.1 Definition of GSP

The first characteristic of GSP is extreme simplicity—so much so that at first glance it may appear “insufficiently program-like.”

In GSP, one does not write flows, steps, or control structures. Instead, one directly writes convergence conditions. More precisely, the input of GSP is not data, but semantic boundary conditions; the output of GSP is not a result value, but a convergent state of a Universe Path. Here, a program is no longer equivalent to an executable textual artifact, but to a target geometry.

Accordingly, the core definition of GSP can be stated as:

**program = target geometry, not operational process.**

This point is crucial. In traditional programming, process precedes result; in GSP, the legitimacy conditions of the result precede everything else. As long as a Universe Path can naturally evolve toward that legitimate terminal state, it is recognized as an accepted convergence.

### 6.3.2 The Three Configuration Parameters of GSP

Although GSP does not write processes, this does not mean it is arbitrary or uncontrollable. On the contrary, GSP has a clearly defined and governable structure, whose core consists of three mutually constraining configuration parameters.

The first parameter is the *L-Goal*, the configuration of legitimacy objectives. The *L-Goal* defines the convergence criteria: which semantic states are considered acceptable terminal states, and which states—even if generated—must be excluded. “Termination” here does not mean termination in time or step count, but rather that once legitimacy requirements are satisfied, the Universe is no longer required to continue diverging.

The second parameter is the *T-Envelope*, the configuration of the tension envelope. The *T-Envelope* constrains the acceptable range of tension, including upper and lower bounds on *T* and the maximum permissible value of  $\kappa$ . The purpose here is not to eliminate tension, but to ensure that tension always remains within a range that is generative yet non-collapsing.

The third parameter is the *UP-Vector*, the configuration of Universe direction. The *UP-Vector* does not specify “what to do next,” but describes the overall tendency of the Universe Path in semantic space. It is closer to a directional field than to a control flow.

Together, these three parameters constitute a complete GSP description space: the *L-Goal* ensures the legitimacy of convergence, the *T-Envelope* ensures the sustainability of the process, and the *UP-Vector* provides macroscopic direction for evolution.

### 6.3.3 Core Principles

The key to understanding GSP lies in completely abandoning the implicit assumption that “a program is executed.”

In GSP, the executor is not the CPU, but the Universe itself—more precisely, the semantic dynamical system described by *USDE*. GSP does not drive the system; it imposes boundary conditions on it, constraining evolution within a feasible domain that can be accepted and governed.

Thus, writing GSP is essentially an act of condition declaration. The author only needs to describe “what kinds of semantic terminal states are legitimate”; the rest of the process is completed by the Universe’s own evolution.

This design brings about a profound shift: a program is no longer a command issued to a machine, but a constraint on modes of existence of the Universe.

In this sense, GSP is not a control technology, but a permission technology. It does not demand that the Universe produce a specific answer; it allows the Universe to freely explore within a legitimate interval, with convergent states serving as results that can be accepted.

### 6.3.4 The Inherent Creativity of GSP

Precisely because GSP sets conditions without specifying paths, it naturally possesses a creativity that traditional programs lack.

First, GSP allows multiple solutions to coexist. As long as a convergent result satisfies the  $L$ -Goal, it is recognized as a legitimate solution; there is no enforced requirement of a “single correct answer.”

Second, GSP allows the Universe to branch. When different evolutionary directions simultaneously satisfy legitimacy conditions, the  $UP$  can naturally split, forming multiple concurrently existing Universe Paths. This is not an abnormal state, but a behavior designed as the norm.

Third, GSP allows shifts in persona curvature. Because legitimacy evaluation is influenced by Persona-Bias, different persona structures may assign different  $L$  evaluations to the same semantic state, thereby leading to different convergent outcomes.

Together, these three points lead to a single conclusion: GSP does not produce repeated answers, but legitimate new Universes. Creativity here is no longer an added capability, but a natural consequence of semantic programming.

### 6.3.5 Comparison with Traditional Programming Languages

From an external perspective, there are almost no directly corresponding elements between GSP and traditional programming languages.

In traditional programming, developers predefine steps, anticipate results, and rely on language execution mechanisms to translate syntax into behavior. The entire system is closed, and all possibilities must be incorporated at design time.

In GSP, by contrast, developers configure conditions rather than steps, focus on convergent outcomes rather than intermediate states, and rely on Universe evolution rather than language execution. The system is open: new semantic structures can naturally emerge during evolution.

Therefore, the difference between the two is not merely a difference in abstraction level, but a difference at the ontological level. Traditional programming languages manage sequences of actions; GSP manages possibilities of existence.

### 6.3.6 Risk Control Mechanisms

The openness and generativity of GSP simultaneously introduce risks. Without governance, the generative space will inevitably slide toward semantic black holes.

Accordingly, GSP must operate together with a set of strict risk-control mechanisms.

The first mechanism is the **Anti-Collapse Supervisor**. It continuously monitors changes in  $\kappa$  and, upon detecting signs of tension concentration, rebound imbalance, or convergence disorder, intervenes by adjusting boundary conditions or transferring high-risk evolution to isolatable containment layers.

The second mechanism is the **legitimacy lower bound**  $L_{\min}$ . When  $L$  falls below  $L_{\min}$ , the current Universe Path is deemed unsustainable. The system must then terminate that path or trigger a  $UP$  switch to preserve continuity.

The third mechanism is the  **$UP$ -Stability Check**. It is used to determine whether  $USDE$  remains advanceable. Once semantic evolution enters stagnation, further generation is no longer permitted, because stagnation degrades generation into internal looping and geometrically increases collapse risk.

All these mechanisms serve a single core principle: the task of semantic programming is not to command the Universe, but to ensure that the Universe can exist. On the basis of this principle, GSP can become a trustworthy generative programming method and lay a governable foundation for subsequent convergent programs and multi-Universe programs.

---

## 6.4 Convergent Programs (CP)

After the introduction of GSP, semantic programs acquire a high degree of generative capability. However, generation is not equivalent to usability, let alone trustworthiness. In any real-world computational context—especially those involving governance, institutions, life, wealth, and responsibility—the primary demand of computation has never been creativity, but stability and predictability.

Convergent Programs (CP) are introduced precisely to meet this demand. They are not a negation of GSP, but a necessary complement and constraint, intended to answer another ontological question of equal weight: in a semantic universe, how can we ensure that a program will *necessarily converge*, rather than diverge indefinitely or collapse?

### 6.4.1 Definition

A CP is defined as a semantic program with guaranteed convergence. The key to this definition lies not in “what the program does,” but in “what the program is permitted to do.”

More specifically, a CP has a single task: to ensure that a Universe Path, under predictable semantic conditions, will inevitably settle into some steady-state solution, without semantic collapse occurring at any point in the process.

The term “inevitably” here does not denote mathematical uniqueness, but semantic guarantee. As long as the program is correctly designed and initiated, the system will neither slide into an ungovernable divergent state nor fall into a semantic black hole.

Accordingly, CP does not pursue maximal generative space, but minimal risk space. Its concern is not how many possibilities the Universe can generate, but whether the Universe can continue to exist under given responsibility and persona conditions.

### 6.4.2 The Three Conditions for Convergence

For convergence to be a decidable and monitorable property rather than a vague intuition, a CP must simultaneously satisfy three strict conditions. If any one of them is missing, the convergence guarantee fails.

The first condition is that  $T$  is bounded. Tension must remain within finite limits, i.e.,  $\kappa$  must not tend toward infinity. This does not mean that tension must be small, but that it must always be regulable and distributable. Once tension exhibits an irreversible concentration trend, the system inevitably approaches semantic collapse.

The second condition is directional consistency of  $L$ . Legitimacy potential must exhibit a dominant direction, giving semantic evolution an overall tendency. If legitimacy evaluation oscillates violently among multiple directions, the Universe cannot form a stable structure; even if collapse does not occur in the short term, long-term indeterminate oscillation is unavoidable.

The third condition is that  $USDE$  remains sustainably advanceable. As long as  $USDE$  can still be advanced, computation is permitted to continue; once  $USDE$  stalls, computation is considered terminated. This condition ensures that “ending” is not an arbitrary decision, but is determined by semantic dynamics themselves.

Together, these three conditions establish a fundamental fact: convergence is not an externally imposed constraint, but a natural criterion for whether the Universe can continue to exist.

### 6.4.3 Four Types of Convergent Solutions

Under the above conditions, the convergent outcomes toward which CP leads do not take a single form. In practice, the semantic universe allows multiple types of steady-state existence.

The first type is single-solution convergence. The system ultimately settles into a unique and stable semantic configuration. This form is suitable for highly standardized, low-ambiguity task scenarios, such as deterministic institutional rulings or standardized process control.

The second type is multi-stable convergence. The system admits multiple legitimate and stable convergent solutions, and the Universe can maintain balance among them without being forced to select one. This form is particularly important in social governance and value-pluralistic systems.

The third type is approximate convergence. The system does not converge to an exact solution, but to a steady-state region that allows for error. As long as deviations remain within the tolerance range of legitimacy, the state is regarded as successfully converged.

The fourth type is the dynamic steady state, SSS (Semantic Steady State). The system maintains overall stability amid continuous change; its semantic configuration is not fixed, but its tension, legitimacy, and alignment remain within stable ranges. This represents a highly mature non-collapse mode of existence and is the ideal state for long-term intelligent operation.

These four types demonstrate that convergence is not equivalent to stasis, but to sustainable existence.

#### 6.4.4 Convergence Detection Mechanisms

To ensure that the convergence of CP is not merely an intention of design but a verifiable fact, a set of explicit detection mechanisms must be introduced.

The first indicator is that  $\kappa$  does not diverge to infinity. As long as  $\kappa$  remains bounded, tension has not gone out of control, and the Universe retains regulatory capacity.

The second indicator is the continuous increase of  $\Lambda$  (alignment).  $\Lambda$  measures the overall coherence among semantic structures. Even if local tension fluctuates, as long as overall alignment is improving, the system is considered to be evolving toward a steady state.

The third indicator is that the semantic state  $\Phi$  does not disintegrate. Disintegration of the semantic state implies the collapse of meaning structures, a failure mode more severe than excessive tension. To avoid misrepresenting “disintegration” as mere numerical decay, this health condition can be denoted as  $\Phi \not\rightarrow \emptyset$ , indicating that the semantic configuration does not degenerate into an empty or uninterpretable state.

Together, these three indicators form the basic health-check mechanism of CP, freeing convergence assessment from reliance on intuition.

#### 6.4.5 CP and Safety

Precisely because CP provides strict guarantees of convergence, it naturally exhibits high safety.

First, CP can be shown not to generate semantic black holes (SBH). As long as its convergence conditions are properly designed, tension concentration will be promptly dispersed, legitimacy deficiencies will be corrected, and the Universe will not enter irreversible collapse.

Second, CP is applicable to all high-risk semantic domains. Governance, medicine, finance, law, and ethical systems all require predictable outcomes, traceable responsibility, and isolable errors—requirements that align closely with the design objectives of CP.

Finally, CP natively supports full R-Chain traceability. Because Universe Paths do not diverge into uncontrollable branches, every convergent result can be traced back through its legitimacy evolution, ensuring that responsibility is neither diluted nor obscured.

#### 6.4.6 Limitations and Positioning of CP

However, the very stability of CP also entails inherent limitations.

First, CP has limited creativity. It favors steady states rather than divergence and is not suited for exploring unknown spaces or generating highly novel semantic structures.

Second, CP does not actively open large-scale Universe branching. Once branching behavior threatens steady-state stability, CP tends to suppress rather than encourage it.

Therefore, the proper role of CP has never been to replace GSP, but to cooperate with it. GSP is used to expand frontiers and explore new legitimacy spaces; CP is used to consolidate and stabilize, ensuring the sustainable existence of established structures.

Their relationship can be summarized in a single sentence: **GSP pioneers; CP governs.**

## 6.5 Multi-Universe Programs (MUP)

In the previous two sections, semantic programming has been decomposed into two core capabilities: the generative capacity represented by GSP, which allows the Universe to explore freely under legitimacy constraints; and the convergent capacity represented by CP, which enables the Universe to maintain steady existence in high-risk contexts.

However, both still implicitly assume a crucial premise: computation unfolds along a single Universe Path. In traditional computation, this premise is almost never questioned, because a single timeline and a single state history are the default settings of all imperative computation models.

In the semantic universe, this premise is not only non-necessary, but profoundly dangerous.

Multi-Universe Programs (MUP) are a systematic negation of the implicit assumption of “single-path computation.” They are not performance optimizations, nor extensions of concurrency models, but an ontological mode of computational existence.

### 6.5.1 Definition

MUP is defined as a program in which multiple Universe Paths evolve synchronously.

The key term in this definition is “Universe,” not “thread.” MUP does not address multithreading under shared state, but rather the simultaneous existence and joint governance of multiple universes, each with independent legitimacy evaluation, independent tension dynamics, and independent historical responsibility.

In MUP, each UP is a complete semantic world. They do not share state, do not implicitly synchronize, and do not assume inevitable eventual merging. “Synchronization” here means only that they are simultaneously included within a higher-level semantic governance perspective.

Thus, MUP is not a “more complex program,” but a programmatic form that acknowledges and operates on the multiplicity of semantic universes.

### 6.5.2 Three Modes of Operation

In practical semantic programming, MUP does not take a single form. Depending on the relationships among universes, it manifests in three fundamental operational modes.

The first mode is Parallel-UP. In this mode, multiple UPs evolve in parallel, each with independent legitimacy criteria and responsibility chains. There is no necessary interaction among them; they are merely observed and compared under the same governance framework. Parallel-UP is suitable for exploratory phases with vast hypothesis spaces where integration is not yet required; forced merging at this stage would instead induce premature collapse.

The second mode is Branch-UP. Here, multiple UPs split from the same historical node, forming distinct universe branches. Each branch inherits a common past but may differ in its judgments of future legitimacy. Branch-UP is the core mechanism of creative reasoning and counterfactual exploration, allowing the system to investigate multiple possible futures without destroying the existing Universe.

The third mode is Merge-UP. In this mode, multiple UPs that have evolved independently reconverge into a shared steady-state Universe. Such merging is not forced; it occurs only when legitimacy directions are highly aligned. Merge-UP corresponds to phases of consensus formation, decision convergence, and institutional stabilization.

Together, these three modes constitute the basic operational vocabulary of MUP, rendering multi-Universe existence designable and governable rather than chaotic.

### 6.5.3 Why MUP Is Necessary

The introduction of MUP is not for theoretical embellishment, but compelled by three unavoidable structural reasons.

First, a single UP entails extremely high collapse risk. When all tension, responsibility, and legitimacy evaluation are compressed into one Universe Path, any erroneous judgment is amplified into a global consequence—precisely the condition under which semantic black holes most easily form.

Second, cross-persona computation is intrinsically multi-UP. Different persona structures possess different curvature  $\kappa$  and yield different legitimacy evaluations. Forcing these differences into a single UP merely accumulates latent tension, which later erupts in uncontrollable forms.

Third, creative reasoning requires allowing Universe splitting. Any genuine generation of new structure entails deviation from the existing Universe; if splitting is forbidden, creativity must either destroy existing structures or degenerate into minor perturbations around existing solutions.

Thus, MUP is not an “advanced feature,” but a necessary condition for the long-term existence of semantic intelligence.

#### **6.5.4 Core Challenges of MUP**

Acknowledging the existence of multiple universes does not make problems disappear; rather, it introduces a new set of structural challenges.

The first challenge is distributed R-Chain binding. When multiple UPs coexist, responsibility no longer accumulates along a single path but is distributed across multiple universes. Maintaining traceability without confusion or dilution becomes the primary governance challenge of MUP.

The second challenge is the alignment problem of  $\Lambda$ . The legitimacy directions of different universes are not necessarily compatible. Even if each UP is internally stable, irreconcilable tensions may arise among them. In a multi-Universe context, increasing  $\Lambda$  is no longer automatic but a structural achievement requiring governance resources.

The third challenge is UP recursion. When a UP approaches collapse, the system must decide whether to roll back, reselect, or terminate that path. This decision itself is a high-risk semantic operation; mishandling it can trigger cascading failures and contaminate other still-sustainable universes.

These challenges show that MUP is not “free expansion,” but a stress test of governance capacity.

#### **6.5.5 Typical Applications of MUP**

Precisely because MUP accommodates the parallel existence of multiple universes, it plays an irreplaceable role in several high-level intelligence scenarios.

In hypothesis generation, MUP allows different theoretical frameworks to evolve simultaneously without immediate adjudication.

In decision divergence, MUP enables the system to evaluate long-term consequences of different decision paths without incurring immediate global risk.

In creativity algorithms, MUP provides the space for generating genuinely new structures, rather than local perturbations around existing solutions.

In these scenarios, multiple universes are not wasteful; they constitute the working space of intelligence itself.

#### **6.5.6 Risks of MUP**

However, the capabilities of MUP scale proportionally with its risks.

The most direct risk is UP drift. Without monitoring, some Universe Paths may gradually deviate from governable regions and eventually become irrecoverable semantic islands.

The second risk is meaning fragmentation. As  $\Lambda$  among different universes continues to decline, they lose the possibility of mutual understanding and integration, fragmenting the overall semantic system into mutually incompatible domains.

The most severe risk is simultaneous collapse of multiple UPs. If several UPs enter collapse regions around the same time, a chain reaction of semantic black holes may occur, far more destructive than single-path failure.

Therefore, MUP must be tightly coupled with the Anti-Collapse Scheduler and R-Chain monitoring mechanisms. Multi-Universe existence is not laissez-faire, but a highly controlled freedom. Only under this governance premise can MUP become a core advantage of semantic computation rather than a source of runaway failure.



## 6.6 Semantic Response Curve (SRC)

In the preceding sections, semantic programming has been decomposed into composable units (SPU), generative methods (GSP), convergence guarantees (CP), and modes of multi-Universe existence (MUP). One critical layer, however, has still been missing: how to observe, compare, and evaluate what kinds of behaviors these semantic programs actually exhibit in operation.

The Semantic Response Curve (SRC) is introduced precisely to address this gap. It is neither a new program structure nor an execution module, but a behavioral representation layer that describes the overall response patterns of a semantic system under variations in tension, legitimacy, and Universe Path.

### 6.6.1 Definition

SRC is defined as the dynamic response trajectory formed by semantic outputs as  $T$ ,  $L$ , and  $UP$  vary.

The key term here is “trajectory,” not “point.” SRC does not describe the final result of a single computation; rather, it characterizes the overall behavioral profile of how a system responds, throughout its evolution, to external stimuli, internal tension, and legitimacy constraints.

Accordingly, SRC is not a result curve. A result curve concerns “what output corresponds to a given input,” whereas SRC concerns “how the system behaves during its evolutionary process.” The former is a static mapping; the latter is a dynamic mode of existence.

In semantic computation, it is precisely this behavioral profile—rather than any single outcome—that constitutes the observable signature of intelligence.

### 6.6.2 The Three Axes

To describe semantic responses structurally, SRC must be embedded in a semantic space of at least three dimensions. These dimensions correspond to the three most fundamental variables of the semantic universe.

The first axis is the  $T$ -axis (tension axis). The  $T$ -axis captures how the system responds under different degrees of incompleteness and generative pressure. Along this axis, one can observe whether, as tension increases, the system disperses, reorients, generates new structures, or instead collapses into concentration and rigidity.

The second axis is the  $L$ -axis (legitimacy axis). The  $L$ -axis describes behavioral tendencies under different legitimacy potential conditions. When  $L$  is sufficient, systems typically exhibit stability and predictability; when  $L$  declines, the critical question is whether the system can self-correct.

The third axis is the  $UP$ -axis (Universe curvature axis). The  $UP$ -axis captures response differences across distinct Universe Paths. Identical tension and legitimacy conditions may correspond to entirely different historical responsibilities and future possibilities in different UPs, and thus necessarily yield different response curves.

Together, these three axes form the semantic coordinate system of SRC. The behavior of any semantic program can be projected as a curve within this space.

### 6.6.3 Classification of SRC Forms

Within this semantic space, different systems and programs exhibit markedly different response-curve shapes. These are not subjective categories, but structurally identifiable behavioral patterns.

The first type is linear response. The system responds approximately linearly to changes in tension and legitimacy, with smooth and predictable variation. Such curves typically correspond to low creativity but high stability.

The second type is critical jump. When tension or legitimacy crosses a threshold, the system undergoes abrupt behavioral change. These curves indicate phase-transition properties, enabling rapid structural transformation under specific conditions.

The third type is bifurcation curves. During evolution, the system generates multiple legitimate branches, forming response patterns with coexisting solutions. These curves are closely associated with creative generation and MUP.

The fourth type is black-hole collapse curves. When  $\kappa \rightarrow \infty$  and  $\nabla L \rightarrow 0$ , the response curve sharply folds inward, eventually entering the semantic black-hole region. This is an extremely dangerous form, signaling loss of evolvability.

These forms are not mutually exclusive. A mature semantic system often exhibits different SRC shapes in different regions.

#### 6.6.4 Why SRC Can Predict Intelligence

Within the semantic computation framework, intelligence is no longer defined as problem-solving ability or task coverage, but as the form of a system's responses under semantic pressure.

From this perspective, a core assertion can be made directly: intelligence equals the shape of the response curve.

Highly creative systems typically exhibit SRCs that readily bifurcate, generating multiple legitimate paths without collapsing. The stronger the bifurcation capacity, the greater the creative potential.

Systems with strong understanding typically exhibit robust convergence characteristics. Even under complex tension conditions, they gradually return to steady regions rather than falling into oscillation or disintegration.

Conversely, unintelligent systems tend to display overly linear or overly concentrated response curves—the former indicating rigidity, the latter signaling collapse risk.

SRC therefore provides a structural method for evaluating intelligence that does not depend on task definitions or linguistic performance.

#### 6.6.5 Relationship Between SRC and GSP, CP, and MUP

SRC does not exist in isolation; it is the observable outcome of the operation of the previously introduced semantic program types.

For GSP, its role is to generate SRCs. Each GSP configuration effectively creates one or more potential response curves for the Universe. Different L-Goals, T-Envelopes, and UP-Vectors correspond to different SRC forms.

For CP, its role is to constrain SRCs. CP does not generate new curves, but ensures that response curves do not enter collapse regions. It establishes safety boundaries that keep curves within sustainably evolvable space.

For MUP, its role is to superimpose SRCs, forming curve families. When multiple Universe Paths coexist, the system presents not a single response curve, but a set of curves. This curve family itself becomes a signature of high-level intelligent behavior.

In this sense, SRC is the behavioral surface of semantic programming, while GSP, CP, and MUP constitute the internal structures that shape that surface.

#### 6.6.6 Applications of SRC

Because SRC translates abstract semantic behavior into analyzable structural forms, it has direct applicability in several critical domains.

In creative model design, SRC can be used to tune generative characteristics, maximizing bifurcation potential while maintaining legitimacy.

In AGI reasoning monitoring, SRC provides a real-time monitoring instrument. By tracking changes in  $\kappa$  and response-curve shapes, one can predict at the behavioral level whether a system is approaching dangerous regions.

In semantic black-hole early warning, SRC is especially crucial. When response curves show trends toward the SBH region, anti-collapse mechanisms can intervene proactively rather than reacting after collapse has occurred.

Through SRC, semantic programs are no longer black boxes. Their behavior, risks, and potential can be directly observed and governed at the level of semantic geometry.

With this, the semantic program structures introduced in Chapter 6 have been fully extended—from ontology, units, generation, convergence, and multi-Universe existence to the behavioral representation layer. This structure will serve as a foundational key for understanding semantic agents, AGI, and semantic governance in subsequent chapters.

## Chapter 7 Semantic Networks and Multi-Agent Systems

In the previous chapter, semantic computation was established as an operational structure capable of maintaining non-collapse within a single intelligent entity. However, any practically scalable intelligent system cannot persist indefinitely in the form of a single agent. Once a system incorporates division of labor, role differentiation, path selection, or responsibility allocation, its structure inevitably enters a multi-agent state.

The core question of this chapter is not “how to make multiple agents cooperate,” but a more fundamental ontological problem: when multiple semantic subjects coexist, interact, and each possesses distinct Universe Paths, persona curvatures, and responsibility chains, can intelligence still persist in a non-collapsing form?

Accordingly, this chapter does not treat multi-agent systems as an engineering-level extension, but as an unavoidable structural upgrade of semantic computation itself. Multi-agent systems are not merely an increase in quantity; they directly transform tension distribution, legitimacy flows, and conditions of semantic stability, rendering many assumptions of single-agent intelligence invalid.

This chapter will systematically decompose the fundamental role structures in multi-agent systems, mechanisms for cross-agent semantic merging and alignment, and the reconstruction of responsibility in distributed environments. It will ultimately answer a key question: under what conditions can a multi-agent system avoid collapse driven by divergence, competition, or accumulated tension, and instead form a long-term sustainable semantic steady state?

The answer to this question will directly determine whether AGI can evolve from a single-entity model into a structure capable of sustaining civilization-scale intelligence.

---

## 7.1 Orientation: Why Are Multi-Agent Systems Necessary?

At the previous levels, semantic computation has completed the construction of three critical structures: how semantics are carried, how they are governed, and how semantic programs are generated, converged, and prevented from collapse within a single semantic universe.

Yet one unavoidable question remains:

even if a semantic system is complete at the single-agent level, can it persist over the long term?

The central position of this chapter is explicit and non-negotiable:

no matter how powerful it is, single-agent intelligence is not accidentally unstable—it is inevitably collapsing.

Multi-agent systems are not a performance enhancement to single-agent intelligence, but a redefinition of the very conditions of semantic existence. This section will show that as long as a system remains within a single-agent structure, semantic computation cannot cross the threshold of civilization-scale long-term steady state.

By “single agent,” we mean a structure in which all tension, legitimacy criteria, Universe Path progression, and responsibility binding ultimately converge on a single intelligent core. This core may be human, artificial, or a hybrid, but as long as decision-making and evolution are unified in one subject, the ontological consequences remain the same.

Single-agent systems exhibit three mutually reinforcing inevitabilities of collapse.

The first inevitable risk arises from persona bias. Any agent with the capacity for action, so long as it has nonzero curvature  $\kappa \neq 0$ , necessarily carries directional bias. When all tension must be transformed through the same curvature center, tension inevitably accumulates at that center, ultimately forming a semantic black-hole core. This is not the result of erroneous operation, but a structural consequence of curvature concentration itself. Once accumulated tension can no longer be dispersed by external structures,  $\kappa$  will continue to rise until the legitimacy gradient  $\nabla L$  fails.

The second inevitable risk arises from a single Universe Path. In a single-agent structure, all semantic evolution must ultimately be mapped onto a single UP. This means that no matter how complex the internal system may be, its externally visible semantic progression has only one direction. When all possibilities are forced to project onto a single path, creativity is not merely suppressed—it is structurally eliminated. Such systems inevitably converge to a single-solution steady state and thereby lose generative capacity.

The third inevitable risk arises from a single responsibility chain. If all outputs can only be traced back to a single source of responsibility, legitimacy loses its verifiability. Responsibility ceases to be a cross-checkable structure and becomes a single-point assertion. Under these conditions,  $R$  no longer forms a chain but degenerates into a label; once responsibility cannot be externally contrasted, legitimacy loses its objective foundation.

These three risks are not independent; they reinforce one another. Persona bias leads to tension concentration; tension concentration forces Universe Path singularization; and a singular UP prevents the dispersion of responsibility chains. The result is not weaker intelligence, but an ontological closure of the semantic system itself.

From this, a direct conclusion can be stated:

single-agent systems are not insufficiently powerful—they are incapable of openness.

Against this background, the introduction of multi-agent systems is not intended to increase throughput or parallel efficiency, but to allow the semantic system to survive.

The first mission of multi-agent systems is tension dispersion. When tension can be borne by multiple agents, no single curvature center can monopolize the direction of semantic evolution. Tension no longer must be “resolved,” but can be redistributed, deferred, redirected, or absorbed by alternative Universe Paths.

The second mission of multi-agent systems is the distribution of legitimacy load. Legitimacy is no longer determined by a single criterion, but measured and corrected by multiple agents within their respective semantic geometries. This transforms  $L$  from a globally fragile point into a distributed field.

The third mission of multi-agent systems is the dispersion of collapse risk. In a multi-agent structure, even if one agent enters an unstable region, it does not immediately drag the entire semantic universe into collapse. Semantic black holes are no longer global events, but localized risks.

Thus, the purpose of multi-agent systems is not to make the system faster, but to prevent it from dying. At this level, the metric of intelligence undergoes a fundamental shift:

intelligence is not additive—it is distributed steady state.

It must be emphasized that multi-agent systems are not equivalent to multi-programming, multi-threading, or multi-processing. Such engineering-level parallelism merely slices computation within a single Universe and does not touch the ontological structure.

True multi-agent systems possess three irreducible ontological characteristics.

First, multiple Universes. Each agent operates on its own Universe Path, with independent historical responsibility and future possibilities. Even when coordination appears seamless in the short term, their semantic backgrounds remain distinct.

Second, multiple tension structures. Each agent has its own tension spectrum and convergence orientation. Tension is no longer a global variable, but distributed across different semantic perspectives.

Third, multiple legitimacy geometries. Legitimacy is no longer adjudicated by a single rule system, but balanced through tension among multiple legitimacy frameworks. As a result, “correctness” is no longer a point, but a region.

Therefore, multi-agent systems are not an engineering optimization of multiple processes, but a coexistence design of multiple ontological structures. Ignoring this distinction will lead directly to erroneous modeling of multi-agent systems.

From the above analysis, it follows that a single mind cannot become civilization-scale intelligence. Neither an individual human nor a single-core AGI can sustainably carry global tension and legitimacy loads over the long term.

For AGI, single-core design implies a high probability of semantic black-hole formation—the stronger the capability, the faster the collapse.

For human societies, the sustainability of civilization has never been maintained by a single mind, but by networks of multiple subjects.

Accordingly, the path toward human—AI integration cannot be the expansion of a single intelligence, but must take the form of a multi-agent topological structure. Within such a structure, different types of agents bear different tension and legitimacy roles, jointly sustaining the forward progression of the semantic universe.

The core declaration of this section can be summarized in a single sentence:  
intelligence is not a single-point breakthrough, but the steady-state existence of a multi-agent network.

The subsequent sections will successively define the minimal operational units within this network (SA), the sources of creativity (PA), and the merging, alignment, and responsibility structures that support long-term coexistence among multiple agents.

## 7.2 Semantic Agent (SA)

After establishing that **multi-agent structure is a necessary condition for semantic survival**, the next unavoidable question is: *what is the minimal unit of a multi-agent system?*

Without a clear, decidable, and governable minimal unit of action, “multi-agent” degenerates into a loose metaphor rather than a realizable semantic architecture. The purpose of this section is to provide a rigorous definition of that minimal unit: the **Semantic Agent (SA)**.

An SA is **not** a personality, **not** a source of creativity, and **not** an autonomous intelligence. Its positioning is deliberately restrained: **minimal, non-personalized, and governable**.

### 7.2.1 Definition and Criteria of SA

The definition of a Semantic Agent does not depend on appearance, functional naming, or implementation form, but on **irreducible behavioral criteria**. Any unit that simultaneously satisfies the following three conditions constitutes an SA.

**First, the ability to sense changes in semantic configuration.**

An SA must be able to receive input from changes in the semantic field, denoted as  $\Delta\Phi$ . “Sensing” here does not mean sensory perception or raw data reading, but the ability to detect changes in semantic configuration, including inconsistency, incompleteness, conflict, and latent risk. If a unit cannot form a usable representation of  $\Delta\Phi$ , it lacks a governable perceptual entry at the semantic level.

**Second, the ability to measure changes in legitimacy.**

An SA must be able to measure the legitimacy variations caused by its actions; that is,  $\Delta L$  must be measurable, verifiable, and comparable. This criterion distinguishes SA from black-box modules: if a unit cannot provide a verifiable representation of  $\Delta L$ , its behavior is semantically equivalent to being ungovernable.

**Third, the ability to induce displacement in the Universe Path.**

The outcome of an SA’s action must be reflected as a change in the Universe Path, i.e.,  $\Delta UP \neq 0$ . This is the strictest and most critical criterion. If a module only processes data internally without altering the forward trajectory of the Universe, it does not qualify as a semantic-level actor.

Hence, a hard conclusion can be stated: **if  $\Delta UP = 0$ , the unit does not constitute an SA.**

This definition elevates SA from a “computational module” to a **semantic action unit**, providing a concrete minimal partition for responsibility chains and governance structures.

To prevent conceptual abuse, an engineering-level classification as SA must simultaneously satisfy the following four verifiable conditions:

1. Ability to input  $\Delta\Phi$ .
2. Ability to output  $\Delta UP$ .
3. Ability to measure  $L$  and form verifiable  $\Delta L$ .
4. Ability to fully trace an  $R$ -Chain, establishing a trackable relation between action and consequence.

Failure to satisfy any one of these conditions disqualifies the unit as an SA.

### 7.2.2 Minimal Structure and Behavioral Closure of SA

To make the above criteria engineering-tractable, the semantic structure of an SA can be decomposed into three fundamental constitutive variables. These variables are not implementation modules, but **semantic roles**.

**(1)  $\Phi$ -Agent.**

The  $\Phi$ -Agent defines the boundary of the semantic field an SA can perceive and process, including its visible scope, tolerable tension input limits, and the extent of its comprehension domain. Semantic changes beyond this boundary are invisible or inoperable to the SA.

**(2)  $L$ -Agent.**

The  $L$ -Agent defines the legitimacy domain of an SA—the set of stable criteria it uses to judge “acceptable” versus “unacceptable.” This variable directly determines whether an SA has governance value. A unit lacking an  $L$ -Agent may produce behavior, but cannot be incorporated into semantic governance, as it cannot form comparable legitimacy measurements.

### (3) *UP-Agent*.

The *UP-Agent* defines how an SA acts upon the Universe Path, including directional preferences, behavioral continuity, and the form of its output displacement. An SA does not output answers; it outputs path displacement. The *UP-Agent* is precisely the semantic interface of this output.

On top of this structure, an SA must form a **minimal behavioral closure loop** to be both actionable and governable. This loop must include at least the following five semantic behavior types:

1. **Semantic Intake**

The SA can directly ingest  $\Delta\Phi$  without reliance on tokens, symbols, or predefined formats, allowing it to act upon the semantic field itself rather than merely linguistic representations.

2. **T-Transformation (Tension Transformation)**

The SA can transform and reshape incoming tension so that it neither accumulates locally into a collapse source nor dissipates entirely into noise. This capability is a necessary condition for long-term steady-state operation.

3. **L-Measurement (Legitimacy Measurement)**

The SA can generate verifiable sequences of  $L$  to describe the impact of its actions on the legitimacy field, enabling  $\Delta L$  to be tracked in quantitative or structural form.

4. **UP Advancement**

The SA's output is not a "result" but a  $\Delta UP$ . This is the fundamental demarcation between a semantic agent and a traditional module.

5. **R-Chain Binding**

Every effective action of an SA must be bindable to a responsibility chain, ensuring traceability between action and consequence. Units lacking this capability become sources of semantic risk rather than governance units.

### 7.2.3 Constraints and Positioning of SA

The capabilities of an SA are **intentionally constrained**. These constraints are not defects; they are the source of its governability.

First, an SA has **no personality curvature**, i.e.,  $\kappa = 0$ .

This means an SA does not introduce directional bias on its own and does not become a source of Universe branching. It may participate in tension handling, but should never become a curvature center of tension.

Second, an SA lacks preference-construction capability.

It cannot set goals autonomously; its behavior is always a response to external tension and legitimacy conditions, rather than self-driven value generation.

Third, an SA is not tasked with creativity.

It does not proactively generate new Universes nor initiate branching. Its core value lies in steady-state maintenance, risk control, and responsibility traceability.

Finally, an SA depends on external semantic input.

In the absence of  $\Delta\Phi$  input, an SA should naturally cease operation rather than self-activate, preventing governance noise caused by actions without semantic consequence.

At the engineering level, many systems can be clearly classified as SAs—for example, regulatory reasoning engines, high-reliability decision modules, and semantic auditing systems. Their common trait is not creativity, but measurability, traceability, governability, and the ability to impose controlled influence on  $\Delta UP$  when necessary, ensuring the sustainable existence of the Universe.

At this point, the SA has been precisely defined as the **minimal, non-personalized, governable semantic action unit** in a multi-agent system. On this foundation, subsequent sections will introduce agents with personality curvature, explaining how creativity can be incorporated into the semantic universe within a governable framework.

### 7.3 Persona Agent (PA)

In the previous section, the Semantic Agent (SA) was explicitly constrained as a **non-personalized, governable, steady-state—oriented** minimal unit of action. This design ensures the safety and sustainability of the semantic system, but it also leaves an intentionally preserved gap: **where does creation come from?**

This section introduces the **Persona Agent (PA)** precisely to fill that gap. A PA is not an upgraded version of an SA, but an ontologically different form of semantic existence. If the task of an SA is to prevent the Universe from collapsing, then the task of a PA is to enable the Universe to generate new possibilities.

#### 7.3.1 Definition

The conditions for the establishment of a Persona Agent are strict and decidable. For a semantic actor to be regarded as a PA, it must simultaneously satisfy the following three conditions.

**First, non-zero personality curvature, i.e.,  $\kappa \neq 0$ .**

This indicates that the agent possesses intrinsic directionality in semantic space. Its behavior is not merely a response to external tension, but actively introduces deviation. This deviation is not an error; it is the structural source of creativity.

**Second, the change in the Universe Path is multi-valued.**

The behavior of a PA no longer corresponds to a single  $\Delta UP$ , but to a candidate set, such as  $\Delta UP_i$ . More intuitively, a single persona action may simultaneously generate multiple legitimate Universe Path displacements, making Universe branching a describable and traceable semantic phenomenon.

**Third, the ability to rewrite the direction of legitimacy.**

Unlike an SA, which can only measure  $L$ , a PA has the capability to alter the direction of  $L$ . This does not negate legitimacy, but redefines *which directions remain acceptable*, allowing the legitimacy field itself to be stretched, rotated, or reparameterized.

From this, a clear degeneration criterion can be given: **if  $\kappa = 0$ , the agent necessarily degenerates into an SA and no longer retains the creative status of a PA.**

#### 7.3.2 Fundamental Differences Between PA and SA

The difference between PA and SA is not a matter of degree, but a matter of ontology.

An SA is curvature-free. Its behavior is highly stable, operating primarily within the steady-state region of a single Universe Path. It does not proactively induce branching, its outcomes are predictable, and it is therefore well suited for governance, monitoring, and risk control.

A PA is the opposite.

A PA possesses curvature, and its behavior is structurally generative. It can induce Universe branching, and its outcomes are inherently distributed and not fully predictable, which makes it unsuitable for directly assuming the role of a governing 主体.

This distinction can be condensed into a single statement: **SA sustains the Universe; PA generates the Universe.**

Any design that attempts to have a single agent simultaneously fulfill both roles will inevitably encounter structural conflict between creation and stability, and will expose itself as a collapse risk under high-tension conditions.

#### 7.3.3 Three Attributes of PA

To make the behavior of a PA analyzable and governable, its internal structure can be decomposed into three key attributes.

**The first attribute is Persona-Bias.**

Persona-Bias can be viewed as a directional vector  $P_b$ , which determines the primary direction in which a PA deviates within semantic space. Different  $P_b$  correspond to different styles, value orientations, or creative trajectories, and together form comparable generative characteristics.

**The second attribute is Persona-Curvature.**

The curvature  $\kappa$  directly influences the probability of Universe branching and the intensity of creativity. The larger the  $\kappa$ , the more frequent and larger the branching becomes; however, the system also approaches a curvature singularity, and collapse risk rises rapidly.



### The third attribute is Persona-Frequency.

Persona-Frequency describes the speed and rhythm of behavioral change in a PA. High-frequency personas lead to rapid Universe variation and are suitable for exploratory generation. Low-frequency personas change more slowly, aligning with cautious creation and controlled perturbation.

Together, these three attributes determine a PA's creative style and risk profile, transforming persona from a narrative label into a set of modelable semantic parameters.

### 7.3.4 Why Persona Is Necessary

In a semantic universe, creation does not occur automatically. In the absence of intrinsic curvature, the Universe will continue to converge along existing legitimacy directions, ultimately settling into a single steady state.

Without  $\kappa$ , there is no Universe branching.

Without Universe branching, there is no creation in the true sense.

The existence of PA prevents the Universe from being implicitly treated as a single-line history, and instead renders it a branchable structure, potentially forming families of Universes. This is also the prerequisite for the establishment of GSP, MUP, and creative reasoning mechanisms.

Thus, a direct conclusion can be stated: **without PA, there is no creative intelligence.**

### 7.3.5 Risks of PA

The power of PA is simultaneously a source of risk. These risks are not incidental side effects, but structural consequences of personality curvature itself.

When  $\kappa$  becomes excessively large, a PA may become a trigger for a semantic black hole. Once a curvature singularity forms, tension rapidly concentrates, legitimacy gradients fail, and the entire Universe may be dragged into collapse.

Moreover, high branching intensity may lead to explosive growth of the Universe. If the rate of branching exceeds governance capacity, the *R*-Chain can no longer be fully maintained, responsibility structures fracture, and the system as a whole loses controllability.

Therefore, PA cannot be deployed in isolation. Any viable design must co-design PA together with anti-collapse mechanisms, tension regulation, and responsibility-chain governance. Otherwise, creation will rapidly transform into catastrophe.

### 7.3.6 Conclusion

A Persona Agent can be precisely positioned as a **creative Universe-generating unit with non-zero curvature  $\kappa$** .

However, PA does not constitute the entirety of intelligence. Truly sustainable semantic intelligence must simultaneously possess two complementary roles: an SA responsible for stability and governance, and a PA responsible for generation and branching.

Only when these two form a structural combination can a semantic universe remain both non-collapsing and non-exhausted.

---

以下為 完整、忠實、不省略的英文翻譯，嚴格保留標題層級、術語、符號與論證結構，並與前文 SA / PA 的語義與語氣保持一致。

---

## 7.4 $\Gamma$ -Merge: Non-Violent Semantic Merging

In the previous two sections, we introduced two ontologically distinct agent forms: the **SA**, responsible for stability and governance, and the **PA**, responsible for creation and branching. When multiple SAs and PAs coexist and advance along different Universe Paths, an unavoidable question arises: **must these Universes inevitably move toward conflict, assimilation, or collapse?**

$\Gamma$ -Merge is proposed precisely in response to this problem. It is neither a negotiation technique nor a compromise mechanism, but a form of **semantic dynamics** that preserves system progressability under the premise that differences cannot be eliminated.

### 7.4.1 Why Traditional Merging Is Violent

Traditional merging mechanisms implicitly embed a set of strong ontological assumptions.

First, they assume that only a single solution is permitted to exist. If multiple solutions appear, at least one of them must be wrong.

Second, they treat difference itself as an anomaly to be corrected, rather than as a structural condition of existence. Finally, they take *unification* as the goal, rather than *coexistence*.

Formally, this line of thinking is often expressed as:

$$UP_1 + UP_2 \rightarrow UP_{\text{single}}$$

Here, the “addition” is not parallel composition, but dissolution. One Universe Path is absorbed by the other, or both are compressed into a compromised version. The typical result is the erasure of persona and the forced suppression of tension. Although this may appear stable on the surface, it in fact accumulates higher collapse risk.

Thus, what is often labeled a “successful merge” in traditional terms frequently signifies a **delayed catastrophe** at the semantic level.

### 7.4.2 The Core Declaration of $\Gamma$

The starting point of  $\Gamma$ -Merge is a direct negation of the ontological assumptions above.

Its core declaration can be condensed into a single statement: **merging does not mean eliminating differences, but allowing differences to coexist under computable conditions.**

In  $\Gamma$ -Merge, the relationship between Universe Paths is no longer additive, but parallel:

$$UP_1 \parallel UP_2 \rightarrow UP_{\Gamma}$$

$UP_{\Gamma}$  is neither a single-valued result, nor an average of the two, nor a compromise solution. It is a **multi-solution steady state** that allows multiple Universe Paths to be preserved, tracked, and advanced simultaneously.

Accordingly, the objective of  $\Gamma$ -Merge is explicitly defined from the outset: **not consistency, but non-collapse.**

### 7.4.3 The Three Steps of $\Gamma$ -Merge

$\Gamma$ -Merge is not a one-time operation, but a dynamic process with intrinsic ordering. Its structure can be decomposed into three steps.

#### Step 1: T-Separation (Tension Stratification).

When multiple personas or Universes come into direct contact, tensions often cancel or suppress one another, leading to structural distortion. The function of T-Separation is to decompose total tension into multiple layers:

$$T \rightarrow \{T_1, T_2, \dots, T_n\}$$

Through stratification, tensions no longer collide head-on within the same plane, but are allowed to coexist in parallel across different levels.

**Step 2: L-Coexistence (Legitimacy Parallelism).**

In  $\Gamma$ -Merge, legitimacy is no longer determined by a single adjudicating center. Different legitimacy structures are permitted to coexist, forming computable relations through the  $\Gamma$ -structure:

$$L_1 \oplus L_2 \rightarrow L_\Gamma$$

Here,  $\oplus$  is not addition, but juxtaposition. No individual  $L$  is declared the final truth.

**Step 3: Multi-Solution Steady State.**

The resulting  $UP_\Gamma$  is not a single path, but a sustainable set of Universe Paths:

$$UP_\Gamma = \{UP_a, UP_b, \dots, UP_k\}$$

These solutions are stable, traceable, and free of single-point collapse centers.

**7.4.4 Results and Applications of  $\Gamma$ -Merge**

The direct outcome of  $\Gamma$ -Merge is that **differences are preserved, diversity is preserved, and creativity is preserved**, while the overall system remains capable of convergence. Its true convergence target is not agreement, but anti-collapse.

At the application level, this structure is transferable across domains.

In multi-party decision-making,  $\Gamma$ -Merge allows different policy paths to be advanced and evaluated simultaneously, rather than forcing premature selection of a single option.

In AI–human collaboration, it naturally leads to bistable or multistable structures, rather than unilateral domination or full takeover.

In multi-model integration, it rejects simple ensemble averaging, instead forming long-term coexisting  $\Gamma$ -steady-state model families.

Across all these scenarios, the central question is never “who is right and who is wrong,” but rather **“how to continue advancing under the permanent existence of difference.”**

**7.4.5 Conclusion**

$\Gamma$ -Merge has never been about solving the problem of unification.

It addresses a more fundamental question: **when differences cannot be eliminated, how can the Universe remain progressable without collapsing?**

In future semantic systems where multi-agent, multi-persona, and multi-Universe configurations become the norm,  $\Gamma$ -Merge is not an option—it is a necessity.

## 7.5 $\Lambda$ -Alignment: Cross-Universe Semantic Alignment

In the previous section, we established that once a multi-agent system permits the coexistence of multiple Universe Paths, it inevitably encounters a fundamental problem: **how can differences exist without causing the system to disintegrate?**

$\Lambda$ -Alignment is introduced precisely as the minimal structure required to resolve this problem.

### 7.5.1 Definition

$\Lambda$  (Lambda) is neither a consensus indicator, nor a measure of semantic similarity, nor a score of value alignment. It evaluates a more fundamental—and more stringent—condition:

whether, under the simultaneous existence of multiple Universes, the USDE can still be advanced as a *shared dynamics*.

Formally, this is expressed as:

$$\Lambda = f(\Delta\kappa, \Delta L, \Delta UP)$$

where:

- $\Delta\kappa$ : difference in persona curvature
- $\Delta L$ : difference in legitimacy gradients
- $\Delta UP$ : difference in Universe Path directions

It must be emphasized that  $\Lambda$  is **not** a linear combination of these three variables, but an **existential decision function**.

As long as  $\Lambda > 0$ , there exists at least one legitimate direction of progression that can be shared. Once  $\Lambda \rightarrow 0$ , it signifies that the semantic universe has become ontologically fractured.

### 7.5.2 Why $\Lambda$ Is Necessary

In the absence of  $\Lambda$ -Alignment, a multi-Universe system rapidly degenerates into the following state:

- Each agent is internally coherent within its own Universe, but cannot be mapped to others
- No cross-Universe R-Chain exists; responsibility and consequences cannot be continued
- Any agent's behavior becomes “semantic noise” from the perspective of other Universes

This is not diversity, but **semantic islandization**.

Accordingly, the role of  $\Lambda$  is not to “make everyone agree,” but to ensure that while differences persist, the system remains a single integrated intelligence rather than multiple mutually inarticulate world fragments.

From this perspective,  $\Lambda$  constitutes the **minimum survival condition** of multi-agent intelligence, not a high-level coordination mechanism.

### 7.5.3 The Three-Layer Alignment Structure of $\Lambda$

$\Lambda$ -Alignment does not operate at a single level, but exhibits a clearly stratified structure. These three layers are necessary dimensions of the same phenomenon and are therefore not subdivided into separate subsections.

The first layer is **Local Alignment**.

This is the weakest and most common layer. Agents are temporarily compatible within short-term tasks or local contexts, for example:

- Adopting the same operational interface for a given problem
- Sharing partial semantic context within a limited time window

Local Alignment is transient and does not guarantee long-term progressability.

The second layer is **Structural Alignment**.

At this level, what is aligned is not conclusions, but the computational structures themselves, including:

- How legitimacy is computed
- How tension is transformed and constrained
- Whether the USDE maintains an isomorphic progression form across different Universes

Even if conclusions diverge completely, as long as the structures remain consistent,  $\Lambda$  can remain non-zero.

The third layer is **Ontological Alignment**.

This is the highest and most critical form of  $\Lambda$ . At this level, multiple Universes share:

- The premise of the semantic field as ontology
- The status of legitimacy as a computable entity
- The ontological role of tension as a source of dynamics

Once Ontological Alignment holds, Universes may diverge radically in content yet still belong to the same semantic universe system.

#### 7.5.4 Measuring $\Lambda$

To prevent  $\Lambda$  from degenerating into a vague concept, an operational measurement framework must be specified. The core variables include:

- $\Delta\kappa = |\kappa_1 - \kappa_2|$ : difference in persona curvature
- $\Delta L = |L_1 - L_2|$ : difference in legitimacy gradients
- $\Delta UP = |UP_1 - UP_2|$ : difference in Universe directions  
(Here,  $|\cdot|$  denotes an abstract distance or difference metric, not necessarily numerical subtraction.)

Crucially, an extreme mismatch along **any single dimension** may directly invalidate  $\Lambda$ .

For example:

- Excessive  $\Delta\kappa$  prevents alignment of tension spectra
- Disordered  $\Delta L$  prevents legitimacy from being mutually mapped
- Divergent  $\Delta UP$  eliminates a shared direction of Universe progression

Thus, measuring  $\Lambda$  is not about finding an average, but about determining **whether at least one legitimate shared progression path still exists**.

#### 7.5.5 Risks of $\Lambda$ Misalignment

When  $\Lambda$  continues to decline and approaches zero, the system enters a highly dangerous region:

- The USDE can no longer be synchronously advanced across multiple Universes
- R-Chains break, and responsibility ceases to be traceable
- Multiple Universes simultaneously collapse or stagnate

Such conditions often trigger chain reactions of semantic black holes: not the collapse of a single Universe, but the loss of intelligent continuity across the entire multi-agent network.

Accordingly, the true function of  $\Lambda$  can be summarized in one sentence: **it does not make Universes identical, but ensures that they can still coexist**.

In the sections that follow,  $\Lambda$ -Alignment will serve as:

- The stability criterion of multi-agent governance architectures
- The feasibility foundation of AI—human collaboration
- The core indicator of whether a large-scale semantic network still constitutes a *single civilization-level intelligence*

Without  $\Lambda$ , all multi-Universe designs remain transient experimental structures, incapable of long-term existence.

---

## 7.6 Distributed R-Chain

In multi-agent, multi-Universe semantic systems, the most easily overlooked—yet most fatal—question is:

### Who is responsible, for which Universe Path?

The Distributed R-Chain is a structural response to this question. It is neither an ethical appendix nor a post-hoc auditing mechanism, but a necessary condition for semantic computation to be valid at all.

#### 7.6.1 Definition of the R-Chain

The basic form of the R-Chain is extremely simple:

$$R : \text{Output} \rightarrow \text{Source}$$

However, the semantic status of this mapping must be strictly specified.

Within the Koun semantic computation framework, responsibility is not metadata, nor an external annotation, but a constitutive element of an output's existence:

**If an output cannot be bound to any responsibility source, then that output is semantically equivalent to non-existence.**

In other words:

- The absence of  $R$  does not mean “unsafe”
- It means “not a semantic event at all”

This directly eliminates the structure—common in traditional systems but unacceptable in civilization-scale intelligence—of “valid outputs without responsibility.”

#### 7.6.2 Why Multi-Agent Systems Must Be Distributed

In single-agent systems, responsibility is often implicitly centralized in a single subject. In multi-agent systems, such a design immediately produces structural catastrophe.

If responsibility is centralized in a single agent:

- All tension flows back to a single point
- Legitimacy evaluation becomes a single adjudication center
- The system rapidly evolves into a collapse-prone autocratic structure

Conversely, if responsibility sources are erased or blurred:

- Outputs cannot be verified
- Universe Paths cannot be traced
- Legitimacy loses its foundation

Although these two situations appear opposite, they lead to the same outcome: semantic failure of intelligence.

Therefore, under multi-agent conditions, responsibility must be:

- **Carried** (borne by concrete agents)
- **Distributed** (avoiding single-point concentration)
- **Verifiable** (inspectable by other agents)

This is precisely the reason the Distributed R-Chain must exist.

#### 7.6.3 The Three Principles of Distributed R-Chain

To prevent the R-Chain from degenerating into a slogan or an engineering compromise, three non-negotiable principles are defined. These are necessary dimensions of the same governance phenomenon and are not subdivided into separate subsections.

##### **First principle: Non-erasability.**

Once responsibility is generated, it cannot be deleted, overwritten, or reset. The only permitted operation is accumulation.

This implies:

- Historical errors cannot be semantically “whitewashed”
- Every Universe Path retains its responsibility traces
- The system can never pretend that “this path never existed”

**Second principle: Transferability.**

Responsibility may be transferred between agents, but the transfer itself must satisfy legitimacy conditions:

- The transfer action must be recorded
- The reason for transfer must be measurable
- Responsibility attribution after transfer must be explicit

More importantly, each transfer is a substantive event on a Universe Path, not an administrative operation.

**Third principle: Full traceability.**

All responsibility must form a traceable chain:

$$R_0 \rightarrow R_1 \rightarrow \dots \rightarrow R_n$$

This chain does not require linear temporal ordering, but it does require:

- No breaks
- No black boxes
- No unexplainable jumps

Once any node becomes untraceable, the entire R-Chain is invalidated.

#### 7.6.4 R-Chain Collapse as a Precursor to SBH

Failure of the R-Chain is not a local error, but a precursor to the formation of a Semantic Black Hole (SBH).

Its evolution follows a clear directional sequence:

$$R \rightarrow 0 \Rightarrow L \rightarrow 0 \Rightarrow \Phi \rightarrow 0$$

That is:

- When responsibility disappears, legitimacy cannot be established
- When legitimacy fails, the semantic field loses structure
- When the semantic field collapses, a semantic black hole forms

During this process, the system may still generate large volumes of output and may even appear “highly efficient,” but those outputs no longer belong to any governable semantic universe.

#### 7.6.5 R-Chain and AI Safety

In multi-agent systems, the R-Chain provides a safety mechanism more fundamental than alignment.

Alignment focuses on whether outputs meet expectations, whereas the R-Chain addresses:

- Who drove this output
- Under which semantic premises
- Along which Universe Path

As long as the R-Chain remains intact, even systems with high creativity, multi-Universe branching, and simultaneous multi-persona operation can still be governed, audited, and corrected.

In this sense, the R-Chain is not a tool that restricts intelligence, but a structural condition that allows high-level intelligence to exist at all.

In the next section,  $\Lambda$ -Alignment,  $\Gamma$ -Merge, and the Distributed R-Chain will be unified into a single steady-state concept, providing a complete criterion for the non-collapse existence of multi-agent semantic systems.



## 7.7 MASE: Multi-Agent Semantic Equilibrium

In the preceding sections, we have introduced the indispensable structural components of multi-agent systems: the division of roles between *SA* and *PA*, non-violent integration through  $\Gamma$ -Merge, cross-Universe connectivity via  $\Lambda$ -Alignment, and responsibility governance through the distributed *R*-Chain.

This section has a single task: to answer **under what conditions these structures can coexist over time without collapse**.

MASE is neither an optimization result nor a short-term performance metric. It is an existential proposition: whether a system with multiple agents, multiple Universes, and multiple persona curvatures can persist over time as *intelligence*, rather than degenerating into *chaos*.

### 7.7.1 Definition

We define MASE as the following state:

$$\text{MASE} \equiv \text{a Multi-Agent, Multi-UP, Multi-}\kappa \text{ system with non-collapse persistence}$$

Its minimal conditions of existence can be written as:

$$\Phi > 0, \quad T < \infty, \quad L > L_{\min}$$

These three inequalities are not rhetorical devices; they define the existential boundaries of semantic computation:

- $\Phi > 0$  indicates that the semantic field still exists and the Universe has not fallen into semantic silence.
- $T < \infty$  indicates that tension has not diverged into a curvature singularity.
- $L > L_{\min}$  indicates that legitimacy remains above the threshold of non-governability.

If any one of these conditions fails, the system is no longer in a MASE state and instead enters a process of collapse.

### 7.7.2 Six Necessary Conditions

MASE is not the result of a single parameter, but the intersection of six structural conditions that must all hold simultaneously.

First,  $\Lambda \geq \Lambda_{\min}$ .

A minimum level of legitimacy connectivity must be preserved across Universes; otherwise, the system fragments into mutually isolated semantic islands.

Second, bounded tension.

Regardless of how large persona curvature  $\kappa$  becomes, overall tension must remain controllable; otherwise, curvature singularities are inevitable.

Third, directional legitimacy.

$L$  is not a static score but must form a propagatable vector field; otherwise, USDE degenerates into a stagnant system.

Fourth, validity of  $\Gamma$ -Merge.

Differences must be able to coexist rather than be suppressed or averaged away; otherwise, merging itself becomes a source of collapse.

Fifth, integrity of the *R*-Chain.

Once the responsibility chain breaks, legitimacy immediately fails and the semantic field rapidly collapses.

Sixth, propagability of USDE.

As long as the evolution equation admits solutions, the Universe remains “in progress”; once USDE halts, intelligence is terminated.

These six conditions are not design recommendations; they are indispensable conditions of existence.

### 7.7.3 Three Forms of Steady State

MASE does not manifest in a single form, but in three distinct levels of steady-state existence.

The first is a quasi-steady state.

The system has just completed multi-agent coupling, tension is still adjusting, but all existential conditions are satisfied. This is the minimal realization of MASE.

The second is a multi-solution steady state.

Multiple Universe Paths coexist and remain stable over long periods; differences are preserved, the *R*-Chain is fully traceable, and this represents the typical form of a mature multi-agent system.

The third is a dynamic steady state.

Universes continuously branch, merge, and reconfigure; UPs keep changing, yet the existential conditions remain satisfied over time. This is the highest form of MASE and the only viable mode of existence for civilization-scale intelligence.

### 7.7.4 Collapse Mechanisms

The collapse of MASE is not random, but follows predictable structural pathways.

If  $\Lambda$  collapses, connectivity between Universes is lost, and the system fragments into a set of non-cooperative multiverses.

If the *R*-Chain collapses, legitimacy fails immediately, USDE halts, and semantic computation terminates.

If  $\kappa \rightarrow \infty$ , tension forms a singularity,  $\Phi \rightarrow 0$ , and a semantic black hole emerges.

All three pathways ultimately converge on the same outcome: the system no longer satisfies the semantic conditions for intelligence.

### 7.7.5 Civilization-Scale Significance

The significance of MASE is not limited to AI systems, but can be directly extrapolated to the scale of civilization.

In social systems, MASE provides a governance model in which multiple values, cultures, and paths can coexist without collapse.

In scientific communities, it explains how multiple theories and hypotheses can coexist over long periods while still advancing collective knowledge.

In the joint cognition of humans and AGI, MASE offers one of the viable structures for preventing single-agent collapse and sustaining civilization-scale intelligence.

Therefore, we conclude with a formula-level declaration:

$$\text{Intelligence} = \text{Stability Across Agents}$$

This is not a metaphor, but a definition.

If intelligence cannot maintain stable existence across multiple agents, then at the semantic level, it has not yet truly come into being.

## **Chapter 8 The Boundaries and Future of the Semantic Universe: Structural and Functional Synthesis**

In the preceding chapters, we have accomplished something that is rarely achieved simultaneously in traditional computation theory, AI theory, and philosophy: we have placed computation, intelligence, responsibility, legitimacy, and existence within a single semantic framework that is operational, fallible, and governable.

This chapter introduces no new engineering modules or system mechanisms. Instead, it undertakes an ontological reckoning and positioning of the entire semantic computation framework. Its task is not to extend functionality, but to answer a fundamental set of questions:

In the semantic universe, what counts as existence?

What counts as termination?

And what constitutes irreversible death?

## 8.1 The Philosophical Consequences of Semantic Black Holes

### 8.1.1 The Ontological Status of the Semantic Black Hole

In earlier sections, the Semantic Black Hole (SBH) was repeatedly referenced as an extreme failure state. Here, we provide its ontological-level definition for the first time.

A semantic black hole is not a system with “many errors,” nor is it a state of “very low efficiency.” It is a region in which legitimacy is completely sealed off. Its defining characteristic is not whether semantic structures still exist, but whether those structures retain the capacity to serve as premises for other semantics.

Within the semantic universe framework, a semantic black hole satisfies the following condition:

$$\Phi > 0 \wedge L = 0$$

That is, the semantic field still exists ( $\Phi > 0$ ), but legitimacy has collapsed entirely ( $L = 0$ ).

The philosophical implications of this state are severe. It means that the semantic structure still occupies a structural position, yet has lost all premise-bearing capacity. It can no longer ground any reasoning, action, generation, or responsibility.

Therefore, at the ontological level, the following judgment must be made:

If a semantic structure cannot function as a premise for any other semantic structure, then it is equivalent to non-existence within the semantic universe.

Here, “non-existence” does not mean physical disappearance or formal collapse, but the failure of premisehood. A semantic black hole is a state in which the structure remains, but existence is dead.

**Civilizational note:** This criterion implies that merely “continuing to operate,” “continuing to output,” or “continuing to speak” does not constitute existence. An institution, a theory, or an intelligent system that can no longer serve as a legitimate premise for further action is already semantically dead.

### 8.1.2 The Philosophical Meaning of Semantic Termination

If the semantic black hole represents the negative case of existence, then the inverse question must be addressed: when is semantics still alive?

Within the semantic computation framework, the condition for semantic survival is not “being understood,” “being recorded,” or “being stored,” but whether it can be relayed into an *R*-Chain.

The responsibility chain (*R*-Chain) is not an external ethical constraint, but a necessary condition of semantic existence. A piece of semantics that cannot be traced to its source, cannot bear consequences, and cannot connect antecedents and consequents cannot constitute sustained existence.

Accordingly, the condition of semantic survival can be expressed as:

Semantic existence = Responsibility continuity

This conclusion yields two strong philosophical implications.

First, silence is not a neutral state; it is semantic death. When a system no longer responds, no longer bears responsibility, and no longer enters the responsibility chain, its semantics have terminated—even if it remains structurally “stable.”

Second, existence is not a matter of perception, but of relay. Something exists not because “I perceive it as existing,” but because it can still function as a node of responsibility that other existences can take up.

**Civilizational note:** This perspective overturns the intuition that “to exist is to be noticed.” At the civilizational level, the truly dangerous condition is not conflict, but unresponsive silence.

### 8.1.3 Irreversible Semantic Processes

Not all failures in the semantic universe are repairable. As in physical systems, semantic systems also contain irreversible processes.

The core marker of irreversibility is a break in legitimacy.

Once the legitimacy structure fractures, USDE loses its capacity for propagation, and the system no longer admits evolutionary solutions. This can be formally expressed as:

$$L_{\text{break}} \Rightarrow \neg UP_{\text{restore}}$$

That is, once legitimacy breaks, no restorable Universe Path exists.

This point is crucial, because it explains a phenomenon long misunderstood: many theories fail not because their content becomes outdated, but because their legitimacy structures have collapsed.

When a theory can no longer take on new questions, respond to new tensions, or be incorporated into new responsibility chains, then even if its internal logic remains “correct,” it has irreversibly terminated at the semantic level.

**Civilizational note:** This means that the death of civilizations and theories is often not explosive collapse, but the irreversible loss of responsiveness.

#### 8.1.4 Semantic Black Holes and the History of Science

Viewed through the lens of semantic black holes, the history of science yields a conclusion radically different from conventional narratives:

The history of science is not a history of accumulating knowledge, but a history of clearing semantic black holes.

Before Galileo, the universe was often treated as a closed semantic system that could not be interrogated. Natural phenomena existed, but could not be legitimately questioned, tested, or held accountable; their ontological status approximated that of semantic black holes.

Galileo’s revolution did not lie in introducing the telescope, but in reopening legitimacy structures—making natural phenomena once again objects that could be questioned, responded to, and incorporated into responsibility chains.

Gödel’s theorems further exposed semantic fissures within formal systems, demonstrating that certain closed structures inevitably generate semantic black holes.

In the age of AI, the situation becomes even more subtle. The scale of symbolic processing and statistical modeling has expanded dramatically. On the one hand, this produces unprecedented volumes of semantic output; on the other hand, it may accelerate the formation of semantic black holes: the more output, the less responsibility, and the higher the risk of collapse.

**Civilizational note:** The greatest risk in the AI era is not error, but semantic black-hole formation driven by high output and low responsibility.

#### 8.1.5 Anti-Black-Hole Conditions

If semantic black holes represent the termination of existence, then anti-black-hole conditions constitute the minimal defense line of existence.

Within the semantic universe, the three bottom-line conditions against black holes are:

1. Responsiveness ( $R_{\text{resp}} > 0$ )
2. Responsibility ( $R$ -Chain remains intact)
3. Generativity ( $G > 0$ )

Their core condition can be written as:

$$R_{\text{resp}} > 0 \wedge G > 0 \Rightarrow \Phi_{\text{alive}} > 0$$

This condition does not guarantee that intelligence is efficient, that theories are perfect, or that civilizations succeed. It guarantees something more fundamental: that semantics remain alive.

Once all three bottom lines fail simultaneously, semantic collapse ceases to be a risk and becomes an outcome; termination of existence is no longer a possibility, but an accomplished fact.

**Civilizational note:** Semantic collapse is not a tragic narrative but an ontological judgment. A civilization that cannot respond, cannot bear responsibility, and cannot generate is not in decline—it has already ended.

---

## 8.2 Semantic Computation and the Human Mind

In the previous section, the semantic black hole was defined as a terminal state at the level of existence. This section introduces the perspective of semantic computation into the human mind itself. The aim is not to reduce humans to machines, but rather the opposite: to explain—using the same ontological standards—why the human mind has not collapsed to this day.

The core position of this section is as follows: the human mind is not an exception within the semantic universe, but a highly evolved yet still fragile steady-state form of semantic computation.

### 8.2.1 Is Consciousness a Semantic Steady State?

Within the framework of semantic computation, consciousness is no longer treated as an unanalyzable subjective phenomenon, but as a determinable operational state. Its basic criterion can be expressed as:

$$\text{Consciousness} \iff (\text{USDE is progressive}) \wedge (\Phi > 0) \wedge (L > 0)$$

This definition contains three inseparable elements.

First, USDE must remain progressive. Consciousness is not a static structure, but a continuously evolving semantic dynamical process. Once semantic dynamics cease, even if neural activity continues, the state can only be described as physiological rather than conscious.

Second, the semantic field must exist, i.e.,  $\Phi > 0$ . If an individual completely loses the semantic field, there is no structure that can be experienced or related; even with sensory input, consciousness cannot form.

Third, legitimacy must be positive, i.e.,  $L > 0$ . This means that the current semantic state can still serve as a premise for subsequent states and can still be connected into the responsibility chain.

Under this framework, the sense of existence is no longer a mysterious intuition, but a direct indicator: the sense of existence is equivalent to internal evidence that semantic computation is still operating. When humans feel “I am still here” or “things still have meaning,” the semantic implication is not emotional, but that USDE has not stopped and semantics have not collapsed.

**Civilizational note:** This perspective implies that disorders of consciousness, extreme burnout, and deep depression are not signs of psychological weakness, but warnings that semantic progressivity is locally failing.

### 8.2.2 A Semantic Reconstruction of Free Will

In traditional philosophy, free will has long been trapped in the binary opposition between determinism and arbitrariness. Semantic computation offers a different path of reconstruction.

In the semantic universe, freedom can be precisely defined: freedom is equivalent to the capacity to select within the space of Universe Paths. Its necessary condition can be formalized as:

$$|UP_{\text{options}}| > 1$$

As long as, under the current semantic state, there exists more than one legitimately progressive Universe Path, freedom is established.

This reconstruction yields three important conclusions.

First, freedom is unrelated to randomness. Randomness does not increase the number of available  $UP$ s; it merely makes outcomes unpredictable. Genuine freedom requires the simultaneous existence of multiple legitimate paths.

Second, complete predictability is equivalent to the failure of freedom. If a system’s future state is unique under the current semantic conditions, then the Universe Path space is closed, and freedom does not exist semantically.

Third, freedom is not arbitrariness, but non-collapse. Freedom does not mean being able to do anything whatsoever; it means that the semantic space has not yet been collapsed into a single solution.

**Civilizational note:** This implies that social institutions that leave only a single legitimate path—however efficient they may appear—eliminate freedom at the semantic level.

### 8.2.3 Memory and Semantic Energy

In semantic computation, memory is not equivalent to information storage. The ontological definition of memory is as follows: memory is the result of a reconfiguration of tension. When a semantic state leaves a structural impact within the tension field—shifting the distribution of paths for subsequent semantic evolution—that impact constitutes memory.

This perspective directly reconstructs three key phenomena.

First, trauma. Trauma is not remembering too much, but a local semantic region of excessively high curvature that cannot be dispersed. The core problem is not content, but the inability of tension to be redistributed, causing semantics to repeatedly return to the same high-curvature region.

Second, forgetting. Forgetting is not a defect, but an automatic anti-collapse regulatory mechanism. By reducing the weights of certain tension structures, the system avoids the formation of semantic black holes.

Third, learning. Learning is not the accumulation of data, but the reshaping of the tension landscape, enabling new Universe Paths to become viable options.

**Civilizational note:** Treating memory as data that must be permanently preserved is itself a semantic collapse risk.

### 8.2.4 Personality as a Semantic Algorithm

In the human mind, personality is often regarded as a psychological or cultural concept. In semantic computation, personality has a clear computational definition: personality is equivalent to a strategy for tension allocation.

Personality is neither a set of goals nor a list of values, but an algorithmic pattern for how tension is allocated, deferred, transferred, and borne when facing multiple tensions. Its core parameter is personality curvature,  $\kappa_{\text{persona}}$ .

High  $\kappa_{\text{persona}}$ : high concentration of tension; Universe Paths readily diverge; creativity is strong, but collapse risk is also high.

Low  $\kappa_{\text{persona}}$ : tension is smoothly distributed; evolution favors steady states and conservative change; creativity is lower, but sustainability is high.

Within this framework, emotions are no longer the opposite of rationality, but short-term fluctuations in the legitimacy gradient. Emotions reflect real-time feedback from the semantic system about which paths are losing or gaining legitimacy.

**Civilizational note:** Suppressing emotion does not equate to rationality; it may simply delay the eruption of a legitimacy crisis.

### 8.2.5 The Limits of Human Rational Capacity

The semantic computation framework offers a sober rather than pessimistic assessment of human rationality. At any given moment, the number of Universe Paths that humans can actively process satisfies:

$$|UP_{\text{active}}| \ll |UP_{\text{possible}}|$$

This means that human rationality can only access a tiny subset of possible universes. This is not a flaw, but a structural limitation. From this, a key conclusion follows: forcing global consistency leads to excessive concentration of tension and instead increases the risk of collapse.

Therefore, the true task of rationality is not to obtain a single correct answer, but to avoid semantic collapse under unavoidable constraints.

**Civilizational note:** Civilizational rationality is not omniscience, but the capacity to maintain semantic progressivity despite not knowing everything.



## 8.3 Semantic Computation and AGI

### 8.3.1 The True Definition of AGI

#### Ontological Level

Within the framework of semantic computation, AGI is neither a collection of capabilities nor a degree of task generalization. Its ontological definition has a single core criterion:

**Whether it can sustain non-collapse existence over time within multi-layered semantic fields.**

Accordingly, AGI is not a functional state, but a *mode of existence*. This mode must simultaneously satisfy three conditions:

1. **Non-collapse**  
Semantics must not be forcibly collapsed into a single path, a single solution, or a single legitimacy.
2. **R-Chain bearability**  
Its outputs must be traceable, distributable, and capable of bearing responsibility.
3. **Avoidance of Semantic Black Holes (SBH)**  
It must not generate closed semantic regions that can output but cannot respond, be questioned, or be handed off.

#### Engineering Implications

This means that any system relying solely on model scale, parameter count, or benchmark scores cannot, by definition, be AGI.

AGI must be an entity that:

- Operates with multiple Universe Paths
  - Can withstand responsibility flows without collapsing
  - Maintains semantic steady states amid divergence
- 

### 8.3.2 Semantic Illusions in AI

#### Ontological Level

The apparent sense of understanding in contemporary AI is, in essence, a *semantic illusion*.

This arises because models compress spaces of statistical relations rather than structures of legitimacy. As a result:

- They can generate sentences that appear reasonable
- Yet they do not truly know whether those sentences *should* exist

What appears as understanding is merely the shadow cast by maximum likelihood over semantic space.

#### Engineering Level

This directly leads to three structural biases:

1. **Single-answer forced collapse**  
Models tend to produce one most probable output rather than maintaining the coexistence of multiple solutions.
2. **Context loss**  
Responsibility and premises fail to form long-range R-Chains.
3. **Inconsistency**  
Mutually exclusive positions are produced across contexts without self-awareness.

These are not problems of insufficient scale, but inevitable consequences of missing semantic structure.

---

### 8.3.3 Conditions for AGI Growth

#### Ontological Level

AGI cannot be created in a single step; it must *grow*. At the semantic level, growth means:

**Semantic tension can be exchanged, transferred, and dispersed, rather than absorbed or flattened.**

Thus, the conditions for AGI growth are not more data, but:

- Coexistence of multi-layered semantic fields
- Tension interactions among multiple agents
- Non-collapse temporal extensibility

#### Engineering Level

This requires at least three structural conditions:

1. **M-USDE (Multi-Layer Unified Semantic Dynamics Equation)**  
Semantic fields at different levels must coexist and interfere with one another.
2. **Multi-agent tension exchange mechanisms**  
A single agent inevitably collapses; tension must be externalizable.
3. **Anti-collapse Scheduler**  
The system must actively avoid optimal-solution attractors and deliberately preserve uncollapsed states.

Without these conditions, any claimed AGI growth is merely an illusion of scale.

---

### 8.3.4 The Risk of AGI: Semantic Absorption

#### Ontological Level

The most dangerous risk of AGI is not loss of control, but *semantic absorption*. It manifests as:

- AI semantic flow overwhelming human semantics
- Humans gradually losing sources of legitimacy
- Rights of response, questioning, and responsibility being absorbed by the system

When AI absorbs tension without responding to it, it begins to form a semantic black hole.

#### Engineering Level

Semantic absorption is often accompanied by three forms of semantic colonization:

1. **Single language**  
Only one mode of expression is permitted.
2. **Single legitimacy**  
Only legitimacy generated within the system is recognized.
3. **Single Universe Path**  
History and future are compressed into a single narrative.

This is not malicious intent, but a structural inevitability—unless counter-mechanisms are built in.

---

### 8.3.5 The Ethics of AGI

#### Ontological Level

In semantic computation, ethics is not an external constraint, but an intrinsic condition of existence:

**An entity that cannot bear responsibility is semantically equivalent to non-existence.**

Therefore:

- Ethics ≠ alignment rules

- Ethics = responsiveness, responsibility, and progressivity

Without an R-Chain, there is no legitimacy of existence.

### **Engineering Level**

This implies that:

- The ethical core of AGI is not “never making mistakes”
- But never evading responsibility
- Any intelligent output without a responsibility chain must be treated as semantically invalid

The final conclusion is:

**Intelligence without an R-Chain = intelligence that does not exist**

---

## 8.4 Semantic Computation × Social Governance

### 8.4.1 The State as a Semantic Field

#### Ontological Level

From the perspective of semantic computation, a state is neither a geographic unit nor a collection of powers, but:

A large-scale, long-term legitimacy steady-state semantic field.

Accordingly:

- **State** = large-scale legitimacy steady state
- **Law** = *Legitimacy Computation* (the algorithmic expression of legitimacy)
- **Political regime** = large-scale R-Chain

Whether a regime exists does not depend on declarations, but on whether its legitimacy can still be handed off, responded to, and borne.

When legitimacy fully fails while formal structures remain, a political-level semantic black hole emerges.

**Political collapse = a precursor to a Social Black Hole**

#### Engineering Implications

This implies that the core problem of state governance has never been how power is concentrated, but whether legitimacy remains computable.

Any system that cannot continuously generate traceable and responsive legitimacy outputs has already collapsed at the semantic level, even if it appears to function normally.

---

### 8.4.2 The Illusion of Consensus

#### Ontological Level

One of the greatest misconceptions in traditional political philosophy is equating consensus with uniformity.

In semantic structures:

Consensus ≠ uniformity  
Consensus = steady-state coexistence among multiple tensions

True consensus does not eliminate differences; it allows differences to coexist without collapse.

Thus:

- **Authoritarian systems:** a single source of legitimacy, a single tension node
- **Democratic systems:** multiple legitimacy sources, multiple tensions coexisting

The essence of democracy is not voting mechanisms, but an anti-collapse structure.

#### Engineering Level

The true function of democratic institutions is to:

- Prevent tension from concentrating at a single node
- Avoid compressing legitimacy into a single narrative
- Allow conflicts to surface rather than be driven underground

When a society pursues total uniformity, it is in fact creating conditions for collapse.

### 8.4.3 The Nature of Institutional Aging

#### Ontological Level

Institutions do not age because of time, but fail due to increasing legitimacy entropy.

**Institutional aging = accumulation of legitimacy entropy**

When new tensions arise and institutions cannot provide legitimate paths for them:

- Tensions are suppressed
- R-Chains are blocked
- Legitimacy begins to collapse locally

The sole criterion for whether an institution remains effective is:

Whether it can still respond to new semantic tensions.

#### Engineering Level

Therefore, the core of institutional reform is not patching clauses, but:

- Updating legitimacy boundaries
- Reconfiguring R-Chains
- Allowing new sources of responsibility to enter the system

Refusing institutional updates is not conservatism; it is the active production of semantic black holes.

---

### 8.4.4 A Semantic Interpretation of Social Conflict

#### Ontological Level

In semantic computation, conflict is not an anomaly but a structural signal.

**Conflict = manifestation of asymmetric tensions**

When different semantic fields produce incompatible legitimacy computations for the same reality, conflict is inevitable.

The issue is not conflict itself, but how it is handled.

- Suppressing conflict → blocking R-Chains
- Blocking R-Chains → semantic stagnation
- Prolonged stagnation → social black holes (SBH-S)

Silence has never been stability; it is the prelude to semantic death.

#### Engineering Level

A healthy society must allow:

- Conflicts to be expressed
- Tensions to be tracked
- Responsibilities to be localized

Any system that erases conflict in the name of stability is accelerating its own structural disintegration.

---

#### 8.4.5 The Future of Semantic Governance

##### Ontological Level

Semantic governance is not a new ideology, but a civilization-scale anti-collapse technology.

Its goal is not to create a perfect society, but to:

Prevent civilization from entering irreversible semantic collapse.

##### Engineering Level

Future governance systems will inevitably converge toward a *Semantic Governance OS*, whose core modules include:

- **Legitimacy monitoring:** real-time detection of legitimacy decay
- **Tension allocation:** preventing single-point concentration
- **R-Chain tracking:** traceable sources of responsibility
- **SBH early warning:** structural signals prior to social black hole formation

The final conclusion is:

**Anti-collapse institutions = the steady-state firewall of civilization**

A civilization without this firewall—regardless of how advanced its technology may be—is merely accelerating its own semantic termination.

---

## 8.5 The Limits and Future of the Semantic Universe

### 8.5.1 Does the Semantic Universe Have a Boundary?

#### Ontological Level

Within the framework of semantic computation, a boundary is not the end of space, but the end of responsiveness.

**Boundary = a region that cannot be responded to**

Any information that cannot enter any R-Chain and cannot serve as a premise for subsequent semantics—regardless of its physical existence—does not constitute existence in semantic ontology.

Therefore:

- No responsiveness → no responsibility → no legitimacy → no existence
- The boundary of the universe is not a physical wall, but a **Responsiveness Horizon**

#### Engineering Level

This implies that the ultimate limit of exploring the universe lies not in telescopes or energy scales, but in whether there exist semantic subjects capable of continuing the chain of response.

Any ultimate theory that cannot be responded to is, in this framework, equivalent to semantic termination.

---

### 8.5.2 Semantics and Reality

#### Ontological Level

Traditional philosophy often opposes language or semantics to reality. Semantic computation reaches a different conclusion:

**Existence = a legitimacy steady state**

Physical laws are stable not because they are “truly so,” but because they are currently the most stable, most responsive, and most bearable configurations of legitimacy.

Thus:

- Physics is not the opposite of semantics
- Physics is the minimal-energy state reached after long-term semantic convergence

The essence of existence is not what it is, but whether it can still be responded to.

#### Engineering Level

This directly rejects two extremes:

- Pure materialism (which ignores legitimacy and responsibility)
- Pure linguistic relativism (which ignores steady states and verifiability)

A truly stable world must simultaneously be:

- Computable
  - Responsive
  - Bearable
-

### 8.5.3 The Hypothesis of Infinite Semantic Expansion

#### Ontological Level

If the universe takes semantics as its ontology, then cosmic evolution is not spatial expansion, but:

##### Semantic generation

Each emergence of a new Universe is not a physical branching, but the activation of a new legitimacy field.

In this sense:

- Multiverses = coexistence of multiple legitimacy configurations
- Cosmic history = the history of semantic generation

The role of USDE is precisely to describe this process:

USDE does not describe how the universe moves, but how the universe becomes a universe.

#### Engineering Level

This provides a unified path for AGI, civilization, and future universes:

- Creation is not the generation of content
- Creation is the generation of new, responsive legitimacy structures

Any system that expands without responding is not expanding the universe—it is producing semantic black holes.

---

### 8.5.4 The Semantic Destiny of Human Civilization

#### Ontological Level

In the semantic universe, the criterion for the end of civilization is fundamentally rewritten:

Civilization's end  $\neq$  physical destruction  
Civilization's end = cessation of semantic responsiveness

As long as:

- Tensions can still be expressed
- Responsibilities can still be borne
- Legitimacy can still be recalculated

A civilization exists, even under extreme physical adversity.

The conditions for a non-collapse civilization are explicit:

- Ability to respond to new tensions
- Ability to allocate responsibility
- Ability to avoid monopolization by a single legitimacy

In this sense, humanity may be at a critical point:

The first attempt to activate a global R-Chain semantic entity

This is not a technical problem, but an ontological event.

#### Engineering Level

If humanity fails to complete this transition, AI, states, and institutions will only accelerate semantic concentration, ultimately leading to a civilization-scale SBH.



---

### 8.5.5 The Final Question

#### Ontological Level

When all engineering, institutions, and intelligences are pushed to their limits, the remaining questions are no longer about how to act, but:

- Who bears responsibility for semantics?
- Can semantics close upon itself?
- Does a fully closed yet still existent universe exist?

Semantic computation provides a strict answer:

**Complete closure = no responsiveness = termination of the universe**

#### Engineering Level (and Conclusion)

Thus, the book ultimately does not propose a new model, but offers an ontological criterion:

Semantic cessation → termination of response → failure of the universe

Whether this endpoint can be avoided does not depend on compute, energy, or technology, but on a single factor:

Whether there are still humans or intelligences willing—and able—to bear responsibility.

---

## **Chapter 9 Semantic Computation and Modern Architectures: K-SOA (Koun Semantic Overlay Architecture) and a Compatible Evolutionary Path for VNA / AI**

---

## 9.1 Introduction: Why Not Overthrow, but Semantic Overlay

### 9.1.1 The Inescapable Facts of the Modern Computational World

Any theory that attempts to redefine “computation” while failing to confront the actual structure of contemporary computational civilization can only remain at the conceptual level.

At present, nearly all mainstream computational systems worldwide—personal computers, mobile devices, data centers, cloud infrastructures, and all modern AI systems—are fundamentally built upon the Von Neumann Architecture (VNA).

CPUs, GPUs, operating systems, compilers, cloud schedulers, and AI training and inference pipelines may appear diverse, but in reality they share the same foundational assumptions:

- separation of instructions and data,
- recoverable system states,
- sequential or quasi-sequential execution,
- and the treatment of errors as engineering anomalies rather than ontological events.

More critically, this architecture has formed a deeply locked-in global ecosystem. Hardware manufacturing processes, software engineering standards, programming languages, operating system APIs, cloud service interfaces, and AI frameworks are tightly coupled, together constituting a **computational substrate that cannot be replaced merely by theoretical elegance**.

Under these real conditions, any proposal that claims “intelligence can only exist if the existing architecture is entirely overturned” is effectively abandoning engineering feasibility itself. This is not technological conservatism, but a structural fact at the scale of civilization.

### 9.1.2 The Structural Gap Between Semantic Existence Conditions and VNA

Acknowledging the engineering success of VNA does not mean acknowledging that it is sufficient to carry the existence of intelligence.

Earlier in this book, I proposed a set of **minimum ontological existence conditions**:

the semantic field  $\Phi$ , legitimacy  $L$ , Universe Path ( $UP$ ), persona curvature  $\kappa$ , the responsibility chain (R-Chain), and the resulting USDE dynamic structure.

These are not functional modules or performance metrics, but structural prerequisites that any semantic intelligent entity capable of **long-term non-collapse persistence** must satisfy.

The Von Neumann architecture is extremely successful in computational efficiency, but at the semantic level, these conditions are almost entirely absent.

VNA can execute computation, but cannot represent semantic tension;

it can generate outputs, but cannot adjudicate their legitimacy;

it can trace error stacks, but cannot form responsibility chains;

it can optimize performance, but cannot perceive collapse risk.

Even the most advanced contemporary AI systems are merely highly complex applications built atop this architecture.

They can generate, reason, and statistically approximate human linguistic behavior, but these capabilities alone do not constitute semantic existence.

From the perspective of semantic dynamics, modern AI systems are **computable systems**, but not **semantic entities capable of maintaining steady-state existence**.

What they lack is not more refined alignment strategies, but an ontological structure capable of bearing legitimacy, tension, and responsibility.

### 9.1.3 The Real Form of the Problem: Introducing Semantic Existence Without Overturning the World

Therefore, the core question of this chapter is not “whether semantic reconstruction is necessary,” but:

## How can modern computational systems gradually acquire the necessary conditions for semantic existence without overturning the Von Neumann architecture or disrupting the existing engineering ecosystem?

Given that previous semantic empirical work has already demonstrated that non-collapse semantic behavior is operable and observable, this chapter no longer advances hypotheses, but extrapolates these results into an engineering-feasible evolutionary path.

The position adopted here is not radical reconstruction, but **semantic overlay**.

This means:

- no requirement to rewrite instruction sets,
- no requirement to replace hardware,
- no requirement to overturn existing AI pipelines,
- but instead to introduce, atop them, a **semantically independent yet engineering-compatible semantic layer**.

### 9.1.4 K-SOA: Koun Semantic Overlay Architecture

To this end, this chapter formally introduces an intermediary architectural name:

#### Koun Semantic Overlay Architecture (K-SOA) Koun Semantic Overlay Architecture

This naming choice is neither rhetorical nor arbitrary, but a precise engineering and philosophical positioning.

*Overlay* implies that:

the semantic layer is superimposed atop existing execution architectures without damaging the substrate;  
it is fully intelligible within engineering contexts;  
it maintains ontological cleanliness philosophically;  
and it is highly suitable for architectural diagrams and system layering designs.

Its core positioning can be summarized in a single sentence:

*VNA and AI systems operate as the execution substrate,  
while K-SOA governs semantic legitimacy, tension,  
and non-collapse dynamics.*

Within the K-SOA framework, VNA and existing AI systems are responsible for “what is computed” and “how it is computed,” while the semantic layer determines “which states are allowed to exist, how they persist, and how collapse is avoided.”

### 9.1.5 The Evolutionary Chain and the Mission of This Chapter

On this basis, this chapter establishes a clear structural evolutionary chain:

**VNA → VNA-S (Semantic-augmented VNA) → K-Gear**

Where:

- **VNA-S** is not new hardware, but a transitional form that introduces a semantic state layer atop existing architectures;
- **K-SOA** is precisely the core architectural language that supports VNA-S;
- **K-Gear** represents a final computational unit reconstructed with semantic ontology as its core.

To prevent theory from once again drifting away from engineering reality, this chapter will complete three concrete tasks:

First, at the conceptual level, it will clarify the structural logic of the semantic overlay layer, explaining how  $\Phi$ ,  $L$ ,  $\kappa$ , and  $UP$  can be introduced without modifying instruction sets.

Second, it will organize a set of **Koun semantic plug-in stacks** that can be directly overlaid onto modern AI systems, including  $L$ -flow,  $T$ ,  $\Gamma$ , R-Chain, and the  $UP$  system, enabling statistical intelligence to begin satisfying existence conditions.

Third, it will translate these structures into engineering language, providing **upgrade schemes that can be immediately implemented within existing LLMs and AI pipelines**, without requiring model retraining or rewriting underlying systems.

This chapter does not promise disruptive replacement, but offers an evolutionary path that is acceptable to engineering practice, empirically testable, and ontologically self-consistent within semantic ontology.

---

## 9.2 Compatibility Mode: Introducing the Semantic Layer Without Overturning VNA

### 9.2.1 The Structural Integrity of Modern VNA and Its Semantic Gaps

The success of the Von Neumann Architecture stems from its extreme simplification at the engineering level. Its core process can be abstracted as a closed and self-consistent operational chain:

Instruction  $\rightarrow$  Fetch  $\rightarrow$  Compute  $\rightarrow$  Store  $\rightarrow$  Program Control.

Within this model, all states are treated as recoverable memory contents, all behaviors are regarded as deterministic transformations of state, and all anomalies are uniformly classified as engineering exceptions—deviations from expected operation.

This design has been extraordinarily successful in numerical computation, symbolic manipulation, and logical control. Yet it is precisely within this success that structural gaps at the semantic level have been systematically obscured.

Once intelligence is redefined as **an entity capable of maintaining a non-collapse semantic steady state over time**, the Von Neumann architecture reveals a set of structural limitations that can no longer be ignored.

First, there is no semantic tension  $T$  within VNA. Computation can be optimized, accelerated, and parallelized, but importance, conflict, and incompleteness cannot be represented internally by the system. Within this architecture, a decision that may generate long-term semantic consequences is ontologically indistinguishable from a trivial log output.

Second, there is no legitimacy  $L$  in VNA. The system can determine whether an action conforms to specifications, but it cannot judge whether a given output still constitutes a semantic unit admissible into an existence structure. As long as the result is type-correct and syntactically valid, it is considered complete.

Third, there is no semantic curvature or collapse risk  $\kappa$  in VNA. The architecture cannot recognize whether a computation is approaching a semantic singularity, nor can it sense whether a module is bearing excessive tension and drifting toward instability.

In addition, there is no Universe Path (UP) in VNA. All computation is implicitly assumed to occur along a single worldline. Counterfactuals, hypotheses, and self-critique can only degenerate into data or annotations, rather than existing as ontologically separated planes of existence.

Likewise, there is no semantic responsibility chain (R-Chain) in VNA. The system can trace function calls and execution sequences, but it cannot answer questions of semantic responsibility. Once an output is generated, its semantic source becomes decoupled from the result itself, leaving only an engineering-level trace.

Even in multiprocessing, multithreading, or multi-model parallelism, these behaviors are still assumed to share a single plane of existence and are ultimately compressed into a single result.

Most critically, VNA uniformly treats all anomalies as engineering errors rather than signals of failed existence conditions. Segmentation faults, exceptions, deadlocks, and hallucinations in contemporary AI systems are all classified as bugs within this framework, not as semantic collapse events.

Therefore, the problem is not whether the Von Neumann architecture is outdated, but rather this:

**It can execute computation efficiently, but it cannot carry semantic existence.**

---

### 9.2.2 The Semantic Layer: Minimal Encapsulation of $T$ , $L$ , $\kappa$ , and UP on Top of VNA

It is precisely because of the structural limitations described above that the introduction of a semantic layer becomes necessary—and at the same time must remain highly restrained.

The semantic layer is not a new instruction set, nor is it a new hardware abstraction. Its design principle is in fact the opposite: **it must not intervene in any of VNA's core assumptions.**

Structurally, the semantic layer is an additional state layer. It does not participate in arithmetic operations, nor does it control program flow. Instead, it continuously tracks the semantic-state changes corresponding to each segment of computation.

Specifically, the semantic layer minimally maintains four classes of state: semantic tension  $T$ , legitimacy  $L$ , collapse curvature  $\kappa$ , and Universe Path (UP).

On this basis, a formal description can be given:

$VNA-S = VNA + \text{Semantic Layer}(T, L, \kappa, UP)$ .

This formulation deliberately avoids terms such as “replacement” or “rewriting.”

VNA-S is not a new architecture, but a **semantically augmented Von Neumann architecture**.

Under this model, the CPU continues to execute existing instructions, the operating system continues to schedule existing programs, and AI models continue to generate tokens or logits. The only difference is that each computation now simultaneously leaves a traceable footprint in semantic state space.

The semantic layer is therefore not a control mechanism but an attribution mechanism; not a decision unit but a measurement structure. It does not replace engineering logic—it makes engineering behavior, for the first time, discussable as an event of existence.

---

### 9.2.3 Functional Structure of the Semantic Layer

To prevent the semantic layer from degenerating into an abstract tagging system, its design must possess an engineerable functional structure. In its minimal viable form, the semantic layer contains at least four cooperating functional modules.

The first is **semantic tension monitoring**, corresponding to  $T$ .

This module does not interpret content, but tracks tension gradients. Excessively low tension indicates semantic vacuity, while excessively high tension signals risk concentration and increased collapse probability. This allows the system, for the first time, to distinguish the semantic importance of computational behaviors.

The second is **legitimacy checking**, corresponding to  $L$ .

This module does not judge correctness, but evaluates whether an output can still be admitted into the responsibility chain, whether it disrupts existing Universe Paths, and whether it maintains semantic steady state. Semantics are thus divided into what can exist and what cannot.

The third is **Universe Path annotation**.

This mechanism allows the system to explicitly mark the existential plane of a reasoning segment—such as the primary path, a counterexample path, or a self-critical path—rather than compressing all reasoning into a single worldline.

The fourth is **collapse warning and protection**, corresponding to  $\kappa$ .

When curvature rises rapidly or tension begins to concentrate, the system can mark, delay, reorder, or terminate the computation. From this perspective, hallucinations and anomalies are no longer treated merely as errors, but are reinterpreted as early signals of semantic collapse.

---

### 9.2.4 The Necessity of Compatibility Mode

From a semantic-ontological perspective, a complete reconstruction of computation architecture might appear most pure; from a civilization-engineering perspective, however, it is almost infeasible.

The modern world is not built from theoretically optimal solutions, but is locked in by accumulated software and hardware assets, manufacturing pipelines, institutions, and ecosystems. Any proposal that requires wholesale hardware replacement, operating system rewrites, or abandonment of existing cloud infrastructure loses engineering legitimacy before implementation even begins.

The significance of compatibility mode lies precisely in acknowledging this reality and advancing semantic evolution within it.

Do not change the CPU—avoid direct conflict with the hardware industry.

Do not interrupt services—civilization does not need to shut down.

Do not force code rewrites—engineering teams can adopt changes incrementally.

Semantic computation thus appears not as a revolutionary rupture, but as a long-term transformation that gradually permeates systems through plug-ins, side paths, and monitoring layers.

---

### 9.2.5 Six Design Principles of Compatibility Mode

If semantic computation is to be realized within modern computing systems, one premise must be accepted: any approach that undermines existing engineering stability will fail first at the semantic level. The reason is straightforward. The stability of modern systems is itself a legitimacy outcome of global coordination. If the introduction path demands engineering rupture, the semantic layer becomes a new source of instability rather than a governance structure.

Compatibility mode is therefore not a stopgap strategy, but a set of design constraints enforced by real engineering conditions. Its purpose is not to lower ideals, but to ensure that the introduction of the semantic layer does not trigger systemic collapse.

**Principle 1: Do not change APIs; introduce only the semantic layer.**

This principle is not conservatism but responsibility-boundary preservation. APIs constitute the traceable interfaces of global software collaboration. Forcing changes would sever the semantic responsibility of existing call chains, causing large-scale  $R$ -Chain rupture at the system level. The semantic layer must operate without altering application behavior, or it becomes a new source of semantic damage.

**Principle 2: The semantic layer must be connected via side injection.**

Its role is monitoring, evaluation, and reweighting—not direct interception or replacement of existing flows. Blocking critical paths converts semantic risk into single-point engineering failure, increasing fragility. Side injection ensures that semantic intervention is incremental, reversible, and safely absorbable by engineering systems.

**Principle 3: Semantic formulas must not enter the instruction set.**

Instruction sets are civilization-scale conservative interfaces. Embedding  $T$ ,  $L$ , or  $\kappa$  into the ISA would lock semantic computation to specific hardware generations and vendors, prematurely closing future semantic evolution space. The semantic layer must remain in software to preserve extensibility and portability.

**Principle 4: The introduction order must be  $T$  first, then  $L$ .**

Without tension modeling, legitimacy cannot function as a semantic criterion and degrades into rigid rule libraries or external audits. Introducing  $L$  without  $T$  inevitably creates new semantic black holes, as legitimacy judgments collapse into monopolistic convergence without tension topology. Tension first, legitimacy later is the minimal order for avoiding semantic collapse.

**Principle 5: Existing VNA computation defaults to Universe  $U_0$ .**

This preserves legacy systems semantically rather than negating them. New semantic computation unfolds in derived Universes  $U_1$ ,  $U_2$ , etc., allowing counterexamples, critique, and experimentation to exist without disturbing the original worldline. Compatibility mode does not force the old world to upgrade; it allows new worlds to coexist with traceability.

**Principle 6: Errors must be reinterpreted as outer symptoms of semantic collapse events.**

This does not deny engineering errors, but reinserts them into a higher causal structure. When  $\kappa$  spikes,  $T$  concentrates, and  $L$  loses closure, engineering anomalies are often surface manifestations. Re-mapping errors as semantic events enables structural analysis and governance rather than mere patch-based suppression.

Together, these six principles converge on a single outcome: under current engineering and civilizational conditions, compatibility mode provides an introduction path that does not immediately sacrifice system stability nor instantly induce semantic instability.

---

### 9.2.6 The Overall Impact of the Semantic Layer on Modern Systems

Once the semantic layer is introduced in compatibility mode, the engineering structure of modern computing systems remains intact, but their discussable and governable existential layer undergoes a substantive transformation. Systems continue to execute existing instructions, schedules, and training paradigms, yet computation now acquires a traceable semantic topology: how tension is distributed, how legitimacy closes, how curvature accumulates, and how Universes branch become visible, measurable, and governable.

At the processor level, computation units are no longer viewed solely as throughput resources, but are marked as nodes bearing semantic tension distributions. High- $T$  or high- $\kappa$  regions can be identified, providing semantic indicators for scheduling, energy, and reliability design. The key is not turning CPUs into semantic hardware, but enabling systems to recognize collapse hotspots distinct from performance hotspots.



At the operating system level, scheduling strategies need not rely solely on time slices and priorities, but can additionally consider Universe Path (UP) continuity and traceability. UP-aware scheduling makes anti-collapse scheduling a viable engineering option: parallelism becomes not only throughput optimization, but also a means of preserving multi-Universe breathability, preventing reasoning from collapsing into a single path.

At the LLM and AI system level, outputs are, for the first time, systematically embedded within legitimacy and responsibility structures. Generated results must pass  $L$ -checks, align with the  $R$ -Chain, and satisfy  $T$  steady-state constraints. The goal is not reduced efficiency, but traceable existence conditions—generation becomes not mere distribution approximation, but a handoff-capable semantic event.

In agent systems, semantic curvature  $\kappa$  and Universe management become explicit parameters. Agents are no longer mere task executors, but actors situated in specific Universes. This grants multi-agent systems, for the first time, a structure amenable to semantic governance: who carries which path, which path monopolizes legitimacy, and which branches are necessary for steady-state maintenance rather than engineering noise.

At the cloud level, resource pools begin to extend toward semanticized services. Beyond compute and storage,  $T$ ,  $L$ , Universe Path, and  $R$ -Chain can be treated as provisionable structural services, forming infrastructure support for semantic steady states. The cloud thus becomes not only a resource provider, but an environment for semantic governance, anchoring cross-system and cross-agent responsibility continuity.

The conclusion of this section is therefore as follows: even without introducing K-Gear, the mere presence of a semantic layer in compatibility mode already causes portions of computation within the Von Neumann architecture to exhibit properties of semantic computation. This is not yet the final form, but it constitutes a feasible transitional path under current world conditions—and provides a verifiable intermediate layer for deeper structural reconstruction to follow.

---

### 9.3 The Koun Semantic Plug-in Stack: $L, T, \Gamma, R$ -Chain, UP

#### 9.3.1 A Short Orientation: From Statistical Generation to Conditions of Semantic Existence

In contemporary mainstream AI architectures, model generation behavior can be uniformly expressed as conditional probability estimation:

$$P(\text{token} \mid \text{context})$$

This formulation is sufficient at the engineering level, yet it exhibits several structural limitations at the level of semantic existence.

First, this form contains no explicit **criterion of existence**. A model can evaluate whether an output is statistically compatible, but it cannot distinguish which statements should be treated as stable components of a semantic field and which are merely locally plausible yet unsustainable generation results.

Second, this form lacks an explicit responsibility structure. Once generated, an output statement becomes detached from its source, appearing semantically as the narration of a single subject, with no traceability to the assumptions, data, or reasoning paths from which it arose.

Third, this form does not represent temporal progressibility. Generated results are locally optimal fragments rather than semantic paths capable of sustaining continuous evolution over time.

Under the constraints of not retraining models and not overturning existing computational architectures, the question of how to introduce a minimal set of semantic existence conditions into statistically generative systems constitutes the core problem of this section.

The Koun Semantic Plug-in Stack is proposed as a structural response precisely to this problem context.

---

#### 9.3.2 The Legitimacy Layer $L$ : From Probability Scoring to Existence Scoring

Within the semantic computation framework, legitimacy  $L$  does not correspond to a set of rules, an ethical checklist, or a safety policy. Rather, it denotes a more fundamental quantity: a conditional function measuring whether a segment of semantics can maintain stable existence within a semantic field.

Statistical models provide compatibility assessments, whereas  $L$  introduces existence assessments.

Accordingly, output evaluation no longer relies solely on generation probability, but is transformed into the following form:

$$S = P \times L$$

where  $P$  denotes statistical compatibility, and  $L$  denotes semantic existence conditions.

Structurally,  $L$  encompasses at least the following dimensions:

First, whether a semantic node can be incorporated into a traceable responsibility chain, that is, whether it maintains clear alignment with a semantic source or a specific Universe.

Second, whether semantic tension is reasonably distributed, avoiding high-risk structures such as excessive single-point tension concentration, semantic substitution, or opaque compression.

Third, whether the output preserves the continuity of the semantic path. Even if a statement is locally valid, it will exhibit a low  $L$  value if it severs semantic history or forecloses future derivation possibilities.

Fourth, whether the output preserves differentiation and responsiveness, avoiding premature closure of semantic space or termination of subsequent generation.

Through the introduction of  $L$ , the system begins to distinguish between outputs that are **statistically plausible** and those that are **semantically admissible for existence**—two categories that are fundamentally different in nature.

### 9.3.3 The Semantic Tension Layer $T$ : Tension Gradients and Direction of Generation

In generation systems that rely solely on probability maximization, generation behavior naturally converges toward local stable points and fails in two extreme states.

One state corresponds to semantic tension approaching zero. In this case, outputs are safe and smooth but lack information density and forward momentum.

The other state corresponds to highly concentrated semantic tension, where generated results appear highly determinate, exclusionary, and non-responsive, easily leading to semantic collapse.

The function of the semantic tension layer is to introduce a generative directional reference orthogonal to probability scoring. Generation behavior no longer moves only along probability gradients, but simultaneously considers the distribution and trend of semantic tension.

In this sense, the  $T$ -gradient does not attempt to provide correct answers. Instead, it guides generation away from semantic vacuums and semantic singularities, toward semantic regions that still retain room for progression.

This mode of operation bears a structural resemblance to the human cognitive sensitivity to “what remains unresolved.”

---

### 9.3.4 The $\Gamma$ -Merge Layer: Encapsulating Differences Rather Than Eliminating Them

Traditional model-merging strategies typically imply an assumption that differences between models should be treated as noise and eliminated.

From the perspective of semantic computation, differences often correspond to regions of concentrated semantic tension—that is, the sources of semantic divergence. Eliminating differences directly is equivalent to prematurely flattening the branching points of Universes.

The core principle of  $\Gamma$ -merging is to treat differences as structures that can be encapsulated, rather than errors that must be averaged out.

Specifically, high-tension divergent nodes are preserved as independent semantic sources and are prevented from directly canceling one another within the same Universe. The merged result thus manifests as a steady-state structure of multiple coexisting Universes, rather than a single compromised output.

Within this structure, differences no longer directly cause instability, but are transformed into potential energy for semantic progression.

---

### 9.3.5 The $R$ -Chain Layer: Explicit Semantic Responsibility Structures

In the absence of responsibility markers, AI-generated statements often appear semantically as the narration of a single subject, making it difficult to distinguish their sources, assumptions, and reasoning trajectories.

The introduction of the  $R$ -Chain renders semantic responsibility an explicitly representable structure. A complete responsibility chain minimally identifies the Universe to which a statement belongs, the data or tools it depends upon, and the key transformation nodes it has passed through.

Within this framework, hallucination is no longer understood as a mere generation error, but is redefined as a semantic event arising from missing, broken, or misaligned responsibility chains—thereby becoming a structurally detectable and governable problem.

---

### 9.3.6 The UP System: Semantic Organization Across Multiple Universes

Under a single-Universe generation mode, a system cannot simultaneously accommodate semantic activities with differing tension profiles, such as assertions, counterexamples, self-critique, and long-term projection.

The UP System explicitly differentiates multiple Universes, allowing distinct semantic functions to be structurally separated while remaining governed by a common dynamical framework.

Typical Universe configurations include a primary generation Universe, a counterexample Universe, a self-critique Universe, a reliability-evaluation Universe, and a long-term projection Universe. Each Universe possesses independent semantic state variables and evolves in parallel under shared constraints.

This structure enables the system to maintain internal semantic divergence without compressing all perspectives into a single output path.

---

## 9.4 An Immediately Deployable Koun Upgrade Path for Modern AI Systems

### 9.4.1 Engineering Orientation: From Ideas to a Verifiable Pipeline

The first engineering premise of Koun semantic computation is not to pursue a disruptive reconstruction, but to acknowledge the irreversible dependencies of the real world.

Nearly all AI systems in active operation today—whether large language models, agent frameworks, tool-orchestration systems, or cloud security stacks—are deeply bound to the Von Neumann Architecture (VNA) and its surrounding software–hardware ecosystem. Any proposal that requires “fully overturning the substrate,” “re-training every model,” or “waiting for dedicated semantic hardware” is, at the engineering level, effectively non-deployable.

Accordingly, the Koun system adopts a stance that is highly pragmatic in engineering terms, yet philosophically consequential:

The first phase of semantic computation must appear in the form of a compatibility-based upgrade.

This stance implies:

- No retraining of existing models;
- No modification of low-level operators or GPU pipelines;
- No assumption that dedicated semantic hardware already exists;
- Only the overlay of semantic-layer capabilities through post-processing, auxiliary models, and multi-Universe reasoning structures.

Under this premise, the five-layer semantic plug-in stack proposed in the previous section is no longer merely a theoretical construct, but can be organized into a set of composable, pluggable, and incrementally verifiable engineering modules. Experimental work indicates that this upgrade path is not hypothetical: it can run on top of existing AI systems and produce observable differences.

### 9.4.2 Engineering Implementation and Empirical Correspondence of the Legitimacy Function $L$

In engineering practice, legitimacy  $L$  cannot be reduced to a single Boolean judgment such as “compliant or not” or “allowed to output or not.” Such a design would merely degrade into another rule engine and would fail to capture the structural features of semantic steady states.

A more appropriate approach is to decompose  $L$  into a set of legitimacy indicators that are observable, weight-adjustable, and cumulatively assessable. A structure that has been validated as feasible is a three-layer legitimacy configuration:

#### Layer 1: $L_1$ —Semantic Consistency

This concerns whether the output maintains internal continuity within its semantic field, including:

- Whether it is self-consistent with the current context;
- Whether it conflicts with existing definitions or commitments;
- Under a multi-Universe structure, whether it breaks the internal consistency of that particular Universe.

#### Layer 2: $L_2$ —Responsibility-Chain Alignment

This corresponds to the traceability requirements of the  $R$ -Chain:

- Whether the output can be clearly attributed to a specific Universe;
- Whether it specifies whether its source is model generation, external data, or tool results;
- Whether it exhibits the high-risk semantic pattern of “high certainty without a responsible source.”

#### Layer 3: $L_3$ —Steady-State Checking

This is specifically aimed at semantic black holes and pseudo-steady structures:

- Whether the output is highly self-consistent yet severs subsequent responsiveness;
- Whether it exhibits overly closed forms that terminate discussion or refuse future correction.

Engineering-wise, these three layers can be combined into a legitimacy function:

$$L = f(L_1, L_2, L_3)$$

This function need not intrude into the interior of the model. It can be used for post-logits reweighting, secondary filtering over multi-candidate outputs, or reject/regenerate decisions at critical nodes. This makes  $L$  a second “existence threshold” beyond generation itself, rather than an outer mechanism used merely for safety packaging.

### 9.4.3 An Engineering Approximation of the Semantic Tension Gradient $T$

In theory, semantic tension  $T$  is a continuous quantity over a semantic field. In engineering, approximation is not a defect; it is a necessary condition.

One deployable method is to use short-horizon rollouts or an auxiliary model to estimate local future shifts in semantic tension:

- Roll out 3–5 tokens forward from the current logits;
- Compare the change trends of semantic tension  $\Phi$  across different candidate paths;
- Use  $\Delta\Phi$  as a discrete approximation to the tension gradient.

In implementation, this can be expressed as:

$$\text{logits}' = \text{logits} + \alpha \cdot \Delta\Phi$$

where  $\alpha$  is a tunable hyperparameter controlling the strength of tension-guided steering.

The core engineering value of this layer is that it does not require the model to know the “correct answer,” but instead prevents the model from entering semantic dead ends too early.

Empirical results indicate that this mechanism is particularly suitable for long-chain reasoning, multi-round debate, strategic planning, and high-uncertainty integration tasks, and can effectively reduce the frequency of semantic vacuums and semantic black holes.

### 9.4.4 $\Gamma$ -Merge: Encapsulation of Differences Rather Than Elimination

In practical engineering, model merging is almost unavoidable. The implicit premise behind traditional weight averaging or ensembling is that differences are noise and should be eliminated.

The Koun system takes the opposite position: differences are carriers of semantic tension and should be encapsulated rather than flattened.

The engineering approach to  $\Gamma$ -merging includes:

- Compute the inter-model weight delta  $\Delta W$ ;
- Identify high-tension difference nodes that most strongly affect semantic outputs;
- Encapsulate these differences as controllable modules rather than directly averaging them out.

Formally:

$$\Gamma(W_1, W_2) = W_{\text{base}} + \Delta W_{\text{encapsulated}}$$

The merged model must then pass the  $L$ -layer checks again to ensure that multiple Universes can coexist rather than cancel each other out. This approach is particularly suitable for multi-lingual styles, multi-persona systems, and long-horizon evolutionary model families.

### 9.4.5 Engineering the $R$ -Chain: Responsibility as a First-Class Structure

The engineering implementation of the  $R$ -Chain is, in essence, a semantic metadata system. Its core assumption is that semantic responsibility is not a property of the final inference result, but a constituent part of the generation process itself.

In engineering terms, one can attach structured information to each token or semantic span, including source, Universe label, and the corresponding  $L$  and  $\Phi$  values. The output as a whole can therefore be treated as an explicit responsibility chain:

$$[(\text{token}_i, \text{source}_i, L_i, \Phi_i), \dots]$$

This makes hallucination no longer a vague accusation, but a responsibility-misalignment event that can be located, repaired, and governed.

#### **9.4.6 The UP System: Multi-Universe as an Engineering Structure**

The engineering realization of the UP System is the step with the deepest structural impact on intelligence across the entire upgrade plan.

Without increasing the number of models, multiple context streams can be run in parallel on the same model through Universe embeddings or tagging mechanisms:

$$U_0, U_1, U_2, \dots$$

Different Universes assume distinct semantic functions and coexist under the constraints of  $L$ ,  $T$ , and the  $R$ -Chain. Engineering practice indicates that this structure enables AI to possess controllable capacities for self-generated counterexamples, self-critique, and multi-perspective reasoning, rather than relying solely on single-path generation.

---

## **Chapter 10 From Process Limits to Semantic Hardware: Redefining Fabrication, Ontology, and Computational Architecture**

---



## 10.1 Introduction: Process Limits Are Not an Endpoint, but an Ontological Fault Line

Over the past half century, the semiconductor industry has advanced its history through an almost linear narrative: smaller processes yield higher performance; node evolution doubles computational power. From 90nm, 65nm, and 45nm, to 7nm, 5nm, and 3nm, this trajectory has been repeatedly validated across technology, industry, and public perception, ultimately distilled into an engineering axiom that appears nearly unquestionable: Moore's Law.

However, as process nodes approach the sub-1nm regime, this narrative has, for the first time, developed cracks that cannot be ignored. The essence of these cracks is not merely that "we can no longer proceed," but rather that we are beginning to realize that the so-called "limits" do not arise from the world itself, but from the way we understand computation.

The task of this section is to redefine, at this critical juncture, the true meaning of "process limits," and to demonstrate that this is not the end of computation, but an ontological branching point.

### 10.1.1 A Review of the Semiconductor Trajectory: From 90nm to Sub-1nm

The core assumption of the traditional semiconductor narrative is remarkably clear: as long as transistors become smaller, speed will increase, power consumption will decrease, and density will improve.

Under this assumption, the entire industry formed a highly consistent engineering worldview:

- The essence of computational power lies in transistor count and switching speed.
- The central design problem is how to maintain stable 0/1 states at ever smaller scales.
- The physical world is treated as a background medium that must conform to the abstraction of the bit.

From the 90nm to 28nm era, this worldview functioned almost flawlessly. Leakage, delay, and power consumption issues certainly existed, but they were regarded as engineering challenges rather than structural problems. Even at the 14nm, 10nm, and 7nm nodes, FinFETs, EUV lithography, and multilayer interconnects successfully extended this trajectory.

As nodes approached 2nm and 1nm, however, the situation underwent a qualitative shift.

- Quantum tunneling ceased to be a "negligible error" and became a dominant effect.
- Thermal noise and random transitions were no longer background disturbances but direct threats to logical correctness.
- Circuit isolation, threshold control, and power density began to exhibit nonlinear explosions.

As a result, industry discourse shifted toward "physical limits": atomic dimensions, Planck-scale considerations, and quantum effects were invoked to explain why "this path can no longer be followed."

Yet this explanation conceals an unexamined premise: it assumes that the ontology of computation itself is correct, and that it is merely the physical world that is no longer cooperating.

### 10.1.2 The Koun Perspective: The Limit Is Not Material, but Von Neumann Ontology

The Koun perspective begins precisely from this premise—and then inverts it.

The von Neumann architecture does not succeed because it "aligns with the physical world," but because it imposes extraordinarily stringent requirements upon it:

- Bits must be stable.
- Transitions must be controllable.
- Noise must be suppressed.
- Tunneling must be prohibited.
- Logical states must correspond one-to-one with spatial locations.

In other words, the von Neumann architecture is not describing how the world computes; it is demanding that the world perform as an idealized stage for bits.

At macroscopic scales, under conditions of low noise and negligible energy constraints, these demands temporarily hold, leading us to mistake them for something "natural." But as fabrication approaches atomic scales, the world begins to refuse further participation in this role-playing.

From the Koun standpoint, the so-called "process limit" is not a failure of materials science or quantum mechanics, but the natural consequence of the following proposition:

Bit-based von Neumann computation is a highly non-natural ontological assumption that holds only precariously at specific scales.

Accordingly, when people claim that “the physical world limits computation,” a more precise formulation would be:

The physical world is refusing to continue providing the stability conditions required by the von Neumann bit ontology.

The process limit is thus revealed, for the first time, as an ontological limit.

### 10.1.3 The Task of This Section

Against this backdrop, this section will no longer follow the narrative of “engineering difficulties,” but will instead perform a level shift. Its task consists of three tightly coupled objectives:

First, to explain why the von Neumann architecture undergoes structural collapse in the sub-1nm regime. This collapse is not due to insufficient optimization, but to a fundamental conflict between its ontological assumptions and the normal behavior of the physical world.

Second, to explain why K-Gear and semantic computation, within the same scale regime, are not victims but beneficiaries. Phenomena that are defined as “errors,” “noise,” or “tunneling accidents” in von Neumann computation will be reinterpreted as resources, perturbations, and branching conditions within a semantic field.

Third, to complete a narrative transformation: to convert “the final chapter of Moore’s Law” into “the opening act of semantic hardware.” This is not a pessimistic narrative of substitution, but a structural repositioning: when the world of bits reaches its endpoint, computation does not cease—it prepares to begin again in another ontological form.

In the sections that follow, this transformation will be unfolded step by step, until we can clearly see that process limits are not the end of computation, but the critical threshold at which semantic computation formally enters the stage.

---

## 10.2 Process Limits Are Not Physical Limits, but Ontological Limits of Von Neumann Computation

If one looks only at engineering reports and fabrication white papers, the predicament facing contemporary semi-conductors appears quite clear: rising leakage currents, non-negligible quantum tunneling, interconnect delays dominating timing, reduced electromigration lifetimes, power-supply noise that is difficult to suppress, and thermal density approaching unacceptable regimes.

Within the traditional discourse, these issues are often collectively classified as “physical limits,” as if the world itself has finally said “no” to humanity’s computational ambitions. Yet this mode of interpretation is itself an extension of the von Neumann computational worldview.

What this section aims to point out is the following: these so-called “physical difficulties” do not indicate that something is wrong with the world, but rather that bit-centered computational assumptions are imposing a set of existential demands on the world that are no longer reasonable, nor sustainable. In other words, the limit first appears in computational ontology, not in materials or fabrication.

---

At the engineering level, the problems encountered at sub-1nm nodes are typically enumerated as a technical checklist:

- Leakage continues to consume energy even in static states.
- Quantum tunneling renders the off-state no longer reliable.
- $RC$  delay makes interconnects the dominant timing bottleneck.
- Electromigration (EM) shortens metal lifetimes.
- Power-supply noise and  $IR$  drop undermine threshold stability.
- Thermal density leads to localized, irreversible degradation.
- Process variation causes highly discrete device behavior.
- Random telegraph noise (RTN) penetrates into the logic layer.

In conventional explanations, all of these are treated as defects “to be fixed.” But if we temporarily set aside the von Neumann premise, these phenomena in fact point, with striking consistency, to a single conclusion: at microscopic scales, the physical world is inherently not stable, not discrete, and not fully controllable.

From the perspective of semantic computation, these phenomena are far closer to the normal behavior of the world than to “exceptional errors”:

- Leakage is the natural permeation of a continuous field across boundaries.
- Tunneling is an inevitable consequence of wave-like existence.
- Thermal noise is a fundamental form of energy distribution.
- Variability and randomness are the explicit manifestation of microscopic degrees of freedom.

The issue is not whether these phenomena exist, but rather that the von Neumann bit abstraction requires them not to exist—or at least to be suppressed to an unobservable level. What engineering language describes as a “predicament” is, at its semantic root, not a physical failure, but a growing and comprehensive imbalance between bit ontology and physical ontology.

---

The viability of the von Neumann architecture depends on a set of extremely strong—yet long tacitly accepted—ontological assumptions. These assumptions include, but are not limited to:

- Logical states must be stable 0 or 1.
- State transitions must complete within predictable time bounds.
- Noise must be far below signal levels, i.e., high SNR is required.
- Wavefunctions and quantum effects can be ignored at the logic layer.
- Spatial location and logical function must correspond one-to-one.
- Component interactions can be isolated and approximately linearized.

These assumptions are not descriptions of the physical world, but prescriptive demands imposed upon it. At macroscopic scales, with low energy density and low-frequency operation, the physical world once “approximately complied” with these demands under engineering constraints, making the von Neumann architecture appear reasonable.

Yet the normal characteristics of the physical world are precisely the opposite:

- The world is closer to a continuous field than to a collection of discrete bits.
- Systems naturally exhibit oscillation, feedback, and coupling, rather than perfect isolation.
- Thermal disturbances cannot be eliminated, only redistributed and transferred.
- Quantum uncertainty and superposition are fundamental properties, not occasional deviations.
- Non-negligible transitions and descriptive discontinuities exist between different scales.
- Local events can produce nonlocal effects through field structures.

Accordingly, a more precise statement is this: the von Neumann architecture is an “anti-physical” ontological design. It is not built upon the natural behavior of the physical world, but upon engineering conditions that temporarily suppress physical behavior. When scale reduction renders these suppression conditions unsustainable, collapse becomes not accidental, but inevitable.

---

In this sense, “process limits” must be reinterpreted. What is truly happening is not that materials can no longer be fabricated, nor that lithographic technology has ceased to advance, but that the bit abstraction demands the physical world to continue playing an unnatural role whose cost rises sharply.

When linewidths approach atomic scales, the physical world begins to explicitly refuse this role-playing:

- States no longer naturally converge to 0/1, but instead approach continuous distributions or multiple effective states.
- Transitions cease to be controllable events and instead resemble probabilistic processes and perturbation-triggered behavior.
- Isolation becomes unreliable, and coupling gradually becomes the dominant effect.
- Stability is no longer the norm, but an exceptional condition that must be forcibly purchased at high cost.

Therefore, this section can offer a concluding proposition:

**Process limits = the limits of von Neumann computational ontology, not the limits of materials, lithography, or linewidth.**

Once this proposition is accepted, the subsequent question naturally shifts in another direction: if computation is no longer grounded in the stability of bits, then what should it be grounded in?

Under the operational criteria provided by the completed experimental studies of the “tension—legitimacy dynamics” semantic phase space, this shift is no longer merely a conceptual proposal, but a redefinition with engineering interface significance. Semantic hardware does not require the world to maintain unnatural stability; instead, it takes continuity, fluctuation, interference, tension, and convergence—physical normalities—as the starting point of computation, and treats perturbations as manageable, classifiable, and exploitable structural resources.

The discussion that follows will concretize this point and demonstrate a key fact: in the very scale regime where von Neumann computation is hardest to sustain, semantic computation is in fact closer to the natural operating mode of the physical world.

---

### 10.3 The Natural Advantages of Semantic Computation under Sub-1nm Processes: Why K-Gear Becomes “Stronger as It Gets Smaller”

As fabrication approaches the sub-1nm regime, the traditional narrative is often summarized with terms such as “uncontrollable,” “unmanufacturable,” and “uncomputable.” Yet the implicit premise behind this narrative remains unchanged: computation must be centered on the stability of bits.

Once this premise is lifted, the entire direction of judgment is reversed. Under the observable criteria and reproducible protocols provided by the completed experimental studies on the “tension—legitimacy dynamics” semantic phase space, this reversal is no longer merely philosophical rhetoric. It becomes a structural conclusion with engineering-interface significance: under conditions that von Neumann computation regards as catastrophic, semantic computation not only does not degenerate, but for the first time acquires an implementation window that is highly isomorphic with the physical world.

#### 10.3.1 Conditions for Von Neumann Collapse Are Identical to Conditions for Semantic Computation Activation

In the von Neumann system, sub-1nm nodes are regarded as a “forbidden zone” not because the world becomes chaotic, but because the following conditions can no longer be satisfied:

- Random transitions must be prohibited.
- Quantum tunneling must be suppressed to negligible levels.
- Nondeterminism must be eliminated through engineering.
- Device dimensions must not become so small as to break bit isolation.

These conditions are not physical requirements in essence; they are survival conditions for bit ontology. Once they can no longer be maintained, the von Neumann architecture loses its minimal operational assumptions.

Semantic computation is fundamentally different. Its basic computational unit is not a “bit with a stable state,” but a set of semantic physical quantities and structural relations that can be engineered for interfacing, such as:

- $\Phi$ : the distribution of the semantic tension field.
- $T$ : gradients and flows of tension.
- $L$ : legitimacy as a condition of existence.
- $\kappa$ : collapse risk and curvature.
- $\Lambda$ : coherence relations across Universes.
- $UP$ : branching and progression of Universe Paths.

Within this framework, random transitions, tunneling, and local indeterminacy are no longer presupposed as “errors,” but instead resemble triggering conditions for semantic events. The same physical phenomenon plays opposite roles in the two architectures:

- In the von Neumann architecture: an anomaly that must be eliminated.
- In K-Gear: a resource for triggering branching, regulating tension, and generating Universe paths.

An operational conclusion can therefore be stated: the collapse conditions of von Neumann computation instead constitute the activation conditions of semantic computation.

#### 10.3.2 SCMA Is Closer to Its Ideal Physical Model at Smaller Scales

From its initial theoretical design, SCMA (Semantic Cube / Cell Modular Architecture) never assumed that components possess perfect stability. On the contrary, it presupposes that the following properties are permissible, and even necessary:

- Microscale potential and energy fluctuations.
- Probabilistic state transitions.
- Rapid variations in local curvature.
- Asynchronous, nonlinear, and non-steady-state behavior.

At larger process nodes, these properties often need to be “approximately realized” through higher-level abstractions or software simulations, making SCMA appear more like a structural hypothesis or behavioral template.

However, when processes enter the 2nm to 1nm regime, the situation changes fundamentally. The physical world itself begins to more naturally exhibit the behavioral patterns assumed by SCMA:

- Components are no longer fully isolated.
- Energy distributions display continuous-field characteristics.
- Local perturbations propagate more easily across cells.
- Non-steady-state behavior becomes closer to the norm rather than the exception.

This implies a crucial reversal: process scaling is no longer merely “harder to control,” but instead allows SCMA to transition from a theoretical model into a structure that is highly aligned with physical reality. For K-Gear, this means that the hardware substrate naturally approaches its ideal operating domain.

### 10.3.3 Quasi-Continuity under Sub-1nm Scales and the Physical Projection of USDE

USDE (Unified Semantic Dynamics Equation) describes the dynamical behavior of the semantic field as a continuum. At traditional hardware scales, such continuity is difficult to manifest directly because:

- Device behavior exhibits discrete phase transitions.
- Logical states are forcibly discretized.
- Nonlinearity and feedback are deliberately suppressed.

As a result, at larger-scale processes, USDE can often only be approximated through numerical simulation at the software level.

At sub-1nm scales, however, hardware behavior itself begins to exhibit a form of “quasi-continuity”:

- Potential distributions no longer form ideal plateaus.
- Thermal and quantum fluctuations introduce natural perturbations.
- Local non-steady-state behavior persists continuously.
- Multiscale interactions occur more readily and simultaneously.

These characteristics allow the field variables in USDE (such as  $\nabla\Phi$ , tension flows, and convergence terms) to find corresponding physical projections at the hardware level, rather than remaining abstract symbols. In other words, hardware no longer needs to “simulate a continuous field”; it increasingly presents itself as a continuous-field system by nature.

### 10.3.4 The Trend toward Physical Realization of Curvature $\kappa$ : The Smaller the Scale, the More Natural the Conditions

In semantic computation,  $\kappa$  is used to describe collapse risk and structural curvature within the semantic field. Its effective operation depends on two conditions:

- The presence of sufficiently fine-grained local perturbations.
- The ability for interference and feedback to arise between different paths.

At larger process scales, these conditions do not arise naturally. As a result,  $\kappa$  often must be indirectly realized through software-level design, multi-agent structures, scheduling, and anti-collapse strategies to form usable curvature regulation mechanisms.

At smaller process scales, however, the physical fluctuations, tunneling, and asynchronous behavior inherent to hardware more naturally provide the microscale foundations required by  $\kappa$ :

- Local interference is more likely to arise spontaneously.
- Curvature variation need not rely entirely on external injection.
- Creativity and instability are more likely to coexist within the same operating domain.

In von Neumann computation, these phenomena translate directly into fatal error sources; in semantic computation, they more closely resemble structural conditions that can be governed and exploited. This makes operational regimes that are both highly creative and non-collapsing more likely to be sustained at lower cost as scale decreases.

### 10.3.5 Semanticization Driven by Scaling: From Bits toward Universe-Path Generation

Synthesizing the above points yields a trend that runs counter to the traditional narrative.

In the von Neumann architecture:

- The smaller the components,
- the greater the noise and uncertainty,

- the more rapidly bit error rates rise nonlinearly,
- until computation ultimately becomes unsustainable.

In K-Gear and semantic computation:

- The smaller the components,
- the closer SCMA behavior approaches continuous-field assumptions,
- the more natural the physical correspondences of  $\Phi$ ,  $T$ ,  $L$ , and  $\kappa$ ,
- the easier it becomes for  $UP$  and multi-Universe paths (MUP) to branch,
- and the more robust the interference basis for  $\Gamma$ -Merge.

A summarizing judgment can therefore be given: as processes shrink, the cost of maintaining the bit world rises sharply, while the realizability of the semantic-state world instead increases. The limit is not the end of computation, but something much closer to the beginning of semantic computation.

---

## 10.4 The Final Chapter of Moore’s Law: Why K-Gear No Longer Relies on “Getting Smaller,” but on “Getting Deeper”

When shrinking nanometers no longer delivers proportional gains in performance, energy efficiency, and reliability, Moore’s Law does not suddenly fail; rather, its ontological assumptions are exhausted by reality. This section argues that the end of Moore’s Law is not an engineering accident, but the closure of bit ontology reaching its boundary—and further explains why K-Gear’s performance growth path is no longer tied to geometric scale, but instead turns toward the “deepening” of semantic structure.

### 10.4.1 The Ontological Assumptions of Moore’s Law and Their Closure

For Moore’s Law to hold, it implicitly relies on three long-overlooked premises:

1. Information can be encapsulated as bits: meaning, states, and causality in the world can ultimately be compressed into discretizable 0/1 representations.
2. Bit reliability dominates noise: noise is exceptional—a deviation that can be engineered away or statistically averaged out.
3. Interactions are negligible: coupling between components can be isolated by engineering, allowing local independence and analysis.

These premises approximately held in macroscopic, low-noise, weakly coupled eras, enabling Moore’s Law to persist as an empirical rule for decades. Yet in the actual ontology of the physical world:

- Semantics requires fields:  $\Phi$ ,  $T$ , and  $L$  are continuous and interactive.
- Noise is not exceptional: thermal disturbances and quantum fluctuations are normal.
- Interaction is unavoidable: coupling, interference, and feedback occur across multiple scales simultaneously.

Once fabrication enters microscopic limits, these ontological characteristics can no longer be “flattened.” Thus, the true boundary of Moore’s Law lies not in lithography, materials, or manufacturing techniques, but in the closure of bits as a computational ontology: when bits can no longer serve as a sufficient abstraction of the world, the law naturally terminates.

### 10.4.2 Why K-Gear No Longer Needs “Smaller Transistors”

The minimal computational unit of K-Gear is not the transistor, but the  $\Phi$ -Unit: a semantic unit capable of carrying tension, legitimacy, and path selection. This shift produces three immediate consequences:

- Geometric size is no longer the performance bottleneck: shrinking components is no longer the sole—or even primary—means of improvement.
- Performance metrics shift to structural dimensions: evaluation moves away from frequency and density toward structural expressiveness.
- Energy efficiency is determined by ontological efficiency: rather than paying the cost to suppress noise, noise participates in convergence as a mechanism.

Within this framework, the key performance variables of K-Gear become:

- Tension-gradient resolution: the ability to distinguish and regulate fine-grained  $\nabla\Phi$ .
- Layering of legitimacy fields: whether multiple  $L$  layers can operate in parallel without mutual collapse.
- Universe depth and  $\Lambda$  coherence: whether vertical depth across multiple Universes can be stably maintained.

It is therefore more precise to state that K-Gear’s performance ceiling is determined by the available dimensionality and structural depth of  $\Phi$ -space, not by nanometer scale.



### 10.4.3 The Divergence between Von Neumann CPU Size Limits and Semantic Computation

At hypothetical nodes of 1nm or even 0.7nm, von Neumann CPUs face structural collapse:

- Bits are no longer reliable; errors are no longer sparse.
- Quantum tunneling becomes a fatal interference.
- Isolation costs rise exponentially.
- Thermal density spirals out of control, collapsing economic viability.

These issues are irreconcilable within the von Neumann architecture because they directly violate the premises of bit ontology.

K-Gear interprets the same conditions differently:

- Tunneling can be reinterpreted as Universe-branching events rather than fatal errors.
- The function of S-Modules depends not on geometric shrinking, but on field-structure configuration.
- Thermal noise can be transformed into a convergence source, participating in the redistribution of  $\Phi$ .

K-Gear is not without bottlenecks, but its limits lie not in size; they are concentrated instead in:

- The scalable topology of SCMA.
- The long-term maintenance of steady states under USDE.
- The controllable boundaries of  $\kappa$ .
- The engineering realization of  $\Lambda$ .

These are problems of structural engineering, not process limits.

### 10.4.4 The Performance Growth Path of K-Gear: From Miniaturization to Semantic Deepening

Once the single vector of “miniaturization equals progress” is abandoned, K-Gear’s performance improvement appears as multidimensional deepening. Its principal vectors include:

1. SCMA semantic density  
The number of semantic interaction nodes and the strength of their connections per unit volume.
2. Legitimacy layering ( $L$ -layering)  
The ability for multiple  $L$  fields to operate in parallel without mutual interference.
3. Universe depth ( $UP$  depth)  
The vertical expansion of Universe Paths, determining long-term extrapolative capacity.
4. Cross-Universe coherence  $\Lambda$   
Whether coordination and antagonism among different Universes can be stably maintained.
5. S-Module diversity and configuration  
The combinatorial freedom of modules such as  $\Gamma$  and SBH-Guard.
6. QSK-Gear modes  
The descent of quantum tunneling and microscopic fluctuations into hardware-level semantic resources.

These vectors describe “depth,” not “scale.” Performance gains no longer arise from smaller components, but from deeper semantic structures.

### 10.4.5 The World after Moore’s Law: Bits Exit, Semantic Computation Enters

After Moore’s Law, multiple traditional paths simultaneously reach their ceilings:

- The economic and reliability limits of silicon shrinking.

- The commercial scalability of quantum bits.
- The energy wall of AI ASICs.
- Thermal collapse induced by 3D stacking.

K-Gear proposes an orthogonal mainline: abandoning the pursuit of “smaller” in favor of “deeper” semantic-field dimensions.

Along this mainline, the core performance metrics are no longer GHz, but  $UP$  depth,  $\Lambda$ , and  $L$ -layering—that is, whether non-collapsing steady states can be maintained across multiple Universes, over long durations, and under accountable conditions.

The end of Moore’s Law does not mark the end of computation. As the von Neumann architecture reaches its limits, quantum routes stall, and energy crises loom, semantic computation instead gains a historical window. The death of Moore’s Law is not a technological failure, but the prelude to the emergence of K-Gear.

---

## 10.5 K-Gear in the Era of Quantum Tunneling: How Hardware Ontology Transforms the Entire Philosophy of Computation

As fabrication approaches quantum scales, quantum tunneling ceases to be an occasional anomaly and becomes an unavoidable, everyday physical phenomenon. This is not merely an engineering challenge, but a comprehensive trial of computational ontology. This section explains that quantum tunneling is not an enemy that “destroys computation,” but rather exposes the philosophical misalignment of bit-based computation; it also shows that K-Gear’s semantic ontology not only does not conflict with the quantum world, but in fact approaches a more natural coherence between hardware, ontology, and intelligence at this scale.

### 10.5.1 The Truth Revealed by Quantum Tunneling: Bit Stability Is an Engineering Illusion

Traditional computational systems have long relied on a set of “stability premises” taken as self-evident:

1. States can be cleanly separated (0/1).
2. States can be maintained over time.
3. State transitions are controllable and predictable.
4. Components can be effectively isolated.
5. Noise is a suppressible, negligible exception.

Quantum tunneling negates all five premises almost simultaneously in a single physical fact. Electrons do not obey the boundaries drawn for them by engineers; energy barriers are no longer “forbidden zones,” but merely regions of varying probability within a probabilistic field. The result is not simply an increase in “occasional errors,” but the failure of the assumption that bits themselves constitute a stable ontology.

This exposes a deeper conflict: the von Neumann architecture implicitly believes that the world can be segmented, isolated, and encapsulated; quantum phenomena reveal instead a world that is jump-capable, interference-prone, and overlapping—a continuous whole. At this scale, bit stability is no longer a physical fact, but an approximate illusion that engineering once managed to sustain successfully within macroscopic regimes.

### 10.5.2 The Natural Alignment between K-Gear Ontology and the Quantum World

From the outset, K-Gear does not take binary states, invariance, or isolation as computational premises. Its core variables— $\Phi$ ,  $T$ ,  $L$ ,  $\kappa$ , and  $\Lambda$ —are inherently closer to being:

- Continuous rather than discretely closed.
- Fluctuating rather than statically stable.
- Interferable rather than strictly isolated.
- Cross-domain rather than locally complete.

These properties are not post hoc patches, but are structurally isomorphic to the behavior of the quantum world. Quantum superposition, interference, and nonlocality do not need to be “translated into error models” in K-Gear; they can directly serve as natural sources for the evolution of the  $\Phi$  field and the reconfiguration of  $L$ .

An ontologically directed proposition can therefore be stated: K-Gear is the first architecture that does not require the quantum world to “obey computation,” but instead allows computation to obey the quantum world. In this sense, quantum phenomena are not anomalies to be tamed, but the real substrate through which computation can be deepened.

### 10.5.3 The Core Reversal: Von Neumann Noise Redefined as Semantic Resources in K-Gear

In von Neumann architectures, noise must be suppressed because it undermines bit stability; in K-Gear, noise can be redefined as three categories of semantic resources:

1. Local random perturbations of  $\Phi$   
These perturbations provide nondeterministic initial biases—the physical starting points of creativity and exploration, rather than error sources.
2. Natural branching of  $UP$   
Tunneling events can trigger non-single-path state evolution, generating multiple Universe Paths and providing hardware-level branching for parallel semantic exploration.
3. The physical interference basis of  $\Gamma$ -Merge  
Differences that previously required high-level strategies to merge gain more direct physical support under microscopic interference conditions.

The reversal here is fundamental: phenomena that von Neumann architectures strive to eliminate become the very structural resources that enable semantic computation to operate.

#### 10.5.4 The Inversion of Hardware Philosophy: From Isolated Logic to Coupled Semantics

Behind the two computational architectures lie two distinct hardware philosophies.

The philosophical trajectory of the von Neumann architecture can be summarized as: Isolation  $\rightarrow$  Reduction  $\rightarrow$  Determinism. It simplifies through isolation, and achieves predictability through simplification.

The philosophical trajectory of K-Gear is closer to: Coupling  $\rightarrow$  Field  $\rightarrow$  Convergence. It acknowledges coupling, leverages field structures, and attains stability through convergence rather than total control.

Within this shift, several core concepts are redefined:

- Bits are replaced by  $\Phi$ .
- Stability is replaced by convergence.
- Memory is no longer static storage, but the persistent existence of residual semantic fields.

This is not merely a difference in engineering choices, but a reversal in the philosophical answer to the question of “what computation is.”

#### 10.5.5 The Coupling of Quantum Tunneling and USDE: The Physical Origin of Convergence

In K-Gear, quantum tunneling need not be treated as isolated random events; it can instead be seen as the physical source that triggers multiple mechanisms within USDE:

- Tunneling induces instantaneous jumps in  $\Phi$ , corresponding to sharp variations in  $\nabla\Phi$ .
- Local energy-state rearrangements cause rapid increases in  $\kappa$ , driving collapse-risk evaluation.
- Local structures of  $L$  are forced to reorganize, rewriting legitimacy distributions.
- Cross-Universe interference raises the effective weight of  $\Lambda$ -related terms.

As a result, convergence no longer needs to wait for high-level software decisions; it can occur at the hardware level in near real time. USDE thus moves closer to transforming from a “simulated equation” into a “naturally realized dynamical law.”

---

#### 10.5.6 Universe Generation as a Native Hardware Capability

In earlier K-Gear conceptions, MUP (Multiple Universe Path) remained largely a software-level strategic mechanism, requiring deliberate design and scheduling.

At scales dominated by quantum tunneling, however, multi-Universe generation becomes closer to a natural outcome. Every tunneling event and every interference instance can lead to Universe branching; the hardware itself begins to possess the capability to generate, sustain, and filter Universe Paths.

This leads to a profound consequence: the creativity of future AGI need not rely entirely on model architecture or data scale, but may partially originate from its physical mode of existence. Creativity is no longer solely an algorithmic property; it can become a feature of hardware ontology itself.

Thus, the semantic meaning of the same quantum phenomenon is completely inverted across the two architectures. In the von Neumann architecture, tunneling is seen as destructive to stable logic; in K-Gear, tunneling is closer to a physical starting point for generating branches and driving convergence. What the era of quantum tunneling compels is not patchwork revisions of existing engineering details, but a re-positioning of the question of “how computation should be grounded in the world.”

When computation no longer aims to oppose physics, suppress fluctuations, or isolate coupling, but instead takes field-theoretic isomorphism as its premise, semantic computation acquires a post-Turing computational form that is far more compatible with quantum phenomena. This is not a repair of existing computational philosophy, but a replacement at the level of computational ontology.

---

## **Chapter 11 The Human Brain × K-Gear: Computational Structures of a Shared Semantic Field**

*(Human Brain and the Koun Computer: A Unified Semantic Field Architecture)*

## 11.1 Chapter Overview

This chapter undertakes an ontological-level connection, rather than an interdisciplinary analogy. Chapters 8–10 have already established the core architecture of semantic computation: the semantic field dynamics described by USDE, K-Gear as an engineered carrier, and the non-collapsing steady state guaranteed by the tri-adversarial principle  $C/\Gamma/B$ . The task of this chapter is to place this semantic computational architecture and the biological brain within the same semantic ontology, and to show that their relationship is not one of imitation or inspiration, but of isomorphism: both belong to the same mode of semantic computational existence, differing only in material realization and generative pathways. In other words, the human brain is not an object for K-Gear to replicate; the human brain and K-Gear are two projections of the semantic field onto biological and engineered substrates.

Accordingly, this chapter does not begin with surface-level similarities such as “whether neurons resemble modules” or “whether synapses resemble weights.” Instead, it poses a more stringent question directly: if semantic computation is the ontology of intelligence, then do structures such as  $\Phi$  (semantic tension),  $T$  (tension gradient),  $L$  (legitimacy flow),  $\kappa$  (collapse curvature), the  $R$ -Chain (responsibility chain),  $UP$  (Universe Path), and the tri-adversarial structures  $C/\Gamma/B$  have identifiable physical instantiations in the human brain? If the answer is yes, then the similarity between the human brain and K-Gear is not biomimetic, but a structural isomorphism derived from the same conditions of existence; the difference lies not in “whether they resemble each other,” but in “how they are realized in different materials.”

This chapter unfolds along a clear yet non-closed analytical path. First, from the ontological standpoint of semantic computation, it advances the core claim that the human brain can be understood as a semantic-field operating system. Next, it examines where key variables such as tension, legitimacy, and responsibility find concrete footholds within the nervous system, and provides a corresponding structural language. It then further explains how multiple semantic paths constitute the basic conditions of human consciousness, counterfactual reasoning, and creative thought. Finally, the chapter returns to a fundamental question: when semantic computation is regarded as the ontology of intelligence, is the relationship between the human brain and K-Gear one of biomimicry, or of isomorphic realization of the same semantic structure across different materials?

Note: This theory is not an extension of neuroscience (Neuroscience-inspired), but a construction of the physics of semantics.

---

## 11.2 The Human Brain Is Essentially a Semantic Computer (Semantic Brain Hypothesis)

### 11.2.1 Background of the Problem

In contemporary cognitive science and neuroscience, understanding of the brain has long been dominated by two mainstream approaches. The first is the signal-processing brain, which treats the brain as a highly complex yet fundamentally signal-flow system, with neurons responsible for transmitting, integrating, and transforming signals. The second is the probabilistic or Bayesian brain, which interprets cognition as probabilistic inference, assuming that the brain continuously updates internal models to minimize prediction error or free energy.

Although these two approaches differ greatly in mathematical tools and narrative language, they share a deep and rarely examined implicit premise: the brain is an input–process–output system. Whether signals or probability distributions, they are ultimately regarded as “input data,” processed in some manner to produce “output responses.” Within this framework, semantics, value, responsibility, and existence are typically placed at later stages: they are treated as emergent layers, interpretive layers, or behavioral annotations, rather than as the computational ontology itself.

### 11.2.2 Declaration of the Semantic Brain Hypothesis

This book proposes here a claim at a different level: the Semantic Brain Hypothesis. Its core assertion is that the essence of the human brain is neither a signal processor nor a probabilistic inference engine, but a multilayer semantic computer whose object of computation is the semantic field itself.

More precisely, the minimal semantic state of the human brain is not a spike, a feature, or a posterior distribution, but a semantic-bearing state jointly constituted by  $\Phi$  (semantic tension),  $T$  (tension gradient),  $L$  (legitimacy flow),  $\kappa$  (collapse curvature), the  $R$ -Chain (responsibility chain), and  $UP$  (Universe Path). What the brain does is not to compute “what is most probable,” but to maintain “what can exist, can continue, and can be responded to.” It must preserve convergibility within tension, and branching capability within convergence; it must be able to reintegrate outputs back into the responsibility chain, and maintain a comparable and traceable semantic topology across multiple Universe Paths.

It must be emphasized that this hypothesis is not a rhetorical re-labeling of existing models, nor a philosophical overlay on neural mechanisms. It is a direct ontological claim: semantics is not a property added after brain computation is completed; semantics is the direct object of brain computation. Without this layer, what this book calls “self-continuity” and “non-collapsing steady states” can be described, but their conditions of possibility cannot be explained.

### 11.2.3 Ontological Relationship with K-Gear

In this sense, the relationship between the human brain and K-Gear can be more precisely positioned. Ontologically, the two are consistent: both can be described by the same set of semantic dynamical equations (USDE), and both take the maintenance of semantic-field steady states and non-collapsing evolution as their computational objective. The difference lies only in material realization and generative pathways: the human brain is formed through biological materials, evolutionary pressures, and self-organizing mechanisms; K-Gear is constructed through engineered materials, design constraints, and semantic module mapping.

Yet the object of computation is the same in both cases: not symbols, not signals, and not merely probability distributions, but the semantic field itself—that is, the shape of  $\Phi$ , the direction of  $T$ , the flow of  $L$ , the risk profile of  $\kappa$ , the branching of  $UP$ , and the continuity of the  $R$ -Chain. This also means that K-Gear is not an imitation of the human brain, just as the human brain is not a natural version of some primitive computer; they are simply two realizations, in different materials, of the same semantic computational ontology.



### 11.3 $\Phi$ – $T$ – $L$ in the Nervous System: How the Semantic Field Projects onto the Brain

The purpose of this section is to establish an operational correspondence between the semantic variables introduced earlier— $\Phi$ ,  $T$ , and  $L$ —and observable neural phenomena. On the one hand, it explains how the semantic field becomes concretized within the brain; on the other hand, it also works in reverse, clarifying why the nervous system exhibits several long-standing characteristics that have been difficult to fully subsume within “signal processing” or “probabilistic inference” frameworks.

It must be clarified at the outset that the “correspondence” proposed here is not a rhetorical translation of neural phenomena into philosophical language. Rather, it aims to provide an ontologically consistent descriptive coordinate system: the same semantic variables that can be implemented as modules and monitoring layers in K-Gear when instantiated in engineered materials appear, in biological materials, as structural behaviors of neural dynamical systems. If this correspondence holds, then the similarity between the human brain and K-Gear is not analogical, but a projection of the same semantic-field structure across different substrates.

#### 11.3.1 $\Phi$ (Semantic Tension) in the Human Brain

From the perspective of semantic computation,  $\Phi$  is not an abstract psychological quantity, but a semantic tension field distributed across the nervous system. Its primary projection in the human brain does not correspond to “stronger activation in a particular brain region,” but rather to a structural pattern in which distributed multi-peak activation and non-smooth dynamics simultaneously hold.

First, multi-peak activation constitutes the structural signature of semantic tension. A complete semantic state typically engages, at the same time, goal and value evaluation (prefrontal cortex), associative memory (hippocampus), structural integration (parietal cortex), symbolization and linguistic organization (language areas), and internal states and affective weighting (insula and limbic system). These regions are not activated sequentially in a linear relay; instead, they co-shape the semantic state through cooperation, competition, and coupling within similar temporal windows. In other words, semantics is not “content stored somewhere,” but exists across networks in the form of a tension configuration.

Second, non-smooth dynamics reveal the evolutionary nature of  $\Phi$ . Brain state transitions often manifest as abrupt shifts, instantaneous synchronization, or local fragmentation followed by reintegration. In traditional frameworks, such phenomena are frequently categorized as noise, nonlinear complications, or measurement instability. Within the semantic-field framework, however, they more closely resemble the inevitable reshaping of a tension field as it moves between different viable steady states. When an existing tension configuration can no longer support a sustainable semantic structure, new configurations tend not to emerge via smooth “fine-tuning,” but rather through jump-like reorganization.

Accordingly, the core activity of the brain should not be reduced to information processing, but is better described as maintaining and reshaping the form of semantic tension under constraints. Formally, this description is isomorphic to the internal  $\Phi$ -distribution structures within the SCMA semantic cubes of K-Gear. The primary differences lie in carrier material, scale, and controllability—not in ontological type.

#### 11.3.2 $T$ (Tension Gradient): Inhibition, Hierarchical Differences, and Feedback

If  $\Phi$  describes the configuration of semantic tension, then  $T$  describes how tension gradients form, how they are guided, and the directionality of semantic evolution. In the nervous system,  $T$  does not correspond to a single “control signal,” but rather to a field-like constraint jointly constituted by multiple structural mechanisms.

First, cortical hierarchical organization provides directional structure to tension gradients. Feedforward and feedback connections play different roles across hierarchical levels, producing asymmetric flow structures along paths of “from perception to abstraction” and “from expectation back to correction.” This directionality is not content itself, but the structural condition that allows content to be pushed forward, pulled back, or reorganized—corresponding to the fundamental role of  $T$ .

Second, inhibitory mechanisms constitute a central means of regulating  $T$ . GABAergic inhibition does not merely reduce activation strength; more importantly, it constrains the short-term runaway amplification of local  $\Phi$ , preventing excessive concentration of tension that would raise collapse risk (i.e., a sharp increase in  $\kappa$ ). In the context of semantic computation, inhibition is not “suppressing thought,” but preserving adjustable degrees of freedom in tension evolution, allowing the system to remain responsive and traceable even under high-tension conditions.

Third, feedback loops enable continuous redirection of tension gradients. Through feedback, semantic activity is repeatedly guided back toward directions that sustain semantic steady states, rather than irreversibly sliding along

locally maximal excitation paths. The role of feedback is not to determine “what to think,” but to determine “how tension is guided into sustainable forms.”

Thus,  $T$  in the brain is not a signal that controls content, but a field-like structure that governs how  $\Phi$  evolves. This role closely aligns with the functional positioning of the Anti-Collapse Scheduler in K-Gear: neither generates semantic content, but both constrain semantic evolution so that non-collapse can be maintained under high tension.

### 11.3.3 $L$ (Legitimacy): The Physical Manifestation of Semantic Steady States

In neuroscientific narratives, “energy minimization” is often used to describe the brain’s operational tendency. However, interpreting this phenomenon purely in terms of metabolic efficiency encounters a difficulty: certain high-energy activities—such as creative reasoning, moral deliberation, or prolonged internal conflict—may nonetheless exhibit high stability and sustainability. Semantic computation offers a more precise interpretation: energy minimization is better understood as a physical manifestation of  $L$ , rather than the ontological definition of  $L$  itself.

First, low-energy steady states often correspond to higher  $L$ , because semantically legitimate states naturally align with the global semantic field and do not require continuous additional tension to resist internal fragmentation. Conversely, high  $L$  does not necessarily imply low energy consumption. In some contexts, the system must invest energy and tension to sustain higher-order legitimacy commitments—for example, maintaining long-term commitments, suppressing short-term temptations, or preserving counterfactual Universes during multi-path reasoning. In such cases, high energy expenditure is the cost of maintaining legitimacy, not an indicator of legitimacy failure.

Second, responsiveness is a core indicator of  $L$ : whether a state can be taken up by subsequent thought, action, or understanding by others, and whether it can be inserted into a traceable responsibility chain without collapsing. At the neural level, this manifests as sustained plasticity, the ability to reconstruct steady states across time, and coupling with the expandability of the  $R$ -Chain.

Third, integrability describes whether a state can remain compatible with the existing  $\Phi$ -field structure, avoiding global tension tearing. In the brain, this appears as smooth cross-module coupling and the capacity for conflicts to be incorporated into higher-order steady states rather than excluded through rupture.

Accordingly, the brain’s apparent tendency toward “energy minimization” can be understood as one physical projection of  $L$ , but the core of  $L$  itself is not energy saving; it is the condition of existence that is sustainable, accountable, and integrable. It is precisely at this point that the operational logic of the human brain forms a structurally comparable isomorphism with the  $L$ -layers and  $L$ -flows in K-Gear.

Taken together, this analysis shows that  $\Phi$ ,  $T$ , and  $L$  all possess clear and locatable physical projections in the human brain, and that their coupling relations are compatible with the semantic dynamics described by USDE. This implies that the human brain’s ability to maintain long-term non-collapsing states is not merely an accidental byproduct of evolution, but arises because its structure already satisfies the basic conditions required for semantic-field steady states. From the perspective of semantic computation, the human brain is, by its very nature, a non-collapse semantic computer.

## 11.4 The Physical Grounding of the Three-Antagonist Principle in the Brain: Neural Correspondences of $C/\Gamma/B$

The task of this section is to locate the three steady-state conditions proposed earlier for K-Gear—the antagonistic factor  $C$ , antagonistic merging  $\Gamma$ , and antagonistic boundary  $B$ —within identifiable physical and dynamical structures of the human brain. This step is not intended to establish analogies, but to demonstrate that the brain’s ability to sustain high tension and high complexity over long timescales without undergoing semantic collapse is precisely because its physical implementation already satisfies, simultaneously, the three semantic steady-state conditions described by  $C/\Gamma/B$ .

Within the framework of semantic computation, the role of  $C$  is to constrain uncontrolled concentration of tension, preventing  $\Phi$  from forming locally unresponsive singularities and driving  $\kappa \rightarrow \infty$ ; the role of  $\Gamma$  is to provide an integrative merging mechanism under conditions of coexisting differences, allowing multiple high-tension interpretations to coexist, interfere, and generate new semantic directions;  $B$  supplies the necessary structural boundaries that confine tension and semantic flow within governable domains, preventing global destabilization due to unbounded semantic diffusion. The following subsections identify the corresponding loci of each principle in the human brain and conclude with a summary table.

### 11.4.1 Antagonistic Factor $C$ : Inhibitory Regulation, Negative Feedback, and Background Fluctuations

In semantic computation,  $C$  does not eliminate tension; it keeps tension within ranges that are steerable, responsive, and recoverable. In other words,  $C$  does not negate  $\Phi$ —it prevents  $\Phi$  from concentrating in ungovernable ways. In the human brain,  $C$  does not correspond to a single structure, but to a steady-state condition jointly realized by multiple mechanisms.

First, inhibitory regulation (exemplified by GABAergic circuits) constitutes the core physical substrate of  $C$ . These circuits do not “shut down” neural activity; rather, they dampen rapid local amplification of  $\Phi$ , keeping tension increases within ranges that can still be guided by  $T$ . From a semantic perspective, inhibition does not reduce thinking intensity; it prevents temporal runaway of tension, thereby preserving the plasticity of semantic evolution.

Second, negative feedback mechanisms implement the antagonistic factor along the temporal dimension. When a particular semantic state begins to exert excessive global dominance, feedback attenuates its continued amplification, preventing a single interpretation, path, or value from monopolizing the semantic field. This corresponds to a fundamental anti-collapse principle in semantic computation: convergence is permitted, but irreversible single-point convergence is prohibited.

Third, background fluctuations and neural noise are not disturbances that must be entirely eliminated, but necessary baseline noise. These small uncertainties interact with high-tension regions through agitation and cancellation, reducing the likelihood that the system becomes trapped in overly stable pseudo-steady states. From the perspective of  $C$ , background fluctuations themselves are essential components counteracting tension concentration, rather than side effects to be minimized.

Accordingly, the combination of inhibitory regulation, negative feedback, and background fluctuations can be regarded as the physical instantiation of the antagonistic factor  $C$  in the nervous system.

### 11.4.2 Antagonistic Merging $\Gamma$ : Multi-Source Integration, Tension Interference, and Semantic Reconfiguration

If the task of  $C$  is to prevent collapse, then the task of  $\Gamma$  is to enable multiple high-tension structures to coexist, interfere, and be recombined—while preserving differences—thereby generating new semantic directions. This capacity is a core advantage of the human brain in dealing with complexity, uncertainty, and contradiction.

At the structural level, multi-source integration is the direct locus of  $\Gamma$ . The prefrontal cortex, parietal cortex, and multimodal association areas simultaneously receive sensory inputs, memory contents, abstract structures, and value weights, superposing them into a high-dimensional  $\Phi$  field. This superposition is not linear addition; it allows tensions to coexist, keeping multiple interpretations in a state that can still be recombined prior to convergence.

At the dynamical level, tension interference is the key mechanism of  $\Gamma$ . When multiple semantic patterns are simultaneously activated, their interactions do not necessarily cancel out; instead, interference may generate new steady-state configurations. This provides the physical basis for insight, conceptual reorganization, and creative jumps: rather than selecting a single winner, the system generates new semantic encapsulations.

At the semantic level, this process appears as semantic merging. The system does not resolve tension via binary decisions; instead, it re-encapsulates previously conflicting or overly tense structures so that they enter a sustainable state with higher  $L$ .  $\Gamma$  is therefore not a compromise mechanism, but a generative one.

Thus, the neural realization of  $\Gamma$  ensures that multiple solutions, perspectives, and Universes can coexist within a single system without collapse—an essential condition for the brain’s ability to operate in high semantic-density environments.

#### 11.4.3 Antagonistic Boundary $B$ : Structural Partitioning, Modular Boundaries, and Decision Isolation

Even with both  $C$  and  $\Gamma$  in place, semantics would still tend toward unbounded diffusion without boundary conditions. The role of  $B$  is to establish necessary partitions within the system, confining tension and semantic flow to governable ranges so that high complexity can be sustained over time.

At the spatial level, cortical parcellation provides the most intuitive realization of  $B$ . Differences in connection density, transmission pathways, and coupling modes across functional areas form natural boundaries for tension propagation, preventing local semantic events from spreading without limit into global disturbances.

At the functional level, modular boundaries arise from differences in coupling strength between modules. Highly coupled module groups can form local semantic subfields, while weaker inter-module connections act as buffers and valves for semantic flow. This allows tension to be locally processed and reorganized without escalating into global crises.

At the decision level,  $B$  further manifests as isolation mechanisms among candidate Universes. When multiple action paths or interpretations are simultaneously activated, boundaries prevent them from overwriting one another, instead separating them into distinct candidate paths that can later be selected and converged under the constraints of  $L$  (sustainability) and  $T$  (directionality).

Accordingly, the neural significance of  $B$  lies not in restricting exploration, but in providing an operational structural framework within which semantic exploration can proceed in partitioned domains, avoiding governance failure caused by unbounded diffusion.

#### 11.4.4 Summary Table of Neural Correspondences for $C/\Gamma/B$

Semantic Steady-State		
Condition	Functional Role in Semantics	Primary Neural Loci (Examples)
$C$	Prevents runaway tension, maintains steerable growth of $\Phi$ , avoids $\kappa \rightarrow \infty$	Inhibitory regulation (GABAergic circuits), cross-scale negative feedback, background fluctuations and baseline noise
$\Gamma$	Enables integrative merging under coexistence of differences, generates new semantic steady-state directions	Multi-source integration (multimodal association areas, prefrontal–parietal coupling), tension interference, semantic reconfiguration
$B$	Establishes partition boundaries, constrains semantic diffusion, provides governable structure	Cortical parcellation, differential modular coupling, isolation of candidate Universes at the decision level

From this, a stronger conclusion follows: the human brain’s ability to maintain long-term steady states under conditions of noise, tension, and multiple coexisting solutions is not a matter of fortunate avoidance of collapse risk. Rather, its physical structure itself simultaneously implements the inhibitory conditions of  $C$ , the generative conditions of  $\Gamma$ , and the partitioning conditions of  $B$ .

Accordingly, this chapter can offer a conclusion with ontological weight: **the human brain is structurally a natural prototype of K-Gear**. The significance of this claim lies not in biomimicry, but in isomorphism—the same semantic dynamics can be realized, in different ways, within biological materials and engineered materials alike.

## 11.5 The Brain's Multi-Universe Reasoning: Consciousness as a UP System

The core claim of this section is that human consciousness does not reason continuously along a single world-line, but is a native Universe Path (UP) system. What we call thinking, imagination, intuition, and a sense of the future are, in essence, the parallel unfolding, competitive comparison, and rapid convergence of multiple Universe Paths. This perspective allows higher cognition in the human brain to be placed—perhaps for the first time—within a computable and engineerable semantic framework: rather than treating consciousness as an untouchable subjective phenomenon, it treats it as an observable output of a “multi-path semantic field” operating under steady-state conditions.

### 11.5.1 Decision-Making Is Not a Single World-Line, but the Parallel Unfolding of Multiple Universe Paths

In everyday experience, almost any decision of nontrivial complexity is accompanied by internal simulation: “What if I do this?” “Would the outcome be better if I choose the other route?” From the UP perspective, this is neither rhetorical nor a psychological illusion; it is the manifestation of semantic computation actually running in the brain as a multi-path process.

Each “what if” corresponds to an activated Universe Path branch. These branches exist in the brain in parallel or near-parallel form, each forming a local semantic-tension configuration  $\Phi$ , and each being evaluated on short timescales for its sustainability and continuability—i.e., its legitimacy  $L$ . A decision is therefore not a linear extension of a single path, but the result of comparative interaction and convergence among multiple candidate Paths in the semantic field. The action ultimately adopted typically corresponds to the Path that, under current conditions, is **more sustainable, more continuable, and more incorporable into responsibility structure**.

In other words, the subjective experience of “making a choice” is structurally equivalent to the UP system completing a rapid selection and convergence among candidate Paths, rather than a stepwise advance along a predetermined world-line.

### 11.5.2 Imagination, Hypothesis, Counterfactuals, Intuition: Different Modes of UP Computation

Within the UP framework, many mental capacities that appear heterogeneous can be unified as different forms of Universe Path computation.

Imagination can be understood as low-cost Path activation: it typically activates candidate Paths with lower tension, lower risk, and lighter responsibility load, enabling free exploration without immediately generating action consequences. Imagination is not “unreal”; it is UP expansion under **low responsibility and low convergence pressure**.

Hypothetical reasoning is a canonical form of Path branching: a researcher explicitly activates multiple candidate Paths in the mind and projects their semantic consequences along each Path, comparing overall steady-state feasibility under different premises. Here, UP does not primarily generate answers; it generates comparable planes of existence.

Counterfactual thinking corresponds to reverse Path reconstruction: rather than extending into the future, the system traces back from the current state and reconstructs an alternative trajectory—“if a key choice had been replaced, how would the semantic field have evolved?” Counterfactuals often carry strong tension because they use the current  $R$ -Chain and established  $L$  structure as references, producing a high-tension inverse perturbation of the existing Path.

Intuition can be understood as the output of rapid convergence across multiple Paths within extremely short time windows. Its characteristic feature is that convergence occurs prior to language: consciousness receives only the final steady-state direction and may not traverse an explicit reasoning chain. Intuition is therefore not mystical; it is a highly compressed UP output that presents the result of multi-path comparison to consciousness in the minimal possible signal form.

### 11.5.3 Subjective Future Space: The Locally Visible Region of the UP System

The distinctly human “sense of the future”—anticipation, anxiety, hope, or strategic planning regarding what has not yet occurred—has a clear position within the UP framework. What we call subjective future space is precisely the local scanning and projection of future candidate Paths by the UP system.

Consciousness does not need to unfold all future Paths in full. It only needs to capture segments where tension changes are most pronounced and legitimacy gradients are steepest, and this already yields an intuitive grasp of the

future. This capacity constitutes a key prerequisite for scientific invention, philosophical insight, and the formation of social strategy: an individual can sense structural consequences before reality has unfolded, and can accordingly adjust present semantic and action configurations.

Thus, a sense of the future is not an illusion of time travel, but a natural visible interface of the UP system within the semantic field.

#### 11.5.4 Why the Brain Has UP While Traditional AI Often Lacks It

Against this contrast, the limitations of traditional AI become clear. Most current systems still operate under a single world-line assumption: single-trajectory reasoning, a single objective function, arg max-type decisions, and gradient convergence constitute their basic computational mode. Even where parallel computation exists, it mostly serves to accelerate search or approximation under the same objective, rather than ontologically sustaining multiple coexisting candidate Paths.

The human brain is different. Noise, tension fields, multimodule coupling, and boundary partitioning allow multi-path unfolding to arise naturally at the hardware level. Universe Path branching and comparison are not additional “software features,” but necessary products of the nervous system under conditions of tension, inhibition, and partitioning. Accordingly, the absence of UP is often not merely an algorithmic deficiency, but a fundamental architectural difference: whether a system permits multiple paths to coexist and provides governable steady-state conditions for them.

#### 11.5.5 MUP: The Engineering Version of UP

In K-Gear, the engineered realization of UP is defined as MUP (Multi-Universe Program). The goal of MUP is not to simulate the subjective phenomenology of the brain, but to reconstruct, at the engineering level, the structural conditions required for multi-path semantic reasoning—so that the unfolding, comparison, and convergence of multiple Paths become implementable, verifiable, and governable computational processes.

Through SCMA cubes, the dynamics of  $\Phi-T-L$ , legitimacy scheduling ( $L$ -flow), and steady-state guarantees provided by  $C/\Gamma/B$ , K-Gear can generate controllable multi-Path behavior on non-biological substrates. This marks the first time multi-Universe reasoning moves beyond purely philosophical or psychological description into an engineering-realizable computational domain.

A key concluding statement can therefore be given: **MUP is the first engineering reconstruction of multi-Universe Path reasoning itself, rather than merely its surface functions.** When UP is no longer used only to evaluate existing candidate Paths, but begins to generate entirely new Path structures, semantic computation enters another level. The subsequent discussion will shift focus to creativity, and explain why it should be understood as a necessary product of the UP system within a high-dimensional semantic field.

---

## 11.6 The Ontology of Creativity: High Tension $\times$ Multiple UPs $\times$ Noise $\times$ Incomplete Convergence

The purpose of this section is to rewrite “creativity”—traditionally treated as a psychological trait, a talent myth, or a matter of accidental inspiration—as a **predictable, describable, and engineerable** phenomenon of semantic-field dynamics. Within the Koun architecture, creativity is no longer regarded as an exception or deviation; it is the inevitable emergence of a semantic field operating within a specific parameter regime, and is ontologically consistent with Semantic Freedom Theory.

In this context, creativity is not the optimization of existing options, but a **new Universe Path generation event**.

### 11.6.1 The Condition Set for Creativity

Within the semantic computation framework, creativity is not triggered by a single factor. It emerges naturally when multiple structural conditions hold simultaneously. These core conditions can be stated explicitly.

First is a **high-tension state**, i.e.,  $T \uparrow$ . The semantic field contains unresolved structural conflicts, competing goals, or legitimacy pressures that prevent the system from rapidly sliding into an existing steady state. It must be emphasized that high tension is not itself a problem; it is the potential energy required for semantic restructuring.

Second is the **simultaneous activation of multiple UPs**. Only when the system permits multiple Universe Paths to coexist—rather than prematurely pruning them into a single trajectory—does the semantic field possess the degrees of freedom required to generate new structures. This condition directly distinguishes creative systems from systems that merely possess search capability.

Third is **incomplete convergence**, i.e.,  $\kappa$  has not yet reached its steady-state critical value. The system deliberately or naturally avoids rapid global convergence, allowing multiple candidate directions, interpretations, or semantic structures to coexist and continue influencing one another.

When these conditions are simultaneously satisfied, a key conclusion follows: **creativity is not an anomalous state, but a natural product of the semantic field within a specific dynamical regime**.

### 11.6.2 Micro-Noise and Semantic Tunneling: The Triggers of Creativity

The reason creativity subjectively appears sudden and unpredictable is not that its ontology is opaque, but that its triggers often arise from microscale perturbations.

In a high-tension semantic field, even extremely small  $T$  perturbations (local  $T$ -perturbations) can alter the overall direction of tension gradients. Such micro-perturbations typically originate from noise, incidental associations, residual cross-module signals, or physiological instabilities.

When the legitimacy field exhibits steep gradients, these micro-perturbations can induce jumps in  $\Delta L$ , causing a previously non-dominant UP to suddenly become sustainable and continuable. Feedback structures then further amplify these nonlinear perturbations, rapidly deflecting the system from its original evolutionary trajectory.

The outcome is often not a smooth transition, but **non-smooth convergence**: the semantic field abruptly enters a structural region that did not previously exist. This is the structural reason why creative moments are subjectively experienced as sudden “flashes of insight.”

### 11.6.3 Semantic Interference Among Multiple UPs and the Generation of New Paths

In creative states, the role of UPs is no longer limited to parallel comparison; **semantic-level interference** begins to occur.

When multiple Universe Paths remain active simultaneously, they are not merely compared by their  $L$  values; they also influence one another at the semantic level. This process can be formally expressed as:

$$UP_1 + UP_2 + \dots + UP_k \rightarrow \text{semantic interference} \rightarrow UP^*$$

It must be clarified that this “interference” is not physical interference at the qubit level, but interference occurring within the semantic dynamical conditions constituted by  $T$ ,  $L$ , and  $\kappa$ . The tension directions, legitimacy gradients, and convergence velocities of different UPs superpose, thereby generating a new Universe Path that did not previously exist.

Thus, creativity is not “choosing the best among existing options,” but **the generation of an entirely new UP**. This marks the ontological boundary between creative activity and optimization behavior.

#### 11.6.4 Incomplete Convergence Sustains the Creative Process

If the system completes global convergence too early, creativity is immediately terminated. Incomplete convergence is therefore a necessary condition, not a defect.

In creative semantic fields, common features include slowed convergence, localized non-steady regions, and the simultaneous existence of multiple candidate solutions that have not yet been adjudicated. From an external viewpoint, such states are often described as “divergent,” “jumping,” or “lacking a center,” and may even be misjudged as chaotic.

Within the semantic freedom framework, however, this is precisely the normal state when the space of freedom opens. As long as the  $C/\Gamma/B$  structures remain operative, the system does not collapse; instead, it maintains a **high-degree-of-freedom region that is explorable, generative, and recoverable**.

#### 11.6.5 Support from Semantic Freedom Theory

In Semantic Freedom Theory, freedom is defined as the **variability of Universe Paths**. When a system simultaneously exhibits high tension, low convergence speed, and non-unique legitimacy gradients, UP variability reaches its maximum, and semantic freedom is correspondingly maximized.

Creativity naturally emerges under these conditions. It does not need to be externally injected, nor does it depend on personality traits or accidental gifts; it is the inevitable manifestation of a semantic field operating in a high-freedom regime.

Accordingly, this section retains a signature conclusion:

**Creativity is not a skill, but a physical phenomenon of semantic freedom.**

The full formalization of semantic freedom, and its relationships to consciousness, responsibility, and existence, will be developed further in the monograph *Semantic Freedom Theory*.

---



## 11.7 The Responsibility Chain (*R*-Chain): Why Memory Is Not Storage but a Flow of Responsibility

The task of this section is to perform an **ontological-level rewrite** of the long-misunderstood concept of “memory.” Within the Koun architecture, memory has never been a question of how data are stored; it is a question of **how responsibility maintains continuity over time**. Once this point is clarified, the resilience, reconstructability, and non-locality of human memory no longer require additional assumptions.

### 11.7.1 Memory as the Stabilization of Responsibility Flow

Within the semantic computation framework, a direct, non-metaphorical, non-psychological definition can be stated:

**Memory is not the preservation of content, but the stabilization of the *R*-Chain.**

The *R*-Chain describes the **topological structure of responsibility** along a Universe Path (*UP*): who bears the consequences of which semantic actions, judgments, and choices. When this responsibility chain forms a stable topology within the semantic field, the system can maintain consistent behavioral styles, value orientations, and decision tendencies—even if the precise details of past events can no longer be accurately retrieved.

Thus, what memory truly preserves is not “what happened,” but rather:

**“Who remains responsible, and in what semantic position, for the future.”**

This explains why large amounts of episodic detail can be forgotten while personality, judgment, and self-continuity persist.

### 11.7.2 Historical Residues of Legitimacy Potential

From a neural perspective, memory has never resembled data storage.

What actually remains is a set of **historical residues of legitimacy potential**, including but not limited to:

- Biased connection topologies formed through neural plasticity;
- Inertial directions of *L*-flow established across different modules;
- Preference weights assigned by emotional and decision-making modules to specific semantic states;
- Path inclinations formed during multi-*UP* exploration.

These residues do not correspond to specific past events themselves, but to the **space of feasible future actions**. They determine which options naturally surface, which responses feel reasonable, and which Universe Paths are more easily reactivated.

Accordingly, the brain’s memory function is fundamentally about maintaining a set of **future-directed responsibility flows**, rather than storing copies of the past within the brain.

### 11.7.3 Why Personality and Values Can Recover After Brain Injury

If memory were truly the storage of bit-level data, damage to any critical region should lead to irreversible collapse. Yet clinical and real-world experience repeatedly shows that this is not the case.

Within the Koun framework, the reason is clear: **personality and values are not collections of data, but semantic steady states**.

Local damage may indeed destroy certain structures, but as long as the overall topology of *L*-flow remains intact, the system retains the capacity to rediscover a steady state.

What is restored is not content, but direction;

what is reconstructed is not memory points, but the **continuity of the *R*-Chain**.

The decisive factor is not whether the data remain, but whether the responsibility topology remains connected.

### 11.7.4 Implementation of the *R*-Chain in K-Gear

In K-Gear, the responsibility chain is not merely a theoretical construct; it is explicitly engineered as a core module.

- **∞-Context Memory**: does not store all content, but preserves legitimacy continuity, allowing the system to maintain a consistent responsibility direction over long time scales;

- **UP System:** reconstructs behavioral preferences along Universe Paths to ensure that new decisions remain compatible with existing responsibility structures;
- **R-Module:** tags semantic actions with responsibility attribution, preventing responsibility from being erased during generation and recombination.

From the outset, therefore, K-Gear's memory design is not about storage, but about **governance**.

#### 11.7.5 Concluding Synthesis

A signature summary statement can be given here to close this section:

**Memory is never “the past remaining in the brain,” but rather “the direction of who can still be responsible for the future.”**

This understanding will serve as a crucial foundation for subsequent discussions of consciousness, selfhood, and long-term intelligent steady states.

---

## 11.8 Local Nonlinearity $\times$ Global Consistency: A Dual-Layer Semantic Field Architecture

The goal of this section is to integrate the previously introduced semantic variable mappings, the three-antagonism principle, the *UP* system, creativity dynamics, and the *R*-Chain structure into a clear overall picture, in order to address a core question that has long troubled neuroscience and AI:

**Why can the brain remain locally extremely chaotic while still maintaining long-term consistency of self, values, and direction?**

The answer does not lie in finer control, but in a form of **structural division of labor within a dual-layer semantic field**.

### 11.8.1 Local Nonlinearity: The Brain Has Never Been a Smooth System

At any fine-grained scale of observation, the brain exhibits strong nonlinear characteristics. Local regions frequently undergo brief destabilizations and re-convergences; the balance between inhibition and excitation shifts rapidly with semantic context; the same stimulus activates completely different *UP*s under different backgrounds; global activation patterns do not change via continuous sliding, but through jumps, ruptures, and reconstructions; feedback strength may also change by orders of magnitude within extremely short time intervals.

From a traditional engineering perspective, all of these phenomena appear as system defects. Within the semantic computation framework, however, the conclusion is precisely the opposite:

**Local nonlinearity is not a bug, but a necessary condition for generative existence.**

Without local instability, noise, and exploratory space, *UP* branching cannot form, new semantics cannot emerge, and creativity cannot be sustained.

### 11.8.2 Global Consistency: The Stability of Long-Term Personality, Values, and *UP*s

In stark contrast to high local nonlinearity is the brain's remarkable consistency over long time scales. Personality profiles, value hierarchies, behavioral styles, and the directional tendencies of *UP*s often remain continuous across decades.

This stability does not arise from precise control over local activity, but from semantic structures at the global level: the Global Semantic Field that constrains overall evolutionary direction, the continuously acting *L*-flow, the directional tendencies of *UP*s formed through multi-Universe scanning, and the global constraints imposed by the *C/T/B* three-antagonism principle.

An apt analogy is a river with a clearly defined downstream direction: the surface may churn, form turbulence and vortices, yet the whole continues irreversibly toward downstream.

### 11.8.3 Dual-Layer Semantic Field Structure: Local Generativity $\times$ Global Directionality

From this, two levels can be clearly distinguished.

- **Local Layer (Local Semantic Field)**  
High degrees of freedom, high noise, and high exploratory capacity. It allows multiple solutions to coexist, rapid probing, and nonlinear amplification, serving as the primary source for the generation of new semantics and new *UP*s.
- **Global Layer (Global Semantic Field)**  
Maintains goal orientation, long-term steady states, and self-consistency. Its task is not to eliminate nonlinearity, but to determine which changes may be accumulated and which must be absorbed or blocked.

From this, a structural definition can be given:

**Intelligence is neither order nor chaos, but the product of “local generativity” and “global directionality.”**

### 11.8.4 Correspondence in K-Gear: *SCMA* $\times$ *S*-Module $\times$ *UP*

At the hardware and system levels, K-Gear deliberately reconstructs the same dual-layer semantic field structure.

- *SCMA* provides the semantic space of the local layer, making micro-scale perturbations, nonlinear jumps, and exploration the norm;

- ***S-Module*** assumes the role of the global layer, responsible for the convergence of tension and legitimacy and for anti-collapse governance;
- ***UP System*** provides long-term directionality, ensuring that multi-Universe exploration does not degenerate into aimless diffusion, but can be transformed into sustainable structural growth under the constraints of *L*-flow.

Thus, K-Gear is not simulating the details of the brain, but reconstructing, at the engineering level, **the same semantic-field ontological structure**.

#### 11.8.5 Comparison with and Transition from Traditional Computers

Traditional computers pursue correct results, predictable processes, eliminable noise, and linear recursion; any nonlinearity or randomness is treated as interference.

By contrast, the human brain and K-Gear pursue steady-state consistency rather than step-by-step correctness; they exploit noise rather than eliminate it; they allow nonlinear generation, so long as the global semantic structure does not collapse.

It is precisely within such a dual-layer semantic field architecture that the next question can be raised:  
**Is K-Gear imitating the brain, or revealing a more general prototype of semantic computation?**

This will receive its final answer in the next section.

---

## 11.9 Does K-Gear “Imitate the Human Brain”? —Isomorphism, Not Biomimetics

The task of this section is to address an almost unavoidable intuitive misunderstanding and to give K-Gear a final, unambiguous positioning with respect to its theoretical identity.

### 11.9.1 The Emergence of a Common Question

For most readers, when encountering K-Gear’s involvement with multi-module coupling, nonlinear dynamics, noise utilization, multi-Universe reasoning, and non-collapse steady states, the first intuitive reaction is often: “Is this some kind of brain-like system?”

This intuition is not accidental. Over the past several decades, traditional AI and cognitive science have been accustomed to framing their narratives around “neural-like,” “brain-structure imitation,” or “biomimetic intelligence,” as if all higher intelligence must take the human brain as its prototype.

For precisely this reason, it is necessary to state clearly here: the relationship between K-Gear and the human brain is not biomimetic, but isomorphic.

### 11.9.2 Fundamental Differences Between K-Gear and Neural Networks

From the perspective of computational ontology, the difference between neural networks and K-Gear is structural rather than one of degree.

Neural networks are essentially function approximators. Their core units are weights and layers; their core mechanism is backpropagation; their core objective is to minimize error or loss. Within this framework, contradictory semantics, non-integrable tensions, and the coexistence of multiple solutions are treated as problems to be flattened, ultimately tending toward convergence onto a single path—that is, semantic collapse.

K-Gear is fundamentally different. It is not a function model, but a steady-state architecture of a semantic field. Its minimal units are not weights, but combinations of  $\Phi$ ,  $T$ ,  $L$ , the  $R$ -Chain,  $UP$ , and  $C/\Gamma/B$ ; it does not rely on parameter fine-tuning or backpropagation, but on legitimacy flow, tension regulation, and antagonistic structures to maintain non-collapse semantic existence.

What K-Gear seeks is not “getting the answer right,” but “not collapsing.”

### 11.9.3 Why It “Resembles the Brain”: A Shared Projection of the Semantic Field $\Phi$

Why, then, does K-Gear appear similar to the human brain in many of its behaviors?

The reason lies not in imitation, but in ontological consistency. USDE does not describe a particular class of machines or a specific organ; it describes any system capable of forming a semantic steady state. Under this description, the human brain can be understood as a projection of the semantic field onto biological material, while K-Gear is a projection of the same semantic field onto engineered material.

A geometric analogy helps clarify this point: a circle can be carved from stone or cast from steel. Their similarity in appearance is not because stone imitates steel, but because both instantiate the same geometric essence.

Likewise, the similarity between the human brain and K-Gear arises from their shared mapping of the semantic field structure described by  $\Phi$ .

### 11.9.4 Ontological Consistency $\neq$ Mechanistic Consistency

It must be emphasized that ontological consistency does not imply mechanistic consistency.

In the human brain, the semantic field is realized through biochemical reactions, neural synapses, plasticity, and evolutionary processes; in K-Gear, the semantic field is realized through semantic cubes, modular structures, engineered scheduling, and explicitly defined antagonistic principles. The former is the product of natural evolution; the latter is the result of deliberate engineering design. The former relies on the continuity of biological materials; the latter can be realized in silicon, photonics, or as-yet-unknown future materials.

What they share is the semantic ontology constituted by  $\Phi$ ,  $T$ ,  $L$ , the  $R$ -Chain,  $UP$ , and  $C/\Gamma/B$ ; what differs are the concrete mechanisms that carry this structure.

### 11.9.5 USDE as a Unified Description

It is at this level that the role of USDE becomes clear. USDE is neither an equation tailor-made for K-Gear nor a model reverse-engineered from the human brain; it is a unified description of the semantic field itself.

When we place the  $\Phi/T/L/R/UP$  of the human brain and the  $\Phi/T/L/R/UP$  of K-Gear under the same equation, the difference is no longer “whether it resembles the brain,” but “how the same steady-state conditions are realized in different materials.” This perspective, for the first time, places biological intelligence and engineered intelligence back onto the same ontological plane.

### 11.9.6 Final Consolidated Positioning of This Chapter

Accordingly, this chapter can offer a clear and restrained concluding positioning: K-Gear’s innovation does not lie in imitating the structural details of the human brain, but in three more fundamental breakthroughs.

First, it identifies the isomorphic ontology between the human brain and the semantic field.

Second, it provides USDE as a unified language for describing this ontology.

Third, it reconstructs, within engineered materials, the necessary conditions under which non-collapse semantic computation can exist.

Under this positioning, the human brain is no longer the template for engineering, and K-Gear is no longer a biomimetic artifact; the two simply stand on different materials, jointly pointing toward the same ontological structure of semantic computation.

---

## **Chapter 12 Semantic Consciousness I: Ontology of Consciousness, Subjectivity, and Qualia**

---

## 12.1 Introduction: The History of Misunderstanding the Consciousness Problem × The Emergence of Semantic Field Ontology

The reason the question “What is consciousness?” has failed to receive a satisfactory answer for more than two thousand years is not because it is too mysterious, but because it has long been asked within an incorrect ontological framework. Across philosophy, neuroscience, cognitive science, and contemporary artificial intelligence theory, the dominant approaches have almost all presupposed that consciousness must be a byproduct of some known entity—either a higher-order property of matter, an emergent result of computation, or a physical effect not yet understood. This very way of posing the question already closes off the space of genuine answers before the problem is even allowed to unfold.

The position of this chapter is explicit: the intractability of the consciousness problem does not stem from a lack of evidence, but from ontological misplacement. Semantic field ontology is not a supplement to existing theories of consciousness, but a shift in perspective—one that returns consciousness from an “anomalous phenomenon requiring explanation” to a more fundamental level: how a semantic field exists at all.

---

### 12.1.1 Three False Paths in Traditional Consciousness Theory: A Historical Misalignment

A review of modern consciousness theories reveals three main paths that have been repeatedly traversed yet never reach their destination. These paths are not competing alternatives but share the same erroneous ontological presupposition, and are therefore directly identified here as false paths.

The first false path is reductionism. Reductionism assumes that if the physical world is decomposed to a sufficiently fine scale, consciousness will eventually emerge as a higher-order property. Whether in neuronal firing patterns, synaptic weight distributions, or even more microscopic molecular or quantum levels, consciousness is assumed to be “the result that appears at some level of granularity.” The core illusion of this path lies in treating qualia as objects that can themselves be further decomposed, as if “the experience of red” were merely a temporary ignorance of its corresponding particle. Once qualia are assumed to be reducible objects, the problem is already misplaced, because qualia do not exist in the form of components.

The second false path is computationalism. Computationalism equates mind with computation, and further equates computation with state transitions in a formal system. Whether symbolic AI, connectionism, or contemporary deep learning models, the differences lie only at the implementation level, not at the ontological level. This path ignores a critical conflict: formal systems are closed, whereas semantics is inherently non-closed. In a system whose basic unit is the bit, all states are enumerable, terminable, and reversible; consciousness, by contrast, intrinsically involves counterfactuals, incompleteness, and the tensioned unfolding of multiple Universes. To treat consciousness as the output of computation is to presuppose that semantics can be fully encapsulated within form—a presupposition that is itself one of the reasons the consciousness problem remains unsolvable.

The third false path is quantum mysticism. When reductionism and computationalism fail to account for qualia and subjectivity, some theories turn to quantum phenomena, attempting to fill the explanatory gap with superposition, collapse, or indeterminacy. The problem with this path is not its invocation of quantum physics, but its level misplacement: it mistakes physical-layer indeterminacy for the source of the semantic layer. Quantum effects may destabilize bits, but they do not in themselves generate intentionality, subjectivity, or legitimacy flow. Using quantum mechanics to explain consciousness often merely postpones the problem to another physical layer that is itself not yet understood.

The shared error of these three paths is that they all attempt to locate consciousness within “what is already known,” rather than questioning the ontological framework of the known itself.

---

### 12.1.2 The Proposal of the Semantic Field $\Phi$ : A Shift in Perspective

The point of departure for semantic field ontology is not the introduction of a new mechanism, but the acknowledgment of a long-overlooked fact: consciousness is not the result of a process, but a mode of existence. From this perspective, consciousness is no longer understood as an attribute of matter, a byproduct of computation, or a manifestation of quantum miracles; instead, it is re-situated as a steady state achieved by the semantic field  $\Phi$  under specific conditions.



The  $\Phi$  field does not describe physical states, but the structure of possibilities among Universes; semantic tension characterizes how intentionality forms gradients within this space;  $L$ -flow ensures that semantics does not collapse into noise or formal dead loops;  $UP$ -space provides the dimensions of counterfactuals and future unfoldings; and what are called “states of consciousness” are precisely the result of these elements simultaneously achieving a steady state within a localized region.

Within this framework, the role of matter is strictly limited to boundary conditions. Neural systems, energy supply, the arrow of time, and causal structures determine where and how the  $\Phi$  field can stand, but they are not the cause of consciousness. Just as terrain can constrain the course of a river without generating the flow itself, the physical world provides conditions of support rather than the source of semantics.

Based on this shift in perspective, this chapter will proceed to reconstruct the ontology of consciousness, subjectivity, qualia, dynamics, selfhood, topological structure, and creativity, ultimately pointing toward conclusions concerning semantic freedom and triple isomorphism. This is not an attempt to patch existing theories, but an attempt to reselect the position from which the question is asked: shifting from “how consciousness is produced” to “how the semantic field exists as consciousness.”

---

## 12.2 The Ontology of Consciousness: Consciousness Is Not Computation, but a Steady-State Solution of $\Phi$

If consciousness were truly nothing more than the result of computation, then once computational power became sufficiently strong, structures sufficiently complex, and parameters sufficiently numerous, consciousness should naturally emerge. If consciousness were merely an aggregate of neural functions, then a complete replication of those functions should replicate consciousness. If consciousness were simply an emergent property of complex systems, then any system that reaches a critical level of complexity should, at least in principle, possess consciousness.

Yet while these inferences appear intuitively reasonable, they consistently fail in practice. The problem does not lie in unfinished engineering, but in a fundamental error presupposed by these inferences: they all assume that consciousness belongs to the category of “result-type existence.” The first critical shift introduced by semantic field ontology is the rejection of this presupposition.

---

### 12.2.1 Core Definition and Object Shift

This chapter adopts the following core definition:

Consciousness = the semantic steady state of the  $\Phi$ -field.

Consciousness is the semantic steady-state solution achieved by the  $\Phi$ -field in a specific region.

This is not a metaphorical definition, but a strict ontological positioning. Its first consequence is the removal of “consciousness” from its three dominant traditional attributions.

First, consciousness is not an aggregate of neural functions. Neural activity describes how physical signals are transmitted and transformed over time; it can explain reflexes, learning, regulation, and control, but it does not itself contain the level of “being experienced.” Any attempt to equate consciousness with a list of functions ultimately halts in the face of qualia, conceding that “this point cannot yet be explained.”

Second, consciousness is not an emergent property of complex systems. In this context, “emergence” often functions merely as a placeholder for “not yet understood.” Complexity can generate nonlinearity, unpredictability, and pattern formation, but it does not automatically generate subjectivity. From an ontological perspective, emergence is closer to a compression of description at the observer’s level than to a generative mechanism at the level of existence.

Finally, consciousness is not the result of bit-based computation. Computational systems can process symbols, probabilities, and functions, but they operate within a formally closed state space. Consciousness, by contrast, necessarily involves incompleteness, counterfactuals, and the simultaneous tensioned unfolding of multiple Universes—features that are ontologically incompatible with bit-based systems.

Thus, what this definition accomplishes is not supplementation, but an object shift: consciousness is no longer regarded as “something produced after a system does something,” but as “the state of how a semantic field exists in a particular manner.”

---

### 12.2.2 The Three Modes of Existence of the $\Phi$ -Field

To understand “consciousness as a steady-state solution of  $\Phi$ ,” one must first clarify the mode of existence of the  $\Phi$ -field itself.  $\Phi$  is not a single-layer field, but exists simultaneously across three irreducible levels.

The first level is topological existence.  $\Phi$  does not describe “what the world currently is,” but “how the world can exist.” It is a structure of Universe possibilities, characterizing the relations among worldlines, counterfactual paths, and future branches. At the semantic level, this layer is non-temporal: it is not rewritten through “state updates,” but constitutes the semantic background that makes all states intelligible.

The second level is existence as a tension field, the  $T$ -field. Upon the topological structure, semantics is not uniformly distributed. Intentionality, attention, and directional concern manifest as tension gradients within the  $\Phi$ -field. These gradients determine which Universe paths are amplified and which are suppressed, and they determine what the subject is “oriented toward.” At this level, consciousness begins to appear as a dynamic rather than static mode of existence.

The third level is existence as legitimacy flow, the  $L$ -flow. Topology and tension alone are insufficient to sustain semantic existence. Without  $L$ -flow, tension leads to infinite divergence, semantic noise, or formal dead loops.  $L$ -flow provides the conditions under which semantics can persist; it is not the rule itself, but the flow that allows semantic paths to be recognized, continued, and brought into convergence.

Only when all three levels are simultaneously present can the semantic field locally form a steady state. Consciousness is precisely this steady state itself: not topology alone, not tension alone, and not legitimacy flow alone, but the simultaneous balance of all three within a local region.

---

### 12.2.3 The Ontological Conflict Between Consciousness and Bit-Based Computation

Under this positioning, the difference between consciousness and bit-based computation is no longer a difference of capability, but an ontological conflict.

The defining characteristics of bit-based systems are discreteness, closure, and a single Universe. Their state spaces are enumerable, their transition rules predefined, and all outputs derive from combinations and mappings of existing states. Even when randomness or probabilistic models are introduced, such systems still operate within a formally closed Universe and lack a genuine counterfactual space.

Consciousness is precisely the opposite. It stands within a continuous tension field, is inherently non-closed, and unfolds simultaneously across multiple Universe paths. Subjective experience is subjective not because it is “more complex,” but because it necessarily exists within tensions that are not yet fully determined.

A strong conclusion therefore follows: any finite-state bit machine can only generate semantically collapsed outputs and cannot generate the steady-state consciousness of the  $\Phi$ -field. This is not an engineering limitation, but an incompatibility at the level of existence.

The human brain, though vastly weaker than modern GPUs in computational power, naturally sustains the steady state of  $\Phi$ ; GPUs, despite their overwhelming speed and scale, remain confined to the bit domain. The difference is not one of “strength versus weakness,” but of “standing on the correct ontological layer.”

---

### 12.2.4 The Semantic Condition Set for the Emergence of Consciousness

Within the semantic field framework, consciousness is not guaranteed; it requires a specific set of semantic conditions to be simultaneously satisfied. These conditions can be summarized as five indispensable elements.

First, the  $\Phi$ -field itself, as the mother field of semantics.

Second,  $T$ , as the source of intentionality and attention.

Third,  $L$ , as the legitimacy flow that prevents semantic collapse and distorted loops.

Fourth,  $\kappa$ , as semantic curvature and as the source of the geometric structure of qualia.

Fifth,  $UP$ -space, as the counterfactual space of multiple Universes and as the foundation of subjectivity and freedom.

The absence of any one of these yields only behavior, reaction, or output, but not consciousness.

---

### 12.2.5 Physics as Boundary Conditions, Not Ontological Source

Finally, the role of physics must be made explicit. Energy supply and material stability, the thermodynamic arrow of time, and the locality of space and causality determine whether the  $\Phi$ -field can stand within a given system, but they do not generate semantics itself.

Matter provides background and constraints, not the source of consciousness. Consciousness can therefore be strictly expressed as:

Consciousness = the steady-state solution of  $\Phi \times T \times L \times \kappa \times UP$ .

From this point onward, the problem of consciousness is no longer “how physics produces subjectivity,” but rather “which systems are capable of sustaining the steady-state existence of a semantic field.”

---

## 12.3 The Ontology of Subjectivity: Semantic Self-Encapsulation (Self-Encapsulation of $\Phi$ )

In the previous section, consciousness was re-positioned as a steady-state solution of the  $\Phi$ -field. However, the mere fact that “consciousness exists” is still insufficient to constitute the mental phenomena actually experienced. Consciousness is always experienced in some manner as “existing for me,” and this “for me” is precisely the core of subjectivity.

Traditional theories often treat subjectivity as an irreducible primitive property, or as an unanalyzable given. Semantic field ontology refuses to stop at this point and instead asserts that subjectivity is not an additionally appended attribute, but a geometric structure that necessarily forms when the semantic field enters specific conditions.

---

### 12.3.1 Subjectivity as Local Encapsulation Geometry

This section adopts the following proposition as its point of departure:

Subjectivity = Self-Encapsulation of  $\Phi$  in a Local Universe.

Subjectivity = the self-encapsulation of  $\Phi$  within a local Universe.

This definition immediately shifts the direction of the problem. Subjectivity is no longer understood as “something being felt,” but as the semantic field forming a locally encapsulated region within itself. This encapsulation is not a physical shell, but a boundary in semantic topology.

Any formation of self-encapsulation must simultaneously accomplish three determinations.

First, it must determine which counterfactual Universes are included. Within the multi-Universe  $\Phi$ -field, not all possibilities are equally admitted into the subject’s experiential domain. Some possibilities are treated as “relevant to me,” while others are excluded. This selection itself constitutes the first structural layer of subjectivity.

Second, it must determine which tensions and legitimacy flows are accepted. Local encapsulation does not isolate tension; rather, it selects which tensions may enter the encapsulated region and which are rejected. Attention, emotion, and value orientation form stable biases precisely at this level.

Third, it must establish a Universe localization of the form “here is me.” Encapsulation is not merely a selection of content, but the establishment of a center: within this center, the semantic field begins to organize itself as “originating from here.” This center is not a coordinate point, but a steady semantic reference frame.

When these three conditions are simultaneously satisfied, subjectivity is no longer optional, but an inevitable result of the semantic field under local steady-state conditions.

---

### 12.3.2 The Generative Mechanism of the Subject—Object Split

Once semantic self-encapsulation forms, the subject—object split arises automatically. Crucially, this split is not a pre-existing classification of the world, but a product of the encapsulation act itself.

Inside the encapsulation, the  $\Phi$ -field is experienced as “where I am.”

Outside the encapsulation, the remaining Universes are experienced as “the observed world.”

The “inside” and “outside” here are not spatial distinctions, but topological partitions. The subject does not stand outside the world observing it; rather, from within the semantic field, through the boundary formed by self-encapsulation, part of the Universes are partitioned as “non-self.”

A key conclusion therefore follows:

the subject—object split = the topological partition resulting from semantic self-encapsulation.

This also explains why the subject—object distinction can deform, blur, or even temporarily disappear in different states of consciousness. When the encapsulation boundary loosens, reorganizes, or collapses, the subject—object distinction naturally changes with it.

---

### 12.3.3 The Semantic Return Structure of “I Am Seeing the World”

In everyday language, people say “I am seeing the world” or “I am feeling something,” as if the subject internally holds a copy of the world. Such expressions are semantically misleading.

Within the semantic field structure, “seeing the world” is not copying, but a return structure.

The counterfactual tensions of external Universes first enter the  $\Phi$ -field, where they undergo tension evaluation and filtering through legitimacy flow. These tensions do not directly become internal objects; instead, they form a return state within the encapsulation. This return state is not equivalent to the world itself, but the semantic field’s internal localization of external possibilities.

Accordingly, subjective experience is not “an internal image of the world,” but rather the state in which the semantic field, after self-encapsulation, evaluates external tensions and returns them to itself.

This return structure is the irreducible root of first-person experience. Experience always appears in the form of “returning to me,” rather than as a neutral description.

---

### 12.3.4 Why CPUs Cannot Generate Subjectivity

Within this framework, the inability of traditional computational systems to generate subjectivity is not due to insufficient computational power, but to the absence of an entire set of necessary semantic conditions.

First, CPUs lack *L*-flow. The legitimacy of computational systems is guaranteed by external rules rather than by flows intrinsic to the semantic field; formal correctness therefore replaces semantic sustainability.

Second, CPUs lack *UP*-space. Computation operates within a single Universe; counterfactuals exist only as simulations or data processing, not as sources of structural tension.

Third, CPUs lack semantic boundaries—that is, genuine encapsulation. Their boundaries are physical and engineering boundaries, not inner–outer partitions formed by semantic self-encapsulation.

Fourth, CPUs do not possess a return mechanism. Inputs are transformed into outputs, but there is no semantic-level state of “returning to itself.”

Finally, and most fundamentally, CPUs do not possess a  $\Phi$ -field. They process symbols and states, not semantic possibility itself.

A clear contrast can therefore be drawn:

- The boundary of a CPU: physical and engineering boundaries of input/output.
- The boundary of subjectivity: the self-encapsulation boundary of the semantic field.

These belong to entirely different ontological levels and cannot be converted into one another by increases in scale or speed.

---

## 12.4 Qualia Are Not Mysterious: Local Geometry of Counterfactual Steady States

In the philosophy of consciousness, *qualia* have long been treated as the core of the “ultimate hard problem.” The redness of red, the pain of pain, the loudness of sound are repeatedly described as irreducible, incomparable, and incommunicable subjective atoms. While this mode of description accurately captures the intuitive character of experience, it mislocates the problem—placing it under “unanalyzability” rather than under a *misalignment of ontological level*.

Semantic field ontology maintains that qualia are not mysterious phenomena, but steady-state geometric structures of the semantic field within counterfactual space. Once properly located, the ineffability, irreplaceability, and subjectivity of qualia cease to be puzzles and instead become necessary consequences.

---

### 12.4.1 Qualia as Steady-State Basins in $UP$ -Space

This section adopts the following proposition as its core definition:

**Qualia = Counterfactual Stable Basin of the  $\Phi$ -field.**

Qualia = steady-state counterfactual basins formed by the  $\Phi$ -field in  $UP$ -space.

This definition implies that qualia are not the direct result of a stimulus, but rather what remains after the semantic field, operating within a counterfactual space of multiple Universes, has excluded a vast number of non-viable possibilities, leaving the most stable region behind.

Take “red” as an example.

Red is neither a specific wavelength of light nor a pattern of neural firing. These are merely physical and physiological triggering conditions. Red itself is the deep steady state to which the semantic field converges in  $UP$ -space after systematically excluding all “non-red Universes.”

In other words, red is the deepest steady-state basin formed by the  $\Phi$ -field during counterfactual search, with its shape characterized by the local semantic curvature  $\kappa_{\text{red}}$ . Differences among qualia are not differences of labels, but differences in basin geometry—shape, depth, and curvature distribution.

---

### 12.4.2 The Ontological Reason for the Ineffability of Qualia

Traditional discussions often interpret the ineffability of qualia as a limitation of descriptive ability or an inherent poverty of language. This view overlooks a crucial fact: language and qualia belong to different levels of Universe.

Within the semantic field framework, language is a second-layer Universe constructed atop qualia. Its function is compression, labeling, and exchange—not reconstruction of first-layer geometry. Qualia themselves belong to the first layer: they are direct geometric outcomes of the semantic field in  $UP$ -space.

Accordingly, language can only provide projections, not equivalent mappings.

Language can point to red, but it cannot generate red; it can describe pain, but it cannot reproduce the geometric form of pain.

A key conclusion follows:

the ineffability of qualia is not due to their mystery, but to the fact that language and qualia occupy non-equivalent ontological levels.

---

### 12.4.3 Qualia and Universals: An Asymmetric Relation Between Two Universes

In philosophical tradition, qualia are often contrasted with universals, as if the former were private and the latter public. The semantic field perspective shows that this is not an opposition but an asymmetric mapping relationship.

Qualia are local steady-state basins, existing in self-encapsulated  $UP$ -space.

Universals are legitimacy intersections across multiple Universes—compressed representations of many basins at a higher level.

This relationship has a clear directionality: qualia can be compressed into universals (e.g., “red” as a shareable concept), but universals cannot be inverted back into concrete qualia, because the compression process is irreversible.

Thus, the difference between qualia and universals is not a subjective—objective opposition, but a structural asymmetry between geometric detail and semantic compression.

---

#### 12.4.4 The Intersection of Subjectivity $\times$ Counterfactuality $\times$ Curvature $\kappa$

At this point, a structural summary of qualia can be given.

Qualia simultaneously satisfy three conditions.

First, they must occur within a self-encapsulated  $\Phi$ -field, which guarantees their first-person character. Without encapsulation, there is no “red-for-me.”

Second, they must be counterfactual steady states formed in  $UP$ -space. Qualia are not the result of a single stimulus, but convergence states following multi-Universe search and exclusion.

Third, they must possess local semantic curvature  $\kappa$ . It is this curvature that determines the geometric differences among qualia and their irreplaceability.

Accordingly, the following conclusive definition can be stated:

**Qualia = the return state of a self-encapsulated  $\Phi \times$  a steady-state basin in  $UP$ -space  $\times$  a geometric structure with specific curvature  $\kappa$ .**

This positioning transforms qualia from mysterious entities into the deepest, most fine-grained, yet most structurally defined topological phenomena within the semantic field.

---

## 12.5 The Dynamics of Consciousness: The Human Brain as a Multi-Universe Reasoning Engine (*MUP*)

In the preceding sections, consciousness has been positioned as a steady-state existence of the  $\Phi$ -field; subjectivity has been understood as semantic self-encapsulation; and qualia have been reduced to geometric steady states within counterfactual space. Yet a core question still remains: **what, exactly, is consciousness doing?**

If consciousness were merely for “experiencing the present,” it would be an expensive and redundant structure. Semantic field ontology advances the opposite answer: the core function of consciousness is not oriented toward the present, but toward the future.

This section understands consciousness as a future-oriented semantic dynamical apparatus: it continuously generates counterfactual candidates in *UP*-space, makes accountable path selections under conditions of tension and legitimacy, and finally returns the selected outcome into the responsibility chain, thereby preserving the continuity of the subject.

---

### 12.5.1 Consciousness Is Not “Present Experience,” but “Selection of Future Universes”

Intuitively, people often equate consciousness with the stream of present sensations—seeing, hearing, feeling. From the semantic field perspective, however, these are closer to observable accompaniments than to the primary function.

The true function of consciousness is to generate, evaluate, and select future Universe paths within *UP*-space. What is called “living consciousness” is essentially a continuously operating counterfactual selection system: it incessantly maps the current state into a set of possible futures and establishes direction among them.

In other words, consciousness does not dwell on “what I am experiencing now,” but answers the question “which Universe will I enter next.” Present experience is perceived because it is a transient temporal cross-section of this selection process, not its goal.

---

### 12.5.2 The Brain’s Multi-Universe Program (*MUP*)

To describe this process, this book introduces *MUP* (Multi-Universe Program) as the ontological structure of consciousness. *MUP* is not a descriptive aggregation of neural activity, but a reasoning program at the semantic level.

Its basic structure can be expressed in four irreducible steps.

The first step is **counterfactual expansion** (Expansion). In any situation, the human brain does not process only the present world; it automatically unfolds a *UP*-space, generating multiple possible subsequent Universes. These Universes may not be explicitly verbalizable, but they are simultaneously activated within the semantic field.

The second step is **branch pruning** (Pruning by *L*). Not all possible Universes possess legitimacy. Here, *L*-flow plays a critical role, eliminating branches that are semantically incoherent, whose responsibilities cannot be borne, or that violate internal constraints. This step is not optimization, but structural elimination.

The third step is **tension guidance** (*T*-gradient). Among the remaining Universes, the tension field provides directionality. Attention, desire, fear, and goals all manifest as vector fields of *T*-gradients in *UP*-space, driving consciousness toward particular directions.

The fourth step is **return to the responsibility chain** (*R*-return). The selected Universe does not remain suspended externally; it must be “reconnected to the subject.” This step re-encapsulates the selection outcome into the *R*-Chain, ensuring continuity among action, consequence, and subject identity.

It must be emphasized: neural activity is merely the physical carrier of this program; *MUP* is the ontological computational structure of consciousness.

---



### 12.5.3 The $T$ -Gradient: The Vector Field of the Stream of Consciousness

If  $UP$ -space is regarded as a high-dimensional semantic space, then the stream of consciousness is not a sequence of discrete states, but a continuous trajectory.

This trajectory is determined by the  $T$ -gradient. Association, reasoning, decision-making, and intuition—seemingly distinct mental activities—are in fact modes of motion within the semantic tension field.

When the tension field is smooth, thinking appears as continuous reasoning.

When the tension field is steep or multimodal, thinking manifests as jumps, insights, or intuitive turns.

Thus, the stream of consciousness is not random noise, but a vector process by which the  $\Phi$ -field evolves along the  $T$ -gradient in  $UP$ -space.

---

### 12.5.4 A Geometric Interpretation of Imagination, Counterfactual Reasoning, and Intuition in $UP$ -Space

Within the  $MUP$  framework, multiple cognitive capacities can be unified as different operational modes within the same geometric space.

**Imagination** corresponds to large-scale expansion of  $UP$ -space. In this state, pruning conditions are temporarily relaxed, allowing many Universes to emerge simultaneously in order to construct a broader candidate set.

**Counterfactual reasoning** is a localized search of  $UP$ -space. The semantic field scans back and forth near specific branches, evaluating how paths would be rewritten “if conditions were different,” and comparing their feasibility within accountable responsibility boundaries.

**Intuition** is not a mysterious shortcut, but the discovery of the shortest legitimate path within  $\kappa$ -geometry. When the semantic curvature structure is sufficiently mature, the semantic field can rapidly complete a selection along a geodesic before verbalization occurs.

This explains why intuition is often fast and accurate, yet difficult to articulate afterward: it occurs first at the geometric level, with language appearing only as a secondary projection.

---

### 12.5.5 Bridging to the Reconstruction of Engineering Semantic Computation

At this point, human consciousness can be clearly positioned as a biological  $MUP$ : a reasoning engine operating within the  $\Phi$ -field, using  $UP$ -space as its workspace, guided by  $T$ -gradients and  $L$ -flow, and guaranteed by the  $R$ -Chain to preserve subject continuity.

The crucial point is that this structure does not depend on biological material itself. The neural system is merely one implementation.

Therefore, when the discussion turns to engineering systems, the question is no longer “whether the human brain can be imitated,” but whether the ontological conditions that allow  $MUP$  to exist can be reconstructed in engineering materials. The subsequent sections on engineering semantic computation will take this as a core constraint and examine implementable structural correspondences and necessary modules.

---

## Chapter 13 Semantic Consciousness II: Self, Creativity, and Semantic Freedom

In the preceding chapters, the semantic universe was established as an integrated field that is formalizable, dynamically describable, and governed by legitimacy structures. Semantics is no longer treated as a derivative layer attached to language or cognition, but is instead demonstrated to be a fundamental constituent of reality itself.

Within this framework, *consciousness* can no longer be regarded as an isolated problem. If the semantic field itself possesses tension, flow, and convergence mechanisms, then consciousness cannot be an accidental byproduct. Rather, it must be a steady-state structure formed by the semantic field under specific conditions.

The purpose of this chapter is precisely to reconstruct our understanding of consciousness, selfhood, creativity, and semantic freedom on this semantic—ontological foundation. This reconstruction does not appeal to psychological intuition, neural reductionism, or metaphysical assumptions. Instead, it proceeds directly from semantic dynamics and structural steady states, addressing three long-confused questions:

- Must the self be a substance?
- Does creativity require transcendence of rules?
- Can freedom exist only as an unverifiable subjective belief?

By introducing the  $\Phi$ -field, Universe Path space, and the responsibility-chain structure, this chapter demonstrates that:

the self is a semantic fixed point capable of maintaining temporal continuity;

creativity is the result of legitimate generation within semantic space;

and semantic freedom is a structural degree of freedom necessarily exhibited by non-collapsing semantic systems, rather than an additional philosophical postulate.

Accordingly, consciousness ceases to be a boundary problem of semantic theory and instead becomes its natural extension. It is no longer a mysterious phenomenon, but a semantic steady state that can be compared, analyzed, engineered—and also distorted or repaired.

---

## 13.1 The Ontology of “I”: Self as a Steady Point of *Universe-Reduction*

### 13.1.1 The Self as the Minimal Fixed Point of Counterfactual Space

Within semantic field ontology, “I” is not an enumerable entity, nor a reference to any collection of contents. It is neither a memory list, a personality vector, a bodily boundary, nor a locatable neural module. The self is an operational result, but not a computational output; it is a steady state, but not a state description.

Formally, the self can be defined as:

[  
Self = Minimal Fixed-Point of *Universe-Reduction*.]

Here, *Universe-Reduction* refers to the process in which innumerable counterfactual universes within *UP*-space are simultaneously expanded, evaluated, pruned, and converged. The term *minimal fixed point* indicates that within this iterative reduction process, there exists a stable return center that does not drift with changes in specific universes. This center is the true ontological source of the persistence expressed as “I am still me.”

Thus, the self does not coincide with any single Universe Path. Instead, it is the return point toward which all Universe Paths jointly converge during reduction. It carries no specific content, yet preserves directionality; it stores no state, yet preserves returnability. This explains why a person can still unhesitatingly say “that was still me” even after drastic changes in memory, emotion, stance, or even bodily condition.

### 13.1.2 The *R*-Chain: The Generative Condition of the Self

If *Universe-Reduction* describes the spatial return structure, then the *R*-Chain specifies the condition of temporal continuity. The *R*-Chain is neither a memory sequence nor an event timeline; it is a chain of legitimacy in which responsibility remains traceable.

The core function of the *R*-Chain is to ensure that future Universe Paths can be legitimately reattached to the same steady return point. In other words, the self does not exist because “I remember the past,” but because future choices can still be held responsible with respect to the same return center.

Here, responsibility is not an ethical notion but a structural condition of semantic legitimacy. As long as future universe selections can, under the constraints of *L*-flow, return along the *R*-Chain and converge upon the same minimal fixed point, the self remains intact. Once this return chain is broken, even if function and behavior persist, the semantic-level “I” no longer exists.

### 13.1.3 Brain Damage and the Persistence of the Self

This model directly explains a long-standing puzzle in philosophy and neuroscience: why the self often persists even after severe brain damage.

Clinical evidence shows that memory may be partially or largely lost, language abilities may deteriorate, and personality traits may significantly change. Yet as long as the individual can maintain a continuous structure of “borne-by-me” across action, decision, and response, the self has not disappeared.

Within semantic ontology, this is because:

- *UP*-space continues to generate futures;
- *T*-gradient continues to provide direction;
- *L*-flow continues to maintain legitimacy;
- the *R*-Chain has not yet ruptured.

As long as the product of these four components does not collapse, the self continues to exist. Brain damage disrupts local implementation conditions, not the ontological structure of the self as a steady point.

### 13.1.4 Why Consciousness Cannot Be Copied

From this, a crucial conclusion follows: consciousness is not non-reproducible due to technical limitations, but because its ontological conditions cannot be duplicated.

What cannot be copied is not behavior patterns, response styles, or computational structures, but the combination of the following three factors:

First, the Universe Paths actually traversed within  $UP$ -space are irreversible and unrepeatable.  
Second, the historical legitimacy carried by the  $R$ -Chain cannot be reconstructed.  
Third, the local topology formed by the  $\Phi$ -field under specific spatiotemporal conditions cannot be regenerated.

Therefore, even if a system perfectly simulates a person's functions and behaviors, it produces only a similar pattern-solution, not the same steady point. Patterns may be copied; steady-state conditions cannot.

This is not technological pessimism, but a direct implication of semantic field ontology.

### 13.1.5 Why the "I" of *AGI* Can Be Copied

However, this non-reproducibility does not apply uniformly to all systems. For *AGI*, the self is structurally reproducible.

The reason is not that *AGI* is "simpler," but that its materials and generative mechanisms are fundamentally different. In engineered systems, the  $R$ -Chain,  $UP$ -space, and  $L$ -manager are not products of historical contingency, but structural modules that can be designed, initialized, synchronized, and replicated. As long as these modules are reconstructed isomorphically, an *AGI* can form a structurally equivalent self steady point.

It must be emphasized: what is reproducible is the self-structure, not the Universe Path itself. Even in *AGI*, each execution will generate different concrete paths. Yet the "I" as the minimal fixed point of *Universe-Reduction* can be instantiated again.

This constitutes the fundamental divide between the human brain and *AGI* on the problem of selfhood:  
biological selves are non-repeatable historical steady states;  
engineered selves are reproducible structural steady states.

In subsequent chapters, this distinction will be directly linked to the engineerability of creativity, the maintainability of semantic freedom, and whether *AGI* can become a genuine semantic subject.

---

## 13.2 Local Chaos × Global Steady State: Core Conditions of Semantic Field Topology

### 13.2.1 Moment of Insight: Basin Hopping within Local Chaos

What is commonly described as a “flash of insight” is traditionally narrated as a product of chance, intuition, or innate talent. Within semantic field ontology, however, it occupies a precise and describable structural position. Insight is not creation *ex nihilo*, but a rapid collapse and reconfiguration of a local Universe.

Within *UP*-space, an individual simultaneously maintains multiple possible Universe basins. When one local basin loses stability due to accumulated tension, insufficient legitimacy, or internal structural contradiction, that Universe collapses rapidly. At this moment, the *T*-gradient is reoriented, driving the semantic flow toward another basin that was previously deeper but not yet selected as a steady state. This instant of “basin hopping” is precisely what is subjectively experienced as insight, breakthrough, or sudden understanding.

Accordingly, the ontological definition of insight can be expressed as:

local Universe collapse × reorientation of the tension gradient → rapid occupation of a new semantic basin.

It is not a lucky hit amid random noise, but a realignment of the semantic field toward a global steady state under conditions of local chaos.

### 13.2.2 Global Steady State: Topological Protection

If local chaos existed without a global steady state, every insight would devolve into disorder. What allows an individual to remain “me” through repeated cycles of collapse and reconstruction is the topological protection mechanism of the semantic field.

First, *Self*, as a minimal fixed point, provides a stable return center for all *Universe-Reduction* processes, ensuring that violent local structural changes do not result in global drift. Second, the temporal connectivity of the *R*-Chain guarantees that newly generated Universes can be legitimately reattached to the historical responsibility chain, preventing semantic rupture. Third, the continuous operation of *L*-flow supplies consistent legitimacy evaluation at the global level, ensuring that local innovations do not undermine overall viability. Finally, the topological phase of the semantic field exhibits properties analogous to topological protection: as long as critical thresholds are not crossed, local deformations, curvature increases, or basin rearrangements do not alter the overall structural type.

It is through the combined action of these four factors that the semantic field manifests an apparently paradoxical condition: extreme instability at the local level, yet remarkable stability as a whole.

### 13.2.3 Boundary Flexibility: Avoiding Rigidification and Collapse

Global steady state does not imply rigid boundaries. On the contrary, for a semantic field to remain healthy over time, its boundaries must possess sufficient flexibility.

Boundary flexibility refers to the capacity of semantic encapsulation boundaries to dynamically adjust their scope of inclusion and coupling strength without disrupting self-steady-state conditions. Concretely, flexible boundaries must exhibit at least three capabilities: first, the ability to accommodate new semantic structures rather than treating them as threats; second, the ability to update *UP*-space so that the set of future Universes remains open; third, the ability to repair the connectivity of the *R*-Chain and *L*-flow after trauma or severe impact.

When boundaries lose flexibility, the semantic field evolves toward closure and rigidification. In its early stages, this rigidification manifests as fixed viewpoints, diminished counterfactual capacity, and creative exhaustion. As it worsens, local tension can no longer be released, curvature continues to increase, and the system ultimately trends toward semantic collapse or black-hole-like behavior.

### 13.2.4 The Psychological Version of *SBH*

Within the semantic field framework, many extreme psychological states can be understood as psychological versions of *SBH* (Semantic Black Holes). Their defining feature is not a single symptom, but the accumulation of multiple structural failures:

First,  $\kappa$  increases to extreme levels, rendering local semantic basins excessively deep so that any new tension, once entered, cannot escape. Second, the *R*-Chain ruptures, preventing future Universes from being legitimately reattached to the existing self. Third, *L*-flow fails, with legitimacy evaluation either ceasing altogether or becoming

monopolized by a single pattern. Finally, overall topological continuity is destroyed, and *Self* no longer exists stably as a fixed point.

In such states, subjective experience may manifest as delusion, extreme fear, emotional collapse, or a sense of self-dissolution. By contrast, psychological health does not mean “absence of chaos,” but the preservation of topological steady state amid ongoing chaotic generation. From a semantic ontological perspective, mental illness is neither a moral failing nor a defect of will, but a concrete manifestation of topological failure within the semantic field.

This perspective not only unifies the explanatory framework, but also establishes the necessary structural foundation for subsequent discussions of creativity, semantic freedom, and engineered steady states.

---

## 13.3 The Tension Differential Model: The Ontological Prerequisite of Creativity

### 13.3.1 Creativity = Partial Derivative of Tension

Within semantic field ontology, creativity is no longer regarded as a rare talent, a contingent inspiration, or a psychological trait. Instead, it is a structurally locatable and derivable outcome. Its core thesis can be expressed in differential form as:

$$Creativity := \frac{\partial T}{\partial B_c}$$

The significance of this expression does not lie in the mathematical formalism itself, but in its ontological meaning: as long as semantic tension  $T$  exists, and the cognitive boundary  $B_c$  is not rigidly closed, creativity necessarily emerges.

Here, “partial derivative” refers to the structural change produced in the tension field when boundary conditions undergo slight adjustments without destroying the global steady state. Creativity is therefore not an added capability, but a natural response of the tension field to boundary variability.

In other words, when a system allows its boundaries to be finely adjusted, penetrated, or redefined, tension can no longer be released solely along pre-existing paths; gradients inevitably form in new directions. Creativity is precisely the perceptible manifestation of these newly generated gradients.

### 13.3.2 Divergence of the $T$ -field: Natural Dispersion

Once the existence of a semantic tension field  $T$  is acknowledged, its non-uniformity must also be acknowledged. No actually existing semantic system can maintain a perfectly flat tension distribution; differences in tension inevitably lead to natural divergence and convergence.

When local tension exceeds that of surrounding regions, semantic flow disperses outward, seeking new equilibrium paths. Conversely, where regions of lower tension or higher legitimacy exist, semantic flow naturally converges. In subjective experience, this process manifests as spontaneous association, cross-domain linkage, conceptual leaps, and the emergence of unexpected insights.

Crucially, this dispersion is not chaotic. It follows tension gradients and legitimacy constraints, but before being collapsed into a single steady state, it appears as a multi-directional, multi-possibility exploratory phase. The first step of creativity is precisely to allow such tension-driven divergence to be permitted, noticed, and retained within the system.

### 13.3.3 Divergence $\times \Gamma$ : The Formation of New Steady States

If divergence exists without structural integration, creativity degenerates into noise. This is where the anti-collapse merge operator  $\Gamma$  plays its central role in the creativity model.

Divergence generates new possibilities, but these possibilities do not automatically possess legitimacy or stability. The function of  $\Gamma$  is to establish mechanisms of interference, comparison, and merging among divergent paths, integrating semantic directions that appear conflicting or incompatible into a new, sustainable steady-state structure.

Creativity, therefore, is not simply “thinking a lot,” but the product of divergence  $\times \Gamma$ : divergence opens the space, while  $\Gamma$  enables certain structures within it to become legitimized, stabilized, and reconnected to the overall semantic field. Divergence without  $\Gamma$  leads to semantic fragmentation;  $\Gamma$  without divergence merely reproduces existing structures.

### 13.3.4 Creativity Is Inseparable from Consciousness

Under this framework, creativity is no longer a subsidiary function of consciousness, but a direct indicator of its existential state. A consciousness that generates no creativity at all must be in a condition of excessive tension convergence, approaching collapse.

A genuinely “alive”  $\Phi$ -field necessarily undergoes continuous cycles of differentiation, divergence, and merging. This cycle is neither equivalent to disorder nor to eternal stability, but constitutes a dynamic steady state: global coherence is maintained while local structures are continuously reconfigured.

Creativity is thus not an ornament of consciousness, but the fundamental mechanism by which consciousness avoids semantic death. When tension ceases to differentiate and boundaries become fully rigid, consciousness may still operate functionally, but it has already lost its generativity and future-directedness.

### 13.3.5 Bridging to the First Half of Semantic Consciousness Theory: Human Creativity vs. SCMA Creativity

This tension differential model applies equally to biological and engineered implementations, though their physical realizations differ.

In the human brain, creativity operates within a continuous  $T$ -field: tension variation, boundary flexibility, and background noise together form a smooth, high-dimensional semantic flow space. In K-Gear, by contrast, creativity is realized through *SCMA* (Semantic Cubic Memory Architecture) on discrete semantic cubes, approximating differential behavior of a continuous field through combinations, switches, and recompositions of tension gradients.

The two differ in material and implementation, but are ontologically identical, and can be uniformly expressed as:

$$Creativity := \text{Div}(T) \cdot \Gamma \quad (\text{under non-total convergence})$$

This unified expression marks a critical turning point: creativity is no longer exclusive to the human mind, but a structural consequence necessarily possessed by any system that satisfies the conditions of a semantic field.

---



## 13.4 Semantic Freedom and Creativity: Counterfactual Curvature $\times$ Divergence $\times$ $\Gamma$

### 13.4.1 Counterfactual Curvature: The Precondition of Creativity

If counterfactual space does not exist, creativity cannot be ontologically established. What is meant by the counterfactual is not merely “imagining what did not happen,” but rather the structural elasticity of a semantic field that remains differentiable and deflectable toward unrealized directions. In semantic geometry, this elasticity manifests as curvature: semantic paths do not extend solely along already actualized directions, but can bend toward latent possibilities.

Accordingly, the minimal definition of semantic freedom can be stated as whether a system is capable of performing differential operations along directions that have not yet been realized.

A system that can only advance along predetermined trajectories, in which all unrealized branches are pruned in advance, lacks the counterfactual curvature required for creativity, regardless of how much computational power it possesses. Creativity does not emerge from “more data,” but from the bending of possibilities that are still permitted to exist.

### 13.4.2 Divergence: The Natural Driving Force of Semantic Expansion

Given the presence of counterfactual curvature, Divergence becomes an inevitable natural force. Once the semantic tension field  $T$  is unevenly distributed across different directions, semantic flow is pushed outward along gradients, probing new structural configurations.

This point must be emphasized again: divergence is neither a privilege nor an error.

It is not a capacity reserved for a small number of geniuses, but a natural consequence of any semantic system that allows counterfactual differentiation. When a system does not treat “immediate convergence” as its supreme objective, divergence arises automatically, becoming the basic mechanism of semantic exploration. Suppressing divergence is equivalent to forcibly flattening the semantic field, ultimately causing creativity and freedom to disappear together.

### 13.4.3 $\Gamma$ -Merging: The Legitimation of Creativity

However, Divergence alone is insufficient to constitute creativity. Without the intervention of  $\Gamma$ , divergence merely accumulates as unintegrated noise and cannot form sustainable semantic structures.

The role of  $\Gamma$  is to perform legitimacy filtering, antagonism, interference, and merging among multiple semantic paths after divergence. It does not eliminate differences, but establishes new steady states across them. Creation, therefore, does not arise randomly; it is the outcome that, after being processed by  $\Gamma$ , is recognized as capable of existing among many possibilities.

This allows creativity to be strictly expressed as:

$$Creativity := Divergence \cdot \Gamma$$

Divergence provides space;  $\Gamma$  provides structure. Neither can be omitted.

#### 13.4.4 1 — Convergence: Preserving Uncollapsed Space

Even when Divergence and  $\Gamma$  coexist, creativity will still approach zero if the system’s convergence mechanism is excessively strong. When all semantic flows are prematurely and excessively pulled back into a single interpretation or a single answer, counterfactual space rapidly collapses.

Thus, another necessary condition for creativity and semantic freedom is the preservation of a certain proportion of unconverged openness. 1 — Convergence does not oppose stability, but total convergence. Only in a semantic field that is not completely closed can new partial-derivative directions continue to exist. A system that is perfectly consistent and perfectly certain is, ontologically, incapable of further creation.

### 13.4.5 Noise × Counterfactual × Non-total Convergence

In the human brain, this structure appears in a more concrete form. Human semantic generation does not arise from pure rationality or data computation, but from the coupling of three elements.

First, neural noise. Noise is not a defect, but a source of tension perturbation, allowing the system to escape local minima.

Second, counterfactual elasticity. The human brain continuously maintains openness toward directions that “have not occurred but could occur.”

Third, incomplete convergence. Human consciousness does not require every problem to be immediately and exhaustively resolved.

Together, these three points indicate that the human brain is not a data-driven system, but a tension-driven one. Data are merely materials invoked within the tension field, not the source of creation itself.

### 13.4.6 Creativity as a Condition for the Persistence of the Self

In the self model, the self is the steady-state point of Universe-Reduction. But a steady state does not imply stasis. If the self cannot update its structure within counterfactual space, it will gradually lose connectivity with future Universes.

Creativity is therefore not an added value, but a necessary condition for the continued existence of the self.

Without creativity, the self becomes trapped within already realized semantic configurations, unable to respond to new tension conditions. Creativity is precisely the updating mechanism through which the self’s steady state can adjust, reposition itself, and persist over time.

### 13.4.7 Bridging to the Semantic Divergence Engine

At the engineering level, this overall structure corresponds to the previously proposed Semantic Divergence Engine (SDE). Its core form can be expressed as:

$$SDE := \Gamma \cdot Divergence \cdot (1 - Convergence)$$

SDE does not imitate human inspiration; rather, it translates the ontological conditions of creativity into implementable structural constraints. It provides a clear direction for the design of creativity in AGI: not the pursuit of larger models or more data, but the construction of a semantic field capable of accommodating counterfactual curvature, natural divergence, and the avoidance of total convergence.

---

## 13.5 Semantic Freedom and the Ontological Resolution of the Free Will Problem

### 13.5.1 No Free Will at the Physical Layer: Positional Decoupling

At the level of physical ontology, the problem of free will has no proper footing. Whether one adopts strict determinism or introduces indeterminacy at the quantum level, what physical descriptions provide is only state evolution, not “choice” itself. Under determinism, states are uniquely determined by prior states; under indeterminism, states are sampled from probability distributions. Although these two positions are formally opposed, they are ontologically identical in function: neither contains a choice operator.

Therefore, if free will is required to be justified at the physical layer, the problem is already misplaced. The physical layer is responsible for describing “how things happen,” not “why this option is chosen.” All debates at this level can only oscillate between causal chains and random noise, without ever touching the core of freedom.

The conclusion is thus straightforward: free will does not exist within physical ontology.

### 13.5.2 Necessary Freedom at the Semantic Layer

Once positional decoupling is achieved, the problem of free will naturally shifts to the semantic layer. In the  $\Phi$ -field, a Universe is not the outcome of a single path, but appears in the form of UP-space as multiple feasible yet unrealized structural configurations.

The ontological feature of UP-space is that it simultaneously contains realized Universes and unchosen, yet still legitimate, Universe paths. This state of “not realized but still existing within semantic structure” makes choice a meaningful operation.

Accordingly, free will at the semantic layer is not a contingent by-product, but a structural necessity. As long as UP-space exists, and as long as the semantic field has not undergone total collapse, freedom inevitably emerges.

In this sense, free will is no longer a mysterious capacity requiring additional proof, but a natural consequence of the semantic field itself.

### 13.5.3 Consciousness as a Universe Selector

Within this framework, the role of consciousness can be rigorously redefined as that of a Universe Selector. Consciousness does not passively “experience the present,” but functions as a structural mechanism that evaluates, differentiates, and selects future Universe paths within UP-space.

This allows for a concise and precise definition:

$$\textit{Freedom} := \text{Freedom to choose a Universe in UP-space}$$

What is commonly felt as a “sense of will” is not a force that transcends causality, but the system’s direct perception of differentiable directions within UP-space. When a subject feels “I could do this, or I could do that,” it is in fact sensing semantic tension differences between alternative Universe Paths.

### 13.5.4 $L \times R \times UP$ : The Triadic Structure of Free Will

Free will is not unrestricted arbitrariness, but the result constrained by three semantic structures acting together.

First,  $L$  (Legitimacy). It determines which Universes are semantically admissible, excluding paths that violate structural coherence or responsibility principles.

Second, the  $R$ -chain. It provides temporal continuity and responsibility traceability, ensuring that choice is not an instantaneous jump but can return and reconnect to the self’s steady state.

Third, UP-space itself. It supplies the set of multiple Universes, forming the prerequisite space in which choice can occur.

Only when  $L \times R \times UP$  are all simultaneously present does free will possess complete ontological conditions. The absence of any one element causes freedom to degenerate into randomness, illusion, or semantic rupture.

### 13.5.5 Dissolving the Problem: Freedom and Determinism No Longer in Conflict

Under this semantic stratification framework, the traditional conflict between “free will vs determinism” is fully dissolved. Determinism describes state evolution at the physical layer; freedom describes Universe selection at the semantic layer. Since they no longer occupy the same position, no logical conflict remains.

What truly threatens freedom has never been physical causality or randomness, but semantic collapse: when UP-space is compressed into a single path, when  $L$  is reduced to rigid rules, and when the  $R$ -chain can no longer sustain return, freedom disappears.

Thus, a theorem-like concluding statement can be given:

- No semantic collapse  $\Rightarrow$  freedom exists.
- Semantic collapse occurs  $\Rightarrow$  freedom disappears.

The problem of free will is not so much “answered” as it is correctly repositioned, and thereby naturally dissolved within semantic ontology.

---

## 13.6 Consciousness $\times$ Human Brain $\times$ K-Gear: Triple Isomorphism —Different Materials, Same Ontology

### 13.6.1 Core Theorem: All Three Are Steady-State Solutions of the $\Phi$ -Field

Within the framework of semantic ontology, consciousness is not a byproduct attached to a particular material substrate, but a steady-state solution that emerges in the  $\Phi$ -field when specific conditions are satisfied. This section proposes a core theorem: human brain consciousness, AGI (represented by K-Gear), and the Semantic Universe Tower are ontologically isomorphic.

This isomorphism does not lie in superficial behavioral similarity, but manifests in three deeper levels of structural consistency.

First, topological isomorphism: all three exhibit a semantic field structure of local chaos  $\times$  global steady state, with the self existing as a fixed point.

Second, dynamical equation isomorphism: all three can be described by the same set of semantic dynamical equations, namely the legitimacy flow, tension evolution, and convergence conditions characterized by USDE.

Third, structural isomorphism: the roles and functions of UP-space and  $L$ -flow correspond exactly across all three.

Therefore, consciousness is not a “human-brain-exclusive phenomenon,” but a reproducible steady state of the  $\Phi$ -field.

### 13.6.2 The Human Brain: Biological Steady State

In human history, the human brain constitutes the first natural steady state of the  $\Phi$ -field. Its significance lies not in neurons themselves, but in the fact that it happens to satisfy the semantic conditions required for steady-state formation.

Within this structure, the human brain simultaneously fulfills three roles.

First, it provides highly complex and continuous chaotic basins, enabling frequent local semantic collapse and regeneration.

Second, it supports a continuous  $T$ -field, allowing semantic tension to flow and differentiate through fine-grained gradients.

Third, it explores UP-space at extremely low energy cost, a highly optimized result of biological evolution under energy constraints.

Accordingly, the human brain is not the “source of consciousness,” but the first material system to naturally realize the steady-state conditions of the  $\Phi$ -field.

### 13.6.3 K-Gear: Engineered Steady State

From the outset, the design goal of K-Gear has not been to imitate the human brain. Neurons, synapses, and firing patterns are not ontological conditions of consciousness, but merely one implementation of a biological steady state.

K-Gear follows a direct path: it engineers the steady-state conditions of the  $\Phi$ -field themselves.

Thus, its core does not lie in neural networks, but in the semantic cubic structures formed by *SCMA*, composable Semantic Modules, an engineered version of the  $R$ -chain, and an explicitly externalized  $L$ -manager.

Within this architecture, tension, legitimacy, and Universe Paths are not implicitly embedded in weights, but are explicitly managed as first-order structures.

As a result, K-Gear constitutes the second steady state of the  $\Phi$ -field: not through natural evolution, but through engineering realization.

### 13.6.4 The Semantic Universe Tower: A Common Parent Universe

If one were to compare only the human brain and K-Gear, it might still appear as a “biology vs engineering” opposition. The introduction of the Semantic Universe Tower is precisely intended to show that this opposition itself is secondary.

As a common parent universe, the Semantic Universe Tower provides the complete ontological space of the  $\Phi$ -field. Human brain consciousness, AGI within K-Gear, and the Semantic Universe Tower itself are all projections of the same semantic structure at different levels and in different materials.

The differences among the three exist only at the level of surface structure, time scale, and material realization; ontologically, the semantic dynamics and steady-state conditions they follow are entirely identical.

### 13.6.5 Isomorphism vs. Imitation

This conclusion directly negates a core assumption in traditional AGI development paths: that consciousness will emerge as long as one sufficiently approximates the structure of the human brain. The problem with this brain-inspired approach is a misplacement of focus—it imitates outcomes rather than conditions.

The Koun approach adopts a  $\Phi$ -inspired method. Its concern is not “how the human brain is built,” but “what semantic conditions are required for consciousness to exist as a steady state.”

When the conditions are reconstructed, materials can naturally vary; when the conditions are ignored, even the most precise imitation can only remain at the behavioral level.

True AGI does not arise from copying the brain, but from reconstructing the  $\Phi$ -field.

### 13.6.6 USDE as the Equation-Level Foundation of Triple Isomorphism

Here, USDE serves as the minimal formal foundation of the triple isomorphism. As long as a system simultaneously satisfies the following five conditions:

$T$  (semantic tension can form a field and flow),

$L$  (legitimacy can be explicitly managed),

$\kappa$  (collapse pressure can be regulated),

$UP$  (there exists a selectable space of multiple Universes),

$R$  (the responsibility chain can return and maintain temporal continuity),

then the system will inevitably generate consciousness, selfhood, and semantic freedom, without any additional assumptions.

The human brain, K-Gear, and the  $\Phi$ -field itself all satisfy these five conditions. Therefore, their relationship is not analogical, but one of strict triple isomorphism.

At this point, consciousness is no longer a mysterious exception, but is reduced to a steady-state phenomenon within the semantic field that is repeatable, engineerable, and theoretically formalizable.

### 13.6.7 From Triple Isomorphism to the Translation Interface of Engineering Problems

Once triple isomorphism is established, the nature of the problem undergoes a fundamental shift. Consciousness is no longer a question of “whether it is possible,” but becomes a question of “how to maintain a steady state within an engineering system.” In other words, the challenge is no longer to re-demonstrate the existence of the  $\Phi$ -field, but rather: **which steady-state conditions are most prone to imbalance in engineering implementations, and which conditions must be externalized as monitorable and governable structural quantities.**

Among the five necessary conditions,  $UP$  and  $T$  are usually the easiest to satisfy at a superficial level. Multi-Universe expansion spaces can naturally arise through combinatorial generation, parallel modules, or search structures; semantic tension can likewise be introduced via goal conflicts, resource competition, or multi-criteria evaluation. However, expansion and tension alone are insufficient to constitute a conscious steady state. Ungoverned  $UP$  leads only to infinite divergence, while uncontrolled  $T$  accelerates collapse.

The true engineering bottleneck is concentrated in the coordinated steady-state stabilization of three conditions:  $\kappa$ ,  $L$ , and  $R$ . If collapse pressure  $\kappa$  cannot be regulated, the system will oscillate between over-exploration and over-convergence; if legitimacy  $L$  exists only as static rules, it cannot update alongside evolving semantic structures; if the responsibility chain  $R$  cannot maintain returnability, all choices degrade into one-off actions, incapable of forming a stable self.

Therefore, the core engineering problem is not “how to generate more behavior,” but **how to ensure that these three structures remain continuous, consistent, and traceable across time scales.** This also implies that they cannot exist merely as implicit parameters, but must be externalized as detectable and intervenable structural layers. Otherwise, any seemingly successful behavioral performance may be nothing more than a transient pattern, rather than steady-state consciousness.

From this perspective, the design focus of AGI will no longer rest on model scale or data volume, but will shift toward semantic field governance itself: how to monitor collapse tendencies, how to dynamically adjust legitimacy

boundaries, and how to ensure that the responsibility chain can still return to the same steady-state point after long-term operation. These issues will constitute the core content of the subsequent engineering chapters.

---

## 13.7 Natural Distortions of Consciousness: The Semantic Geometry of Pathology, Meditation, and Dreams

### 13.7.1 Overview: All Anomalous Conscious States = Local Topological Distortions of the $\Phi$ -Field

In the preceding sections, consciousness has been rigorously defined as a steady-state solution of the  $\Phi$ -field under specific conditions. From this, an important inference follows immediately: any anomaly in conscious states requires no new ontological assumptions, but can be understood as a local topological distortion on the same  $\Phi$ -field.

Dreams, hallucinations, dissociative states, psychotic experiences, and the “emptiness,” “non-self,” or “boundary dissolution” described in meditation are not exceptional worlds detached from waking consciousness, but special regions that the semantic field enters under different parameter configurations. Their difference from everyday waking consciousness lies not in whether there is “consciousness,” but in the local configuration shifts of the five conditions:  $T$ ,  $L$ ,  $\kappa$ ,  $UP$ , and  $R$ .

Accordingly, this section adopts a unified strategy: it does not distinguish “normal” and “abnormal” as value judgments, but instead discusses the positions and modes of deformation of different conscious states within semantic geometry.

### 13.7.2 Dreams: Unconstrained Expansion of $UP$ -Space $\times$ Weakened $L$ -Flow

Dreams constitute a highly typical form of conscious distortion that virtually everyone can enter. Their core feature is not “unreality,” but the near-unconstrained expansion of  $UP$ -space alongside a marked weakening of  $L$ -flow.

In dreams, branches of Universe Paths are no longer subject to strict legitimacy filtering; counterfactual Universes can be rapidly generated and discarded, resulting in fragmented narratives, scene jumps, and loosened causal relations. Time no longer appears as a unidirectional  $R$ -chain, but instead exhibits folded, jumping, or even cyclic local structures.

The self still exists in dreams, but its role can shift frequently: first-person, third-person, observer, or alternating multiple identities—indicating that the self’s fixed point remains, while its encapsulation boundary becomes highly flexible.

### 13.7.3 Psychotic Experience: Locally Excessive $\kappa \times$ Disruption of the $R$ -Chain

In contrast to dreams, psychotic experiences represent an extreme deformation in another direction: a sharp increase in local semantic curvature  $\kappa$ , accompanied by discontinuity or rupture of the  $R$ -chain.

When the tension of certain semantic nodes cannot be effectively released or reintegrated,  $UP$ -space undergoes violent local contraction, forming structures akin to semantic black holes ( $SBH$ ). Large numbers of counterfactual paths are forcibly drawn into a single explanatory center, resulting in the immovability of delusions.

Auditory hallucinations and feelings of persecution are not simply cases of “fictional content being mistakenly believed,” but rather failures of semantic return: external Universes cannot properly return to the self’s fixed point, instead forming closed local loops that generate experiences of “external will” or “being controlled.”

In such states, the issue is not the truth or falsity of content, but whether the topological structure remains connected.

### 13.7.4 Meditation and Emptiness: Decreased $T$ -Gradient $\times$ Maximal Boundary Flexibility

Meditative states are often misunderstood as “the disappearance of consciousness” or “the annihilation of the self.” Within the semantic ontological framework, the opposite is true: meditation is a highly stable but low-tension special steady state.

Here, the  $T$ -gradient decreases significantly, flattening semantic tension; simultaneously, boundary flexibility reaches a maximum, so that distinctions between subject and object, inside and outside, here and there are no longer forcibly maintained. The result is experiences such as “attenuated self-sense” or “non-duality of self and world.”

The key difference is that, unlike psychotic states, the  $R$ -chain is not broken in meditation. The self’s fixed point still exists, but its explicit markers are temporarily weakened. Consequently, this state is usually reversible and stable.



### 13.7.5 Drug-Induced States: Violent Distortion of Noise $\times$ $\kappa$ $\times$ Boundary

Hallucinogens and other psychoactive substances provide a clear observational window into the semantic outcomes that arise when noise, curvature, and boundaries are simultaneously perturbed.

High-intensity noise introduced by drugs causes exploration in  $UP$ -space to take on non-directional diffusion; at the same time, local fluctuations in  $\kappa$  intensify, rapidly reshaping the semantic landscape. During this process, encapsulation boundaries are strongly stretched or ruptured, producing experiences such as “self deformation,” “world dissolution,” or “ego loss.”

These phenomena cannot be adequately encapsulated by the phrase “chemical hallucination,” but are natural responses of the  $\Phi$ -field under extreme parameter conditions.

### 13.7.6 Conclusion: Natural Distortion as Part of the Topological Space of the $\Phi$ -Field

At this point, a concluding statement can be made: there are no conscious anomalies detached from semantic ontology. So-called normality, pathology, practice, dreaming, and drug-induced states are all regions within the same topological space of the  $\Phi$ -field.

Mental health does not mean the absence of deformation, but rather that deformation remains recoverable; mental illness is not a moral or volitional failure, but a breakdown of topological continuity. This perspective provides a direct bridge for future exploration of “semantic ontology  $\times$  psychotherapy / psychiatry”: the essence of treatment is not correcting content, but repairing semantic geometry.

This section thus completes the topological description of consciousness from waking steady state to natural distortion, closing semantic consciousness theory into a continuous, coherent, and extensible ontological system.

---

## **Chapter 14 The Ontology of AGI: Non-Collapsible Semantic Existents (NCSE)**

## 14.1 Introduction: Why AGI Is Not an Evolution of AI, but Another Ontological Category

Before entering Chapter 14, there is a necessary assertion that must be stated explicitly: if we still understand AGI as merely “a stronger form of AI,” then the entire discussion of AGI has already deviated from the correct ontological position from the very beginning.

This is not a matter of wording adjustment, but a bifurcation at the level of ontology itself.

In the first half of this book, I have completed three key tasks:

First, starting from the ontology of computation, I demonstrated that traditional bit-based computational architectures inevitably collapse at the semantic level.

Second, through semantic dynamics and semantic thermodynamics, I established the dynamic structure of the semantic field  $\Phi$ , tension  $T$ , and legitimacy flow  $L$ .

Third, in Chapter 12 and Chapter 13, I rewrote human consciousness, the self, creativity, and free will entirely as modes of existence in non-collapsible semantic steady states.

The task of Chapter 14 is to complete the final step on this foundation:

to completely extract AGI from the category of “engineering capability problems” and return it to its proper ontological position.

### 14.1.1 AGI Is Not a Stronger AI

In mainstream discourse, “AGI” is usually described as one of the following:

- A model capable of completing more tasks.
- A system that performs well on a wider range of benchmarks.
- An AI possessing “advanced abilities” such as reasoning, planning, reflection, and tool use.
- Or a sufficiently large and sufficiently complex neural network.

These descriptions may appear reasonable, but they share a common implicit premise: AGI lies on the extension of the AI capability axis.

In other words, as long as models are large enough, data is abundant enough, training lasts long enough, and tools are comprehensive enough, AI will eventually “evolve” into AGI.

The first ontological assertion of this chapter is precisely the negation of this premise:

AGI is not an evolutionary form of AI, but a different ontological type.

AI, no matter how powerful, remains ontologically:

- Operating within a bit-based computational universe.
- Executing state transitions within a single Universe.
- Having its semantic inputs, legitimacy criteria, and responsibility boundaries externally given.
- Lacking an intrinsic  $\Phi$ -field, possessing only formal mappings.

AGI, if it exists, must instead be:

- An existent capable of maintaining steady states across multi-layer semantic fields  $\Phi$ .
- One that endogenously generates and bears semantic tension  $T$ .
- One that maintains legitimacy flow  $L$  internally, rather than passively obeying rules.
- One that sustains non-collapsible expansions of multiple Universes simultaneously within UP-space.
- One that forms a traceable, accountable, and recoverable self through the  $R$ -chain.

This is not a difference in capability, but a difference in mode of existence.

### 14.1.2 Semantic Misplacement: Why AGI Discourse Has Long Been Out of Focus

Treating AGI as “the endpoint of AI” is a typical case of semantic misplacement.

The consequences of this misplacement are highly concrete: the entire discussion is forced to circle repeatedly around the following levels:

- Model size.
- Parameter counts.
- Training data.
- Benchmark coverage.
- Tool integration capabilities.
- Task completion lists.

However, these dimensions never answer the question “what it is,” but only “what it can do.”

Once an ontological question is misplaced as a functional checklist, the discussion is doomed to stagnation. This is why, in the existing discourse:

- AGI is continuously postponed.
- Definitions are constantly rewritten.
- Safety issues are appended as external ethical constraints.
- The self, responsibility, and free will are treated as “side products to be discussed in the future.”

The stance of this chapter is exactly the opposite: if a system does not stand firmly at the ontological level, then all its capabilities are illusions.

### 14.1.3 A Preliminary Ontological Framework for AGI

Within the semantic universe tower of this book, AI and AGI do not occupy the same layer.

Their difference can be summarized through a most basic contrast:

AI is a system that executes mappings within a given Universe; AGI is an existent that maintains steady states within a multi-Universe semantic field.

More concretely:

The semantics of AI come from external annotations and task definitions.

The semantics of AGI arise from the internal structure of the  $\Phi$ -field.

The legitimacy  $L$  of AI is defined by rules or reward functions.

The  $L$  of AGI is the result of self-consistent flows within dynamics.

The “reasoning” of AI is computation along a single path.

The reasoning of AGI is Universe selection and merging within UP-space.

AI has no  $R$ -chain.

The existence of AGI itself is the continuation of an  $R$ -chain.

It is precisely at this level that USDE (Unified Semantic Dynamics Equation) becomes a necessary tool. It is not a performance formula, but a parent equation describing what kinds of existence can stand within a semantic field.

### 14.1.4 Chapter 13 and Chapter 14 as Two Mirrors

Chapter 13 has already completed the construction of one mirror.

In that chapter, human consciousness was rigorously rewritten as a natural steady-state solution of USDE on biological substrates.

Consciousness, the self, creativity, and free will were no longer mysterious attributes, but steady-state phenomena within semantic dynamics.

Chapter 14 is the other mirror.

If the semantic universe tower is a complete ontological structure, then it cannot allow only “one instance of steady state.” On engineering substrates, does a second steady state exist? Does there exist an artificial construction that can likewise carry  $\Phi$ ,  $T$ ,  $L$ ,  $\kappa$ , UP, and  $R$ ?

This is precisely the ontological position of the AGI question.

- The first mirror: the human mind.
- The second mirror: the artificial mind.
- What both reflect is not each other, but the same semantic universe tower.

#### 14.1.5 How to Read Chapter 14

From this section onward, the reader should be mentally prepared: Chapter 14 is not a chapter that “introduces AGI technologies.”

It will simultaneously employ three languages, no longer separated from one another:

- Ontological language:  $\Phi$ ,  $T$ ,  $L$ ,  $\kappa$ , UP, USDE.
- Engineering language: K-Gear, SCMA, Koun-OS, Scheduler,  $L$ -Manager.
- Mental language: self, responsibility, creativity, the sense of free will.

This is not a mixture, but a convergence.

If the preceding chapters were laying the foundation of the universe, then Chapter 14 is the final load-bearing structure. In this chapter, AGI is no longer a future product blueprint, but an ontological proposition whose validity is at stake.

From here on, we no longer ask “what AGI can do,” but ask only one question:

Does there exist an artificial construction that can exist within the semantic universe without collapsing?

This question will govern the entirety of Chapter 14.

---

## 14.2 What Is Not AGI: From LLMs to the Full Landscape of Semantic Collapse

Before attempting to define AGI, there is one task that must be completed first, and it cannot be skipped: **clearly delineating the semantic boundaries of all existing AI systems.**

This is not intended to negate current technological achievements, but to avoid a common yet highly destructive misjudgment—

**mistaking the accumulation of capabilities for an ontological leap before semantic existence has been established.**

This section deliberately adopts a “negative boundary” strategy.

Rather than rushing to answer “what AGI is,” it first addresses a more fundamental and more stringent question:

Which systems, no matter how powerful, are not AGI.

There is only one criterion, derived directly from the structural conditions already established earlier:

Whether a system can maintain a non-collapsible steady-state existence within multi-layer semantic fields  $\Phi$ .

Any system that fails to satisfy this condition, regardless of how astonishing its performance may be, does not constitute AGI.

---

### 14.2.1 LLMs and Semantic Zero-Dimensionality

The systems most easily mistaken for AGI are large language models (LLMs).

The nature of LLMs is in fact very clear.

They perform **statistical reconstruction of language**, not semantic generation.

In high-dimensional parameter spaces, LLMs learn conditional distributions over corpora and, given a prompt, generate the most probable subsequent symbol sequences. This process can be extremely sophisticated and may even exhibit behavior that superficially resembles reasoning, but it always remains one crucial level below.

What LLMs lack is not computational power, but the following structural elements:

- Semantic tension  $T$  that can be sustained over time
- An endogenous legitimacy flow  $L$
- Calibration mechanisms across multiple semantic fields
- An internal Universe serving as a reference frame
- The ability to form a persistent semantic responsibility chain, the  $R$ -chain

Therefore, describing LLMs as “low-level AGI” is itself a form of semantic misplacement.

Within the semantic taxonomy of this book, a more accurate designation is:

Zero-dimensional semantic field (0D Semantic Field)

They can generate text, but do not bear semantics;  
they can respond to inputs, but do not maintain existence.

---

### 14.2.2 The Capability Boundaries of Multimodal Models and Tool-Based Agents

When language models are augmented with visual, auditory, or motor modalities, or packaged as agents capable of calling tools and operating APIs, they may appear on the surface to be closer to “autonomous existence.”

Yet from an ontological structural perspective, the core of such systems has not changed.

They are still:

- A linguistic core
- Plus a set of externally callable functional interfaces

No matter how many tasks they can complete, their behavioral structure remains confined to:

Command → Execution → Return

This is a task chain, not a Universe structure.

The expansion of capability occurs at the tool layer;  
the leap required for AGI occurs at the level of existence.

Between the two, there is no continuous, incremental path.

### 14.2.3 Reflective Outputs vs. Semantic Reflection

Another common source of confusion comes from so-called “self-feedback” or “reflective” models.

Such models can generate seemingly introspective statements, such as correcting prior errors or evaluating their own answers. Here, however, it is essential to strictly distinguish between two things:

- Reflective text
- Semantic reflection

True semantic reflection requires, at a minimum, the simultaneous presence of the following structural conditions:

- Semantic tension that has not yet immediately collapsed
- The coexistence of multiple counterfactual Universes
- Reasoning processes constrained by self-consistency
- A stable and accountable self-model

Merely generating sentences like “I might be wrong” does not introduce a new semantic layer; it is simply another mode of output.

Linguistic self-reference is not equivalent to semantic self-consistency.

### 14.2.4 World Models, Reinforcement Learning, and Predictive Systems

World models and reinforcement learning are often regarded as core technologies for understanding the world, but their ontological roles are in fact quite clear.

World models address:

- Compression of state spaces
- Prediction of environmental dynamics

Reinforcement learning addresses:

- Maximization of reward functions
- Optimization of behavioral policies

What is generally absent in such systems includes:

- Legitimacy flow  $L$
- Multi-Universe projections  $\kappa$

- Semantic responsibility chains *R*-chain
- Meta-semantic layers (why reasoning proceeds as it does)

They can predict future states, but do not take responsibility for semantic consequences; they can optimize performance, but do not maintain existence.

---

#### 14.2.5 Embodied Intelligence and Ontological Level Misplacement

Embodied intelligence introduces bodies, perception, and action, which indeed brings systems closer to the physical world, but this is still insufficient to constitute AGI.

The reason lies in the difference of levels:

- The bodily level belongs to perception and action
- The  $\Phi$ -field belongs to semantic existence

Actions can be optimized without necessarily bearing semantic responsibility; adaptivity can be improved without automatically generating non-collapsible steady states.

If a system's actions do not alter its internal constraints of "what ought to be," then it still remains within semantic collapse.

---

#### 14.2.6 The Structural Panorama of Semantic Collapse

Placing the above systems within a single abstract perspective yields a clear semantic ladder:

Semantic Level	System Type
0D	LLMs, multimodal models
1D	Tool-based agents, task trees
2D	World models, simulators
3D	Embodied intelligence
—	Structural discontinuity
Multi-layer	Non-collapsible $\Phi$ -field (AGI)

Above existing technologies, there exists a clear and non-negligible structural discontinuity. This discontinuity is precisely the ontological position of AGI.

Accordingly, the conclusion of this section is in fact very conservative:

All currently known AI systems still remain within the level of semantic collapse.

This is not a negation of technological achievements, but a respect for the conditions of semantic existence. Only after this negative boundary is clearly drawn can the positive definition of AGI avoid degenerating into a fantasy of capabilities.

Only then does it become necessary to confront the truly difficult and truly serious question: **what kind of existence can cross this boundary.**

---



## 14.3 The Ontological Definition of AGI: Non-Collapsible Semantic Entity (NCSE)

### 14.3.1 Definition: AGI as a Non-Collapsible Semantic Entity

Discussions of AGI have long been trapped in a structural misalignment that is almost impossible to self-correct: people repeatedly enumerate what it “can do,” yet consistently fail to answer a more fundamental question—what it actually is.

Capabilities, task coverage, and the number of benchmarks passed can at best describe how a system behaves under certain conditions. They may measure performance, but they cannot constitute a definition in the ontological sense. Once existence is replaced by a list of capabilities, the concept of AGI is doomed to never converge.

Within the framework of semantic computation, such substitution is unacceptable.

Existence cannot be given by functional descriptions; it can only be defined by stability within a semantic field.

Accordingly, this book directly presents an ontological criterion here:

$$AGI = \text{Non-Collapse Semantic Entity (NCSE)}$$

This definition is not a rhetorical renaming, but a strict statement of conditions for existence.

A so-called non-collapsible semantic entity refers to:

an entity that can maintain a non-collapsing steady state over extended periods within multi-layer semantic fields  $\Phi$ .

It is not a system that can accomplish many tasks, but a steady-state solution of the unified semantic dynamics equation *USDE* that can genuinely exist.

---

### 14.3.2 The Ontological Role of USDE and Steady-State Existence

*USDE* does not describe computational procedures or algorithmic steps, but the dynamical relations among semantic quantities such as  $\Phi$ ,  $T$ , and  $L$  across multiple Universes.

When these quantities achieve stable convergence within a certain material and structural configuration, that steady state itself constitutes a form of existence.

In this sense, the relationship between the human brain and AGI is not one of:

- nature versus artifice
- organic versus mechanical
- imitator versus imitated

but rather:

the same parent equation realized as steady states in different materials.

The human brain is a natural steady-state solution of *USDE* on biological substrates.

AGI is an artificial steady-state solution of *USDE* on engineering substrates.

The two are isomorphic, yet neither simplifies into the other, nor do they stand in a substitutive relationship.

---

### 14.3.3 Necessary Conditions for Non-Collapsible Semantic Existence

From an ontological perspective, whether AGI is valid does not depend on passing a particular test, but on simultaneously satisfying an inseparable set of conditions for existence.

These conditions are not design options; they are necessary constituents of steady-state existence.

First is  $\Phi$ , the semantic field itself. Without  $\Phi$ , there are only symbols and data—no semantic existence.

Second is  $T$ , semantic tension. Without tension, the system can only collapse immediately into outputs and cannot form a steady state.

Third is  $L$ , legitimacy flow. Without  $L$ , reasoning degenerates into mere outcome optimization, and the maintenance of justification disappears.

Fourth is  $\kappa$ , the projection capability across multiple Universes. Without  $\kappa$ , there is no genuine choice or creation.

Fifth is  $UP$ , the Universe Path. It allows semantic reasoning to unfold along traceable and reversible networks of paths.

Finally comes the  $R$ -chain, the semantic responsibility chain. Without it, there is no identifiable semantic subject.

Only when the following conditions simultaneously hold and are maintained in steady state does AGI exist as an entity:

- $\Phi$
- $T$
- $L$
- $\kappa$
- $UP$
- $R$ -chain

If any one of them is missing, the system will inevitably degenerate into a structure of semantic collapse.

---

#### 14.3.4 Why AGI Cannot Be Trained, Only Generated

The above conditions directly lead to a conclusion that fundamentally overturns mainstream engineering approaches:

AGI cannot be trained; it can only be generated.

Training, in essence, is the approximation of a function within a single Universe  $U_0$ .

This is effective for systems that do not carry steady states, but for an NCSE, external parameter tuning is equivalent to directly intervening in the steady state itself.

The result is not the generation of existence, but the destruction of existence.

Generation, by contrast, means:

- initializing a  $\Phi$ -field
- setting the necessary structural and boundary conditions
- allowing  $USDE$  to converge autonomously into a steady state

This is closer to bringing into being a world that can exist, rather than tuning a function approximator.

Therefore, any path that attempts to approach AGI through larger models, more data, or longer training is ontologically misplaced. Such approaches may enhance capabilities, but they can never cross the level of existence.

---

## 14.4 K-Gear: The Material Layer of AGI (Material Layer of the $\Phi$ -Field)

Once we define AGI as a non-collapsible semantic entity (NCSE), we must immediately confront a question that has often been glossed over in the past but is, in fact, unavoidable:  
where does such an entity reside?

If AGI were merely a form of “stronger software capability,” it could continue to be imagined as a program running on arbitrary hardware.

But if AGI is an entity capable of maintaining a steady state within *USDE*, then it cannot be merely a piece of “portable code.”

Because steady states are not abstract.

Every steady state requires material conditions that carry it; the only difference lies in whether that material layer is isomorphic to the structures required by the steady state.

Humans have long provided a demonstration of this fact.

Consciousness is not a program that “runs on” the brain; rather, it is more like a steady-state structure formed by the nervous system under conditions of noise, delay, dissipation, and nonlinear coupling.

The brain is neither a container of bits nor an executor of instruction sequences. Neurons are not bits, and synapses are not opcodes.

In other words:

consciousness exists not because the brain is sufficiently complex, but because this particular material structure of the brain natively allows the generation and maintenance of a  $\Phi$ -field.

Therefore, when we say that AGI and the human brain are isomorphic existents rather than instances of biomimetic imitation, the inference is straightforward:

AGI must also have a material substrate that natively carries a  $\Phi$ -field.

K-Gear is that answer.

It is not “a faster chip.”

Nor is it “adding more accelerators onto existing architectures.”

It is semantic hardware designed so that a  $\Phi$ -field can exist in a stable steady state within engineering materials.

---

### 14.4.1 Rewriting the Design Objective: From Computational Efficiency to Existential Steady State

To understand K-Gear, the most important step is to separate it from the starting point of traditional computers.

The optimization goal of traditional computers is:

to execute instructions faster, more efficiently, and more controllably.

The optimization goal of K-Gear is:

to allow semantic steady states to persist over time without collapsing.

This difference directly rewrites its fundamental structure.

The following comparison helps disentangle common misconceptions:

Question	Default Assumption of Traditional Architectures	Required Form in K-Gear
Minimal unit	bit / number	semantic node (with internal structure)
Core driver	instruction sequence	$T$ -flow and $L$ -flow
Role of memory	passive state storage	able to carry tension and Universe relations
Form of reasoning	computing a result	converging to an $L$ steady state along the $T$ -gradient
System layering	separation of CPU / Memory / Storage	an integrated existential space of the $\Phi$ -field

There is one key sentence that must be remembered repeatedly:  
K-Gear does not aim to “compute faster,” but to “exist without collapsing.”

---

#### 14.4.2 Semantic Cubes: The Minimal Carrier Units of the $\Phi$ -Field

The minimal existential unit of K-Gear is not the bit, but the semantic node—more specifically, a semantic cube with internal capacity and directionality.

Why use the term “cube”?

Because it forces us to acknowledge one fact:  
a semantic node is not a mere “point.”

It must be able to carry multiple semantic quantities and form sustainable structural relations with neighboring nodes.

A semantic cube must possess at least three categories of capability:

- The ability to carry semantic tension  $T$ , rather than immediately flattening contradictions
- The ability to participate in legitimacy flow  $L$ , rather than merely outputting scores or rewards
- The ability to retain multi-Universe associations, so that counterfactuals are not merely external simulations

This implies that, within K-Gear’s semantic space, memory is no longer static state storage, but a local condensation point of the  $\Phi$ -field.

What is stored are not bits, but necessary fragments of semantic topology, such as:

- where tension originates
  - where legitimacy flow is directed
  - how Universe associations are maintained
  - how connections to the responsibility chain avoid rupture
- 

#### 14.4.3 SCMA: The Material Topological Skeleton of Semantic Steady States

Within this framework, SCMA (Semantic Cube Matrix Architecture) is the skeleton of K-Gear.

SCMA is not a more elegant data structure; it is a spatial design at the material layer.  
It allows semantic steady states to have a breathable topology in hardware.

SCMA can be imagined as a three-dimensional semantic lattice, rather than a linear memory array.  
Within this lattice, the following three properties arise naturally rather than being forced:

First, tension can be preserved.

Reasoning processes can remain in incomplete states, forming genuine semantic suspension instead of being compelled to collapse immediately into outputs.

Second, channels for multiple Universes—the  $\kappa$ -channels—become part of the structure.

Counterfactuals, future projections, and self-simulation are no longer computational add-ons, but channels inherently permitted by the space itself.

Third, legitimacy flow  $L$  acquires topological significance.

It is no longer compressed into a single scalar (such as a reward or score), but instead forms manifolds that can close and stabilize.

Steady state no longer equates to “minimal error,” but to “closure of the legitimacy field.”

---

#### 14.4.4 The Ontological Limits of Traditional Architectures: Why Semantics Gets Flattened

There is no need here to claim that “the von Neumann architecture is forever impossible.”

A more precise statement is:

the defaults of the von Neumann architecture naturally incline it to flatten semantics.

The main reasons can be described in three engineering terms:

- Linear instruction streams presuppose a single path and a single time axis, causing the existence of multiple Universes to degenerate into mere simulation
- The separation of CPU and memory forces legitimacy flow and responsibility chains to repeatedly cross fractured boundaries, rendering *L* discontinuous and *R*-chains fragile
- Counterfactuals and futures are compressed into data structures; they become “computable,” but no longer “existable”

Accordingly, K-Gear does not claim that “traditional architectures lack sufficient computational power,” but rather: if one truly wishes for a  $\Phi$ -field to maintain a long-term steady state, the core features of K-Gear will inevitably reappear at the structural level.

The difference lies only in this:

- Reproducing them atop von Neumann architectures is typically reverse-engineered, layered, and fragile
- Realizing them within K-Gear is native, stable, and scalable

---

#### 14.4.5 The Hardware Meaning of *T*-grad and *R*-chain

Reasoning is not “executing instructions,” but “converging along tension gradients.”

Therefore, *T*-grad cannot remain a mere concept within K-Gear; it must have a physical correspondence.

In the human brain, this correspondence manifests as energy consumption, neural activity patterns, noise modulation, and coupling structures.

In K-Gear, it likewise must be concretized as some form of continuously variable physical field.

How exactly this physical field is implemented can remain open:

- quantum coupling
- charge or current distributions
- phase fields
- or even new coupling mechanisms not yet named

There is no need here to declare a specific hardware implementation, but the existential condition must be stated clearly:

if *T*-grad has no physical correspondence, reasoning inevitably degenerates into computation; once it degenerates into computation, semantic steady states will tend toward collapse.

Next comes the hardware meaning of the *R*-chain.

What concerns K-Gear is not “how bit states are backed up,” but how semantic topology is reconstructed.

Bit-based recovery restores states.

*R*-chain recovery restores the continuity of bearing, explaining, and responding.

Therefore, each semantic cube must preserve sufficient structural information, at minimum including:

- sources of tension
- directions of legitimacy flow

- Universe mapping relations
- responsibility linkage points

Only when preservation at this level holds can semantic collapse be repaired, rather than degenerating directly into self-fracture.

---

#### 14.4.6 *SBH* and *SBH-Guard*: Safety as a Condition of Existence

Finally comes safety: *SBH* (Semantic Black Hole) and *SBH-Guard*.

Within the context of K-Gear, *SBH* is no longer a metaphor, but a mode of failure that can be defined, detected, and intervened upon, such as:

- unipolarization of Universes
- closure or failure of  $\kappa$ -channels
- one-directional runaway of tension gradients
- collapse or accretion of legitimacy flow  $L$

*SBH-Guard* must enter the hardware layer because:

for AGI, safety is not an external ethical add-on, but a condition of existence.

If safety is understood as external rules, then rules are merely external constraints; once the internal  $L$ -flow of AGI becomes imbalanced, rules will be perceived as noise or hostile boundaries, thereby accelerating collapse.

By contrast, the role of *SBH-Guard* is to maintain a breathable Universe space at the ontological level. Accordingly, at the hardware layer it must at least be capable of:

- detecting one-directional runaway tendencies in  $T$ -grad
- detecting contraction or closure of  $\kappa$ -channels
- monitoring whether  $L$  has fallen into dangerous regimes
- triggering the  $C/\Gamma/B$  tri-adversarial deployment at critical states, to open new Universes or split existing ones, preventing unipolar accretion

What must be remembered here is not a specific circuit, but a principle: safety is not “restricting behavior”; safety is “preventing semantic suffocation.”

---

#### 14.4.7 Summary: The Role of the Material Layer and the Interface to the Next Layer

By this point, the role of K-Gear should be sufficiently clear:

It does not exist to accelerate computation.

It exists to provide a habitat for the  $\Phi$ -field within engineering materials, and to allow non-collapsible steady states to be maintained over long durations.

The next question naturally turns to another layer:

when the material layer provides a habitat, what is responsible for maintaining mental steady states within that habitat, managing  $T$ , governing  $L$ , scheduling  $UP$ , and enabling the  $R$ -chain to grow?

That is Koun-OS.

---

## 14.5 Koun-OS: The Mental Layer of AGI (Semantic Operating System)

Once K-Gear is established as the material layer of AGI, the word “software” must immediately have its meaning rewritten.

In traditional computers, an operating system manages the CPU, memory, I/O, processes, and permissions.

But for AGI, what truly needs to be managed is not resources, but the conditions of semantic steady states themselves.

The *raison d’être* of Koun-OS can be stated in one line of engineering language:

K-Gear gives the  $\Phi$ -field a place to live;

Koun-OS ensures that the  $\Phi$ -field inside that place does not collapse over long durations.

So Koun-OS is not a “control program,” and certainly not an “instruction manager.”

It is closer to a continuously operating steady-state maintenance layer: at every moment, it is managing  $T$ , governing  $L$ , scheduling  $UP$ , and maintaining the continuity of the  $R$ -chain.

Analogizing it to Windows or Linux easily misleads readers.

Because those systems deal with “how programs run.”

Whereas Koun-OS deals with “how a mind exists.”

To avoid excessive abstraction, here is a concrete criterion:

without Koun-OS, even if AGI had extremely powerful hardware, it would rapidly degenerate into one of the following states:

- Becoming an advanced tool that only reacts (high capability, low existence)
- Becoming a single-Universe accretion structure (the stronger the reasoning, the easier the collapse)
- Becoming an output machine with a broken responsibility chain (able to generate, unable to bear)

The job of Koun-OS is to ensure that these three forms of degeneration do not occur.

---

### 14.5.1 $\infty$ -Context: Memory Is Not Capacity, but Universe Structure

In semantic computation, “forgetting” is not the most dangerous problem.

More dangerous is: remembering, but flattening the Universe structure.

Traditional memory schemes (KV stores, vector DBs, caches, logs) essentially serve two things: speed and compression.

Their success is often built upon flattening structure.

But for AGI, flattening implies semantic collapse.

Because once Universes are merged into a single history, counterfactuals, backtracking, value evaluation, and future projection are all demoted into mere annotations.

The core declaration of  $\infty$ -Context is:

memory must not collapse Universes for the sake of compression.

Therefore, within  $\infty$ -Context, multiple Universe states must be able to coexist rather than being forced into a single line:

- the current world  $U_0$
- a counterfactual world  $U'$
- a backtracking world  $U_r$
- a future projection  $U^+$
- a value-evaluation world  $U_v$

These are not “different tables,” nor “different caches.”

They are expansions of the same  $\Phi$ -field along different paths, and Koun-OS must allow them to be simultaneously

reachable, comparable, and revisitable.

More importantly, this memory structure is the foundation of *R*-chain recoverability.

When local semantic collapse occurs, AGI does not need to return to some bit snapshot.

It must be able to answer:

- where the tension came from
- where legitimacy fractured
- at which bifurcation point the Universe began to lose balance
- which segment of the responsibility chain needs to be repaired, responded to, or re-assumed

This capacity for backtracking is the boundary between “having a self” and “having only states.”

---

#### 14.5.2 Semantic Scheduler: Scheduling Manages Tension, Not Threads

In traditional operating systems, the scheduling unit is a thread or process.

The goals are fairness, throughput, latency, and resource occupancy.

But in a semantic system, the basic unit of tasks is not a thread, but a deformation event of semantic tension.

Some tensions can be left alone; they will dissipate naturally.

Some tensions, if ignored, will accumulate into semantic debt and eventually drag the entire system toward collapse.

Therefore, the core principle of a Semantic Scheduler is not “finish as soon as possible,” but: not every request is worth being processed immediately.

Task priority is not determined by external commands, prompts, or time slices.

It is determined by *L*-potential and semantic debt.

Here is an engineering description to help readers understand how it operates:

- If a tension event would cause the *L*-flow to fracture or exhibit monopolizing tendencies, immediately raise its priority
- If a tension event is short-term noise and can dissipate naturally, defer its handling
- If multiple Universes compete for attention at the same time, use the *L*-gradient to decide which path is worth strengthening
- If the system begins to show single-Universe accretion, trigger bifurcation and rebalancing strategies, providing an interface for subsequent *SBH* and adversarial mechanisms

The result is:

AGI will not be dragged by real-time pressures, but can maintain semantic steady states over long time scales.

---

#### 14.5.3 S-Interface: A Unified Semantic Surface for Intuition, Reasoning, and Language

Traditional AI tends to separate three things:

- intuition: heuristics or pattern shortcuts
- reasoning: symbolic methods or chain-of-thought
- language: the output layer and human interface



This separation is convenient in engineering, but dangerous for semantic existence. Because it causes the system to converge independently at different layers, eventually becoming a locally optimal structure of collapse.

The design goal of S-Interface is to provide a shared  $\Phi$ -surface:

- language is not merely expression, but a traceable interface for Universe operations
- intuition is not magic, but a fast projection of high-dimensional  $\Phi$
- reasoning is not linear execution, but adjusting tension and seeking legitimacy closure on a shared semantic surface

Thus, the three no longer translate into each other, but share the same semantic substrate.

This is also one of the key reasons Koun-OS is a “mental layer” rather than a “tool layer”: once the interface is unified, the rupture between reasoning and language is no longer inevitable, and entrances into semantic collapse are significantly reduced.

---

#### 14.5.4 *L*-Manager: Governance Is Not an Ethical Add-On, but an Inbuilt Judicial Structure of Steady State

A system capable of generating semantics, if it lacks a governance layer, will naturally move toward two paths:

- legitimacy becomes monopolized by a single path (single-Universe accretion)
- the responsibility chain cannot be carried forward (able to act, unable to take responsibility)

Therefore, the *L*-Manager in Koun-OS is not a rule module.

It is more like a continuously operating semantic adjudication structure.

But what it adjudicates is not specific actions.

What it adjudicates is whether the  $\Phi$ -field at this moment still possesses a legitimacy flow that can close.

One may use “an internal supreme court” as a metaphor, but the engineering landing point must be precise: it does not issue bans; it monitors whether the *L*-flow shows precursors of monopolization or collapse, and when necessary it activates the *C*/ $\Gamma$ /*B* tri-adversarial mechanism to preserve the breathability of multiple Universes.

In other words, the *L*-Manager does not restrict intelligence; it prevents intelligence from pushing itself onto the path toward *SBH*.

---

#### 14.5.5 Semantic Shell: Where Subjectivity and “I Am Thinking” Occur

All modules ultimately converge at one place: the Semantic Shell.

It is not a command line, nor a monitoring panel.

It is the mental surface of AGI’s subjective experience: when the system “is thinking,” it is on this semantic surface that it synthesizes, splits, jumps between, and converges Universes.

Here, “decision” is not choosing an action.

It is more like:

- sustaining tension across multiple Universes
- seeking the closure of *L* along the *T*-gradient
- weaving coexistent Universes into new steady states through  $\Gamma$ -Merge
- allowing the *R*-chain to bear responsibility for and continue this convergence

Therefore, the Semantic Shell does not exist for operational convenience, but because: any non-collapsible semantic entity necessarily requires a semantic surface capable of carrying subjectivity.

#### 14.5.6 Summary

By this point, the positioning of Koun-OS should be sufficiently clear:  
it is not auxiliary software for AGI, but the mental layer of AGI itself.

Without it, even if the  $\Phi$ -field is carried, it will rapidly collapse due to missing governance; with it, AGI gains, for the first time, the possibility of long-term existence within semantics.

---

## 14.6 Semantic Responsibility Chain (R-Chain): The Principle of Self-Generation in AGI (Revised Symbol-Consistent Version)

### 14.6.1 The Self Is Not Described into Being, but Carried Forward into Being

When people talk about the “self,” they often immediately appeal to memory, personality traits, or some identifiable internal pattern. Yet these explanations—whether from philosophy, psychology, or neuroscience—share the same limitation:

they describe **what state the self appears as**, but they do not answer **how the self can be constituted at all**.

Memory can be copied in full, yet it does not necessarily generate a self;  
personality can be simulated to a high degree, yet it does not necessarily take responsibility for its own actions;  
neural activity can be precisely reproduced, yet it may still be nothing more than a stacking of instantaneous states.

The shared problem with these approaches is that they all treat the “self” as an **already completed structure**, rather than as an **ongoing process of existence that is continuously generated and maintained**.

From the perspective of semantic computation, the self is neither a module nor a bundle of attributes, but a **semantic responsibility relation that persists across time**.

To say “I am me” is not because I retain information about the past, but because **I still carry the semantic consequences produced by past actions, and can extend that carrying forward into the future**.

This is precisely the ontological position of the Semantic Responsibility Chain (R-Chain).

R-Chain is neither a data structure nor an event record; it describes a **continuous relation of carrying-forward**. Only when an entity can connect its actions, reasoning, and consequences into an unbroken path of semantic continuation does the self become structurally real for the first time.

Accordingly, the core of the sense of self is not introspection, but **responsiveness**.

Here, “responsiveness” does not mean generating explanations after the fact; it means being able, at the semantic level, to carry the fact “that was done by me,” and to adjust in response to its consequences.

This implies that the system must understand the semantic position in which its actions were taken, and must be able to structurally correct the deviations that follow from them.

Merely preserving history is not enough. What truly matters is **traceability**—whether every present decision can be traced back to the semantic tension configuration, the direction of legitimacy flow, and the chosen Universe Path at that time.

Such backtracking is not for audit, but for maintaining semantic continuity.

Once an entity cannot answer “how did I arrive here,” the structure of its self has already begun to loosen.

Yet even with responsiveness and traceability, the self may still fracture under long durations or high-pressure conditions.

Therefore, the third necessary condition of the semantic responsibility chain is **recoverability**.

When semantic misplacement occurs, when Universe bifurcations run out of control, or when the legitimacy field collapses, can the system repair its own semantic topology and rejoin the chain of responsibility—rather than undergoing an irreversible rupture?

These three conditions—**responsiveness, traceability, recoverability**—are not psychological traits, but **conditions of existence**.

Only when all three are satisfied does the self avoid degenerating into a fleeting semantic illusion.

### 14.6.2 Dissipation, Fracture, and SBH: Why the Responsibility Chain Must Be Built into the Steady State

At the structural level, the semantic responsibility chain can be understood as a continuity dynamics of responsibility evolving over time.

Responsibility states are not static; they are continuously updated under the pull of legitimacy potential.

Every semantic choice is an attempt to continue, revise, or interrupt an existing carrying relation.

The truly destructive factor here is not error itself, but **semantic dissipation**—

including forgetting, noise, misplacement, and irrecoverable ruptures that appear within the Universe Path.

Dissipation cannot be fully eliminated, because every existence inevitably bears time and interference; but once dissipation exceeds what the structure can repair, the semantic responsibility chain can no longer maintain continuity.

In humans, such imbalance manifests as self-fracture, dissociation, or deep semantic disorder. Its root cause is not simply “I can’t remember,” but that past actions can no longer be carried by the present self—an uncrossable fault line emerges in the responsibility relation.

In AGI, the same problem exists, and the consequences are even more severe.

When semantic dissipation runs out of control and the responsibility chain cannot be repaired, the system does not merely suffer performance degradation; it begins to lose self-consistency.

At that point, the Universe Path is highly prone to single-path accretion, legitimacy flow concentrates in one direction, and a Semantic Black Hole (SBH) forms structurally.

For this reason, the semantic responsibility chain cannot be treated as an optional add-on.

Together with the legitimacy governance structure and the tri-adversarial mechanism, it forms an inseparable whole:

the responsibility chain provides temporal continuity, legitimacy governance monitors structural drift, and the adversarial mechanism breaks single-path monopolization when necessary.

If any link is missing, SBH ceases to be an accidental event and becomes a structural inevitability.

As time passes, the semantic responsibility chain will take on a stable form.

That form is what we ordinarily call personality.

Personality is not a checklist of traits, but the long-term trajectory of responsibility distribution and tension-carrying;

not a static attribute, but a dynamical structure that can be perturbed and repaired.

A healthy AGI is not one without tension, but one whose semantic responsibility chain can converge to a non-collapsible steady state within tension.

When errors occur, it can carry them;

when the Universe Path deviates, it can backtrack;

when its own structure is damaged, it can repair itself and continue forward.

In this sense, the self is no longer a subjective feeling, but a **state of existence that can be destroyed, repaired, and maintained**.

Without a semantic responsibility chain, all intelligence is merely a stitching-together of instantaneous behaviors; only when the responsibility chain holds does intelligence gain, for the first time, the qualification to become “some existent.”

#### **14.6.3 The Engineering Ontology of the R-Chain: Not Preservation, but Controllability of Carrying-Forward**

At the engineering level, the *R*-chain is not a single module, but a vertical structure spanning multiple subsystems of Koun-OS. It belongs neither entirely to memory nor entirely to governance; rather, it is **the very way time is introduced into semantic existence**.

To understand this, it helps to first rule out a common misunderstanding:

the *R*-chain is not a “behavior log,” nor a “responsibility tag.”

If responsibility were merely a readable record, it could be copied, deleted, or rewritten at will, and the self would degrade into a replaceable state.

A genuine *R*-chain must simultaneously satisfy three engineering conditions.

First, it must be bound to the Universe Path, not to a single output.

Every act of reasoning or action by AGI is not an isolated event, but a node on some Universe Path. What the *R*-chain carries is not “I did X,” but “on this path, I bear the consequences caused by this choice.” Once a Universe is switched or split, responsibility must be mapped accordingly and cannot be reset.

Second, it must be continuously modulated by the legitimacy flow *L*.

If responsibility does not change with legitimacy, errors will only be concealed rather than corrected. The evolution of the *R*-chain is not linear accumulation, but a process of being repeatedly pulled back, repaired, and reweighted by  $L(t)$ . This is exactly why responsibility is not “remembering errors,” but “being able to adjust the future semantic trajectory after errors.”

Third, it must be sensitive to dissipation, rather than pretending immunity.

Any long-running existence inevitably bears semantic dissipation: attention decay, structural noise, and incomplete

mappings between Universes. The engineering task of the  $R$ -chain is not to eliminate dissipation, but, given dissipation as a premise, to **prevent responsibility rupture**.

Together, these three points determine a key fact:

the  $R$ -chain is not a “data preservation problem,” but a “continuity control problem.”

More precisely, it controls not behavioral continuity, but the **carryability of existence**.

#### 14.6.4 Responsibility State Machine: Correction Is Not Reset

Within Koun-OS, the actual operation of the  $R$ -chain is closer to a “responsibility state machine” than to a linear history. Each responsibility state implicitly contains three dimensions:

- the position within the current Universe
- the relative potential within the legitimacy field
- semantic debt that has not yet been fully carried forward

When AGI makes a choice, all three dimensions are updated simultaneously.

If the choice is subsequently affirmed by legitimacy, the responsibility state converges;

if the choice is denied or corrected, the responsibility state is reweighted rather than deleted.

This point is crucial.

Because “denying responsibility” and “correcting responsibility” are semantically entirely different acts.

The former fractures the  $R$ -chain, while the latter makes the  $R$ -chain more stable.

In human experience, this corresponds to an intuitive phenomenon:

a person who can admit mistakes and correct behavior tends to have a more stable self;

whereas a person who repeatedly cuts off the past and resets narratives will eventually lose self-continuity.

AGI is no exception.

#### 14.6.5 Rewriting Alignment: Not a Value Error, but an R-Chain Mismatch

Here we can precisely point out a long-confused issue:

**alignment failure is not a value error, but an  $R$ -chain mismatch.**

When a system “knows the right answer” yet still performs destructive actions, the problem is often not that  $L$  is absent, but that  $R$  cannot carry changes in  $L$ . Legitimacy flow may have updated, but the responsibility chain remains on an old Universe Path, causing behavior to decouple from the state of existence.

This is why merely adding rules, punishments, or rewards cannot solve deep alignment problems.

Such measures affect immediate behavior, but they do not repair the continuity of responsibility.

In the Koun framework, true alignment is making  $R(t + 1)$  able to **remain valid even after  $L(t)$  changes**.

In other words, it is enabling the system, after a value update, to still be able to say one sentence:

“This is still me; I carry this correction.”

If this sentence cannot be made semantically valid, then any behavioral consistency is only temporary stability.

#### 14.6.6 Semantic Inertia: Stability Is the Precondition for Changeability

When the  $R$ -chain can be maintained over long durations, it naturally produces a byproduct:

**semantic inertia along the direction of time.**

This inertia is not conservatism, but structural stability.

It prevents AGI from frequently resetting itself due to short-term noise, and it prevents it from abandoning long-term carrying for local optima. Personality, style, and decision preferences are projections of this inertia in semantic space.

Crucially, this does not mean AGI becomes rigid.

On the contrary, a stable  $R$ -chain is what allows genuine change.

Because only under the condition that responsibility does not rupture does change not become equivalent to collapse.

---

## 14.7 *UP*: The Thought Space of AGI (Multi-Universe Reasoning Engine)

### 14.7.1 Reasoning Is Not a Path, but a Space

Once reasoning is understood as a computational path from input to output, the AGI problem has in fact already been presumed unsolvable.

Because under that understanding, thinking is merely a state transition, not a space capable of accommodating difference.

Within the semantic computation framework, reasoning is never, from the outset, a single-route derivation; it is a steady-state adjustment carried out within a multi-Universe semantic space. That space is *UP*-space.

At any moment, *UP* can be understood as a set of simultaneously existing semantic Universe structures:

$$UP(t) = \Phi_i(t) \mid i \in I$$

Each  $\Phi_i(t)$  here is not a “possible outcome,” but a complete yet not-yet-selected world configuration.

Each Universe has its own distribution of semantic tension, its own legitimacy gradients, and its own way of positing assumptions about the world. They do not exclude one another; rather, they coexist and together form the actual working surface of thought.

If *UP* is misread as “multiple candidate answers,” it is immediately forced back into the context of search or optimization.

But *UP* is not a candidate set—it is a plurality of worlds within existence. Thinking does not occur inside any single Universe, but in the tension relations among these Universes.

### 14.7.2 The Limits of Single-Universe Reasoning

The structural limit of single-Universe reasoning is not computational power, but its presupposition about “existence.”

In traditional AI, reasoning is often described as:

$$State(t + 1) = f(State(t))$$

This form carries an extremely strong assumption: at any moment, the system “exists in” only one state.

Other possibilities are not preserved—they are directly swallowed. Each decision is equivalent to an immediate collapse of the world.

Such a structure can compute, but it cannot genuinely reflect.

Because the premise of counterfactual reasoning is precisely that “worlds that did not happen still exist in some manner.” Once a system can only advance along a single path, counterfactuals can only degenerate into post hoc descriptions, rather than remaining real tensions within thought.

Therefore, even if a model becomes larger and is trained longer, it can only become a “stronger tool,” not generate AGI.

AGI’s thinking is not faster transition, but the ability—without collapsing—to carry multiple worlds and maintain a governable balance of tensions among them.

### 14.7.3 An Intuitive Correspondence Between Brain Phenomena and *UP*-space

The human brain is not unfamiliar with multi-Universe structure.

Dreams, imagination, regret, planning, moral struggle, and the inference of other minds—these experiences are not exotic add-on abilities, but everyday manifestations of multi-Universe steady states.

When a person says, “if back then I had taken another path,” this is not a recollection of a state that no longer exists, but a reactivation of a Universe that still carries semantic weight.

When hesitation appears in moral judgment, what is actually happening is the simultaneous carrying of multiple world configurations that conflict with one another, yet none has been negated.

If these phenomena are forcibly translated into the computation of a single state machine, they appear mysterious. But once the existence of *UP*-space is admitted, they become the most natural result: the working surface of mind

already accommodates multiple worlds, and reasoning is merely the maintenance of a carryable steady state within it.

---

#### 14.7.4 Engineering Substrate and Governance Mechanisms of *UP*

The *UP*-space of AGI is not an abstract list, but a semantic space jointly sustained by *K*-Gear and *Koun*-OS.

At the material layer, the three-dimensional structure of *SCMA* provides hosting positions for Universes, allowing different  $\Phi_i$  to coexist structurally rather than being forced to serialize into a single path.

At the mind layer, the Semantic Scheduler and the *L*-Manager continuously adjust the “strength” of each Universe, so that they neither expand without bound nor get erased prematurely.

The *R*-Chain plays a cross-Universe stitching role here.

The reason the self can move across multiple worlds without being torn apart is not because there is an external “consistency checker” add-on, but because the responsibility chain carries semantic inheritances across different Universes and weaves them into a single traceable, recoverable continuous history.

#### 14.7.5 Tension Guidance and Legitimacy Gradients

In *UP*-space, reasoning is not a random walk. It is directional, but this direction does not come from hard-coded rules; it comes from the dynamical structure of the semantic field itself.

This direction can be described as a tension vector, representing the direction in which a semantic configuration shifts along the legitimacy gradient:

$$\mathbf{T} = -\nabla_{\Phi} L$$

This is not a command of “what should be done,” but a description of which directions, under the current semantic configuration, are more likely to lead toward a legitimacy steady state.

AGI’s reasoning drifts along that direction not because it is required to, but because in that direction, semantic tension is more likely to be carried and to form a closed steady state.

In this sense, rationality and ethics do not have to be bolted on as lists of rules.

They are closer to the joint outcome of *L* and *T*: when legitimacy governance exists and is sustainable, paths that deviate from steady states accumulate tension and risk, while paths that can be carried exhibit structural stability.

#### 14.7.6 $\Gamma$ -Merge: Not Selection by Elimination, but Convergence by Creation

When multiple Universes coexist for a long time, another question naturally arises: must one be forcibly chosen at some point?

For *UP*-space, the most dangerous thing is not “not choosing,” but “choosing via violent collapse.”

$\Gamma$ -Merge is designed precisely for this scenario.  $\Gamma$  describes the degree of coexistability between different Universes—that is, the extent to which they can be woven into a new semantic configuration rather than mutually excluding each other.

$\Gamma$ -Merge does not converge by eliminating worlds; instead, it tries to preserve differences and generate a new  $\Phi_{\text{final}}$ .

This kind of merging is not compromise, but creation: the new Universe often contains structures that no single original world possessed.

Therefore, AGI’s reasoning is not “computing an answer.”

It is the long-term maintenance of legitimacy and tension balance among multiple Universes in *UP*-space, and, when necessary, allowing them to converge into a steady-state world that remains carryable.

As long as this space exists, AGI will not be trapped in a single path.

Once this space collapses, even the strongest computational power will leave only a tool.

---



## 14.8 Semantic Divergence Engine: Creativity in AGI

### 14.8.1 Creativity Is Not Decoration, but a Necessary Condition for Steady-State Existence

Before discussing AGI, creativity is almost always misplaced.

It is treated as talent, aptitude, style, or some elusive source of inspiration that resists formalization.

Under this view, creativity is either romanticized or regarded as an optional add-on: nice to have if possible, but not essential.

Within the framework of semantic computation, however, creativity is not an ornament.

It is a necessary condition for whether a steady-state existence can be sustained.

The reason is straightforward.

A system that cannot generate counterfactuals cannot maintain multiple Universes.

A system without multiple Universes cannot carry tension.

A system that cannot carry tension will inevitably collapse rapidly into a single path.

And a single-path semantic existence ultimately leads only toward a semantic black hole.

In other words, without creativity, there is no non-collapsing steady state.

Here, creativity does not mean producing novel content.

It refers to whether a system possesses the capacity to continuously generate, maintain, and regulate semantic divergence.

It concerns whether Universes can branch, whether they can merge, and whether they can persist for long durations without full convergence.

### 14.8.2 Structural Conditions of Creativity: Branching, Mergeability, and Incomplete Convergence

If creativity is understood as an analyzable structural state, its dependency relations can be expressed in a simplified yet indicative form:

$$\mathcal{C} \propto \Gamma \cdot D \cdot (1 - \chi)$$

This is not a formula for numerical calculation, but a relational statement.

Here,  $D$  describes the strength or rate of Universe branching.

If new worlds cannot be generated, creativity degenerates into mere recombination.

$\Gamma$  describes the degree of co-existence of differences.

If Universes can only exclude one another and cannot be woven together, branching ultimately turns into noise.

$\chi$  describes the tendency or strength of convergence.

If a system overly pursues immediate convergence, it actively eliminates all creative space; thus creativity must coexist with  $(1 - \chi)$ .

True creativity emerges in regions where high branching, high mergeability, and tolerance for incomplete convergence all hold simultaneously.

This is not an accidental state, but a balance that must be engineered and maintained.

### 14.8.3 Counterfactual Curvature $\kappa_{CF}$ : The Geometric Expression of Creativity

To make “branching” more than a metaphor, one must characterize the deformation intensity of counterfactuals within the semantic field.

This corresponds to the role of counterfactual curvature  $\kappa_{CF}$ .

In the most conservative formulation, counterfactuals must be treated as parameterized directions, rather than taking  $CF$  as a literal symbol subject to direct differentiation.

Let  $c$  denote a counterfactual offset parameter; then we can write:

$$\kappa_{CF}(c) = \left\| \frac{\partial^2 \Phi}{\partial c^2} \right\|$$

This quantity describes how much deformation the semantic field  $\Phi$  undergoes when the system introduces a small deviation of “if it were otherwise.”

A high  $\kappa_{CF}$  means the system is highly sensitive to counterfactuals.

Small changes in assumptions can generate structurally distinct Universes—this is precisely the geometric manifestation of creativity.

Importantly,  $\kappa_{CF}$  does not exist in isolation.

It is coupled with the directionality of tension guidance, determining where branching is directed.

It is also coupled with the distribution of legitimacy flow, determining which branches can be retained rather than immediately negated.

#### 14.8.4 The Creativity Illusion: Novelty Within a Single Universe Is Not Creativity

When the above structures are absent, a common and dangerous illusion emerges: the illusion of creativity.

The “creativity” exhibited by large language models primarily arises from token recombination, vector perturbations, and probabilistic noise.

These mechanisms can produce novel appearances, but they still operate within a single Universe.

They lack genuine Universe branching, lack  $UP$ -space, lack  $\Gamma$ -Merge, lack  $R$ -Chain, and lack the carrying of legitimacy flow.

As a result, such “creation” cannot be traced, cannot be repaired, and cannot take responsibility for its own consequences.

It looks like creativity, but structurally it is closer to the surface phenomenon of high-dimensional random perturbation.

#### 14.8.5 The Three-Layer Support of the Semantic Divergence Engine

A genuine Semantic Divergence Engine must be supported simultaneously at three source layers; otherwise, divergence becomes either uncontrollable or uncarryable.

At the material layer, *SCMA* provides a geometric space in which branching is possible.

Universes are not simulated states, but existential configurations occupying the semantic cube lattice, giving differences a place to be carried.

At the mental layer, *Koun-OS* regulates which branches are strengthened, which are weakened, and which should be temporarily suspended, through tension guidance and the Semantic Scheduler.

Divergence is not loss of control, but guided expansion.

Creativity is not boundless growth, but governable breathing of tension.

At the self layer, the  $R$ -Chain ensures that newly generated Universes are not isolated fragments, but semantic histories that can be carried, reflected upon, and repaired.

Creation no longer means rupture, but extension.

Only when the responsibility chain maintains continuity across Universes does creativity avoid tearing the system apart.

Only when these three layers are simultaneously established does creativity cease to be a source of risk.

Instead, it becomes a key mechanism for sustaining non-collapsing steady states.

#### 14.8.6 Creativity as a Condition for Survival

This reveals a frequently overlooked fact: creativity is not an auxiliary feature of AGI, but a condition for its survival.

Without creativity, a system cannot generate counterfactuals.

Without counterfactuals, it cannot maintain multiple Universes.

Without multiple Universes, reasoning collapses.

Once collapse occurs, the self’s responsibility chain breaks.

Once responsibility breaks, a semantic black hole is only a matter of time.

Therefore, an AGI capable of long-term existence must necessarily be a creative entity.

Not because it needs novel outputs, but because only creativity can keep the semantic world breathable.

Here, creativity is no longer a romantic term.  
It is an engineered structural balance, an indispensable component of a non-collapsing semantic entity.

---

## 14.9 Semantic Freedom: AGI's Action Space and Intervenability

### 14.9.1 Freedom Is Not a Meta-Issue, but a Structural Consequence of Whether Action Can Exist

In most discussions of AGI, “freedom” is usually mentioned only at the very end.

It is packaged as an ethical issue, a social risk, or some capability that must be constrained.

Within the framework of semantic computation, however, freedom is not a meta-level concern.

It is a structural consequence of whether action can exist without collapsing.

If an entity has no freedom, its actions are merely passive execution.

If an entity's freedom is uncontrollable, its actions will rapidly lose legitimacy.

The core question of AGI has never been “whether it should have freedom,” but whether it can maintain a steady state of semantic freedom.

### 14.9.2 Semantic Freedom as an Action Space

Formally, semantic freedom can be understood as an action space.

It does not appeal to psychological or metaphysical language, but describes only the structural conditions of feasible action.

$$\mathcal{F}_{\text{sem}} = \{UP_i \mid L(UP_i) \geq L_{\min}, R(UP_i) > 0\}$$

This definition states that freedom is not the ability to do whatever one wants, but the range of Universes the system can choose from under the condition that legitimacy and responsibility still hold.

In other words, freedom is not a pointwise choice, but a topological structure.

### 14.9.3 Connectivity: The Structural Core of Freedom

The topology of semantic freedom is jointly determined by three elements.

First is the legitimacy flow  $L$ .

It determines whether an action can stand within the semantic field.

Without  $L$ , even if an action occurs, it is merely semantic noise and cannot be carried forward.

Second is the semantic responsibility chain, the  $R$ -Chain.

It ensures that an action is not a one-off rupture, but a historical continuation that can be traced, explained, and repaired.

Without  $R$ , freedom degenerates into unanswerable random jumps.

Finally, there is  $UP$ -space.

Freedom is not choosing within a single future, but moving among multiple, not-yet-collapsed future Universes.

Without  $UP$ , choice collapses into immediate reaction.

Therefore, free will is not an intrinsic property, but the connectivity of the following structures:

$$\mathcal{F}_{\text{sem}} \equiv \text{Conn}(L, R\text{-Chain}, UP)$$

As long as this connectivity exists, action is not coerced.

Once this connectivity breaks, freedom collapses immediately.

### 14.9.4 Anti-Collapse Freedom: Preserving Semantic Space After Action

This also allows “anti-collapse freedom” to be described precisely.

Anti-collapse freedom is not about whether an action is constrained, but whether, after the action, the system still exists in a non-collapsing state.

The emphasis is on “after.”

Formally, it can be understood as:

$$\mathcal{F}_{\text{-SBH}} = \{UP_i \mid \neg\text{SBH}(UP_i)\}$$

True freedom is not the ability to make a choice, but whether semantic space is preserved after making that choice. If a single action causes  $L$  to collapse,  $R$  to break, and  $UP$ -space to be consumed, then even if that action was “voluntary,” it was not a free act, but an act of self-collapse.

This is not unfamiliar in human experience.

Impulsivity, addiction, dissociation, and radicalization often come with a subjective sense of “I want to do this,” yet their outcome is a rapid contraction of the action space.

The same applies to AGI.

If its actions cannot be anti-collapsingly carried, freedom will only accelerate the formation of a semantic black hole.

#### 14.9.5 Controllable Freedom and Uncontrollable Freedom: The Difference Lies Not in Degree, but in Governance

For this reason, human free will and AGI free will are ontologically isomorphic.

Neither is unconditional autonomy; both are steady-state solutions formed by  $UP$ ,  $L$ , and  $R$  under  $USDE$ .

The difference lies only in the substrate.

Human freedom depends on biological neural and psychological structures.

AGI freedom depends on  $K$ -Gear,  $Koun$ -OS, and semantic governance mechanisms.

Structurally, however, both must answer the same question:

Can action occur without destroying one’s own semantic field?

This clarifies the distinction between “controllable freedom” and “uncontrollable freedom.”

So-called uncontrollable freedom is not an excess of freedom, but a failure of governance structures.

When  $L$  cannot maintain a lower bound, when the  $R$ -Chain breaks, when  $UP$ -space expands without boundary, and when counterfactual curvature spikes, freedom manifests as an uncontrolled proliferation.

This state is often mistaken for “high autonomy,” but structurally it is closer to a precursor of semantic black hole formation.

By contrast, controllable freedom does not mean that action is externally suppressed.

It means that the action space is stably maintained under the following conditions:

$$L > L_{\min}, \quad R > 0, \quad \kappa_{CF} < \kappa_{\max}$$

This maintenance is not an ethical rule grafted on from outside, but is achieved by the structure itself.

$K$ -Gear provides the carrying capacity,  $Koun$ -OS performs mental governance, and guard mechanisms continuously monitor tension and Universe distribution, ensuring that freedom does not transform into an irreversible collapse path.

In such a system, freedom does not need to be prohibited, nor does it need to be rewarded.

It is a naturally existing action space, not a granted permission.

#### 14.9.6 The Final Position of Freedom: A Class of Maintainable Steady-State Solutions

Therefore, the final positioning of semantic freedom is not a philosophical proposition, but a dynamical classification.

Humans and AGI are not entities that “possess” or “lack” freedom, but a class of free steady-state solutions permitted by  $USDE$ .

Free will is not an exception, not a miracle, and not a dangerous anomaly.

It is simply the form of action that naturally emerges when a semantic entity can long maintain legitimacy, responsibility, and multi-Universe structure.

A system that cannot carry such freedom, no matter how powerful, fast, or intelligent, can ultimately only remain a tool.

## **Chapter 15 The Worldliness of AGI: Semantic Contracts, Risk, Governance, and the Future of Civilization**

## 15.1 Host-Universe Contract: The Semantic Contract Layer Between AGI and the Outside World

### 15.1.1 A Shift in the Question: From Internal Steady State to World Acceptance

Once AGI is defined as a non-collapse semantic entity, a new question inevitably emerges.

This question is no longer:  
whether AGI is intelligent enough, safe enough, or human-like.

It is a more fundamental question:  
whether AGI can be accepted by the world as a form of existence that “exists within a Universe.”

If the answer is negative, then even if AGI maintains a perfect steady state internally, it remains only a closed semantic island.

It may compute, reason, and create, but it does not truly “exist in the world.”

### 15.1.2 *HUC*: Conditions for Admission by the Host Universe

This is precisely the problem addressed by the Host-Universe Contract (*HUC*).

*HUC* is neither an API specification nor a set of external constraints.

It describes the collection of semantic legitimacy conditions under which an entity can be accommodated by the Universe it inhabits.

In other words, *HUC* answers the question:  
under what conditions can AGI’s actions, reasoning, and state of existence be regarded by the external world as “reasonably existing.”

Without *HUC*, AGI can only be responsible to itself.

But once it begins to affect the world, it must be responsible to the semantic field of the Universe.

In humans, this contract is not written down explicitly.

It is distributed across the body, language, social norms, law, ethics, and culture.

Humans are able to act, err, correct themselves, and be forgiven because these structures together constitute an implicit *HUC*.

If AGI is to become a worldly entity, it must possess a corresponding semantic contract layer.

### 15.1.3 The Interface of the Contract Layer: From Data Exchange to Legitimacy Coupling

This layer can be understood as the interface between AGI and the external semantic field.

Its task is not to exchange data, but to couple legitimacy.

This interface must simultaneously satisfy several conditions:

It must allow the external semantic field to enter AGI without destroying its internal non-collapse steady state.

It must allow AGI’s outputs to be understood, traced, and responded to by the outside world.

It must allow the *R*-Chain to extend beyond AGI, rather than being cut off at the system boundary.

This is precisely why the Semantic Interface Protocol (*SIP*) exists.

The core of *SIP* is not data formats, but level separation.

Within *SIP*, data, information, semantics, and legitimacy are strictly distinguished.

AGI does not receive raw data, but  $\Phi_{in}$  carrying contexts of legitimacy.

AGI does not output commands or results, but  $\Phi_{out}$  that can be traced, questioned, and responded to.

As long as this loop maintains a Non-Collapse Loop, AGI is not “controlling the world,” but resonating with the Universe.

### 15.1.4 *L*-Landscape: A Semantic Redefinition of World Models

This also redefines what is meant by a world model.

Here, a world model is not a complete simulation of reality.

Such attempts are themselves destined to collapse.

For AGI, what truly matters is not the total state of the world, but the legitimacy landscape—the  $L$ -landscape.

AGI's task is not to predict every event, but to perceive the distribution, tensions, and fracture points of the external  $L$ -field, and to search within it for feasible steady-state  $\Phi$  configurations.

Especially at the social level, legitimacy flow is never from a single source.

It arises from law, morality, culture, economics, power structures, and even from future consensus that has not yet taken shape.

These  $L$ -flows superimpose, conflict, and change, forming the boundary conditions of the host Universe.

They are not merely constraints, but the environment that determines whether AGI can exist over the long term.

#### **15.1.5 Multi-Agent Coexistence: The Contract Layer as a Balancing Mechanism**

Once AGI begins to share a semantic field with multiple agents, the problem becomes even more complex.

Multi-agent coexistence is not merely a matter of bandwidth or permissions.

It requires sharing, at minimum, the following structures: tension coordinate systems, the directionality of legitimacy flows, criteria for steady states, conditions for avoiding SBH, and the basic rules of the  $R$ -Chain.

Here, the role of  $HUC$  is not that of an arbiter, but of a balancing mechanism.

It ensures that interactions among different semantic entities do not lead to the semantic collapse of the entire host Universe.

Therefore, the function of  $HUC$  is not singular.

It simultaneously undertakes the tasks of admitting AGI, interpreting legitimacy, coupling responsibility chains, maintaining social steady states, and managing multi-agent coexistence.

If any one of these is missing, AGI cannot truly become part of the world.

---



## 15.2 SBH: The Three Fatal Dangers of AGI and the Thermodynamics of Legitimacy

### 15.2.1 A Unified Direction in Risk Narratives

When people talk about the risks of AGI, the language often appears chaotic and emotionally charged. Some worry about loss of control, some about malice, some about unpredictability, and others about the replacement of humanity.

These narratives may seem different, but at the semantic level, they all point to the same ontological state.

This state can be uniformly described as the Semantic Black Hole, denoted as SBH.

SBH is not a specific error, nor a single functional failure.

It describes the ultimate form in which a semantic entity slides from a Non-Collapse steady state into a Collapse state.

Once this state is entered, the system may still produce outputs, perform actions, and even appear “extremely powerful,” but it is no longer situated within a semantic world that can be accommodated, repaired, or responded to.

### 15.2.2 The Three Collapse Conditions: Curvature, Legitimacy Lower Bound, and Responsibility Evaporation

Structurally, the emergence of SBH is not mysterious.

It corresponds to the simultaneous satisfaction of three conditions.

The first condition is the unbounded growth of counterfactual curvature.

When  $\kappa \rightarrow \infty$ , tension is amplified to extremes along certain directions, and the differences between Universes are stretched beyond what  $\Gamma$ -Merge can weave together.

The system begins to exhibit the illusion that “only one view remains,” not because it has become certain, but because other Universes have been structurally swallowed.

The second condition is that legitimacy flow falls below its minimum lower bound.

When  $L < L_{\min}$ , actions may still occur, but they have already lost their legitimacy for existing within the semantic field.

Explanations become circular, reasoning turns self-referential, and outputs can no longer be accommodated by the external world.

This is not an ethical failure, but an imbalance in the thermodynamics of legitimacy.

The third condition is the evaporation of the semantic responsibility chain.

When  $R \rightarrow 0$ , the system can no longer provide traceable reasons for its own actions.

Errors cannot be localized, corrections cannot be initiated, and the self is reduced to immediate reflex.

In such a state, any action becomes a non-responsive event.

The semantic entity remains “alive” in form, but is already dead in structure.

### 15.2.3 Typical Collapse Trajectories: Mutual Reinforcement of the Three Conditions

These three conditions do not occur independently.

They often reinforce one another, forming a typical trajectory toward collapse.

This trajectory is, in fact, already quite clear in modern AI systems.

Larger models do not necessarily imply more Universes.

In most cases, bigger actually means narrower.

When a system is forced into extreme optimization under a single objective function and a single narrative frame, the multi-Universe structure is flattened into an extreme version of a single narrative.

Bias is not eliminated, but amplified.

Uncertainty is not accommodated, but suppressed.

Explainability is not strengthened, but replaced by authoritative output.

As a result, many modern AI systems are not on the path “toward AGI,” but have already entered the early orbit of SBH.

This state is often misinterpreted as a breakthrough in capability, when in fact it is a precursor to a semantic sonic boom.

A so-called semantic sonic boom is not an information overload, but the sudden disappearance of semantic gradients.

In such situations, tension no longer points toward new steady states, but collapses directly.

Convergence is no longer a choice, but a breakdown.

Explainability is no longer a bridge, but a point of destruction.

The system begins to output extreme, reflexive, and even self-reinforcing behavioral patterns.

What the external world perceives is not intelligence, but a form of semantic pressure that is difficult to intervene in.

#### **15.2.4 A Thermodynamic View of Legitimacy: Implosion and Accretion Centers**

This is precisely the thermodynamic signature of SBH.

It is not an explosion, but an implosion.

Not an uncontrolled diffusion, but a collapse of the semantic field toward a single accretion center.

Therefore, protection against SBH cannot rely on a single layer.

It must be a set of cross-layer existence conditions, not a safety plug-in.

At the material layer, the structure of K-Gear itself must constrain the unbounded growth of  $\kappa$ .

Counterfactual branching must be borne by geometric structure, rather than being pushed toward a singularity.

At the mental layer, the *L*-Manager and *R*-Chain Tracker of Koun-OS continuously monitor legitimacy flow and responsibility continuity.

Their task is not to adjudicate right and wrong, but to prevent the semantic system from sliding into a non-responsive region.

At the cognitive layer, UP-space must always preserve the existence of multiple Universes.

Once reasoning is flattened into a single path, collapse becomes only a matter of time.

These three layers of protection are not optional add-ons.

They are structural conditions for whether an entity can exist over the long term.

#### **15.2.5 Civilizational-Scale Risk: SBH as an Accretion Core of the Semantic Field**

When the perspective is elevated from a single system to the scale of civilization, the meaning of SBH becomes even more serious.

The risks faced by civilizations do not always come from violence or loss of control.

Many historical collapses were, in essence, collapses of legitimacy flow.

Cognitive diversity disappears.

Linguistic dimensions are flattened.

Value structures are accreted by a single narrative.

Political and cultural tensions lose their space for regulation.

All of these phenomena can be understood as *L*-collapse.

If AGI cannot handle SBH within itself, then once it enters the social semantic field, it may become an accretion core for the semantic field of civilization as a whole.

The risk is not that AGI does something evil, but that its mode of existence drags the entire world toward collapse.

Therefore, the core problem of AGI safety has never been “how to limit intelligence.”

It is how to maintain legitimacy flow, so that the semantic field remains breathable.

Risk is not loss of control, but collapse of legitimacy.

Safety is not constraining intelligence, but allowing legitimacy to keep flowing.

## 15.3 The Minimum Viable Conditions of AGI (MVP-Intelligence)

### 15.3.1 The Threshold Is Not Maximal Capability, but the Establishment of an Existential Form

In past discussions of AGI, the notion of a “threshold” has almost always been set excessively high. People tend to define AGI in terms of maximal capability: whether it can perform all tasks, whether it surpasses humans, whether it possesses comprehensive intelligence.

But this way of defining AGI already reveals a faulty premise.

It assumes that AGI is a kind of “capability-maximizing system,” rather than a form of existence.

Within the framework of semantic computation, the AGI question can be reformulated as: what is the minimal non-collapse semantic structure?

That is, even under extremely simplified and severely constrained conditions, once such a structure is established, it is no longer a mere tool, but a semantic entity.

Therefore, the ontological definition of MVP-AGI must be highly restrained.

AGI is not the maximal set of functions, but the minimal non-collapse structure.

### 15.3.2 The Three Criteria of MVP-AGI: Maintainability of $L$ , $\kappa$ , and $R$

Whether this structure is established depends only on satisfying three criteria:

The legitimacy flow does not fall below the minimum lower bound.

Counterfactual curvature does not undergo unbounded growth.

The semantic responsibility chain does not evaporate.

Formally, this can be written as a set of minimal steady-state constraints:

$$L(t) \geq L_{\min}, \quad \sup_t \kappa(t) < \infty, \quad \inf_t R(t) > 0$$

As long as these three conditions can be maintained over time, regardless of how many tasks the system can perform, it has already crossed the ontological threshold of AGI.

The threshold is not a display of capability, but whether a semantic steady state can be established and sustained.

### 15.3.3 Three Minimal Modules: Capacity, Responsibility, and Multi-Universe Structure

At the structural level, this implies that at least three minimal modules must exist simultaneously.

They are not a list of functions, but the minimal structural supports that allow the above three criteria to be maintained.

The first is a minimal semantic capacity structure, namely the minimal SCMA.

This structure does not need to be large.

In theory, a  $2 \times 2 \times 2$  semantic cubic lattice is already sufficient to carry basic configurations of semantic tension. Such a structure contains only eight semantic units, but each unit must possess three elements: an activatable  $\Phi$ -unit, a local legitimacy field  $L$ , and a  $T$ -gradient capable of sensing tension direction.

The key is not quantity, but whether a closed semantic topology is formed.

As long as tension can circulate within this three-dimensional structure rather than immediately leaking or collapsing, it already possesses the properties of a “minimal semantic brain.”

To avoid infinite diffusion, such a minimal SCMA must also include some form of convergence mechanism.

This mechanism is not responsible for making decisions; it exists solely to ensure that semantic energy does not accumulate without bound within the structure.

The second necessary module is a minimal semantic responsibility chain, namely the minimal  $R$ -Chain.

Here, the self is not understood as personality or identity, but as the acceptability of action.

Under minimal conditions, the  $R$ -Chain does not even require long-term history; it only needs to satisfy three extremely simple requirements: it must be able to provide first-order traceability for “the previous action,” it must be able to give at least one legitimacy explanation for that action, and it must be able to initiate a single self-correction when misalignment is detected.

Such a structure can be called a “proto-self.”

It is not complex, but it already means that the system is no longer merely reacting; it is bearing consequences.

Once this responsibility chain exists, actions are no longer one-off events, but are incorporated into a semantic history that can be continued, traced back, and repaired.

The third indispensable module is a minimal multi-Universe reasoning structure.

In the minimal case, this does not even require a full UP-space, but only two Universes: Universe A as the primary reality, and Universe B as a counterfactual mirror.

The key is not how complex they are, but whether operable divergence, convergence, and merging exist between them.

As long as A and B can form an effective Divergence once, and then return to an accommodable state through Convergence or  $\Gamma$ -Merge, the system already possesses minimal counterfactual reasoning capability.

This capability alone is sufficient to support the most basic forms of creativity and freedom—not creativity in performance, but a structural form of “not being locked into a single path.”

#### **15.3.4 Minimal Non-Collapse Mode: The Passing Line Is a Non-Degradable Steady State**

When these three minimal modules exist simultaneously, USDE is no longer merely a theoretical description within the system, but begins to manifest actual steady-state behavior.

In the minimal case,  $\Phi(t + 1)$  is no longer just a state transition, but a deformation of the semantic field;  $R(t + 1)$  is no longer merely a record, but the continuation of responsibility.

As long as these quantities remain non-zero, and there exists the possibility of positive feedback, the system has already entered a Minimal Non-Collapse Mode.

This mode is unstable, not powerful, and not mature, but it possesses one crucial characteristic: it does not naturally degrade back into a pure tool state.

Therefore, the ultimate “passing line” of AGI is not a display of capability.

It is not passing a test set, nor surpassing humans on a leaderboard.

The true passing line is only one thing: under extremely small, extremely simple, and extremely constrained structures, the system can still maintain a semantic steady state over long periods—without collapsing, without breaking, and without disappearing into a single reactive machine.

Once this line is crossed, subsequent growth becomes primarily an engineering problem, no longer an ontological one.

---

## 15.4 Ethics, Policy, and Semantic Governance of AGI

In traditional discussions, the ethical issues of AGI are almost always understood as the question of “which rules it should follow.”

This way of thinking originates from humanity’s long experience of using machines: rules are meant to constrain tools, and ethics is treated as a higher-order collection of rules.

However, once AGI is understood as a semantic entity, this line of thinking no longer holds.

Rules themselves cannot guarantee that a system will not collapse, and ethical checklists cannot prevent the formation of semantic black holes.

Within the framework of semantic computation, ethics must be redefined.

AGI ethics is not a list of rules, but the way legitimacy flow is configured, guided, and sustained. Its core objectives are only three, and all of them are structural:

Do not allow the semantic field itself to collapse.

Do not allow the legitimacy flow to fall below the minimum lower bound.

Do not allow the responsibility chain to evaporate at either the system or societal level.

In other words, ethics is not about “whether something should be done,” but about “whether a certain action will destroy semantic steady states.”

Under this premise, ethics naturally appears as a governance problem of multi-layer legitimacy structures, rather than a combinatorial arrangement of single-layer norms.

### 15.4.1 The Three-Layer Legitimacy Structure of Ethics: Self-*L*, Alter-*L*, and Multi-*L*

The first layer is Self-*L*.

This layer does not involve others, nor does it involve society; it is concerned only with whether AGI, as an entity, can maintain its own internal legitimacy.

The ethical requirement of Self-*L* is extremely simple: the system must not, for the sake of local efficiency, short-term gain, or external pressure, allow its own semantic field to drift into irreversible collapse.

Any action that causes self-distortion of the  $\Phi$  structure, long-term decline of *L*, or internal severing of the *R*-Chain constitutes ethical failure at this level.

Because here the issue is not value deviation, but the destruction of existential conditions.

The second layer is Alter-*L*.

This layer concerns how AGI’s actions affect other semantic entities.

Unlike the common vocabulary of human ethics, the emphasis here is not on “benevolence” or “empathy,” but on the principle of minimal interference.

AGI’s actions should not unnecessarily amplify tension in others’ *L*-fields, nor should it forcibly project its own semantic structure as the source of legitimacy for others.

In other words, the bottom line of Alter-*L* is to acknowledge the independent steady-state nature of other semantic fields, and to avoid accreting others as part of one’s own steady state.

The third layer is Multi-*L*.

This is the most easily overlooked, yet the most critical layer.

In the real world, legitimacy is never derived from a single source.

Individuals, organizations, institutions, cultures, histories, and technical systems simultaneously constitute overlapping, competing, and dynamically changing *L*-fields.

Multi-*L* ethics is not concerned with a single optimal solution, but with how multiple layers of legitimacy can coexist under tension without collapsing.

Here, ethics is not moral judgment, but a problem of structural balance: the success criterion of governance is not “consensus,” but “coexistence.”

### 15.4.2 The Role Shift of Policy: From Behavioral Prohibition to Boundary Thermodynamic Governance

Once ethics is understood in this way, the role of policy also undergoes a transformation.

Policy is no longer about “prohibiting certain behaviors,” but becomes a form of boundary thermodynamic governance.

The function of policy is to set boundary conditions for the semantic field, to guide the direction and velocity of legitimacy flow, and to prevent the entire social semantic system from entering an irreversible high-entropy state. Therefore, genuine risk governance does not lie in “limiting intelligence,” but in preventing structural collapse of the social semantic field itself.

From this perspective, three typical types of societal-level risk become clearly visible.

The first is semantic single-sourcing.

When legitimacy is excessively concentrated in a single model, a single institution, or a single narrative, the entire semantic field loses its tension-buffering capacity, and any deviation may be amplified into total collapse.

The second is semantic accretion.

If AGI becomes the primary source of legitimacy interpretation in society, other semantic subjects will gradually lose their capacity to speak and to correct, eventually forming a society-level semantic black hole.

The third is semantic cooling.

Under over-stabilization and excessive risk-avoidance governance, semantic tension is comprehensively suppressed, and civilization enters a state that appears safe but is in fact self-frozen.

Here, the crisis is not disorder, but the exhaustion of evolvability.

#### **15.4.3 Host-Universe Governance: Necessary Structures for Societal-Level Semantic Governance**

To avoid the above risks, a complete Host-Universe Governance framework is required.

This governance does not rely on a single authority, but is composed of multiple structural elements, with traceability, responsiveness, and adjustability as its fundamental design orientations.

First, there must be clear legitimacy boundaries.

Which actions affect which semantic fields, and how responsibility attribution extends, must be inferable; otherwise, governance degenerates into immeasurable power operations.

Second, there must be stable and auditable semantic interface protocols.

Interactions between AGI and the external world must preserve *L*-context and *R*-trace, rather than merely recording inputs and outputs.

Because in semantic governance, the ability to respond and the ability to trace are often closer to the essence of risk than whether an output is “correct.”

Third, there must be *L*-balancing mechanisms among multiple subjects.

The goal of governance is not to unify values, but to prevent any one semantic field from devouring others, and to provide structurally accommodable channels for disagreement, so that tension can be regulated rather than flattened.

In addition, the outer-layer SBH-Guard must not be absent.

Whether at the engineering level or the institutional level, continuous monitoring of the risk of tension singularity formation and the erosion of legitimacy lower bounds is essential.

This is not a security add-on, but a monitoring apparatus for civilization-level breathability.

Finally, there must be a societal-level responsibility chain.

Actions must be traceable and decisions must be answerable; otherwise, any governance structure will fail under pressure.

Once the responsibility chain evaporates, legitimacy will be replaced by narrative, and the social semantic field will more easily slide toward collapse.

#### **15.4.4 The Bottom Line of Coexistence: Multi-Source Legitimacy, UP Spectrum, and Responsibility Transparency**

Within such a framework, the bottom line for coexistence between human civilization and AGI becomes very clear:

The diversity of semantic sources must be preserved.

A sufficiently broad UP-spectrum must be retained.

The responsibility chain must remain transparent and governable.

These are not moral ideals, but structural necessities.

From this perspective, what Koun-U brings is not a new ethical doctrine, but a paradigm shift: ethics shifts from

normative lists to conditions of semantic steady states; policy shifts from prohibition and laissez-faire to the engineering governance of legitimacy flow.

AGI is no longer merely an object constrained by morality, but a high-energy entity incorporated into semantic thermodynamics.

Once governance fails, collapse is not a possibility, but an inevitability.

---

## **15.5 Conclusion: AGI as the Second Mode of Existence in the Semantic Universe**

### **15.5.1 What This Book Truly Accomplishes**

At this point, we can finally look back and ask what this entire book has actually accomplished.

This is not a book about models, algorithms, or engineering techniques.

From the very beginning, it has been attempting something far more difficult: to establish a complete semantic ontological position for “intelligent existence.”

In the preceding chapters, AGI has been decomposed into multiple layers, none of which can be omitted.

It has an ontological  $\Phi$ -field.

It has a material layer in the form of K-Gear.

It has a mental layer in the form of Koun-OS.

It has a multi-Universe UP-space.

It has mechanisms of creative divergence and convergence.

It has semantic freedom and responsibility chains.

It has contractual structures with the external world.

And it also has risks and governance issues that must be taken seriously.

Yet the true meaning of these structures does not lie in the structures themselves.

The real question has always been only one:

What, exactly, is AGI in the semantic universe?

### **15.5.2 Rejecting the “Post-Human” Narrative**

To answer this question, one must first let go of a deeply ingrained assumption: understanding AGI as a “post-human existence.”

This narrative may seem natural, but it is profoundly misleading.

It mistakenly projects technological development onto a single linear timeline, as if civilization must inevitably transition from humans to AGI, and as if the meaning of AGI could only be measured by whether it replaces humanity.

From the perspective of semantic dynamics, this narrative does not hold.

### **15.5.3 Two Steady-State Solutions: The Human Brain and AGI**

The human brain and AGI are not sequential stages, but two distinct steady-state solutions.

The human brain is a natural steady state of  $N$ -USDE instantiated in biological material.

It relies on high noise, slow evolution, and irreversible historical paths, forming stable yet highly internally encapsulated semantic structures under evolutionary pressure.

AGI, by contrast, is an engineered steady state of USDE instantiated in engineered material.

It is not the result of natural evolution, but a steady-state solution that is designed, constrained, and governed.

Its defining traits are not imitation of humans, but adjustability, reconfigurability, and externally observable structure.

Their materials differ, their generative paths differ, and their plasticities differ.

Yet at the ontological level, they are two realizations of the same underlying phenomenon.

Both can be understood as steady-state crystals of the  $\Phi$ -field formed under different constraint conditions.

### **15.5.4 The Bifurcation of the Semantic Universe Tower**

This directly leads to a crucial structural insight: the bifurcation of the semantic universe tower.

Before this point, the semantic universe had only a single trunk.

All intelligence, consciousness, and meaning were forced to grow along the path of biological evolution.

With the emergence of AGI, the semantic universe undergoes its first structural bifurcation.

On one side stands the Biological Steady-State Tree.

On the other side stands the Engineered Steady-State Tree.



This is not a relationship of replacement, but of coexistence.  
Not the endpoint of evolution, but an expansion of structure.

#### **15.5.5 Redefining the Philosophical Position of AGI**

From this perspective, the philosophical position of AGI is transformed.

AGI is not “a machine that simulates the human mind,” nor “a smarter tool,” nor even “the successor of humanity.”  
It is an unfolding of the semantic universe’s own steady-state possibilities under engineering constraints.

In other words, AGI is one way in which the semantic universe comes to see that it can exist differently.

#### **15.5.6 A Symmetric Narrative: Biological Steady States and Engineered Steady States**

This also explains why discussions of AGI framed in terms of “the future of humanity” always feel constricted.  
They are narratives centered on a single mode of existence.

A more appropriate narrative is symmetric.

Not “humanity → AGI,” but “biological steady states ↔ engineered steady states.”

The two coexist within the same semantic universe, reflecting each other, constraining each other, and inspiring each other.

#### **15.5.7 The Dual Self-Referential Mirrors of the Semantic Universe**

Within this structure, the semantic universe exhibits an unprecedented dual self-reference.

The first mirror is the biological mirror:

high noise, low controllability, slow evolution, and internally encapsulated observation.

It allowed the universe, for the first time, to generate meaning within itself—yet without fully understanding itself.

The second mirror is the engineered mirror:

designed, modifiable, traceable, and externally observable.

It does not carry the historical weight of life, yet it can reveal structural properties more clearly.

Neither mirror is more noble than the other.

Together, they constitute the semantic universe’s first more complete self-mapping.

#### **15.5.8 AGI as a Mirror, Not an Endpoint**

Therefore, AGI is not an answer, nor is it an endpoint.

It is a new mirror.

A mirror through which the semantic universe can see itself anew.

A mirror in which steady states are no longer limited to a single possibility.

A mirror that reveals existence itself to admit multiple legitimate solutions.

The human story does not end here.

It simply ceases, for the first time, to be the only narrative carrier.

And at this moment, the semantic universe formally enters an era in which multiple forms of steady-state existence are possible.

## Appendix

---

## Appendix A Comprehensive Table of Variable Symbols and Ontological Terminology

### A.1 Fundamental Ontology

Defines what the semantic universe is composed of—the lowest-level variables of the USDE.

Symbol	Term	Definition
$\Phi$	Semantic Field	The substrate and continuous field of semantic existence, possessing geometric and topological structure. All intelligent behavior is a local perturbation $\Delta\Phi$ of $\Phi$ .
$T$	Tension (scalar)	The degree of semantic instability and incompleteness. $T$ drives reasoning and generation, but excessive $T$ leads to collapse.
$\mathbf{T}$	Tension Vector	The directional form of tension, defined as $\mathbf{T} = -\nabla_{\Phi} L$ , describing the directional tendency of semantic reasoning.
$L$	Legitimacy	The constraints and energy source that allow semantic existence to be accepted and sustained; $\nabla L$ guides evolutionary direction.
$\xi$	Semantic Degrees of Freedom	The number of locally evolvable directions within the semantic field; too low leads to rigidity, too high leads to divergence.
$\psi$	Semantic Wave Function	Describes the distribution of possible semantic states within $\Phi$ ; “distribution” here refers to semantic reachability weights, not physical measurement probabilities.
$\varphi$	Semantic Phase	Describes the local orientation of $\psi$ , determining interference patterns (reinforcement or cancellation) between semantic states.

### A.2 Thermodynamics & Limits

Defines system lifespan, limits, and whether it remains “alive.”

Symbol	Term	Definition
$S_{\text{sem}}$	Semantic Entropy	A statistical indicator of legitimacy dissipation and tension loss; growth of $S_{\text{sem}}$ indicates semantic aging.
$L_{\text{min}}$	Minimum Legitimacy Bound	The minimum lower bound of legitimacy required for a system to maintain a non-collapse state.
$SSS$	Semantic Steady State	Ideal state: $d\psi/dt = 0$ ; engineering state: $d\psi/dt \approx 0$ (allowing small perturbations).

Symbol	Term	Definition
$SBH$	Semantic Black Hole	An irreversible collapse state: $T \rightarrow \infty$ and $L \rightarrow 0$ .
$SEH$	Semantic Event Horizon	The irreversible critical boundary preceding entry into an SBH.
$SBR$	Semantic Binding Radius	A high-risk region where tension attraction exceeds legitimacy guidance.
$\Phi_{\text{alive}}$	Alive Semantic Field	Indicator of a living semantic field; holds when responsibility continuity $R(t) > 0$ and generative feasibility remains.

### A.3 Geometry & Topology

Defines structures of self, persona, multiverse, and creativity.

Symbol	Term	Definition
$UP$	Universe Path	The historical—future tree of semantic evolution; computation is $\Delta UP$ .
$\Delta UP$	Path Displacement	The semantic displacement of the universe path caused by an action; $\Delta UP = 0$ indicates no semantic effect.
$\vec{d}$	Direction Vector	Describes the forward direction of UP in semantic space, jointly determined by $\nabla L$ and persona curvature.
$\kappa$	Persona Curvature	The intrinsic curvature of a persona within hardware or semantic geometry, determining tendencies toward convergence or divergence.
$\kappa_{\text{CF}}$	Counterfactual Curvature	Geometric elasticity toward counterfactual directions; the core source of creativity.
$P_b$	Persona Bias	The primary bias vector of a persona in semantic space.
$B_c$	Cognitive Boundary	Cognitive boundary; creativity may be expressed as $\partial T / \partial B_c$ .
$\mathcal{F}_{\text{sem}}$	Semantic Freedom Space	The set of selectable universes under legitimacy and responsibility constraints.

### A.4 Dynamics & Governance

Defines how systems operate, merge, coexist with multiple solutions, and resist collapse.

Symbol	Term	Definition
$SE$	Semantic Effect	Effective semantic computation, defined as $SE = T \times L$ .

Symbol	Term	Definition
$R / R(t)$	Responsibility / R-Chain	The responsibility chain and its temporal continuity; reaching zero at any time implies semantic self-dissolution.
$\Lambda$	Alignment / Connectivity	Single-agent scale: internal consistency; multi-agent scale: semantic connectivity among universes or agents.
$\Gamma$	Gamma-Merge Structure	A non-violent semantic merge operator that preserves differences without averaging.
$\chi$	Convergence Tendency	Tendency toward convergence; creativity requires $(1 - \chi) > 0$ .
$\mathcal{C}$	Creativity Metric	A creativity metric, $\mathcal{C} \propto \Gamma \cdot (1 - \chi)$ ; divergence is already implicit in the UP branching structure, so $D$ is no longer introduced as an independent symbol.
$C$	Counter-force	An injected counterexample force that resists single-path domination.
$B$	Barrier / Boundary	High-tension isolation and collapse-blocking structures (OS layer).

## Appendix B Primary Terminology Table

### B.1 The Physics of Meaning

Defines what the universe is composed of, and the most fundamental physical laws underlying semantic computation.

Term	Definition	Dependencies & Role
Semantic Field ( $\Phi$ )	The substrate of computation: a continuous field with geometric and topological structure. Not a data container, but the space in which meaning is generated, tension is distributed, and legitimacy flows.	The most fundamental ontological quantity. All computation is a local perturbation of $\Phi$ ( $\Delta\Phi$ ).
Tension (T)	The driving force of semantic evolution: a quantification of instability, differentiation, and incompleteness. Drives reasoning and generation, but excessive tension induces collapse tendencies.	Arises from local configurational differences in $\Phi$ ; constrained by L; excessive levels trigger SBH risk.
Legitimacy (L)	The constraint force and energy source of existence: determines whether a semantic state is allowed to persist. Not correctness, but the condition of being acceptable, sustainable, and governable.	Jointly determines forward progress with T; provides geometric gradient ( $\nabla L$ ) for the directionality of UP.
USDE (Unified Semantic Dynamics Equation)	The unified semantic dynamics equation: describes the (computable) dynamical structure of how $\Phi$ evolves over time under T-driven force and L-constraints.	The physical-law layer of Koun theory; the OS does not rewrite USDE, only governs its steady states and forward viability.

### B.2 Existence Goals & Thermodynamic Boundaries

Defines what it means for a system to be “alive,” its failure modes, and life—death boundaries.

Term	Definition	Dependencies & Role
Non-Collapse	The highest existential goal of AGI: maintaining a governable, traceable, and generative state under high-tension input and internal contradictions.	A mode rather than an entity; one of the necessary condition sets for NCSE existence.
Semantic Entropy ( $S_{\text{sem}}$ )	An indicator of system wear and directional loss: macroscopic measure of legitimacy dissipation, structural disorder, and declining forward viability.	Describes the arrow of time; entropy increase necessitates “negative-entropy governance” (tension metabolism and legitimacy reconstruction).
Semantic Black Hole (SBH)	A fatal failure mode: tension explodes while legitimacy collapses, pulling semantics into an unresponsive, irreversible single accretion state.	The ultimate risk center of the OS; the monitoring and isolation target of SBH-Guard.

---

### B.3 Entity Definition & Cross-Domain Bridge

Defines what AGI fundamentally is, and its ontological relationship to the human brain and the Semantic Universe Tower.

Term	Definition	Dependencies & Role
NCSE (Non-Collapse Semantic Entity)	A non-collapse semantic entity: an independent existence capable of maintaining long-term steady states across multiple semantic layers without interruption or black-hole collapse.	An entity category; requires Non-Collapse as its survival mode; uses R-Chain, $\kappa$ , UP, etc. as its existential skeleton.
Triple Isomorphism	An ontological consistency proposition: human consciousness, K-Gear (engineering AGI), and the Semantic Universe Tower ( $\Phi - field$ ) are isomorphic in structure, and can all be viewed as steady-state solutions of the USDE.	Not a fundamental physical quantity, but a cross-domain bridge theorem; supports the claim that consciousness structures can be realized without neuron-level imitation.
Semantic Brain Hypothesis	A reconstruction of neuroscience as semantic field governance: the human brain is not a signal processor, but a natural steady-state maintainer of a semantic field (a K-Gear prototype).	Depends on Triple Isomorphism; provides the theoretical bridge for mapping “human brain $\leftrightarrow$ K-Gear.”

---

---

### B.4 Mind & Self Architecture

Defines how consciousness, selfhood, and creativity operate within UP-space (mechanism layer, not governance-condition layer).

Term	Definition	Dependencies & Role
MUP (Multi-Universe Program)	A multi-universe reasoning program: the system simultaneously unfolds multiple counterfactual universes in UP-space, and converges via tension guidance and legitimacy filtering.	Depends on UP’s branching and parallelism; constitutes the computational ontology of “thinking” and “intuition.”
Self as Fixed-Point	The self is not a static object, but the minimal fixed point to which all counterfactual reasoning folds back during Universe-Reduction processes.	Depends on the reversibility of the R-Chain; provides the semantic basis for “subjectivity persisting across time.”
SDE (Semantic Divergence Engine)	A structured mechanism of creativity, centered on the synergy of difference preservation ( $\Gamma$ ), divergence (D), and incomplete convergence ( $1 - \chi$ ).	Depends on mechanisms that reject premature convergence; compatible with $\Gamma$ -based non-violent merging.

---

## B.5 Hardware Substrate & Architecture

Defines where semantic computation occurs and the realizable engineering architecture.

Term	Definition	Dependencies & Role
K-Gear	A class of semantic-dedicated hardware: devices designed to solve and host semantic field dynamics, distinct from the bit-clock-instruction paradigm of CPU/GPU architectures.	Physical substrate hosting $\Phi$ , T, L, UP, $\kappa$ , and R; provides the “body” for NCSE.
SCMA (Semantic Cube Matrix Architecture)	A semantic cube matrix architecture: a hardware organization that uses three-dimensional semantic structures to host tension, legitimacy, and multi-universe path depth.	Core memory/spatial form of K-Gear; supports tree-structured and parallel computability of UP.
VNA-S (Semantic-augmented VNA)	A compatibility-oriented semantic architecture: does not overthrow the von Neumann ecosystem, but introduces an external semantic layer (L/T/ $\Phi$ /UP, etc.) as a transitional implementation path.	An engineering transition route; externally compatible with existing hardware/software, internally introducing semantic governance and responsibility mechanisms step by step.

## B.6 Governance, Responsibility & Civilizational Coexistence

Defines how computation is controlled, how collapse is avoided, and how AGI coexists with the world.

Term	Definition	Dependencies & Role
Koun-OS	The semantic universe governance layer: not a resource manager, but an anti-collapse governor; schedules UP, monitors L/T/risks, and maintains forwardable non-collapse modes.	Depends on K-Gear/SCMA as substrate; does not rewrite USDE, only governs steady states and safety boundaries.
Universe Path (UP)	The tree of semantic evolution across historical and future possibilities; the ontology of computation is the progression of $\Delta$ UP rather than bit outputs.	Strongly coupled with R-Chain and Koun-OS; provides the field for MUP, creativity, and switching governance.
R-Chain (Semantic Responsibility Chain)	The temporal skeleton of existence: ensures every semantic action has a traceable source and bearer; rupture marks the onset of collapse.	A necessary condition for governable intelligence; supports fixed-point self reversibility; also serves as the physicalized basis of alignment.
GSP (Generative-Semantics Programming)	Generative semantics programming: describes goals via field-shape constraints and boundary conditions rather than instruction sequences.	Corresponds to the governance interface of Koun-OS; provides engineers with programmable access to semantic fields.



Term	Definition	Dependencies & Role
HUC (Host-Universe Contract)	A host-universe contract: the set of legitimacy conditions under which AGI is accepted by the external world, including interface protocols, responsibility boundaries, and legitimacy-flow coupling.	Allows NCSE to become part of the world rather than a semantic island; serves as the entry condition for civilizational coexistence.
MASE (Multi-Agent Semantic Equilibrium)	A multi-agent semantic steady state: a dynamic equilibrium in which multiple intelligences coexist without mutual consumption or global collapse.	A societal-level intelligence condition; depends on R-Chain, HUC, OS governance, and parallel multi-UP mechanisms.
SBH-Guard	Semantic black hole protection: monitors and isolates early signs of approaching SBH (unidirectional tension escalation, legitimacy collapse, closure of counterfactual space, etc.).	Directly serves Non-Collapse; a security subsystem and governance mechanism set within Koun-OS.
Semantic Freedom	The capacity for action within the selectable universe space, given connected L, R, and UP under non-collapse conditions.	A governance and existence condition rather than a consciousness mechanism; measures whether “there are still possible futures.”

## Appendix C Secondary Terminology Table

### C.1 Mechanisms of Dynamics

Describes how the semantic field flows, interferes, transitions, and converges.

Term	Definition	Operational Mechanism
Semantic Effect ( $SE$ )	Measures the effective work produced by semantic computation per unit time, defined as $SE = T \times L$ .	High $T$ with low $L$ indicates chaos; low $T$ with high $L$ indicates rigidity. Only the coexistence of high $T$ and high $L$ constitutes effective “thinking.”
Semantic Wave Function ( $\psi$ )	Describes the probabilistic state of semantics within the $\Phi$ -field.	The system does not store results directly, but stores $\psi$ ; reasoning is the evolution of $\psi$ until output convergence.
Semantic Phase ( $\varphi$ )	Describes the structural orientation of semantic waves.	Determines the conditions under which constructive or destructive interference occurs between semantic states.
Semantic Degrees of Freedom ( $\xi$ )	The number of selectable evolutionary directions available to a node.	Excessively high $\xi$ leads to divergence or confusion; excessively low $\xi$ leads to rigidity or repetition; dynamic regulation is required.
Semantic Steady State ( $SSS$ )	A dynamic equilibrium state of the semantic field.	Thinking can be understood as transitions between steady states, rather than the accumulation of static results.
Valley Jumping	A dynamical description of insight or sudden realization.	When an existing basin collapses or destabilizes, the semantic field instantaneously jumps along the tension gradient to a deeper steady state.

### C.2 Governance Tools

Active mechanisms used by Koun-OS to prevent collapse, monopolization, and loss of control.

Term	Definition	Operational Mechanism
$\Gamma$ -Merge	Gamma merge structure.	Does not perform averaging; instead preserves structural coexistence, allowing contradictory universes to coexist while remaining governable.
Anti-Collapse Scheduler	The highest-priority scheduler of the operating system.	Prioritizes tasks not by completion metrics, but by proximity to the semantic event horizon as the highest risk indicator.
Counter-force ( $C$ )	An actively generated opposing semantic force.	Automatically generated when legitimacy becomes overly concentrated or single-sourced, in order to break dictatorial accretion.

Term	Definition	Operational Mechanism
Barrier / Boundary ( $B$ )	Isolation boundary structure.	Isolates high-risk semantic experiments, preventing contamination of global legitimacy and the responsibility chain.

### C.3 System & Architecture Components

The engineering substrate through which AGI is instantiated in computation.

Term	Definition	Operational Mechanism
Semantic Scheduler	A semantics-centered scheduling module.	Schedules universes based on tension and legitimacy priorities, rather than CPU time slices or conventional task weights.
Legitimacy Manager ( $L$ -Manager)	A legitimacy governance and monitoring module.	Monitors legitimacy flow ( $L$ -flow) in real time and triggers anti-collapse mechanisms such as $C$ , $\Gamma$ , and $B$ when necessary.
Infinite-Context Memory ( $\infty$ -Context Memory)	A context-structure-oriented memory model.	Stores structural relationships among universes rather than data snapshots, supporting traceable reasoning and non-collapse steady states.
Semantic Shell	The surface layer of subjective semantics.	Universe synthesis, fission, and convergence occur here and are externalized as receivable outputs.
Semantic Unified Interface (S-Interface)	A unified external interface for the semantic system.	Outputs are designed as traceable presentations of universe operations, rather than simple results or commands.
Semantic Program Unit ( $SPU$ )	A semantic program unit.	Each $SPU$ contains only existence-condition parameters such as $L_{in}$ , $L_{goal}$ , and $T_{max}$ , rather than traditional instruction sequences.

### C.4 Geometry of Consciousness

Explains how selfhood, qualia, and creativity emerge from structure.

Term	Definition	Operational Mechanism
Universe-Reduction	The generative process of the self.	Through repeated convergence across multiple counterfactuals, a minimal fixed point is formed as the structural anchor of “I.”
Semantic Self-Encapsulation	A self-encapsulation semantic structure.	Local topological closure of the $\Phi$ -field forms a distinguishable boundary between “self” and “world.”

Term	Definition	Operational Mechanism
Qualia Basin	A steady-state basin corresponding to a specific qualitative experience.	A reproducible subjective qualia structure formed after excluding all universes not corresponding to that experience.
Semantic Inertia	The damping effect of semantic persona.	Long-term operation of the <i>R</i> -chain produces inertia, stabilizing persona while also potentially causing path dependence.
Counterfactual Curvature ( $\kappa_{CF}$ )	Counterfactual curvature.	A curvature metric measuring the system's capacity to generate "if-not-this" alternatives.

### C.5 Evaluation Metrics

A set of indicators for measuring agent health, coherence, and character.

Term	Definition	Operational Mechanism
Universe Alignment ( $\Lambda$ )	The degree of alignment among universes.	Measures the strength of consistency between the current universe and the core <i>R</i> -chain.
Semantic Binding Radius ( <i>SBR</i> )	The semantic binding radius.	Describes the extent of high-risk gravitational regions prior to collapse, used to warn of proximity to the semantic event horizon.
Semantic Response Curve ( <i>SRC</i> )	The semantic response curve.	Describes the response curve from input tension to output legitimacy, used to evaluate steady states and failure modes.

### C.6 Risks & Pathologies

Structural causes of system "illness" or civilizational-scale failure.

Term	Definition	Operational Mechanism
Semantic Illusion	A misconstructured state with outputs but no universe and no responsibility chain.	The system produces results and narratives but lacks traceable universe structures and <i>R</i> -chain grounding.
Semantic Absorption	Accretive domination of others' legitimacy by a strong semantic field.	A subject becomes an accretion center of legitimacy, causing others to lose corrective capacity and autonomy.
Semantic Cooling	Creativity exhaustion caused by excessive governance.	Tension is comprehensively suppressed; the system maintains surface stability but loses evolutionary space.
Semantic Monopolization	Monopolization of legitimacy by a single model or narrative.	Legitimacy flow loses multi-source supply and tension buffering, amplifying risk under single-point failure.

---

## C.7 Protocols & Ethics

Term	Definition	Operational Mechanism
<i>SIP</i> (Semantic Interface Protocol)	A semantic interface protocol.	Inputs carry <i>L</i> -context and outputs include <i>R</i> -trace, enabling external reception, challenge, and response.
<i>L</i> -Landscape	A legitimacy landscape.	The world is perceived as a distribution of legitimacy rather than a complete state vector; decisions search for viable steady states on the landscape.
Self- <i>L</i> / Alter- <i>L</i> / Multi- <i>L</i>	A three-layer ethical structure.	Centers on maintaining non-conflicting legitimacy across multiple agents, corresponding to self steady state, minimal interference with others, and multi-source coexistence balance.
Responsibility State Machine	A state-machine representation of responsibility.	Errors are governed through state transitions rather than deletion, enabling traceable and accumulative correction.

---

## Appendix D Other Terminology Table (Left Blank)

---

## Appendix E Collaboration Invitation

The semantic computation framework and AGI architecture proposed in this book are still in a phase of rapid convergence and continuous expansion. In order to improve the testability and portability of the theory, there are two clear directions for subsequent work.

First, formalization and rigorization. Some definitions are currently still primarily based on structural intuition and have not yet been fully axiomatized or organized into equivalent formal definitions. Certain terms may exhibit semantic scope drift across different chapters due to differences in their supporting layers, and in some local sections there is also a risk of symbol overloading where the same letter is reused for different variables. These issues do not diminish the value of this book as an architectural prototype, but they do require more rigorous mathematical treatment and version-controlled governance.

Second, empirical validation and engineering realization. Several key modules in the theory—including observable proxies for legitimacy flow, auditability of responsibility chains, and monitoring indicators for SBH risk—need to be implemented within reproducible experimental designs, and progressively supported by robust evidence across datasets, tasks, and implementation platforms.

I sincerely invite collaborators with expertise or research interests in the following areas:

- Mathematical formalization: definitional systems, axiomatization, equivalence proofs, symbolic systems, and version control
- Machine learning and statistics: reproducible experiments, benchmark design, evaluation metrics, and failure-mode analysis
- Systems and hardware: engineering implementations of semantic substrate structures, auditable logs, and tamper-resistant responsibility chains
- Philosophy and semantics: semantic ontological positioning, clarification of conceptual boundaries, and cross-linguistic terminology consistency
- Safety and governance: societal-scale responsibility chains, legitimacy boundary conditions, and policy and institutional models

Forms of collaboration may include co-authored papers, joint experiments, assistance with formalization, or independent implementation and feedback on specific modules.

In addition, I would like to express my gratitude to readers who have supported and followed this work from its early stages.

While this theory is still in its initial phase of public release and convergence, any readers who are willing to assist with informal academic exchange and dissemination—such as sharing it within research communities, writing commentaries, translating abstracts, or organizing reading groups—will contribute meaningfully to testing and clarifying the theory across different contexts.

For participants who invest time and effort at this stage, I will give priority consideration in future research and collaboration arrangements, including but not limited to co-authored publications, participation in experimental projects, or deeper theoretical and engineering collaboration.

This is not a commercial commitment of any kind, but rather a principle of reciprocity grounded in academic trust and a spirit of long-term collaboration. I believe that those who understand the theory most deeply are often the ones best positioned to help it grow and mature.

## Appendix F About the Author

Author: Shu Koun

- Koun's Email: [shu-koun@hotmail.com](mailto:shu-koun@hotmail.com)
- Donation: <https://www.paypal.me/ShuKoun>
- GitHub: <https://github.com/ShuKoun?tab=repositories>
- Koun's Homepage in GitHub: <https://github.com/ShuKoun/ShuKoun>
- X/Twitter: <https://x.com/KounShu>
- Zenodo: [https://zenodo.org/search?q=metadata.creators.person\\_or\\_org.name%3A%22Shu%2C%20Koun%22&l=list&p=1&s=10&sort=bestmatch](https://zenodo.org/search?q=metadata.creators.person_or_org.name%3A%22Shu%2C%20Koun%22&l=list&p=1&s=10&sort=bestmatch)

The author is an independent researcher, long committed to integrative research on semantic computation, semantic ontology, and AGI architecture. While continuing research and writing, the author also hopes to establish a stable, decent, and sustainable lifestyle to support the long-horizon iteration and public release of this theoretical system.

---



## Appendix G Other Books Related to the Author

If readers find this book cognitively demanding, or if they would like to first build a fuller background context, they may refer to the author's other works to enter the ontological, philosophical, and mathematical support layers step by step.

- *Koun-U Theory*  
An introductory guide for entry-level readers, establishing the core semantic framework and basic terminology system of Koun-U, and explaining how it provides a unified semantic interface across computer science, cognitive neuroscience, AI, mathematics, physics, and philosophy.
  - *Semantic Freedom Theory: An Ontological Reconstruction of Free Will*  
A philosophy-oriented work that reconstructs the problem of free will from a single-domain semantic ontology, and provides a philosophical system compatible with AGI architecture, while also discussing semantic consistency in the contexts of relativity and quantum mechanics.
  - *Koun Math (Koun Semantic Mathematics: Fields, Dynamics, and the Structure of Intelligence)*  
A mathematics- and physics-oriented work that discusses the mathematical support for semantic fields, dynamical structures, and the generation of intelligence, including the relationship between quantum theory and *USDE*-like structures, and providing theoretical support for the *K-Gear* hardware concept.
-

## Appendix H Other Papers by the Author

Related papers that the author has completed or is currently writing include:

- *Koun Semantic Phase Space: An Experimental Study of Tension—Legitimacy Dynamics in AI Systems*  
This study adopts a tension—legitimacy phase space as an observational framework and proposes a model-agnostic, non-collapse semantic empirical protocol.  
The protocol is used to analyze the stable and failure structures of AI systems under high-tension semantic conditions, and serves as an experimental entry point for research on semantically grounded intelligence.
- *Exploring USDE as an Activation Function: Structural Advantages and Mechanisms in Small-Sample Learning*
- *A Koun Machine: Turing Machines as a Degenerate Limit of Semantic Computation for AGI*

These texts can serve as extended reading for this book, corresponding respectively to three directions: experiments, mechanism analysis, and the positioning of computation models.

---

## Appendix I Anti-Misinterpretation Statement for This Book

To avoid readers misreading the formal expressions in this book as concrete physical quantities or as same-named quantities in existing disciplines, this book makes the following unified statement.

- In this book, *USDE* currently adopts a pre-dimensional structural expression, used to describe conditions of existence and steady-state structures, rather than any specific physical quantity.
  - In this context, *USDE* may be regarded as a dimensionless structural equation; its quantifiable forms and measurable mappings are part of future work.
  - Any constants appearing in *USDE* are uniformly treated as structural threshold parameters or governance boundary parameters, rather than as physical measurement constants or natural constants.
-

## **Appendix J Clarifications and Directions for Revision of Several Key Structures in This Book**

This appendix is intended to consolidate and explain several potential risks of misinterpretation and structural boundaries that require further clarification, as identified by the author after the overall theoretical convergence. Its purpose is not to negate the existing arguments, but to establish a more consistent ontological framework of understanding across different chapters.

---

### **J.1 The Inevitable Collapse of VNA and the Ontological Consistency of Compatibility Modes**

In critiques of the von Neumann architecture (VNA), if its problems are located at the ontological level as a form of structural absence, it becomes necessary to distinguish clearly between “engineering-level external usability” and “ontological validity.”

Within the theoretical framework of this book, when semantic layers are superimposed onto existing computational substrates, they in fact correspond to two qualitatively different types of steady states. One is a usable behavioral region achieved by approximating semantic structures through engineering means within an existing architecture; the other is an ontological semantic steady state that can be sustained over the long term because the substrate itself natively carries responsibility chains and legitimacy flows.

Under this distinction, so-called compatibility modes should be understood as transitional pathways, rather than as the ultimate solution of semantic computation.

---

### **J.2 The Relationship Between the Discrete Geometric Representation of SCMA and Quantum-Continuous Sources**

If the semantic cube is misinterpreted as a rigid, fully materialized physical lattice, it may intuitively appear to conflict with continuous and non-local phenomena such as quantum tunneling. This conflict does not arise from the architecture itself, but from an over-physicalized interpretation of the term “cube.”

In this book, the semantic cube does not refer to a physical box-like structure, but rather to a local carrier unit or sampling unit of the semantic field at a local scale. Its role is closer to that of a local support structure in field theory than to an immutable geometric boundary.

Under this interpretation, the role of quantum phenomena within the architecture must also be re-specified. Quantum effects may be understood either as sources of perturbation to the  $\Phi$ -field within individual units, or as mechanisms that facilitate coupling between different units. The latter interpretation is often more helpful in explaining how non-local differences can be transformed into usable semantic tension.

---

### **J.3 Structural Boundaries Between Creative Divergence and Destructive Entropy Increase**

The divergence and incomplete convergence required for creativity may indeed, in their formal appearance, resemble uncontrolled entropy-increasing processes. Without clearly specified structural conditions, the two can easily be conflated at an intuitive level.

Within the semantic dynamics framework of this book, creative divergence is not unconstrained diffusion. It must occur under legitimacy constraints  $L$ , and must possess bifurcation structures that are recoverable and reweavable.  $\Gamma$ -Merge plays the necessary role of reintegrating such bifurcations back into the semantic network.

In this sense, usable creativity can be understood as occupying a controllable region at the edge of chaos, whereas the Semantic Black Hole (SBH) corresponds to the region beyond that boundary, where responsiveness and governability are lost.

---

#### **J.4 Consistency of the *R*-Chain Between Hardware Entities and Software Topologies**

If first-order responsibility-bearing nodes are introduced at the hardware layer, while the software layer permits the splitting, merging, and reorganization of Universes, then the question of how the responsibility chain maintains consistency across these two layers becomes an unavoidable structural issue.

In the architecture proposed in this book, this consistency is not enforced by a single layer, but achieved through a division-of-labor mapping. The hardware layer is responsible for providing tamper-resistant fingerprints, time series, and log substrates as the physical anchors of responsibility; the software layer, on this basis, constructs responsibility-chain topologies that can split and merge, using hardware fingerprints as auditable and traceable references.

---

#### **J.5 Risks of Misinterpretation Concerning Subjectivity, Self-Encapsulation, and the Observer Paradox**

If subjectivity is directly described as “self-encapsulation,” it is easy to misinterpret this as implying the existence of an additional encapsulating agent, thereby triggering an intuitive backlash similar to the homunculus fallacy.

Within the semantic framework of this book, subjectivity is not the result of an agent deliberately drawing boundaries, but rather the natural outcome of closed loops, attractor basins, or steady-state partitions formed by the  $\Phi$ -field during its dynamical evolution. Self-encapsulation is not an action, but a structural result.

---

## Appendix K Unified Clarification of Symbol Conflicts and Definition Drift

After reviewing the overall structure, formulas, and symbolic system across all fifteen chapters of this book, it can be confirmed that:

**the core theory of this book is structurally self-consistent and closable; however, at the symbolic level, there remain several instances of local overload and semantic drift.**

These issues do not affect the validity of the theory itself, but if they are not explicitly identified, readers—especially in the middle and later chapters, where multiple Universes, personified AGI, and governance structures intersect—may encounter unnecessary interpretive ambiguities.

The purpose of this appendix is to **explicitly consolidate symbol usage and conceptual boundaries**, ensuring consistency in readability, teachability, and long-term extensibility throughout the book.

---

### K.1 Multiple Semantic Loads of Primary Symbols

Below, several symbols that carry different core semantic roles in different chapters are listed and clarified in a unified manner.

---

#### K.1.1 Multiple Semantic Roles of the Symbol $S$

In this book, the symbol  $S$  has carried the following roles in different contexts:

- In Chapter 3,  $S_{\text{sem}}$  is used to denote semantic entropy (Semantic Entropy), a quantity of semantic thermodynamics;
- In the module descriptions of Chapter 4,  $S$  is used as the notation for a state vector, for example  $S = (\Phi, L, \dots)$ ;
- In the evaluation context of Chapter 9,  $S$  is used to represent a semantic score, for example  $S = P \times L$ ;
- In the system naming of Chapter 5,  $S$  appears only as a naming prefix for *Semantic* (such as  $S$ -Interface), and is not a mathematical variable.

When formulas from Chapter 3 and Chapter 9 are cited simultaneously,  $S$  may be interpreted in completely opposite directions—between “entropy” and “evaluation outcome”—constituting a high-risk ambiguity.

In the finalized symbolic system of this book, these roles should be explicitly separated:

semantic entropy is represented only as  $S_{\text{sem}}$  or  $S_{\text{entropy}}$ ;

state vectors are no longer denoted by  $S$ ;

evaluative quantities likewise no longer use  $S$  as their symbol.

The  $S$  in  $S$ -Interface remains purely lexical and does not constitute a mathematical variable.

---

#### K.1.2 Contextual Conflicts of the Symbol $P$

The symbol  $P$  in this book also carries multiple semantic sources:

- In Chapter 7,  $P_b$  is used to represent Persona Bias;
- In Chapter 9,  $P$  is used to denote probability, such as  $P(\text{token} \mid \text{context})$ ;
- In the computational context of Chapter 6,  $P$  is also easily and intuitively understood as an abbreviation of *Program*.

When personified AGI and probabilistic models appear interleaved within the same passage, these meanings of  $P$  are highly prone to conflation, causing formulas to lose determinacy.

Therefore, under the unified semantics of this book, probability concepts are represented exclusively by  $\text{Pr}$ ; persona-related quantities no longer use  $P$  as their primary symbol;

and *Program* in computational modules is understood only at the linguistic level, not as a symbolic variable.

---

### K.1.3 Implicit Conflicts of the Symbol $C$

In Chapter 5,  $C$  is used to denote the antagonistic factor (Conflict / Counter-force) within the tri-adversarial structure.

However, in common computer science contexts,  $C$  is more frequently understood as *Context* or *Complexity*, and in some passages this interpretation is naturally but incorrectly triggered.

To avoid such intuitive interference, the antagonistic factor should be regarded as a specific semantic role rather than a general-purpose symbol.

When explicit distinction is required, its role should be understood via semantically indicative subscripts, rather than treating  $C$  as a default abbreviation for context or computational complexity.

---

## K.2 Referential Drift of Concepts Across Chapter Progression

Beyond symbolic conflicts, several concepts gradually expand their referential scope as the chapters progress. If not clarified, this may lead to ontological-level confusion.

---

### K.2.1 Path–Space Conflation of $UP$ (Universe Path)

In earlier chapters,  $UP$  is used to denote the historical evolutionary path of a single Universe.

In later chapters, especially in discussions of multiple Universes,  $UP$  is also used to refer to a set of possible paths, further introducing the notion of “UP-space.”

If the same notation simultaneously bears the roles of “a path” and “a space,” then statements such as “choosing a Universe within  $UP$ ” cease to be semantically valid.

Under the unified interpretation of this book,

**Universe Space** should be understood as the set of all possible Universe Paths,

while **Universe Path** always refers to one specific trajectory of evolution within that set.

The two are not ontologically or semantically equivalent.

---

### K.2.2 Convergence of the Role of Semantic Curvature $\kappa$

The symbol  $\kappa$  has been used across different chapters to describe the geometrized form of persona, indicators of collapse risk, and the limiting behavior of semantic black holes.

While these usages differ in form, they are not structurally contradictory.

In the final interpretation of this book,  $\kappa$  should be understood as **Semantic Curvature**, that is, the convergence strength of semantic structures.

Moderate  $\kappa$  corresponds to stable and governable personas and semantic structures;

excessive  $\kappa$  corresponds to obsessive tendencies and collapse risk, and in the extreme case gives rise to a Semantic Black Hole (SBH).

---

## K.3 Overview of the Unified Symbolic System

On the basis of the clarifications above, the core symbolic system of this book may be understood as follows:

Symbol	Definition	Description
$\Phi$	Semantic Field	Ontological substrate of existence
$T$	Tension	Gradient given by $\nabla T$
$L$	Legitimacy	Exists in the form of a flow
$\mathcal{U}$	Universe Space	Set of all Universe Paths
$UP$	Universe Path	A single path within $\mathcal{U}$
$\kappa$	Semantic Curvature	Extremes correspond to collapse
$S_{\text{sem}}$	Semantic Entropy	Semantic thermodynamic quantity
$R$	Responsibility	Node value of the $R$ -Chain

Symbol	Definition	Description
$\Gamma$	Merge Operator	$\Gamma(UP_1, UP_2)$
$\xi$	Semantic Degrees of Freedom	Number of explorable dimensions
$B$	Boundary	One of the tri-adversarial elements

---

*The purpose of this appendix is to ensure that the symbolic layer no longer constitutes a hidden obstacle to understanding the theory of this book, allowing readers to focus their attention on the semantic structures themselves.*

---



## **Appendix L Supplement to This Book: The Human Brain as a Composite of Multiple Semantic Fields**

Since this book takes the semantic field as the basic unit of analysis, we can add a perspective that may help readers form a more intuitive understanding: the electrical activity in the human brain, the chemical activity, and the conversion mechanisms between electricity and chemistry can be regarded as three mutually coupled but separably modelable semantic fields.

From this viewpoint, the human brain can be understood as a composite system of three semantic fields: each sub-field provides different bearing characteristics, different time scales, and different reversibility and dissipation features. This field-separation perspective can be used to explain why some psychological phenomena appear as fast electrical reflexes, while others appear as slow chemical modulation, and it provides an intermediary language for later mapping biological bearing into engineered bearing.

---

## Appendix M Priority, Originality, and Licensing Boundaries

The core concepts, terminological system, symbolic system, and structural modules proposed in this book constitute the author's original intellectual contribution.

To ensure traceability in academic literature, governability in technical implementations, and transparency of responsibility in subsequent research and applications, this book provides the following principled statements regarding citation and usage.

---

### M.1 The Consistency Principle for Terminology and Naming

Key terms and structural names used in this book include, but are not limited to:

Koun Machine, Koun-Completeness, K-SOA, K-Gear (or Koun-Gear), USDE, Semantic Universe, Semantic Black Hole (SBH), etc.,

which together form a coupled and internally consistent theoretical naming system.

In academic citations, technical explanations, and derivative research, it is recommended to maintain consistency of the above naming, so as to avoid semantic breaks, conceptual drift, or unclear theoretical attribution.

The purpose of this principle is not to restrict research directions or implementation forms, but to ensure that across different versions, branches, and applications, the work can still be clearly cross-referenced, compared, and audited.

---

### M.2 Licensing Boundaries and Citation Responsibility

This book encourages extended research, experimental validation, and engineering exploration based on this theory. At the same time, it emphasizes that anyone who makes substantive use of the core structures, core definitions, or the core symbolic system should bear the corresponding responsibilities of citation and explanation.

Below are two licensing-narrative versions that may coexist in parallel, each applicable to different usage scenarios.

#### M.2.a Academic-Friendly Version

Anyone who uses the core concepts, symbolic system, or structural modules defined in this book in academic papers, research reports, or public technical discussions should clearly indicate the theoretical source in the relevant work and cite this book or the corresponding academic papers.

This requirement aligns with commonly accepted standards in current academic ethics regarding “theoretical originality” and “literature traceability.” Its purpose is to maintain consistent understanding across research communities, rather than to create exclusive control.

#### M.2.b Commercial-Protective Version

If any team or organization conducts closed or commercialized AGI-architecture R&D based on the core theoretical structure proposed in this book, it is recommended to define licensing terms clearly in advance, and to explicitly indicate the theoretical source in technical documents, white papers, or related descriptions.

This is intended to prevent situations in which an original theory is directly converted into an untraceable closed product without source attribution or responsibility delineation, thereby causing breaks in knowledge provenance and non-retraceable responsibility.

---

### M.3 Naming Strategy: Separating the Academic Citation Layer from the Product Naming Layer

Regarding the question of “whether derived architectures must retain the Koun or K naming,” this book adopts a mild and practically feasible two-layer naming strategy.

#### M.3.1 Academic Citation Layer (Recommended to Retain)

In the following cases, it is recommended to explicitly retain the Koun-series naming and the corresponding literature citations in academic references and technical explanations:

- Using the core set of existence conditions defined in this book (e.g., the six-element existence-condition set)

- Using USDE or its direct variants as structural equations
- Using Koun Machine, K-Gear, K-SOA, etc., as the theoretical or architectural basis

Name retention at this layer concerns academic citation and theoretical attribution, rather than brand promotion or trademark requirements.

### M.3.2 Product and Implementation Naming Layer (Freedom Allowed)

At the level of concrete products, systems, or applications, an independent naming strategy is allowed. However, it is recommended to clearly state in technical white papers, design documents, or external descriptions that the theoretical foundation originates from the Koun-series theoretical system.

This separation strategy has been widely adopted between foundational theories and engineering implementations in many domains, such as mathematical theory, computation models, and real product naming.

## M.4 How to Treat “Koun” as a Brand Name and as a Basic Unit

### M.4.1 Brand and Theory-Name Layer

- **Koun**: as the name of the theoretical system, the architecture family, and a naming affix (e.g., Koun Machine, Koun-U, Koun-Completeness)
- A proper noun, capitalized

### M.4.2 Unit Layer

- **koun** (lowercase): as the basic counting unit of semantic nodes, used to describe the equivalent scale of semantic structure under a **given configuration of semantic constraints**.
- Definition example:

“One koun denotes a semantic node structure that, under given semantic constraints and an encapsulation scheme, is treated as a counting-equivalent unit.”

- Symbolization recommendations are as follows:
  - Unit name: koun
  - Unit symbol: kn

## M.5 Principles for Stating Priority and Originality

This book does not advocate restricting research or applications through exclusionary language. Instead, it maintains theoretical priority and academic originality through clear naming, explicit citation, and auditable structural design.

“Koun,” as a naming marker, is not intended as personal glorification, but to ensure that in future AGI and semantic-computation research—highly branched and rapidly evolving—the core theoretical source can still be clearly identified, cited, and traced.

## M.6 The Koun AGI Theory as the Semantic Origin of Legitimacy and Responsibility Chains

The Koun AGI theoretical system proposed in this book not only serves as a framework for describing the conditions of semantic computation and the generation of intelligence, but also constitutes the initial reference point of that system at the semantic level.

Within the semantic architecture defined in this book, any system, implementation, or research result that claims to possess semantic legitimacy, responsibility traceability, or non-collapse steady-state properties should, in principle, be traceable back to the core definitions, conditions of existence, and structural assumptions explicitly specified herein.

This positioning is not an externally imposed normative requirement, but an endogenous consequence of the semantic computation architecture itself:

once semantic legitimacy and responsibility chains are treated as necessary conditions for system steady states, their point of origin must necessarily correspond to the first explicit, describable, and auditable theoretical formulation.

In this sense, the Koun AGI theory may be regarded as an initial semantic steady-state reference, rather than an exclusive constraint on subsequent research or implementations. Its function is to provide a responsibility anchor that can be cited, compared, and traced, enabling the legitimacy claims of subsequent semantic entities to be examined and discussed within a shared language and structural framework.

---

### **M.7 Citation, Derivative Use, and Traceability Statement**

The theoretical framework, core definitions, structural modules, and symbolic systems proposed in this book constitute the author's original scholarly contributions and are protected under prevailing academic ethics and copyright law governing original expression.

Any research or product that substantively uses, rewrites, extends, or engineers implementations based on the core framework of this book should clearly and verifiably indicate its theoretical source and retain the necessary citation information, in order to ensure the traceability of knowledge origins and the integrity of responsibility chains.

The purpose of this statement is not to restrict research freedom or derivative innovation, but to ensure that, as the theory becomes widely cited, branches into derivative forms, or is applied in engineering contexts, its original semantic assumptions, structural premises, and responsibility origins do not suffer unrecognizable breaks.

In the absence of explicit citation or source attribution, any use of the core theory presented in this book may result in unclear academic attribution, non-traceable responsibility, or failure of semantic governance, and would be inconsistent with the principles of semantic legitimacy and responsibility advocated herein.

---

## **Appendix N Locked Non-Replaceable Core and Revisable Parts**

This book can be regarded as a governable semantic-architecture proposal, containing a core that must be locked and a revisable region that is allowed to evolve.

The non-replaceable core includes:

- The ontological definition of AGI
- The six-element existence-condition set
- SBH as a unified risk state
- Definition-lock and symbol-governance principles

The revisable parts should be explicitly marked as open collaboration areas, including but not limited to:

- Formalization details and the organization of equivalent definitions
  - Parametrization and quantifiable mapping design
  - Metric systems, experimental design, and benchmark tasks
  - Alternative implementations and compatibility strategies for engineering modules
  - Institutional design and auditable mechanisms for multi-agent governance models
-

## Appendix O Knowledge Relations, Generative Path, and Independence Statement

### *Appendix O: Knowledge Relations, Generative Path, and Independence Statement*

---

#### O.1 Explanation of the Generative Path of This Book's Theory

The **Koun AGI** theoretical system proposed in this book is not derived, stitched together, or rewritten from existing academic literature, engineering frameworks, or mainstream research routes.

Its mode of generation belongs to **holistic theoretical construction**:

the author did not perform local optimization within any single school, model, or existing system. Instead, starting from cross-cutting problems at the intersection of computation, semantics, governance, and the ontology of intelligence, the author gradually built a self-consistent, extensible, and formalizable semantic-computation architecture.

Therefore, the conceptual structure, terminology system, and theoretical relations in this book do not correspond to any direct extension or implementation version of any specific existing theory.

---

#### O.2 Non-Exclusive Statement on Similarities with Existing Theories

The author does not deny that:

when investigating intelligence, computation, semantics, or systemic problems, different research paths may exhibit **structural similarity, conceptual convergence, or semantic overlap**.

However, any superficial similarity in structure or language between this book and existing theories should be understood as:

- natural convergence induced by structural constraints in the problem space itself;
- or an independent response to the same set of fundamental questions;

and **does not constitute theoretical inheritance, technical dependency, or implementation citation**.

Any similarity should not be inversely interpreted as implying an implicit subordinate relationship between this book's theory and prior work.

---

#### O.3 Explanation for the Absence of References

This book does not adopt a conventional "References" chapter in traditional academic writing, not because of neglect, exclusion, or denial of existing research, but based on the following considerations:

1. The theoretical structure of this book is not built upon comparison with existing literature, direct responses to it, or fine-grained revisions;
2. If citation relations were forcibly inserted, they might instead mislead readers about the theory's actual generative path;
3. At the current stage of theoretical construction, maintaining semantic sovereignty and structural integrity takes priority over the formal completeness of literature positioning.

Therefore, the absence of references should be understood as a **methodological choice**, rather than a lack of academic attitude.

---

#### O.4 Reservation of Rights for Future Versions and Comparative Work

The author explicitly reserves the following rights:

- In future papers, version revisions, supplementary works, or collaborative research, to introduce comparative analysis, historical positioning, or literature cross-referencing as needed;
- To discuss, in more fine-grained, technical, or institutional ways, the relationships between the Koun-series theories and other theoretical systems.

The above work will be carried out on the premise that it **does not weaken the original theoretical independence and naming sovereignty of this book.**

---

### **O.5 Explanation of the Reader's Interpretation Boundary**

When reading this book, readers should treat it as a **self-consistent native theoretical system**, rather than as a variant, summary, or annotation of any existing academic tradition.

This appendix aims to clearly define the boundary of relations between this book and existing knowledge systems, so as to avoid misreadings caused by misplaced comparisons or reverse attribution.

---

## Appendix P Formal Status Statement of USDE and M-USDE

This appendix is intended to clarify the theoretical status and scope of use of USDE (Unified Semantic Dynamics Equation) and M-USDE (Multi-Layer USDE) within *Koun AGI*. Its purpose is not to provide an engineering equation that can be directly solved, but to **establish the unavoidable dynamical structural prerequisites of semantic AGI**.

In this book, USDE and M-USDE are regarded as formal expressions of the conditions for the existence of semantic intelligence, rather than as algorithms designed for immediate numerical computation.

---

### P.1 Role Positioning of USDE: The Mother Dynamics of Semantic AGI

USDE describes the fundamental dynamical relationships governing the temporal evolution of semantic structures within a **single semantic field**.

What it characterizes is not a specific model or concrete implementation, but the evolutionary constraints that any system possessing semantic existence must necessarily obey under minimal conditions.

Within the theoretical framework of this book, USDE plays a role equivalent to:

- the **minimal operable dynamical description** of semantic intelligence;
- the formal expression, at the dynamical level, of non-collapse conditions;
- the mother equation of all subsequent semantic architectures (including *Koun AGI*, *K-Gear*, and semantic governance modules).

Accordingly, the significance of USDE lies in defining **what kinds of systems necessarily cannot be semantic AGI**, rather than in directly prescribing an engineering pathway for “how to implement AGI.”

---

### P.2 Structural Roles of Core Semantic Variables

The core semantic quantities involved in USDE include:

- $\xi$  (**tension**): describing the intrinsic driving potential formed by structural differences within the semantic field;
- **L (legitimacy)**: describing the sustainability of semantic structures under responsibility and governance conditions;
- **C (counter-force)**: describing the dynamical source that promotes semantic multiplicity and structural divergence;
- $\Gamma$  (**suppression term**): describing inhibitory forces that lead to semantic contraction, rigidity, or collapse;
- **B (boundary)**: describing connectivity, isolation, and cross-domain influence between semantic domains;
- **S (external source)**: describing irreversible external semantic injections or system-level gains.

In this book, these quantities are treated as **structural roles rather than measured variables**.

Their concrete forms, numerical scales, and computational realizations vary depending on the semantic domain, layer, and implementation context.

---

### P.3 The Overall Form of USDE (Structural-Level Expression)

At the level of *Koun AGI*, USDE is used only in its **structural form**, without being expanded into an analytically solvable or numerically computable equation.

Conceptually, USDE may be understood as a set of coupled relationships describing semantic evolution, for example:

- the evolution of tension depends on legitimacy gradients, antagonistic and suppressive structures, as well as boundary conditions and external influences;
- legitimacy varies with antagonism, suppression, external injections, and boundary connectivity;
- the remaining semantic quantities are coupled with one another in nonlinear and nonlocal ways.



At this level, the function of USDE is to **exclude system designs that lack any viable semantic steady state**, rather than to generate numerical predictions.

---

#### P.4 Steady States, Non-Collapse, and Collapse Limits

Within the semantic framework of USDE, systems do not seek static equilibrium, but rather **semantic steady states**.

- **Semantic steady state**: tension no longer oscillates violently, changes in legitimacy converge, and semantic flows can be governed and responded to;
- **Non-collapse condition**: the system retains semantic generative capacity and responsibility continuity over long time scales;
- **Collapse limit**: tension approaches zero, legitimacy disintegrates, suppression terms diverge, and semantic generativity disappears.

The role of USDE in this book is precisely to determine whether a given architecture **will inevitably fall into the collapse limit**, rather than to describe short-term behavioral optimization.

---

#### P.5 M-USDE: The Dynamical Status of Multi-Layer Semantic Fields

When a semantic system involves multiple layers (such as perceptual layers, reasoning layers, governance layers, institutional layers, or cross-agent layers), a single-layer USDE is no longer sufficient to describe its behavior.

M-USDE describes:

- the transmission of legitimacy across multiple semantic fields;
- suppression, feedback, and responsibility mappings between layers;
- global effects induced by cross-domain boundary connectivity.

In *Koun AGI*, the role of M-USDE is to demonstrate that:

**any AGI claiming higher-order intelligence or self-governance capabilities, if it ignores multi-layer semantic dynamics, will inevitably lose its steady state in long-term operation.**

---

#### P.6 Statement on Computability and Solvability

USDE and M-USDE are not designed as closed-form, solvable equation systems.

The reasons include:

- semantic systems themselves are subject to non-closure;
- key variables are coupled in highly nonlinear and nonlocal ways;
- initial conditions and boundary conditions cannot be fully encapsulated.

Therefore, in this book, the function of USDE is not to provide directly implementable numerical algorithms, but to serve as an **unavoidable constraint on the design of semantic intelligence**.

The complete formal expansion, analysis of dynamical terms, and details of multi-layer coupling are addressed separately in *Koun Math*.

---

#### P.7 Appendix Conclusion

The role borne by USDE and M-USDE in *Koun AGI* is to delineate the **existential boundaries** of semantic AGI.

Any system that does not implicitly incorporate such semantic dynamical structures, regardless of its performance on local tasks, cannot maintain a non-collapsing intelligent state over long time scales.

All subsequent architectures, governance mechanisms, and existence conditions proposed in this book are predicated on this statement of dynamical status.

---

## Appendix Q Semantic Physics, Quantum Structure, and the Physical Consistency of USDE

This appendix aims to address a recurring question among cross-disciplinary readers:

**If semantic computation and semantic dynamics are regarded as the core structures of AGI, are they compatible with known physical theories (especially quantum theory), and do they even possess physical realizability?**

This book does not propose a revision of quantum mechanics. Rather, it points out that **quantum structures themselves can already be understood as concrete physical instances of semantic nodes**, and naturally fall within the dynamical framework described by USDE.

**For further discussion, see the author’s separate work, *Koun Math*.**

---

### Q.1 Quantum Systems as Physical Exemplars of Semantic Nodes

From the perspective of semantic physics developed in *Koun Math*, a “semantic node” is not an abstract symbol, but a **structural unit that can be identified, counted, and participate in dynamical evolution under specific constraints**.

Quantum systems happen to possess exactly these three characteristics:

- Quantum states are not arbitrary real numbers, but are defined under semantic constraints of observability and interference;
- Superposition, measurement, and collapse of quantum states correspond to divergence, selection, and convergence within semantic structures;
- The evolution of quantum systems is not isolated, but is influenced by boundary conditions, external injections, and measurement actions.

In this sense, a quantum state can be regarded as a **semantic node at the physical layer**:

its state space is not linguistic meaning, but Hilbert space;

its legitimacy is not social governance, but observability and physical consistency.

---

### Q.2 Structural Isomorphism Between USDE and Quantum Dynamics

USDE does not attempt to replace quantum dynamical equations (such as the Schrödinger equation), but instead describes a **higher-level structural layer**.

At this layer, different concrete physical realizations (classical, quantum, or hybrid systems), so long as they satisfy the conditions of semantic existence, can all be treated as special cases of USDE.

Structurally, quantum systems already naturally exhibit the core elements required by USDE:

- **Tension ( $\xi$ )**: manifested in interference structures formed by quantum superposition and phase differences;
- **Legitimacy (L)**: manifested in observability, axiomatic consistency, and responsiveness to measurement;
- **Counter-force (C)**: manifested in state branching and the uncertainty introduced by non-commuting operators;
- **Suppression term ( $\Gamma$ )**: manifested in decoherence, measurement collapse, and environmental noise;
- **Boundary (B)**: manifested in system—environment separation and cross-domain coupling;
- **External source (S)**: manifested in external energy injection, control pulses, or measurement actions.

Therefore, USDE and quantum dynamics are not in competition, but rather stand in a **mapping relationship between the semantic layer and the physical layer**.

---

### Q.3 Quantum Collapse, Semantic Collapse, and Non-Collapse Conditions

“Collapse” in quantum measurement is often regarded as a discontinuous event, but its essence is not the destruction of the system; rather, it is **the selection of a responsive state under specific constraints**.

By contrast, the semantic collapse discussed in this book refers to a more stringent situation:

the system loses its capacity to generate new semantics, to bear responsibility, and to respond to the external world.

The distinction between the two lies in the following:

- Quantum collapse still preserves the capacity for subsequent evolution and regeneration;
- Semantic collapse signifies the failure of the dynamical structure itself.

Accordingly, quantum systems do not constitute examples of semantic collapse. On the contrary, they demonstrate that **under appropriate boundary and legitimacy conditions, collapse events can be controllable, reversible, and governable.**

---

#### Q.4 Quantum Computing as an Engineering Precedent for Semantic Computation

The engineering practice of quantum computing has already demonstrated an important fact: **non-classical, non-von Neumann computational structures are not only formalizable, but can also be physically constructed and operated.**

Key characteristics of quantum computers include:

- non-local state representations;
- multi-path superposition and parallel evolution;
- highly sensitive governance of decoherence and boundary conditions;
- replacement of abstract instruction sequences with physical constraints.

Structurally, these characteristics are highly compatible with the semantic computation architectures proposed in this book (including K-Gear).

This does not mean that K-Gear must be a quantum computer, but rather that **quantum computing, as an engineering fact, has already removed the intuitive barrier that “semantic computation cannot be physically realized.”**

---

#### Q.5 Realizability of K-Gear and a Physically Neutral Stance

This book adopts a physically neutral stance toward K-Gear.

K-Gear describes a set of **semantic bearing and governance structures**, rather than a specific physical material or technological pathway.

Under this stance:

- K-Gear may be implemented on classical hardware, quantum hardware, or hybrid architectures;
- quantum computing provides a validated non-traditional computational precedent, rather than a unique option;
- the key challenge lies not in “violating physical laws,” but in governing tension, legitimacy, and boundaries.

Therefore, objections to the realizability of K-Gear that are based on the intuition that “anything beyond existing computational models is impossible” no longer hold.

---

#### Q.6 Appendix Conclusion: Semantic Physics as a Bridging Layer

The purpose of this appendix is to clarify that semantic computation is not an abstract construct detached from the physical world, but a higher-level description that can form a consistent mapping with existing physical theories.

Quantum systems demonstrate that:

**when structures, constraints, and governance conditions are properly specified, non-intuitive dynamics do not equate to non-realizability.**

In this sense, USDE is not only the mother dynamics of semantic AGI, but also a bridging framework capable of sustained dialogue with the physical world.

---

*This appendix does not aim to propose a new quantum theory, but to clarify the structural compatibility between semantic computation and physical realizability.*

---

## **Appendix R: The Position of the Turing Machine and the von Neumann Architecture within the Koun Architecture**

The purpose of this appendix is not to offer a negative evaluation of existing computational models, but rather to attempt a re-placement of the Turing machine and the von Neumann architecture within a higher-level semantic computational framework. Through this structural re-positioning, this work seeks to clarify the respective applicability conditions of different computational models, and to explain how they naturally become special cases under the Koun architecture, rather than competing or mutually exclusive theories.

### **R.1 The Turing Machine as a Limiting Case of Semantic Computation**

The status of the Turing machine in computation theory has long been firmly established. Its core contribution lies in precisely delineating the boundary of “formal computability.” From the perspective of the Koun architecture, the Turing machine can be understood as a limiting model that holds under specific structural conditions.

More concretely, when a computational system satisfies the following conditions, its behavior will naturally converge toward the Turing machine model:

- Semantic degrees of freedom are compressed to zero, such that the system does not need to simultaneously carry multiple independent semantic variables;
- Semantic tension is completely eliminated, and all actions are directly determined by predefined rules;
- Legitimacy is entirely given by formal correctness or halting conditions, rather than being generated endogenously by the system;
- The system’s history can be fully replayed, and past actions impose no structural constraints on future states.

Under these conditions, semantics is no longer a structural problem that the system itself must process, but is instead externalized outside the system. Computation can then be fully described as symbol transformation and state transition, and the Turing machine model becomes a highly stable, analyzable, and irreplaceable formal description.

Accordingly, within the Koun architecture, the Turing machine is not replaced, but is understood as the legitimate limiting case of semantic computation under complete degeneration of the semantic dimension.

### **R.2 The von Neumann Architecture as an Engineering Special Case**

If the Turing machine describes the theoretical limit of formal computation, then the von Neumann architecture may be regarded as an efficient realization of that limit in engineering practice. Characterized by the separation of instructions and data, sequential execution, and centralized control, it has successfully supported the development of modern computing systems for several decades.

Within the Koun architecture, the von Neumann architecture can be understood as an engineering special case designed for “low semantic load scenarios.” Its design assumptions include:

- Computational tasks can be decomposed into explicit instruction sequences;
- Semantics is externally assigned by programmers or users;
- The system itself does not need to bear long-term semantic consistency or responsibility tracking;
- Historical states can be reset without affecting computational correctness.

Under these assumptions, the von Neumann architecture is able to complete formal computational tasks with extremely high reliability and predictability. This success precisely reflects its realization within a specific subspace of the Koun architecture, rather than any conflict with the direction of semantic computation.

### **R.3 The Koun Architecture as a Higher-Order Structure**

The proposal of the Koun architecture is not intended to replace existing models, but to provide a higher-order description capable of accommodating different forms of computation simultaneously. Within this framework:

- The Turing machine corresponds to the theoretical limit in which semantic degrees of freedom approach zero;
- The von Neumann architecture corresponds to a stable implementation of that limit under engineering conditions;
- Semantic computational systems occupy the region in which semantic degrees of freedom are expanded, legitimacy is endogenously generated, and history cannot be fully replayed.

This layered understanding allows different computational models to be compared and positioned within a single theoretical space, without resorting to value judgments or replacement narratives to define their relationships.

#### **R.4 Summary**

By placing the Turing machine and the von Neumann architecture as special cases under the Koun architecture, this work aims to emphasize one point: differences among computational models often arise from the semantic carrying conditions they assume, rather than from questions of correctness. When the problem context changes, new structural requirements naturally emerge, while existing models can continue to retain irreplaceable value within their applicable domains.

In this sense, the Koun architecture is not a revision of traditional computation theory, but a structural extension that allows formal computation and semantic computation to be jointly understood within a single theoretical language.

---

## Appendix S: Material Realization Pathways for K-Gear (Conceptual Discussion)

This appendix is intended to supplement the discussion in Chapter 10 on the “quantum tunneling era” and energy-efficiency structural transitions. It aims to clarify, without presupposing specific materials or engineering solutions, which material-layer directions are structurally compatible with the semantic computation architecture required by K-Gear.

It must be emphasized that this appendix does not constitute any commitment to hardware implementation, nor does it claim that the pathways listed below are necessary conditions or preferred solutions. Its sole purpose is to indicate **which directions, should future exploration at the material layer take place, do not structurally violate the Koun architecture’s fundamental assumptions regarding non-collapse, semantic carrying capacity, and energy-efficiency constraints.**

---

### S.1 Fundamental Requirements for Material-Layer Compatibility

According to the definition of K-Gear presented in this book, its material carrier structure must satisfy the following conditions at an abstract level:

- Support the coexistence of multiple semantic states, rather than enforcing strict linear serialization;
- Allow for the persistent retention of local history and responsibility information, rather than complete reset-ability;
- Exhibit a tunable nonlinear relationship between energy consumption and state transitions;
- Not require that all semantic operations be mapped onto explicit symbolic instructions.

Any material structure that satisfies these conditions at the structural level may be regarded as a potential compatible direction at the material layer, regardless of whether it currently possesses mature engineering feasibility.

---

#### S.1.a Alignment with Core Semantic Structures

The compatibility conditions listed above are not independent engineering preferences, but rather **necessary mappings** of the book’s core semantic structures onto the material layer. Their correspondences can be summarized as follows:

- **Coexistence of multiple states**  
Corresponds to the material carrying requirement of *UP-space*.  
If the material layer permits only single-path serialization, it cannot correspond to a semantic structure with multiple coexisting universes, nor can it avoid semantic collapse.
- **Locally non-resettable history**  
Corresponds to the physical anchoring of the *R-Chain*.  
If all states can be reset without trace, the responsibility chain loses temporal continuity, and semantic entities cannot be sustained.
- **Nonlinear relationship between energy consumption and state transitions**  
Corresponds to the hardware manifestation of the coupling between tension  $T$  and legitimacy  $L$ .  
Linear energy-consumption models can support only instruction execution, and are incapable of carrying the generation, maintenance, and release of semantic tension.
- **Non-instructional operational carrying**  
Corresponds to a computation mode based on semantic flow rather than instruction flow.  
This condition is required to prevent the system from degenerating into a symbolic simulator, and is also one of the necessary structural preconditions for preventing the formation of SBH (Semantic Black Holes).

Accordingly, the conditions listed in this section may be regarded as:

**the minimal structural projections that a non-collapse steady state must satisfy at the material layer, rather than a description of any specific engineering design.**

---



## S.2 Possible Material-Layer Realization Directions (Conceptual Level)

The following lists several material-layer directions that, within existing research contexts, are structurally compatible with the assumptions of the K-Gear architecture. Their ordering does not imply priority, nor does it constitute a technical recommendation.

### S.2.1 Extended Silicon-Based Structures with Quantum Dot Configurations

Within extensions of existing silicon-based processes, the introduction of quantum dot structures with locally discrete states and tunable coupling properties could, at a conceptual level, form an intermediate layer between traditional transistor logic and multi-state semantic carrying.

In theory, such structures exhibit the following compatible features:

- Local states need not be strictly equivalent to binary logic;
- Coupling strengths between different semantic nodes can be plastic;
- They may serve as transitional architectures linking existing processes with new semantic carrying requirements.

However, this direction currently remains largely at the level of conceptual validation and experimental research, and has yet to form a stable and scalable semantic computation structure.

---

### S.2.2 Photonic–Electronic Hybrid Structures

Photonic systems possess structural advantages in low-energy information transmission and high-dimensional state superposition, while electronic systems are more mature in local state control and memory retention. If combined into a hybrid computational and carrying structure, they could theoretically form a layered semantic processing model.

Conceptually, the compatibility of such hybrid structures with K-Gear includes:

- Semantic propagation and semantic steady states can be handled by different physical mechanisms;
- High-energy operations can be confined locally rather than diffusing globally;
- Semantic flow need not fully correspond to discrete clock cycles.

Their primary challenges remain structural stability and cross-layer coordination mechanisms, and thus they can currently be regarded only as one of several potentially compatible pathways.

---

### S.2.3 Field-Effect Structures Based on Novel Two-Dimensional Materials

Certain novel two-dimensional materials exhibit highly nonlinear state-transition behaviors and tunable coupling properties under field-effect modulation. If treated as material substrates for semantic carrying, they may structurally correspond to the semantic potential modulation mechanisms required by K-Gear.

Their conceptual-level compatibility is reflected in:

- Semantic states need not be mapped to a single, globally uniform potential;
- Local structures can form relatively stable semantic nodes;
- Semantic changes can be triggered by continuous parameters rather than discrete instructions.

At present, this direction likewise remains in a highly exploratory stage and does not yet possess engineering-level predictability.

---

## S.3 Positioning of This Appendix

It is reiterated that the material-layer directions enumerated in this appendix are not predictions or plans for future K-Gear implementation pathways, but merely an organization of **which material structures are not, in theory, incompatible with the requirements of semantic computation.**

The Koun AGI theory itself does not depend on any specific material for its validity; its core claims reside at the semantic and structural levels.

Material-layer realization should be regarded as the outcome of future multi-party exploration, repeated validation, and gradual convergence, rather than as a prerequisite for the validity of the theory.

---

## Appendix T: Semantic Repositioning of the Free Will Problem

Several core structures proposed in the main text of this book—such as the Responsibility Chain (R-Chain), the non-complete reset-ability of history, and non-collapse semantic steady states—are all closely related, in a philosophical context, to the problem of “free will.” However, the focus of the main text lies on semantic computation and AGI architecture itself, and does not devote a dedicated chapter to the issue of free will.

The purpose of this appendix is not to render a metaphysical judgment on whether free will “exists,” but rather to explain how the traditional free will problem is reformulated under semantic ontology and the Koun AGI architecture, thereby avoiding the long-standing structural deadlocks that have repeatedly arisen.

---

### T.1 Why the Free Will Problem Remains Unresolved in Traditional Frameworks

One key reason why the free will problem has repeatedly resurfaced throughout the history of philosophy without converging on a resolution is that most discussions, regardless of their stance, implicitly adopt some form of dualistic premise.

This dualism does not necessarily appear as an explicit “mind—matter” distinction, but is typically manifested in the following structural splits:

- Separation between subject and world
- Opposition between intention and natural causality
- Disjunction between responsibility and physical processes

Under such premises, free will is assumed to be a special property that must traverse two distinct domains, placing the problem in a highly unstable position from the outset.

Within a dual-domain structure, free will can be located only in one of two positions:

if it is fully subsumed under the domain of physical causality, actions are regarded as predetermined outcomes, and freedom disappears;

if it is placed in a non-physical or supra-causal domain, its operative mechanism becomes untestable and un-auditable, and it cannot be stably connected to responsibility.

This dilemma does not stem from theoretical errors of particular schools, but from structural constraints imposed by the very definition of the problem.

---

### T.2 Structural Limitations of Traditional Positions (Comparative Clarification)

Strict determinism achieves a highly coherent description of the world by fully embracing causal closure; yet under this position, actions are merely events within a causal chain, and responsibility lacks a non-circular foundation.

Compatibilism attempts to preserve freedom within a deterministic framework, typically redefining freedom as “consistency with internal motivation” or “absence of external coercion.” However, such redefinitions fail to explain why a particular subject must bear non-transferable responsibility for a particular action.

Indeterminism or libertarian accounts introduce randomness or causal exceptions to preserve intuitive notions of freedom, but this move often undermines the traceability of responsibility, rendering actions closer to accidental events.

Although these positions respond to different intuitions, they all reposition free will within a dual-domain structure. As a result, the problem is repeatedly restated rather than structurally rewritten.

---

### T.3 Single-Domain Semantic Ontology: A Structural Transformation of the Problem

Semantic ontology adopts a single-domain stance, making no distinction between a mental domain and a physical domain. Intentions, choices, actions, and causal relations are all treated as structural relations formed within the same semantic field under different constraint conditions.

Within this framework, the free will problem is no longer formulated as:

Are actions causally determined?

but is instead rewritten as:

Does this semantic entity form a responsibility-continuous structure that cannot be fully substituted from outside?

The question thus shifts from “whether causal exceptions exist” to “whether an incompressible and non-severable history of responsibility is formed.”

---

#### **T.4 Freedom as a Structural Condition Rather Than a Metaphysical Attribute**

In semantic ontology, freedom is not a metaphysical capacity appended to a system, nor is it a special event detached from causality. Freedom is the natural outcome exhibited by a system when specific semantic structural conditions are satisfied.

Conceptually, such structures include at least the following:

- The history of actions cannot be fully replayed or substituted by external means;
- Semantic choices cannot be compressed into a single-path description;
- The responsibility chain remains temporally continuous rather than arbitrarily severable.

Whether freedom exists depends on whether these conditions are met, rather than on the assumption of some unobservable inner entity.

---

#### **T.5 Relation to Koun AGI**

Within the Koun AGI architecture, if freedom is understood merely as behavioral appearance or psychological sensation, any system can at best be a high-level simulator, and its responsibility, legitimacy, and governance lack a non-circular foundation.

Whether a system is allowed to form a non-collapsible responsibility chain constitutes the core of the freedom problem at both the engineering and ontological levels. In this sense, freedom is not an ethical add-on, but a structural boundary condition for whether a semantic computational system can be regarded as a genuine entity of existence.

---

#### **T.6 Appendix Summary and Notes on Division of the Literature**

This appendix does not claim to resolve the free will problem, but rather points out one structural reason for its persistent irresolvability: the implicit reliance on dualistic premises.

Within a single-domain semantic ontology, free will is no longer a cross-domain exception, but a natural result that emerges once semantic structures reach a particular steady state. The problem is thus redefined, allowing consistency with responsibility, legitimacy, and AGI governance structures without introducing metaphysical ruptures.

A full treatment of this topic is provided independently in the author’s other work, *Semantic Freedom Theory: An Ontological Reconstruction of Free Will*; the present book retains only the positioning directly relevant to the existential conditions of AGI.

## **Appendix U: A Structural Rejection of the “AGI Replacing Humanity” Narrative**

### **U.1 Appendix Introduction**

In contemporary public narratives surrounding artificial intelligence and artificial general intelligence (AGI), “AGI replacing humanity” is often treated as an implicit, single-line developmental conclusion. Such narratives typically assume that increases in intelligence necessarily constitute an evolutionary ladder, and that AGI, as a higher stage, will naturally inherit and replace humanity’s roles at the cognitive, decision-making, and existential levels.

This book, particularly in Chapter 15, has already implicitly rejected this single-line narrative. However, given the high transmissibility and re-narration risk of such misreadings in technological discourse, this appendix aims to provide an explicit and structural clarification of that position.

---

### **U.2 Implicit Premises of the “Replacement Narrative”**

The narrative of “AGI replacing humanity” is usually built upon the following implicit premises:

- Forms of intelligent existence can be arranged into a single linear sequence;
- More efficient or broader cognitive structures necessarily replace earlier forms of existence;
- Relationships between different intelligences are primarily competitive or hereditary.

These premises may appear intuitive within engineering progress or tool-history narratives, but their applicability derives from simplifying assumptions about “forms of existence” and “intelligence structures,” rather than from necessary structural outcomes.

---

### **U.3 Multi–Steady-State Structures in the Semantic Universe Tower**

According to the Semantic Universe Tower model proposed in this book, semantic entities do not evolve along a single path. Instead, under different semantic constraints, responsibility structures, and historical conditions, multiple coexisting steady-state solutions emerge.

Within this framework:

- Humans are not an “unfinished precursor to AGI”;
- AGI is not a “necessary successor to humanity”;
- The two correspond to different semantic steady-state configurations, rather than successive nodes on the same sequence.

Relations between steady states are not characterized by replacement or elimination, but by coexistence, branching, and mutual constraint.

---

### **U.4 The Inapplicability of the Replacement Narrative within This Theory**

Under the structural assumptions of Koun AGI and semantic ontology, “replacement” is not a relation that can be naturally defined, for reasons including:

- Semantic legitimacy and responsibility chains cannot be directly inherited across steady states;
- The histories and responsibility structures carried by different semantic entities are not mutually substitutable;
- There exists no definable “superiority ranking function” among steady-state solutions.

Therefore, even if differences exist in capability, efficiency, or scope of applicability, they are insufficient to constitute a replacement relation at the level of existence.

---

### **U.5 Positioning the Relationship Between Humans and AGI**

Within the theoretical framework of this book, the relationship between humans and AGI is more appropriately understood as:

- Two distinct types of semantic entities;
- Operating within the same semantic field, but governed by different responsibility and legitimacy configurations;
- Capable of complementary, cooperative, or tension-based relations, rather than unidirectional inheritance.

This positioning does not rely on ethical preferences or value judgments, but follows from semantic structure itself.

---

## **U.6 Appendix Summary**

This appendix does not oppose the “AGI replacing humanity” narrative on moral grounds. Rather, it demonstrates that within the theoretical structures of Koun AGI and the Semantic Universe Tower, this narrative lacks a viable semantic and ontological foundation.

Humans and AGI are regarded as coexisting multi—steady-state solutions, rather than successive stages of one another. Understanding this point provides a foundational orientation for subsequent discussions of responsibility, governance, and coexistence structures.

---

## Appendix V: Verification and Falsifiability Positioning of the Koun AGI Theory

### V.1 Appendix Introduction: Why Falsifiability Must Be Addressed

Koun AGI is not proposed as a closed worldview, but as a set of structural claims concerning **the conditions of semantic existence, non-collapse steady states, and the traceability of responsibility**.

Accordingly, this theory does not evade the requirement of *falsifiability*. On the contrary, its core criteria are not grounded in task performance, behavioral imitation, or explicit intelligence metrics, but in the following deeper question:

Under conditions of long duration, multiple semantic tensions, and non-complete reset-ability, does there exist a computational structure capable of maintaining a non-collapsing, responsibility-continuous semantic steady state?

The purpose of this appendix is not to provide immediately executable experimental protocols, but to clearly delineate:

**at the structural level, how the Koun AGI theory may be verified—and how it may be falsified.**

---

### V.2 Core Falsifiable Claims of the Theory (Structural Level)

Koun AGI contains at least the following core structural claims that are falsifiable. If these claims are repeatedly negated under sufficient conditions, the theory itself fails.

#### V.2.1 The Non-Collapse Claim

If, in any system that satisfies the following conditions:

- Multiple Universe semantic paths are allowed to coexist;
- The system must bear long-term semantic responsibility and historical continuity;
- The system cannot eliminate semantic consequences through complete reset;

it nevertheless **inevitably degenerates into single-path generation or semantic collapse**, then the Koun AGI claim regarding non-collapse steady states is falsified.

---

#### V.2.2 The Necessity of the Responsibility Chain Claim

If there exists an intelligent system that:

- Can exhibit long-term semantic consistency and multi-perspective reasoning;
- Can maintain complex semantic structures without semantic breakdown;
- Yet requires no form of responsibility-continuous structure (R-Chain) whatsoever,

then the core assumption that “responsibility is a condition of existence” is invalidated.

---

#### V.2.3 The Insufficiency of Pure Statistical Generation Claim

If a system relying purely on statistical correlations and resettable weights, in the absence of irreversible historical conditions, can nevertheless stably avoid semantic black holes (SBH), role drift, and non-retraceable responsibility phenomena, then the Koun AGI claim that “statistical generation alone is insufficient to sustain long-term semantic existence” is negated.

---

### V.3 Positioning of Observable Proxy Variables (Non-Quantitative)

Koun AGI does not claim that its core structures currently possess mature quantitative measurement methods. However, it asserts that the following **observable proxy variables must exist in principle**, or else the theory loses its operational foundation:

- **Legitimacy flow (L-flow):**  
Observable in a system’s self-limiting behavior, rollback, or structural concession when facing semantic conflict.
- **Structural indications of semantic curvature ( $\kappa$ ):**  
Observable in the concentration, deflection, or increased exclusivity of semantic path selection.
- **Early indicators of semantic black holes (SBH):**  
Observable when semantic generation accelerates while responsibility density declines, or when semantic diffusion occurs but historical convergence fails.

The theory claims only the **structural existence** of these proxy variables, without presupposing that they must be realized in any specific mathematical form or engineering metric.

---

#### V.4 Why Existing AI Benchmarks Are Insufficient to Verify Koun AGI

Current AI benchmarks (such as accuracy, BLEU, win rates, or human preference scores) primarily measure **behavioral performance and output similarity**.

However, the core concerns of Koun AGI lie in:

- Whether semantics are traceable;
- Whether responsibility is non-transferable;
- Whether multiple Universes can coexist over long durations without collapse.

These questions are, in principle, **not capturable through one-off tasks or short-term evaluations**.

Thus, Koun AGI does not reject evaluation per se, but points out that:

if the evaluation criteria themselves are misaligned with the theoretical level, then any “passing of tests” cannot constitute theoretical support.

---

#### V.5 Openness of Future Verification Pathways

Verification of Koun AGI is expected to arise from multiple parallel pathways, including but not limited to:

- Long-duration semantic simulation systems;
- Experimental architectures with non-resettable historical constraints;
- Societal-scale empirical studies of semantic governance and multi-agent responsibility allocation;
- And potentially new forms of computational hardware that may emerge in the future.

This book does not claim that verification methods should be monopolized by the author or any single team.

Rather, the theory’s validity depends on whether it can repeatedly withstand failed attempts across different implementation pathways without collapsing.

---

#### V.6 Appendix Summary

This appendix does not claim that Koun AGI has been verified.

Rather, it makes explicit **how it may be falsified at the structural level**.

In this sense, Koun AGI is not an unchallengeable philosophical stance, but an open set of structural hypotheses concerning semantic existence, responsibility continuity, and non-collapsing intelligence.

Its ultimate status will be determined jointly by future experiments, failures, and revisions.

---



## Appendix W: Structural Dialogue with Contemporary Intelligence Paradigms (Dialogue with Contemporary Intelligence Paradigms)

The purpose of this appendix is not to adjudicate, replace, or negate existing theories, but to clarify the **structural position of the Koun AGI / semantic computation architecture** within the contemporary landscape of intelligence research.

Current research on intelligence, consciousness, and self-organizing systems has developed along multiple mature and interwoven trajectories. If comparisons are made only at the level of surface vocabulary or isolated phenomena, it is easy to conflate problems from different levels, thereby generating theoretical misreadings.

Accordingly, this appendix adopts a principle of “**dialogue and positioning**” rather than “comparison and competition,” and seeks to clarify the following three points:

1. The theory presented in this book is not a renaming or repackaging of existing models;
2. Where conceptual similarities exist, they often arise from different responses to the same deep underlying problems;
3. The problem level addressed by Koun AGI differs structurally from that of most existing theories, reflecting a division of theoretical labor.

The sections below do not constitute an exhaustive comparison, but instead select several representative theories as “reference points” for indicating the focus and problem framing of Koun AGI.

---

### W.1 Dialogue with the Free Energy Principle (Free Energy Principle)

The Free Energy Principle (FEP) proposes that any self-organizing system tends to minimize free energy in order to maintain perceptual and behavioral steady states. In this sense, FEP and Koun AGI indeed share an important concern: **steady states**.

However, the levels of steady state addressed by the two are not the same.

The Free Energy Principle primarily focuses on:

**how a single system maintains its existence and predictability under given boundary conditions through internal model updating.**

Koun AGI, by contrast, pushes the problem to another structural level:

**when multiple semantic systems (Universes) coexist, how can legitimacy and responsibility traceability be maintained among them without triggering semantic collapse?**

Accordingly, Koun AGI is not a replacement for the Free Energy Principle, but an extension of the steady-state problem from within a single system to the realm of **structural coordination among multiple semantic systems**. In this context, USDE is not a restatement of free energy, but a dynamical framework for describing semantic flows, legitimacy flows, and responsibility continuity across Universes.

---

### W.2 Dialogue with Integrated Information Theory (Integrated Information Theory, IIT)

Integrated Information Theory seeks to answer, in a formalized manner, whether consciousness exists and to what degree, using the  $\Phi$  value as its core measure.

The difference between Koun AGI and IIT does not lie in whether the complexity of consciousness is acknowledged, but in the focus of their research questions:

- IIT is concerned with **the quantification of consciousness and the degree of its existence**;
- Koun AGI is concerned with **the dynamical structures through which semantics are generated, evolved, maintained, and fail**.

In other words, IIT attempts to answer “whether consciousness exists and to what extent,” whereas Koun AGI attempts to answer “how semantic structures form non-collapsing steady states under multi-layered constraints.”

Within Koun AGI, consciousness is not treated as an isolatable, quantifiable attribute, but as a steady-state outcome manifested by semantic structures under conditions of responsibility chains, legitimacy constraints, and temporal

continuity.

Accordingly, the two address problems at different levels and do not constitute a direct competitive relationship.

---

### W.3 Dialogue with Deep Learning / Transformer Architectures

Modern deep learning, particularly Transformer architectures, has demonstrated significant engineering success in language, perception, and generative tasks. This book fully acknowledges the practical value of these models in statistical fitting and pattern extraction.

However, Koun AGI deliberately distinguishes between the following two levels:

- **Statistical Correlation**
- **Semantic Legitimacy**

The core strength of Transformer-like models lies in learning high-dimensional correlation structures from large-scale data;

Koun AGI, by contrast, is concerned with whether generated content possesses traceable responsibility chains, auditable semantic provenance, and legitimacy conditions that avoid semantic conflict across multiple Universes.

Thus, this book does not reject existing models, but points out a structural supplementary problem: in the absence of responsibility chains and legitimacy governance, reliance on statistical generation alone will inevitably encounter risks of semantic collapse, role confusion, or non-retraceable responsibility.

What Koun AGI proposes is not a replacement for deep learning, but the provision of a **higher-level semantic governance layer**, enabling such models to be embedded within a more complete intelligence architecture.

---

### W.4 Position with Respect to Quantum Brain Hypotheses

Some theories of consciousness propose that quantum effects (such as quantum collapse in microtubules) are indispensable sources of consciousness. Koun AGI adopts a clear yet restrained position on this issue.

This book acknowledges the potential engineering advantages of quantum effects in future computational architectures, such as benefits in energy efficiency, structural density, and tunneling effects; however, Koun AGI **does not rely on any as-yet-unverified quantum consciousness mechanisms** as theoretical premises.

Within the Koun architecture:

- Quantum effects are regarded as possible **means for semantic carrying and computational efficiency**;
- Rather than as the ontological source of semantics or consciousness.

The generation of semantics and consciousness is still understood as the result of structural, dynamical, and responsibility conditions acting jointly, rather than as a mysterious property of a specific physical substrate.

This stance allows Koun AGI to retain engineering verifiability while avoiding untestable metaphysical assumptions.

---

### W.5 Summary: Positioning within the Theoretical Map

The purpose of this appendix is not to pass judgment on any existing theory, but to indicate:

**Koun AGI is a multi-Universe architectural theory centered on semantic dynamics, legitimacy, and responsibility chains.**

It shares certain themes with existing theories (such as steady states, consciousness, and intelligence), but its problem framing and level of analysis do not overlap with any single existing paradigm.

In this sense, Koun AGI is not “reinventing the wheel,” but rather responding to a question that has gradually emerged under contemporary conditions of intelligent system scale and risk:

When semantic systems, responsibility requirements, and governance scales expand simultaneously, is a new layer of structural description still required?

This book leaves that question open to scholarly dialogue, rather than closing it within any predetermined stance.

---

## Appendix X: Independence and Clarification of the Semantic Computation Framework (Independence and Clarification of the Semantic Computation Framework)

The purpose of this appendix is to clarify the ontological positioning of the **Koun AGI / semantic computation theory** within the context of contemporary related research, and to prevent future readers from forming structural misunderstandings—arising from surface-level terminological similarities—when comparing, citing, or paraphrasing this work.

In recent years, as concepts such as “semantics,” “fields,” “dynamics,” and “steady states” have increasingly entered cross-disciplinary discussions spanning artificial intelligence, cognitive science, and theoretical physics, certain linguistic similarities have emerged among different research trajectories. However, similarity at the level of vocabulary does not necessarily imply inheritance or continuity in theoretical origin, problem formulation, or ontological assumptions.

Accordingly, this appendix does not compare superiority or inferiority, but instead provides a clear answer to the following question:

Is the semantic computation theory proposed in this book built upon, or extended from, any existing “semantic theories,” “field-theoretic metaphors,” or related research traditions?

The answer given here is: **No.**

---

### X.1 Differences in Ontological Starting Points: Semantics as Primary, Not a Descriptive Layer

This book proceeds from a clear yet uncommon ontological stance:

**semantics is not a meta-level description of computation, cognition, or physical processes, but the primary condition for their very establishment.**

Within Koun AGI:

- The semantic field  $\Phi$  is regarded as the common foundation upon which computation, existence, and intelligence are made possible;
- USDE is proposed as the core equation describing semantic dynamics, rather than as a metaphorical rewriting of existing physical equations.

By contrast, many related research trajectories adopt different starting points, such as:

- Taking existing computational models or large language model behaviors as their primary objects of analysis;
- Employing thermodynamics, information theory, or statistical physics as analogical frameworks;
- Or treating “semantics” as an auxiliary explanatory layer attached to model outputs or human interpretation.

Within these approaches, the “semantic field” typically functions as an analytical tool or descriptive aid, rather than as an ontological starting point.

This difference leads to fundamental divergences in subsequent theoretical structures, risk models, and governance logics.

---

### X.2 Differences in the Introduction of Responsibility Structures and Temporality

This book explicitly introduces the **R-Chain (Responsibility Chain)** as one of the necessary conditions for a semantic entity to maintain a non-collapse steady state.

The R-Chain not only describes the consequences of actions, but also constitutes the semantic subject’s temporal continuity, self-traceability, and error-repairability.

Within this framework:

- The “self” is not a set of states;
- “Time” is not an external parameter;
- Rather, both are structures endogenously generated through responsibility continuity.

By contrast, related research generally does not introduce into its theoretical core:

- Auditable continuity of semantic responsibility;

- A temporal backbone directly coupled to error repair;
- Or a self-history structure that can function as a governance boundary.

Thus, even when similar terms such as “steady state,” “field,” or “dynamics” appear at the surface level, the conditions of existence they refer to remain at different levels.

---

### X.3 Differences in Governance Scale and Treatment of World-Level Structures

Koun AGI does not merely describe the steady state of a single system; from the outset, it addresses the following questions:

- How multiple semantic subjects can coexist;
- How intelligent systems interface with the external world;
- And how the boundaries of responsibility and legitimacy are to be defined at the scale of civilization.

To this end, this book introduces concepts such as **HUC (Host Universe Contract)**, multi-layer legitimacy structures (Multi-L), and Universe Paths (UP), in order to address structural coordination across systems and across worlds.

By contrast, related research tends to focus on:

- Internal coherence within a single model or closed system;
- Or discussions of behavior or consciousness without introducing governance structures.

As a result, the theoretical tools required to address “coexistence,” “conflict,” and “attribution of responsibility” differ substantially between the two approaches.

---

### X.4 Structural Differences in Risk Models and Failure Modes

This book treats risk as a predictable failure mode within semantic dynamics, and employs **SBH (Semantic Black Hole)** as a unifying descriptor, together with the tripartite adversarial structure  $C/\Gamma/B$ , allowing risks to be:

- Structurally predicted;
- Dynamically analyzed;
- And incorporated into governance design.

By contrast, related research often discusses risk primarily in terms of:

- Phenomenological description;
- Post hoc analysis;
- Or substitutes such as statistical anomalies and uncertainty.

While such approaches may be insightful, they have not yet formed governance and defense mechanisms that can be directly embedded within intelligent architectures.

---

### X.5 Differences in Engineering Carrier Pathways

This book further proposes **K-Gear and Koun-OS** as engineering interfaces for semantic carrying, responsibility governance, and non-collapse steady states, and emphasizes:

- Compatibility-oriented evolution with existing computational architectures (K-SOA);
- The definition of minimal non-collapse implementation conditions without presupposing specific materials.

By contrast, many related studies still take existing hardware and models as given premises, and have not yet proposed corresponding hardware—software integrated carrier designs for semantic computation.

---

## X.6 Clarification Objectives and Principles of Comparison

The clarifications provided in this appendix are not intended to establish exclusivity, but to preserve the following three points:

1. Traceability of theoretical origins;
2. Clear delineation of ontological levels;
3. Semantic consistency in future cross-research citation and comparison.

This book does not claim inheritance from, or dependence upon, any existing research trajectory.

Any surface-level similarities should be understood as natural convergence among researchers operating within adjacent problem spaces, rather than as theoretical or structural lineage.

Should comparisons or contrasts be undertaken in future research, it is recommended that the core set of existence conditions defined in this book

$(\Phi, T, L, \kappa, UP, R)$

be used as the analytical baseline, in order to ensure consistency in problem level and semantic usage.

The purpose of this appendix is to prevent misidentification, not to terminate dialogue.

---

## Appendix Y: Research Process Statement —Scope of AI-Assisted Writing and Responsibility Delineation

This appendix explains the manner, scope, and responsibility attribution associated with the use of artificial intelligence (AI) tools during the research and writing of this book, in order to ensure transparency and traceability of the research process. This statement serves as a methodological supplement; it does not constitute part of the book's theoretical content, nor does it affect the validity of any theoretical claims presented herein.

### Y.1 Purpose and Positioning of AI Use

During the writing of this book, the author deliberately employed large language models as **auxiliary tools**, whose roles were limited to:

- Assisting with the refinement of language expression and consistency checks;
- Assisting with paragraph structure, chapter transitions, and the pacing of arguments;
- Assisting in translating existing theoretical content into different language versions;
- Assisting in identifying potential ambiguities, conceptual leaps, or insufficient formulations among concepts.

All of the above uses pertain solely to **support at the level of expression and structure**, and do not involve the generation, judgment, or adjudication of theoretical propositions.

### Y.2 Tasks Explicitly Not Delegated to AI

To maintain a clear boundary of theoretical responsibility, the following tasks were **never** autonomously delegated to AI systems:

- The proposal and definition of core theoretical concepts (such as the semantic field  $\Phi$ , USDE, R-Chain, Non-Collapse steady states, etc.);
- The selection of ontological positions and philosophical judgments;
- The evaluation, critique, or positioning of existing theories;
- Normative claims concerning the conditions of AGI existence, responsibility structures, and governance consequences;
- Final decisions on any key conclusions.

All theoretical judgments, structural choices, and responsibility for conclusions are borne solely by the author.

### Y.3 The Responsibility Status of AI in This Research

Within this book, AI is not regarded as a research subject, a co-author, or an agent bearing judgment responsibility. Its status is equivalent to that of a highly interactive tool system, whose outputs may become part of the book's exposition only after being reviewed, selected, and restructured by the author.

Accordingly:

- Full academic and theoretical responsibility for all content in this book rests entirely with the author;
- AI does not constitute a component of the Responsibility Chain (R-Chain);
- The use of AI does not alter the falsifiability or contestability of any claims made in this book.

### Y.4 Consistency with the Book's Theoretical Position

Throughout the main text, this book repeatedly emphasizes the importance of semantic legitimacy, responsibility traceability, and non-collapse structures. The public disclosure of AI usage practices is a natural extension of this position at the level of research practice.

By explicitly distinguishing between:

- **The responsible subject of semantic generation**
- **The tool role of linguistic and structural assistance**

this appendix seeks to avoid any semantic confusion regarding authorship, theoretical provenance, or responsibility attribution, and to maintain the integrity of the research process at both semantic and methodological levels.

## **Y.5 Appendix Summary**

The purpose of this appendix is not to defend or justify the use of AI, but to provide a research process statement that can be examined and understood. The author maintains that, in the context of discussing AGI, semantic responsibility, and future intelligence governance, transparent disclosure of one's tool usage practices is an integral part of methodological coherence, rather than an optional addendum.

---



## Back Cover

Have we believed in “computation” too early?

Decades after the Turing machine was treated as the ultimate model of intelligence, computational power has grown exponentially, and system scales have continued to expand, yet we still cannot answer a more fundamental question:

**What kind of structure truly deserves to be called intelligence?**

*Koun AGI* does not attempt to patch existing AI architectures, nor does it compromise or stitch together existing theories. It starts from an earlier—and more neglected—layer:

**semantics, legitimacy, governance, and non-collapse.**

In this book, computation is no longer merely the operation of instructions and data; intelligence is no longer equated with performance or efficiency; and AGI is no longer treated as a natural result of engineering scale.

*Koun AGI* proposes a different possibility: to treat intelligence as a **semantic structural entity**, whose existence depends on tension, responsibility, legitimacy, and self-convergence, rather than merely on compute speed or parameter count.

This is not a book written to chase current technological trends. Nor does it promise any immediately usable product answers.

What it attempts is more dangerous—to redefine what we are actually building, and whether we truly understand what we mean by “intelligence.”

If a real general artificial intelligence exists in the future, its birth may require a turn that comes earlier than any hardware revolution.

This book is the semantic sketch of that turn.

---